THE UNIVERSITY OF CHICAGO

PLACEBO EFFECTS, SELF-SELECTION AND THE EXTERNAL VALIDITY OF CLINICAL TRIALS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY
ANUP MALANI

CHICAGO, ILLINOIS
AUGUST 2003

ProQuest Information and Learning Company
300 North Zeeb Road
PO Box 1346
Ann Arbor, MI 48106-1346

To my wife, Rachel Cantor, and my parents, Brij and Prema Malani.

ACKNOWLEDGEMENTS

I am grateful to Scott Baker, Gary Becker, Harvey Cantor, Albert Choi, Feifang Hu, Brij Malani, Willard Manning, Ron Michener, Charles Mullin, Tomas Philipson, Andrew Sellgren and participants in the University of Virginia Public Economics Workshop for helpful comments. I am indebted to David Greene and the staff of the University of Virginia Law Library for excellent research assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In this thesis I examine the effect of voluntary enrollment and placebo effects on the external validity of randomized controlled trials (RCTs), i.e., on the extent to which estimates of treatment response among trial subjects can be employed to draw inferences about treatment response among populations of non-subjects. My analysis starts with the proposition that, where individuals have a choice whether to enroll in an trial, they will enroll only if the expected utility of enrolling is greater than that of not enrolling. An RCT is essentially a lottery with two possible consumption states: treatment or control. The alternative to the trial is certain consumption either of no treatment or of conventional treatment, if available. Each consumption state is in turn a lottery over different health outcomes.

As probabilities regarding consumption state and health outcomes given consumption state may be uncertain, individuals rely on their beliefs about these probabilities to make choices based on expected utility. If all individuals have identical beliefs, all individuals either enroll or do not enroll. If individuals have heterogeneous beliefs, some individuals may enroll and others may not. In general, only those individuals who believe that a test treatment is superior to no treatment will consider enrolling. Of these, only those who believe the test treatment is *so much better* than conventional treatment that, even with the risk of obtaining the control, enrollment is

superior will take the leap and actually enroll. This thesis assumes that individuals have heterogenous beliefs.[1]

If there exist placebo effects, which I define as a positive relationship between health outcomes and not just the specific effects of a treatment strategy[2] but also *expectations* about the specific effects of that treatment strategy, then health outcomes of individuals will depend on beliefs about the efficacy of treatment. For subjects in a blinded trial, outcomes will also depend on beliefs about consumption state, i.e., the probability of randomization into test treatment. I assume subjects learn this probability in trials with informed consent.[3]

---

1. I do not directly address the question of how these beliefs are derived. I simply assume that there is a population distribution of beliefs. I place restrictions on this distribution—at various points that it is log-concave or log-convex or that beliefs about no treatment are independent of beliefs about active treatments. But I do not explicitly justify the existence of a distribution. My analysis of placebo effects does not require me to do so: if beliefs affect outcomes, the mere fact that they are heterogenous and that trials attract only those who believe that the test treatment is superior to alternatives permits testing for placebo effects and causes the standard estimator of treatment response to be biased. When I discuss self-selection I assume that there is a population distribution of treatment efficacy and that treatment response is positively correlated with beliefs about treatment response. The latter rules out the possibility that individuals generate beliefs that are just wrong and is not a significant restriction on the development of beliefs.

2. Abusing medical terminology a bit, I will call the sum of the physiological effects of a treatment and the natural progression of disease the specific effect of that treatment state. Traditionally the specific effect of treatment excludes the natural progression of disease, which, along with placebo effects, are called the non-specific effects of treatment.

3. This assumption is reasonable. Under U.S. law, trial subjects muse be asked for informed consent before they can be enrolled in a clinical trial. See 21 C.F.R. §7.3(f); 45 C.F.R. §46.101(a). Informed consent requires that subjects be informed not only about the test treatment and alternative therapies, but also about the structure of the trial. See, e.g., 21 C.F.R. §20.25. The laws of other developed countries are similar.
It should be acknowledged that the precision with which subjects are informed about the probability of receiving the test treatment varies. In many cases, subjects are explicitly told there is a 50% chance of receiving the control [1]. However, in some cases they may simply be told the number of arms in the trial. Nevertheless, I find strong evidence to support my assumption regarding the content of informed consent: the relationship I find between the share of subjects given test treatment and outcomes in trials with informed

That beliefs affect outcomes implies two things. First, one can test for the existence of placebo effects by checking whether outcomes in the test-treatment or control group of blinded RCTs with a higher share treated are superior to outcomes in the same group of trials with lower a lower share treated. Outcomes will increase with the share treated because only subjects who believe the test treatment is superior to the control will enroll in a trial and because a higher share treated raises the expected benefits of the trial among this subpopulation. I apply this test to a data set of 150 trials of anti-ulcer medications in chapter two and find robust evidence of placebo effects.[4]

The second implication is that placebo effects cause the standard estimator of treatment response—the difference in average outcomes in the treatment and control groups—in blinded RCTs to have a negative bias. A subject in the test-treatment group manifests a lower outcome than would a subject given test treatment but told her treatment allocation because the former thinks that there is a chance that she is in the control group. A subject in the control group has the opposite reaction because she thinks there is a chance she is in the test-treatment arm. The difference in outcomes across groups will underestimate the difference in outcomes if subjects in each arm knew their treatment allocation. The latter difference is the parameter of interest because, outside the context of a trial, individuals know the treatment they consume. I explore this bias and propose an experimental design that permits estimation of treatment response without bias due to placebo effects in chapter three. That design randomizes individuals not just across treatment groups, but also across the probability of being assigned to the test-treatment group. The variation in those

consent vanishes in trials without informed consent.

4. The test is valid and evidence of placebo effects persist despite the presence of self-selection as discussed in chapter four.

probabilities generates variation in beliefs that allows separate identification of the effects of beliefs on outcomes.

If the specific effects of treatment vary across the population, self-selection is also threat to the external validity of clinical trials. Chapter four examines the implications of self-selection under the assumption that subjects' beliefs about treatment response are positively correlated with their treatment response. This assumption is consistent with the view that individuals self-sample and base their beliefs on observations of their own response to treatment. It is also consistent with the assumption—common in economics—of rational expectations.[5] A positive relationship between beliefs and efficacy implies that, because voluntary trials attract subjects who are optimistic about the test treatment, such trials also attract subjects who respond to test treatment better than non-subjects. However, the higher the proportion of subjects treated in a trial, the less optimistic an individual must be to view enrollment as utility maximizing because the costs of the trial—randomization into the control group—are lower. This has two implications. First, one can test for self-selection by checking whether the standard estimator produces lower estimates of treatment response in trials with higher shares of subjects treated. I apply this test to my data set of ulcer trials and find significant evidence of self-selection. Second, because a trial with a share treated of one is the same as certain consumption of the test treatment outside the trial context, self-selection causes the standard estimator of treatment response to have a positive bias.

---

5. This restriction on beliefs does not rule out the possibility that individuals systematically overestimate their response to treatment. All individuals may be overly optimistic but at the same time individuals who are more responsive to treatment may be more optimistic than those who are less responsive.

CHAPTER 2

A FORMAL TEST FOR PLACEBO EFFECTS IN CLINICAL TRIALS, WITH AN APPLICATION TO ULCER TRIALS

Placebo effects can roughly be defined as that component of health outcomes that cannot be attributed to the physiological effects of treatment or to the natural progression of disease. There is a lively debate in the medical literature about whether placebo effects actually exist. On one side of the debate there is, for example, a recent New England Journal of Medicine article by Hrobjartsson and Gotzsche [2] that examined 114 studies with both a blinded placebo-control group and an unblinded no-treatment group. The authors found few systematic differences in outcomes between these groups across their sample. Although widely publicized, this result does not conclusively disprove the existence of placebo effects. It is consistent with the plausible theory that members of unblinded no-treatment groups seek out alternative medication, which elevates their health outcomes.

On the other side of the debate are, e.g., studies by Kirsch and Sapirstein [3] and Kirsch et al. [4] that point to evidence that members of the placebo-control group in a given double-blinded trial manifested substantially improved health outcomes. These findings are weak support for placebo effects because the improvements could be due to the natural progression of disease [5]. Better studies employ a balanced-placebo design wherein subjects are first randomized across treatment states and then across instructions about treatment state, with one group in each treatment state being told

they were given active treatment and the other being told they were given placebo.[1] Such studies generally find evidence in support of placebo effects [8], although in more recent studies the results are mixed [9]. More importantly, these studies are ethically questionable and perhaps even illegal.

One weakness of studies on both sides of the debate is that they do not begin with a formal model of placebo effects that can be clearly falsified. Therefore, it is unclear how powerful their evidence on the existence of placebo effects really is. This chapter addresses these shortcomings by presenting a simple model of how clinical trials are conducted, formalizing the dominant medical theory for how placebo effects operate, and using these two structures to explore how placebo effects alter health outcomes observed in trials.

My model of clinical trials, which focuses on the gold-standard randomized, parallel-group, placebo-controlled trial (RPCT) design, posits that candidates view a trial as a lottery to obtain a new treatment. If a subject wins the lottery, she gets the new treatment. If she loses, she gets a placebo control.[2] If the trial is blinded, the subject never knows which treatment she gets. Due to informed consent, however, she may learn her probability of obtaining the new treatment. Although a trial is the only way a patient can obtain the new treatment, the patient can always choose to forego any treatment or, if available, employ a conventional treatment, with certainty about

1. Moreover, Penick & Hinckle [6] and Penick & Fisher [7] have performed related experiments that randomize across treatment and instruction about treatment efficacy (as opposed to about treatment state). These trials yielded mixed results and, like the balanced-placebo design, are ethically questionable.

2. The analysis in this thesis may be extrapolated to conventional-control trials. The critical difference is that, while one conventional alternative to the test treatment can trigger self-selection in placebo-controlled trials, a second conventional alternative or potential subjects who believe no treatment superior to the conventional control are required to trigger self-selection in conventional-control trials.

her treatment state.

My model of placebo effects is based on the so-called expectancy theory of placebo effects. This theory, which is the most widely supported in the medical literature [10], posits that the more optimistic a patient is about the efficacy of a treatment, the more positive will be her health response to that treatment [11, 12]. Moreover, if the patient is told she is being administered a treatment she thinks will prove helpful, she will manifest an improved health outcome even if she is never in fact given treatment.

Combining my models of RPCTs with the expectancy of placebo effects reveals two features of trial outcomes. First, because subjects are blinded, they do not know their treatment allocation. Therefore, outcomes are a function not just of subjects' beliefs about payoffs in different treatment states, but also their assessment of the probability of being in different treatment states. Second, because enrollment in trials is voluntary, trials attract individuals who are more optimistic than average about the efficacy of the new treatment.

I use these insights to determine the conditions under which my models generate different, testable predictions whether one assumes placebo effects do or do not exist. These predictions can be used to test for the existence of placebo effects. Whether predictions differ depends, it turns out, on the impact of the share given new treatment on self-selection into clinical trials. Consider two different RPCTs of the same new treatment, with the first (HPT trial) offering individuals higher probability of being randomized into the new-treatment group than the second (LPT trial).

Suppose there are only two treatment options available to patients: the new treatment or no treatment. In this case, my models *always* generate different predictions with and without placebo effects. RPCTs attract everyone who believes that the new treatment is more effective than no treatment. Altering the probability of obtaining

the new treatment in a trial does not affect selection into trials. If placebo effects do not exist, one would predict that outcomes in the HPT trial would not significantly differ from those in the LPT trial, conditional on treatment group assignment. If placebo effects exist, outcomes are also a function of expectations about the trial. The fact that trials attract subjects who are more optimistic about the new treatment than the placebo means that an increase in the probability of getting the new treatment raises expectations about the trial among enrollees. Therefore, the HPT trial will produce *better outcomes* than the LPT trial, conditional on treatment group assignment.

If there is a conventional treatment option, my models continue to generate different predictions with and without placebo effects, but these prediction require additional (but reasonable) conditions on beliefs. If placebo effects do not exist, I predict that outcomes will fall or stay constant as the probability of randomization into the new-treatment group increases. The intuition is that an RPCT only attracts individuals who believe that the new treatment is so much better than conventional treatment that, even with the risk of obtaining no treatment at all, enrolling in the trial is a superior strategy to opting for conventional treatment. If one increases the probability of obtaining the new treatment, a patient who is marginally not optimistic enough about the new treatment or no treatment to have risked randomization into the placebo-control group before may now be willing to take that risk because it is smaller. The result will be lower average beliefs about new treatment or no treatment among enrollees. In the absence of placebo effects, more pessimistic beliefs about new treatment or no treatment among enrollees have no effect on outcomes unless beliefs and physiological effects are correlated, e.g., due to subjects' own assessments of the efficacy of different treatments. Because these assessments surely imply a positive

correlation between beliefs and physiological effects, outcomes observed in the HPT trial will be inferior to outcomes in the LPT trial.

Once placebo effects are introduced into the analysis, however, one cannot be sure that outcomes will fall in the probability of consuming the new treatment. While self-selection exerts negative pressure on outcomes, the fact that all enrollees believe that the new treatment is superior to no treatment implies that an increase in the probability of obtaining the new treatment raises expectations about the trial among those who enroll. If expectations influence outcomes, this elation will place upward pressure on outcomes due to blinding. The net effect is ambiguous. If, however, one observes that outcomes in the HPT trial are *superior* to outcomes in the LPT trial, it must be that there exist placebo effects.

I apply this test for the existence of placebo effects to data from over 150 RCTs of anti-ulcer medications. Each of the trials is parallel arm and double blinded. In trials where patients were asked for informed consent and thus had some indication of their odds of obtaining active treatment, I find a positive correlation between the probability of obtaining active-treatment and outcomes in each treatment group of trials of H_2-blockers (e.g., Zantac, Tagamet, and Pepcid) and of proton-pump inhibitors (e.g., Prilosec, Nexium, and Prevacid) controlling for available group-level clinical covariates and study-level design covariates. This positive correlation is significant and robust to covariate specification. In trials without informed consent, this correlation vanishes. Under very reasonable assumptions, this result can only be explained by the existence of placebo effects.

Section 2.1 presents a simple model of RPCTs. Section 2.2 reviews the expectancy theory of placebo effects. Section 2.3 derives testable predictions of the model from Section 2.1 with and without placebo effects. Section 2.4 tests these predictions

against data from ulcer trials. The appendices present extensions of my analysis to different assumptions about individuals' formation of beliefs, health outcomes and different types of trials and present proofs to the propositions in the text.

2.1 A Model of Medical Trials

Suppose there are three treatments, indexed by subscript k: no treatment ($k = 0$), the new or test treatment ($k = 1$), and a conventional treatment ($k = 2$). The y_{ki} is individual i's health outcome given treatment k. Initially assume that outcomes are univariate and binary, with $\bar{y} > \underline{y}$ indicating recovery from illness and $\underline{y}$ indicating continued illness.[3] Define $p_{ki} = \Pr\{y_{ki} = \bar{y} | \text{no placebo effects}\}$. This probability is a function of the physiological effects of treatment k and the natural progression of individual i's ailment. Abusing medical terminology a bit, I will call the sum of physiological effects and natural progression the specific effect of a treatment state.[4]

There are four treatment strategies, indexed by s, that individual i can follow: she can refuse treatment ($s = 0$); obtain the test treatment ($s = 1$); obtain conventional treatment ($s = 2$); or she can enroll in a RPCT ($s = BT$). My analysis starts with the assumption that the trial takes place before the test treatment is approved for use by the population at large. Once the test treatment is approved, no additional trials are conducted. These assumptions imply that the trial, when conducted, is the only opportunity an individual has to obtain the new treatment. However, the trial is a lottery. An individual will obtain the test treatment only if she is randomized into the test-treatment group.

---

3. The analysis can easily be extended to the case of continuous outcome variables.

4. Traditionally the specific effect of treatment excludes the natural progression of disease, which, along with placebo effects, are called the non-specific effects of treatment.

RPCTs take place in three stages. First, individuals are recruited. Enrollment is voluntary and subject to informed consent. As part of this disclosure, I assume subjects are given information about the probability that they will receive the new treatment. Individuals who choose to participate are either called enrollees or subjects. In the second stage, subjects are randomized into the new-treatment or a placebo-control group. Let $d \in [0, 1]$ indicate the share of subjects randomized into the new-treatment group and $D_i = 1$ or 0 indicate whether subject i was assigned to the new-treatment group or not, respectively. Subjects in unblinded trials are told, after randomization, the group to which they are assigned. In blinded trials, the trials with which I am concerned, they are not. In the third stage of medical trials, subjects in the new-treatment group receive the new treatment; subjects in the control group receive a placebo.

In order to focus the analysis on selection *into* medical trials, I make two assumptions about subject behavior. First, there is no unblinding in blinded RPCTs. Subjects do not discover their group assignment through, e.g., subject sampling [13] on outcomes or side effects. This controls attrition from blinded trials. Second, although subjects are free to exit a trial at any time, if they remain, they do not consume any treatment other than that which they have been assigned. This stops subjects from obtaining conventional treatment in addition to their assigned treatment. Given these assumptions, voluntary enrollment ensures subjects consume the treatment they have been assigned. Subjects enter only if they believe that their expected treatment is better than any individual outside option.

Individual i's belief about the probability of recovery in treatment state k is π_{ki}. For convenience I assume that individuals do not take placebo effects into account

when formulating their beliefs regarding treatment.[5] This assumption is not essential to the analysis. In Appendix B, I present the conditions under which the conclusions of this chapter hold even if individuals are aware of placebo effects. Individuals i's belief about the probability that she is in treatment state k is δ_{ki}, where $\sum_{k=0}^{2} \delta_{ki} = 1$. Outside the context of a trial, an individual knows for sure the treatment she consumes, so $\delta_{ki} = 1$ for $k = 0$ or 2. In the context of a blinded trial, the individual knows she will obtain either the new treatment or no treatment, but does not know which. Because of informed consent, the individual knows she will receive the new treatment with probability d. Given full compliance, the individual will set $\delta_i = (\delta_{0i}, \delta_{1i}, \delta_{2i}) = (1-d, d, 0)$ upon enrollment.

Individuals draw utility from health and other items, including wealth. I assume, for simplicity, that all individuals have identical utility functions and that these functions are additively separable in health and other items. Let $u(y)$ be the utility from health outcome y, with $u' > 0$, $u'' < 0$. The expected utility of strategy s to individual i is weighted sum of the utility from each health outcome, with the weights being her subjective beliefs about the probability of each health outcome given s: $U_i^s = \pi_i^s u(\bar{y}) + \left(1 - \pi_i^s\right) u\left(\underline{y}\right)$.[6] Given that an individual knows her treatment state outside the context of a trial, her subjective belief is $\pi_i^s = \pi_{ki}$, where $k = s$, for $s = 0, ..., 2$. Belief about the probability of recovery given the strategy of enrolling in a trial depends on the probability of being in the test-treatment. Informed consent reveals the latter probability to be d, so the former is $\pi_i^{BT} = d\pi_{1i} + (1-d)\pi_{0i}$.

In order to determine the sorting of individuals to strategies, one must know the

5. This assumption draws support from Berthelot et al.'s [14] finding that three-quarters of the 300 rheumatology patients they surveyed did not know about placebo effects.

6. This model of treatment choice and the model of health outcomes with placebo effects in the next section are compatible with Savage's axiomatization of subject expect utility. I demonstrate this in Appendix A.

distribution of beliefs about the efficacy of each treatment state among the population. Since the object of these beliefs is the actual efficacy of each treatment, let $\mathbf{g}_p$ give the probability distribution function of $\mathbf{p}_i = (p_{0i}, p_{1i}, p_{2i})$ across the population. Let $\mathbf{g}_\pi$ give the probability distribution for $\boldsymbol{\pi}_i = (\pi_{0i}, \pi_{1i}, \pi_{2i})$. I assume that treatment efficacy and beliefs about efficacy are related by $p_{ki} = f(\pi_{ki}) + \varepsilon_{ki}$, where either (a) f is an affine transformation with $f' > 0$ or (b) f is strictly increasing, differentiable, concave and $\left(f^{-1}\right)'$ is log-concave, and where ε_{ki} is independent of $\pi_{k'i'}$ for all $\left(k', i'\right)$ and of $\varepsilon_{k'i'}$ for all $\left(k', i'\right)$ except $\left(k' = k, i' = i\right)$. That $f' > 0$ is reasonable: $f' < 0$ implies that patients do not merely over or underestimate efficacy, but rather that they guess better treatments are worse. The error term ε_{ki} reflects error in predictions of relative efficacy by individual i. I assume that all distributions discussed are well-defined. Whenever I employ the expectations operator $E(\cdot)$, I do so with respect to the joint distribution of $(\mathbf{p}_i, \boldsymbol{\pi}_i)$ unless otherwise indicated.

2.2 The Expectancy Theory of Placebo Effects

According to the expectancy theory of placebo effects, patients manifest changed health outcomes in response to expectations regarding treatment. In particular, the more effective a patient expects a treatment to be, the better her response to it. Moreover, the more likely a patient thinks she is to get a beneficial treatment the better is her health outcome holding constant whether or not she receives treatment.[7]

7. There are two alternative theories of placebo effects that have received substantial attention in the medical literature. The conditioning theory suggests that, if the body has previously experienced a specific health response after consumption of a medication, consumption of a substance close in appearance or smell to that medication will trigger a similar non-specific health response [15, 16]. The reaction is at the sub-conscious level. No well-formed beliefs about efficacy are involved. The motivation theory posits that individuals with a stronger desire to respond to treatment experience placebo effects when treated [10].

While a significant number of studies have documented these responses, the studies in Table G.1 are representative.

These studies, and others like them, provide important insights into the nature of placebo effects. They do not, however, provide compelling evidence in support for the existence of placebo effects. First, the studies have small sample sizes, as measured by the number of trials in which placebo effects are observed. Second, many of the trials examine subjective rather than objective criteria. For example, the studies cited by Skovlund and Pollo et al. relied upon patient self-reports of pain levels. One of the goals of this chapter is to test for the existence of placebo effects in trials with an objective measure of outcomes.

2.2.1 Definition of placebo effects

Following the expectancy theory of placebo effects, I define a positive placebo effect to exist for individual i if two conditions hold. First, conditional on information the individual has about her treatment state and the efficacy of treatments, the individual's expectations about her health outcome, including only the specific effects of treatment, are greater than some arbitrary cut-off q_{ki} in state k. In the notation from the previous section, this implies that, e.g., if an individual is in the new-treatment state, then $\pi_i^s > q_{1i}$ given treatment strategy s. If such an individual is in the no-treatment state, then the condition is $\pi_i^s > q_{0i}$. The second condition is that the individual manifests a health outcome, including both specific effects and placebo effects, greater than the specific effects of her treatment state. In the new-treatment state, this means that $y_{1i} > p_{1i}$. In the no-treatment state, this means $y_{0i} > p_{0i}$.

Thus, the individual can be said to have experienced a positive placebo effect in the new-treatment state if she knew this ($\delta_{1i} = 1$) but believed that the new

treatment is sufficiently effective ($\pi_{1i} > q_{1i}$), *and* she experienced an outcome greater than p_{1i}. The individual can be said to have experienced a positive placebo effect in the no-treatment state if she thought that she was given the new treatment ($\delta_{1i} = 1$), believed that the new treatment is sufficiently effective ($\pi_{1i} > q_{0i}$), *and* she actually experienced an outcome greater than p_{0i}.

I define a negative placebo (or the so-called "nocebo") effect to exist under opposite conditions. First, conditional on the information the individual has about her treatment state and the efficacy of treatments, the individual's expectations about her health outcome, including only specific effects, are less than the arbitrary cut-offs mentioned earlier. This means that $\pi_i^s < q_{1i}$ in the new-treatment state and $\pi_i^s < q_{0i}$ in the no-treatment state, given strategy s. Second, the individual must manifest a health outcome, including specific and placebo effects, in each state that is less than the specific effect of that treatment state. So, in the new-treatment state, $y_{1i} < p_{1i}$ and, in the no-treatment state, $y_{0i} < p_{0i}$.

The purpose of requiring expectations to exceed a certain tipping point q_{ki} in state k before placebo effects are defined to exist is to permit the existence of nocebo effects and to relate such effects in a simple manner to positive placebo effects. Without a tipping point, it is unclear how one would define nocebo effects to exist. An obvious candidate for the cut-off is $q_{ki} = p_{ki}$, i.e., the specific effects of treatment k. It seems reasonable to suppose that a positive placebo effect exists if a patient expects more from a treatment strategy than the specific effect of treatment k ($\pi_i^s > p_{ki}$) and manifests a health outcome after treatment k that is greater than the specific effect of that treatment ($y_{ki} > p_{ki}$).[8] Table G.2 summarizes, for an individual in

8. The difficulty with this definition is that p_{ki} may not be known to the patient. (Knowing p_{ki} would not rule out placebo-type effects, but would confine them to the blinded trial setting. If the p_{ki} were known, the subject would set $\pi_{ki} = p_{ki}$. In a blinded trial, however,

a blinded trial, the conditions under which positive and negative placebo effects are said to exist.

2.2.2 Model of health outcomes

Suppose that the health outcome for individual i in treatment state k is given by

$$\Pr\left\{y_{ki} = \bar{y} | a_k, p_{ki}, \pi_i^s\right\} = p_{ki} + a_k\left(\pi_i^s - p_{ki}\right), \tag{2.1}$$

where $a_k \in [0,1]$, for all s, k. This parameterization posits that health outcomes are the sum of the non-placebo effects of treatment state k plus a placebo effect driven by expectations given strategy s. Whether the placebo effect is positive or negative depends on whether the individual's beliefs given her treatment strategy, π_i^s, are greater or less than the specific effects of her treatment state. Beliefs are determined by strategy and not merely treatment state because in a blinded trial, e.g., subjects do not know their true state. The total placebo effect is the difference between an individual's beliefs and reality, weighted by the parameter a_k in treatment state k.

---

the subject may be uncertain about her treatment state, thus π_i^{BT} would deviate from p_{ki} for $k = 0, 1$. Thus placebo effects are possible. It should be noted, however, that there are countless examples of placebos successfully being used in medical interactions outside the trial setting [17].) Moreover, there is no proven psycho-physiological reason for why positive placebo effects would only exist when beliefs are greater than specific effects. Why would the body wait until expectations exceeded the specific effect of treatment to manifest a health outcome greater than the specific effect of treatment?

Nevertheless, in the next subsection I present a model of health outcomes that takes the specific effect of a treatment state as the cut-off for beliefs in that state before positive or negative placebo effects are triggered. I do so for two reasons. First, the test I recommend for the existence of placebo effects are valid even with an arbitrary cutoff, so long as the correlation between the cutoff and beliefs is not too negative. Second, using specific effects as cutoffs permits me to model observable health outcomes as a weighted sum of the specific effects of treatment states and beliefs about treatment state. This is simple and intuitive assumption about the relationship between outcomes, treatment, and beliefs.

