

NORTHWESTERN UNIVERSITY
MINI-COURSE—“ECONOMICS AND EVOLUTION”

March 10-14, 2008

Arthur J. Robson

MONDAY

Where do preferences come from? What determines their properties? Though traditionally religiously shy about asking such questions, some economists have recently broken the taboo and turned to biological models for illumination.

ADAPTATION OF UTILITY

Robson, A.J. “The Biological Basis of Economic Behavior,” J. Econ. Lit. 29 (2001), 11-33.

A biological perspective suggests that the utility function we like to refer to in such an offhand way, as really only preferences that is, IS really hedonic in character. We eat lunch because being hungry is slightly unpleasant and eating feels good. This raises all sorts of awkward questions that motivated the stripping of any vestige of cardinality, let alone pleasure or pain, from the notion of utility.

For example, psychologists (see Frederick and Lowenstein) discuss how a burst of intense pleasure stems from winning the lottery, but this pleasure subsides soon, ending up only a little higher. Analogously, the intense sadness that arises from becoming the victim of a crippling accident fades, so that you end up only a little sadder than you were. This seems potentially awkward because it seems to predict that you would not mind running the risk of a catastrophic accident or would not buy a lottery ticket.

But this issue of adaptation is not so awkward upon closer examination. More than that, adaptation of hedonic utility is exactly what you would expect as being biologically optimal. Furthermore, the behavior that it generates need not be very different from what

the conventional model generates.

The issue is analogous to an old-fashioned voltmeter, one with a needle. To get an accurate reading from an old-fashioned voltmeter, it is necessary to first estimate the range into which the unknown voltage falls. Only if the range is set right, can you obtain information. The problem is that real measuring devices have limited sensitivity, although more expensive instruments have more. Similarly, biological organisms have limited perceptual discrimination, although this can be improved at a biological cost.

Consider the following example. An individual organism must choose between two “arms.” These arms generate numbers drawn independently from the same known continuous cumulative distribution function, cdf, F . The individual must choose after the draws are made, but can only perceive where each realization falls relative to a set of N threshold values

$$c_1 < c_2 < \dots < c_N.$$

These perceptual possibilities are experienced as hedonic utility. Suppose, for example, each threshold generates a utility increment of $1/N$ so that utility is normalized to run in equal steps from 0 to 1. If the choices both lie in the same range, mistakes arise, with probability $1/2$. What values of the c 's then minimize the probability of error? This probability of error is

$$\begin{aligned} PE(N) &= (1/2)(F(c_1))^2 + \dots + (1/2)(F(c_{n+1}) - F(c_n))^2 + \dots + (1/2)(1 - F(c_N))^2 \\ &= (1/2)(y_1)^2 + \dots + (1/2)(y_{n+1} - y_n)^2 + \dots + (1/2)(1 - y_N)^2, \end{aligned}$$

where $y_n = F(c_n)$, for $n = 1, \dots, N$

It can be shown that $PE(N)$ is a convex function of (y_1, \dots, y_N) so that the first-order conditions are sufficient for a global minimum.

These first-order conditions are

$$\begin{aligned}\frac{\partial PE(N)}{\partial y_1} &= 0 \text{ so } y_2 - y_1 = y_1 - 0 \\ \frac{\partial PE(N)}{\partial y_n} &= 0 \text{ so } y_{n+1} - y_n = y_n - y_{n-1}, \text{ for } n = 2, \dots, N - 1 \\ \frac{\partial PE(N)}{\partial y_N} &= 0 \text{ so } 1 - y_N = y_N - y_{N-1}\end{aligned}$$

That is,

$$y_1 - 0 = k, \quad y_{n+1} - y_n = k, \text{ for } n = 2, \dots, N - 1 \text{ and } 1 - y_N = k.$$

It must then be that $k = 1/(N + 1)$ so that

$$y_n = F(c_n) = n/(N + 1), \text{ for } n = 1, \dots, N.$$

For example, if $N = 9$, the thresholds should be at the deciles of the distribution. It is optimal therefore for the thresholds to adapt to the circumstances, as represented by F .

What is the probability of error $PE(N)$ when the thresholds are chosen optimally like this?

$$PE(N) = \overbrace{\frac{1}{2(N+1)^2} + \dots + \frac{1}{2(N+1)^2}}^{N+1 \text{ terms}} = \frac{1}{2(N+1)} \rightarrow 0, \text{ as } N \rightarrow \infty.$$

Suppose now that there is a cost that is proportional to the probability of error, is this probability then with a harmless choice of constant, but there is another component of cost, a cost of complexity, that depends directly on N , as $c(N)$. The total cost is then

$$PE(N) + c(N),$$

which should be minimized over choice of N . If $c(N) \rightarrow 0$, in some reasonable way, it follows readily that $N \rightarrow \infty$, and that $PE(N) + c(N) \rightarrow 0$. Although utility remains adaptive, even in this limit, the resulting choice behavior is exactly as conventionally predicted. It is only because there is a cost of computation that there is any difference at all; the lower this cost, the better an approximation the biological model is to the usual one.

It is implicit in the conventional approach that there are no costs of computation. However, any reasonable economist should presumably admit that such costs exist and that they will make a difference. The reason that they are left out of the conventional account is that it is hard to say exactly what form these costs would have, combined with a belief or a hope anyway that these costs are not overwhelmingly large.

Netzer, N “Evolution of Time Preference and Attitudes to Risk,” Universität Zürich WP (2008)

The probability of error is not the most convincing objective here. After all, some errors involve a very big gap between the x you chose and the x you could have had, and some involve a very small gap. A more plausible objective would be to maximize the expected value of the x you receive, assuming x is fitness.

In the simplest case where there is one threshold, c , the appropriate choice is $c = \int x f(x) dx$, the mean of F . Things are tougher if there are an arbitrary number, N , say, of thresholds. Netzer investigates this issue in the limit as $N \rightarrow \infty$. Each threshold is again interpreted as an increment in utility of $1/N$ and utility again ranges between 0 and 1. Netzer further simplifies the problem by taking this limit as $N \rightarrow \infty$ for a step density approximation of the original density. He then takes the limit as the number of steps, S , tends to infinity, so the step density converges uniformly to the original density function. For the probability of error criterion, he shows that

$$V(x) = \lim_{S \rightarrow \infty} \lim_{N \rightarrow \infty} V_{N,S}(x) = F(x) = \int^x f(y) dy,$$

as before. For the expected fitness criterion he shows that

$$U(x) = \lim_{S \rightarrow \infty} \lim_{N \rightarrow \infty} U_{N,S}(x) = c \int^x f(y)^{2/3} dy \text{ where } c = \left(\int f(y)^{2/3} dy \right)^{-1}.$$

That is, with this criterion, since the size of mistakes now matters, it pays to make the thresholds more evenly spaced than under the probability of error criterion.

The assumption that fitness is x is without loss of generality, in that, if fitness were a monotonically increasing function of x , we could easily find the cdf over fitness that is implied by the given cdf over x . It is trivial to modify these expressions to reflect this.

Netzer applies this model to intertemporal choice. Suppose that the choices all yield a fitness payoff of 1 after a time t which is distributed according to a density g on $[0, T]$. The fitness of an alternative with waiting time t is devalued on this account to $\psi(t) = e^{-\delta t}$. This might be because of a constant mortality rate or population growth, as we will discuss. It follows that, making the trivial modification mentioned a moment ago, and where t is a bad,

$$U'(t) = -cg(t)^{2/3}(-\psi'(t))^{1/3} < 0.$$

It seems plausible that $g(t)$ is quite a bit larger for small values of t than it is for larger values. That is, most of the decisions that an individual has to make involve excursions into the immediate future at most. It follows that the rate of time discounting of utility may well be higher at first than it becomes for more distant times. That is, there is apparently a present bias along the lines of behavioral economics.

In order to generate actual reversals, however, some element of evolutionary irrationality must be introduced. Suppose, for simplicity, that an individual must either choose among immediate alternatives OR among alternatives in the more or less distant future, but never has to choose between a present and a future alternative. Accordingly, the individual may have evolved a mechanism for utilizing the range of utility either only in one context or only in the other. Suppose the immediate choices are at $t = 0$ and the future choices are at $t = 1$ and $t = 2$, for example. That is, there is a interpretation switch that makes it possible to use the same utility meter either in the first set of circumstances or the second without ambiguity. Suppose then you choose an alternative at $t = 2$ over one at $t = 1$. If, despite the restriction we just described, you are suddenly unexpectedly offered a chance

to reconsider after one period, you may well prefer the now immediate alternative. This is certainly true if you are accustomed to sufficiently low immediate payoffs, since this will make a given positive immediate payoff exciting.

