A SIMPLE MODEL OF PLACEBO LEARNING WITH SELF-REMITTING DISEASES[1]

Daniel Carpenter
Institute for Quantitative Social Science
Department of Government
Harvard University

November 4, 2005

[1]Preliminary. Comments and corrections welcome. A more general version of this paper was presented at the RWJ Core Seminar, Harvard Medical School, March 2004; and at the Center for Advanced Study in the Behavioral Sciences, Stanford, May 2004. I thank Robert Ader, James Alt, Nicholas Christakis, Erica Field, Richard Frank, Ted Kaptchuk, Dan Kessler, Arthur Kleinman, Gary King, Kenneth Kendler, and Robert Scott for very helpful comments and insightful discussions. All errors, omissions and interpretations are mine.

Abstract

I develop a simple stochastic model of inference and therapeutic utilization in the presence of placebo effects, when the underlying medical condition may be self-remitting. In the model, expectations generate a "felt" health state which can mimic the medically cured health state even when the treatment in question has no real curing power. This effect may be augmented by self-limitation of the medical condition for which the treatment is utilized. A human agent then applies Bayes' rule to the felt history as if it were generated pharmacologically. A more sophisticated agent knows of placebo effects but does not know the precise extent to which they contribute to curing. I describe the bias that attends inference and the under - or overutilization of therapies under such a model. A central result of the model is that human placebo learning is generally subject to greater bias in estimating treatment efficacy when diseases are self-limiting. Human agents may commit several types of decision errors under placebo learning. They may continually choose a more costly (expensive, hazardous) treatment when a less costly one would work as well, or they may continually use inferior treatments for life-threatening illnesses. When diseases are self-limiting, both these types of error are *more* likely when the human agent has high initial beliefs about the treatment. Possible applications of the model include the patent medicine industry, the robustness of markets for herbal and nutritional supplements, and the contemporary stability of counterfeit drug operations.

There is now increasing consensus that expectation-induced placebo response is a genuine neurological phenomenon. This agreement is attributable in part to experiments where placebo analgesia has been reversed by the opioid antagonist naloxone (Levine et al 1978; Amanzio and Benedetti 1999), and in part to recent neuroimaging studies (Petrovic et al 2002; Wager et al 2004). Yet the salience of placebo effects *in clinical and behavioral settings* remains under criticism and question for at least two reasons. First, it remains possible that self-limitation of disease – alternatively, "spontaneous remission" or "natural history" – can explain clinical and behavioral phenomena that many attribute to placebo mechanisms (Hrobjartsson and Gotzsche (2001)). While inclusion of a "natural history" or "no treatment" arm can validate placebo inferences, these remain rare in clinical trials. A second issue concerns the mechanism of placebo response, and whether it draws upon expectations or upon conditioning or some combination of the two (Amanzio and Benedetti 1999; Stewart-Williams and Podd 2004).

Might self-limitation stand not as a substitute for placebo effects, but as a complement? Is it possible, in other words, that the false inferences induced by self-limiting diseases could work together with placebo effects to further complicate human learning about the efficacy of medical and pharmaceutical treatments? There would appear to be no theoretical or empirical research that addresses this question.

In this paper I develop a model of "placebo learning" and show how the learning bias induced by placebo effects may, at least theoretically, be greatly exacerbated by self-limiting conditions. The model of placebo learning is the first stochastic analytic model of its kind to my awareness.[1] Placebo effects are a subject of considerable study in almost every field of medicine, and in many fields of psychology, but there exist few (if any) mathematical models dedicated to their analysis. This is rather surprising, given the ubiquity of placebo effects in modern medicine and their intensive study in neuroscience (Amanzio and Benedetti 1996; Guess, Kleinman, Kusek and Engel 2002).

The model is premised upon the following thought experiment: Consider the possibility that a human agent, upon becoming ill, utilizes an inert treatment and that (1) the treatment itself has little or no therapeutic value, but that (2) her illness naturally recedes and would have done so in the absence of the treatment. If the agent is not aware of facts (1) and (2), she might wrongly attribute therapeutic power to the treatment, and might then place greater belief in the efficacy of the treatment than she would if she had (randomly) declined it at the time when she became sick. If expectations-mediated placebo analgesia is operative, the agent's inflated expectations may then exercise *additional* curative power themselves, and valid inference regarding the "true" (pharmacological) efficacy of the treatment may be further hindered.

This question is important not merely for scientific inference but also for the operation of

[1]The closest analog to this model appears to be the computational model of Redish (2004). I discuss the potential nesting of Redish's model in mine, and the potential nesting of mine in his, below.

pharmaceutical markets. Even though drugs approved for marketing in the United States and other nations have usually proven their therapeutic value versus a placebo regimen, it remains possible that many of the observed clinical responses to these products *after* regulatory approval and market entry are nonetheless due (partially or wholly) to placebo effect. In other words, controlling for placebo effects in clinical research does nothing to prevent the operation of placebo effects in the health care system. Despite the ubiquity of these phenomena in medical and other settings, some interesting questions remain unexplored, especially by medical and social scientists.

◦ Do placebo effects potentially complicate inference by patients and their doctors about the efficacy of medical treatments? Might this be true even if the existence of placebo effects is known for a given therapy?

◦ How might placebo effects interact with the cyclic or self-limiting nature of diseases to influence human inference about the efficacy of medical or pharmaceutical treatments?

◦ Might placebo effects complicate research-based inference on the returns to medical and pharmaceutical investment, advertising, and other expenditures?

To address these and other questions, I offer here a simple mathematical model of "placebo learning," or of how otherwise rational human beings might learn from their own experience about the efficacy of medical or pharmaceutical treatments in the presence of placebo effects. I model placebo learning as a Bayesian inference process in which the agent's observed health is directly (but unknowingly) affected by the expectation of treatment. In the model, "expectation enhancements" generate a "felt" or experienced history that mimics the health state that would have been generated pharmacologically, and the agent uses Bayes' rule to make inferences from the felt history as if it were the "true" one.[2] I describe the erroneous long-term inferences that can be generated by a simple Bayesian inference process in the presence of placebo effects. I then discuss how a human agent, fully rational except for her unawareness of placebo effects and their influence upon her learning, chooses over time from a battery of medical treatments (drugs, for example), by solving a multi-armed stochastic dynamic programming problem.

For purposes of this paper I define a placebo effect as an observable improvement in an agent's health state that is generated entirely by subjective expectations of a "treatment." This improvement prevails in the absence of the actual treatment but in the presence of a consumed experience (the "placebo") that has similarities to the treatment in any or all respects *but* the treatment's curative (e.g., pharmacological) mechanism. The clinical and scientific literature on such effects is enormous (see Harrington 1997 and Kleinman, ed, 2002, for a summary and partial bibliography). However, mathematical and computational analysis of placebo effects and their possible influences upon human behavior have not been directly attempted.

[2]See Mullainathan (2003) for a similar substitution of perceived for real histories in the context of a memory-based model of bounded rationality.

The paper also offers a simple technology for simultaneously embedding the role of expectations and the role of conditioning in placebo learning. In this respect, I emphasize that the neurobiology of placebo effects is not the focus of analysis here; rather it is the potential influence of placebo effects upon learning algorithms and dynamic programming that is assessed. From recent studies, it would appear that one key to placebo learning lies in the initial expectations that the agent (patient) attaches to the prospect of treatment and the essentially unobservable character of the treatment's curing mechanism (Amanzio and Benedetti 1999; Wager et al, 2004; Finniss and Benedetti 2005).[3] It would seem a crucial property of placebo learning that most human agents cannot directly test the curative power of the consumed treatment and cannot (without significant cost) ascertain whether it is a genuine or effective treatment.

1. Inference under Placebo Contamination

1.1 Structure

An infinitely-lived single agent ("Patient") faces an infinitely-repeated problem and experiences one of two exclusive health states, illness $X_t = X^{ILL}$ or wellness $X_t = 0$. I assume that the patient starts healthy and that at some time $t \equiv 0$ she eventually experiences sickness. Sickness can be of two types, *chronic stable* (under which the sickness persists unless treated and perhaps even if treated), or *cyclic* (under which there is reversion to the healthy state after some elapsed interval of time).

Under transition to the sick state, the patient may be induced to try up to $Q + 2$ treatments, where no more than one treatment may be undertaken in any period. Throughout treatments (drugs) are indexed by i and diseases by j. The agent's action in any given period is to choose actions σ_t over these treatments. For one treatment, called the "default" ($\sigma_t = 0$), the agent is perfectly informed about the value of a treatment. Without loss of generality, we say that the agent begins with an unknown "incumbent" treatment $\sigma_t = 1$. There is also a battery of alternative treatments $\sigma_t = q$, $1 \leq q \leq Q < \infty$. For all but the default treatment, the agent is imperfectly informed about a treatment's efficacy and can learn it only from experience (utilization) of that treatment. For the incumbent treatment, efficacy may be denoted $\gamma_{1j} \in \Gamma \equiv [0, 1]$. It is convenient to think of γ_{1j} as a curing probability, or in a clinical context, the probability in a sequence of Bernoulli trials that the patient "responds" to medication at time t. The agent has prior beliefs over γ_{1j} that are represented by a Beta distribution; thus, $\gamma_{1j} \sim \beta(m, n)$. The Bernoulli properties of the state space and the Beta-distributed posterior offer a highly generalizable and flexible framework

[3] Unobservable, that is, to the agent who makes dynamic utilization decisions.

for analytic solution of the model (Gittins 1979).[4] Still, some complexities require computational simulation, which I discuss below.

One and only one treatment may be used in any given period, but in any period, all treatments are potentially available for use. The human agent, knowing that all (uncertain) treatments have Beta-distributed efficacy, chooses that treatment for which the prior (initially expected) distribution of curing efficacy is highest. If this arm is eventually abandoned, then the agent's best alternative is that treatment with the next highest prior, then so on down to the point where next highest prior is that of the default, which if tried once is kept forever. The problem is benchmarked by the availability of this default treatment whose curing probability is known with certainty as $\alpha \in (0, 1)$. The known quality of α then becomes a decision-theoretic "certainty equivalent" against which the value of continued utilization can be assessed. The patient can also forgo treatment, leaving her sickness "untreated" entirely.

We represent pharmacological curing by any given treatment by a Bernoulli variable G_{ijt} (with $i = 1$ for the incumbent). Let $G_{ijt} = 1$ represent the outcome that the individual with disease j, who utilizes treatment i, exhibits a healing response at time t, with 0 scoring no response. G_{ijt} is a Bernoulli variable whose mean is γ_{ij}. Then the individual's pharmacologically-determined health state may be represented by $Y_{ijt} = (1 - G_{ij})\, X_{ij}^{ILL}$, such that Y_{ijt} is the human agent's health state in period t, after onset of disease j and utilization of the ith treatment.

1.2 Inference

If placebo analgesia were impossible and sickness states were not self-remitting, the agent's observed history would be sufficient for optimal (Bayesian) inference as to the pharmacological efficacy γ_{1j} of the incumbent therapy. For any stopping time τ, the optimal estimate is (DeGroot 1970)

$$\widehat{\gamma}_{1j,\tau+1}^{*} = 1 - \frac{(n-m) + \sum_{t=0}^{\tau} Y_{1jt}}{n + \tau X^{ILL}} \tag{1}$$

The central problem in placebo learning is that the human agent observes not Y_{ijt}, but something else, something that is "contaminated" by her own expectations about the incumbent treatment. Her own expectations can influence the health state that she sees.