The simple linear formulation of expectancy theory in (2.1) implies that health outcomes are a weighted average of the specific effects of a treatment and beliefs about treatment strategy. It is easy to see this when the model is re-written as $\Pr\left\{y_{ki} = \bar{y} | a_k, p_{ki}, \pi_i^s\right\} = (1 - a_k)\, p_{ki} + a_k \pi_i^s$. The weighting parameter a_k indicates the relative importance of beliefs in determining the health outcome in treatment state k. If $a_k = 0$, beliefs have no effects on outcomes in state k. If $a_k = 1$, outcomes are completely determined by beliefs in state k.[9] I assume the influence of placebo effects may vary across treatment states but is constant across individuals. The former assumption acknowledges that the influence of beliefs may depend on the chemical process by which a treatment operates. For example, a treatment may be so effective that a patient will be cured regardless of beliefs or a treatment may actually inhibit the mechanism by which beliefs affect outcomes. The purpose of the latter assumption is to facilitate application of my test for the existence of placebo effects to data from ulcer trials. Those data are aggregated to the level of treatment groups so estimation of a random effects model is not feasible.[10]

9. I assume that $a_k < 0$ is not possible. Such a parameterization would be consistent with a model of health outcomes where individuals manifest nocebo effects if a drug does not live up to expectations. I have not found any empirical support for this sort of regret-based model of outcomes in the literature on placebo effects.

10. One implication of this assumption is that the influence of placebo effects is independent of individuals' beliefs about the efficacy of different treatment states (though not of treatment state per se). This implicit assumption is consistent with my earlier assumption that individuals form beliefs about the efficacy of treatment states without taking into account placebo effects. Moreover, there seems to be no medical basis for thinking that individuals who are more or less responsive to treatment, who are more less likely to recover naturally from an ailment, or who are more or less optimistic about the efficacy of treatment, manifest outcomes that are more or less driven by placebo effects. Nor can I think of any logical reason why individuals whose outcomes are more or less driven by expectations would believe that treatment is—from a physiological perspective—more or less effective.

2.3 Tests for the Existence of Placebo Effects

My models of RPCTs and health outcomes can be used to generate predictions regarding health outcomes observed in trials given parameters that define trials, e.g., blinding, informed consent and the probability of randomization into the new-treatment arm. If predictions about the relationship between outcomes and observed parameters depend on whether one assumes there exist placebo effects, then the difference in predictions may be used to test for the existence of placebo effects. This section applies the foregoing logic to changes in the probability of randomization into the new-treatment arm.[11]

2.3.1 Role of self-selection

In a blinded trial, the outcomes the investigator observes in the new-treatment and placebo-control groups are

$$\begin{aligned} E\left[y_{ki}|s=BT,d\right] &= (1-a_k)\,E\left[p_{ki}|s=BT\right] \\ &\quad +a_k\left\{dE\left[\pi_{1i}|s=BT\right]+(1-d)\,E\left[\pi_{0i}|s=BT\right]\right\} \end{aligned}$$

for $k=0,1$, respectively, where the dependence of s on d has been suppressed for notational convenience. A change in the probability d of randomization into the new

11. This parameter is readily observed by investigators in already-completed trials and is a control variable in the design of new trials. Due to informed consent, the parameter is typically observed by subjects in completed trials and can easily be revealed to subjects in future trials. Moreover, manipulation of d generates fewer ethical problems than, e.g., elimination of informed consent, and more manageable problems of attrition than, e.g., unblinding. Finally, manipulation of d generates more controlled and replicable changes in the beliefs of subjects than manipulation of beliefs $\boldsymbol{\pi}_i$ about the efficacy of treatment states through, e.g., the content of informed consent or the balanced-placebo design.

treatment group has both direct and indirect effects on observed outcomes.

$$\frac{\partial E\left[y_{ki}|s=BT,d\right]}{\partial d}=\frac{\partial E\left[p_{ki}|s=BT\right]}{\partial d} \tag{2.2}$$

$$+d\frac{\partial E\left[\pi_{1i}|s=BT\right]}{\partial d}+\left(1-d\right)\frac{\partial E\left[\pi_{0i}|s=BT\right]}{\partial d} \tag{2.3}$$

$$+E\left[\pi_{1i}-\pi_{0i}|s=BT\right] \tag{2.4}$$

The direct effect (2.4) is positive. An increase in d increases subjects' expectations for the clinical trial. Only those individuals who are more optimistic about the efficacy of the new treatment than of no treatment enroll in a trial. (The rest forego treatment.) An increase in d raises the expectations of enrollees in proportion to the what is believed to be the advantage of the new treatment over no treatment. Since health outcomes in blinded trials are proportion to expectations for the trial, an increase in d should boost outcomes.

An increase in the share of subjects given new treatment has two additional, indirect effects. Both operate through the self-selection of subjects into clinical trials. Because only individuals who believe that the new treatment is better than no treatment ever consider enrollment in a medical trial, an increase in d will be seen as an increase in the value of the trial to this population. Because the trial now offers a higher probability of "winning" the new-treatment lottery, it will attract a larger number of subjects than before. This expanded population (2.3) may not have the same beliefs (π_{0i},π_{1i}) about the efficacy of the new treatment or no treatment. This is the first indirect effect of a change in d.

The second indirect effect (2.2) depends on the relationship between beliefs about the efficacy of treatment and the specific effects of treatment. If beliefs are correlated with specific effects, then selection, which takes place on the basis of beliefs, will alter

the specific effects experienced by subjects, and thus the outcomes manifested by those who enroll in a trial. It seems reasonable to suppose that beliefs about the efficacy of a given treatment state are, if anything, positively correlated with the specific effects of that state. (This is implied by my assumption that $f'(x) > 0$.) While patients may make systematic errors in judgment, few people would argue that individuals who are physiologically the most responsive to a given treatment will have the most pessimistic expectations for that treatment. A positive correlation between beliefs and specific effects implies that if, for example, an increase in d raises the average expectations about the efficacy of new treatment or of no treatment in the enrolling population, then it will also raise the specific effects manifested among that population who are ultimately randomized into the new-treatment group or the placebo-control group, respectively.

If there are placebo effects, the effect of a change in d depends on the net impact of the direct and indirect effects just described. If there are no placebo effects, a change in d has no direct effect and only the second indirect effect on outcomes. Because the direct effect is always positive, predicting the net effect of a change in d in the case with placebo effects versus the case without placebo effects depends on the effect of a change in the share of subjects given new treatment, due to selection, on the average beliefs of subjects regarding the efficacy of new treatment and of no treatment: $\partial E[\pi_{ki}|s = BT]/\partial d$.

If this derivative is zero, then d only has a direct effect and outcomes rise with placebo effects and remain constant without. If the derivative is negative, a change in d has a negative effect on outcomes without placebo effects. When there exist placebo effects, the direct and indirect (selection) effects of d work in opposite directions. In this case, a strong test for the existence of placebo effects is whether outcomes improve

after an increase in d. Given the null hypothesis that placebo effects do not exist, the drawback of this test is a large risk of type II error. There may be placebo effects but the test cannot reject the null. Finally, if the derivative is positive, the effect of d on outcomes is positive regardless of whether there are placebo effects. In this case, there is no good test for placebo effects based on changes in the share of subjects randomized into the new treatment.

2.3.2 Conditions for identification of placebo effects

Without further information about $\mathbf{g}_\pi$ or the structure of clinical trials one cannot determine the sign of the relationship between changes in the probability of randomization into the new-treatment group and beliefs. In this subsection I explore a number of plausible assumptions that ensure the sign of the derivative of beliefs with respect to d is non-positive and thus that one can test for the existence of placebo effects by checking for a positive relationship between outcomes and d.

The following proposition describes the most obvious assumption that permits testing for the existence of placebo effects.

Proposition 1 *If the distribution of* π_{2i} *is degenerate at* $\overline{\overline{\pi}}_2 \leq \min \pi_{0i}$, *then* $\partial E\left[\pi_{ki} | s = BT\right] / \partial d = 0$.

In this case the self-selection equation for the trial is $d\pi_{1i} + (1-d)\,\pi_{0i} \geq \pi_{0i}$ or, equivalently, $\pi_{1i} \geq \pi_{0i}$. Any individual who prefers the new treatment to no treatment will enroll. Increasing d does not sway any individual who would not have previously enrolled in the trial to do so. Therefore, d has only direct effects on observed outcomes.

There are a number of circumstances in which one may assume the distribution of π_{2i} is degenerate at $\pi_{2i} \leq \min \pi_{0i}$. One is where there is no conventional treatment

option. Here the distribution of π_{2i} is degenerate at 0. Another circumstance is where clinical trials compare a new treatment to conventional treatment, but not to no treatment and it is reasonable to suppose that everyone in the population believes the new or conventional treatment is superior to refusing any treatment. In this context, one can simply switch the labels on $k = 0$ and 2 and proceed. A final circumstance is where the trial compares new treatment to no treatment, but both the new-treatment and the placebo-control groups also receive the conventional treatment. Again, one would also have to assume that everyone in the population believes the new treatment or conventional treatment is superior to no treatment. The outside options with such a trial are no treatment or conventional treatment. However, the control group is at least as beneficial as the conventional treatment and by assumption no one prefers no treatment to conventional treatment.

A second assumption which I conjecture permits testing for the existence of placebo effects is that the correlation ρ_{12*} between beliefs about the new treatment and conventional treatment is sufficiently small. Exactly how low ρ_{12*} and $E(\pi_{2i})$ must be can be difficult to pin down. They depend on other features of the distribution of beliefs. I will solve for exact thresholds in a simple example in a moment. In any case, the second assumption is probably reasonable where, for example, the physiological mechanism by which the new treatment operates is different than the physiological mechanism by which the conventional treatment operates. This assumption ensures that average beliefs about the new or no treatment among enrollees do not rise with the share randomized into the new treatment.

The intuition behind the second assumption is that an RPCT only attracts individuals who believe that the new treatment is so much better than conventional treatment that, even with the risk of obtaining no treatment at all, enrolling in the

trial is a superior strategy to opting for conventional treatment. If one alters a trial to increase the probability of obtaining the new treatment, a patient who is marginally not optimistic enough about the new treatment to have risked randomization into the placebo-control group before may now be willing to take that risk because it is smaller. Moreover, because trials only attract subjects who believe that the new treatment is superior to no treatment, altering a trial to increase the probability of obtaining the new treatment also means that the trial will be attractive to individuals who were previously just marginally too pessimistic about no treatment to enroll because there is now less risk of obtaining a placebo. So long as beliefs about each treatment state are positively correlated with the physiological effects of that treatment state—an implicit but reasonable assumption, selection alone implies that outcomes observed in trials with higher d should be superior to outcomes observed in trials with lower d.

This intuition is valid only if individuals who are optimistic about the new treatment aren't too optimistic about conventional treatment as well. If individuals who are more optimistic about the new treatment are also (sufficiently) more optimistic about conventional treatment, individuals who are more optimistic about the new treatment are not more likely to join a trial at any given level of d. Although the value of trial is higher given these individuals' optimism about the new treatment, so is their optimism about the conventional alternative. If these individuals are sufficiently optimistic about the alternative, they may prefer it to enrollment in the trial despite their high expectations for the new treatment.[12]

12. Consider the following numerical example. Suppose that one-half the population has beliefs $(\pi_{0i}, \pi_{1i}, \pi_{2i}) = (0.5, 0.5, 0.5)$ and one-half has beliefs $(0.75 - 1.5\varepsilon, 0.75 + \varepsilon, 0.75)$. A trial with $d = 0.5$ will attract only the first group of individuals. A trial with $d = 0.6$, however, will attract both groups. Yet the average subject in the second trial would be more optimistic than the average person in the first trial with respect to the efficacy of both the new treatment and no treatment. The reason is that correlation of π_{1i} and π_{2i} in the population is very high (specifically, greater than $d = 0.5$).

To support my conjecture, consider the simple case where the physiological effects of new or conventional treatment and the natural progression of disease are thought to be additive and independent, so $\pi_{ki} = \pi_{0i} + \varepsilon_{ki}$ and $E(\varepsilon_{ki}|\pi_{0i}) = E(\varepsilon_{ki})$ for $k = 1, 2$. Additivity implies that subjects for whom $\varepsilon_{1i} > \varepsilon_{2i}/d$ enroll. Clearly, independence of ε_{ki} and π_{0i}, for $k = 1, 2$, implies that selection pressures due to changes in d do not affect the distribution of beliefs about no treatment among enrollees. The following proposition indicates precisely how low the correlation between the log of physiological effects of new and conventional must be to ensure that average beliefs about the new treatment among enrollees do not rise with d.

Proposition 2 *Suppose* $\varepsilon_{ki} > 0$*, for* $k = 0, 1$*, and* $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$ *have non-degenerate log-concave or log-convex densities with mean* (μ_1, μ_2) *and variance* Σ*. Define* $u_{ki} = \ln \varepsilon_{ki} - \mu_i$*,* $W_i = u_{1i} - u_{2i}$*,* $\sigma = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$*,* $a_1 = (\sigma_{11} - \sigma_{12})/\sigma$*,* $a_2 = a_1 - 1$*, and* $V_i = a_1 u_{2i} - a_2 u_{1i}$*. By construction* $u_i = a_i W_i + V_i$*, where* W_i *and* V_i *are uncorrelated. Suppose* W_i *and* V_i *are independent. Define* $\rho_{12} = corr[\ln \varepsilon_{1i}, \ln \varepsilon_{2i}]$*. If* $\rho_{12} < \sigma_1/\sigma_2$*, then* $\partial E[\pi_{1i}|\varepsilon_{1i} > \varepsilon_{2i}/d] \leq 0$*.*

This result is fairly general. First, the class of log-concave or log-convex densities is quite large. It includes, e.g., the bivariate normal distribution. In that case, the fact that W_i and V_i are (by construction) uncorrelated implies they are also independent. If $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$ are normal, $(\varepsilon_{1i}, \varepsilon_{2i})$ are log normal. Second, truncation does not alter the result. See Proposition 6 in An [18].

According to Proposition 2, so long as the covariance between (log) beliefs about the physiological effects of the new and conventional treatments is less than the variance of (log) beliefs about the new treatment, changes in d, if anything, reduce enrollees' expectations regarding the new treatment. From the definition of the correlation coefficient, it is obvious that the condition on the covariance is satisfied whenever

the variance of beliefs about the physiological effects of the new treatment is greater than the variance of beliefs about the conventional treatment. In this light, the condition on the covariance does not appear at all unreasonable. The new treatment, by virtue of being new, will be associated with greater uncertainty in beliefs among the patient population than with the conventional treatment.[13]

In addition to the restrictions on $\mathbf{g}_\pi$ above, there are two other assumptions which permit testing for the existence of placebo effects. One assumption is that investigators are able to suppress selection based on d or estimate the impact of changes in d on observed outcomes, holding constant selection effects. For example, if one offers a monetary participation incentive greater in value than recovery from illness, at least for the duration of the trial, for every individual i, then selection ceases to be a problem. The trial will, in effect, draw a random sample from the population. Alternatively, if one were able to gather data on individuals who choose not enroll in two trials with different d, one could use this information to estimate the direct effect of the change in d on outcomes across the two trials. This could be accomplished, e.g., by gathering demographic and clinical data from individuals who respond to advertisements for the trials but choose not to enroll after having the trial risks described as part of an informed consent procedure.

---

13. Using a Taylor-series approximation, one can approximate the condition that $\rho_{12} < \sigma_1/\sigma_2$ for $(\ln \varepsilon_1, \ln \varepsilon_2)$ with the condition that $\rho_{12*} < (\mu_{2*}/\mu_{1*})\,(\sigma_{1*}/\sigma_{2*})$, where (μ_*, Σ_*) are the mean and variance of $(\varepsilon_{1i}, \varepsilon_{2i})$. Where the mean belief about the physiological effect of new treatment is greater than the mean belief about that of the conventional treatment, the condition on the correlation between those beliefs is more constriaing.

2.4 Testing for Placebo Effects in Ulcer Trials

In this section I test for the existence of placebo effects in clinical trials of anti-ulcer medications that promise to heal ulcers. I choose ulcer trials because they offer objective measures of health outcomes. Ulcers are erosion of the mucous lining in the stomach or small intestine due to acid buildup. Ulcers can objectively be judged healed via endoscopy, which examines the stomach lining for evidence of damage. Moreover, because ulcers are such a common problem throughout the world, a large number of trials have been conducted.

The majority of ulcer trials examine three types of medication. The first type, H_2-blocker, was introduced in 1977. The most popular brands are Tagamet (cimetidine), Zantac (ranitidine), and Pepcid (famotidine). H_2-blockers prevent the production of acid in the stomach.[14] The second type of medication, prostaglandin, was introduced in 1987. The most common prostaglandins are misoprostil and enprostil. These drugs build up and thus repair the mucous lining of the stomach and intestine. The third class, proton-pump inhibitor, and were introduced after prostaglandins. The most popular brands are Prilosec (omeprazole), Nexium (esomeprazole) and Prevacid (lansoprazole). Like H_2-blockers, these medications prevent the production of acid in the stomach.[15]

---

14. In contrast, antacids are alkali that absorb acid in the stomach. They reduce the amount of acid available to damage the stomach lining, but not enough to permit the healing of damage to that lining.

15. A second reason I examine ulcer trials is that de Craen et al. [19] and Moerman [20] claim to find evidence of placebo effects in such trials. The de Craen study finds that outcomes in the placebo group of trials with a four-times-a-day (q.i.d.) regimen of placebo treatment are 6 - 8 percent higher than in placebo groups of trials with a twice-a-day (b.i.d.) regimen. This is highly suggestive of placebo effects. However, the result does not differentiate between types of ulcer medication. When the authors examined only trials of H_2-receptor antagonists, the difference dropped to 3 percent. (There was no indication whether this estimate is statistically significant.) Moreover, the authors do

2.4.1 Data

My data set includes the published results from over 150 clinical trials studying treatment for pyloric, pre-pyloric and duodenal ulcers.[16] Each of the trials is randomized, parallel-armed, and double-blind, and employs either a placebo, antacid, bismuth subcitrate or conventional control. If conventional controls are employed, they are from either the same or previous class of medication as the test treatment. Importantly, subjects in 110 of the trials were asked for informed consent prior to enrollment. Hence it is reasonable to suppose that subjects in those trials had some indication of their chance of obtaining the new treatment.

I gathered data on the characteristics of trials and subjects. Data on subjects are aggregated to the arm- or group-level. For example, I have data on the average age of subjects assigned to any given treatment group, but not the age of each subject

---

not relate their results to any specific models of placebo effects. Nor do they formally consider the interaction between the structure of the randomized, double-blind, placebo-controlled trials they examine and any theory of placebo effects they purport to test. This is particularly important if, e.g., patients in q.i.d. trials receive the same total dosage as or different treatment than those in b.i.d. trials. (In the sample of ulcer trials I consider, all 12 of the b.i.d. trials examined ranitidine while 29 of 30 q.i.d. trials examined cimetidine. Both are H_2-blockers, but require substantially different dosages. Moreover, the one q.i.d. trial of ranitidine involved the same total daily dosage as the 12 b.i.d. trials of the drug.) In that case, the result would be difficult to explain even with the model of health outcomes in Section 2.2. Nevertheless, the de Craen study's findings are very interesting, and ought to be subject to the type of analysis in this paper.

Moerman examines the same set of ulcer trials and finds a significant positive correlation (0.49) between outcomes in the placebo group and the treatment group. Unfortunately, Moerman's finding has little size. As I demonstrated in the last section, the positive correlation can be explained simply by selection. Such pressures tend to depress outcomes in both the new treatment and placebo groups. Moreover, Moerman does not confine his sample to trials with informed consent. It is unclear what information he thinks subjects had or why they responded as they did. A better test would be to check if an increase in the probability of obtaining the new treatment increased outcomes separately in the new-treatment and placebo-control groups.

16. These are ulcers that arise just before the pyloric tract, in that tract, or in the duodenum. These are located at the start of the small intestine.

assigned to that group. Moreover, I have data on subjects in a group as of the date that they are randomized into the group. I do not have precise information on how the group changes due to attrition. Table G.3 provides summary statistics for the data, which are analyzed at the treatment-group level. Data from groups examining H_2-blockers, prostaglandins, and proton-pump inhibitors are presented separately.[17]

2.4.2 Empirical model

I assume that the specific effects of treatment k on individual i enrolled in trial j are a deterministic, linear function of the vector x_{ij}, which includes a constant, clinical and demographic variables on individual i, and structural features of trial j:

17. There are several things to note about the data and Table G.3. First, there are more groups given the new treatment than given the control. The reason is that each trial typically involves one control arm but multiple new treatment arms. Typically these arms will vary the total daily dosage or the daily frequency of medication.

Second, I measure the probability of active treatment by one minus the probability of randomization into the placebo, antacid or bismith subcitrate group. The reason is that, while individuals may have detectably different beliefs about the efficacy of some active treatment versus no treatment, I am skeptical that individuals have sufficiently refined beliefs about different dosages of a given active treatment that one can identify the responsiveness of outcomes to changes in the probability of randomization into each dosage arm. Thus variation in the number of new-treatment arms across trials generates variation in the share treated in my data. Importantly, I assume a trials with same-class controls has a probability of new treatment equal to one.

Third, one-third, forty percent, and 80 percent of H_2-blocker, prostaglandin, and proton-pump inhibitor trials, respectively, have lower-class drug conventional controls.

Fourth, the antacid-permitted variable is coded from 1 to 5. One indicates that subjects were prohibited from taking antacids, two that subjects were discouraged from taking antacids, three that subjects were permitted to take antacids (or the study did not counsel subjects on antacids), four that antacids were provided, and five that antacids were required.

Fifth, the difference in total dosage and dosage frequency between different classes of anti-ulcer medications has little significance. Because the classes have different chemistries and modalities, their recommended dosages are not comparable. Moreover, the total dosage of placebo, antacid or bismuth subcitrate controls is omitted because it has little meaning given that the control is either inert or subjects are typically permitted to take antacid.

$p_{kij} = \beta'_k x_{ij}$. This is a strong assumption, but because my data are rather course there is little benefit from a more nimble parameterization of p_{kij}. I assume that beliefs regarding specific effects are given by $\pi_{kij} = \gamma'_k x_{ij} + \varepsilon_{kij}$.[18] I assume ε_{kij} is independent of x_{ij} and is i.i.d. mean-zero across individuals, trials, and treatment states, with mean zero and variance σ_ε. This last assumption implies that individual errors in estimating specific treatment response do not depend on the treatment.

Because trials often take multiple measurements on each individual, I interpreting a treatment's effect as a hazard rate.[19] Assuming it is constant over time, (2.1) and my parameterization of $\left(p_{kij}, \pi_{ij}\right)$ imply $-\ln S_{ijk}(t)/t = \theta_{(x)k} x_{ij} + \theta_{(xd)k} d_j x_{ij} + \eta_{ijk}$, where $S_{ijk}(t)$ gives the probability of still having an unhealed ulcer on date t, $\theta_{(x)k} = (1 - a_k)\beta_k + a_k\gamma_0$, $\theta_{(xd)k} = a_k(\gamma_1 - \gamma_0)$, and $\eta_{ijk} = a_k d_j \varepsilon_{1ij} + a_k(1-d_j)\varepsilon_{0ij}$. Summing over individuals and dividing by n_{jk}, the number of subjects enrolled in treatment arm k of trial j, yields the regression equation $-\overline{\ln S_{jk}(t)}/t = \theta_{(x)k}\bar{x}_{jk} + \theta_{(xd)k} d_j \bar{x}_{jk} + \bar{\eta}_{jk}$. I approximate $\overline{\ln S_{jk}(t)} = \sum_i \ln S_{ijk}(t)/n_{jk}$ with $\ln(\bar{S}_{jk}(t))$, a first-order Taylor approximation around $\bar{S}_{jk}(t)$, the average probability of remaining ill at t. Approximating $\bar{S}_{jk}(t)$ is difficult because I do not have data on subjects who attrite out of the trials. Therefore, I calculate $\bar{S}_{jk}(t)$ under three assumptions: individuals who attrite out heal at the same rate as those who remain, individuals who attrite out do not heal, and these individuals all heal.

---

18. If $\gamma_k = \beta_k$, then my assumption implies rational expectations. The assumption of linearity limits the scope of cognitive errors that may plague individual projections, but should not otherwise be controversial. The error term measures mistakes in prediction by an individual.

19. Survival analysis is better able to employ information on the timing of treatment response than, say, a simple qualitative dependent-variable framework. Viewing (2.1) as a hazard function does not affect selection into the trial so long as the hazard function is time-invariant.