Finally, this could be the basis for the chicane-shaped utility of prospect theory. That is, if the distribution of fitness is unimodal, then the evolutionary U will be qualitatively similar to the cdf, as needed. This utility will furthermore automatically generate this shape relative to the initial wealth.

Rayo, L., and Becker, G. “Evolutionary Efficiency and Happiness,” JPE 115 (2007), 302-337.

Rayo and Becker also reconsider the evolutionary basis of adaptive hedonic utility. The adaptation considered has two aspects—

i) habituation—getting used to a permanent shift, positive or negative, in your circumstances

ii) peer comparisons—a concern with relative income or wealth

What these have in common is a description in terms of a reference point, that is determined by your own past outcomes, or by the past and present outcomes of peers, respectively.

Rayo and Becker also view utility as hedonic, as a biological device that induces appropriate actions by an individual. In particular, Mother Nature chooses the mapping from material outcomes into pleasure in the most effective way possible. In the present context, this most effective way involves the construction of a reference point that reflects the individual’s expectations of the world.

Again, there is a metaphorical principal-agent problem here, with MN as the principal and the individual as the agent. MN “wishes” the individual to be maximally fit, she has the extraordinary ability to choose the utility function of the agent to her best advantage. Key ingredients of the model are a limited range of utility levels that are possible, and a

limited ability to make fine distinctions.

STATIC MODEL

We will limit attention here to the basic atemporal model, which precludes getting to some of the applications mentioned. Consider a hunter-gatherer who is undertaking a one shot project. She first observes a state s specifying the physical state of the world—the availability of game animals, the presence of sabre-toothed tigers, etc. Next, she chooses $x \in X$ as the strategy adopted—the method of hunting, whatever.

$f(y | x, s)$ = the pdf of output $y \in \mathfrak{R}$ given x, s where this pdf is known to the agent.

y represents goods such as money or food. The agent here is guided by the coupling of emotional rewards to this output. The agent then chooses x so as to maximize her expected emotional payoff.

Why might it be advantageous for MN to do this rather than to simply prescribe the agent's choice as a function of s ? Perhaps MN is daunted by the sheer complexity of the second option. Better yet, perhaps there are novel s 's and novel f 's, but the individual is in a position to still estimate the expected emotional reward.

$V(y)$ = the hedonic utility of income y ; $V'(y) > 0$.

The bottom line for the agent is then to maximize

$$E[V | x, s] = \int V(y)f(y | x, s)dy$$

over choice of $x \in X$.

MN has the ability to choose V optimally. But what does MN want? It is assumed that MN wants to maximize $E(y)$. This would make sense if y were expected offspring, since this says MN wishes to maximize overall expected offspring. If y is money or food, it's straightforward to extend this analysis if $U(y)$, say, is expected offspring from income y ,

so that MN wants to maximize $E[U(y)]$. That is, the individual's problem can be redefined so that it is given in terms of $z = U(y)$. It is wlog to assume $E(y)$ is the criterion, that is.

A key element of the model is that there are bounds on V so that

$$V \in [\underline{V}, \bar{V}], \text{ so that } V \in [0, 1], \text{ wlog.}$$

This stems from the finite number of neurons in the brain, most basically. These upper and lower constraints will be binding. MN would benefit from a wider range of emotional responses. It is expensive, however, to enlarge the range, and so this range must be finite and MN must use whatever this range is optimally.

A second key element in the model is limited discrimination. This takes the precise form that, if

$$|E[V | x_1, s] - E[V | x_2, s]| \leq \varepsilon$$

then the individual cannot rank x_1 and x_2 . Hence all choices within ε of $\max_{x \in X} E[V | x, s]$ are "optimal." It is assumed that the agent randomizes uniformly, or at least using a continuous distribution with full support, over this satisficing set. Of course MN would also prefer a smaller value of ε , but this is expensive, and she will have to optimize given some optimal $\varepsilon > 0$.

Output is given by

$$\tilde{y} = m(x, s) + \tilde{z}, \text{ where } \tilde{z} \text{ has a continuous unimodal distribution, } E(\tilde{z}) = 0, \text{ so } E\tilde{y} = m(x, s).$$

That is, the effect of changes in x or s is to shift the entire distribution of output up or down without changing its shape. A choice of x that increases $m(x, s)$ entails a new distribution that FOSD the old one.

The following trick simplifies the notation

$$\varphi(x, s) = \frac{E[y | x, s] - \min_x E[y | x, s]}{\max_x E[y | x, s] - \min_x E[y | x, s]} \in [0, 1].$$

This index captures the relative efficiency of choice x in state s . Expected output can then be rewritten as

$$E[y | x, s] = \varphi \max_x E[y | x, s] + (1 - \varphi) \min_x E[y | x, s] \equiv E[y | \varphi, s]$$

where $E[y | \cdot, s]$ is monotone increasing. Output can be expressed as

$$\tilde{y} = m(\varphi, s) + \tilde{z} \text{ so the pdf of } y \text{ is now } f(y | \varphi, s).$$

Consider the simplest case that the state does not matter so

$$E[y | \varphi, s] = E[y | \varphi]$$

and

$$E[V | \varphi] = \int V(y) f(y | \varphi) dy.$$

The satisficing set is now

$$\{\varphi | E[V | \varphi] \geq \max_{\varphi} E[V | \varphi] - \varepsilon\} = [\varphi_{\min}(V, \varepsilon), 1].$$

Now MN's problem becomes

$$(I) \quad \max_{V(\cdot) \in [0,1]} \varphi_{\min}(V, \varepsilon)$$

Proposition 1. This problem has as its essentially unique solution

$$V^*(y) = \begin{cases} 1 & y \geq \hat{y} \\ 0 & y < \hat{y} \end{cases}$$

where \hat{y} solves

$$f(\hat{y} | 1) = f(\hat{y} | \varphi^*).$$

Consider now the limit as $\varepsilon \rightarrow 0$. Of course, $\varphi^* \rightarrow 1$ and $\hat{y} \in [E(y | \varphi^*) - E(y | 1)] \rightarrow E(y | 1)$. This V penalizes the agent at the maximal rate for choosing suboptimal x . This optimal assignment of utility/pleasure to income is a limiting case of S-shaped V .

This S-shape might be more plausible than a cliff-shape. R&B show it would arise if deviations from a given reference level V_0 were costly. Alternatively, it might be that the agent knows more about the output technology than does MN. Now MN might not be able to target $E(y | 1)$, since she doesn't know this. Now MN might need to smooth out V to provide strong incentives over a range of possible $E(y | 1)$'s.

TUESDAY

RISK

Now I want to turn to the implications of an evolution for two key attributes of preferences—attitudes to risk and time preference. These seem somehow two of most eligible candidates for an evolutionary explanation, partly because they are so general rather than specific. Integrating these arguments completely with the adaptive hedonic view is yet to be done, so let's put that on hold.

Robson, A.J. "A Biological Basis for Expected and Non-Expected Utility,"
J. Econ. Theory, 1996, 68, 397-424.

It is straightforward to derive the EUT from biological evolution, as long as all the reproductive risk is independent across individuals—"idiosyncratic," that is. Consider the following simple model. Individuals have a short and not so merry life. They are born at age 0, survive for sure to age 1, reproduce, and die immediately. The bad news is: No sex.

If all the risk concerning reproduction is idiosyncratic, and the population is "large," the average number of offspring will be the expected offspring per capita, and this is the growth rate of the population. A type achieving a higher growth rate will be more blessed in the eyes of Darwin. But maximizing expected offspring is equivalent to maximizing expected utility of consumption if expected offspring is a function of consumption.

But there seems no compelling reason why all risk should be idiosyncratic like this. Perhaps some of the risk in a hunter-gatherer society concerned the weather, for example, which clearly has a shared component. In a modern setting, the stock market remains

a shared risk. The opposite polar extreme to idiosyncratic risk is aggregate risk. If idiosyncratic risk means a separate personal coin is flipped for each individual, aggregate risk means one big public coin is flipped—heads everyone wins, tails everyone loses.

What difference would this make? Let's warm up for this by considering this mild puzzle. I have a deck of 52 cards, 26 red and 26 black. I offer you the following gamble: \$1 to begin with, and then we go through the deck doubling the amount on each red card, leaving it alone on each black. You are risk neutral, say. What is your expected wealth after one iteration? \$1.50. After that things get more complicated, but suppose I offer you the alternative of $(3/2)^{52} \simeq \$1,400,000,000$. Which is better? The gamble is not a gamble at all; it pays exactly $2^{26} = (\sqrt{2})^{52} \simeq \$67,000,000$. The “gamble” is worse because $\sqrt{2} < 3/2$.

What would happen if the deck were infinite, but we just looked at the first 52 draws, or T draws more generally? Now the gamble really is a gamble, and the expected amount after T draws is $(3/2)^T$, because every draw IS like the first one. The two options are then precisely equivalent in terms of expected values.