To model this process, let us first represent healing via endogenous opioids. Let the parameter λ_{ij} $(0 \leq \lambda_{ij} \leq 1)$ represent "suggestibility," or the degree to which the agent's ailment is susceptible to curing by expectancy. Suggestibility of pain and other symptoms of disease to expectations lies

[4] One alternative possibility is to consider health as a continuous stochastic process (a Wiener or Poisson process) and then to consider treatments as possible alterations of the process's first moment or limiting distribution. Work is currently underway using Wiener process realizations of health, but continuous state-spaces and continuous time both complicate the model in numerous ways, and I have decided to commence analysis of this general problem using the simple but quite flexible framework here.

at the core of placebo-effect phenomena (Amanzio and Benedetti 1999; Guess, Kleinman et al 2002; Wager et al 2004). I first consider suggestibility as an exogenous constant, to keep the focus on Bayesian learning under influence of placebo effects. However, recent evidence suggests that "endogenous opioids" (such as dopamine) exhibit phasic response patterns whose movements are substantive and important to understand, and I model this possibility below. To model healing by placebo effects, further let A_{ijt} represent the event that the individual using treatment i for disease j responds to either medical curing *or* a placebo effect at time t. A_{ijt} is a binary variable ($A_{ijt} = 1$ implies a healing response; 0 scores no response) with mean $E_{zt}[A_{ij,t+1}] = \alpha_{ijt} = (1 - \lambda_{ij})\gamma_{ij} + \lambda_{ij}\widehat{\gamma}_{ijt}^{felt}$, where E_{zt} is the expectation operator conditioned upon $T = t$ and $Z = z$. The variable $\widehat{\gamma}_{ijt}^{felt}$ is the agent's estimate of the efficacy of the incumbent treatment based upon (placebo-contaminated) self-observation. Given this estimate, we can represent the individual's experienced health state Z_{ijt} as a psychosomatically-weighted combination of response to actual treatment and placebo effect, or $Z_{ijt} = (1 - A_{ijt})\, X_j^{ILL}$.

Under placebo learning, then, the human agent observes Z_{1jt}, not Y_{1jt}. If the agent uses Bayes' rule in every other respect, but (unknowingly) substitutes Z_{1jt} for Y_{1jt} in Bayesian learning, she estimates the efficacy of the treatment as follows.

$$\widehat{\gamma}_{1j,\tau+1}^{felt} = 1 - \frac{(n-m) + \sum_{t=0}^{\tau} Z_{1jt}}{n + \tau X_j^{ILL}} \tag{2}$$

Along with the equations for A_{ijt} and Z_{ijt}, equation (2) specifies a recursive structure that creates the possibility for "self-fulfilling prophecies" in medical treatment. If initial beliefs about a treatment are (wrongly) high, expectancy-based healing can keep these beliefs artificially inflated even when the agent is learning rationally from her own medical history. So too, wrongly deflated beliefs about treatment can remain depressed. Placebo learning thus places a causal premium on what these initial beliefs are and where such expectations come from.

Beliefs: Confidence, Prognosis and Advertising. Intuitively, we can imagine initial beliefs about product quality as a function of (i) beliefs about the curability of the agent, (ii) beliefs about the curability of the disease with which the agent is afflicted, and (iii) beliefs about the match between product, disease and patient. Let the Beta variate μ represent *confidence*, the individual's faith that *she* can be cured by any given therapy. Let the Beta variate μ^j represent *prognosis*,or the individual's belief that a given disease is curable, whatever the therapy. This may be a function of messages received from one's physician ("You have three months left"; see Christakis 1999), family members and associates ("That disease took Harold quickly") or from memories (the daughter whose mother died young of heart disease might be less willing to exercise because she believes it her lot to die young). Finally one can also imagine hype, the subjectively expected probability that a match of drug i to disease j will result in a cure. As with prognosis, hype might be affected by

therapeutic messages, but unlike prognosis may also be affected by advertising. Then initial beliefs are given by $\widehat{\gamma}_{ij0}^{felt} = \inf\left[\mu, \mu^j, \mu^i\right]$. This simple functional form simply embodies the notion that, for instance, an advertising blitz directed at consumers and doctors is of little help in raising patient beliefs about the likelihood of being cured when the patient is told that she has three months to live.

We can imagine most human agents (patients, physicians) as confusing Z_{1jt} and Y_{1jt} entirely, though it is possible that a more sophisticated agent knows of the possibility of placebo effects, and I consider this possibility shortly. If Z_{1jt} and Y_{1jt} are conflated, then the observed ("felt") health state for individual using the incumbent treatment, with expectation enhancements, is $E\left[Z_{1jt}\right] = X_j^{ILL}\left(1 - (1-\lambda_{1j})\,\gamma_{1j} - \lambda_{1j}\widehat{\gamma}_{1jt}^{felt}\right)$. Again for any stopping time τ, the behavior of the "felt" estimate in the τth period can be described as

$$E_0\left[\widehat{\gamma}_{1ij,t=\tau}^{felt}\right] = \frac{m + X_j^{ILL}\left[\tau\gamma_{1j} + \lambda\left(\sum_{t=0}^{\tau-1}\psi_{1jt}^{felt}\right)\right]}{n + \tau X_j^{ILL}} \tag{3}$$

where ψ_{1jt}^{felt} is the error of the felt estimate for the incumbent treatment at time t. The long-run behavior of the felt estimate will then depend upon the convergence (if any) of the series $\left\{\psi_{1jt}^{felt}\right\}$. The difference between the optimal Bayes estimate of $\widehat{\gamma}_{1jt}^{*}$ and the felt Bayes estimate $\widehat{\gamma}_{1jt}^{felt}$ can be expressed in terms of the "observables" Y_{1jt} and Z_{1jt}.

As a benchmark case, it is possible to characterize aspects of the asymptotic distribution of estimated curing when the agent naively uses the felt history as if it were the true one, *without* self-remitting conditions. It turns out that placebo-learning can induce highly ' 'inefficient" estimates of the pharmacological efficacy of treatments, and that the "felt" estimate of efficacy converges more slowly to the true value γ_{ij} than does the optimal estimate in (1).

Comment 1. Asymptotic Consistency and Inefficiency of Bayes-Estimated Curing based on the Felt Curing History. Let the history $H_t^{felt} = (Z_1, Z_2, ...Z_{t-1})$. If $\gamma_{ij0} \neq \gamma_{1j}$, then $\forall H_t^{felt}$, $\lim E_t\left[\psi_t \mid H_t^{felt}\right] = 0$. For any history, the estimate $\widehat{\gamma}_{1jt}^{felt}$ is inefficient in that its variance is greater than that of the optimal estimate $\widehat{\gamma}_{1jt}^{*}$. *Proof.* Proofs of all comments and propositions are in the Appendix.

1.3 Inference under Self-Conscious Awareness of Placebo Effects While most patients in medical settings will not know of their own susceptibility to placebo effects, it is quite possible that some will, or (more likely) that their doctor will consider this possibility.[5] Consider then the case of a sophisticated agent, who knows of the existence of suggestibility (λ) but does not know its

[5]To the extent that incentive compatibility issues matter here, I assume that the agency problem between doctor and patient is costlessly solved. This problem is one obviously deserving of a separate modeling effort.

precise value. An example of this would be a health care provider who knows of the existence of placebo effects in a given population, and therefore knows of $\overline{\lambda}_j$ (which we consider as the population-averaged extent of psychosomaticity) but does not know the patient-specific value λ_{ij}. Until this point, in other words, the agent has proceeded under the assumption that $\lambda = 0$. The self-conscious agent now supposes that a population average $\overline{\lambda}_j$ exists and then substitutes this quantity for λ_{1j} in the Bayes equation (1), iterating expectations throughout. Then the self-conscious agent estimates the curing power of the therapy by solving for γ_{1j} in the following equation: $E_{t\overline{\lambda}}[\widehat{\alpha}_{1jt}] = \frac{(n-m)+\left[(1-\overline{\lambda})\gamma+\overline{\lambda}\widehat{\gamma}^{felt}_{1j,t-1}\right]t}{n+t}$. Then the optimal Bayes estimate of γ, given all available information, is $\widehat{\gamma}_{\overline{\lambda}_j,t} = \frac{t^{-1}\sum A_{1jt}-\overline{\lambda}\widehat{\gamma}^{felt}_{1j,t-1}}{1-\overline{\lambda}_{1j}}$

Expressed in terms of true parameters λ_{1j} and γ_{1j}, this is

$$E\left[\widehat{\gamma}_{\overline{\lambda}_{1j}}\right] = \frac{\gamma_{1j} - \lambda_{1j}\gamma + \left(\lambda_{1j} - \overline{\lambda}_{1j}\right)\tau^{-1}\sum_{t=0}^{\tau}\widehat{\gamma}^{felt}_{1jt}}{1-\overline{\lambda}_{1j}} \tag{4}$$

Equation (4) shows essentially that the self-conscious agent still retains an inconsistent estimate of the curing power of the consumed therapy, except under several knife-edge conditions (the simplest of these is perfect estimation of the placebo effect, or $\overline{\lambda}_j = \lambda_{1j}$). Let $\psi^{\overline{\lambda}_{1j}}_{it}$ be the augmented bias of the "self-conscious" felt estimate at time t. By a argument similar to that of Comment 1, it can be shown that $\psi^{\overline{\lambda}_{1j}}_{1jt}$ has asymptotic inefficiency for any $\varepsilon_\lambda \neq 0$.

2. Placebo Learning with Self-Limiting Diseases

2.1 The Complementarity of Placebo Learning and Self-Remission

Consider self-limitation as a binary variable $W_{jt} \in [0,1]$. Here $W_{jt} = 1$ corresponds to the event that the j^th disease self-remits in period t and 0 corresponds to the event that the disease does not remit (though the unremitted disease may be subject to curing by genuine treatment or placebo). We consider a self-limitation cycle by writing the first moment of W_{jt} as ω_{jt}. Then a self-limitation cycle is a convergent sequence $\{\omega_{jt}\}$ with length t^W. The remission probability ω_{jt} begins at zero (the disease cannot remit during the period of onset) and by definition reaches $\omega_j^{\max}$ at t^W. Once self-remission begins, convergence is monotonic such that per-period improvement is always non-negative ($\omega_{jt} - \omega_{j,t-1} \geq 0, \forall t$). We begin by considering the agent's inference problem with one cycle only, and consider multiple cycles as an extension in the following section. The agent now observes

$$Z^{felt,W}_{ijt} = X^{ILL}_{ij}\left(1 - A_{1jt}\right)\left(1 - W_{jt}\right)$$

If the naive agent substitutes the history of $Z^{felt,W}_{1jt}$ for the history of Y_{1jt} in the Bayes equation 1, then we can describe the limiting distribution of the estimator. Whereas Comment 1 establishes

the main problem with "pure" placebo learning as inefficiency of estimation, Comment 2 shows that adding a single cycle of self-limitation to placebo learning can lead to biased and inconsistent estimation of treatment efficacy in the long run.

Comment 2. Asymptotic Inconsistency of Bayes-Estimated Curing based on the Felt Curing History with Self-Limitation. Assume that any one of the following four conditions holds:(a) $\omega_{j1} \geq \gamma_{ij}$, (b) $\omega_{j1} \geq \frac{\widehat{\gamma}_{1j0}^{felt,W} - \widehat{\alpha}_{ijt}}{1-\widehat{\alpha}_{1j0}}$, (c) $\lambda_{ij}\left(\widehat{\gamma}_{1j1}^{felt,W} - \gamma_{ij}\right) \geq \frac{\widehat{\gamma}_{1j0}^{felt,W} - \omega_{j1}}{1-\omega_{j1}} - \gamma_{ij}$, (d) in period $\tau > t$, $\omega_{j(\tau)}^{\max} \geq \gamma_{ij}$, or $\omega_{j(\tau)}^{\max} \geq \frac{\widehat{\gamma}_{ijt}^{felt,W} - \widehat{\alpha}_{ijt}}{1-\widehat{\alpha}_{ijt}}$. Let the history $H_{1t}^{felt,W} = \left(Z_{1,j,1}^{felt,W}, Z_{1,j,2}^{felt,W}, ...Z_{1,j,t-1}^{felt,W}\right)$. If $\mu_{1j0} > \gamma_{1j}$, then $\forall H_{1t}^{felt,W}$,

$$\liminf E_t\left[\psi_{1jt}^{felt,W} \mid H_{1jt}^{felt,W}\right] \geq \left(1-\omega_j^{\max}\right)\widehat{\alpha}_0^{felt,W} + \omega_j^{\max} - \mu_{1j0} > 0$$

Inspection of the conditions shows that if $\omega_{j,t=1} > \gamma_{i=1,j}$, then the agent's upward bias is increasing in ω_j and λ_{1j}. A crucial property of placebo learning under self-limitation of disease is that the self-limiting force need not be great in order to induce false inference. Nothing in Comment 2 requires that the series ω_{jt} converge to one. In other words, an appreciable probability of recurrence can exist in any given period and placebo learning will still give rise to biased estimation of treatment efficacy.