The regression equation I ultimately estimate is

$$-\frac{1}{t}\ln\left(\bar{S}_{jkt}\right) = \theta_{(x)k}\bar{x}_j + \theta_{(xd)k}d_j\bar{x}_j + \omega_{jk}, \tag{2.5}$$

for $k = 0, 1$, where $\bar{S}_{jkt}$ is approximated in one of the three methods discussed in the last paragraph, $\bar{x}_j = \left(1/n_{jk}\right)\sum_i x_{ij}$, $\omega_{jk} = \bar{\eta}_{jk} + u_{jkt} + v_{jkt}$, and u_{jkt} is the error from approximating $\sum_i S_{ijk}(t)$ with some function of y_{jkt}. The Lagrange remainder from approximating $\sum_i \ln S_{ijk}(t)$ with $n_{jk}\ln\sum_i S_{ijk}(t)$ is absorbed into the coefficient on the constant.[20] For simplicity I left this out of the definition of $\theta_{(x)k}$. The error term v_{jkt} captures the variation in the remainder across arms and trials.[21]

Under the conditions set forth in Section 2.3, a test for placebo effects is whether there is a positive relationship between d_j and survival. This is complicated by self-selection, which may imply that $E\left(\varepsilon_{kij}d_j|s = BT\right) \neq 0$. I address this problem in three different ways. The first method, which I label the "no-selection-subsample" approach, involves estimating (2.5) on a subsample of H_2-blockers for which it is reasonable to believe there will be no self-selection pressure. I hypothesize that this includes all H_2-blocker trials before 1987, when prostaglandins are first introduced. During this period, there were no conventional alternatives that promised

20. This approach bears some resemblence to Amemiya and Nold's [21] adaptation of the logit model to grouped data.

21. Because the hazard rate is $\Pr\{y_{ki} = \bar{y}|s\} \in [0, 1]$, my parameterization requires estimation of a linear probability model. While that model has been criticized for, e.g., potentionally generating predicted values outside the 0-1 range, I do not believe the model is wholly inappropriate for my application. As a theoretical matter, the dependent variable in (2.5) can range from $(0, \infty)$. Moreover, because I aggregate across individuals in an arm, the error term is more likely to resemble a normal distribution [22]. Finally, I modified and estimated (2.5) as a generalized linear model with the dependent variable obeying the binomial distribution and a logistic link function. However, nearly every specification performs worse on Pregibon's [23] link test than the linear probability model.

ulcer healing.[22] Because there is no obvious selection-free subsample for the other class of drugs, I cannot test for placebo effects employing this first approach to selection. The second method, which I call the "x-captures-selection" approach, is to employ the entire sample of trials for each class of anti-ulcer drug but assume that selection pressures are fully captured by observable variables. With the first two approaches, I test for placebo effects by counting the number of treatment arms for which the predicted $\hat{\theta}_{(xd)k}$ is significantly greater than zero. A third method, which I label the "additional-d" approach, is to assume $E\left(\varepsilon_{kij} d_j | s = BT\right) = 0$ but partition $\bar{x}_j = (\bar{x}_j^o, \bar{x}_j^u)$, where $\bar{x}_j^o$ is observable but $\bar{x}_j^u$ may not be. I choose $\bar{x}_j^o$ such that, as a theoretical matter, it ought to be $\theta^o_{(xd)k} d_j \bar{x}_j^1 > 0$. Given that selection pressures depend on d_j, I assume $\theta^u_{(xd)k} d_j \bar{x}_j^u = \phi_{(xd)k} d_j^2 + e_{jk}$, where e_{jk} is independent of d_j and $\bar{x}_j^o$. If selection is a problem but there are no placebo effects, $\theta^u_{(xd)k}$ and thus $\phi_{(xd)k}$ should be zero. Thus, if the estimate of $\phi_{(xd)k}$ is significantly different from zero, then there must exist placebo effects or selection is not a problem.

2.4.3 Results

Table G.4 presents results employing the x-captures-selection method of testing for placebo effects in H_2-blocker trials.[23] The first four columns present coefficient

22. Antacids are a poor substitute because they cannot heal an ulcer. Moreover, most trials permit subjects to consume antacids as well as assigned treatment.

23. Estimation was by feasible GLS. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. My regression model suggests that the variance of error terms depends on the share randomized into each arm. However, I measure only one randomization share per trial, namely the share not given a lower class or non-healing control. Therefore I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. I only report estimates where the dependent variable is calculated assuming subjects who attrite out

estimates for treatment arms and different specifications of $\bar{x}_j$; the last four do the same for control arms. Below the coefficient estimates for each specification are F-tests of the joint significance of certain subsets of regressors: those on d_j, on trial variables interacted with d_j and on subject variables interacted with d_j. Their purpose is to test whether placebo effects operate through any of these subsets of variables. Below the F-tests are counts of the number of measurements that manifest placebo effects at different levels of confidence. I calculate these by computing a prediction for $\theta_{(xd)k}\bar{x}_j$ for each observation and counting those that are significantly greater than zero employing a one-sided test at specified levels of confidence. All arms in the first specification manifest placebo effects at the 95% confidence level. Around three-quarters of measurements on treatment arms in other specifications manifest placebo effects at the 90 percent confidence level; over one-half do so at the 95 percent level. In contrast, measurements on control arms do not manifest evidence of placebo effects. (This does not mean that placebo effects do not affect these arms, only that I cannot verify their influence. Placebo effects may exist but be masked by selection effects.) In specifications of that manifest evidence of placebo effects in treatment arms, the results of the F-tests do not permit one to conclude that placebo effects fail to operate through any natural subset of variables.

Table G.5 presents further results for treatment arms, but for all three classes of

---

heal at the same rate as those who are evaluated. Results from regressions which assume that those who attrite out either all heal or all do not heal are not materially different.

I estimate four specifications of $\bar{x}_j$. Specification (1) includes a constant and d_j; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with d_j; (3) adds subject-level variables (sex, smoker, and age) and their interactions with d_j; and (4) removes the trial-level variables and interactions from (3). When I employ the instrument approach to test for placebo effects, I add d_j^2 to each specification. I test each specification with Pregibon's link test[23]. Those that fail are marked with a dagger (†). I also checked the residuals for patterns but did not find any serious problems.

anti-ulcer drug, by specification and approach to selection. Table G.6 does the same for control arms of trials. The tables present only the results of tests for placebo effects; coefficient estimates for the entire model are omitted. Arms treated with H_2-blockers manifest strong evidence of placebo effects. With the no-selection-subsample approach approximately one-third of arms manifest evidence of placebo effects at the 95 percent confidence level. The one specification where this fails to hold does not pass the link test. The additional-d approach supports an inference of placebo effects in the first, third and fourth specifications, which return a coefficient on d_j^2 that is significantly different from zero. Arms treated with proton-pump inhibitors manifest moderate evidence of placebo effects. Between one-quarter and one-third of measurements manifest evidence of placebo effects with the x-captures-selection approach. However, only the second specification suggests the existence of placebo effects under the additional-d approach. Arms treated with prostaglandins do not manifest much evidence of placebo effects.

The control arms of H_2-blocker trials manifest little evidence of placebo effects. They appear in the fourth specification under the additional-d approach, but under no other specifications or approaches. The control arms of proton-pump inhibitors manifest moderate evidence of placebo effects. Under the x-captures selection approach, the third and fourth specifications suggest one-third of measurements are affected by placebo effects. Under the additional-d approach, the first two specification also manifest evidence of placebo effects. The control arms of prostaglandin trials also manifest moderate evidence of placebo effects. Under the x-captures-selection approach, the first and fourth specifications suggest the existence of placebo effects in all and one-third of measurements, respectively. Under the additional-d approach the first specification supports the existence of placebo effects. However, the second and

third specifications of the x-captures-selection approach, which cannot be rejected under the link test, offer minimal evidence of placebo effects.[24]

Importantly, in twenty trials—all of H_2-blockers—in my sample that fail to confirm in their published results a requirement of informed consent for enrollment, there is significantly diminished evidence of placebo effects except in the first specification under the instrument approach.[25] (These results are omitted.) This does not imply that placebo effects have no influence, but rather that I cannot confirm their influence via a relationship between share treated and outcomes. This is consistent with my assumption that individuals learn the probability of receiving the new treatment via informed consent. Without this disclosure, individual's beliefs about the probability of treatment are unrelated to the actual share treated. Therefore, outcomes, even if they depend on beliefs, should be invariant to the actual share treated.

2.5 Conclusion

This chapter provides evidence of placebo effects in blinded, parallel-arm RCTs of H_2-blockers and proton-pump inhibitors. The findings are summarized in Table G.7. The results are notable because they are derived from a model with clear assumptions and falsifiable predictions. These features make interpretation of findings easier and lend credibility to the conclusion that placebo effects affect ulcer healing.

---

24. I also estimate the proportional hazard model

$$\ln\left(-\ln\left(\bar{S}_{jkt}\right)\right) = \theta_{(t)k}t + \theta_{(x)k}\bar{x}_j + \theta_{(xd)k}d_j\bar{x}_j + \omega_{jk}$$

and obtain similar results. The main difference is that there is somewhat stronger evidence of placebo effects. The coefficient on time is positive and significant, suggesting the risk of healing rises over time. I do not report the results from the proportional hazard model because far fewer specifications pass Pregibon's link test.

25. Subsamples of trials that fail to confirm informed consent or that were conducted outside the U.S., Canada, Western Europe, Australia or New Zealand yield similar results.

However, a greater deal of further investigation is required before placebo effects can be labelled a serious medical phenomenon. These effects may be ailment, drug and context specific. Future research ought to proceed in two directions. One is theoretical. This chapter assumes health outcomes follow a simple linear model. Beliefs may affect outcomes, but individuals do not take placebo effects into account when formulating their beliefs. A more general test for placebo effects would consider the possibility of more general models of health outcomes and the possibility that individuals are aware of placebo effects. In appendix B I examine whether the test for placebo effects proposed in this chapter is robust, e.g., to a multiplicative model of outcomes or to individual knowledge of placebo effects. But this analysis is just a start. This chapter also examines only one type of trial. Many other designs are employed in the medical literature. Appendix C explores tests for placebo effects in unblinded and cross-over trials. It would also be useful to have models for non-randomized trials or trials with monetary incentives. A second avenue for research is application to other medical contexts, such as trials for other ulcer drugs, especially antibiotics, and trials of drugs for other ailments, such as hypercholesterolemia.

If placebo effects are found to have an important influence on outcomes, two questions will naturally follow. First, why do beliefs affect outcomes? Some economists may suspect that placebo effects are product of a rational allocation of the body's resources by the mind. The body has to allocate, e.g., immunoglobulins, white blood cells, macrophages and phagocytes, and fibroblasts to address multiple concurrent infections. In the absence of assistance from pharmacological agents, the mind "chooses" some efficient assignment of these resources to ailments. But if the mind learns that assistance is forthcoming, it may modify its allocation of the body's resources. This adjustment may produce outcomes that are called placebo effects

because, on occasion, the body does not actually receive assistance but outcomes change because of shifts in the body's resource allocation. If drugs and bodily resources are compliments, the result may be an improvement in health triggered by beliefs. While this model is appealing because its similarity to, say, the standard model of resource allocation within firms, it draws skepticism from many doctors with whom I spoken. They typically argue that resource allocation is involuntary, not controlled by the mind. However, that still leaves the question: why do beliefs affect outcomes? Finding an answer requires medical testing of the investment model as well as the development of alternative theories.

A second question triggered by evidence of placebo effects is, can they be used for therapeutic purposes? For example, can a doctor cure a patient by suggesting that a drug is more effective than it really is or by falsely telling the patient that she is receiving treatment? If so, and if fooling patients is less costly than producing drugs, then placebo effects may be able to reduce the costs of health care. If health outcomes are a function of the specific effects of treatment, if the beliefs that affect outcomes are those concerning treatment strategy, and if individuals have rational expectations, then there is reason to be skeptical. If a doctor exaggerates the efficacy of a drug or misleads a patient into believing she has been treated, the patient will observe an outcome that is inferior to what she expected. She will revise her beliefs downwards. This will reduce the influence of the false instruction on beliefs and thus on outcomes. Indeed, it may ultimately render future attempts by the doctor to manipulate beliefs ineffective. If, however, individuals can be fooled, then the placebo may be a useful therapy. For example, if a patient does not have strong priors about the efficacy of a treatment and cannot correct mistaken beliefs by self-sampling or learning from others' experiences, then a doctor may be able to manipulate the

patient's outcome with suggestion. If faith—as opposed to beliefs about the specific effects of treatment—can influence outcomes, then it may be rational for individuals to believe that a given treatment will completely cure them. In fact, it would be rational to believe this about every treatment. Of course, we rarely observe individuals with such beliefs, a fact which suggests that only beliefs tied to reality—say the specific effects of treatment—affect outcomes.

CHAPTER 3

THE EXTERNAL VALIDITY OF CLINICAL TRIALS IN THE PRESENCE OF PLACEBO EFFECTS

This chapter examines the impact of placebo effects on the external validity of estimates of treatment response from data generated by clinical trials and proposes an experiment design for trials that can address complications. Employing the model of RPCTs and health outcomes from chapter two, I offer two insights concerning the impact of placebo effects on the external validity of clinical trials. First, the typical parameter of interest in trials—the physiological effect of treatment—may not always be appropriate. When there exist placebo effects, patients with particularly high expectations will respond better to treatment than those with low expectations. While continuous self-sampling by patient may mean expectations become more accurate over time, in the short-term the appropriate parameter of interest may need to take should take into account subjects' perhaps-erroneous expectations.

Second, the standard method of calculating treatment effects—taking the difference of average outcomes in the treatment and placebo-control groups—yields biased estimates of the effects of treatment outside the trial context. The bias can be decomposed into two major parts. One is due to blinding. When trial subjects are blinded, they do not know whether they are in the treatment or control group. Those in the active-treatment group doubt that they are being treated, so they typically experience a negative expectancy effect which worsens their observed health outcomes.

Conversely, those in the placebo-control group experience a positive expectancy effect because they believe that they might be treated even though they are not. This improves their observed health outcomes. The overall result is to contract estimates of treatment effects, biasing these estimate towards zero.

The other component of bias is the result of self-selection. The population of interest includes only individuals who believe that the new treatment is superior to alternatives. The remainder would not choose new treatment even if it were available outside the context of a trial. However, clinical trials do not attract this entire population. They only attract those individuals who believe that the new treatment is so much better than alternatives that, even with the risk of obtaining placebo, they believe enrollment in a trial to be a superior strategy. This composition effect may introduce bias, although the sign of this bias is not known a priori.

The final contribution of this chapter is an alternative design for RCTs that can generate unbiased estimates of treatment response. The experiment involves randomizing subjects not just over treatment and control groups, but over different probabilities of being randomized into treatment.[1] Randomization over the probability of being treated is valuable because it manipulates subjects' beliefs about the value of the trial in a measurable and objective manner. The data generated by this design can identify the short-run parameter of interest without bias from placebo effects.

This chapter relates to the medical literature on placebo effects and the biostatistics literature on experimental design. A few authors [24, 25, 26] discuss whether the

1. For example, in the first stage of the trial, subjects are randomized evenly into two arms. Those randomized into arms 1 and 2 are informed that they will be randomized into treatment with a probability of 3/4 and 1/4, respectively. At stage 2 subjects are randomized into treatment or placebo-control at the rate they have been told, but are not informed of their ultimate treatment state.

physiological effects and placebo effects of treatment are additive. They note that, if placebo effects are not additive, the fact that there is no difference between outcomes in active-treatment and placebo-control groups does not demonstrate that the active treatment is ineffective. Some authors [14, 25, 27] hint at the bias from blinding. They state, without explanation, that expectation effects in placebo-control groups of blinded trials make it difficult to demonstrate the efficacy of active treatment. Finally, there is some [5, 28] exploration of experimental designs to identify the magnitude of placebo effects or to identify treatment effects without bias due to placebo effects. These designs either compare placebo-control groups to no-treatment groups or screen out trial subjects with characteristics that suggest strong placebo-response. The designs have serious flaws—attrition bias in the former and poor observable proxies for strong placebo response. More importantly, the design recommendations are, from a statistics perspective, very informal.

There is no economics literature on placebo effects, perhaps because it is rarely possible to conduct a blinded experiment to address empirical questions traditionally of interest to economists. The experimental design proposed in this chapter is similar, however, to that proposed by Philipson [29] to address the problem of external effects, i.e., externalities that treatment (such as vaccination) may have on subjects in the control group.[2] My design, like Philipson's, is a variant of the standard two-factor design familiar to the statistics literature. The feature that distinguishes the experiment in this chapter is that the two factors are not independent: whether a

2. When the share treated in a trial affects the size of this externality and differs from the share that will be treated once a program is fully implemented, then the trial will produce biased estimates of the effect of treatment under full implementation. Philipson proposes a two-stage randomized design to generate variation in the share treated that can used to predict how treatment effects change as the share treated is modified. I modify Philipson's design to take into account selection and attrition.

subjects in a given arm is treated is a function of the share treated in that arm. This complicates optimal selection of design parameters such as treatment allocation and sample size.

Section 3.1 discusses the appropriate parameter of interest when placebo effects are present. Section 3.2 examines how placebo effects bias the standard estimator of treatment response. Section 3.3 present an experimental design for clinical trials that permits unbiased estimation of the appropriate parameter of interest. (Appendix D examines the $\mathcal{D}$-optimal design of this experiment.) The final section concludes.

3.1 Parameter of Interest

The parameter of interest in clinical trials is ostensibly $E(p_{1i}) - E(p_{0i})$. That trials attract only those individuals who are optimistic about the new treatment is typically overlooked. When the issue of self-selection is raised, investigators either minimize its effect or respond that the parameter of interest is actually the (specific) effect of treatment on the treated. However, the population of interest includes everyone who would choose a strategy of new treatment ($s = 1$) if it were available outside the context of a trial. This population includes all individuals who believe that new treatment is superior to alternatives, i.e., $\pi_{1i} \geq \max\{\pi_{0i}, \pi_{2i}\}$.[3] Thus the proper parameter of interest is

$$E(p_{1i}|s = 1) - E(p_{0i}|s = 1). \tag{3.1}$$

3. The subset of entire population that would enroll in a blinded trial is a subset of the interested population. To see this, note that the condition that the condition $V_i^{BT} \geq \max\{V_i^0, V_i^2\}$ implies $\pi_i^{BT} = d\pi_{1i} + (1-d)\pi_{0i} \geq \max\{\pi_{0i}, \pi_{2i}\}$. The constraint that $\pi_i^{BT} \geq \pi_{1i}$ implies that $\pi_{1i} \geq \pi_{0i}$. The constraint that $\pi_i^{BT} \geq \pi_{2i}$, combined with the result that $\pi_{1i} \geq \pi_{0i}$, implies that $\pi_{1i} \geq \pi_{2i}$.

An initial difficulty with this parameter is that beliefs change over time. Indeed, the population's beliefs may change in response to the investigator's study. This updating may alter the composition of the population of interest such that the investigator's findings about the effect of treatment upon the ex ante population of interest may not characterize the effect of treatment upon the ex post population of interest. This problem, which is the medical trial analogue of the Heisenberg uncertainty principle in physics, exists even if placebo effects do not. In this chapter I assume away this problem by stipulating that $\mathbf{g}_\pi$ is fixed over time. This assumption may be valid for obscure studies, but not for those whose findings are widely publicized.

A second problem with the standard parameter of interest is the existence of placebo effects. In the blinded trial setting, because subjects do not know their treatment state, there will be a gap between individuals' expectations and the specific effects of their actual treatment state: $\pi_{ki}^{BT} = d\pi_{1i} + (1-d)\,\pi_{0i}$. This gap will trigger placebo effects. Outside the context of a trial, however, individuals know their treatment state. Nevertheless, they continue to experience placebo effects unless they estimate the specific effects of each treatment without any error. Under the model of outcomes in chapter two, $\Pr\{y_{ki} = \bar{y} | s = k\} = p_{ki} + a_k(\pi_{ki} - p_{ki}) \neq p_{ki}$.[4] This suggests that the appropriate parameter of interest ought to be

$$E(y_{1i}|s=1) - E(y_{0i}|s=1). \tag{3.2}$$

4. This insight is not rendered academic if individuals are assumed to have rational expections so that $E(\boldsymbol{\pi}_i) = E(\mathbf{p}_i)$. Individuals who self-select into trials are a subset of the interested population, which in turn believes that new treatment is superior to alternatives, whether that calculation is accurate or due to manifestiation, e.g., of positive error in prediction of the efficacy of the new treatment. Thus trial subjects will be a subset of the population who on average overestimate the efficacy of the new treatment and underestimate the efficacy of alternatives. So while the population at large can be characterized as having rational expectations, the subpopulation of trial subjects cannot be.

The one argument that can be made on behalf of the parameter of interest in (3.1) is that, if individuals repeatedly self-sample their response to treatment outside the trial context and use this information to gradually update their beliefs, their beliefs may eventually converge to the specific effects of treatment. For example, suppose individual i starts at time $t = 0$ with arbitrary beliefs, $\pi_{ki0} = \hat{\pi}_{k0}$, is aware of the existence of placebo effects, and updates her beliefs according to $\pi_{kit} = y_{kit} + \alpha_k \left(\pi_{ki(t-1)} - y_{kit}\right)$, for all k, where α_k is her beliefs about the influence of beliefs on outcomes. (I assume, for simplicity, that $\alpha_{k'}$ is independent of $\pi_{k'i't'}$ and $y_{k'i't'}$ for all k'.) At time t, her beliefs will be $\pi_{kit} = [\sum_{\xi=0}^{t-1} \alpha_k^{\xi}] (1 - \alpha_k) y_{kit} + \alpha_k^t \hat{\pi}_{k0}$. So long as $\alpha_k \in (0, 1)$, $\lim_{t\to\infty} E(\pi_{kit}) = p_{ki}$, where expectations are taken over the distribution of outcomes for individual i. Thus, in the long-run, the proper parameter of interest may properly be (3.1).

Unfortunately, in the next section I demonstrate that it may be impossible to estimate the long-run parameter of interest with the standard estimator. Even the experimental trial design I propose in Section 3.3 can only estimate short-run parameter of interest.

3.2 Bias Due to Placebo Effects

The standard estimator used to calculate treatment effects in clinical trials is $(\sum_{i\in I_1} y_{1i}/n_1) - (\sum_{i\in I_0} y_{0i}/n_0)$, where I_k is the set of subjects randomized into the group k and n_k is the number of subjects in I_k, for $k = 0, 1$. This estimator measures $E(y_{1i}|s = BT, \delta_{1i} = d) - E(y_{0i}|s = BT, \delta_{0i} = 1 - d)$, i.e., the difference in outcomes between individuals who choose to enroll in a blinded clinical trial and are randomized into new treatment and into no treatment.

The standard estimator produces biased estimates of the long-run parameter of interest in clinical trials. This bias can be broken down into three parts:

$$E\left(y_{1i}|s=BT,\delta_{1i}=d\right)-E\left(y_{0i}|s=BT,\delta_{0i}=1-d\right) \tag{3.3}$$
$$-E\left(y_{1i}|\underline{s=1},\delta_{1i}=d\right)-E\left(y_{0i}|\underline{s=1},\delta_{0i}=1-d\right)$$

$$E\left(y_{1i}|s=1,\delta_{1i}=d\right)-E\left(y_{0i}|s=1,\delta_{0i}=1-d\right) \tag{3.4}$$
$$-E\left(y_{1i}|s=1,\underline{\delta_{1i}=1}\right)-E\left(y_{0i}|s=1,\underline{\delta_{0i}=1}\right)$$

$$\left[E\left(y_{1i}|s=1,\delta_{1i}=1\right)-E\left(y_{0i}|s=1,\delta_{0i}=1\right)\right] \tag{3.5}$$
$$-\left[E\left(\underline{p_{1i}}|s=1,\delta_{1i}=1\right)-E\left(\underline{p_{0i}}|s=1,\delta_{0i}=1\right)\right]$$

(I have underlined the important change for each component.) The first component of the bias is due to self-selection. Subjects who enroll in a clinical trial may not respond to a new treatment in the same manner as randomly selected individuals from the population of interest. The second component is due to blinding. Subjects in blinded clinical trial do not know their treatment state. They only know the odds of being in the new-treatment state. Therefore, their expectations differ from individuals outside the trial setting who are fully aware of their treatment state. If there exist placebo effects, this difference in expectations alters observed outcomes. The third component of bias is due to the existence of placebo effects, that is, the mere fact that expectations affect outcomes. If the investigator is interested in the specific effects of treatment, even an accurate measure of the short-run parameter of interest would be biased because expectations affect outcomes outside the trial

context.

While the long-run parameter of interest focuses on the specific effect of new treatment, the standard estimator picks up placebo effects. The short-run parameter admits these effects and, therefore, does not suffer placebo-effect bias. However, it does suffer blinding and selection bias. Finally, even if the parameter of interest is the effect of treatment upon the treated in the long run, the standard estimator still suffers bias due to blinding. In the next section I examine in reverse order each of these biases in the context of the model of health outcomes from chapter two.

3.2.1 Expectations bias

Given my simple additive model of health outcomes and my assumption that a_k is independent of π_{ki} for all k, i, the bias in the standard estimator due solely to the existence of placebo effects, i.e., $a_k > 0$ for all k, can be written

$$a_1 (\pi_1 - p_1) - a_0 (\pi_0 - p_0), \tag{3.6}$$

where, $\pi_k = E(\pi_{ki}|s = 1)$ and $p_k = E(p_{ki}|s = 1)$. The bias has two parts. Although the long-run parameter of interest includes only the specific effect of treatment, the standard estimator measures outcomes including placebo effects. The resulting bias is the difference between placebo effects in the new-treatment group (the first term in (3.6)) and in no-treatment group (the second term in (3.6)) assuming subjects know their treatment state. If the new treatment group experiences greater (smaller) placebo effects, the sign of the bias is positive (negative).

3.2.2 Blinding bias

The bias due to blinding is

$$-a_1 (1-d)(\pi_1 - \pi_0) - a_0 d (\pi_1 - \pi_0) < 0. \tag{3.7}$$

This bias has two parts. The first is a negative expectancy effect among subjects in the treatment group. If a subject could obtain the new treatment outside the context of a trial, she would experience a placebo effect induced by the discrepancy between her expectations about the efficacy of the *new treatment* and the specific effect of that treatment: $a_1 (\pi_{1i} - p_{1i})$. In the context of a blinded trial, the placebo effect is driven by the discrepancy between the efficacy of the *trial* and the specific effect of the subject's actual treatment state because the subject does not know her exact treatment state: $a_1 \left(\pi_i^{BT} - p_{1i}\right)$. Because the subject is drawn from the population of interest, which believes the new treatment to be superior to alternatives, the expectation-induced health effect in the context of a blinded trial is less than the expectation-induced health effect outside the context of a trial, i.e., $a_1 \pi_i^{BT} = a_1 (d\pi_{1i} + (1-d)\pi_{0i}) < a_1 \pi_{1i}$. The difference is the first term in (3.7). The intuition is that a subject randomized into the new-treatment state fears that she may be in the no-treatment state even though she is actually in the new treatment state, so she suffers a negative placebo effect.

For similar reasons, subjects in the control group will likely experience a positive expectancy effect equal to the second term in (3.7). Because of blinding, a subject randomized into the no-treatment state believes there is a chance that she is in the new-treatment group, which she believes is superior, even though she is actually in the no-treatment group. This inflates the subject's health outcomes. The standard

estimator takes average outcomes in the new-treatment group, which manifest depressed outcomes due to a negative placebo effect, and subtracts average outcomes in the placebo-control group, which manifest inflated outcomes due to a positive placebo effect. The net result is negative bias.