There is still a strong sense in which the gamble is much worse than the certain deal. Let's switch to biological evolution as the background for the mathematics. There are two types. These types have a short life but a merry one. Type 1 has 2 offspring with probability 1/2, or 1 offspring also with probability 1/2. All the risk here is idiosyncratic, independent across all individuals and dates. Type 2 also has 2 offspring with probability 1/2, or 1 with probability 1/2, but now the risk is aggregate. That is, all the type 2 individuals alive at a particular date either get a draw of 2, or they all get a draw of 1.

Given a “large” population, the number of type 1's at date T is $x(T) = (3/2)^T$, assuming $x(0) = 1$. A large population of size one is a phrase that makes undergraduates laugh, as perhaps they should. Note that $\frac{1}{T} \ln x(T) = \ln(3/2)$ measures the rate of exponential growth of this type.

The number of type 2's is inescapably random. The exact number of type 2's is $y(T) = 2^{n(T)}$, where $y(0) = 1$. Now $n(T)$ is the number of heads in a sequence of T flips of a fair coin. It follows that $\frac{1}{T} \ln y(T) = \frac{n(T)}{T} \ln 2 \rightarrow \frac{1}{2} \ln 2 = \ln \sqrt{2}$, w.p. 1, by the strong law of large numbers. This shows that there is a limiting exponential rate of growth for this type and derives it. In terms of a deck of cards, the worse growth rate was imposed by the equal number of red and black cards in a finite deck, but the law of large numbers has much the same implication with an infinite deck.

It follows that $x(T)/y(T) \rightarrow \infty$, w.p. 1. In a strong sense, the first type outdoes the second, despite the fact that the expected values of the two are exactly the same. As we saw already, however, $E(y(t)) = (3/2)^T$, so that, in this same sense, the type 2 population falls short of its own mean in the limit.

This is an example of the martingale convergence theorem, which allows just this sort of perversity. The original meaning of the word martingale was the following betting strategy, which undergraduates still ask you about. Bet \$1 on a fair coin. If you win, quit, now up \$1. If you lose, bet \$2 on the next throw. And so on, doubling each time you continue to lose until you win, ad infinitum. Although your expected change in wealth after any finite number of flips must always be 0, because every flip is fair, there is a well defined limiting distribution that puts weight 1 on \$1. That is, the limit of the means is not equal to the mean of the limiting distribution. The reason this happens is that the distribution after a finite number of flips puts a very small probability weight on a very large loss. Together, this yields a non-vanishing contribution to the mean. In the limit, however, the probability converges to zero, and so the mean jumps up.

In a similar way, the mean of the type 2 population is held up by very small probabilities of very large populations. In the limit, these probabilities vanish, and all that is left is the smile of the Cheshire cat, in that the growth of the population is overestimated by the mean.

The bottom line is that individuals should be more averse to aggregate risk than they

should be to equivalent idiosyncratic risk. The example can be recast as an economic choice as follows. Suppose that bundles b_1 and b_2 induce the offspring levels 1 and 2 as above, so $\Psi(b_1) = 1$ and $\Psi(b_2) = 2$, where Ψ is the common production function for expected offspring from commodities. Now individuals must choose between gamble 1 and gamble 2, as follows. Gamble 1 yields b_1 and b_2 each with probability $1/2$, where all this risk is independent. Gamble 2 also yields b_1 and b_2 each with probability $1/2$, but now all this risk is aggregate. The above theory implies that individuals should prefer gamble 1 to gamble 2.

From a EUT point of view, these two gambles should be equivalent. Indeed, even from the perspective of any decision theory based on “probabilistic sophistication,” these two gambles should be equivalent. It is not enough here to consider only one’s own payoffs and the probabilities. One must also consider how the uncertainty affects others. That is, in contrast to other DT approaches, preferences here are interdependent.

The most general case that can easily be analyzed is as follows. Given an aggregate environments s , each type i faces an idiosyncratic economic gamble where $q_k^{i,s}$ = the probability of receiving a commodity bundle $x_k^{i,s}$. It is assumed that $\Psi(x)$ = the expected offspring from bundle x for any state and any type. Hence $\sum_k q_k^{i,s} \Psi(x_k^{i,s})$ = expected offspring of type i in state s . If each state s has probability π_s then the long run limiting exponential growth rate of type i is

$$\sum_s \pi_s \ln \left(\sum_k q_k^{i,s} \Psi(x_k^{i,s}) \right).$$

Hence the type that maximizes this expression should be favored by natural selection.

I would love to find someone willing to do experiments aimed at discerning whether this distinction matters. Better yet, someone who might look at investment behavior to check whether instruments such as stocks that are subject to aggregate risk are depreciated in investors’ affections on that account relative to those that are not.

Cooper W. S., R. H. Kaplan. 1982. Adaptive “coin-flipping”: a decision-theoretic examination of natural selection for random individual variation. *Journal of Theoretical Biology*. 94:135–151.

This distinction between idiosyncratic and aggregate risk implies a new motivation for gambling; a motivation that seems unrelated to convexity of the vNM utility function. The basic reason is very simple. Consider the fate of an Arctic Hare with global warming. Suppose the probability of a snowy winter is $p \in (0, 1/2)$; so the probability of a clear winter is $1 - p \in (1/2, 1)$. Animals that keep dark coats survive clear winters for sure but die in snowy winters; those that develop white coats survive snowy winters but die in clear ones. Either pure type is doomed. Consider a type that randomizes choosing a white coat with probability π and a dark coat with probability $1 - \pi$. The growth rate is then

$$r = p \ln \pi + (1 - p) \ln(1 - \pi).$$

This is maximized by choosing $\pi = p$.

The most controversial thing about this paper was how C&K described this situation. They asserted that individuals who choose a white coat after the flip of the coin are (coin-flipping) “altruistic.” Why? Because the probability of such an individual dying is higher than the probability for an individual with a dark coat. It is interesting that these biologists seemed to see the situation the way economists might at first blush. But the issue here is: What is the evolutionarily correct notion of fitness, as a biologist would put it, or the correct utility function, as an economist would? The correct notion of utility must be interdependent.

Curry, PA. “Decision making under uncertainty and the evolution of interdependent preferences.” *Journal of Economic Theory*, 98: 2001, 357-369.

Grafen, A. “Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging,” *Proceedings of the Royal Society B: Biological Sciences* 266,

799-803.

Consider a continuum of animals of size 1, say. Suppose π choose white and $1 - \pi$ choose dark. Now consider the choice of a small mass of individuals of size ε . If they choose white, the expected fraction of the population they will constitute is $\frac{p\varepsilon}{\pi} = \varepsilon$ if $p = \pi$. If they choose dark, the expected fraction of the population they will constitute is $\frac{1-p}{1-\pi}\varepsilon = \varepsilon$, if $p = \pi$. It follows that each individual of the type that randomizes $(\pi, 1 - \pi)$ is maximizing its expected relative weight in the population, and this is the right notion of fitness. Dead is just dead, zero fitness, but when you survive it matters how large you loom in the population. White is equivalent to dark just because the survival probability of white is generally worse than that of dark, so one individual looms larger in the smaller white remnant.

From an economic point of view, the usual selfish preferences are inadequate here, since it is not enough to consider just the probability of death, as these selfish preferences would also entail. It is necessary to consider how well you are doing relative to others. Curry translates Grafen's point into economic terms. He shows that it is a Nash equilibrium for individuals to mix $(\pi, 1 - \pi)$ where each individual's payoff is the ratio of a selfish term to a social term reflecting the average selfish payoff in the population. Individuals are indifferent between the pure strategies involved if the right proportions obtain socially.

WEDNESDAY

TIME PREFERENCE

We turn now to intertemporal preferences. That is: Why do we discount future rewards? How exactly do we do this? Today's talk is based mainly on joint work with Larry Samuelson. We derive intertemporal preferences as the evolutionary implications of different age structured reproductive strategies. We will eventually focus on the case where some of the uncertainty involved in reproduction is aggregate. Inevitably enough, evolution still selects the life history that maximizes the appropriately defined growth rate of the associated group of individuals.

Robson, A and Samuelson L. “The Evolution of Intertemporal Preferences,”
American Economic Review 97 (2007), 496-500.

Idiosyncratic risk, age structured population.