2.2 "Endogenous" Placebo Learning with Self-Limiting Diseases: The Case of Multiple Cycles

Another point of complementarity between expectancy-based placebo mechanisms and self-limiting conditions arises when there are multiple cycles of disease, potentially accompanied by multiple cycles of utilization. I index "cycles" of diseases by c. When the human agent becomes ill for the first time ($c = 1$) and uses the incumbent treatment, then upon remission of the disease ($\omega_{ijt}^{c} = \overline{\omega}_j^{c}$ she re-enters a state of "health" ($X_{ijt} = 0$). We assume with certainty that the agent will experience illness again, such that for some $t > t_{\overline{\omega}_j^{c}}$, $X_{ijt} = X^{ILL}$. We then say that a second cycle of illness ($c = 2$) has begun, and we can write the variables observed by the human agent as $X_{ijt}^{c=2}$ and $Z_{ijt}^{c=2}$.

In the present model, expectancy-based placebo effects influence human learning if and only if prior expectations depart from the true pharmacological value of the treatment ($\widehat{\gamma}_{1j0}^{felt,c} \neq \gamma_{1j}$). Yet no such condition is required in order for self-remission to influence human estimation of efficacy. We can then imagine a situation in which the human agent "starts" with a correct estimate of the incumbent treatment's efficacy ($\widehat{\gamma}_{1j0}^{felt,c=1} = \gamma_{1j}$) and then ask what occurs to the "starting" values $\widehat{\gamma}_{1j0}^{felt,c}$ as the cycles of disease progress and c increases. Comment 3 demonstrates that under a rather flexible set of conditions, the human agent can start with a "true" prior and end up with a false one, thereby setting the inefficiency-based dynamics of Comment 1 into motion.

Comment 3. Asymptotic Inconsistency of Curing Priors Across Cycles. Assume that any one of the four conditions (a) - (d) in Comment 2 holds. Assume "accuracy" of the prior such that $\widehat{\gamma}_{1j,t=0}^{felt,c=1} = \gamma_{1j}$. Then $\forall c, H_{1t}^{felt,W}$,

$$E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij0}^{felt,W,c+1}\right] \geq E_{c-1,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij0}^{felt,W,c}\right]$$

and

$$\lim_{c\to\infty} \inf E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij0}^{felt,W,c+1}\right] > \gamma_{1j}$$

Comment 3 shows that when there are multiple cycles of illness and multiple cycles of utilization of the incumbent treatment, then an agent can start with the truth (or something less optimistic, or more) and still arrive at an optimistic estimate of the incumbent treatment's curing efficacy.

3. Utilization under Placebo Learning

3.1 Drug Utilization as a Dynamic Learning Problem. The results established in the previous section can be employed to examine the case of a rational agent who decides at each time whether or not to use a given medical treatment. The utilization problem may be approached as a special case of the multi-armed bandit problem (Gittins 1979; Banks and Sundaram 1992) The agent has per-period health valued at $I \in \Re^{+}$, and discounts future periods by a factor δ $(0 < \delta < 1)$. Let utilization $\sigma_{i,t} \in (0,1)$ indicate whether the agent uses treatment i in period t. The ith treatment has fixed cost $k_i > 0$. Let F_i (with primitive f_i) be the value function for the ith drug, conditional on the agent choosing that drug. The agent's problem is to maximize the difference between known per-period health and a stochastic loss function L, where the agent is presumed "risk-neutral."

$$\sup E_{t=0} \sum \delta^t F\left[I_t - X_{ij}^{ILL}\left(1 - A_{ijt}\left[\gamma_{ij}, \lambda_{ij}, \widehat{\gamma}_{ijt}^{felt}\right]\right) - k_i\sigma_{it}\right] \tag{5}$$

The agent starts with the incumbent treatment, and after t periods of utilization, given m successes from $n \geq m$ trials, then $\widehat{\gamma}_{ijt} = \frac{m}{n}$. The problem is then a multi-armed bandit case, with all but one $(Q+1)$ arms having unknown payoffs, and the default $(\sigma_t = 0)$ treatment having known curing probability β, which functions as a terminal payoff for the agent's problem. As long as the default is not too costly, the agent knows that she can do at least as well as utilizing the default treatment when sick. Drug-specific, per-period value functions are then as follows. If $\sigma_t = 0$, $E[F_t] = \beta X^{ILL} - k_0$. If $\sigma_t = 1$, $E[F_t] = \widehat{\gamma}_{1jt} X^{ILL} - k_1$. If $\sigma_t = q$, $E[F_t] = \widehat{\gamma}_{qjt} X^{ILL} - k_q$

Under standard optimization, $\widehat{\gamma}$ satisfies Bayes' rule, or equation (1). From dynamic programming, $F_t^i(\sigma_{t-1}, \widehat{\gamma}_t) = \sup\left\{F_t^0(\sigma_{t-1}), F_t^1(\sigma_{t-1}, \widehat{\gamma}_t), F_t^q(\sigma_{t-1}, \gamma_t)\right\} = \sup\left\{\frac{\beta X - k_0}{1-\delta}, F_t^1(\sigma_{t-1}, \widehat{\gamma}_t)\right\}$, and

$F^*(\sigma_{t-1}, \widehat{\gamma}_t) = \lim_{t\to\infty} F_t(\sigma_{t-1}, \widehat{\gamma}_t)$. For the default option, health valuation satisfies the following functional (Bellman) equation: $F_t^0(\sigma_{t-1}) = (\beta X - k_0) + \delta F_{t+1}^0(\sigma_t)$, and for the uncertain option,

$$F_t^1(\widehat{\gamma}_t, \beta) = f_t^1 + \delta \int_{\Re} f_{t+1}^1(\gamma)\,\mu(d\gamma)$$
$$= (\widehat{\gamma}_t X - k_1) + \delta\left[\widehat{\gamma}_t F_{t+1}^1\left(\sigma_t = 1, \widehat{\gamma}_{t+1}^+\right) + (1-\widehat{\gamma}_t) F_{t+1}^1\left(\sigma_t = 1, \widehat{\gamma}_{t+1}^-\right)\right]$$

where $\widehat{\gamma}_{t+1}^+ = \frac{m_t+1}{n_t+1}$ and $\widehat{\gamma}_{t+1}^- = \frac{m_t}{n_t+1}$. Then $F^*(\sigma_{t-1}, \widehat{\gamma}_t) = \sup\left\{\frac{\beta X - k_0}{1-\delta}, F_t^1(\sigma_{t-1}, \widehat{\gamma}_t)\right\}$

The solution to a problem of this nature is well known and entails the human agent's calculation of a dynamic allocation index (DAI) or "Gittins index." Intuitively, the Gittins index of an uncertain treatment is the minimum certain reward that an agent would choose over that treatment, given everything that the agent knows about the uncertain treatment. For any given β, let $\eta = \frac{\beta X^{ILL} - k_0}{1-\delta}$ stand for the expected terminal reward associated with choosing the default option forevermore. For the ith treatment (arm), and a (possibly optimal, possibly biased) estimate of curing γ_{ijt}, the agent would calculate the DAI as $R^i(\widehat{\gamma}_{ijt}) = \inf\{\eta \in \Re^+ | \sup[\eta, f_t^i + \delta \int f_{t+1}^i(\widehat{\gamma}_{ijt})\,\mu(d\gamma)] = \eta\}$. Given a calculable Gittins index for any treatment, then the agent's optimal strategy is to choose the treatment with the maximum Gittins index. This, then, is how a rational agent would dynamically utilize a battery of $Q+2$ medications, where one has known curing probability. Again, this technology is well-known and is a special case of bandit learning developed in statistical decision theory (Gittins 1979; Banks and Sundaram 1992). What has not been analyzed is the behavior of such a rational, learning agent when bias from placebo-learning unknowingly affects optimization. This is the aim of the following subsection.

3.2 Errors in Dynamic Utilization with Placebo Learning: Type I Error as Overpayment for Disguises and Sugar Pills. We may approach the possibility of suboptimal choice by human agents as a form of Type I or Type II decision error. A Type I error occurs when the agent uses the incumbent treatment but should have used another instead. Such overutilization can happen in one of two ways: the agent either overpays for a medication whose "work" is being done by beliefs, or she uses the wrong medication because her felt history conceals the true state of physiological damage. First, if placebo curing is separable from the treatment used, such that the agent could have taken a disguised pill identical to the incumbent treatment at lower cost and still have experienced the same curing history, then the agent is overpaying for the incumbent treatment.[6] As a benchmark, suppose that the default treatment could be so disguised. If the default treatment could be disguised in this way, then its efficacy could be raised under placebo effects. Yet it remains possible that the agent, by paying for the more expensive treatment, could be worse off than if placebo effects were not a possibility at all. The period-t value of the disguised default treatment under placebo

[6]This separability is intuitive but requires a particular topological concept for development in the model. See the Appendix and Summers (1972, Theorem 4.2) for the notion of topological separability used.

learning is $V^{\chi}_{t,k=k_0} = \widehat{\gamma}^{felt,W}_{\chi,t} X^{ILL}_{\chi j} - k_0 + \delta \int_{\Re} f^{\chi}_{t+1}(\gamma)\,\mu(d\gamma)$. The period-$t$ value of the incumbent treatment under placebo learning is $V^{i}_{t,k=k_i} = \widehat{\gamma}^{felt,W}_{i,t} X^{ILL}_{ij} - k_i + \delta \int_{\Re} f^{i}_{t+1}(\gamma)\,\mu(d\gamma)$. Then we can define by $\Phi^{I,B}_t = 1$ the "benchmarked" Type I error that occurs when $\sigma_t = i$ and $V^{i}_{t,k=k_i} - V^{\chi}_{t,k=k_0} > V^{\chi}_{t,k=k_0} - \frac{\beta X - k_0}{1-\delta}$. Fixing k_0 and k_i, this overpayment error is increasing in λ and in the value of the default treatment β.

Another possibility is that the disguised pill is a "sugar pill" with no pharmacological value. For any incumbent medication q, let treatment $\sigma_t = \chi(q, \widehat{\gamma}^{felt}_t)$ represent the cheapest possible sugar pill that could yield the same curing history, but with cost $k_\chi < k_q$. This scenario has been suggested by clinicians and scientists (Kirsch and Sapirstein 2002; Kirsch et al, 2002) who suggest, on the basis of a meta-analysis clinical trials, that the aggregate efficacy of SSRIs is not superior to placebo treatments. The ethical implications of intentional "placebo prescriptions" are subject to considerable debate (Hjobartsson and Gotzsche 2004). I sidestep the normative usse here. Still, the possibility that rather expensive SSRIs are not demonstrably superior to inert treatments suggests that alternative treatments may be a less costly means of ameliorating depressive and anxiety disorders.