Note that blinding bias does not disappear even if the influence of beliefs on outcomes are identical in each treatment state, i.e., $a_1 = a_0$.[5] Blinding still depresses outcomes to the tune of $a(\pi_1 - \pi_0)$. The reason is that the standard estimator purges the entire effect of beliefs on outcomes. It differences outcomes in the treatment and control groups, but beliefs have the same effect on outcomes in each group because blinding assures that subjects in each group have identical beliefs. The difficulty is that, if the short-run parameter is of interest, the investigator seeks the effect of treatment on outcomes taking beliefs into account. Even if the investigator is interested in the specific effects of treatment, that $a_1 = a_0$ does not fully eliminate bias. Expections bias—$a[(\pi_1 - p_1) - (\pi_0 - p_0)]$—may persist if there the difference

5. In order to test whether $a_0 - a_1$ in the data set results from ulcer trials describe in Section 2.4, I estimated the following regression model

$$\begin{aligned} -\frac{1}{t}\left[\ln\left(\bar{S}_{j1t}\right) - \ln\left(\bar{S}_{j0t}\right)\right] &= \theta_{(x)1}\bar{x}_{1j} + \theta_{(xd)1}d_j\bar{x}_{1j} + \omega_{j1} \\ &\quad -\theta_{(x)0}\bar{x}_{0j} - \theta_{(xd)0}d_j\bar{x}_{0j} - \omega_{j0}. \end{aligned}$$

(My estimation method is the same as in Section 2.4. However, there is one observation for each paired treatment-control arm. Thus, there are the same number of observations as in Table G.5.) I found that the overall marginal effect of the share treated is negative. This finding is inconclusive. It is consistent with the view that placebo effects are identical ion each group or that placebo effects in the treatment group are larger but self-selection pressures overwhelm placebo effects once group outcomes are differenced. (The latter possibility is supported by the model of treatment choice in Chapter 4.) The finding could be because placebo effects are larger in control arms. However, this is unlikely because treatment arms generally manifest stronger evidence of placebo effects than control arms. Compare Table G.5 and G.6. (The only case where the regression model above is conclusive is where the estimated relationship between share treated and outcomes is positive. That suggests that $a_1 > a_0$.

between beliefs and specific effects differs across the different treatment states. Even if it were positive and offset blinding, the standard estimator would still be biased downwards by $a\,(p_1 - p_0)$ because placebo effects reduce the influence of specific effects on observed outcomes. That being said, if the influence of beliefs were identical in all treatment states, it would be possible to compare the specific effects of the test treatment and an alternative by taking the ratio of the standard estimator of treatment response in a RPCT of the test treatment to the standard estimator of response in a RPCT of the alternative.

Figure H.1 illustrates the bias in the long-run parameter of interest due solely to placebo effects and the bias in both parameters of interest due to blinding. The x-axis gives the share of trial subjects randomized into the new-treatment group, and the y-axis gives the probability of recovery. The top and bottom dash-dotted lines indicate the average specific effect of the new treatment and of no treatment, respectively, among the population of interest. The top solid line presents average outcomes given placebo effects among individuals from the population of interest not enrolled in a trial but able somehow knowingly to secure the new treatment; the bottom solid line presents average outcomes among individuals in the population of interest who knowingly forego any treatment.

The bias due to placebo effects is the difference of the distance between the two solid lines and the distance between the two dash-dotted lines. I have drawn Figure H.1 to reflect a positive placebo-effect bias. One would expect a positive bias if, e.g., individuals had rational expectations and their errors in estimating the specific-effect of a treatment were independent of that treatment and their estimation of the specific-effects of other treatments. While the population of interest consists of individuals who believe the new treatment is superior to alternatives, these beliefs could be due

to overestimation of the value of new treatment or underestimation of the value of no treatment. While these errors may not effect the specific-effect of treatment among the population of interest, it does affect outcomes given placebo effects.[6] In that case, overestimation of the relative value of new treatment elevates outcomes in the new treatment group and depresses outcomes in the no-treatment group.

The top and bottom dashed lines represents average outcome among individuals selected at random from the population of interest, enrolled in a blinded trial, and then randomized into the new-treatment and no-treatment states, respectively. Each of these lines is upward sloping because an increase in the share treated in a blinded trial raises expectations about the trial among members of the population of interest and thus outcomes in each treatment group. Mathematically, $\partial E\left(y_{ki}|\mathbf{p}_i, \boldsymbol{\pi}_i\right)/\partial d = a_k\left(\pi_1 - \pi_0\right) > 0$, for $k = 0, 1$. At $d = 0$, the blinded trial is equivalent to unblinded administration of no treatment. Therefore, outcomes of the population of interest when forced to enroll in a blinded trial with $d = 0$ and randomized into the no-treatment group converge with outcomes when the population of interest is administered no treatment without blinding. Similar arguments explain why the top solid and dashed lines converge at $d = 1$. The bias due to blinding is the difference of the distance between the two dashed lines and the distance between the two solid lines. If the influence of beliefs on outcomes varies across states, the bias due to blinding may vary with the share given new treatment. In particular, if $a_1 > (<) a_0$, then the bias will rise (fall) with d.

---

6. Formally, if one assume $\pi_{ki} = p_{ki} + u_{ki}$ where $E\left(u_{ki}\right) = 0$, $u_{ki} \perp u_{k'i'}$ for all $k' \neq k$ and $i' \neq i$, $u_{ki} \perp p_{k'i}$ for all k', then $E\left(p_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) \geq E\left(\pi_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right)$ and $E\left(p_{1i}|\pi_{0i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) \leq E\left(\pi_{0i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right)$ because $E(u_{1i}|p_{1i} + u_{1i} \geq \max\{\pi_{0i}, \pi_{2i}\}) \geq 0$ and $E(u_{0i}|\pi_{1i} \geq \max\left\{p_{0i} + u_{0i}, \pi_{2i}\right\}) \leq 0$, respectively.

3.2.3 Selection bias

The bias due to self-selection is

$$\begin{aligned}&(1-a_1)\left[p_1^*-p_1\right]-(1-a_0)\left[p_0^*-p_0\right] \\ &+(a_1-a_0)\left[d\left(\pi_1^*-\pi_1\right)+(1-d)\left(\pi_0^*-\pi_0\right)\right],\end{aligned} \tag{3.8}$$

where $\pi_k^* = E\left(\pi_{ki}|s=BT\right)$ and $p_k^* = E\left(p_{ki}|s=BT\right)$, for $k=0,1$.

The first line of (3.8) corresponds to classic self-selection bias. The members of the population of interest that self select into a blinded trial may differ systematically from the population of interest at large. The latter includes anyone who believes that the new treatment is even marginally better than alternatives, whereas voluntary enrollees believe that new treatment is so much better than alternatives that, even with the risk of obtaining no treatment, enrollment is superior to alternatives. If individuals are even mildly adept at predicting their own specific response to treatment, the correlation between beliefs and the specific effect of a treatment is positive and selection affects outcomes in blinded trials. The precise bias is the difference between impact of selection on the new-treatment and no-treatment groups, with the impact in each group weighted by the importance of specific effects to observed outcomes in the state corresponding to that group. The sign of this bias is ambiguous.

Classic selection bias does not depend on the existence of placebo effects. When placebo effects exist, however, selection has a second, more direct impact on outcomes. Whereas classic selection bias requires some correlation between beliefs and specific effects, selection bias due to placebo effects requires only a difference in the beliefs of enrollees and of the population of interest. With placebo effects, beliefs directly affect outcomes. The precise bias is given in the second line of (3.8). Because of

blinding, the expectations of each group about the trial is identical. As evidenced by their decision to enroll in a trial, enrollees' expectations about the trial are greater than the expectations of the population of interest at large. However, the net bias from selection due to placebo effects depends on whether beliefs matter more in the new-treatment or the no-treatment state. If beliefs matter more (less) in the new treatment state, then selection bias due to placebo effects will be positive (negative). If beliefs affect outcomes identically in each state, the impact of selection due to placebo effects in each group exactly cancel each other out.

Figure H.2 provides a graphical example of the bias due to selection. The top and bottom solid lines capture outcomes given placebo effects among individuals in the population of interest who are administered, without blinding, the new treatment and no treatment, respectively. The top and bottom dashed lines capture average outcomes among individuals from the population of interest who are selected at random to participate in a blinded trial and are randomized into the new and no treatment groups, respectively. I explained the slopes of the dashed lines and the intersection with the solid lines in my discussion of Figure H.1.

The dash-dotted and dash-double-dotted lines present outcomes among the subset of the population of interest who would voluntarily enroll in a blinded trial. The top and bottom dash-dotted lines present outcomes among enrollees who are administered the new and no treatment, respectively, in an unblinded fashion. I have drawn these lines to be downward sloping, although they need not be. My depiction is consistent with my conjecture in chapter two that an increase in the share treated will usually lower the average beliefs about the efficacy of both new treatment and no treatment among enrollees. The intuition is that a medical trial only attracts individuals who believe that the new treatment is so much better than conventional

treatment that, even with the risk of obtaining no treatment at all, enrolling in the trial is a superior strategy to opting for conventional treatment. With an increase in the probability of obtaining the new treatment, a patient who is marginally not optimistic enough about the new treatment or no treatment to have risked randomization into the placebo-control group before may now be willing to take that risk because it is smaller. The result will be lower average beliefs about new treatment and no treatment among enrollees unless individuals' beliefs about new treatment and conventional treatment are highly positively correlated.[7] If increases in d always lower beliefs among enrollees, then the fact that, for any given population, outcomes in the no treatment (new treatment) group of a blinded trial converge to outcomes with unblinded administration of no treatment (new treatment) implies that selection bias must be positive.

The top and bottom dash-double-dotted lines present outcomes among enrollees randomized into the new treatment and no treatment groups, respectively, of a blinded trial. These lines are the only set of potentially observable outcomes in a clinical trial. The exact outcomes actually observed in a particular trial are the points of intersection between the dash-double-dotted lines and vertical dotted line indicating the specific share $\hat{d}$ randomized into new treatment in that trial. The dash-double-dotted lines represent outcomes incorporating both positive selection bias and negative blinding bias. I have drawn the graph such that selection bias offsets blinding bias, but not completely. The result is that the standard estimator (SE), which measures the

7. In that case an increase in the probability of being randomized into the new-treatment group may attract not just marginally pessimistic patients, but also lots of patients who are very optimistic about the new treatment if the latter are also very optimistic about the conventional treatment. Although the latter feel the new treatment is better than conventional treatment, they did not previously enroll because the expected payoff of a trial was less than the value of conventional treatment.

distance between the two dash-double-dotted lines at $\hat{d}$, reports a treatment response that is smaller than the short-run parameter of interest (POI).

3.2.4 Empirical evidence of bias

There are a number of studies which suggest that the biases just described cause the standard estimator to underestimate the short-run parameter of interest. To begin with, in chapter two I demonstrated that outcomes in both the test-treatment and the control groups of ulcer trials rose with the share randomized into test-treatment. This result is consistent with my prediction that outcomes in each treatment arm of a blinded trial rise with the share randomized into the new-treatment group because subjects expectations for the trial have risen.

Further, Schulz et al. [30] examined 250 trials from 33 meta-analyses published by the Pregnancy and Childbirth Group of the Cochrane Database. They found that outcomes in trials with inadequate concealment (21 trials) or unclear concealment (150 trials) of treatment state yielded estimates of treatment effects that were 41 and 30 percent higher, respectively, than trials with adequate concealment. Their findings confirm earlier work by Chalmers et al. [31] and are consistent with my prediction (in Appendix C) that blinded trials bias the standard estimator towards zero.

Finally, Bergmann et. al. [32] conducted a conducted a blinded crossover trial of the effect of naproxen on cancer pain in which they first randomized subjects into two groups, with and without informed consent. They found that the difference in the level of pain reported by the treatments and controls was smaller in the informed group. This study suggests that informing subjects of their probability of receiving active treatment and examining voluntary enrollees depresses observed outcomes. The one caution with this study is, because patients were already admitted in an

oncology ward for treatment of mild or moderate cancer pain, the no informed consent group was akin to the population as a whole, not just the population of interest.

3.3 Experimental Design

The discussion above suggests three different ways to address the problem of bias due to placebo effects. First, assume that placebo effects do not exist, i.e., that $a_k = 0$ for all k. Unfortunately, this approach does not address the problem of classic selection bias. More importantly, there is strong and growing evidence that placebo effects do exist. Second, assume that beliefs have identical influence on outcomes in each treatment state, so the direct effect of selection on outcomes is zero, and constrain trials such that exactly 1/2 of subjects are randomized into new treatment and 1/2 into no treatment, so that $d = (1 - d)$ and the bias due to blinding is zero. These restrictions would permit estimation of the short-run parameter of interest, but with classic selection bias.[8]

In this section, I offer a third alternative: an experiment that is able to produce estimates of the short-run parameter of interest without bias due to blinding or selection and without restrictions on the existence or nature of placebo effects.[9]

8. If one wanted to estimate the long-run parameter of interest and assumed that beliefs had the same influence on outcomes in each state, one could estimate $(1 - a)(p_1^* - p_0^*)$. The estimator would suffer not just classic selection bias, but also attenuation bias that reflected the diminished influence of specific effects of treatment on outcomes in the presence of placebo effects.

9. My design is not the only one that has been proposed to estimate treatment response in the presence of placebo effects. However, these other design have siginificant flaws. The most common design that has been proposed is a balanced-placebo design that randomizes subjects, first, across treatment states and, then, across instructions about treatment state, with one group in each treatment state being told they were given active treatment and the other being told they were given placebo. This design holds promise for controlling placebo effects and blinding bias, but it does not account for selection bias. Moreover, this design cannot be used with any frequency. Otherwise potential subjects will come

3.3.1 Proposed experiment

One reason the standard estimator for treatment response is biased is that important variables are omitted, namely information on enrollees' beliefs. Beliefs about efficacy are difficult to measure in an objective fashion. However, the investigator has a good proxy for enrollees' beliefs δ about their treatment state. If subjects are informed of the share d of enrollees that are randomized into the new-treatment group, subjects will set $\delta = d$. The experiment I propose generates predictable variation in subjects' beliefs about treatment state by randomizing subjects, first, across the share given new treatment, and then across treatment state. While subjects are never told their ultimate treatment state, they are updated on their personal probability of receiving treatment.

A second reason why the standard estimator is biased is that existing clinical trials do not provide the data required to adjust the estimator for bias due to self-selection. Few trials vary their parameters across subjects so as to generate variation in participation incentives. Such variation would permit inferences about the attributes of non-enrollees. Moreover, clinical trials rarely gather data on individuals who

---

to expect that investigators lie about treatment assignment and ignore their instructions. Subsequent attempts to manipulate beliefs about treatment state will fail. Finally, the design is unethical. It requires investigators to lie to subjects about their treatment state. Rosenthal (1985) and Ross and Buckalew (1985) proposed a design where investigators vary the dose of placebo given to individuals in the placebo control group. The idea is that changes in treatment response due to variation in placebo dosage can be used to predict outcomes in the no treatment state outside the context of a trial. There are a number of problems with this approach. First, if there are fixed effects due to placebo consumption, it may not be possible to extrapolate outcomes with no treatment outside a trial setting from variable effects due to placebo consumption in the no-treatment group of a trial. Second, varying the dosage of placebo does not address the effects of blinding on outcomes in the new-treatment group. Varying dosages in the latter group does not address the problem if there are fixed effects to consumption of that treatment. Third, observed variations in dosage of placebo, which is how beliefs are manipulated in this experiment, may trigger attrition that renders outcomes across dosage groups incomparable.

decline to enroll. Such data provide direct information on non-enrollee attributes. Together with data on how enrollee attributes correlate with treatment response, information on attributes of non-enrollees may permit inferences on how non-enrollees would respond to treatment. The proposed experiment provides variation in a trial parameter—the probability of receiving the new treatment—that generates variation in selection pressure. Moreover, by providing subjects with a low cost option to obtain treatment with high probability if not virtual certainty, the proposed experiment offers an opportunity to gather data on non-enrollees from nearly the entire population of interest.

The experiment I propose is very much like the standard parallel-arm design described in Section 2.1, but with one additional stage. First, individuals are recruited. Enrollment is voluntary and subject to informed consent. As part of this disclosure, candidates are informed of the two-step randomization process of the trial, included the probability of randomization into each arm of the trial in stage two and the probability of randomization from each arm of the trial into the new treatment state in stage three. Subjects are also told they will be free to exit after stage two. If they choose to enroll, the investigator gathers demographic and clinical data $\mathbf{z}_i$ on each subject. Let n be the total number of enrollees.

In the second stage, subjects are randomized across m arms of the trial. In particular, n_j subjects are randomized into arm j, for $j = 1, ..., m$. Each arm corresponds to a different probability d_j of being given treatment. Subjects are informed of their odds of being randomized into the new treatment from their arm and are told they are free to leave the trial. For expositional purposes, let $i = 1, ..., r_j$ index subjects in arm j that remain part stage two and $i = r_{j+1}, ..., n_j$ index subjects who exit that arm.

In the third stage, the r_j subjects who remain in arm j are randomized into new-treatment or placebo-control groups according to d_j. Let $D_{ij} = 1$ or 0 indicate whether subject i in arm j is ultimately given the new treatment or not, respectively. Note that $\left\{D_{ij}\right\}$ and $\{d_1, ..., d_m\}$ are linked. It must be that the $\sum_{ij} D_{ij} = \sum_{j=1}^{m} n_j d_j$. Subjects are not informed of their ultimate treatment state. In the final stage of the trial, subjects in the new-treatment group receive the new treatment; subjects in the control group receive a placebo. At the end of the trial, outcome data y_{ij} on subjects who remain are recorded.

Figure H.3 provides a simple example of the experiment I propose. At the recruitment stage, candidates are informed of each detail about the trial. At stage two, enrollees are randomized into three arms of the trial, with equal numbers of subjects in each arm. Subjects in arm j are informed that they will be randomized into treatment with probability d_j, with $d_1 = 3/4$, $d_2 = 1/2$, and $d_3 = 1/4$. Subjects are informed of the arm to which they have been assigned and their corresponding chance of receiving the new treatment. At stage three, subjects are in arm j are randomized into treatment with probability d_j. Subjects are not informed whether are in the new-treatment group or the placebo-control group of an arm.

To recap, the important innovations in the proposed experiment are that subjects are first randomized across the share treated; that they are updated on their revised probability of treatment, although they remain blind to their ultimate treatment state; and that they are informed about these innovations before enrollment. Unblinded variation in the share treated generates variation in beliefs about treatment. This, in turn, generates variation in placebo response as well as variation in selection pressure.[10]

10. Informing recruits that they will be randomized across the share treated also allows the investigator to attract a broader swath of individuals from the population of interest.

3.3.2 Estimation and identification

Assume for the moment that changes in the share given new treatment do not trigger selection effects, i.e., no subjects leave the trial after stage two of the proposed experiment. This may be because there is no viable conventional treatment alternative or a fortunate coincidence of the population distribution of beliefs about treatment.[11] In this case, $r_j = n_j$ for all arms j and the model of health outcomes in the first essay of my thesis can be written $y_{ij} = [\left(1 - a_{1j}\right) p_{1ij} + a_{1j}(d_j \pi_{1ij} + \left(1 - d_j\right) \pi_{0ij})] D_{ij} + [\left(1 - a_{0j}\right) p_{0ij} + a_{0j}(d_j \pi_{1ij} + \left(1 - d_j\right) \pi_{0ij})] \left(1 - D_{ij}\right) + e_{ij}$, for subjects $i = 1, ..., n_j$, in arms $j = 1, ..., m$, where e_{ij} reflects error due to the inability to observe underlying hazard rates. This is a random coefficients model that can be rewritten as

$$y_{ij} = \boldsymbol{\beta} \mathbf{x}_{ij} + w_{ij}, \tag{3.9}$$

where $\mathbf{x}_{ij} = \left(1, D_{ij}, d_j, D_{ij} d_j\right)'$, $\mathbf{v}_{ij} = \boldsymbol{\beta}_{ij} - \boldsymbol{\beta}$, $E\left(\mathbf{v}_{ij}\right) = 0$, $E\left(\mathbf{v}'_{ij} \mathbf{v}_{ij}\right) = \Sigma_v$, and $w_{ij} = e_{ij} + \mathbf{v}_{ij} \mathbf{x}_{ij}$. Note that the assumption that $E\left(\mathbf{v}_{ij}\right) = 0$ is only possible if $\boldsymbol{\beta} = (\beta_0, \beta_D, \beta_d, \beta_{Dd})$ estimates the mean $\boldsymbol{\beta}_{ij}$ among the population of interest. Due to the distribution of y_{ij} and to random coefficients, the error term is a function of $\left(d_j, d_j^2\right)$ and thus heteroskedastic. Therefore, (3.9) should be estimated via FGLS along the lines suggested by, e.g., Swamy [33].

To see this, compare the proposed experiment to an ordinary parallel-arm trial with a constant probability $d_0 = \sum_j n_j d_j$ of being treated. The proposed experiment attracts all individuals i for whom $\bar{d} \pi_{1i} + (1 - \bar{d}) \pi_{0i} \geq \max\{\pi_{0i}, \pi_{2i}\}$, where $\bar{d} = \max\{d_1, ..., d_m\}$, whereas the ordinary trial attracts only those individuals i' for whom $d_0 \pi_{1i'} + (1 - d_0) \pi_{0i'} \geq \max\{\pi_{0i'}, \pi_{2i'}\}$. Since each enrollee, regardless of experiment, believes that new treatment is better than alternatives, the fact that $d_0 \leq \bar{d}$ suggests that each individual who would enroll in an ordinary trial would also enroll in a trial based on the proposed experiment.

11. See chapter 3 for problems with the use of particpation incentives to mitigate selection based on the share randomized into the new-treatment group.

Simple algebra confirms that an unbiased estimator for the mean short-run parameter of interest for the population of interest, $(1-a_1)\,p_1+a_1\pi_1-(1-a_0)\,p_0-a_0\pi_0$, is $\hat{\beta}_{SR}=\hat{\beta}_D+\hat{\beta}_d+\hat{\beta}_{Dd}$. The mean long-run parameter of interest for the population of interest, p_1-p_0, cannot be identified. There are only four equations and at least six unknowns: p_1-p_0, p_0, $a_1p_1-a_0p_0$, $(a_1-a_0)\,\pi_0$, $a_0\,(\pi_1-\pi_0)$, and $a_1\,(\pi_1-\pi_0)$.[12]

Now relax the assumption that there is no selection on the basis of the share administered new treatment in each arm. Assuming the cost of enrollment and exit is small,[13] subject i in arm j will remain after stage 2 so long as

$$d_j\pi_{1ij}+\left(1-d_j\right)\pi_{0ij}\geq\max\left\{\pi_{0ij},\pi_{2ij}\right\}. \tag{3.10}$$

This fact, combined with the fact that the investigator does not observe outcomes among non-enrollees, leaves, e.g., a censored multinomial probit model.

To facilitate estimation, partition the data vector $\mathbf{z}_{ij}$ gathered from recruit i in arm j into $\mathbf{z}_{ij}^p$, variables that affect the specific effect of treatment, and $\mathbf{z}_{ij}^\pi$, variables that affect beliefs about the specific effects of treatment. Assume $p_{kij}=\theta_k^p\mathbf{z}_{ij}^p$ and $f\left(p_{kij}\right)=\gamma_k p_{kij}$, for all i, j, k. If $\gamma_k=1$, then individuals have rational

12. An assumption that individuals have rational expectations does not solve this problem. As I explained in Section 3.1, even if individuals in the population have rational expectations, the subpopulation that enrolls in the trial will not. Rather they will include individuals who, due to mean zero but nevertheless real error, overestimate the value of the new treatment or underestimate the value of no or conventional treatment.

13. This is not unreasonable. An investigator can drive the cost arbitrarily close to zero by reducing the amount of information z gathered from enrollees at stage one or by foregoing z and immediately jumping to the stage two randomization after informed consent. In the latter case, instead of informing subjects of the different arms of the trial and their ex ante probability of being given the chance of being assigned to each arm, the investigator could simply inform subjects of their (randomly-selected) stage two probability of being treated. For some this number could be close to one, which ought to attract nearly every member of the interested population.

expectations. Even if the population at large has rational expectations, however, the population of interest likely does not. Thus $E\left(\varepsilon_{kij}|s=1\right)\neq 0$. However, one might assume $\varepsilon_{kij}=\boldsymbol{\theta}_k^{\pi}\mathbf{z}_{ij}^{\pi}+\upsilon_{kij}$, where υ_{kij} is i.i.d. normal with mean zero and variance σ_{υ}^2 and $E\left(\upsilon_{kij}|s=1,z_{ij}\right)=E\left(\upsilon_{kij}|s=1\right)=0$, for all i, k. Thus, beliefs can be written $\pi_{kij}=\boldsymbol{\theta}_k\mathbf{z}_{ij}+\upsilon_{kij}$ for all i, j, k, where $\boldsymbol{\theta}_k=\left(\boldsymbol{\gamma}_k\boldsymbol{\theta}_k^p,\boldsymbol{\theta}_k^{\pi}\right)'$ and outcomes can be written $y_{ij}=\boldsymbol{\beta}\mathbf{x}_{ij}+w_{ij}$, where $\mathbf{x}_{ij}=\left(\mathbf{z}_{ij},\mathbf{z}_{ij}D_{ij},\mathbf{z}_{ij}d_j,\mathbf{z}_{ij}D_{ij}d_j,\right)$, and $w_{ij}=\lambda_{ij}\left(d_j\upsilon_{1ij}+\left(1-d_j\right)\upsilon_{0ij}\right)+e_{ij}$, where $\lambda_{ij}=\left(D_{ij}a_1+\left(1-D_{ij}\right)a_0\right)$. Note, $\sigma_w^2=\lambda_{ij}^2\left(1-2d_j-d_j^2\right)\sigma_{\upsilon}^2$.