Suppose that agents live for at most ℓ periods, producing x_i offspring in each period $i = 1, \dots, \ell$. There is still no sex in the model. Not until tomorrow will it be rediscovered. A life history is then a collection $(x_1, x_2, \dots, x_\ell)$. Given a large population, its evolution can be described in terms of the Leslie matrix

$$L = \begin{bmatrix} e^{-\delta}x_1 & e^{-\delta} & 0 & \dots & 0 & 0 \\ e^{-\delta}x_2 & 0 & e^{-\delta} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ e^{-\delta}x_{\ell-1} & 0 & 0 & \dots & 0 & e^{-\delta} \\ e^{-\delta}x_\ell & 0 & 0 & \dots & 0 & 0 \end{bmatrix}.$$

$$N(t+1)^T = N(t)^T L = \dots = N(0)^T L^{t+1}$$

The Perron-Frobenius Theorem shows that such a population settles down into steady state growth. That is, the age structure is constant in the limit and there is a constant geometric growth rate. This growth rate is the dominant eigenvalue $\tilde{\lambda}$ —the unique positive root of the characteristic equation—

$$1 = \frac{x_1}{e^\delta \lambda} + \frac{x_2}{(e^\delta \lambda)^2} + \dots + \frac{x_\ell}{(e^\delta \lambda)^\ell}.$$

Note for future reference that the steady state growth rate result implies

$$\frac{1}{t} \ln \|N(t)\| \rightarrow \ln \tilde{\lambda}.$$

Evolution selects the life history that maximizes this growth factor. Along any evolutionary indifference surface, $e^\delta \tilde{\lambda}$ is constant and hence we have a linear function of the values x_1, \dots, x_ℓ . The evolved indifference surfaces are thus hyperplanes in a space of ℓ dimensions. Figure 1 illustrates these indifference curves. One such evolutionary IC includes the

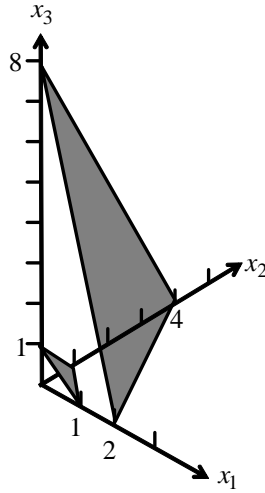


Figure 1: Indifference curves

strategies $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, for example. These reproductive profiles all maintain the parthenogenetic population at a constant level. Furthermore any convex combination of these unit vectors also does so. Another such IC includes $(2,0,0)$, $(0,2^2 = 4,0)$ and $(0,0,2^3 = 8)$ since all of these strategies effectively generate a doubling of the population in each period. Again, all convex combinations of these vectors also give effective doubling in each period. Thus the hyperplanes that represent the IC's twist as they get better to reflect greater discounting of future fertility. Success breeds impatience.

Induced Preferences Over Consumption

Economists are typically interested in preferences over consumption. The simplest, but certainly not the only, way to generate these is to assume that current births are a function of current consumption, so that preferences over consumption are induced by the underlying preferences over births.

Let $f_\tau(c_\tau)$ give age- τ births as a function of period- τ consumption c_τ . Suppose that all the f_τ are strictly increasing and concave.

For any consumption vector c , an evolutionary indifference curve is defined by

$$1 = \frac{f_1(c_1)}{\theta} + \dots + \frac{f_\tau(c_\tau)}{\theta^\tau} + \dots + \frac{f_\ell(c_\ell)}{\theta^\ell},$$

where $\theta > 0$ is constant on a particular indifference surface. A higher value of θ indicates a higher indifference curve. Although there can be no additively separable representation, it follows readily that evolution's indifference surfaces over consumption bundles (c_1, \dots, c_ℓ) have the usual shape, in the sense that $\theta(c_1, \dots, c_\ell)$ (and hence $\lambda(c_1, \dots, c_\ell)$) is strictly increasing and quasi-concave.

Geometric Discounting

Agents can be induced to make optimal choices via exponentially discounting offspring, using the optimal growth rate. Letting (f_1, \dots, f_ℓ) be the optimal utility profile and $\theta = e^{\delta \tilde{\lambda}}$, where $\tilde{\lambda}$ is the leading eigenvalue of the corresponding Leslie matrix, we have

$$1 = \frac{f_1}{\theta} + \frac{f_2}{\theta^2} + \dots + \frac{f_\ell}{\theta^\ell}.$$

Now suppose an alternative utility profile (f'_1, \dots, f'_ℓ) gives a smaller growth rate. Then

$$\frac{f_1}{\theta} + \frac{f_2}{\theta^2} + \dots + \frac{f_\ell}{\theta^\ell} = 1 > \frac{f'_1}{\theta} + \frac{f'_2}{\theta^2} + \dots + \frac{f'_\ell}{\theta^\ell}.$$

The agent can thus be induced to make the optimal decision by geometrically discounted preferences. Cf Figure 1. Since $\ln \theta = \delta + \ln \tilde{\lambda}$, this shows that the rate of time discount is equal to the rate of population growth plus the mortality rate.

In an abstract sense, this is not much of a result. I am trying to capture your preferences by means of expected discounted fertility/utility. In order for me to do this, you first have to tell me your best alternative in terms of the true evolutionary preferences. And I use your answer to obtain the optimal growth rate which I use to construct your standard preferences.

As a related matter, the connection between the true evolutionary preferences and your standard preferences is a limited one. Your standard preferences are only guaranteed to

correctly rank the best choice relative to any suboptimal choice. They might well not get the order of second and third correct, for example.

However, interest in this argument is buttressed by the observation that we do have a good estimate in advance of the optimal growth rate—zero. That is, if we began with Adam and Eve 1.8 mya, and we have now reached 6-7 billion, or whatever, the average growth rate must have been essentially zero. So maybe I don't need to wait for your answer as to what's the best choice. Furthermore, this presumptive maximal growth rate of zero and the relevant mortality rate might have been rather stable, so they could safely be programmed into individuals.

One difficulty is that the mortality rate might be rather too small to account for time preference. Even in hunter-gatherer societies, mortality rates are quite modest—1-2%, or so, as long as we stay away from the very beginning and end of life. I am not sure how seriously one should take the estimates I have seen of the pure rate of time preferences, but these are often rather higher than this—maybe 10% or even 20%. This theory seems then to be sticking its neck out.

AGGREGATE RISK, AGE STRUCTURED POPULATION

Robson, A and Samuelson L. “The Evolution of Time Preference with Aggregate Uncertainty” WP (2008).

General problem.

Aggregate risk in an age structured population means that each Leslie matrix is chosen at random. Hence, where $\tilde{L}(s)$ is the random matrix chosen at date s —

$$\tilde{N}(t+1)^T = N(0)^T \tilde{L}(1) \tilde{L}(2) \dots \tilde{L}(t)$$

It is remarkable that the basic result of the two simpler cases still goes through. There still is limiting exponential growth. This is ensured by the “sub-additive ergodic theorem,” under weak assumptions—that the process is stationary, in particular. That is, this theorem

guarantees the existence of Λ such that

$$\frac{1}{t} \ln \left\| \tilde{N}(t) \right\| \rightarrow \Lambda, \text{ wp } 1.$$

Example. The point of this example is not its inherent plausibility, but to show that we are not in Kansas anymore. That is, with aggregate risk, the idiosyncratic lessons do not hold. This example involves a stark form of aggregate uncertainty. Suppose there are N aggregate states, one for each of the possible ages. These states arise randomly, in an iid fashion, with equal probabilities. In state $k = 1, \dots, N$, only age k parents have any offspring, and they have expected offspring x_k .

$$L_1 = \begin{bmatrix} x_1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

$$\vdots$$

$$L_N = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ x_N & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

For simplicity, we have set the death rate equal to zero, but nothing relies on that. The issue then is: How do the x_i trade off? With the comparable idiosyncratic risk, there would be a single Leslie matrix, with entries x_k/N in the first column. In this case, there would be discounting of future offspring, assuming positive growth. Is it also better to have offspring early, with aggregate risk?

It is enough to study the evolution of new borns, since the number of any other age class can be derived from that. They were left out before...

Proposition. If $z_0(t)$ = number of 0-yr-olds at t , then $\frac{1}{t} \ln z_0(t) \rightarrow \frac{\sum_{k=1}^N \ln x_k}{\sum_{k=1}^N k} = \Lambda$.

That is, there is no systematic discounting of the reproductive efforts of the elderly.

Proof. It's not hard to get the basic intuition for the result. Fix T . If k_T is state at T , only age k_T parents reproduce, with x_{k_T} offspring, so that the number of new borns now derives from the number of new borns k_T periods ago or $z_0(T) = x_{k_T} z_0(T - k_T)$. This process can be iterated. That is, if k_{T-k_T} is state at $T - k_T$, then the number of new borns at $T - k_T$ can in turn be related to the number of new borns k_{T-k_T} before that. $z_0(T - k_T) = x_{k_{T-k_T}} z_0(T - k_T - k_{T-k_T})$. It follows that the number of new borns at date T is the product of the associated $x_{k_T}, x_{k_{T-k_T}}, \dots$. These x 's arise from randomly choosing who will reproduce. But each age of parent is chosen iid with equal probability. That is, each of the possible x 's is equally likely. The growth rate depends symmetrically on the x 's.