A placebo-constrained human agent applies Bayes' rule to the observed series $\left\{ A^{felt,W}_{1jt} | \lambda_{1j} > 0 \right\}$ as if it were $\{G_{1jt}\} = \{A_{1jt} | \lambda_{1j} = 0, \omega_{jt} = 0\}$. Define by $\Phi^{I,\chi}_t = 1$ the Type I error that occurs when the agent utilizes the pharmacological treatment when she could pay less for a sugar pill and do at least as well in health. Formally, this occurs when $\sigma_t = 1$ and $\widehat{\gamma}^{felt,W}_{1,t} X^{ILL}_{j} - k_1 + \int f^{1}_{t+1}(\gamma)\,\mu(d\gamma) \leq \widehat{\gamma}^{felt,W}_{\chi,t} X^{ILL}_{\chi j} - k_\chi + \int_{\Re} f^{\chi}_{t+1}(\gamma)\,\mu(d\gamma)$. Similarly define by $\phi^{I,\chi}_{1jt}$ the probability of a Type I error at time t.

Comment 4: *Asymptotic Type I Error (Overpayment) under Placebo Learning, relative to Disguised Default.* For any chosen period t, let $\sigma_t = 1$ and $\widehat{\gamma}^{felt,W}_{1,jt} X^{ILL}_{j} - k_1 + \delta \int_{\Re} f^{1}_{t+1}(\gamma)\,\mu(d\gamma) \leq \widehat{\gamma}^{felt,W}_{\chi,t} X^{ILL}_{\chi j} - k_\chi + \delta \int_{\Re} f^{\chi}_{t+1}(\gamma)\,\mu(d\gamma)$.

Then

$$\Pr\left[\sup \Phi^{I,B}_t \left(H^{felt,W}_t \right) = 1 \right] > \Pr\left[\sup \Phi^{I,B}_t \left(H_t | \lambda_{1j} = 0 \right) = 1 \right],$$

and

$$\limsup \phi^{I,B}_{1jt} \left(H^{felt,W}_t \right) \geq \limsup \phi^{I,B}_{1jt} \left(H_t | \lambda_{1j} = 0 \right).$$

Comment 5: *Asymptotic Type I Error (Overpayment) under Placebo Learning, relative to Sugar Pill.* For any chosen period t, let $\sigma_t = 1$ and $\widehat{\gamma}^{felt,W}_{1,jt} X^{ILL}_{j} - k_1 + \int_{\Re} f^{1}_{t+1}(\gamma)\,\mu(d\gamma) \leq \widehat{\gamma}^{felt,W}_{\chi,t} X^{ILL}_{\chi j} - k_\chi + \int_{\Re} f^{\chi}_{t+1}(\gamma)\,\mu(d\gamma)$. Then $\Pr\left[\sup \Phi^{I,\chi}_{it} \left(H_t | \widehat{\gamma}_{1j0} = 0 \right) = 1 \right]$ is weakly increasing in $\widehat{\gamma}_{1j0}$, and for any $\widehat{\gamma}_{1j0} > \widehat{\gamma}_{2j0}$, $\limsup \phi^{I,\chi}_{ijt} \left(H^{felt}_t | \widehat{\gamma}_{1j0} \right) \geq \limsup \phi^{I,\chi}_{2jt} \left(H_t | \widehat{\gamma}_{2j0} \right)$.

The intuition here is that if placebo learning were not operative, the agent would eventually abandon the incumbent treatment in favor of the default, and could do no worse than this. With placebo effects, the value of the default treatment is raised, but the agent instead fails to abandon an incumbent treatment whose cost is so high that the agent would have been better off in a world without placebo effects (or where she knew of placebo effects and the separability of beliefs from the incumbent treatment). The "amount" of overutilization – here conceived as the overpayment induced by the (integrated or average) deviation of the series $\left\{\sigma_t^*(\widehat{\gamma}_{1jt}^*)\right\}$ from the series $\left\{\sigma_t^*(\widehat{\gamma}_{1jt}^{felt})\right\}$– is increasing in λ_{1j} and in the quantity $(\mu_{1j0} - \gamma_{1j})$. Under self-limitation, the series $\left\{\sigma_t^*(\widehat{\gamma}_{1jt}^{felt,W})\right\}$ diverges infinitely from the series $\left\{\sigma_t^*(\widehat{\gamma}_{1jt}^*)\right\}$.

As a plausible application of these results, consider again Branthwaite and Cooper's landmark (1981) study of branding and response to aspirin versus placebo. If branded aspirin were considerably more expensive than branded placebo, and the analgesic effect of expectations were separable from the aspirin, then a human agent could do nearly as well at much less cost if her meta-rational doctor prescribed her branded placebo instead of branded aspirin. Alternatively, if the incumbent treatment was a highly expensive SSRI (expensive not merely in terms of price, but also in terms of side-effects and risks), then if there were some less costly means of manipulating expectations (meditation, spiritual regimens, exercise, positive thinking, affirmations, therapy), the depressive agent could use these methods and possibly be better off. The placebo-learning model suggests that she will *not* abandon the incumbent, however, because expectations are inflated and utilization under placebo learning cannot falsify those expectations. The degree to which the human agent could do better would be an *increasing* function of the self-remission probability of the disease. Because pain, depression, and other conditions for which placebo responses have been studied are highly cyclic conditions with respect to pain/misery/anxiety, human agents with these maladies are *more* likely to find themselves in suboptimal utilization situations the more more cyclic are their underlying conditions.

3.3 Errors in Dynamic Utilization with Placebo Learning: Undertreatment. A second form of Type I error occurs when the agent's "felt" history wrongly (but unknowingly) understates the true physiological damage being done by a disease. One salient historical example of this sort of overutilization may lie in the vast market for "patent medicines" in the nineteenth- and twentieth-century United States (Young 1967). Doctors prescribed (and consumers eagerly took and bought) sham cancer "cures" which often made them feel better but did not address the underlying malignancy. Or they gave their children opium-laced tonics for influenza. All the while, they avoided treatments (imagine these as the "default" option of the present model) that would have presented a higher actual curing probability than the incumbent.

Undertreatment can occur if, unknown to the agent, the "felt" history conceals a true pathology.

Suppose then that the the agent's self-observable health follows $Z_{1jt}^{felt,W}$ but that an undetected disease process (diffusion of a malignancy, hypertension, arterial occlusion) continues according to Y_{1jt}. If the sum of Y_{1jt} is greater than an exogenous cutoff point ξ_j, which the agent knows, then the agent "dies," such that behavior is stopped and a known terminal penalty $D_j > 0$ is incurred.[7] Again, the placebo-constrained agent observes only $Z_{1jt}^{felt,W}$, not Y_{1jt}. Let $\sigma_t = q^*$ represent the strategy whose pharmacologically determined Gittins index is highest. Given sickness X^{ILL}, the agent using the q^*th treatment would have died at τ^{real}. But the placebo-learning agent using the incumbent treatment expects, at the beginning of the problem, to die at $\tau^{felt} = E_0^{felt,W}[\tau|\gamma_{1j,t=0}]$. Then define by $\Phi_t^{I,D} = 1$ the Type I error that occurs when the agent utilizes the incumbent treatment when she would have been better off using the q^*th alternative (at a minimum, the default). Formally, this occurs when

$$\Pr\left[\sup\sum_t Y_{1jt}(\sigma_t = q^*) > \xi_j\right]\delta^{\tau^{real}}D_j + \sum_{t=0}^{\tau^{real}}\delta^t(Y_{1jt}(\sigma_t = q^*) - k_{q^*})$$

$$> \Pr\left[\sup\sum_t Z_{1jt}^{felt,W} > \xi_j\right]\delta^{\tau^{felt}}D_j + \sum_{t=0}^{\tau^{felt}}\delta^t(Z_{1jt} - k_1)$$

where $E[\tau] = E\left[\tau|\inf\tau|\sum_t Y_{1jt} > \xi_j\right]$. Similarly define by $\phi_{1jt}^{I,D}$ the probability of a "deadly" Type I undertreatment error at time t.

Comment 6: *Asymptotic Type I Error (Undertreatment) under Placebo Learning.* For any stopping time τ, with $\sigma_t = i$, *let* $\Pr\left[\sup_{t\to\tau}\sum_t Y_{1jt} > \xi_j\right]\delta^{E[\tau]}D_j >$ and $\Pr\left[\sup_{t\to\tau}\sum_t Z_{1jt}^{felt,W} > \xi_j\right]\delta^{E[\tau]}D_j$.

Then

$$\Pr\left[\sup_{t\to\tau}\Phi_\tau^{I,D}\left(H_\tau^{felt,W}\right) = 1\right] > \Pr\left[\sup\Phi_\tau^{I,D}\left(H_\tau|\lambda_{1j} = 0\right) = 1\right],$$

and

$$\lim_{\tau\to\Theta}\sup\phi_{1j\tau}^{I,D}\left(H_\tau^{felt,W}\right) \geq \lim_{\tau\to\Theta}\sup\phi_{1j\tau}^{I,D}\left(H_\tau|\lambda_{1j} = 0\right).$$

At its core, then, placebo learning can lead to failure to abandon incumbent treatments when doing so would enhance the agent's health or utility. Abandonment might be optimal because the human agent could achieve an equal health state at much less cost with another treatment (where "cost" can be interpreted as expense, side-effects or risk). Or abandoning treatment might be optimal because the incumbent treatment is inferior at curing a deadly illness. This second sort of failure also intimates how different sorts of decision errors are related to one another. An agent's

[7] This requires relaxation of the one of the earlier assumptions of the model, namely that agents are infinitely lived. This assumption is necessary to characterize the $t = \infty$ case for a single cycle, but it can be relaxed for the present sub-analysis without loss of generality, if we assume that even a finitely-lived agent can feasibly compute expectations over an infinite time-horizon.

long-run failure to abandon the incumbent treatment can be correlated with long-run failure to experiment with other, possibly superior therapies. That is, a Type I error can generate Type II errors with respect to other treatments, with appreciable health consequences.

Of course if suggestibility is absent ($\lambda_{1j} = 0$) and self-remission is not a possibility or is unknown, estimation of γ_{1j} is unbiased and efficient. Then the human agent eventually abandons the incumbent therapy according to an optimal stopping program.

3.4. The Priority of Initial Beliefs Under Placebo Learning with Self-Limiting Diseases. For diseases with self-limitation, human error does not necessarily occur if the human agent places low faith in uncertain treatments. Intuitively, if the human agent becomes ill, but possesses such low confidence in available treatments that she leaves her condition untreated entirely, then then the proper inference from "natural" healing of the disease is that no treatment was necessary. In other words, false inferences about the efficacy of uncertain treatments are not possible if those treatments are not utilized to begin with.

Although it is a rather straighforward result from the model, it is worth pointing to the existence of some histories under which the agent pessimistically abandons the incumbent treatment before self-limitation probabilities rise sufficiently, leading to optimal utilization. One interpretation of the model is then that human learning and decision making is particularly affected by economic and social mechanisms that lead initial beliefs to be higher so that agents utilize uncertain and untried therapies when they first get ill.

Comment 7: *Utilization and Initial Beliefs.* Both mean utilization $E_t[\sigma_t]$ and mean overutilization $E_t\left[\Phi_t^I\right]$ are non-decreasing in γ_{1j0}. Both relationships are strictly conditional upon (increasing in) suggestibility.

4. Discussion and Extensions

I have presented a dynamic stochastic model of treatment utilization by a human agent whose rationality is constrained by placebo effects. Analysis of this model suggests that, far from being competing mechanisms, expectancy-based placebo effects and self-limitation of disease may reinforce one another in their frustration of valid human inference about the efficacy of medical or pharmacological treatments. The model elaborated here is of course limited by its simplicity, which comes mainly in the functional (probabilistic) forms adopted for representation of human inference. Still, the functional forms here are reasonably flexible, and the structure elaborated in the model offers a tractable foundation for building more complex models where some of the following properties might be explored.