Of course, due to selection and censoring $E\left(w_{ij}\right)\neq 0$ within the population of interest even if $\boldsymbol{\beta}=(\boldsymbol{\beta}_z,\boldsymbol{\beta}_{zD},\boldsymbol{\beta}_{zd},\boldsymbol{\beta}_{zDd})$ estimates the mean $\boldsymbol{\beta}_{ij}$ among the population of interest. To account for this the likelihood function must adjust for the fact that outcomes are only observed for individuals for whom (3.10) holds. The probability of observing outcomes is the probability that $\pi_{0i}\geq\pi_{2i}$ and $\pi_{1i}\geq\pi_{0i}$ plus the probability that $\pi_{2i}\geq\pi_{0i}$ and $d\pi_{1i}+(1-d)\,\pi_{0i}\geq\pi_{2i}$. For expositional convenience, define $\eta_{2i}=\upsilon_{2i}-\upsilon_{0i}$, $\eta_{1i}=\upsilon_{1i}-\upsilon_{0i}$, $\boldsymbol{\psi}=\boldsymbol{\theta}_2-\boldsymbol{\theta}_0$, and $\boldsymbol{\psi}_d=\boldsymbol{\theta}_1-\boldsymbol{\theta}_0$. Note that $\sigma_{w\eta_1}=\lambda_{ij}\left(2d_j-1\right)\sigma_{\upsilon}^2$ and $\sigma_{w\eta_2}=-\lambda_{ij}\left(1-d\right)\sigma_{\upsilon}^2$. Because every individual in the population of interest meets the condition that $\pi_{1i}\geq\pi_{0i}$, the probability of observing outcomes y_{ij} for individual i given she remains in arm j can be written

$$\begin{aligned}
l_{ij} \;=\; & \int_{\infty}^{-\boldsymbol{\psi}\mathbf{z}_{ij}} f\left(y_{ij}-\boldsymbol{\beta}\mathbf{x}_{ij},\eta_{2i};0,0;\sigma_w^2,2\sigma_{\upsilon}^2,\sigma_{w\eta_2}\right)d\eta_{2i} \\
& +\int_{-\boldsymbol{\psi}\mathbf{z}_{ij}}^{\infty}\int_{\left(\eta_{2i}/d_j\right)-\boldsymbol{\psi}\mathbf{z}_{ij}+\boldsymbol{\psi}_d\left(\mathbf{z}_{ij}/d_j\right)}^{\infty} f\left(y_{ij}-\boldsymbol{\beta}\mathbf{x}_{ij},\eta_{1i},\eta_{2i};0,0,0;\right. \\
& \qquad \left.\sigma_w^2,2\sigma_{\upsilon}^2,2\sigma_{\upsilon}^2,\sigma_{w\eta_1},\sigma_{w\eta_2},\sigma_{\upsilon}^2\right)d\eta_{1i}d\eta_{2i},
\end{aligned}$$

where f is the multivariate normal pdf. The probability of not observing outcomes

for those who exit after stage two is

$$\begin{aligned} L_{ij} &= 1 - F\left(-\boldsymbol{\psi}\mathbf{z}_{ij}; 2\sigma_{\upsilon}^2\right) \\ &\quad - \int_{-\boldsymbol{\psi}\mathbf{z}_{ij}}^{\infty} \int_{\left(\eta_{2i}/d_j\right)-\boldsymbol{\psi}\mathbf{z}_{ij}+\boldsymbol{\psi}_d\left(\mathbf{z}_{ij}/d_j\right)}^{\infty} f\left(\eta_{1i}, \eta_{2i}; 0; 0; 2\sigma_{\upsilon}^2, 2\sigma_{\upsilon}^2, \sigma_{\upsilon}^2\right) d\eta_{1i} d\eta_{2i}, \end{aligned}$$

where F is the univariate normal cdf. The likelihood function for the model with selection is $L = \prod_{j=1}^{m}(\prod_{i=1}^{r_j} l_{ij} \prod_{i=r_j+1}^{n_j} L_{ij})$. Again, simply algebra confirms that $(\hat{\boldsymbol{\beta}}_{zD} + \hat{\boldsymbol{\beta}}_{zd} + \hat{\boldsymbol{\beta}}_{zDd})\bar{\mathbf{z}}$, where $\bar{\mathbf{z}} = \sum_{ij} \mathbf{z}_{ij} / \sum_j n_j$, is an unbiased estimator for the short-run parameter of interest.

3.4 Conclusion

The medical literature on trial design strongly supports blinding subjects in clinical trials. It is argued that blinding promotes external validity by reducing bias due to expectations-driven placebo effects. If subjects are not blinded, then their beliefs about treatment efficacy will cloud estimates of the physiological effects of treatment [34]. This chapter rejects this logic. For one thing, the appropriate parameter of interest when there is exist placebo effects is not simply the physiological effects of treatment, but rather those effects plus placebo effects. After all, placebo effects do not vanish outside the context of trials. Moreover, blinding depresses observed outcomes in the active-treatment group and inflates observed outcomes in the no-treatment group. The result is that the different-in-average-outcomes estimator is biased downwards. Thus, blinding actually limits the generalizability of trials.

It tempting to think that unblinded trials are the solution. However, blinding promotes the internal validity of trials. Unblinded trials are susceptible to attrition by subjects randomized into the placebo-control group, many of whom may believe

a conventional alternative is superior to the control. The results is that subjects in the treatment and control groups are not strictly comparable. Moreover, the challenge that blinding poses for external validity in the presence of placebo effects can be addressed by employing the experimental design I proposed in Section 3.3. Although, unblinded trials are a special case of this design—one where all subjects are either in a zero probability of treatment ($d = 0$) or certain treatment ($d = 1$) group, I demonstrate in Appendix D that such a trial would not be $\mathcal{D}$-efficient design, i.e., it would not minimize the variance of one's estimate of treatment response.

The analysis in this chapter may also lead one to wonder whether informed consent is the problem. Without it, subjects would not learn the exact probability of receiving the new treatment and perhaps beliefs would not interfere with estimation of the physiological effects of treatment. This also is incorrect. In the absence of informed consent, subjects would simply rely on their prior beliefs about the probability of treatment. These beliefs would cloud estimates of physiological effects of treatment to the same extent that beliefs updated with information from informed consent does. Nor would lying to patients about their treatment state—as employed in the balance-placebo design—improve estimates. While subjects can occasionally be fooled by investigators, if lying became a common practice, subjects would begin to ignore investigators' instructions about treatment state and rely on their prior beliefs.

CHAPTER 4

THE EFFECT OF SELF-SELECTION ON CLINICAL TRIALS, WITH EVIDENCE FROM ULCER TRIALS

The external validity of an RCT depends on, among other things, the relationship between the population for whom the investigator wants to determine the benefits of the test treatment and the population studied in the trial. If these populations differ and the treatment response is heterogenous, then the results of the RCT may yield biased estimates of treatment response in the population of interest.

The overlap between the population of interest and the trial population depends on the payoff of the test treatment versus the payoff of a RCT of the test treatment. Because consumption of treatment is voluntary, the population of interest includes only those individuals who would choose a test treatment—if it were available outside the trial context—over no treatment or conventional treatment. Because enrollment in a RCT of a test treatment is voluntary, the trial population includes only those individuals who believe that enrolling in such a trial is superior to alternatives. In this chapter, I employ the model of RPCTs from chapter two to demonstrate that this population includes only those individuals who believe that the test treatment is so superior to alternatives that, even with the risk of obtaining no treatment, enrollment in the trial is superior to alternatives.

The population of interest and the trial population may converge when the probability of a trial enrollee being randomized into the test treatment approaches one. This eventuality—so long as there are no monetary incentives to participation in a

trial—implies that the risk of assignment to no treatment vanishes, and patients who even marginally prefer the test treatment, i.e., the entire population of interest, will enroll in the trial. Studying how outcomes observed in a trial change as the probability of receiving the test treatment increases sheds light on the impact of self-selection on outcomes observed in the trial.

In the simple case where health outcomes are univariate; where there are no monetary incentives to participate in the trial; where the population is even mildly competent at estimating their own response to different treatments; and given a technical but fairly lax condition on the population distributions of treatment effects and of the beliefs regarding these treatment effects, my model suggests that increasing the share of enrollees randomized into the test treatment group will lower estimates of average treatment response. The only individuals who would ever consider enrolling in a trial are a subset of those who believe that the test treatment is superior to alternatives, i.e., the population of interest. Other individuals would prefer to forego treatment. Thus, when a trial increases the probability of assignment to the test treatment, the value of the trial to the population of interest increases. Members who previously were marginally not optimistic enough about the test treatment to justify enrollment given the risk of assignment to placebo will now enroll. This influx of relatively pessimistic enrollees lowers observed average treatment response because beliefs and outcomes are positively correlated.

This result also implies that selection bias is positive. As the share of subjects given the test treatment rises to one, two things happen. Average treatment response falls and the subject population converges to the population of interest. Thus, when the share given the test treatment is less than one, mean treatment response must overestimate actual treatment response in the population of interest.

There are a few caveats to these results. First, there are no selection effects in placebo-control trials when the only alternative to the test treatment is no treatment; when the trial has a conventional-treatment control and no one in the population of interest prefers no treatment to conventional treatment; or when there is no heterogeneity in treatment response or beliefs about treatment response. Second, if there are, e.g., monetary incentives to participate in trials, the trial population does not converge to the population of interest as the share given the test treatment goes to one. Some patients who do not believe the test treatment is superior to alternatives may enroll simply because of the money, which would not be available but for the trial. This implies that, even though outcomes still fall as the share given test treatment approaches one, selection bias may not be positive. Third, when outcomes are multivariate, the result that average treatment response falls with the share given test treatment does not hold without much stronger assumptions.

Data from trials of ulcer medications confirm the prediction that average treatment response falls in the fraction of enrollees assigned to the test treatment. The finding is robust to conditioning average treatment response on an array of individual and trial covariates, to different assumptions about attrition from trials, and to investigator-imposed exclusion and inclusion criteria. Most revealing, however, is the fact that, the predicted relationship is not found in trials where subjects are not given informed consent, and thus do not know the exact probability with which they will be assigned to the test treatment.

Numerous papers study the impact of selection on the external validity of RCTs. However, the term selection covers many different phenomena. Most studies examine the impact of exclusion and inclusion criteria on the generalizability of trial results [35, 36, 37]. A few consider the role of informed consent on participation decisions or

the reasons patients give for their decision to participate or not participate in clinical trials. These studies lend support to the theory that patients compare the benefits of the trial with the benefits of alternative treatment strategies. Among these benefits is the probability of assignment to the test treatment, which is revealed through informed consent [1].[1] Although a handful of studies compare the population of eligible patients who do and do not give their consent to enrollment in RCTs [35, 38], none proposes a theory for why the two groups differ or how these differences impact estimates of average treatment response.

While the analysis in this chapter relies on the model of RPCTs in chapter two, it takes a different approach to health outcomes. First, health outcomes are assumed to be a function only of the specific effects of treatment, $\Pr\{y_{ki} = \bar{y}\} = p_{ki}$.[2] Indeed, a distinguishing feature of this chapter is that it examines the impact of self-selection in the absence of placebo effects. Second, let $\mathbf{g}_{\tilde{p}}$ and $\mathbf{g}_{\tilde{\pi}}$ give the probability distribution function of $\tilde{\mathbf{p}}_i = (\tilde{p}_{1i}, \tilde{p}_{2i})$ and $\tilde{\boldsymbol{\pi}}_i = (\tilde{\pi}_{1i}, \tilde{\pi}_{2i})$, respectively, across the population, where $\tilde{p}_{ki} = p_{ki} - p_{0i}$ and $\tilde{\pi}_{ki} = \pi_{ki} - \pi_{0i}$. I assume that $\mathbf{g}_{\tilde{\pi}}$ and $\mathbf{g}_{\tilde{p}}$ are log-concave. This assumption is not restrictive: the class of log-concave distributions is large and includes, e.g., the multivariate normal. Third, I assume $\tilde{p}_{ki} = f(\tilde{\pi}_{ki}) + \varepsilon_{ki}$, where either (a) f is an affine transformation with $f' > 0$ or (b) f is strictly increasing, differentiable, concave and $\left(f^{-1}\right)'$ is log-concave, and where $\tilde{\varepsilon}_{ki}$ is independent of $\pi_{k'i'}$ for all $\left(k', i'\right)$ and of $\varepsilon_{k'i'}$ for all $\left(k', i'\right)$ except $\left(k' = k, i' = i\right)$. The error term ε_{ki} reflects error in predictions of relative efficacy by individual i.

Section 4.1 examines the impact of self-selection on beliefs of trial participants and outcomes observed in RPCTs. It also presents conditions under which selection bias

1. However, many of these studies, which use survey instruments, also point to the benefits to science as a reason for participation in trials [1].

2. If $a_1 = a_0$ the analysis is unaffected by the existence of placebo effects.

can be signed. Section 4.2 presents data from ulcer trials that validate the predictions from Section 4.1 about the effect of self-selection on trial outcomes. Appendix E extends the analysis to the case of continuous health outcomes. Appendix F presents proofs for propositions in the main text.

4.1 Impact of Self-Selection on Beliefs and Outcomes

4.1.1 Test treatment only available via RPCT and treatments are costless

Initially assume that the test treatment is only available in the context of a trial and that the costs of the trial to enrollees, pecuniary or otherwise, are identical to the costs of alternative treatment strategies to enrollees. The first assumption is reasonable for clinical trials before a drug is approved, e.g., by the U.S. Food and Drug Administration. The second assumption may be reasonable for patients who are well-insured; these patients do not bear the marginal cost of specific treatments.

Because consumption of treatment is voluntary, the population of interest (POI) for the investigator ought to be those patients who, if the test treatment were available outside the context of a trial, would choose the test treatment over non-trial alternatives.[3] If choice is driven by expected utility maximization, the POI includes

3. The difficulty with specifying the POI is that it may change over time, perhaps even in response to the publication of the results of the RCT. The logic of the Heisenberg uncertainty principle suggests the latter problem is unavoidable. I address this by assuming the population distribution of beliefs is time invariant but acknowledging that the discussion below is misleading insofar as it assumes an estimator is "unbiased" if its expected value in the trial sample is the same as its value in the population of interest, given that the beliefs of the latter do not change over time.

all individuals for whom $U_i^1 \geq \max\left\{U_i^0, U_i^2\right\}$ or, equivalently,

$$\tilde{\pi}_{1i} \geq \max\left\{0, \tilde{\pi}_{2i}\right\}, \tag{4.1}$$

where $\tilde{\pi}_{ki} = \pi_{ki} - \pi_{0i}$. A patient will choose the test treatment only if she perceives its advantage over no treatment is positive and greater than the perceived advantage of conventional treatment over no treatment.

The investigator observes health outcomes only for those patients who enroll in the trial. This group is comprised of patients who, if faced with a choice of no treatment, conventional treatment, or enrollment in a blinded RPCT, would choose the last of these options. These are patients for whom $U_i^{BT} \geq \max\left\{U_i^0, U_i^2\right\}$ or, equivalently,

$$\tilde{\pi}_{1i} \geq \max\left\{0, \frac{\tilde{\pi}_{2i}}{d}\right\}. \tag{4.2}$$

A patient will enroll in a trial only if the mere chance d of obtaining the advantage of the test treatment over no treatment is positive and better than the advantage of conventional treatment over no treatment. The population observed is a subset of the population of interest.

There are three conditions under which selection does not affect the external validity of trial results. First, if $\mathbf{g}_{\tilde{p}}$ is degenerate, there may be selection but it has no impact on outcomes. Most trials implicitly make this assumption, which is identical to an assumption of constant treatment effects. Second, if $\mathbf{g}_{\tilde{\pi}}$ is degenerate, there is no selection. There may be heterogeneity in treatment response, but the enrollee population is a random draw from the POI. Third, if there is no conventional treatment alternative, (4.1) and (4.2) both collapse to $\tilde{\pi}_{1i} \geq 0$. The enrollee population and the POI are one and the same.

In order to determine the impact of the disconnect between the POI and the observed population when none of these three conditions hold, the investigator must know something about the population distribution of beliefs and the relationship between beliefs and the actual efficacy of treatments. One approach is to estimate $\mathbf{g}_{\tilde{\pi}}$ and the relationship between $\mathbf{g}_{\tilde{\pi}}$ and $\mathbf{g}_{\tilde{p}}$ directly. Estimating $\mathbf{g}_{\tilde{\pi}}$ may be difficult because beliefs are not easily observable. One may be able to estimate the relationship between $\tilde{\boldsymbol{\pi}}$ and $\tilde{\mathbf{p}}$ by applying a censored or truncated regression model to data from trials with different shares treated. However, trials rarely provide data on non-participants and a truncated regression model cannot validate the specification of the selection equation. An alternative approach is to make assumptions about $\mathbf{g}_{\tilde{\pi}}$ or $\mathbf{g}_{\tilde{p}}$. If these are "reasonable" and confirmed by empirical validation of the resulting predictions, they can be used to make predictions about the effect of selection on trial observations. This is the avenue I pursue.

Proposition 3 *Define* $\tilde{\rho}_{12} = corr\left(\tilde{p}_{1i}, \tilde{p}_{2i}\right)$ *and* $\tilde{\sigma}_k^2 = var\left(\tilde{p}_{ki}\right)$. *If* $\tilde{\rho}_{12} \leq \tilde{\sigma}_1/\tilde{\sigma}_2$, *then* $\partial E\left(\tilde{p}_{1i}|s(d) = BT\right)/\partial d \leq 0$.

This proposition states that, so long as the correlation between the advantage of the test treatment and that of the conventional treatment is not too great, an increase in the share of enrollees randomized into the new treatment group will lower the difference in the average response to the test treatment among enrollees. The intuition is that the trial only attracts those individuals who believe that the advantage of the test treatment is so great than, even with a probability $1 - d$ of obtaining no treatment at all, the trial offers a better expected outcome than conventional treatment. As the probability d of obtaining the test treatment rises, the benefit of the trial increases. The reason is that only those individuals who believe that the test treatment is superior to no treatment are candidates for the trial; the rest would

prefer no treatment are not even part of the POI. As the benefit of the trial rises, some individuals who were not optimistic enough about the advantage of the test treatment to have previously enrolled, will now consider doing so. This will lower the average beliefs about the test treatment among enrollees. Since outcomes are monotonically increasing in beliefs, expected outcomes rise with expected beliefs.

If the assumption that the correlation between the advantages of the test and new treatments is not too high is violated, then a higher probability of obtaining the test treatment may attract not just individuals with more pessimistic beliefs about the test treatment, but also those with more optimistic beliefs about the test treatment who nevertheless have not enrolled because they also have more optimistic beliefs about the conventional treatment. The selection condition (4.2) does not merely require that $\tilde{\pi}_{1i}$ be high, but that it be greater that $\tilde{\pi}_{2i}/d$. The upper bound on $\tilde{\rho}_{12}$ is reasonable. It is satisfied whenever $\tilde{\sigma}_1^2 \geq \tilde{\sigma}_2^2$, i.e., the variance of beliefs about the advantage of the test treatment is greater than the variance of beliefs about the conventional treatment. This is likely because the test treatment is new.

So long as subjects are properly randomized into each group and there is no attrition, the standard estimator of treatment response—the difference in mean outcomes in each treatment group—produces unbiased estimates of $E\left(p_{1i}|s=BT\right) - E\left(p_{0i}|s=BT\right) = E\left(\tilde{p}_1|s=BT\right)$. Proposition 3 suggests that as the share randomized into the test treatment group of a trial rises, the standard estimate of treatment response should fall. I test this prediction, which is the main result of this chapter, in Section 4.2.[4]

4. This prediction is robust to the use of exclusion and inclusion criteria when selecting enrollees so long as these criteria are independent of beliefs about treatment efficacy. In that case, $E\left(\tilde{\pi}_{1i}|s\left(d\right)=BT, x_i \in C\right) = E\left(\tilde{\pi}_{1i}|s\left(d\right)=BT\right)$, where x_i are covariates that affect exclusion and inclusion and C defines the exclusion and inclusion criteria. The prediction is also robust to imperfect informed consent so long as whether candidates are informed of

Proposition 3 also implies that selection bias is positive. As the share d of enrollees given the test treatment converges to one, the selection condition (4.2) converges to the characterization (4.1) of the POI. But as d rises, Proposition 3 predicts that expected treatment response among enrollees falls. Thus expected treatment response must exceed treatment response in the POI so long as $d < 1$. Again, so long as enrollees are properly randomized across treatments and there is no attrition, the standard estimator will also overestimate treatment response among the POI.

4.1.2 Monetary incentives for participation

In this subsection I relax the assumption that the costs of the trial and available treatments outside the trial are identical. This implies that selection may also take place on the basis of cost. To see this, assume individuals draw utility from health and other consumption, including leisure and wealth. In order to highlight just the basic results, assume further that utility is additively separable: $u(y, m) = v(y) - w(m)$, where $v' > 0$, $v'' < 0$, $w' < 0$, $w'' > 0$, and m is other consumption. The expected utility of strategy s is $U_i^s = \pi_i^s v(\bar{y}) + \left(1 - \pi_i^s\right) v\left(\underline{y}\right) - w(m^s)$, where m^s is consumption given the pecuniary and nonpecuniary costs of treatment strategy s. Note two things. First, a lower cost strategy is associated with a higher m^s and thus a lower $w^s = w(m^s)$. Second, I assume consumption depends on strategy but not individual characteristics; this simplifies the analysis below but probably is realistic only if one further assumes that utility is separable in the costs of a treatment strategy and that the cost of each strategy is constant across individuals.

---

d via informed consent is independent of the value of d. In that case, $E(\tilde{\pi}_{1i}|s = BT) = \Pr(\delta_{1i} = d) E(\tilde{\pi}_{1i}|s(d) = BT) + \Pr(d \text{ not revealed}) \int E\left(\tilde{\pi}_{1i}|s\left(\bar{d}_i\right) = BT\right) h\left(\bar{d}_i\right) d\bar{d}_i$, where $\bar{d}_i$, individual priors about the probability of receiving the test treatment in a trial, are distributed among the population according to h.

If the test treatment were available outside the context of a trial, a patient would choose the test treatment if $U_i^1 \geq \max\left\{U_i^0, U_i^2\right\}$. This implies $\tilde{\pi}_{1i} - \tilde{w}^1 \geq \max\left\{-\tilde{w}^0, \tilde{\pi}_{2i} - \tilde{w}^2\right\}$, where $\tilde{w}^s = w^s / \left[u\left(\bar{y}\right) - u\left(\underline{y}\right)\right]$. Normalize utility such that the value of consumption given a strategy of foregoing any treatment is zero. (This can serve as a reminder that the time and money costs of a no treatment strategy may be zero.) Now the POI includes all individuals who satisfy

$$\tilde{\pi}_{1i} \geq \max\left\{\tilde{w}^1, \tilde{\pi}_{2i} - \left(\tilde{w}^2 - \tilde{w}^1\right)\right\}. \tag{4.3}$$

This is similar to (4.1), except that individuals discount the value of each treatment by its cost, in utility terms. Note that, if the costs of the test and conventional treatments exceed that of no treatment, $\min\left\{w^1, w^2\right\} \geq w^0 = 0$.

Enrollees include anyone for whom

$$\tilde{\pi}_{1i} \geq \max\left\{\frac{\tilde{w}^{BT}}{d}, \frac{\tilde{\pi}_{2i} - \left(\tilde{w}^2 - \tilde{w}^{BT}\right)}{d}\right\}.$$

There are now two differences between enrollees and the POI. As before, the former only get the test treatment with probability d. However, now the former may also pay a different cost than the latter. While it may be reasonable to suppose that the time cost to enrollees is similar to the time cost to the POI, unless the investigator charges enrollees the same monetary price for the trial as the POI will be charged for the test treatment after it is approved, $\tilde{w}^1 \neq \tilde{w}^{BT}$.

There are two major modifications to the results from the previous subsection. First, the share given the test treatment triggers selection effects even if there is no conventional treatment. The enrollment condition is $\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d$. Assume that $\tilde{w}^{BT}$ is not a function of d and that $\tilde{\pi}_{ki} = f\left(\tilde{\pi}_{ki}\right) + \varepsilon_{ki}$, where f is strictly increasing

and differentiable and where $\tilde{\varepsilon}_{ki}$ is independent of $\pi_{k'i'}$ for all $\left(k', i'\right)$ and of $\varepsilon_{k'i'}$ for all $\left(k', i'\right)$ except $\left(k' = k, i' = i\right)$. The model predicts the same relationship between d and outcomes as Proposition 3 whatever the distribution of $\tilde{\mathbf{p}}_i$ and $\tilde{\boldsymbol{\pi}}_i$.

Proposition 4 $\partial E\left(\tilde{p}_{1i} | s(d) = BT\right) / \partial d \leq 0$.

If $w^{BT} \geq w^1$, the selection bias will be positive. This is because, even if $d = 1$, the cutoff for enrollment is greater than the cutoff for inclusion in the POI. This logic suggests that if the investigator could set $w^{BT} = dw^1$, then the enrollment condition would be identical to (4.3). This is difficult because costs need to be adjusted in units of utility, but the investigator can only control variables such as time and monetary costs. However, a good approximation is to set the time required to complete the trial and the monetary cost of enrollment equal to d times the analogous costs when consuming the test treatment once it is available.

The second change pertains to the case where a conventional alternative is available. Under the same assumptions as required for Proposition 3, plus the assumption that $\tilde{w}^s$ is not a function of d for $s = 2, BT$, the result in Proposition 3 remains valid.

Proposition 5 *If* $\tilde{\rho}_{12} \leq \tilde{\sigma}_1 / \tilde{\sigma}_2$, *then* $\partial E(\tilde{p}_{1i} | s(d) = BT) / \partial d \leq 0$.

The proof is simplified by the assumption that the costs of each strategy are constant across individuals. If costs varied, assumptions about the distribution of $w_i^{BT} - w_i^2$ would be required to ensure $\left(\tilde{\pi}_{1i}, \tilde{\pi}_{2i}^*\right)$ was log-concave. The assumption that $\tilde{w}^{BT} - \tilde{w}^2$ is not a function of d can be verified by checking that d is not set with an eye towards the cost the trial on the benefits of good health or the costs of treatment. There is a danger that this assumption is violated in practice.

An important risk with the use of monetary incentives is that the population of enrollees may not be a subset of the POI. If pay-outs are such that w^{BT} is

sufficiently smaller than w^1, then some individuals may find it profitable to enroll in a trial even though they would not voluntary pay for the test treatment outside the trial context. In this case it is not possible to determine the sign of selection bias even given Proposition 5. Above some d^*, some individuals enroll who are not part of the POI. Proposition 5 says they manifest lower treatment response. If the trial has $d \gg d^*$, estimates of average treatment response from the trial could be lower than average treatment response in the POI. So selection bias could be negative.

However, if $w^{BT} \geq \max\left\{w^1, w^2\right\}$, Proposition 5 still implies positive selection bias. As d converges to 1, individuals for whom $\tilde{\pi}_{2i} - \tilde{w}^2 \leq 0$ will enroll only if $\tilde{\pi}_{1i} \geq \tilde{w}^{BT} \geq w^1$ and individuals for whom $\tilde{\pi}_{2i} - \tilde{w}^2 > 0$ will enroll only if $\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i} + \tilde{w}^{BT} - \tilde{w}^2 \geq \tilde{\pi}_{2i} + \tilde{w}^1 - \tilde{w}^2$. Thus the population of enrollees is a subset of the POI that satisfies higher cutoffs for $\tilde{\pi}_{1i}$. Thus average $\tilde{\pi}_{1i}$ and, by monotonicity of f, average $\tilde{p}_{1i}$ are higher among enrollees than the POI.

4.1.3 Test treatment available outside RCT and monetary incentives

In this subsection I relax the assumption that the test treatment is only available to patients who enroll in the trial. If there were no differences in cost across the treatment strategies, no one would enroll in a trial. Those who prefer the test treatment would simply choose the test treatment strategy. Therefore, I also assume there may be cost differences across treatment strategies.