There is a sense in which having oldsters reproduce is bad for growth. If Methuselah is chosen, you must go back 969 years or whatever, to when Methuselah was a bawling new born infant. Having to go back so far is bad for growth. This shows up in the term $\sum_{k=1}^N k$. But this slowing down of growth is across the board, affecting all the x_k symmetrically. To be more precise, if we shift probability from the states where the older guys reproduce to the states where the younger guys do, the growth rate increases.

A new motive for discounting.

We are not in Kansas, but Oz has some structure. In particular, there is now a reasonably general expectation that the rate of discounting will exceed the rate of population growth plus the mortality rate. Indeed, suppose that all survival probabilities are subject to

a perfectly correlated aggregate shock so that—

$$L = \tilde{s}\bar{L}, \text{ where } \bar{L} = \begin{bmatrix} x_1 & 1 & 0 & \dots & 0 & 0 \\ x_2 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_{\ell-1} & 0 & 0 & \dots & 0 & 1 \\ x_\ell & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

where \tilde{s} is possibly i.i.d., but in any case has a stationary distribution. It follows that the growth rate, as in the subadditive ergodic theorem, is

$$\Lambda = E \ln \tilde{s}\bar{\lambda}(x) < \ln E\tilde{s}\bar{\lambda}(x) = \ln \bar{s}\bar{\lambda}(x),$$

by Jensen's inequality, where E is taken over the stationary distribution of \tilde{s} , $E\tilde{s} = \bar{s}$, and $\bar{\lambda}(x)$ is the dominant eigenvalue of the matrix \bar{L} . Then the marginal rate of substitution between adjacent ages is

$$-\frac{dx_{i+1}}{dx_i} = \frac{\frac{d\Lambda}{dx_i}}{\frac{d\Lambda}{dx_{i+1}}} = \frac{\frac{d\bar{\lambda}}{dx_i}}{\frac{d\bar{\lambda}}{dx_{i+1}}} = \bar{\lambda} > \frac{e^\Lambda}{\bar{s}}.$$

Thus the discount rate exceeds the growth rate plus the mean mortality rate.

This result extends and makes rigorous the following scenario that might describe human populations in particular. When anthropologists look at modern hunter-gatherer societies, they typically find these are growing. The rates might not sound high, but would be enough to have us exploding into space at the speed of light if they had applied for 1.8 my. Modern HG societies might be atypical. More interestingly, they might be typical because HG societies typically grow, but occasionally decline catastrophically, due to OT phenomena—floods, pestilence, war. If the declines are equal opportunity catastrophes, with no selective impact, we should be adapted to life in the growth regime.

Small aggregate shocks.

The completely general case is rather daunting, because the subadditive ergodic theorem provides no hints on how to find the growth rate, and it is surely a complicated function

of the data. One feasible option is to consider the effect of small amounts of aggregate uncertainty. This remains of interest, since the addition of small amounts of aggregate uncertainty to a reasonably general model suffices to overturn the key results in the purely idiosyncratic case and it seems clear that there is at least a small amount of aggregate uncertainty.

Age specific mortality.

Consider first the case where the aggregate shocks affect the survival rates of the various ages of individuals. Consider then the following system. The mean matrix is

$$X = \begin{bmatrix} \bar{s}x_1 & \bar{s} & 0 \\ \bar{s}x_2 & 0 & \bar{s} \\ \bar{s}x_3 & 0 & 0 \end{bmatrix} = \bar{s} \begin{bmatrix} x_1 & 1 & 0 \\ x_2 & 0 & 1 \\ x_3 & 0 & 0 \end{bmatrix}, \text{ for } \bar{s} \in (0, 1);$$

the perturbation matrix is

$$P = \begin{bmatrix} (\tilde{s}_0 - \bar{s})x_1 & \tilde{s}_1 - \bar{s} & 0 \\ (\tilde{s}_0 - \bar{s})x_2 & 0 & \tilde{s}_2 - \bar{s} \\ (\tilde{s}_0 - \bar{s})x_3 & 0 & 0 \end{bmatrix};$$

so the overall Leslie matrix is then

$$\tilde{X} = X + \varepsilon \tilde{P}$$

Thus the aggregate environment affects the survival probabilities from each age to the next. One interpretation: Perhaps a fraction $1 - \varepsilon$ of the population has survival probability \bar{s} for each age transition; whereas a fraction ε faces aggregate survival probabilities given by \tilde{s}_i for each age i . For simplicity, the $\tilde{s}_i - \bar{s}$ are identically distributed with variance σ^2 , and mean 0, where the correlation between any pair of these is $c < 1$.

What does this imply about the tradeoff between the expected offspring levels x_1 , x_2 , and x_3 ?

We find this tradeoff by deriving the population growth rate as a second-order Taylor expansion around $\varepsilon = 0$, and then taking derivatives with respect to x_1 , x_2 , and x_3 . Define

λ as the dominant eigenvalue of X , u as its left eigenvector, and v as its right eigenvector.

Where E denotes expectation, it follows that

$$\Lambda = \ln \lambda - \epsilon^2 \frac{E(u^T P v)^2}{2\lambda^2} + O(\epsilon^3).$$

There is a need for extensive high-school algebra, which makes you happy you were paying some attention in high school long ago and far away. This shows that, for suitable functions α , β and γ ,

$$\begin{aligned} \frac{d\Lambda}{dx_1} &\simeq \alpha\lambda^2 \\ \frac{d\Lambda}{dx_2} &\simeq \alpha\lambda\bar{s} + \epsilon^2\gamma\lambda - \\ &\epsilon^2\beta\lambda [\bar{s}^2\lambda x_2 + \bar{s}^3x_3 + c\lambda^3 + c\bar{s}^3x_3] \\ \frac{d\Lambda}{dx_3} &\simeq \alpha\bar{s}^2 + 2\epsilon^2\bar{s}\gamma - \\ &\epsilon^2\beta\bar{s} [\bar{s}^2\lambda x_2 + 2\bar{s}^3x_3 + 2c\lambda^3 \\ &\quad + c\bar{s}^2\lambda x_2 + 2c\bar{s}^3x_3] \end{aligned}$$

It then follows that there is a “present bias,” for small ϵ , in that

$$-\frac{dx_2}{dx_1} = \frac{\frac{d\Lambda}{dx_1}}{\frac{d\Lambda}{dx_2}} > \frac{\frac{d\Lambda}{dx_2}}{\frac{d\Lambda}{dx_3}} = -\frac{dx_3}{dx_2}$$

There is no possibility of a preference reversal here, in the behavioral economics sense, simply because discounting is tied to age rather than time into the future. As in the previous example, discounting is no longer directly tied to the growth rate of population and the mortality rate.

THURSDAY

“ALTRUISM” TOWARDS KIN

Bergstrom, T.C. “On the Evolution of Altruistic Rules for Siblings,” *Amer. Econ. Rev.* 1995, 85, 58-81.

Bergstrom, T.C. "Economics in a Family Way," *J. Econ. Lit.* 1996, **34**, 1903-1934.

Does your mother love you? Why? How much? What exactly would she do for you? You may bask more assuredly in your mother's love than in your brother's or sister's, but let's start by considering their love for you, and you for them, because the situation there is more symmetric.

That is, economists generally assume that individuals are individualistic. We derived a slightly surprising result on how attitudes to risk might not adhere to this assumption. A much less surprising context in which even economists would admit the need for interdependent preferences is the household. It seems clear that parents care about their children, that is, would make real sacrifices for them. Parental love is not a mere VELLEITY, that is, like "You're welcome." Brothers and sisters have been known to be fond of one another too beyond mere lip service. Perhaps the same could even be said for love between parents, although the biological story there would be rather different than the one we will spin today.

The interactions between members of the same household have practical significance in terms of modelling demand, for example. Biology implies that, rather than taking some ad hoc formulation of interdependence, we can derive this.

In biology, this topic goes by the name of "kin selection." This concerns the evolutionary basis for altruism between genetic relatives. A gene or genes that promote altruism towards relatives will spread if the benefit to the recipient is large enough relative to the cost to the donor. These ideas are eloquently expressed by Richard Dawkins in the "The Selfish Gene," for example.

Bergstrom makes these ideas accessible to economists. More importantly, he shows that the traditional biological approach to kin selection, due to Hamilton, is inadequate to generate the correct predictions in general strategic interactions between relatives. Bergstrom makes a contribution to biology too, that is.