Non-Convergent Cycles. Some diseases do not remit entirely but wax and wane in their felt

intensity or pain. The model here is assisted greatly by the convergence requirement, and for this reason a useful extension would consider non-convergent illness cycles.

Strategic Firms, Advertising and Pricing. In the presence of placebo learning, consumers and their doctors may be less willing to switch medications in the presence of experienced response to a particular drug. To the extent that placebo effects are operating, the firm then may have fewer incentives to price competitively, and it may have greater incentives to advertise drugs to manipulate beliefs. As a conjecture, this problem may be more appreciable when human learning is constrained by memory limitations.

Dosage-Response Curves. The model here considers utilization as binary – the treatment is used or is not – but more complex relationships between "amount" of utilization per period and human learning should be explored theoretically, as well as empirically. However, this point may also be rendered as a critique of the empirical literature. Very few studies in the placebo literature examine dose-response relationships in the placebo effect, and little to nothing is known about precise functional relationships between placebo doses and placebo responses, particularly when the same human agent is exposed to different doses of the same treatment over time.

Computional Models with Neuro-Dynamic Programming. Further analysis of models along the lines of Redish (2004) can only help to illustrate these dynamics. One direction in which the current analytic model points us is towards the possibilty of diffentiating between expectancy-based placebo healing and reinforcement-learning mechanisms, and to view these mechanisms, too, as potential complements (Amanzio and Bendetti 1999). In addition, the analytic model here suggests that incorporation of self-remitting conditions into computational models would yield fruitful inquiry.

REFERENCES

Amanzio, M., and F. Benedetti. 1999. "Neuropharmacological Dissection of Placebo Analgesia: Expectation-Activated Opioid Systems versus Conditioning Activated Specific Subsystems," *The Journal of Neuroscience* 19(1):484-494.

Benedetti, F., and M. Amanzio. 1997. "The neurobiology of placebo analgesia: from endogenous opioids to cholecystokinin," *Prog Neurobiol* 52: 109-25.

BS Bertsekas, D. P., and Shreve, S. E. 1996. *Stochastic Optimal Control: The Discrete-Time Case* Belmont, MA: Athena Scientific.

Branthwaite, A., and P. Cooper. 1981. "Analgesic effects of branding in the treatment of headaches," *British Journal of Medicine* 282: 1576-1578.

Christakis, N. 1999. *Death Foretold: Prophecy and Prognosis in Medical Care* Chicago: University of Chicago Press.

Cobb, L., Thomas, G.I., Dillard, D.H. Merendino, K.A. and R. A. Bruce. 1959. "An evaluation of internal mammary artery ligation by a double-blind technique," *New England Journal of Medicine* 260: 1115-1118.

Colloca, L., L. Lopiano, M. Lanotte, F. Benedetti. 2004. "Overt versus covert treatment for pain, anxiety, and Parkinson's disease," *The Lancet Neurobiology* 3:679-684.

Das, J. 2001. "Do Patients Learn About Doctor Quality? Theory and Evidence from India." Chapter of Ph.D. Dissertation, Department of Economics, Harvard University.

Desharnais, R., Jobin, J., Cote, C., Levesque, L., and G. Godin. 1993. "Aerobic exercise and the placebo effect: a controlled study," *Psychosomatic medicine* 55: 149-154.

deGroot, M. H. 1970. *Optimal Statistical Decisions.* New York: McGraw-Hill.

Dimond, E.G., Kittle, C. F., and J. E. Crockett. 1960. "Comparison of internal mammary ligation and sham operation for angina pectoris," *American Journal of Cardiology* 5: 483-486.

Gibbons, F., and S. E. Hormuth. 1981. "Motivational Factors in Placebo Responsivity," *Psychopharmacology Bulletin* 17:77-79.

Hahn, R. 1997. "The Nocebo Phenomenon: Scope and Foundations," chapter 3 in Harrington, ed *The Placebo Effect: An Interdisciplinary Exploration* Cambridge: Harvard University Press.

Harrington, A. 1997. *The Placebo Effect: An Interdisciplinary Exploration*, ed. Anne Harrington. Cambridge: Harvard University Press.

Kaptchuk, T. J., Goldman, P., Stone, D.C. and W. B. Stason. 2000. "Do medical devices have enhanced placebo effects?" *Journal of Clinical Epidemiology* 53: 786-792.

Khan, A., Warner, H., and W. A. Brown. 2000. "Symptom Reduction and suicide risk in patients treated with placebo in antidepressant clinical trials: An Analysis of the Food and Drug Administration database," *Archives of General Psychiatry* 57: 311-17.

Kirsch, I., T, J, Moore, A. Scoboria and S. S. Nicholls. 2002. "The Emperor's New Drugs: An Analysis of Antidepressant Medication Data Submitted to the U.S. Food and Drug Administration," *Prevention and Treatment* 5 (23).

Kirsch, I., and A. Sapirstein. 2002. "Listening to Prozac but hearing placebo: An meta-analysis of antidepressant medication," *Prevention and Treatment* 1 Article 2a.

Kleinhenz J, Streitberger K, Windeler J, Gussbacher A, Mavridis G, Martin E. 1999. "Randomised clinical trial comparing the effects of acupuncture and a newly designed placebo needle in rotator cuff tendinitis," *Pain* 1999 November; 83(2):235-41.

Koszegi, B. 2002. "Anticipation in Observable Behavior." manuscript, University of California at Berkeley.

Levine, J.D., N. C. Gordon, and H. L. Fields. 1978. "The Mechanism of Placebo Analgesia" *Lancet* 2:654-57.

Mullainathan, S. 2002. "A Memory-Based Model of Bounded Rationality." *Quarterly Journal of Economics* 117 (August 2002): ZZZ.

Petrovic, P., E. Kalso, K. M. Petersson, M. Ingvar. 2002. "Placebo and Opioid Analgesia – Imaging a Shared Neuronal Network," *Science* 295:1737-1740.

Redish, A. D. 2004. "Addiction as a Computational Process Gone Awry," *Science* 306: 1944-47.

Reich, J. 1990. "The Effect of Personality on Placebo Response in Panic Patients." *Journal of Nervous and Mental Diseases* 178: 699-702.

Shapiro, A. K. and E. Shapiro. 1997. *The Powerful Placebo: From Ancient Priest to Modern Physician* Baltimore: Johns Hopkins University Press.

Stewart-Williams, S., and J.Podd. 2004. "The placebo effect: Dissolving the expectancy versus conditioning debate," *Psychological Bulletin* 130:324-340.

Summers, W. H. 1972. " Separability in the Strict and Substrict Topologies," *Proceedings of the American Mathematical Society* 35 (2): 507-514

Swartzman, L.C. and J. Burkell. 1998. "Expectations and the placebo effect in clinical trials," *Clinical Pharmacology and Therapeutics* 64: 1-7.

Tobler, P. N., C. D. Fiorillo, W. Schultz. 2005. "Adaptive Coding of Reward Value by Dopamine Neurons," *Science* 307: 1642-45.

Wager, T. D., et al. 2004. "Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain," *Science* 303: 1162-67.

APPENDIX

We being with a characterization of the problem. Let X be a topologically complete and separable Borel space with probability measure $P(X)$. Let H^{∞} represent the Hilbert cube, or countably many copies of the unit interval.

Lemma 1: Hilbert Cube Representation. The variables G, Y, Z, A and W, all stochastic integrals of these variables, and all action mappings σ are representable on the Hilbert cube H^{∞}, and all variables and probability functions of these variables are Borel-measurable.

Proof: There are two ways to demonstrate this result. The first and simplest is to constrain X^{ILL} itself to the unit interval, as in $X^{ILL} = 1$. Then the entire outcome space is $(0, 1)$. The second and more general is to assume X to be a topologically complete, separable space with probability measure $P(X)$. Then by Urysohn's theorem (BS, Proposition 7.2) there exists a homeomorphism $\rho : X \to H$, such that $P(X)$ is topologically separable and complete (**BS**, Proposition 7.23), and $P(X)$ is itself a Borel space. We can then characterize the σ-algebra $B_{P(X)}$. Then the binary variables G, Y, Z, A and W are all Borel measurable, as are any convex combinations of those variables. Let all probability functions (cdfs) and probability density functions (pdfs) for G, Y, Z, A and W be Borel-measureable functions on X – this is astraightforward for Beta and Bernoulli variates – then (**BS**, Proposition 7.29) all integrals and stochastic integrals are Borel-measurable.

Proof of Comment 1

The agent observes Z_t, not Y_t. The naive agent confuses the two entirely, while the more sophisticated agent knows that $\lambda > 0$ when this is the case but does not know the exact value of λ. Then

$$\begin{aligned} E_t\left[Z_{ijt}\right] &= E_t\left[X_{ij}^{ILL}\left[1 - A_{ijt}\right]\right] \\ &= X_{ij}^{ILL}\left[1 - (1 - \lambda_{ij})\,\gamma_{ij} - \lambda_{ij}\widehat{\gamma}_{ijt}^{felt}\right] \end{aligned}$$

The behavior of the "felt" estimate in the τth period can be described as

$$
\begin{aligned}
E_0\left[\widehat{\gamma}_{ij,t=\tau}^{felt}\right] &= \frac{m + X_{ij}^{ILL}\left[\tau\left(1-\lambda_{ij}\right)\gamma_{ij} + \lambda\left(\sum_{t=0}^{\tau-1}\widehat{\gamma}_{ijt}^{felt}\right)\right]}{n + \tau X_{ij}^{ILL}} \\
&= \frac{m + X_{ij}^{ILL}\left[\tau\left(1-\lambda_{ij}\right)\gamma_{ij} + \lambda\left(\sum_{t=0}^{\tau-1}\left(\widehat{\gamma}_{ijt}^{felt} - \gamma_{ij}\right) + \gamma_{ij}\right)\right]}{n + \tau X_{ij}^{ILL}} \\
&= \frac{m + X_{ij}^{ILL}\left[\tau\gamma_{ij} + \lambda\left(\sum_{t=0}^{\tau-1}\psi_{ijt}^{felt}\right)\right]}{n + \tau X_{ij}^{ILL}} \qquad (6a)
\end{aligned}
$$

where ψ_{ijt}^{felt} is the error of the felt estimate at time t. The long-run behavior of the felt estimate will then depend upon the convergence (if any) of the series $\left\{\psi_{ijt}^{felt}\right\}$. The difference between the optimal Bayes estimate of $\widehat{\gamma}_{ijt}^{*}$ and the felt Bayes estimate $\widehat{\gamma}_{ijt}^{felt}$can be expressed in terms of the "observables" Y_t and Z_t.

$$
\widehat{\gamma}_{ij,t+1}^{*} = 1 - \frac{(n-m) + \sum Y_{ijt}}{n + tX_{ij}^{ILL}} \qquad (7)
$$

$$
\widehat{\gamma}_{ij,t+1}^{felt} = 1 - \frac{(n-m) + \sum Z_{ijt}}{n + tX_{ij}^{ILL}}
$$

First, there is expected bias for any finite time t.

$$
\begin{aligned}
\widehat{\gamma}_{ij,t+1}^{felt} &= \widehat{\gamma}_{ij,t+1}^{*} + \psi_t^{felt} \\
1 - \frac{(n-m) + \sum Z_{ijt}}{n + tX_{ij}^{ILL}} &= 1 - \frac{(n-m) + \sum Y_{ijt}}{n + tX_{ij}^{ILL}} + \psi_t^{felt} \\
1 - \frac{(n-m) + \sum X_{ij}^{ILL}\left(1 - A_{ijt}\right)}{n + tX_{ij}^{ILL}} &= 1 - \frac{(n-m) + \sum_{ijt} X_{ij}^{ILL}\left(1 - G_{ij}\right)}{n + tX_{ij}^{ILL}} + \psi_t^{felt}
\end{aligned}
$$

From this we can conclude that

$$
\psi_t^{felt} = \frac{\sum X_{ij}^{ILL}\left(\lambda_{ij}\sum A_t\left[\gamma_{ij}, \lambda_{ij}, \widehat{\gamma}_{ijt}^{felt}\right] - \lambda_{ij}\sum G_{ij}\left[\gamma_{ij}\right]\right)}{n + tX} \qquad (8)
$$

The expected bias for any time $t + \tau$ is

$$
E_t\left[\psi_\tau^{felt}\right] = \frac{\tau X_{ij}^{ILL}\left[\lambda_{ij}\left(\widehat{\gamma}_{ijt}^{felt} - \gamma_{ij}\right)\right]}{n + \tau X_{ij}^{ILL}} \qquad (9)
$$

The agent's estimate $\widehat{\gamma}_{ijt}^{felt}$ is unbiased if $\lambda_{ij} = 0$ or under the knife-edge condition that $\gamma_{ij} = \widehat{\gamma}_{ijt}^{felt}$, where in other words the agent just happens to believe that the curing power of the pill is exactly that of the real treatment. As μ_{ij0} is a Beta variate and $\widehat{\gamma}_{ijt}^{felt}$ a Beta variate, this event has Lebesgue measure zero for any t.