The population of interest in this case is identical to (4.3). However, the enrollee population differs. It includes all who satisfy

$$\tilde{\pi}_{1i} \geq \max\left\{\frac{\tilde{w}^{BT}}{d}, \frac{\tilde{\pi}_{1i} - \left(\tilde{w}^1 - \tilde{w}^{BT}\right)}{d}, \frac{\tilde{\pi}_{2i} - \left(\tilde{w}^2 - \tilde{w}^{BT}\right)}{d}\right\}$$

If $\tilde{w}^{BT} \geq \tilde{w}^1 - (1-d)\min_i \tilde{\pi}_{1i} = \tilde{w}^*$, no one will enroll because the test treatment strategy is superior to the trial strategy. Thus $\tilde{w}^*$ defines the minimum incentive required for the enrollee population to be nonempty. If this condition is satisfied, the assumptions required for Proposition 3 hold, and $\tilde{w}^s$ is not a function of d for $s = 1, 2, BT$, then the result in Proposition 3 remains valid.

Proposition 6 *If $\tilde{\rho}_{12} \leq \tilde{\sigma}_1/\tilde{\sigma}_2$, then $\partial E(\tilde{p}_{1i}|s(d) = BT)/\partial d \leq 0$.*

Unfortunately selection bias cannot be signed. Unless $w^{BT} \geq \max\left\{w^1, w^2\right\}$, one cannot be sure that selection bias is positive for the reasons given in the last subsection. However, $w^{BT} \geq w^1$ implies $\tilde{w}^{BT} \geq \tilde{w}^1 - (1-d)\min_i \tilde{\pi}_{1i}$, so no one will enroll in the trial.

4.2 Trials of Ulcer Medications

In this section I test the predictions of Section 4.1 with data drawn from the published results of 150 double-blind RCTs of anti-ulcer medications. The data are described in detail in chapter two.

For simplicity, I assume that the physiological and natural progression effects of treatment k on individual i enrolled in trial j is a linear function of a vector x_{ijk}, which includes a constant, observed clinical and demographic variables on the individual in treatment arm k, and structural features of the trial, and an error term: $p_{ijk} = \beta_k' x_{ijk} + \varepsilon_{ijk}$. The error term captures variation due to unobserved characteristics of trial participants. I assume the errors are independent of x_{ijk} and are i.i.d. normal across individuals with mean zero and variance σ_k. This implies unobserved effects do not depend on the trial, but may depend on the treatment.

Because trials often take multiple measurements on each individual, I interpreting a treatment's effect as a hazard rate. Assuming it is constant over time implies $-\ln S_{ijk}(t)/t = \beta_k' x_{ijk} + \varepsilon_{ijk}$, where $S_{ijk}(t)$ gives the probability of still having an unhealed ulcer on date t. Summing over individuals and dividing by n_{jk}, the number of subjects enrolled in treatment arm k of trial j, yields the regression equation $-\overline{\ln S_{jk}(t)}/t = \beta_k' \bar{x}_{jk} + \bar{\varepsilon}_{jk}$. I approximate $\overline{\ln S_{jk}(t)}$ just as in chapter one. The standard estimator of treatment response is the difference in average outcomes in the test treatment and control groups. Comparing groups k and k' at time t yields $-[\ln \bar{S}_{jk}(t) - \ln \bar{S}_{jk'}(t)]/t = \beta_k' \bar{x}_{jk} - \beta_{k'}' \bar{x}_{jk'} + \bar{\varepsilon}_{jk} - \bar{\varepsilon}_{jk'}$.

Self-selection is a problem only if it is not fully captured by $\bar{x}_{jk}$ and the unobservable covariates that capture selection are correlated with $\bar{x}_{jk}$. Section 4.1 suggests that selection is driven by the share randomized into the test treatment group; indeed, it predicts that there ought to be a negative relationship between treatment response and that share. To test this hypothesis I add powers of d_j, the share of enrollees in trial j not given the control, as a regressor:

$$-[\ln \bar{S}_{jk}(t) - \ln \bar{S}_{jk'}(t)]/t = \beta_k' \bar{x}_{jk} - \beta_{k'}' \bar{x}_{jk'} + \textstyle\sum_{n=1}^{N} \gamma_n d_j^n + \bar{\varepsilon}_{jk} - \bar{\varepsilon}_{jk'}. \qquad (4.4)$$

In the presence of self-selection, $\sum_{n=1}^{N} \gamma_n d_j^{n-1}$ should be negative.

Table G.8 presents full coefficient estimates for (4.4) under three specifications of $\bar{x}_{jk}$ and $\bar{x}_{jk'}$ and with only one power of d_j as a regressor.[5] Trials of H_2-blockers

5. Estimation was by feasible GLS. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. My regression model suggests that the variance of error terms depends on the share randomized into each arm. However, I measure only one randomization share per trial, namely the share not given a lower class or non-healing control. Therefore I permit

and proton-pump inhibitors (PPI) manifest significant evidence of selection: the coefficients on d_j are all negative and significant. Unfortunately, this is not very strong evidence for selection because all but the third specification for the PPI trials fail the link test and misspecified.

To address this problem, I incrementally add powers of d_j (up to five) to each specification until it passes the link test. If the specification never passes the link test I retain as a regressor only the power series that produces the lowest value for the link test statistic. Table G.9 reports the results for all six specifications of trials with informed consent and for four to five specifications for trials where investigators did not confirm that subjects were asked for informed consent before enrollment.[6] Because the value of $\sum_{n=1}^{N} \gamma_n d_j^{n-1}$ depends on the value of d_j for each trial, the table lists the number of measurements on arms for which the estimate of this value is negative at various confidence levels.

Trials of H_2-blockers and PPI's manifest strong and moderate evidence of self-selection, respectively. All measurements in H_2-blocker trials report selection effects

---

group-wise heteroskedasticity at the trial-level, but not at the arm-level. I only report estimates where the dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Results from regressions which assume that those who attrite out either all heal or all do not heal are not materially different.
I estimate six specifications of $\bar{x}_j$. Specification (1) includes a constant and certain powers of d_j; (2) includes (1) and trial level-variables (antacid role, daily frequency of medication, total daily dosage); (3) includes (2) plus subject-oriented covariates (male, smoker, age) for treatment and control arms separately; (4) includes (2) plus subject-oriented covariates averaged over the treatment and control groups, with each group assigned a weight in proportion to the number evaluated per protocol in that measurement. (This is a reasonable approximation for $\bar{x}_{jk}$ and $\bar{x}_{jk'}$ so long as there are no flaws in the random allocation scheme employed by the trials in my sample.) (5) and (6) are the same as (3) and (4) but without the trial-level variables added by (2). I test each specification with Pregibon's link test[23]. Those that fail are marked with a dagger (†). I also checked the residuals for patterns but did not find any serious problems.

6. Specifications with a large number of regressors were dropped because of the small number of trials without informed consent.

at the 95 percent confidence level. In PPI trials, two out of the four specifications that pass the link test report selection effects in every measurement at the 95 percent confidence level. One of the remaining specifications reveals evidence of selection in all measurements at the 85 percent confidence level. More remarkable is the fact that, for both these types of medication, trials without informed consent manifest no evidence at all of selection driven by share treated! This lends strong support to the view that informed consent reveals the share treated and that share treated drives self-selection.[7]

4.3 Conclusion

This chapter offers a theoretical analysis of the effects of self-selection on clinical trials. This analysis is confirmed by data that reveals a negative relationship between treatment shares and estimates of treatment response from trials of two anti-ulcer medications. This finding suggests that self-selection may impair the external validity of clinical trials. Before this conclusion can be fully embraced, however, the findings in this chapter themselves must be internally and externally validated. First, the findings ought to be checked against the results of a censored regression model applied to individual-level data from ulcer trials. This chapter makes assumptions about the relationship between beliefs and outcomes and finds the implications of these assumptions confirmed in trial arm-level data. This is an indirect proof of the assumptions. A censored regression model would provide more direct proof. Second, the analysis in this chapter ought to be extended to trials based on other designs, such as without those without parallel-arms or blinding, and applied to trials of other

7. Prostaglandin trials reveal inconsistent evidence of self-selection. The number of trials without informed consent is too small to draw any meaningful conclusions.

medications. If the conclusion of this chapter is ultimately validated, there will be a demand for alternative trials designs that are robust to selection pressures and can generate estimates of useful parameters for the population of interest.

APPENDIX A

RECONCILING PLACEBO EFFECTS AND SUBJECTIVE EXPECTED UTILITY THEORY

In this appendix I offer a model of the world that is simultaneously compatible with both my model of treatment choices and Savage's axiomatization of expected utility [39]. The only difference between the model in the main text and the model below is that states are defined by *inter alia* whether a treatment works (or how effectively it works), whereas health outcomes are treated as consequences. (There is a one-to-one correspondence between whether a treatment works (state variable) and health outcomes (consequence variable), but this is not problematic under Savage's formulation [40].)

Although the model in the main text assumes individuals do not take placebo effects into account when making treatment decisions, the model below can accommodate the assumption that individuals do take placebo effects into account. In the latter case, states must also be defined by individuals' beliefs about which treatment is being consumed and about the efficacy of treatments. The reason is that, in the absence of states defined by beliefs, states which, one the one hand, have the same consequences under blinded and unblinded trials also have, on the other hand, assigned to them different subjective probabilities given enrollment in a blinded and unblinded trials. Savage's axiomatization, however, require that actions such as enrollment in different types of trials not influence probabilities. Defining states by beliefs solves this problem.

Without loss of generality, suppose there are only three treatments (indexed by $k = 0, 1, 2$) and two possible health consequences (recovery or no recovery). Let Y be the set of health consequences, Θ the set of states of the world, and S the set of all feasible actions. Each $\theta \in \Theta$ is described by whether treatment k would work if taken, by the individual's assessment of the probability of recovery with treatment k, and by the individual's assessment of the probability that she is consuming treatment k, for all k. In other words, states are defined by whether a treatment works and the individual's subjective beliefs about the treatment. Each action $s \in S$ is described by the vector δ giving the probability with which the individual is given each treatment, where $\sum \delta_k = 1$, and by whether the individual will be told her treatment assignment. The latter factor permits enrollment in blinded trials.

The key assumptions behind Savage's axiomatization of subjective expected utility are that

- acts do not influence probabilities (Savage's third and fifth axioms)
- utility is not state-dependent (fourth axiom)
- and preferences conform to the sure-thing principle (sixth axiom).

The sure-thing principle states that if the consequences of one action is preferred those of another in each of some subset of states, then the first action is preferred to the second given that subset of states.

It is easiest to demonstrate that the formulation I propose conforms to these axioms by working through a few examples. First consider the simple case (case 1) where there are only two treatments (0 or 1); the feasible actions are choosing no treatment ($s(\theta, 0)$), choosing treatment 1 ($s(\theta, 1)$) or an unblinded lottery ($s(\theta, UT)$) that administers either treatment 0 or 1 (and no other) with equal probability; beliefs

can either be that treatment k works or does not; and there are no placebo effects. This means there are 8 unique states spanning the scenarios where treatment 0 would work or not; treatment 1 would work or not; and the lottery yields assignment of treatment 0 or 1. (Beliefs do not matter to states because there are no placebo effects.) If (π_0, π_1) are the individual's subjective assessments of the probability that treatments $(0, 1)$ would work and these probabilities are thought to be independent, then the probabilities assigned to each state are given in Table G.10. The table also gives the consequence of each action in each state. So, for example, if the individual's action is enrollment in the lottery and she lands in the state where the lottery assigns her to treatment 1 and where neither treatment 0 or 1 works, then the consequence—given in the bottom right cell of the table—is no recovery.

Note that acts do not influence probabilities since states are defined by the outcomes of all possible acts. This is a bit tautological, but not impermissible for my purposes (which are not to challenge Savage's axioms). Utility is only a function of recovery or no recovery. Utility of either consequence does not depend on the state in which it occurs. Finally, the sure-thing principle holds with respect to actions. For example, for all states θ where treatment 1 works, action $s(\theta, 1)$ yields prizes that are at least as good those of $s(\theta, UT)$. So long as the individual prefers taking treatment 1 to the lottery given that treatment 1 works, the sure-thing principle is not violated. This exercise can be performed with any subset of states in this placebo effect-free version of my model of treatment choices.

Consider a second case (case 2) that is identical to case 1 except that there exist placebo effects but individuals do not consider them in making treatment decisions. The formulation of states, actions and consequences in case 1 would be compatible with case 2. The reason is that, although the individual's beliefs affect the objective

probability that different consequences are observed by the investigator, they do not affect the individual's decision. More precisely, although the objective probability of recovery given consumption of treatment k is some function $f(p_k, \pi_k)$, the individual only assigns recovery given treatment k a probability of π_k.

An important variation on this case (case $2'$) is where the lottery is blinded. This changes the objective probability of the investigator observing various consequences because beliefs differ in blinded and unblinded trials. (In an unblinded trial, the probability a treatment works depends only on the specific effects of that treatment and beliefs about the specific effects of that treatment. In a blinded trial, that probability also depends on beliefs about the specific effects of other treatments to which the trial may assign the individual.) However, blinding does not affect the probabilities that an individual assigns to different states. Thus the description of states, consequences and subjective probabilities in Table G.10 remain valid.

Finally, consider a third case (case 3) that is identical to case 2, except that individuals are aware of placebo effects and consider them in their decisionmaking. In this case the states and consequences in Table G.10 are valid. However, subjective probabilities will change to account for the role of beliefs. The result will be Table G.11, where $\phi_0 = E[f_0(p_0, \pi_0)]$, $\phi_1 = E[f_1(p_1, \pi_1)]$, and f_k is the objective probability of treatment k working given specific effect p_k of treatment and beliefs π_k about the specific effects of treatment. (Note that $\pi_k = E[p_k]$, although ϕ_k may not equal π_k.) Awareness of placebo effects alter subjective probabilities, but nothing else.

The main complication is the variant of this case (case $3'$) where the lottery is blinded. This requires the introduction of additional states that reflect the individual's beliefs about the probability she is consuming treatment k and about the efficacy of treatment k, for all k. The reason is that, in the absence of states also defined

by beliefs, states which have the same consequences under blinded and unblinded lotteries have assigned to them different subjective probabilities given enrollment in a blinded or an unblinded trial. (In the latter, the probability is simply a function of the specific effects of the treatment to which the individual is assigned and beliefs about the specific effects of that treatment. In the former, the probability is also a function of beliefs about which treatment the individuals has been administered and about the efficacy of all the treatments to which the individual may have been assigned.) Savage's third and fifth axioms require, however, that acts not influence probabilities. Defining states by beliefs solves this problem. (This move is not troublesome because beliefs are known to the subject and there is no prohibition on defining states by beliefs since all states can so be defined: state θ, which is assigned a probability of 1/2, can be redefined to be the state that is assigned a probability of 1/2.) To see how this might work, consider a modification of case 3 where the only two possible actions are enrollment in a blinded trial or in an unblinded trial and that each trial offers an equal probability of treatment 0 or 1. Define

$$\begin{aligned}
\phi_0' &= 1-\phi_0, \qquad \phi_1' = 1-\phi_1 \\
\phi_0^{BT} &= E\left[f_0\left(p_0, \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1\right)\right] \\
\phi_1^{BT} &= E\left[f_1\left(p_1, \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1\right)\right] \\
\phi_0^{BT\prime} &= 1-\phi_0^{BT}, \qquad \phi_1^{BT\prime} = 1-\phi_1^{BT}.
\end{aligned}$$

Table G.12 presents all *non-zero* probability states, the consequences of each action in each state, and the subjective probability assigned to each state.

APPENDIX B

TESTS FOR PLACEBO EFFECTS GIVEN ALTERNATIVE MODELS OF OUTCOMES

B.1 Subjects Consider Placebo Effects when Forming Beliefs

The definition of placebo effects in the text assumes that when individuals form beliefs about a strategy, they only consider the specific effects of treatment. It is possible that individuals take placebo effects into account. In that case, a double-blind, parallel-arm, RPCT with a linear model of outcomes and specific-effect cutoffs yields outcomes $\Pr\{y_{ki}=1\} = (1-a_k)\,p_{ki} + a_k\pi_i^{BT}$, for $k=0,1$, where

$$\pi_i^{BT} = d\left[(1-\alpha_{1i})\,\pi_{1i} + \alpha_{1i}\pi_i^{BT}\right] + (1-d)\left[(1-\alpha_{0i})\,\pi_{0i} + a_{0i}\pi_i^{BT}\right]$$

and α_{ki}—beliefs about the influence of beliefs a_k in treatment state k—are assumed independent of the efficacy of the treatment and beliefs about the efficacy of treatment.[1] Define $\tilde{\alpha}_{ki} = (1-\alpha_{ki})$. If $a_1 = a_0$, then $\alpha_{1i} = \alpha_{0i}$ and $\pi_i^{BT} = d\pi_{1i} + (1-d)\,\pi_{0i}$, the same beliefs as individuals who do not take placebo effects

1. I do not take a position on how beliefs about the influence of beliefs are generated. However, it is likely that subjects self-sample to form beliefs as they would for beliefs about efficacy. Assumptions about the distribution of α_{ki} (log-concave or log-convex) are required to generate predictions about the relationship between share treated and outcomes. But these do not place obvious constraints on the generation of α_{ki}.

into account when estimating the value of enrolling in a trial. The tests for placebo effects set forth in the main text remain valid.

Even if $a_1 \neq a_0$, many of the results from the text hold. Simple algebra yields

$$\pi_i^{BT} = \frac{d\tilde{\alpha}_{1i}\pi_{1i} + (1-d)\,\tilde{\alpha}_{0i}\pi_{0i}}{d\tilde{\alpha}_{1i} + (1-d)\,\tilde{\alpha}_{0i}}, \qquad \frac{\partial \pi_i^{BT}}{\partial d} = \frac{\tilde{\alpha}_{0i}\tilde{\alpha}_{1i}}{\left[d\tilde{\alpha}_{1i} + (1-d)\,\tilde{\alpha}_{0i}\right]^2}\left(\pi_{1i} - \pi_{0i}\right),$$

i.e., beliefs about the value of the trial strategy are an increasing function of the share treated. If there is no conventional treatment option, the selection equation is

$$\frac{d\tilde{\alpha}_{1i}\pi_{1i} + (1-d)\,\tilde{\alpha}_{0i}\pi_{0i}}{d\tilde{\alpha}_{1i} + (1-d)\,\tilde{\alpha}_{0i}} \geq \pi_{0i}$$

which collapses to $\pi_{1i} \geq \pi_{0i}$. Selection is not a function of the share treated. Thus, a positive relationship between share treated and outcomes is a valid test for placebo effects. This result also holds in the case of conventional-control trials and a population where patients who believe the new treatment is superior to conventional treatment also all believe conventional treatment is superior to no treatment.

If there is a conventional alternative, the selection equation becomes

$$\alpha_{1i}\left(\pi_{1i} - \pi_{2i}\right) \geq \frac{\alpha_{0i}\left(\pi_{2i} - \pi_{0i}\right)}{\tilde{d}},$$

where $\tilde{d} = d/\left(1-d\right)$ is monotonically increasing in d. If the joint distribution of $(\alpha_{1i}(\pi_{1i} - \pi_{2i}), \alpha_{0i}(\pi_{2i} - \pi_{0i}))$ is log-concave or log-convex, the logic of Proposition 2 suggests that, so long as $\alpha_{1i}\left(\pi_{1i} - \pi_{2i}\right)$ and $\alpha_{0i}\left(\pi_{2i} - \pi_{0i}\right)$ are not too highly correlated, selection exerts negative pressure on outcomes. Thus, a positive correlation between treatment shares and outcomes is conclusive evidence of placebo effects.

B.2 Arbitrary Cutoff for Nocebo Effects

The model of health outcomes in the text assumes that placebo effects are proportional to the difference between beliefs about the efficacy of a treatment strategy and the specific effects of treatment. In other words, the cutoff that determines whether beliefs translate into positive or negative placebo effects is the specific effects of treatment. That need not be so. If the cutoff is arbitrary, then one must consider the impact of selection on the cutoff to determine whether the tests I set forth in the main text are valid. So long as selection pressured due to an increase in the share of subjects give new treatment do not push up the cutoff, however, the tests remain viable. (This implies that the cutoff is either unrelated to beliefs about the efficacy of any treatment state or that, if positively related, one of the sufficient conditions for $\partial E\left[\pi_{ki}|s\left(d\right)=T\right]/\partial d\leq 0$, for $k=0,1$, be true.) In that case an increase in d continues to depresses average observed outcomes in the absence of placebo effects. If outcomes rise with the share randomized into new treatment, it must be due to the direct effects of d on outcomes in the presence of placebo effects.

B.3 Multiplicative Model of Outcomes

The model of health outcomes in the text assumes that the specific effects of treatment and placebo effects are additive. An alternative formulation is they are multiplicative:

$$\Pr\left\{y_{ki}=\bar{y}\right\}=p_{ki}^{1-a_k}\left(\pi_{ki}^{s}\right)^{a_k}.$$

(Note that I have re-defined placebo effects to depend on the ratio of beliefs and specific effects.) Because this model is linear in logs and the log function is monoton-

ically increasing, it implies the same tests for the existence of placebo effects, albeit on a log scale.

A more general model of outcomes is

$$\Pr\{y_{ki} = \bar{y}\} = h\left(p_{ki}, \pi^{s}_{ki}\right).$$

This formulation does not affect the conclusions of chapter two so long as (1) $h_1 > 0$ and $h_2 > 0$, (2) $h_1 < 0$ and $h_2 < 0$, or (3) $h_1 = 0$. If the signs of the derivatives are not either both positive, both negative or zero for p_{ki}, one cannot determine a priori whether an increase in the share given treatment has positive or negative effects on outcomes in the absence of placebo effects. If outcomes are ambiguous with and without placebo effects, no test for placebo effects keyed to changes in d has any power. If the conditions above are satisfied, the appropriate test for placebo effects is whether the observed relationship between the share randomized into the new treatment and health outcomes has the same sign as h_2.

APPENDIX C

TESTS FOR PLACEBO EFFECTS GIVEN ALTERNATIVE TRIAL DESIGNS

C.1 Cross-over Trials

In cross-over trials, each enrollee is given both the new treatment and a placebo control. The random component is which treatment comes first. With probability d, the subject gets the new treatment in period 1 and placebo in period 2. With probability $(1-d)$, she gets placebo in period 1 and then new treatment in period 2. Thus the expected value of the trial is

$$V_i^{BT} = d\left(\pi_{1i} + \beta\pi_{0i}\right) + (1-d)\left(\pi_{0i} + \beta\pi_{1i}\right),$$

where β is the discount rate on second period utility. Observed health outcomes are

$$\begin{aligned} t &= 1 : (1-a_k)\, p_{ki} + a_k \left[d\pi_{1i} + (1-d)\,\pi_{0i}\right] \\ t &= 2 : (1-a_{1-k})\, p_{(1-k)i} + a_{1-k} \left[d\pi_{0i} + (1-d)\,\pi_{1i}\right] \end{aligned}$$

for $k = 0, 1$, where $k = 0$ (1) if the subject received placebo (new treatment) first and new treatment (placebo) second.

If the trial is such that treatment are given in rapid succession, investigators may assume the $\beta = 1$. In this case, there is no selection pressures from changes in d. The selection condition $V_i^{BT} \geq \max\left\{V_i^0, V_i^2\right\}$ collapses to $\pi_{1i} + \pi_{0i} \geq 2\max\left\{\pi_{0i}, \pi_{2i}\right\}$.

Checking to see if outcomes increase in the share given new treatment first is a valid test for the existence of placebo effects.

If $\beta \neq 1$, this test is valid, as with parallel arm trials, only if $\pi_{2i} = \bar{\pi}_2 \leq \min\{\pi_{1i}\}$ for all i or, I conjecture, if the correlation between beliefs about the efficacy of new and conventional treatment is not too high. To support my conjecture, I again turn to the example with $\pi_{ki} = \pi_{0i} + \varepsilon_{ki}$ and $E(\varepsilon_{ki}|\pi_{oi}) = E(\varepsilon_{ki})$ for $k = 1, 2$. In this case, additivity implies that individual i will choose a cross-over trial if and only if $\varepsilon_{1i} \geq \varepsilon_{2i}/\tilde{d}$ where $\tilde{d} = (1+\beta)/[d(1-\beta)+\beta]$. If $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$ are non-degenerate log-concave or log-convex random variables, with mean (μ_1, μ_2) and variance Σ, then proposition 2 applies, replacing d with $\tilde{d}$. If $\rho_{12} < \sigma_1/\sigma_2$, then checking for a positive correlation between outcomes in each treatment group of a trial and the share treated with new treatment first in the trial is a powerful test for the existence of placebo effects, albeit with a risk of type I error.

C.2 Unblinded Trials

In unblinded trials, enrollees learn their treatment state after randomization. If the enrollees feel that conventional treatment is superior to no treatment, they are free to leave the trial. Thus the expected value of the trial is

$$V_i^{UT} = d\pi_{1i} + (1-d)\max\{\pi_{0i}, \pi_{2i}\}.$$

(Henceforth, I will use UT to indicate a strategy of enrollment in an unblinded trial.) Observed outcomes are

$$\Pr\{y_{ki} = \bar{y}\} = (1-a_k)p_{ki} + a_k\pi_{ki}$$

for those in treatment state k, for $k = 0, 1$. In blinded trials, the beliefs of those in each treatment group have the same truncated distribution; it includes all individuals for whom $d\pi_{1i}+(1-d)\,\pi_{0i} \geq \max\{\pi_{0i}, \pi_{2i}\}$. In an unblinded trial, the beliefs of individuals in the different treatment groups have different truncated distributions. Both the new-treatment group and the placebo-control group satisfy the requirement that $d\pi_{1i} + (1-d)\max\{\pi_{0i}, \pi_{2i}\} \geq \max\{\pi_{0i}, \pi_{2i}\}$, i.e., $\pi_{1i} \geq \max\{\pi_{0i}, \pi_{2i}\}$. However, the placebo-control group does not contain any individuals for whom $\pi_{1i} \geq \pi_{2i} > \pi_{0i}$. (I assume the indifferent remain the trial.) Such individuals leave the placebo control group to obtain conventional treatment.

Note that selection pressures do not depend upon d. However, one cannot employ the test for placebo effects in blinded trials. Outcomes are unaffected by the probability of obtaining the new treatment because subjects know their treatment state. Fortunately there is an alternative test. Under certain conditions, if (2.1) is a correct specification of health outcomes and health outcomes in the new-treatment group of unblinded trials are greater than those in blinded trials, there must exist placebo effects.