This paper should be contrasted with the large swathe of papers in evolutionary game

theory, which are more familiar works to most people. The vast majority of these papers take evolution merely as a metaphor for low rationality learning. Bergstrom's paper is a literally biological formulation of a strategic interaction. Genes control strategic choices and rise or fall in response to the payoffs they obtain. There is even sex in Bergstrom's paper; simple mechanical purely mathematical sex, maybe, but real sex, nevertheless, as I promised.

GAMES BETWEEN SIBS

Consider a symmetric game between full siblings. It is assumed that individuals can tell sibs from nonsibs—maybe just because sibs are the ones also hanging around mum.

S = set of pure strategies, $x \in S$.

$\pi(x, y)$ = payoff from choice of x if sib chooses y

$\pi(y, x)$ = sib's payoff in this situation

Choice here is nonconscious—actions are determined by one's "genotype." For simplicity, payoffs are the probability of survival to reproductive age. Mating is random, which your mother wouldn't like, but monogamous, which she would, at least if she is like mine. Altogether, that is, individuals either die, with an endogenous probability, or have a fixed number of offspring.

DIPLOID SEX

There is assumed to be a single "locus," and each individual has two alleles, two possible kinds of the gene, that is, at that locus. This is realistic, with the single overwhelming exception that we have a vast number of genes, 100,000 more or less, and more alleles. The individual's genotype here is then the combination of two alleles that he/she has.

Diploid sex then chooses precisely one allele at random from dad's genotype and precisely one allele at random from mum's. If dad is ab and mum is AB , for example, then junior will be aA , wp $1/4$; aB wp $1/4$; bA wp $1/4$; and bB wp $1/4$. In the application here, there will be just two alleles a and A , say, so the possible genotypes are just aa , aA , and AA .

We assume that this genotype determines choice. There is high degree of abstraction involved in supposing that a complex characteristic like strategic behavior is controlled by a single locus, or even by multiple loci, but some insight can nevertheless be gained from this.

The bottom line will be: You will be more benevolent towards a sib than you would towards a stranger. A more accurate description of what is genetically determined might be fraternal or sororal love. Perhaps a still more accurate description might be a tendency towards such intersibling love among those who all hang round the same mum.

It is assumed that aa choose $\bar{x} \in S$ and that AA choose $x \in S$, $x \neq \bar{x}$. What then does aA choose? There are two salient possibilities—either aA chooses \bar{x} , in which case A is recessive; or aA chooses x , in which case A is dominant, because A determines the outcome. For simplicity, we focus on the second option, where a single new allele, A , is enough to make the individual switch choices. Matters will be more delicate with a recessive mutant, because this will focus on the unlikely event that two mutants are paired when there are not many of them to begin with.

ASEXUAL REPRODUCTION

This is as in the recent case of the Komodo dragon or two in London, although suppose parthenogenesis here cannot be switched on or off, as the dragons may do. Suppose every individual who survives has two children who play the symmetric game. There is essentially a single gene, inevitably inherited from mum. Suppose that \bar{x} is such that $\pi(\bar{x}, \bar{x}) > \pi(x, x)$ for all $x \in S \sim \{\bar{x}\}$. Then everyone chooses \bar{x} in the only stable equilibrium.

Consider a population where everyone chooses \bar{x} in which a single mutant child arises. In the interesting case, $\pi(x, \bar{x}) > \pi(\bar{x}, \bar{x})$, so this individual does well right off the bat. However, all of this mutant's children play x against other descendants who also play x , so they all do badly, getting only $\pi(x, x)$. Hence this mutant line grows less rapidly and dies out.

THE GOLDEN RULE (IMMANUEL KANT)

Choose an action that would be best for you, assuming your sib matches your choice.

This is a matter of looking along the diagonal only. This would “solve” the prisoner’s dilemma, for example. This is a dangerous heresy, exactly what you shouldn’t think, from the conventional point of view. Not that you are thinking here at all, of course.

Sex makes matters more complex, and less cooperative, inevitably enough.

DIPLOID SEX

Consider a monomorphic population, with genotype aa , who play \bar{x} against siblings. Suppose a mutant allele A arises such that AA and aA play $x \neq \bar{x}$ against sibs. Since the A mutant must be rare at first, and mating is random, nearly all the A ’s must initially arise in aA combinations. Furthermore, these aA ’s will typically mate with aa ’s. The crucial issue then is: How well does the allele A do relative to the allele a in this situation? Or equivalently: How well does an aA genotype do in terms of producing viable aA descendants relative to how well an aa genotype does in producing aa descendants? The aa alleles play each other with overwhelming probability.

Suppose then that one parent is aA and one is aa . With sex, offspring are aA and aA , wp $1/4$; aA and aa , wp $1/2$ and aa and aa , wp $1/4$. The viable aA offspring payoff for a aA genotype parent is then (twice) $(2/4 = 1/2)\pi(x, x) + (1/2)\pi(x, \bar{x}) \equiv V(x, \bar{x})$. At the same time, the payoff to aa is (twice) $\pi(\bar{x}, \bar{x}) \equiv V(\bar{x}, \bar{x})$, since aa plays aa , with probability near 1.

A necessary condition for \bar{x} to be stable equilibrium is then that $V(x, \bar{x}) \leq V(\bar{x}, \bar{x})$, for all $x \in S$. This is simply the requirement that (\bar{x}, \bar{x}) be a symmetric Nash equilibrium of the game (S, V) . A sufficient condition for stability is that $V(x, \bar{x}) < V(\bar{x}, \bar{x})$, for all $x \in S \sim \{\bar{x}\}$.

V is the semi-Kantian payoff; it is an average of the Kantian $\pi(x, x)$ and the Nash $\pi(x, \bar{x})$.

SEMI-KANTIAN GR

Act towards your brother as if, wp 1/2, his choice mimics your own. (But with probability 1/2, he ignores your choice, as usual.)

Although the action \bar{x} is, in equilibrium, the only action taken by anyone, this choice is derived from a counterfactual experiment that supposes it is vanishingly rare, so the probability of it arising a sibling, given it arises in you, is 1/2. This is what is needed to make \bar{x} immune to the threat of invasion by mutants, where these mutants must be rare to start with.

INCLUSIVE FITNESS

This is the conventional biological notion of fitness, in circumstances where this should be extended to incorporate a component reflecting your relative's fitness, as well as your own. The weight you should give a relative's fitness is determined by the degree of relatedness, r . This r is probability that the relative shares a rare gene that you have, just as in Bergstrom's semi-Kantian GR.

For full siblings, $r = 1/2$. This is based upon monogamy—with less strict monogamy, and paternal uncertainty, r will be less. The relatedness of you and your mother is also 1/2; that of uncle and you is 1/4; that of cousins is 1/8. Again staying away from any suggestion of hanky-panky in your family tree.

This degree of relatedness is often misrepresented, even by biologists, as the raw proportion of genes you share with an other individual. But humans share over 98% of their genes with chimpanzees and may not favor chimps with the full force of the GR.

This notion of inclusive fitness in the present context, with full sibs, is $H(x, \bar{x}) = \pi(x, \bar{x}) + (1/2)\pi(\bar{x}, x) \neq V(x, \bar{x})$, in general.

HAMILTON'S RULE

Love thy sib one half as well as oneself. "If $r < c/b$, don't even ask."

In general, a NE under V need not be a NE under H and vice versa. That is, H is the wrong payoff function.

Hamilton seems to have had a limited class of games in mind when deriving his rule, but biologists since Hamilton have been inclined to overapply H .

ADDITIVE GAMES

For some odd notational reason, known only to Bergstrom, actions are now denoted $a \in A$. Payoffs are additive as follows. $b(a)$ = the benefit to the sib of the action a ; and $c(a)$ = cost to you of this action. That is, the total payoff is $\pi(a, a') = b(a') - c(a)$, so that $\pi(a', a) = b(a) - c(a')$. That is, the benefit to the sib is independent of the sib's choice; the cost to you is also independent of this choice.

PROPOSITION For additive games, the set of symmetric NE of the game $\{A, V\}$ is the same as the set of NE of the game $\{A, H\}$. That is,

$$SNE(\{A, V\}) = SNE(\{A, H\}).$$

What can be said about more general games? What is crucial is the complementarity/substitutability between the two players' choices.

DEFINITION π has symmetric complementarity iff

$$\pi(x, x) + \pi(y, y) \geq \pi(x, y) + \pi(y, x)$$

It has symmetric substitutability iff

$$\pi(x, x) + \pi(y, y) \leq \pi(x, y) + \pi(y, x).$$

PROPOSITION Under SC,

$$SNE(\{X, V\}) \subseteq SNE(\{X, H\}).$$

With SS,

$$SNE(\{X, V\}) \supseteq SNE(\{X, H\}).$$

For an additive game,

$$\begin{aligned}\pi(x, y) - \pi(y, y) &= -c(x) + c(y) \\ &= \pi(x, x) - \pi(y, x).\end{aligned}$$

ROUSSEAU'S STAG HUNT

Two hunters set out to hunt a stag. Along the way, each of them runs across a rabbit. If an individual sticks to the original stag hunt plan, the payoff to him is $R > 1$. If you deviate and hunt down the rabbit, your payoff is 1. If the other guy chases the rabbit, but you stick to the stag hunt plan, you get 0.