To consider the asymptotic distribution of bias, we conduct a first-step analysis. Consider the felt estimate for any two periods t and $t+1$.

$$\begin{aligned} E\left[\widehat{\gamma}_{ij,t+1}^{felt}\right] &= 1 - \frac{(n-m) + \sum Z_{ijt}}{n + tX_{ij}^{ILL}} \\ &= 1 - \frac{(n-m) + \sum X_{ij}^{ILL}\left\{1 - (1-\lambda_{ij})\, G_{ij}\left[\gamma_{ij}\right] - \lambda_{ij} W_t\left[E_t\left[\gamma_{ij}\right]\right]\right\}}{n + tX_{ij}^{ILL}} \end{aligned}$$

Let m_t and n_t denote the observed number of "successes" and "trials" at time t. Then consider the movement of the estimate during the next two period $t+1$ and $t+2$. Again by the law of iterated expectations,

$$\begin{aligned} E\left[\widehat{\gamma}_{ij,t+1}^{felt}\right] &= E\left[1 - \frac{(n_t - m_t) + Z_{ijt}}{n_t + X_{ij}^{ILL}}\right] \\ &= 1 - E\left[\frac{(n_t - m_t) + X_{ij}^{ILL}\left(1 - (1-\lambda_{ij})\,\gamma_{ij} - \lambda_{ij}\left[E_t^{felt}\left[\gamma_{ij}\right]\right]\right)}{n_t + X_{ij}^{ILL}}\right] \\ & \quad 1 - \frac{(n_t - m_t) + X_{ij}^{ILL}\left(1 - (1-\lambda_{ij})\,\gamma_{ij} - \lambda_{ij}\frac{m_t}{n_t}\right)}{n_t + X_{ij}^{ILL}} \end{aligned}$$

Then for period $t+2$,

$$\begin{aligned} E\left[\widehat{\gamma}_{ij,t+2}^{felt}\right] &= E\left[1 - \frac{(n_t - m_t) + Z_{ijt} + Z_{ij,t+1}}{n_t + 2X_{ij}^{ILL}}\right] \\ &= 1 - E\left[\frac{(n_t - m_t) + 2X_{ij}^{ILL}\left(1 - (1-\lambda_{ij})\, G_{ijt}\left[\gamma_{ij}\right] - \lambda_{ij} W_t\left[E_t\left[\gamma_{ij}\right]\right]\right)}{n_t + 2X_{ij}^{ILL}}\right] \\ &= 1 - \frac{(n_t - m_t) + 2X_{ij}^{ILL}\left(1 - (1-\lambda_{ij})\,\gamma_{ij}\lambda_{ij}\left[1 - \frac{(n_t - m_t) + X_{ij}^{ILL}\left(1 - (1-\lambda_{ij})\gamma_{ij}\lambda_{ij}\frac{m_t}{n_t}\right)}{n_t + X_{ij}^{ILL}}\right]\right)}{n_t + 2X_{ij}^{ILL}} \end{aligned}$$

Then when initial beliefs are "high," or $\gamma_{1j} < \frac{m_t}{n_t}$, a sufficient condition for convergence is that the difference $E\left[\widehat{\gamma}_{ij,t+2}^{felt}\right] - E\left[\widehat{\gamma}_{ij,t+1}^{felt}\right]$ tends asymptotically downward, or, $\forall t$

$$Q_{t,t-1} = \frac{m_t + X_{ij}^{ILL}\left[\gamma + \lambda\left(\frac{m_t}{n_t} - \gamma\right)\right]}{n_t + X_{ij}^{ILL}} - \frac{m_t}{n_t} < 0. \tag{10}$$

This occurs iff $\gamma < \frac{m_t}{n_t}$, which is true by assumption. The opposite case has a symmetric proof.

For efficiency, we examine the variance with which $\widehat{\gamma}_{ijt}^{felt}$ converges to γ_{ij}. For the optimal estimate $\widehat{\gamma}_{ijt}^{*}$, for $\frac{m}{n} > \gamma$, the τ-th-period error is

$$\psi_\tau^* = \frac{m + \sum_{t=0}^{\tau-1} G_{ijt}}{n+\tau} - \gamma_{ij}$$

From this we can write the time-τ posterior variance of the optimal estimator as $Var_\tau^{post}\left[\widehat{\gamma}_\tau^*\right]$. Because G_{ijt} is stationary, $E\left[\psi_\tau^*\right] = \frac{m+\tau\gamma}{n+\tau} - \gamma_{ij}$, and $\lim_{\tau\to\infty}\psi_\tau^* = 0$. The rate of change of the expected estimate is $\frac{n\gamma - m}{(n+\tau)^2}$.

From equation (3), we can write

$$Var_\tau^{post}\left[\widehat{\gamma}_\tau^{felt}\right] = Var_\tau^{post}\left[\widehat{\gamma}_\tau^*\right] + \lambda Var_\tau^{post}\left[\psi_\tau^{felt}\right]$$

Because the rightmost term is strictly positive for any $\lambda > 0$ and any $\tau > 0$, the posterior variance for the felt estimate exceeds that for the estimate where $\lambda = 0$, and the Comment is proved.

Comment 1: The Case of Sophisticated Placebo Learning.

We can discuss the the movement of the bias of the agent's estimate in a way analogous to that of the naive $\left(\overline{\lambda}_j = 0\right)$ case. For these purposes, it will be convenient to define the error of the self-conscious agent's estimate of λ as follows.

$$\varepsilon_\lambda \equiv \lambda_{ij} - \overline{\lambda}_j$$

Iterating expectations, the expected bias behaves as does

$$\begin{aligned}
E_t\left[\psi_\tau^{\overline{\lambda}_j}\right] &= E_{t\overline{\lambda}}[\widehat{\gamma}_{ij\tau}] - E_t^*[\widehat{\gamma}_{ij\tau}] \\
&= E_{t\overline{\lambda}}\left(1 - \frac{(n-m) + X_{ij}^{ILL}\sum Z_{ijt}^{\overline{\lambda}_j}}{n+\tau X_{ij}^{ILL}}\right) - E_t^*\left(1 - \frac{(n-m) + X_{ij}^{ILL}\sum Y_{ijt}}{n+\tau X_{ij}^{ILL}}\right) \\
&= E_{t\overline{\lambda}}\left(\frac{m + X_{ij}^{ILL}\sum A_{ijt}^{\overline{\lambda}_j}}{n+\tau X_{ij}^{ILL}}\right) - E_t^*\left(\frac{m + X_{ij}^{ILL}\sum G_{ijt}}{n+\tau X_{ij}^{ILL}}\right) \\
&= \frac{\varepsilon_\lambda\left(\tau^{-1}\sum_{t=0}^{\tau}\widehat{\gamma}_{ijt}^{felt} - \gamma\right)\tau}{n+\tau}
\end{aligned} \tag{11}$$

There are three cases. Once again perfect estimation of the placebo effect ($\varepsilon_\lambda = 0$) eliminates bias. Above-average placebo response ($\varepsilon_\lambda > 0$) induces positive bias whereas below-average placebo response may induce negative bias.

Proof of Comment 2.

$$\begin{aligned} \widehat{\gamma}_{ij,t+1}^{felt,W} &= \widehat{\gamma}_{ij,t+1}^{*} + \psi_t^{felt,W} \\ 1 - \frac{(n-m) + \sum Z_{ijt}^{W}}{n + tX_{ij}^{ILL}} &= 1 - \frac{(n-m) + \sum Y_{ijt}}{n + tX_{ij}^{ILL}} + \psi_t^{felt,W} \end{aligned}$$

Dropping sub- and superscripts for X_{ij}^{ILL}, this leads to

$$\psi_t^{felt,W} = \frac{\sum XA_{ijt} + \sum XW_{ijt} - \sum XA_{ijt}W_{ijt} - \sum XG_{ijt}\left[\gamma_{ij}\right]}{n + \tau X} \tag{12}$$

where all sums are over the sequence $t = 0, ...\tau$.

Of the variables in (12), only G_{ijt} is stationary, making computation of expectations difficult. We can perform a first-step analysis by describing the expected movement of the variable $\widehat{\gamma}_{ijt}^{felt,W}$ in the first few periods. Start in an arbitrary period t, where priors are given my m_t and n_t. By iterated expectations, the period 1 expectation of the period $t+2$ felt curing estimate with self-limitation is

$$E_1\left[\widehat{\gamma}_{ij,t+1}^{felt,W}\right] = \frac{m_t + X_{ij}^{ILL}\left\{(1-\omega_{ijt})\left[(1-\lambda_{ij})\,\gamma_{ij} + \lambda_{ij}\widehat{\gamma}_{ijt}^{felt,W}\right] + \omega_{ijt}\right\}}{n_t + X_{ij}^{ILL}}$$

Now define by $Q_{t,t+1}^{W}$ the differential movement of the felt curing estimate with self-limitation. This is

$$\begin{aligned} Q_{t,t+1}^{W} &= \frac{X_{ij}^{ILL}\left\{(1-\omega_{ijt})\left[(1-\lambda_{ij})\,\gamma_{ij} + \lambda_{ij}\widehat{\gamma}_{ijt}^{felt,W}\right]\right\} + X_{ij}^{ILL}\left(\omega_{ijt} - \widehat{\gamma}_{ijt}^{felt,W}\right)}{n_t + X_{ij}^{ILL}} \\ &= \frac{X_{ij}^{ILL}\left(1-\omega_{ijt}\right)\widehat{\alpha}_{ijt} + X_{ij}^{ILL}\left(\omega_{ijt} - \widehat{\gamma}_{ijt}^{felt,W}\right)}{n_t + X_{ij}^{ILL}} \end{aligned}$$

There are two cases. The source of the bias may be seen in the expression for ψ_t, namely that where $\widehat{\gamma}_t^{felt} > \gamma_{ij}$, ψ_t is a submartingale and a supermartingale in the opposite case. Consider again the quantity

$$Q_{t,t+1} = \frac{X_{ij}^{ILL}\left\{(1-\omega_{ijt})\left[(1-\lambda_{ij})\,\gamma_{ij} + \lambda_{ij}\widehat{\gamma}_{ijt}^{felt,W}\right]\right\} + X_{ij}^{ILL}\left(\omega_{ijt} - \widehat{\gamma}_{ijt}^{felt,W}\right)}{n_t + X_{ij}^{ILL}}$$

If any of the four conditions in Comment 2 holds, $Q_{t,t+1}$is monotone increasing. The fact that $\widehat{\gamma}_{ijt}^{felt,W}$ cannot surpass one establishes the existence of an upper bound upon $Q_{t,t+1}$, which is sufficient for weak concavity of the series.