If there is no conventional treatment, then the selection condition for unblinded trials is the same as for blinded trials. Expected outcomes in the absence of placebo effects is identical: $E(y_{1i}|\pi_{1i} \geq \pi_{0i})$. In the presence of placebo effects outcomes diverge because expectations differ in the two trials. In unblinded trials, placebo effects are driven solely by subjects' beliefs about the efficacy of their treatment state, which is revealed. In blinded trials, treatment state is never revealed. Thus, placebo effects are driven by subjects' beliefs about the efficacy of the trial. Because the trial only selects individuals who believe the new treatment is superior to no treatment, the expected value of the trial, which poses a risk of randomization into

the no-treatment state, is inferior to the expected value of new treatment in the eyes of all enrollees. Thus, placebo effects will be lower in the new-treatment arm of blinded trials than in the same arm of unblinded trials. (The same arguments also suggest that outcomes in the placebo control arm of blinded trials will be superior to outcomes in the same arm of unblinded trials.)

When there is a conventional treatment alternative, I again conjecture that checking for inferior outcomes in the new-treatment arm of blinded trials is a valid test for the existence if beliefs about new treatment are not too highly correlated with beliefs about conventional treatment. For support, consider again the simplification that $\pi_{ki} = \pi_{0i} + \varepsilon_{ki}$ and $E(\varepsilon_{ki}|\pi_{oi}) = E(\varepsilon_{ki})$ for $k = 1, 2$. The following proposition provides sufficient conditions for ρ_{12*} to ensure the test I propose can demonstrate the existence of placebo effects.

Proposition 7 *Suppose that* $(\varepsilon_{1i}, \varepsilon_{2i})$ *are distributed bivariate normal with parameters*
$(\mu_{1*}, \mu_{2*}, \sigma_{11*}, \sigma_{22*}, \rho_{12*})$. *If* $\sigma_1/\sigma_2 \in (\rho_{12*}, \rho_{12*}/d)$, *then*

$$E(\pi_{1i}|\pi_{1i} \geq \max\{\pi_{0i}, \pi_{2i}\}) > E(F_i|F_i \geq \max\{\pi_{0i}, \pi_{2i}\}),$$

where $F_i = d\pi_{1i} + (1-d)\pi_{0i}$.

This proposition differs from proposition 2 in three respects. First, I assume that $(\varepsilon_{1i}, \varepsilon_{2i})$ are jointly normal rather than that $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$ are jointly normal. The subscript asterisks serve to distinguish the parameters of $(\varepsilon_{1i}, \varepsilon_{2i})$ from the parameters of $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$. Second, this proposition places not just an upper bound, but also a lower bound on the correlation between beliefs about new treatment and conventional treatment. The upper bound is satisfied if the variance of beliefs about new treatment

is greater than the variance of beliefs about conventional treatment. Because the new treatment is, well, new, this should be a reasonable assumption. The lower bound requires that the correlation between beliefs about new and conventional treatment be greater than d times the ratio of the standard deviation in beliefs about new and about conventional treatment. This is satisfied if the variance of beliefs about the new treatment are less than ρ_{12*}^2/d^2 times the variance of beliefs about the conventional treatment. In most trials, $d = 1/2$. So this condition requires that the correlation between beliefs about new and conventional treatment be greater than $1/2$.[1] Third, proposition 2 is true even in the case where π_{ki} are confined to the interval $[0, 1]$, for all k. The proposition above may not hold under truncation. It is valid for outcome variables that are fully normal.

Unless there exist placebo effects and conditions set forth in the proposition above are satisfied, expected outcomes in the new treatment arm of an unblinded cannot exceed expected outcomes in the same arm of a blinded trial. Without placebo effects, expected new treatment outcomes in a blinded trial are greater than expected new treatment outcomes in an unblinded trial: $E\left(p_{1i}|F_i \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) \geq E\left(p_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right)$. The reason is that an unblinded trial attracts all individuals who believe that the new treatment is superior to alternatives, whereas the blinded trial attracts only those individuals who believe that new treatment is better than alternatives, even taking into account the risk of randomization into the no treatment group, which is less than or equal in value to $\max\left\{\pi_{0i}, \pi_{2i}\right\}$, the cutoff for interest in either the blinded or unblinded trial. With placebo effects, outcomes are

1. If one is examining outcomes on a $[0, 1]$ scale and is not too confident in the claim that the correlation in beliefs about new and conventional treatment are "high enough," one may want to exclude from the analysis trials with high d.

a weighted average of what I call specific effects and beliefs:

$$\begin{aligned} E\left(y_{1i}|s=BT\right) &= a_1 E\left(p_{1i}|F_i \geq \max\left\{\pi_{0i},\pi_{2i}\right\}\right) \\ &\quad + \left(1-a_1\right) E\left(F_i|F_i \geq \max\left\{\pi_{0i},\pi_{2i}\right\}\right) \\ E\left(y_{1i}|s=UT\right) &= a_1 E\left(p_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i},\pi_{2i}\right\}\right) \\ &\quad + \left(1-a_1\right) E\left(\pi_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i},\pi_{2i}\right\}\right). \end{aligned}$$

The beliefs in an unblinded trial are π_{1i} while beliefs in a blinded trial are F. Obviously, without selection, $E\left(\pi_{1i}\right) \geq E\left(F_i\right)$. However, the proposition above provides sufficient conditions to ensure that the same is true with selection.

Thus, without placebo effects, specific effects drive outcomes to be greater in a blinded trial. With placebo effects, beliefs may possibly tilt the balance back towards unblinded trials. As with the test for strictly blinded trials in the text, it is possible that there exist placebo effects but that they are not detected by checking for higher outcomes in unblinded trials. In other words, the test I propose for unblinded trials when there is a conventional treatment option is prone to type I error.

A number of studies that compare outcomes in trials with different levels of blinding find better outcomes in inadequately blinded trials. For example, Schulz et al. [30] examined 250 trials from 33 meta-analyses published by the Pregnancy and Childbirth Group of the Cochrane Database. They found that outcomes in trials with inadequate concealment (21 trials) or unclear concealment (150 trials) produced treatment effects that were 41 and 30 percent higher, respectively, than trials with adequate concealment. Trials where trial administrators were not blinded (120 trials) reported outcomes that were on average 17 percent higher than the rest. Their findings confirm earlier work by Chalmers et al. [31]. Given the analysis in this

appendix, these studies tend to support the existence of placebo effects.

APPENDIX D

OPTIMAL DESIGN OF PROPOSED EXPERIMENT

The open parameters in the proposed experiment are the number of enrollees, the number of arms, the probability of randomization into each arm, and the probability of randomization into the new treatment from each arm. These parameters should be chosen to ensure that the trial complies with the investigator's budget constraint and to minimize the variance of the resulting estimate of the short-run parameter of interest.

In computing optimal parameters for the experiment, I ignore self-selection. To factor selection into the parameters, the investigator must know the distribution of individual attributes $\mathbf{z}$ among the population of interest and the relationship between individual attributes, individual beliefs, and trial attributes. However, the investigator needs to perform the experiment to obtain information on the latter. To avoid this catch-22, I assume that enrollment from the population of interest into the experiment is random. This assumption means my design recommendations are valid only for ailments where there are no conventional treatment options or, fortuitously, the population distribution of beliefs negates the impact of selection on estimates of treatment response.[1]

1. Manning et al. [41], in their design of the Los Angeles peak-load pricing experiment address the selection problem by employing Carl Morris' Finite Selection Model [42]. (Alternatively, one can employ the simplistic approach of Conlisk and Watts, who simply adjust the costs of observations to reflect selection or attrition.) In the Manning et al. model, it was possible to guarantee enrollment by the relevant population by employing monetary incentives because they could calculate each household's prior expenditures on

I derive the optimal design of the proposed experiment in two steps. First, I assume that the main determinant of trial cost is the number of subjects. Given a fixed budget, average per-subject costs determine the total number of subjects n^* the trial can enroll and treat. Second, conditional on the total number of subjects, I choose the number of arms, the optimal probability of randomization into each arm and the probability of randomization into the new treatment from each arm so as to minimize the variance of my proposed estimate for treatment response. This approach is similar to that recommended by Conlisk and Watts [43] for experiments to estimate response surfaces.[2]

The proposed experiment design is a variant of the common two-factor experiment. The critical distinction is that in the latter the investigator is presumed to be able to select values for one factor without regard to the other. For example, in an agricultural experiment that crosses crops with pesticides, each crop can be crossed with each pesticide and vice versa. In the proposed design, however, there is a deterministic relationship between the two factors d_j and D_{ij}, namely

$$\sum_i \sum_j D_{ij} = \sum_{j=1}^{m} n_j d_j, \tag{D.1}$$

that renders infeasible certain factor combinations. This relationship constrains the selection of factor values to minimize the variance of treatment response estimates.

It is not infeasible to separately maximize the variance of each of the two factors

---

electricity and compensate them so as to make them indifferent to participation. That may not be possible in the experiment proposed in this paper becuase the new treatment may not have been approved yet and, therefore, the investigator may not be able to observe whether subjects would prefer it. In any case, the fact that a individual prefers the new treatment to alternatives does not reveal the size of the monetary incentive required to make her indifferent to enrollment.

2. See also [44].

and their interaction $D_{ij}d_j$ in one trial. However, this approach does not minimize the variance of treatment response estimates. To maximize the variance of D_{ij}, one wants exactly one-half the subject population treated. To maximize the variance of d_j or the interaction term $D_{ij}d_j$, one wants half the subjects randomized into an arm with zero probability of new treatment, and half randomized into an arm with certain administration of new treatment. The objectives are neither incompatible nor do they violate the constraint (D.1). The problem is that maximization of the variance of regressor in this fashion produces collinearity, which renders the information matrix singular and computation of coefficient estimates impossible.

An alternative approach is to employ the $\mathcal{D}$-optimality criterion. That criterion defines as optimal that set of sampling points for regressors that maximizes the determinant of the information matrix and thus minimizes the determinant of the variance-covariance matrix of parameter estimates. In mathematical terms, the criterion is $\max_{\mathbf{x}_{ij}} \left|\mathbf{X}'\mathbf{X}\right|$, where $\mathbf{X} = (\mathbf{x}_{11}, ..., \mathbf{x}_{n_m m})$ and $\mathbf{x}_{ij} = \left(D_{ij}, d_j, D_{ij}d_j\right)$.[3] One advantage of $\mathcal{D}$-efficient designs is that they often coincide with $\mathcal{A}$-efficient designs, which minimize the maximum variance of predictions[45].

Given the constraint (D.1), no analytic, closed-form solution to $\mathcal{D}$-efficient design problem is available. Computer simulation is required. Figures H.4 and H.5 provide a graphical depiction of efficient designs of the proposed experiment conditional on the total number of subjects n available to investigators and the number of arms m in a trial. The title of each bar plot gives the n and m for that plot. Each column of bar plots in a figure has the same number of subjects. The x-axis of each bar plot gives the shares randomized into new treatment in each of the m arms and the y-axis

3. I ignore the constant because the coefficients that generate an estimate of the short-run parameter of interest are those on $(D_{ij}, d_j, D_{ij}d_j)$. Thus, technically, I am employing the $\mathcal{D}_s$-optimality criterion.

gives the number of subjects randomized into each of the arms. So, for example, the first bar plot in the second column of Figure G.4 presents the optimal design of the proposed experiment when there are 40 subjects and 2 arms. The plot states that the optimal design has one arm with roughly 19 subjects and 0.44 share given new treatment and another arm with 21 subjects and 0.94 share given new treatment.

Two features of the designs in Figures H.4 and H.5 stand out. First, arms tend to have roughly equal numbers of subjects. The constraint (D.1) is met by varying the shares given new treatment across arms rather than the number of subjects assigned to each arm. Second, with smaller numbers of arms, there is a clustering of arms just below $d = 0.5$ and a sparse number of arms by $d = 1$. There are no arms near $d = 0$. The clustering just below $d = 0.5$ balances the arms near one to maximize the variance of the number of individuals given new treatment. The absence of arms near $d = 0$ can be explained by the fact that this would introduce multicollinearity between the d_j and $D_{ij}d_j$. With larger numbers of subjects and arms ($n = 160$, $m \geq 8$), arms tend to be uniformly distributed across the spectrum of d.

Figure H.6 completes the optimal design solution. It sheds light on the $\mathcal{D}$-efficient design conditional on only the total number of subjects. Each plot presents the reciprocal of the determinant of the information matrix (y-axis) from the $\mathcal{D}$-efficient design, conditional on the total number of subjects (title of each plot) and number of arms, as one varies the number of arms (x-axis). Lower values for $1/\left|X'X\right|$ are "better." So, for example, with 160 subjects, the $\mathcal{D}$-efficient design has four arms. (Figure H.5 indicates the optimal choices of $\left\{n_j, d_j\right\}$ for $j = 1, ..., 4$ for this trial.) Two features of Figure H.6 stand out. First, the relationship between the number of arms and the value of the $\mathcal{D}$-criterion is highly non-linear. So is the relationship between the optimal number of arms and the number of subjects in a trial. Second,

the optimal number of arms is small, except with very large numbers of subjects. Then the optimal number of arms is equal to the number of subjects. If I were to plot the relationship between m and $1/\left|\mathbf{X}'\mathbf{X}\right|$ for sample sizes greater than 240, the result would be similar to the plot for a sample size of 240.

APPENDIX E

SELF-SELECTION WITH CONTINUOUS OUTCOMES

In this appendix, I demonstrate that the main result of Section 4.1, which assumed that the test treatment is only available in the context of a RCT and that the cost of every treatment strategy, holds in the case of a continuous, univariate health outcome, $y \in \left[\bar{y}, \underline{y}\right] = Y$. Define $p_{ki}(y)$ as the actual distribution of outcomes for individual i given treatment k and $\pi_{ki}(y)$ as individual i's beliefs about the probability of each outcome. Let $g_{\tilde{p}}$ and $g_{\tilde{\pi}}$ give the population distribution of $\tilde{\mathbf{p}}_i = (\tilde{p}_{1i}, \tilde{p}_{2i})$ and $\tilde{\boldsymbol{\pi}}_i = (\tilde{\pi}_{1i}, \tilde{\pi}_{2i})$, where $\tilde{p}_{ki} = p_{ki} - p_{0i}$ and $\tilde{\pi}_{ki} = \pi_{ki} - \pi_{0i}$, for $k = 1, 2$. Suppose that utility is additively separable in the health outcome and that utility from this outcome, $u(y)$, is strictly increasing in outcomes and constant across individuals.

If patients' choice of strategy is dictated by expected utility maximization, then the population of interest includes all individuals who satisfy

$$\int_Y u(y)\,\tilde{\pi}_{1i}(y)\,dy \geq \max\left\{0, \int_Y u(y)\,\tilde{\pi}_{2i}(y)\,dy\right\}$$

once the test treatment is approved and available outside the trial. However, the population of enrollees includes only those for whom

$$\int_Y u(y)\,\tilde{\pi}_{1i}(y)\,dy \geq \max\left\{0, \frac{\int_Y u(y)\,\tilde{\pi}_{2i}(y)\,dy}{d}\right\}.$$

These conditions are analogous to (4.1) and (4.2).

Assume that $g_{\tilde{p}}(\tilde{\mathbf{p}}_i|y)$ and $g_{\tilde{\pi}}(\tilde{\boldsymbol{\pi}}_i|y)$ are log-concave for all $y \in Y$. Further, assume that $\int_Y y\tilde{p}_{ki}(y)\,dy = f\left(\int_Y y\tilde{\pi}_{ki}(y)\,dy\right) + \varepsilon_{ki}$, where f and ε_{ki} obey the same conditions as in the paragraph preceding Proposition 3, and that $\int_Y y\tilde{\pi}_{ki}dy$ exists.

Proposition 8 *Define* $\tilde{\rho}_{12} = corr(\tilde{p}_{1i}, \tilde{p}_{2i})$ *and* $\tilde{\sigma}_k^2 = var(\tilde{p}_{ki})$. *If* $\tilde{\rho}_{12} \leq \tilde{\sigma}_1/\tilde{\sigma}_2$, *then* $\partial E(\tilde{p}_{1i}|s(d) = BT)/\partial d \leq 0$.

APPENDIX F

PROOFS

Proposition 2. Independence of ε_{ki} and π_{0i} for $k = 1, 2$ implies that $\partial E\left[\pi_{1i} | \varepsilon_{1i} > \varepsilon_{2i}/d\right]/\partial d = \partial E\left[\varepsilon_{1i} | \varepsilon_{1i} > \varepsilon_{2i}/d\right]/\partial d$. Given $\varepsilon_{ki} > 0$, for $k = 1, 2$, $E\left[\varepsilon_{1i} | \varepsilon_{1i} > \varepsilon_{2i}/d\right] = E\left[\varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right]$. Because $\ln x$ is monotone increasing in x,

$$\operatorname{sign}\left(\frac{\partial E\left[\varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right]}{\partial d}\right) = \operatorname{sign}\left(\frac{\partial E\left[\ln \varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right]}{\partial d}\right).$$

Because W_i and V_i are independent,

$$E\left[\ln \varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right] = \mu_1 + a_1 E\left(W_i | W_i > c\left(d, \mu\right)\right),$$

where $c\left(d, \mu\right) = -\left(\mu_1 - \mu_2\right) - \ln d$.

By theorem 5 in Bagnoli and Bergstrom [46] and [what], log-concavity or log-convexity of $(\ln \varepsilon_{1i}, \ln \varepsilon_{2i})$ implies log-concavity or log-convexity, respectively, of W_i. Propositions 1 and 2 in Heckman and Honore [47] demonstrate that log-concavity or log-convexity of W_i implies $\partial E\left[W_i | W_i \geq c\right]/\partial c \geq 0$. Therefore,

$$\frac{\partial E\left[\ln \varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right]}{\partial d} = \frac{\sigma_{11} - \sigma_{12}}{\sigma} \frac{\partial E\left[W_i | W_i \geq c\right]}{\partial c} \frac{\partial c\left(d, \mu\right)}{\partial d}.$$

Because $\partial c\left(d, \mu\right)/\partial d < 0$, $\partial E\left[\ln \varepsilon_{1i} | \ln \varepsilon_{1i} - \ln \varepsilon_{2i} > -\ln d\right]/\partial d \leq 0$ so long as $\sigma_{11} > \sigma_{12}$. This condition is the same as $\rho_{12} < \sigma_1/\sigma_2$.

Proposition 3. The conditional expectations can be written

$$\begin{aligned} E\left(F_i|F_i \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) &= \mu_{0*} + E\left(d\varepsilon_{1i}|d\varepsilon_{1i} \geq \varepsilon_{2i}\right) \\ E\left(\pi_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) &= \mu_{0*} + E\left(\varepsilon_{1i}|\varepsilon_{1i} \geq \varepsilon_{2i}\right). \end{aligned}$$

Since $(\varepsilon_{1i}, \varepsilon_{2i})$ are jointly normal,

$$\begin{aligned} E\left(d\varepsilon_{1i}|d\varepsilon_{1i} \geq \varepsilon_{2i}\right) &= d\mu_{1*} - \frac{d\sigma_{12*} - d^2\sigma_{11*}}{\sigma_*}\frac{\phi\left(z\right)}{\Phi\left(z\right)} \\ E\left(\varepsilon_{1i}|\varepsilon_{1i} \geq \varepsilon_{2i}\right) &= \mu_{1*} - \frac{\sigma_{12*} - \sigma_{11*}}{\tilde{\sigma}_*}\frac{\phi\left(\tilde{z}\right)}{\Phi\left(\tilde{z}\right)}, \end{aligned}$$

where $\sigma_*^2 = \text{var}(d\varepsilon_{1i} - \varepsilon_{2i})$, $\tilde{\sigma}_*^2 = \text{var}(\varepsilon_{1i} - \varepsilon_{2i})$, $z = \left(d\mu_1 - \mu_2\right)/\sigma_*$, and $\tilde{z} = \left(\mu_1 - \mu_2\right)/\tilde{\sigma}_*$. Because $\sigma_{1*}/\sigma_{2*} \in (\rho_{12*}, \rho_{12*}/d)$ implies $\sigma_{11*} > \sigma_{12*} > d\sigma_{11*}$,

$$E\left(d\varepsilon_{1i}|d\varepsilon_{1i} \geq \varepsilon_{2i}\right) < d\mu_1 \leq \mu_1 < E\left(\varepsilon_{1i}|\varepsilon_{1i} \geq \varepsilon_{2i}\right).$$

Thus, $E\left(\pi_{1i}|\pi_{1i} \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right) > E\left(F_i|F_i \geq \max\left\{\pi_{0i}, \pi_{2i}\right\}\right)$.

Proposition 4. An individual will choose the strategy $s = BT$ if and only if $\tilde{\pi}_{1i} \geq \max\left\{0, \tilde{\pi}_{2i}/d\right\}$. Because

$$\begin{aligned} E\left(\tilde{p}_{1i}|s(d) = BT\right) &= \Pr\left\{\tilde{\pi}_{2i} \leq 0\right\} E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq 0\right) \\ &\quad + \Pr\left\{\tilde{\pi}_{2i} \geq 0\right\} E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d\right), \end{aligned}$$

$\text{sign}(\partial E\left(\tilde{p}_{1i}|s(d) = BT\right)/\partial d) = \text{sign}(\partial E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d\right)/\partial d)$. Because $\varepsilon_{ki} \perp \pi_{k'i'}$ for all $\left(k', i'\right)$, $E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d\right) = E(f\left(\tilde{\pi}_{1i}\right)|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d) + E\left(\varepsilon_{ki}\right)$. Because

$f\prime > 0$,

$$sign(\partial E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d\right)/\partial d) = sign\left(\partial E\left(\tilde{\pi}_{1i}|\tilde{\pi}_{1i} \geq \tilde{\pi}_{2i}/d\right)/\partial d\right).$$

Since $g_{\tilde{\pi}}$ is log concave, $\tilde{\pi}_{ki}$ is a log-concave random variable, by Theorem 2 of [48]. The remainder of this proof is analogous to that of Proposition 1.

Proposition 5.

$$E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d\right) = E(f\left(\tilde{\pi}_{1i}\right)|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d) + E\left(\varepsilon_{ki}|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d\right).$$

Because f is strictly increasing,

$$sign\left(\partial E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d\right)/\partial d\right) = sign\left(\partial E\left(f\left(\tilde{\pi}_{1i}\right)|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d\right)/\partial d\right).$$

Because $\tilde{w}^{BT}$ is not a function of d, the cutoff $\tilde{w}^{BT}/d$ falls in d.

Proposition 6.

$$\begin{aligned} E\left(\tilde{p}_{1i}|s\left(d\right) = BT\right) &= \Pr\left\{\tilde{\pi}_{2i} - \tilde{w}^2 \leq 0\right\} E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d\right) \\ &\quad + \Pr\left\{\tilde{\pi}_{2i} - \tilde{w}^2 > 0\right\} E\left(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \left(\tilde{\pi}_{2i} + \tilde{w}^{BT} - \tilde{w}^2\right)/d\right). \end{aligned}$$

By Proposition 2, $\partial E(\tilde{p}_{1i}|\tilde{\pi}_{1i} \geq \tilde{w}^{BT}/d)/\partial d \leq 0$. Define $\tilde{\pi}^*_{2i} = \tilde{\pi}_{2i} + \tilde{w}^{BT} - \tilde{w}^s$. Since $\tilde{w}^{BT} - \tilde{w}^s$ is a constant and $\tilde{\pi}_i$ is log-concave, so is $\left(\tilde{\pi}_{1i}, \tilde{\pi}^*_{2i}\right)$. Given that $\tilde{w}^{BT} - \tilde{w}^2$ is not a function of d, the remainder of the proof is analogous to that for Proposition 3.

Proposition 7. The proof is analogous to that of Proposition 5.

Proposition 8. Suppressing the arguments of $u(y)$, $\tilde{p}_{ki}(y)$ and $\tilde{\pi}_{ki}(y)$ and the range of integration Y,

$$E_g\left(\int y\tilde{p}_{1i}dy|s(d)=BT\right)=\Pr\left\{\int u\tilde{\pi}_{2i}dy\leq 0\right\}$$
$$E_g\left(\int y\tilde{p}_{1i}dy|\int u\tilde{\pi}_{1i}dy\geq 0\right)+\Pr\left\{\int u\tilde{\pi}_{2i}dy>0\right\}$$
$$E_g\left(\int y\tilde{p}_{1i}dy|\int u\tilde{\pi}_{1i}dy\geq\int u\tilde{\pi}_{2i}dy/d\right).$$

Thus,

$$sign\left(\partial E_g(\int y\tilde{p}_{1i}dy|s(d)=BT)/\partial d\right)=$$
$$sign\left(\partial E_g\left(\int y\tilde{p}_{1i}dy|\int u\tilde{\pi}_{1i}dy\geq\int u\tilde{\pi}_{2i}dy/d\right)/\partial d\right).$$

This is equal to

$$sign\left(\partial E_g(\int y\tilde{p}_{1i}dy|\int y\tilde{\pi}_{1i}dy\geq\int y\tilde{\pi}_{2i}dy/d)/\partial d\right) \tag{F.1}$$

because u is strictly increasing. (F.1) is equal to

$$sign\left(\partial E_g(\int y\tilde{\pi}_{1i}dy|\int y\tilde{\pi}_{1i}dy\geq\int y\tilde{\pi}_{2i}dy/d)/\partial d\right) \tag{F.2}$$

because $f'>0$ and $\varepsilon_{ki}\perp\pi_{k'i'}$.

By [48] Theorem 2, because $g_{\tilde{\pi}}$ is log concave, for any given $y\in Y$ and $k=1$ or 2, $\tilde{\pi}_{ki}(y)$ is distributed log-concave in the population. Pick any n members of Y and index them by j. Because the convolution of two log-concave random variables is also a log-concave random variable [18], $\sum_{j=1}^{n}\left(y_j/n\right)\tilde{\pi}_{ki}\left(y_j\right)$ is a log-concave random

variable, for $k = 1, 2$. Because $\int_Y y\tilde{\pi}_{ki}dy$ is the limit of this sum as $n \to \infty$ and (by assumption) exists, $\int_Y y\tilde{\pi}_{ki}dy$ is log-concave, for $k = 1, 2$. By the logic of Proposition 3, then, (F.2) is negative.