Thus the game is

	<i>S</i>	<i>H</i>	
<i>S</i>	R, R	$0, 1$	
<i>H</i>	$1, 0$	$1, 1$	

There are two strict NE— (S, S) and (H, H) , when this game is played as usual between unrelated individuals. The point of the game is that there is a coordination issue. If you believe that the other guy will defect from the stag hunt, so should you.

This game satisfies SC, because the sum of the diagonal terms exceeds the sum of the off diagonal terms, $1 + R > 1$. To precisely analyze the game between two brothers, we need the semi-Kantian preferences $V(x, \bar{x}) = (1/2)\pi(x, x) + (1/2)\pi(x, \bar{x})$ —

	<i>S</i>	<i>H</i>	
<i>S</i>	R, R	$R/2, 1$	
<i>H</i>	$1, R/2$	$1, 1$	

since $V(H, S) = \pi(H, H)/2 + \pi(H, S)/2 = 1$ and $V(S, H) = \pi(S, S)/2 + \pi(S, H)/2 = R/2$. If $R > 2$, this game has a unique NE (S, S) . Thus kin selection might solve the coordination problem entirely.

What would Hamilton say here?

	<i>S</i>	<i>H</i>
<i>S</i>	$(3/2)R, (3/2)R$	$1/2, 1$
<i>H</i>	$1, 1/2$	$3/2, 3/2$

since $H(H, S) = \pi(H, S) + \pi(S, H)/2 = 1$ and $H(S, H) = \pi(S, H) + \pi(H, S)/2 = 1/2$. For all $R > 1$, this game has two strict NE, so the function H never solves the coordination problem.

To my mind, a key contribution that Bergstrom makes here is conceptual. That is, the semi-Kantian GR involves a conceit that your action will induce the same action in your sibling w.p. $1/2$. This is a conceit that is usually considered dangerously fallacious in game theory. This philosophy also contrasts sharply with Hamilton's application of inclusive fitness. With this, you have no such conceit, but simply consider your sib's payoff determined in the usual way. You add this payoff with weight $1/2$ to your own.

FRIDAY

TIME PREFERENCE (AND INTERGENERATIONAL TRANSFERS)

Rogers, A. "Evolution of Time Preference by Natural Selection," *Amer. Econ. Rev.* 1994, 84, 460-481.

This was one of the first papers to take seriously the argument that economic preferences were shaped by natural selection. In particular, Rogers looks at intertemporal preferences and, more particularly, the pure rate of time preference. He is led to the startling claim that the real rate of interest should be around 2% per annum. When you think hard about what he is doing, there are good reasons to be skeptical about key elements of his story. In particular, there is no strong prediction that flows from such a model concerning the real rate of interest. Even more basically, Rogers account of intertemporal preferences is not adequate. But Rogers does deserve a lot of credit for the first attempt at this.

The best settings for the argument are perhaps agricultural societies. These provided a

range of savings decisions that resemble modern ones, although borrowing might have been harder. That is, it is possible to store grain, for example, or to replant it, and reap more than you sowed. Agriculture has been going for about 10,000 years at the most, which is a modest although non-negligible span of time from an evolutionary point of view. That is, genetic variants with very sharp advantages or disadvantages might have been driven to fixation, but there is a range with more modest advantage or disadvantage that might not.

The key phenomena are the intergenerational transfers that are characteristic of human beings. These are remarkably large for hunter-gatherers, for agricultural societies, and for us. Kotlikoff and Summers, JPE 1981, claim that the vast majority of savings are intergenerational transfers rather than consumption smoothing for the saver.

Rogers' model considers the evolution of a sexually reproducing species in which resource availability plays a key role. Each individual uses resources to increase his or her probability of survival until the next period, but, as a simplifying assumption, has exogenous fertility at every age at which he or she is alive. Each parent can save resources for the ultimate benefit of an offspring. His or her interest in doing so is attenuated diluted by the degree of relatedness. Rogers derives intertemporal preferences from this, and the rate of time preference and interest, in particular.

A key biological element that Rogers wishes to import is the degree of relatedness, r . The situation here will be rather like the additive game discussed by Bergstrom, so that Hamilton's rule is appropriate here, as Rogers claims.

Another key biological element that Rogers wishes to press into service here is the concept of "reproductive value." He appeals to the definition due to RA Fisher—

$$v(x) = \frac{\sum_{y=x}^{\infty} e^{-\rho y} l(y) m(y)}{e^{-\rho x} l(x)},$$

where $l(y)$ = prob. of surviving until age y , $m(y)$ = expected offspring at age y , ρ = rate of population growth. This is then the expected discounted fertility of an individual of age

x , conditional on being alive at age x , and where the discount rate is the population growth rate. This concept is appropriate in simple models, but it will not be enough here.

He then claims that an allele is selectively neutral if

$$v_D(x_1)\Delta P(x_1) + e^{-\rho\tau}rv_R(x_2)\Delta P(x_1 + \tau) = 0$$

where $P(x)dx$ (not just $P(x)$) is the probability of survival from x to $x + dx$ v_D and v_R are the RV's of the donor and recipient, x_1 and x_2 are their respective ages, and τ is the time lag. It follows that

$$MRS_P = -\frac{\Delta P(x_1 + \tau)}{\Delta P(x_1)} = \frac{v_D(x_1)}{re^{-\rho\tau}v_R(x_2)}.$$

He is careless with mathematical details here. But a conceptual problem with this is: Why isn't the effect of the benefit discounted by the probability that the recipient is alive to get it?

If it is assumed that the survival rates depend on consumption and age as $P(x, \kappa(x))$, he then derives the marginal rate of substitution—

$$MRS_\kappa = -\frac{\Delta\kappa(x_1 + \tau)}{\Delta\kappa(x_1)} = \frac{v_D(x_1)}{re^{-\rho\tau}v_R(x_2)} \frac{P_\kappa(x_1, \kappa(x_1))}{P_\kappa(x_2, \kappa(x_2))}.$$

This sneaks in the assumption that income can vary only with age and is not growing at each age over time. In equilibrium,

$$MRS_\kappa = e^{i\tau},$$

where i is the real rate on interest. He then restricts attention to the key case that the donor and recipient have the same age so $x_1 = x_2$. It follows that

$$\frac{e^{\rho\tau}}{r} = e^{i\tau},$$

so that given that $r = 1/2$, $\tau = T$, the intergenerational time, and $\rho = 0$, as it must have been on average for 1.8my,

$$i = \frac{\ln 2}{T}$$

If $T = 28.9$, $i = 0.024$, in reasonable agreement with the data.

Robson, A and Szentes, B. “Evolution of Time Preference by Natural Selection: Comment” *American Economic Review* forthcoming.

EXAMPLE

Individuals are born at age $i = 0$ and live to at most to age $i = 3$. Let $P_i(c_i)$ denote the survival probability from age i to age $i + 1$, when consumption is c_i , for $i = 0, 1, 2$. These functions are continuously differentiable, strictly increasing, and strictly concave, and have infinite derivatives at $c_i = 0$. It is assumed that $P_0 \equiv 1$, so that individuals surely survive until the age of one. This is to generate a demand for savings, since any transfer that a new born received would then be saved.

At each age, $i = 1, 2, 3$, each individual is paired with another individual of the same age and the couple has two offspring. There are no Lolitas and no Harolds and Maudes, either. For the sake of simplicity, we suppose that the transfer/savings decision is controlled by a single gene. Individuals are haploid. Haploid sexual reproduction means that each offspring acquires either the paternal allele or the maternal allele, each with probability $1/2$.

Individuals have zero endowment at ages $i = 0, 1, 3$ but have one unit of endowment at age $i = 2$. This assumption produces a central need for savings that illuminates the issues that arise in Rogers’ model. At age $i = 2$, the individual can transfer part of her endowment, in the amount s , say, to her two newborn offspring. Later, we consider another transfer, which will be a “same age transfer,” as in Rogers. Any transfer from each parent is shared equally between each of the two offspring. Although only one of these is also a mutant, on average, it is not possible to distinguish between mutant and non-mutant.

The new born offspring will certainly choose to save all income. It is assumed that the technology for intertemporal transformation is linear, with an exogenous interest factor of R . It follows that each offspring then receives $Rs/2$ at age one, as the return from each parent’s transfer.