To show asymptotic inconsistency of $\widehat{\gamma}_{ijt}^{felt,W}$ under the four conditions, we begin with the observation that, for all t, $\psi_t^{felt,W}$ is absolutely monotonic in $Q_{t,t-1}^W$. Since $Q_{t,t-1}^W$ is bounded and monotone, then $\forall H_t^{felt,W}$, there exists some $\liminf \psi_t^{felt,W}\left(H_t^{felt,W}\right)$. We can characterize the infimum by

$$\begin{aligned}\inf \psi_t^{felt,W} &= \inf \frac{\sum XA_{ijt} + \sum XW_{ijt} - \sum XA_{ijt}W_{ijt} - \sum XG_{ijt}\left[\gamma_{ij}\right]}{n + \tau X} \\ &\leq \left(1 - \omega_{ij}^{\max}\right)\widehat{\alpha}_0^{felt,W} + \left(\omega_{ij}^{\max} - \mu_{ij0}\right)\end{aligned}$$

The last expression is guaranteed positive by assuming any of the four conditions. Because lim inf is an increasing series in t, then for some infinitesimal ϵ, $\liminf \psi_t\left(H_t^{felt}\right) = \left(1 - \omega_{ij}^{\max}\right)\widehat{\alpha}_0^{felt,W} + \left(\omega_{ij}^{\max} - \mu_{ij0}\right) - \epsilon$. By Fatou's lemma,

$$\int \liminf \psi_t\left(H_t^{felt}\right) + \epsilon \leq \liminf \int \psi_t\left(H_t^{felt}\right) + \epsilon$$

and so

$$\left(1 - \omega_{ij}^{\max}\right)\widehat{\alpha}_0^{felt,W} + \left(\omega_{ij}^{\max} - \mu_{ij0}\right) \leq E_t\left[\liminf \psi_t\left(H_t^{felt}\right)\right] \leq \liminf E_t\left[\psi_t \mid H_t^{felt}\right]$$

Proof of Comment 3 We first characterize the movement of W_{ijt} within cycles. In cycle 1, ω_{jt}^1 moves from zero to $\overline{\omega}_j^1$. In cycle 2, ω_{jt}^2 moves from zero to $\overline{\omega}_j^2$. And for cycle c, ω_{jt}^c moves from zero to $\overline{\omega}_j^c$. We let learning stop, sickness remit completely ($X^{ILL} = 0$), and utilization stop, when ω_{jt}^c reaches $\overline{\omega}_j^c$. By Comment 2, with its conditions (a)-(d) assumed, $E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ijt}^{felt,W,c=1}\right]$ moves from γ_{1j} to a quantity $E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij,t=t_{\overline{\omega}_j^1}}^{felt,W,c=1}\right]$, such that

$$E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ijt}^{felt,W,c=1}\right] - E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij,t=t_{\overline{\omega}_j^1}}^{felt,W,c=1}\right] > 0 \tag{13}$$

But since $\widehat{\gamma}_{ij0}^{felt,W,c}$ must be altered across cycles in accordance with Bayes' rule, the Bayesian human agent takes

$$\widehat{\gamma}_{ij,t=0}^{felt,W,c=2} \equiv \widehat{\gamma}_{ij,t=t_{\overline{\omega}_j^1}}^{felt,W,c=1} > 0$$

Taking conditional expectations and using (13),

$$E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij,t=0}^{felt,W,c=2}\right] \geq E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{ij,t=0}^{felt,W,c=1}\right]$$

and the first statement in Comment 3 is proved, as $c = 1, 2$ are arbitrary. For the second statement, note that by repeated application of (13),

$$\lim_{c\to\infty} \inf E_{c,\bigcup H_{1t}^{felt,W}}\left[\widehat{\gamma}_{1j0}^{felt,W,c}\right] \geq \gamma_{1j}$$

By (13), the inequality is strong iff there exists one cycle for which the movement is strictly positive across cycles. As c gets large, the event that no such cycle has occurred has vanishing measure. The second statement is proved.

A stronger result is possible, namely that

$$\lim_{c\to\infty} E_{c,\bigcup H_{1t}^{felt,W}}\left(\lim_{\omega_{jt}\to\overline{\omega}_j^r} \inf \widehat{\gamma}_i^{felt,W,c} jt\right) = \overline{\omega}_j \tag{14}$$

where limits are computed right to left.

Lemma 2. Existence of an Optimal Policy.

Proof. SKETCH: The problem here fits into a general class of models that have been analyzed by Gittins, Gittins and Jones, Berry and Fristedt, and Banks and Sundaram (see references). To show existence of an optimal utilization path, notice that for any t, the cost-continuation array of possible choices is limited to the Hilbert cube, by Lemma 1.

Note that the cost function, being linear, meets the lower right-hand corner of each array. Then any non-decreasing and monotonic continuation value must cross the cost function once and only once. Since the continuation value used here is linear, this is sufficient to show a single-crossing property. Hence the optimum is unique.

Because the unbiased Bayes estimate of γ is Markov (or is assumed Markov by the human agent even when, by placebo effects, it is semi-Markov), standard approaches to the optimal stopping of Markov processes (e.g., Shiryaev 1970) may be employed. Note, however, that while $\widehat{\gamma}_t^*$ is fully Markov, $\widehat{\gamma}_t^{felt}$ is not (though it has a Markovian component). That is, if a fully rational agent knew that $\widehat{\gamma}_t^{felt}$was driving the inference process, she would not apply standard optimal stopping approaches for Markovian processes to the problem.

Lemma 3: Existence of Histories Generating Optimal Abandonment of the Incumbent. Let $\overrightarrow{A}_\tau = \sum_{\tau=0}^{t} A_{ijt}$ denote the sum of observed responses (pharmacological or placebo) to time t. We first demonstrate the existence of histories $H_t(\overrightarrow{A}_\tau)$ that induce abandonment of the therapy $[\sigma_t^*(H_t(\overrightarrow{A}_\tau)) = 0 | \sigma_{t-1}^*(H_{t-1}(A_{t-1})) = 1]$. [Note that the incumbent therapy can always be

abandoned in favor of another unknown ($\sigma_t = q$). By substituting the known value of the prior curing value for the "next-best" treatment for the value of the default in what follows, abandonm,ent of the incumbent in favor of any other available treatment follows without loss of generality.) For any period t, optimal abandonment occurs when the continuation value falls below the value of the default, or $V_t^1 = f_t^1(\lambda_1) + \delta \int_{\Re} F_{t+1}^1(\gamma, \lambda_1) d\mu(\gamma) < \eta = \frac{\alpha X^{ILL} - k_0}{1-\delta}$. Note first that $V_t^1\left(\overrightarrow{A}_\tau\right)$ is an increasing function of $\frac{m+\overrightarrow{A}_\tau}{n+t}$, and that $\frac{m+\overrightarrow{A}_\tau}{n+t}$ is strictly decreasing for any failure. Consider then a large sequence of experiments $\{A_t^\chi\}$, with occasional successes but dominated by failure, such that $\limsup_{t\longrightarrow\infty} V_t^1\left(\overrightarrow{A}_\tau\right) \longrightarrow \chi$, with χ an infinitesimal positive quantity ($\chi < \frac{m}{n} < \eta$). Then $\{A_t^\chi\}$ has first-order stochastic dominance over a sequence $\{A_t^0\}$ for which $\limsup_{t\longrightarrow\infty} V_t^1(A_t) \longrightarrow 0$. Because $V_t^1\left(\overrightarrow{A}_\tau\right)$ is monotonic in $\frac{m+\overrightarrow{A}_\tau}{n+t}$, $\{A_t^0\}$ must exist (though it may have Lebesgue measure zero, or negligible probability). [It can be shown, however (Billingsley 1995: Theorem 19.1), that any sequence with this property has at least one convergent subsequence that behaves the same.] Now assume the contrary: subjectively optimal abandonment *never* occurs. This requires, $\forall\left(t, \overrightarrow{A}_\tau\right)$, $\inf_{0<t\leq\infty} f_t^1\left(\overrightarrow{A}_\tau\right) > \eta$. This condition is of course violated for any sequence $\{A_t^0\}$. But then choose any sequence $\{A_t^\chi\}$ for which $0 < \chi < \eta$and for which $(\frac{m}{n} - \chi) > \chi$. Then the path $\{A_t^\chi\}$, which has positive probability and greater likelihood than $\{A_t^0\}$, yields $\limsup_{t\longrightarrow\infty} V_t^1\left(\overrightarrow{A}_\tau\right) < \eta$.

We now turn to singularity of $\sigma_t^* = 0$, namely that for any period t, one and only one value of $\sigma_t^* = 0$ exists. It is sufficient to show that, no matter how refined the type space, only one type can abandon in each period. Let $t^{\sigma_t^*=0}$ be the first period in which abandonment (by type $A_t^{\sigma_t^*=0}$) occurs. For the $T = \infty$ case, this cannot be the first period, else no utilization could ever occur. So $t^{\sigma_t^*=0}$ must have been preceded by a period in which experimentation occured and in which $f_t^1\left(A_{t-1}^W\right) > \eta$. Let $A_{t-1}^{\sigma_t^*=0}$ be the "do-or-die" type in this previous period – i.e, the type which must succeed or get abandoned – of which there is only one possible such type, by the construction of the Beta distribution. Then with probability $\frac{m_t}{n_t}$ type $A_{t-1}^{\sigma_t^*=0}$ observes a clinical response, and by the monotonicity of $f_t^1\left(\overrightarrow{A}_\tau\right)$, this type ($A_{t-1}^{\sigma_t^*=0} + 1$) must consume at $t^{\sigma_t^*=0}$ (else she would have abandoned at $t^{\sigma_t^*=0} - 1$). With probability ($1 - \frac{m_t}{n_t}$) type $A_{t-1}^{\sigma_t^*=0}$ fails and exits the game. So only one type abandons at $t^{\sigma_t^*=0}$. But then the successful $A_{t-1}^{\sigma_t^*=0}$ is the worst remaining type, and for some $t^{\sigma_t^*=0} + j$ (perhaps $j = 1$) there exists one type $A_{t-1}^{\sigma_t^*=0} + 1$, having failed in every period thereafter, which faces a do-or-die experiment again. But this is structurally equivalent to the experiment at $t^{\sigma_t^*=0}$; hence only one type can withdraw at $t^{\sigma_t^*=0} + j$. But j is arbitrary. Since the spread between any types is $\left[\frac{m+\left(A_{t+j}^W+1\right)}{n+t^{\sigma_t^*=0}+j} - \frac{m+\left(A_t^{\sigma_t^*=0}+1\right)}{n+t^{\sigma_t^*=0}}\right]$, which is non-increasing in j, this is true for type space of any positive refinement. QED.

Lemma 4. For any two treatments 1 and q, the subjectively expected value of the first is superior to that of the second if, conditioning upon state variables and parameters, its subjectively expected

continuation value has first-order stochastic dominance over that of the second.