The question is: What value of s would arise in equilibrium? Consider a population with an allele that selects savings \bar{s} . A small proportion of the population is then replaced by an allele that saves s . The question becomes: For what value of \bar{s} is it true that no matter what s is, the mutant allele saving s grows no faster than the original allele saving \bar{s} ?

Define $\mu_i \geq 0$, for $i = 1, 2, 3$, as the proportion of i -year-olds in the mutant gene population, ignoring newborns, so that $\sum_{i=1}^3 \mu_i = 1$. We have the following survival probabilities—

i) $P_1(0) > 0$ is the probability of survival of a mutant (or any) one year old who had one or three year old parents, and hence got zero transfer.

ii) $P_1(R^{\frac{\bar{s}+s}{2}})$ is the the probability of survival of a mutant one year old who had two year old parents, one mutant and one not, and hence got a total return of $(R^{\frac{\bar{s}+s}{2}})$. Since the mutant gene is rare, essentially all mutant individuals with two year old parents had one mutant and one non-mutant parent. Total resources, $\bar{s} + s$, are shared equally between the two offspring.

iii) $P_2(1 - s)$ is the survival probability of a mutant two-year-old who saved s .

Finally, define $F = [(\mu_1 + \mu_3) P_1(0) + \mu_2 P_1(R^{\frac{\bar{s}+s}{2}})]$, as the overall probability of survival of a mutant individual from age one to two. The mutant population converges to steady state growth with growth factor g satisfying

$$g^3 = g^2 + (g + P_2) F.$$

The steady state population proportions μ_i 's can be obtained. The following equation then implicitly determines the growth factor of the mutant population:

$$g^3 = g^2 + (1 + P_2) P_1(0) + (g - 1) P_1(R^{\frac{\bar{s} + s}{2}}).$$

The g as determined by this equation is maximized if and only if the right hand side is maximized for parametric choice of g . For an interior solution to this maximization problem,

the first-order condition is:

$$P_2'(1-s)P_1^1 = \frac{P_1'(R\frac{\bar{s}+s}{2})(g-1)R}{2}.$$

This characterizes the choice of s that yields the fastest limiting growth rate of the mutant allele, when these are still a small fraction of total population of alleles, most of which choose \bar{s} . The right hand side is strictly concave in s . If $s = \bar{s}$, the growth rate of the mutant alleles must match that of the original type. Hence, if the maximum occurs anywhere except at $s = \bar{s}$, the mutants can invade in the sense of initially growing faster than the original population. But if the maximum is at $s = \bar{s}$, then any nontrivial mutant with $s \neq \bar{s}$ will be outstripped right away by the original allele.

Altogether, that is, the evolutionary equilibrium value of savings satisfies

$$\frac{2P_2'(1-\bar{s})P_1(0)}{P_1'(R\bar{s})(g-1)} = R,$$

where the left hand side is the appropriate marginal rate of substitution, MRS. However, this MRS depends in a complex way on the various survival functions, and is not simply the factor of $2g$ as would be claimed by Rogers in this notation.

The factor of 2 arises here from the assumption that transfers must be made blindly and therefore equally to each of the two offspring, one of whom is a fellow mutant, on average, but one of whom is not. This is a particular derivation of Hamilton's rule.

Reproductive Value

The equilibrium saving behavior is a solution to a problem where the two-year old individual maximizes the total discounted reproductive value of those who are affected by her savings. Indeed, this total reproductive value is the utility function yielding equilibrium behavior.

In general, the reproductive values should be defined as the relative shares attained in the population, in the limit as the time into the future tends to infinity. They form a right

eigenvector of the Leslie matrix, precisely. In the present example, a heuristic approach to finding the correct reproductive values is as follows. Let $v(0_1)$ and $v(1_1)$ denote the reproductive values of individuals of ages 0 and 1, respectively, who receive no transfer; and $v(0_2)$ and $v(1_2)$ as the reproductive values of individuals of ages 0 and 1, respectively, who receive transfer $(\frac{\bar{s}+s}{2})$. In addition, let $v(2)$ and $v(3)$ denote the reproductive values of individuals of age two and three, respectively. Considering the descendants of each type one period ahead, these values must satisfy the recursive relationships

$$\begin{aligned} v(0_1) &= \frac{v(1_1)}{g}; v(0_2) = \frac{v(1_2)}{g} \\ v(1_1) &= v(0_1) + \frac{P_1^1 v(2)}{g}; v(1_2) = v(0_1) + \frac{P_1^2 v(2)}{g} \\ v(2) &= v(0_2) + \frac{P_2 v(3)}{g}; v(3) = v(0_1), \end{aligned}$$

It can be shown that the equilibrium behavior can be derived from each two-year old individual solving the following problem

$$\max_s P_2(1-s)v(3) + \frac{P_1(R\frac{\bar{s}+s}{2})v(2)}{g}.$$

The maximand is the expected reproductive value of the parent, as influenced by the choice of s , plus the expected discounted reproductive value of the age one recipient, also as influenced by the choice of s . Thus these reproductive values generate the utility that underpins the evolutionary equilibrium. Rogers must be given credit for this basic insight. One of our disagreements stems from the appropriate notion of reproductive value to be used in such an expression.

Indeed, the problem as formulated by Rogers entails too little savings. Fisher's notion of reproductive value is inadequate when offspring are distinguished, in particular, by their inheritance, as here.

Zero Same Age Transfer

The previous analysis assumed that the only transfer possible was from a two-year-old parent to her current newborn. Suppose now that this two-year-old parent can also put aside resources for her current one-year-old offspring. These resources must be saved for one period, and are then given to the offspring. Thus, this offspring will be the same age when she consumes the transfer as her mother was when she gave it. This combination is then a same-age transfer, as in Rogers.

We show that, in any evolutionary equilibrium, zero additional transfers like this will be made, given that $R \leq 2g$. We assume that there exists a competitive market for these savings by current one-year-old offspring. This market returns a fixed interest factor, R' . Since only a proportion of P_1^1 of the intended recipients survives until the age of two, and the market is competitive, $R' = R/P_1^1$. This additional transfer is zero in equilibrium if and only the following holds for each two year old parent—

$$0 \in \arg \max_{s'} V(s') \equiv P_2 (1 - \bar{s} - s') v(3) + \frac{P_1^1 P_2 (1 - \bar{s} + R' \frac{s'}{2}) v(3)}{g}.$$

The function V , from a biological viewpoint, is “inclusive fitness”; from an economic viewpoint, it is the utility function that underpins behavior. That is, it is the relevant part of intertemporal preferences for this two-year parent contemplating transfers to her current one year old offspring. The first term is the individual’s own expected survival probability to age three multiplied by her age-three reproductive value. The second term is the survival probability of the current one year old offspring from age one to age three multiplied by the present value of her age-three reproductive value.

The required condition holds simply because $R \leq 2g$. The intuitive reason that an interior solution does not generally exist for the same age transfer is that the evolutionary criterion considers the net income position of all two-year-olds. Such transfers appear both as a debit, and as a credit, the latter multiplied by the appropriate interest and growth

factor, in the single argument of the same survival function.

The most telling argument against the interior solution assumed by Rogers is as follows. In general, in a realistic model of unpleasant complexity, there will be many same age transfers possible. A 20 year old woman might save for new born to receive 20 years from now; or she might save when she is 25 for her new born to receive 25 years from now. According to Rogers, the first savings/transfer implies an interest rate of $\frac{\ln 2}{20}$ the second implies a rate of $\frac{\ln 2}{25}$. But it can't be both. At least one of these must be a corner solution; possibly both.

Zero Population Growth

Rogers assumes zero population growth, on the basis of the argument that the average growth rate for humans must have been close to zero over the two million years of our history. Our example entails a growth rate necessarily greater than zero. But nothing crucial depends on a positive growth rate.

A Generalized Example

One might be concerned that our conclusions hold only because we restricted the possible transfers. Indeed, in a generalized model, the basic philosophy might be that all individuals may transfer any amount contemporaneously to any living relative, and all individuals may save any amount. Individuals cannot borrow, however.

We have reanalyzed the present example in this light. Although the analysis is more complex, the essential points made by the above version of the example are preserved. In addition to transferring resources to their new born offspring, two year old parents now make positive transfers to their one year offspring, who immediately consume it. Two year old parents also save for one period, with the return from this being transferred to the new born offspring they have at age three. These offspring reinvest this transfer and finally consume it at age one. It remains inappropriate to use the simple notion of reproductive value used by Rogers. Furthermore, all of the other myriad contemporaneous transfers and savings that

might be made in principle in this generalized example may well be chosen to be zero in equilibrium. Most significantly, that is, there are now a number of combined transfer and savings options that amount to same age transfers, and all of these may optimally be chosen to be zero.