Proof. By definition, the value of the alternative is greater if

$$f_t^{\sigma_t=1}(\widehat{\gamma}_{1jt},\lambda)+\delta\int_{\Re}F_{t+1}^{\sigma_t=1}(\widehat{\gamma}_{1jt},\lambda)d\mu(\gamma)\geq f_t^{\sigma_t=q}(\widehat{\gamma}_{qjt},\lambda)+\delta\int_{\Re}F_{t+1}^{\sigma_t=q}(\widehat{\gamma}_{qjt},\lambda)d\mu(\gamma) \tag{15}$$

It is immediate that a sufficient condition for (15) to be true is that each left-hand side component of 15 has first-order stochastic dominance over its respective right-hand side component. Or

$$f_t^{\sigma_t=1}(\widehat{\gamma}_{1jt},\lambda)\text{ FOSD }f_t^{\sigma_t=q}(\widehat{\gamma}_{qjt},\lambda)$$

and

$$\int_{\Re}F_{t+1}^{\sigma_t=1}(\widehat{\gamma}_{1jt},\lambda)d\mu(\gamma)\text{ FOSD }\int_{\Re}F_{t+1}^{\sigma_t=q}(\widehat{\gamma}_{qjt},\lambda)d\mu(\gamma)$$

Proof of Comment 4. The "truth" is $\widehat{\gamma}_{1,t}^{felt,W}X_j^{ILL}-k_1+\int f_{t+1}^1(\gamma)\,\mu(d\gamma)\leq\widehat{\gamma}_{\chi,t}^{felt,W}X_{\chi j}^{ILL}-k_\chi+\int_{\Re}f_{t+1}^{\chi}(\gamma)\,\mu(d\gamma)$. We seek the conditions under which $\sigma_t=1$ when this is the case, and the (first-order) characteristics of the limiting distribution for this utilization path. The disguised default has $\widehat{\gamma}_{\chi,t}^{felt,W}X_{\chi j}^{ILL}-k_\chi+\int_{\Re}f_{t+1}^{\chi}(\gamma)\,\mu(d\gamma)=\eta$. Then expressed in terms of the functional relation, the human agent errs if

$$\left(\widehat{\gamma}_{1,t}^{felt,W}X_j^{ILL}-k_1\right)+\delta\int_{\Re}f_{t+1}^1\left(\widehat{\gamma}_{1,t}^{felt,W}X_j^{ILL}-k_1\right)\mu(d\gamma)<\left(\beta X_j^{ILL}-k_0\right)+\frac{\delta}{1-\delta}\left(\beta X_j^{ILL}-k_0\right)$$

Because $\int_{\Re}f_{t+1}^1\left(\widehat{\gamma}_{1,t}^{felt,W}X_j^{ILL}-k_1\right)\mu(d\gamma)$ is increasing in $\widehat{\gamma}_{1,t}^{felt,W}$, then for any stopping period τ this occurs if

$$X^{ILL}\left(\widehat{\gamma}_{1j\tau}^{felt,W,c}-\beta)\right)>k_1-k_0>X^{ILL}(\gamma_{1j}-\beta))$$

If this condition holds, then we know from Lemma 3 that the incumbent would be abandoned by a human agent not subject to placebo learning. The placebo-learning agent would, however, continue to utilize the incumbent treatment if (15) held subjectively. By Lemma 4, this condition holds for any incumbent such that $X^{ILL}\left[\liminf\psi_{1jt}^{felt,W}\right]>k_1-k_0$. By Comments 1 - 3, the cost differential k_1-k_0 can always be chosen small enough that this holds for any $\lambda_{1j}>0$.

To see that the probability of a Type I error is increasing in λ_{1j}, consider the probability that (15) holds when $k_1-k_0>X^{ILL}(\gamma_{1j}-\beta))$. The probability that (15) holds is increasing in its left-hand functional components. By Comments 1 and 2, for any $\lambda_{1j}^1>\lambda_{1j}^2$,

$$f_t^{\sigma_t=1}(\widehat{\gamma}_{1jt}, \lambda_{1j}^1) \text{ FOSD } f_t^{\sigma_t=1}(\widehat{\gamma}_{1jt}, \lambda_{1j}^2)$$

and

$$\int_{\Re} F_{t+1}^{\sigma_t=1}(\widehat{\gamma}_{1jt}, \lambda_{1j}^1) d\mu(\gamma) \text{ FOSD } \int_{\Re} F_{t+1}^{\sigma_t=1}(\widehat{\gamma}_{1jt}, \lambda_{1j}^2) d\mu(\gamma)$$

Then by Lemma 4, $\Pr\left[\sup \Phi_t^{I,B}\left(H_t^{felt,W}\right) = 1\right] > \Pr\left[\sup \Phi_t^{I,B}\left(H_t|\lambda_{1j}=0\right) = 1\right]$ follows as a special case. By Comments 1 and 2, again, $\limsup \phi_{1jt}^{I,B}\left(H_t^{felt,W}\right) \geq \limsup \phi_{1jt}^{I,B}\left(H_t|\lambda_{1j}=0\right)$. QED.

Proof of Comment 5. Again the "truth" is $\widehat{\gamma}_{1,jt}^{felt,W} X_j^{ILL} - k_1 + \int_{\Re} f_{t+1}^1(\gamma)\,\mu(d\gamma) \leq \widehat{\gamma}_{\chi,t}^{felt,W} X_{\chi j}^{ILL} - k_\chi + \int_{\Re} f_{t+1}^{\chi}(\gamma)\,\mu(d\gamma)$. By Lemma 4, the agent still chooses the incumbent if

$$f_t^{\sigma_t=1}(\widehat{\gamma}_{1jt}^{felt,W}, \lambda_{1j}) \text{ FOSD } f_t^{\sigma_t=\chi}(\widehat{\gamma}_{\chi jt}, \lambda_{1j})$$

and

$$\int_{\Re} F_{t+1}^{\sigma_t=1}(\widehat{\gamma}_{1jt}^{felt,W}, \lambda_{1j}) d\mu(\gamma) \text{ FOSD } \int_{\Re} F_{t+1}^{\sigma_t=\chi}(\widehat{\gamma}_{\chi jt}, \lambda_{1j}) d\mu(\gamma)$$

Because utilization of the incumbent and the sugar pill are both subject to placebo learning, these conditions hold for any value of the default treatment. From Comment 1 the expected τ-th period bias for the "felt" estimate $\widehat{\gamma}_{1jt}^{felt,W}$ is weakly increasing in $\widehat{\gamma}_{1j0}$. Then $\Pr\left[\sup \Phi_{i\tau}^{I,\chi}\left(H_\tau|\widehat{\gamma}_{1j0}=0\right) = 1\right]$ is weakly increasing in $\widehat{\gamma}_{1j0}$, and for any $\widehat{\gamma}_{1j0} > \widehat{\gamma}_{2j0}$, $\limsup \phi_{ij\tau}^{I,\chi}\left(H_\tau^{felt}|\widehat{\gamma}_{1j0}\right) \geq \limsup \phi_{2j\tau}^{I,\chi}\left(H_\tau|\widehat{\gamma}_{2j0}\right)$. QED.

Proof of Comment 6. The "truth" is

$$\Pr\left[\sup_{t\to\tau} \sum_t Y_{1jt} > \xi_j\right] \delta^{E[\tau]} D_j > \text{and } \Pr\left[\sup_{t\to\tau} \sum_t Z_{1jt}^{felt,W} > \xi_j\right] \delta^{E[\tau]} D_j$$

The agent still utilizes the incumbent if she subjectively perceives that

$$\Pr\left[\sup_{t\to\tau} \sum_t Y_{1jt}(\sigma_t = q^*) > \xi_j\right] \delta^{\tau^{real}} D_j + \sum_{t=0}^{\tau^{real}} \delta^t (Y_{1jt}(\sigma_t = q^*) - k_{q^*}) <$$

$$\Pr\left[\sup_{t\to\tau} \sum_t Z_{1jt}^{felt,W}(\widehat{\gamma}_{1jt}^{felt,W}) > \xi_j\right] \delta^{\tau^{felt}} D_j + \sum_{t=0}^{\tau^{felt}} \delta^t (Z_{1jt}(\widehat{\gamma}_{1jt}^{felt,W}) - k_1)$$

The likelihood of this inequality isincreasing in the right-hand side functional components $\{Z_{1jt}^{felt,W}\}$. By Comments 1 - 3, the relation $\Pr\left[\sup_{t\to\tau}\Phi_{\tau}^{I,D}\left(H_{\tau}^{felt,W}\right)=1\right] > \Pr\left[\sup_{t\to\tau}\Phi_{\tau}^{I,D}\left(H_{\tau}|\lambda_{1j}=0\right)=1\right]$ holds as a special case, and the second statement follows. QED.

Proof of Comment 7. Note that to rule out trivial cases, (15) must be satisfied. By Comments 1-3 and Lemma 4, under this condition or any other in which $\mu_{ij0} > \gamma$, the series $\{A_{ijt}|\lambda_{ij}>0\}$ has first-order stochastic dominance over the series $\{A_{ijt}|\lambda_{ij}=0\} = \{G_{ijt}\}$. For any large series,

$$\Pr\left[\inf\Phi^{I}\left(H_{t}^{felt}\right)=1\right]=1 \tag{16}$$

by a Borel-Cantelli limit. But if $\lambda_{ij}=0$, then $\widehat{\alpha}_{ijt}=\widehat{\gamma}_{ijt}^{*}$ and Type I error occurs only through sampling error. For any finite series, let the probability of abandonment be

$$\Pr\left[V_{t}^{1}\left(A_{t}\left(H_{t}^{i}\right)\right)<\eta\right]=\Pi_{t}\left(V_{t}^{1}\left(A_{t}\left(H_{t}^{i}\right)\right)<\eta\right)=\int_{0}^{\tau}\pi_{t}\left(V_{t}^{1}\left(A_{t}\left(H_{t}^{i}\right)\right)\right)dA. \tag{17}$$

From Comment 3, we know that $V_{t}^{1}\left(A_{t}\left(H_{t}^{i}\right)\right)$ is monotonic in A_t. But because of the first-order stochastic dominance of $\{A_{1jt}|\lambda_{1j}>0\}$ over $\{A_{1jt}|\lambda_{1j}=0\}$,

$$\Pi_{t}\left(f_{t}^{\sigma_{t}=1}(\widehat{\gamma}_{1jt},\lambda_{1j}=0)+\delta\int_{\Re}F_{t+1}^{\sigma_{t}=1}(\widehat{\gamma}_{1jt},\lambda_{1j}=0)d\mu(\gamma)\right)>$$

$$\Pi_{t}\left(f_{t}^{\sigma_{t}=1}(\widehat{\gamma}_{1jt},\lambda_{1j}>0)+\delta\int_{\Re}F_{t+1}^{\sigma_{t}=1}(\widehat{\gamma}_{1jt},\lambda_{1j}>0)d\mu(\gamma)\right)$$

To see that this result holds identically for a set of equivalent histories (H_t^{Ψ}), note that since $\overrightarrow{A}_{\tau}$ is a sufficient statistic for all H_t^{Ψ}, then $\forall H^{i}\in H^{\Psi}$, $i=1,2...\Psi'$, $\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{1}\right)\right)\right)=\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{2}\right)\right)\right)=\cdots=\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{\Psi'}\right)\right)\right)$, i.e., the density is the same for all histories. But then

$$\int_{0}^{\tau}\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{1}\right)\right)\right)=\int_{0}^{\tau}\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{2}\right)\right)\right)dA=\cdots=\int_{0}^{\tau}\pi_{t}\left(V_{t}^{1}\left(\overrightarrow{A}_{\tau}\left(H_{t}^{\Psi'}\right)\right)\right)dA \tag{18}$$

the cumulative distribution function is also identical.

This proves the first statement. For the second, notice that whenever the first-order stochastic dominance condition is satisfied, then by Comments 1 and 2, $\liminf\Pi_{t}\left(H_{t}^{felt}\right)\geq\liminf\Pi_{t}^{*}\left(H_{t}|\lambda_{ij}=0\right)$. ■

For initial beliefs, note that by Comments 1-4, the path of optimal utilization is non-decreasing in $Q_{t,t-1}$. For any t, the quantity $Q_{t,t-1}$ is an increasing function of μ_{ij0}, by Comments 1 and 2. Since no positive $Q_{t,t-1}$ can generate optimal abandonment, utilization is nondecreasing in μ_{ij0}. ■