APPENDIX G

TABLES

Source	Findings
Skovlund (1991)[49]	Noted that patients in the treatment group of a trial which compared paracetamol to naproxen (conventional control) [50] responded better than patients in the treatment group of a trial which compared paracetamol to placebo [51] . Likewise, patients in the control group of the second trial responded better than patients in the control group of the first trial. Skovlund suggests that the patients in the naproxen-control trial may have responded better because they knew that, no matter what, they would get an active drug, whereas the patients in the placebo-control trial were skeptical because they knew they might receive a placebo. Because a better control group option improves expected outcomes, these trials suggest the importance of beliefs about the efficacy of the control-treatment to outcomes. Because the effect was recorded in both treatment and control groups, the trials underscore the relevance of beliefs about treatment state to outcomes in both states.
Pollo et al. (2001)[11]	Studied the placebo effect among thoracotomized patients, i.e., patients with their chests surgically opened for, e.g., heart operation). They divided patients into three groups and gave each group the pain killer buprenorphine on demand as well as an intravenous infusion of saline solution. The first group (natural history) was told nothing about the saline drip. The second group (double blind) was told the saline drip might be a powerful painkiller or it might be a placebo. The third group (deception) was told the saline drip was a potent painkiller. The authors found that the natural-history group asked for buprenorphine more often than the double-blind group, which asked for buprenorphine more often than the deception group. This study suggests that increasing expectations about the probability of being treated with a powerful painkiller (the saline solution) produced better health outcomes as measured by the level of demand for a second pain killer (buprenorphine).
Marlatt and Rohsenow (1980)[8]	Examined the impact of alcohol on cognition employing a balanced-placebo design wherein subjects are first randomized across treatment states and then across instructions about treatment states, with one group being told that they were given active treatment and the other being told they were given placebo. They find superior outcomes among those told they were administered active treatment, controlling for actual treatment state. This suggests that beliefs about treatment state impact outcomes.

Table G.1: Studies supporting the expectancy theory of placebo effects

	Treatment State	Beliefs	Health Outcome
Positive Placebo Effect	$D_i = 1$	$\pi_i^{BT} > p_{1i}$	$y_{1i} > p_{1i}$
	$D_i = 0$	$\pi_i^{BT} > p_{0i}$	$y_{0i} > p_{0i}$
Negative Placebo Effect	$D_i = 1$	$\pi_i^{BT} < p_{1i}$	$y_{1i} < p_{1i}$
	$D_i = 0$	$\pi_i^{BT} < p_{0i}$	$y_{0i} < p_{0i}$

Table G.2: Definition of placebo effects in a blinded trial.

	Trials of H2-Blockers			Prostaglandins			PPIs		
	Obs.	Mean	SD	Obs.	Mean	SD	Obs.	Mean	SD
Treatment arm									
Date trial results published	225	1987	4.62	46	1988	3.62	97	1991	3.97
Share of subjects not given control	225	0.88	0.20	46	0.56	0.08	97	0.61	0.16
Share of arms in placebo-control trials	225	0.22	0.41	46	0.47	0.50	97	0.08	0.27
Share in antacid-control trials	225	0.06	0.23	46	0.00	0.00	97	0.00	0.00
Share in lower-class drug control trials	225	0.01	0.09	46	0.53	0.50	97	0.84	0.36
Share same-class drug control trials	225	0.72	0.45	46	0.00	0.00	97	0.08	0.27
Average number of measurements	225	1.42	0.58	46	1.34	0.52	97	1.60	0.65
Number enrolled	225	198	147	46	123	110	97	110	40
Number evaluated (per protocol)	225	181	133	46	108	96	97	102	37
Share of subjs. not healed (method 1)	225	0.22	0.17	46	0.36	0.18	97	0.22	0.18
Share not healed (meth. 2)	225	0.29	0.17	46	0.42	0.17	97	0.28	0.17
Share not healed (meth. 3)	225	0.21	0.16	46	0.32	0.17	97	0.21	0.18
- ln(share not healed)/t (meth. 1)	213	0.048	0.01	46	0.04	0.02	82	0.087	0.03
Treatment response of subjs. (meth. 1)	225	0.07	0.15	46	-0.09	0.28	97	-0.13	0.14
- (ln(S1) - ln(S0))/t (meth. 1)	212	0.007	0.01	45	0.00	0.02	82	0.03	0.03
Antacids permitted in trial (1-5)?	225	3.66	0.64	46	3.49	0.61	97	3.10	1.15
Frequency of dosage (times/day)	225	1.74	1.00	46	2.75	0.98	97	1.19	0.66
Total daily dosage (mg)	225	0.44	0.36	46	0.09	0.23	97	0.03	0.02
Share male	212	0.72	0.07	41	0.73	0.09	95	0.70	0.09
Share that smoke	181	0.57	0.09	43	0.54	0.13	91	0.49	0.11
Average age (years) of subjs.	207	46	4	41	44	5	87	46	5

Table G.3: Summary statistics for data from ulcer trials.

Table G.3, continued.

	Trials of H2-Blockers			Prostaglandins			PPIs		
	Obs.	Mean	SD	Obs.	Mean	SD	Obs.	Mean	SD
Control arm									
Date trial results published	86	1986	5.04	40	1989	3.79	67	1991	4.26
Share of subjects not given control	86	0.56	0.11	40	0.53	0.07	67	0.53	0.08
Share of arms in placebo-control trials	86	0.75	0.43	40	0.39	0.49	67	0.04	0.19
Share in antacid-control trials	86	0.32	0.47	40	0.00	0.00	67	0.00	0.00
Share in lower-class drug control trials	86	0.00	0.00	40	0.61	0.49	67	0.96	0.19
Share same-class drug control trials	86	0.00	0.00	40	0.00	0.00	67	0.00	0.00
Average number of measurements	86	1.29	0.53	40	1.35	0.53	67	1.60	0.64
Number enrolled	86	89	85	40	130	116	67	113	39
Number evaluated (per protocol)	86	79	76	40	117	105	67	104	38
Share of subjs. not healed (method 1)	86	0.56	0.23	39	0.36	0.28	67	0.36	0.22
Share not healed (meth. 2)	86	0.60	0.20	39	0.41	0.26	67	0.42	0.20
Share not healed (meth. 3)	86	0.50	0.21	39	0.32	0.25	67	0.33	0.21
- ln(share not healed)/t (meth. 1)	85	0.03	0.02	38	0.04	0.02	66	0.05	0.02
Antacids permitted in trial (1-5)?	86	3.31	1.04	40	3.52	0.60	67	3.16	1.15
Share male	77	0.76	0.10	35	0.77	0.09	65	0.70	0.09
Share that smoke	59	0.63	0.12	37	0.52	0.11	61	0.49	0.14
Average age (years) of subjs.	68	45	5	35	44	5	59	46	6

Notes. Each observation represents a measurement on the indicated arm of indicated trial. Each trial may have more than test-treatment arm and more than one measurement on each arm. Therefore, there are more observations than trials. However, each trial has only one control arm. While all trials are controlled, not all are placebo controlled. Means and standard deviations are calculated weighting each arm in proportion to the number of subjects evaluated per protocol, regardless of the number of measurements on the arm. Frequency of medication and total dosage are not provided for control arms because such variables are meaningless for placebo arms. Share of subjects not given control is very high for H_2-blocker trials because same-class control trials are assumed to have a share treated of one. There is assumed to be no control arm in such trials. The antacid permitted variable takes a value of 1 if antacids were prohibited, 2 if discouraged, 3 if permitted or not discussed, 4 if antacids were provided, and 5 if antacids were required.

Arm	Treatment				Control			
Specification	1	2	3	4†	1	2	3†	4†
Constant	0.038 ***	0.026 ***	0.085	0.249 ***	0.034 ***	0.043 ***	-2.31 ***	-2.4 ***
Std. err.	0.001	0.012	0.096	0.108	0.0005	0.01	0.62	0.535
Share	0.012 ***	0.028 ***	0.004	-0.14	-0.015 ***	-0.03 **	4.304 ***	4.389 ***
treated (d)	0.001	0.013	0.145	0.137	0.001	0.015	1.227	1.052
Antacid role (1-5)		0.002	-0.01 ***			-0	-0.01 ***	
		0.002	0.002			0.003	0.005	
Daily freq. of treatment		-0.007 ***	0.003					
of treatment		0.002	0.003					
Total daily dosage (mg)		0.002	0.097 ***					
		0.036	0.042					
Cimetidine dosage (mg)		0.025	-0.06 *					
		0.032	0.036					
Ranitidine dosage (mg)		0.097 ***	0.082 ***					
		0.023	0.029					
Antacid role * d		-0.004 ***	0.009 ***			0.004	0.02 ***	
		0.002	0.003			0.005	0.009	
Freq. * d		0.007 ***	-0					
		0.003	0.004					
Dosage * d		0.009	-0.1 ***					
		0.037	0.043					
Cimetidine dosage * d		-0.031	0.056 *					
		0.033	0.036					
Ranitidine dosage * d		-0.093 ***	-0.08 ***					
		0.024	0.032					
Share male			-0.08 ***	-0.1 ***			-0.12	-0.19
			0.033	0.026			0.177	0.175
Share smokers			0.015	0.021			0.292 ***	0.325 ***
			0.016	0.017			0.125	0.123
Age (mean yrs)			-0.01	-0.04 *			0.607 ***	0.625 ***
			0.024	0.026			0.141	0.121
Male * d			0.069 *	0.097 ***			0.181	0.301
			0.047	0.034			0.349	0.339
Smoker * d			0.012	0.001			-0.54 ***	-0.62 ***
			0.025	0.023			0.246	0.244
Age * d			-0.01	0.021			-1.11 ***	-1.12 ***
			0.037	0.035			0.276	0.237

Table G.4: Evidence of placebo effects from H2-blocker trials with informed consent employing x-captures-selection approach to self-selection.

Table G.4, continued.

Arm	Treatment				Control			
Specification	1	2	3	4†	1	2	3†	4†
F-tests of joint significance of coefficients on								
d	406917 ***	4.88 **	0.00	1.04	187.86 ***	3.63 *	12.30 ***	17.40 ***
P-value	0.00	0.03	0.98	0.31	0.00	0.06	0.00	0.00
Trial variables * d		83.06 ***	44.32 ***			0.52	4.88 **	
		0.00	0.00			0.47	0.03	
Subject variables * d			2.25	10.41 **			27.71 ***	36.92 ***
			0.52	0.02			0.00	0.00
Trial and subject vars. * d			45.65 ***				28.42 ***	
			0.00				0.00	
d and all interactions		85.79 ***	59.03 ***	20.27 ***		11.09 ***	64.76 ***	65.25 ***
		0.00	0.00	0.00		0.00	0.00	0.00
Studies	58	58	37	37	58	58	37	37
Obs./Meas.	213	213	163	163	85	85	55	55
Arms	133	133	98	98	58	58	37	37
Measurements manifesting placebo effects (by confidence level)								
At all	213	183	130	151	0	0	12	17
At 85%	213	153	107	127	0	0	7	12
At 90%	213	153	103	122	0	0	4	9
At 95%	213	124	87	114	0	0	4	8

Notes. Estimation was by feasible GLS. The dependent variable is the –ln(S(t))/t, where S(t) is the fraction that remain ill after t days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Standard errors are reported below coefficients. Coefficients significant at the 1%/5%/10% level marked with ***/**/*. P-values for F-tests for the joint significance of indicated groups of variables are reported below the F-test statistics. Specifications that fail Pregibon's link test are marked with a dagger.

Treatment	H2-Blocker				Prostagladin			
Specification	1	2	3	4	1	2	3	4
Approach 1			†	†				
Obs./meas.	174	174	133	133				
Studies	38	38	22	22				
Obs. manifesting placebo effects (by confidence level)								
At all	174	83	40	120				
At 85%	174	59	6	68				
At 90%	174	58	5	43				
At 95%	174	56	1	37				
Approach 2				†			†	
Obs./meas.	213	213	163	163	46	46	41	41
Studies	58	58	37	37	21	21	17	17
Obs. manifesting placebo effects (by confidence level)								
At all	213	183	130	151	0	3	2	13
At 85%	213	153	107	127	0	0	0	8
At 90%	213	153	103	122	0	0	0	8
At 95%	213	124	87	114	0	0	0	2
Approach 3		†						
Obs./meas.	213	213	163	163	46	46	41	41
Studies	58	58	37	37	21	21	17	17
d^2	0.14 ***	0.15 ***	0.15 ***	0.18 ***	1.29 **	-1.86 ***	-3.14	0.57
	0.01	0.02	0.03	0.02	0.78	0.80	2.40	0.99

Notes. Estimation was by feasible GLS. The dependent variable is the –ln(S(t))/t, where S(t) is the fraction that remain ill after t days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Specification (1) includes a constant and d; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with d; (3) adds subject-level variables (sex, smoker, and age) and their interactions with d; and (4) removes the trial-level variables and interactions from (3). When I employ the instrument approach to test for placebo effects, I add d^2 to each specification. Coefficients significant at the 1%/5%/10% level marked with ***/**/*. Specification that fail Pregibon's link test are marked with a dagger.

Table G.5: Evidence of placebo effects in treatment arms of ulcer trials with informed consent, by drug type, specification, and approach to self-selection.

Table G.5, continued.

Treatment	PPI			
Specification	1	2	3	4
Approach 1				
Obs./meas.				
Studies				
Obs. manifesting placebo effects (by confidence level)				
At all				
At 85%				
At 90%				
At 95%				
Approach 2				
Obs./meas.	82	82	70	70
Studies	29	29	24	24
Obs. manifesting placebo effects (by confidence level)				
At all	0	26	27	28
At 85%	0	16	21	25
At 90%	0	16	21	25
At 95%	0	16	21	25
Approach 3				
Obs./meas.	82	82	70	70
Studies	29	29	24	24
d^2	0.00	-0.39 ***	-0.03	-0.04
	0.09	0.11	0.18	0.08

Treatment	H2-Blocker				Prostagladin			
Specification	1	2	3	4	1	2	3	4
Approach 1								
Obs./meas.	51	51	30	30				
Studies	38	38	22	22				
Obs. manifesting placebo effects (by confidence level)								
At all	0	0	22	21				
At 85%	0	0	4	5				
At 90%	0	0	3	2				
At 95%	0	0	3	2				
Approach 2			†	†				
Obs./meas.	85	85	55	55	38	38	33	33
Studies	58	58	37	37	21	21	17	17
Obs. manifesting placebo effects (by confidence level)								
At all	0	0	12	17	38	8	25	25
At 85%	0	0	7	12	38	8	9	21
At 90%	0	0	4	9	38	8	6	21
At 95%	0	0	4	8	38	8	3	19
Approach 3						†	†	
Obs./meas.	85	85	55	55	38	38	33	33
Studies	58	58	37	37	21	21	17	17
d^2	0.06	0.11 **	-0.56 ***	-0.63 ***	-1.91 ***	-4.65 ***		-0.57
	0.05	0.07	0.09	0.10	0.67	0.88		0.49

Notes. Estimation was by feasible GLS. The dependent variable is the –ln(S(t))/t, where S(t) is the fraction that remain ill after t days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Specification (1) includes a constant and d; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with d; (3) adds subject-level variables (sex, smoker, and age) and their interactions with d; and (4) removes the trial-level variables and interactions from (3). When I employ the instrument approach to test for placebo effects, I add d^2 to each specification. Coefficients significant at the 1%/5%/10% level marked with ***/**/*. Specification that fail Pregibon's link test are marked with a dagger.

Table G.6: Evidence of placebo effects in control arms of ulcer trials with informed consent, by drug type, specification, and approach to self-selection.

Table G.6, continued.

Treatment	PPI			
Specification	1	2	3	4
Approach 1				
Obs./meas.				
Studies				
Obs. manifesting placebo effects (by confidence level)				
At all				
At 85%				
At 90%				
At 95%				
Approach 2				
Obs./meas.	66	66	56	56
Studies	29	29	24	24
Obs. manifesting placebo effects (by confidence level)				
At all	66	24	24	27
At 85%	0	10	20	22
At 90%	0	10	20	22
At 95%	0	10	18	18
Approach 3				
Obs./meas.	66	66	56	56
Studies	29	29	24	24
d^2	-0.81 ***	-0.37 **	0.62	0.55
	0.18	0.20	1.01	1.00

	Strength of Evidence of Placebo Effects in	
Class of Drug Tested in Trial	Treatment Arm	Control Arm
H_2-Blockers	Strong	Weak
Prostaglandins	None	Moderate
Proton-Pump Inihibitors	Moderate	Moderate

Table G.7: Summary of results.

Treatment	H2-Blocker			Prostaglandin		
Specification	1†	2†	3†	1	2	3
Constant	0.051 ***	0.049 ***	0.149 ***	-0.009	-0.093 ***	0.203
	0.001	0.002	0.013	0.013	0.022	0.160
Share treated	-0.050 ***	-0.044 ***	-0.048 ***	0.021	-0.071 *	-0.072
	0.001	0.001	0.003	0.023	0.049	0.064
Antacid role (1-5)					0.025 ***	0.025 ***
					0.003	0.003
Daily freq. of treatment					0.012 ***	0.009 **
					0.003	0.005
Dosage 1 (mg)		0.000	-0.003		-2.040	50.186
		0.004	0.007		25.504	51.841
Dosage 2 (mg)		-0.005	-0.001		130.499 ***	131.142
		0.004	0.006		33.651	132.116
Dosage 3 (mg)		-0.012 ***	-0.006 *		0.040 ***	0.057 ***
		0.004	0.004		0.012	0.019
Dosage 4 (mg)					0.072 ***	15.979
					0.007	27.363
Treatment group						
Share male			0.037 ***			-0.027
			0.014			0.057
Share smokers			-0.014			0.050
			0.013			0.044
Age (mean yrs)			0.014			-0.041
			0.010			0.055
Control group						
Share male			0.002			-0.119 ***
			0.014			0.040
Share smokers			0.035 ***			-0.062
			0.012			0.057
Age (mean yrs)			-0.048 ***			-0.005
			0.010			0.067
Obs.	212	212	162	45	45	40
Arms	133	133	98	28	28	24

Notes. Estimation was by feasible GLS. Each observation is a measurement on a treatment arm of the trial. Some trials provide multiple measurements on each arm. I weighted the observations such that each treatment arm is weighted in proportion to the number of subjects in the arm, regardless of the number of measurements on the arm. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is $-[\ln(S_1(t)) - \ln(S_0(t))]/t$, where $S_1(t)$ ($S_0(t)$) is the fraction of subjects in the treatment (control) group that remain ill after t days. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. The first dosage variable gives total daily dosage for the test treatment in H_2-blocker and PPI trials. The remaining give the dosage for the dominant drugs in each class. For prostaglandin trials, all dosage variables give the dosage for specific, dominant drugs in the class.

Table G.8: Evidence of self-selection in ulcer trials with informed consent, by drug type and specification.

Table G.8, continued.

Treatment	PPI		
Specification	1†	2†	3
Constant	0.077 ***	0.060 ***	0.286 ***
	0.002	0.011	0.088
Share treated	-0.077 ***	-0.071 ***	-0.029 *
	0.003	0.005	0.019
Antacid role (1-5)		0.000	0.000
		0.002	0.003
Daily freq. of treatment		-0.011	-0.075 ***
		0.011	0.014
Dosage 1 (mg)		1.177 ***	1.644 ***
		0.182	0.280
Dosage 2 (mg)		-0.565 ***	-0.200
		0.272	0.485
Dosage 3 (mg)		-0.049	0.695 ***
		0.196	0.286
Dosage 4 (mg)			
Treatment group			
Share male			0.020
			0.036
Share smokers			-0.091 ***
			0.043
Age (mean yrs)			-0.071
			0.066
Control group			
Share male			-0.039
			0.032
Share smokers			0.022
			0.033
Age (mean yrs)			0.028
			0.065
Obs.	82	82	70
Arms	41	41	35

Notes (continued). Subject-oriented covariates (male, smoker, age) for control groups are based upon the associated control arm for each treatment arm. Because trials often have multiple treatment arms, the control group covariates for those arms will be identical. Coefficients significant at the 1%/5%/10% level are marked with a ***/**/*. Specifications that fail Pregibon's link test are marked with a dagger. A negative coefficient on the share treated variable is evidence of self-selection.

Treat-	Sample	Trials with informed consent						Trials w/o informed consent				
ment	Specification	1	2	3	4	5	6	1	2	4	5	6
H2-	Power of d	5	4	3	5	3	3	5	4	5	3	3
Blockers	Obs.	212	212	162	162	162	162	51	51	26	26	26
	Arms	133	133	98	98	98	98	41	41	22	22	22
	Link test	1.83	0.18	2.76	1.73	4.91	4.87	1.99	0.60	1.01	-0.32	-0.85
	Measurements manifesting placebo effects (by confidence level)											
	At all	212	212	162	162	162	162	0	0	0	0	0
	At 85%	212	212	162	162	162	162	0	0	0	0	0
	At 90%	212	212	162	162	162	162	0	0	0	0	0
	At 95%	212	212	162	162	162	162	0	0	0	0	0
Prosta-	Power of d	1	1	2	3	2	1	1	1		1	1
glandins	Obs.	45	45	40	40	40	40	19	19		15	15
	Arms	28	28	24	24	24	24	13	13		9	9
	Link test	1.17	1.34	1.58	3.62	-1.31	0.90	1.65	-2.73		-0.05	-0.11
	Measurements manifesting placebo effects (by confidence level)											
	At all	0	45	40	40	40	0	0	19		15	15
	At 85%	0	45	0	40	40	0	0	19		15	0
	At 90%	0	45	0	0	40	0	0	19		15	0
	At 95%	0	0	0	0	40	0	0	19		15	0

Table G.9: Evidence of self-selection in ulcer trials, by drug type, specification, and informed consent status.

Table G.9, continued.

Treat-	Sample	Trials with informed consent						Trials w/o informed consent				
ment	Specification	1	2	3	4	5	6	1	2	4	5	6
Proton	Power of d	2	4	1	1	3	1	2	4		3	1
Pump	Obs.	82	82	70	70	70	70	36	36		28	28
Inhibitors	Arms	41	41	35	35	35	35	16	16		12	12
	Link test	0.52	-2.97	0.38	-0.13	-4.76	0.76	-0.83	-0.45		-0.02	0.37
	Measurements manifesting placebo effects (by confidence level)											
	At all	17	82	70	70	0	70	0	0		0	0
	At 85%	0	82	70	70	0	70	0	0		0	0
	At 90%	0	0	0	70	0	70	0	0		0	0
	At 95%	0	0	0	70	0	70	0	0		0	0

Notes. Estimation was by feasible GLS. Each observation is a measurement on a treatment arm of the trial. Some trials provide multiple measurements on each arm. I weighted the observations such that each treatment arm is weighted in proportion to the number of subjects in the arm, regardless of the number of measurements on the arm. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is $-[\ln(S_1(t)) - \ln(S_0(t))]/t$, where $S_1(t)$ ($S_0(t)$) is the fraction of subjects in the treatment (control) group that remain ill after t days. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Subject-oriented covariates (male, smoker, age) for the control group is based upon the associated control arm for each treatment arm. Because trials often have multiple treatment arms, the control group covariates for those arms will be identical. Specification (1) includes a constant and the specified powers of the share treated (d); (2) includes (1) and trial level-variables (antacid role, daily frequency of medication, total daily dosage); (3) includes (2) plus subject-oriented covariates (male, smoker, age) for treatment and control arms separately; (4) includes (2) plus subject-oriented covariates averaged over the treatment and control groups, with each group assigned a weight in proportion to the number evaluated per protocol in that measurement. (5) and (6) are the same as (3) and (4) but without the trial-level variables added by (2). The link test regresses the dependent variables against predicted values and such values squared. The value reported for the link test is the t-statistic for the coefficient on the predicted values squared. Statistical significance indicates misspecification. I selected the power of d by determining the lowest power that enabled the specification to pass the link test. Where this was not possible with less than five powers, I employed the power that yielded the lowest value of the link test statistic.

	0 works, 1 works	0 works, 1 fails	0 fails, 1 works	0 fails, 1 fails
Lottery yields 0	$\frac{\pi_0\pi_1}{2}$ $s(0)=\bar{y}$ $s(1)=\bar{y}$ $s(UT)=\bar{y}$	$\frac{\pi_0(1-\pi_1)}{2}$ $s(0)=\bar{y}$ $s(1)=\underline{y}$ $s(UT)=\bar{y}$	$\frac{(1-\pi_0)\pi_1}{2}$ $s(0)=\underline{y}$ $s(1)=\bar{y}$ $s(UT)=\underline{y}$	$\frac{(1-\pi_0)(1-\pi_1)}{2}$ $s(0)=\underline{y}$ $s(1)=\underline{y}$ $s(UT)=\underline{y}$
Lottery yields 1	$\frac{\pi_0\pi_1}{2}$ $s(0)=\bar{y}$ $s(1)=\bar{y}$ $s(UT)=\bar{y}$	$\frac{\pi_0(1-\pi_1)}{2}$ $s(0)=\bar{y}$ $s(1)=\underline{y}$ $s(UT)=\underline{y}$	$\frac{(1-\pi_0)\pi_1}{2}$ $s(0)=\underline{y}$ $s(1)=\bar{y}$ $s(UT)=\bar{y}$	$\frac{(1-\pi_0)(1-\pi_1)}{2}$ $s(0)=\underline{y}$ $s(1)=\underline{y}$ $s(UT)=\underline{y}$

Table G.10: Subjective probabilities assigned to different states and consquences of actions in these states (case 1).

	0 works, 1 works	0 works, 1 fails	0 fails, 1 works	0 fails, 1 fails
Lottery yields 0	$\frac{\phi_0\phi_1}{2}$	$\frac{\phi_0(1-\phi_1)}{2}$	$\frac{(1-\phi_0)\phi_1}{2}$	$\frac{(1-\phi_0)(1-\phi_1)}{2}$
Lottery yields 1	$\frac{\phi_0\phi_1}{2}$	$\frac{\phi_0(1-\phi_1)}{2}$	$\frac{(1-\phi_0)\phi_1}{2}$	$\frac{(1-\phi_0)(1-\phi_1)}{2}$

Table G.11: Subjective probabilities assigned to different states (case 3).

	0 works, 1 works, $\delta_0 = \delta_1 = \frac{1}{2}$ (π_0, π_1)	0 works, 1 fails $\delta_0 = \delta_1 = \frac{1}{2}$ (π_0, π_1)	0 fails, 1 works $\delta_0 = \delta_1 = \frac{1}{2}$ (π_0, π_1)	0 fails, 1 fails $\delta_0 = \delta_1 = \frac{1}{2}$ (π_0, π_1)
Unblinded lottery yields 0, blinded yields 0	$\frac{\phi_0\phi_1\phi_0^{BT}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0\phi_1'\phi_0^{BT}\phi_1^{BT\prime}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0'\phi_1\phi_0^{BT\prime}\phi_1^{BT}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$	$\frac{\phi_0'\phi_1'\phi_0^{BT\prime}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$
Unblinded lottery yields 0, blinded yields 1	$\frac{\phi_0\phi_1\phi_0^{BT}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0\phi_1'\phi_0^{BT}\phi_1^{BT\prime}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \underline{y}$	$\frac{\phi_0'\phi_1\phi_0^{BT\prime}\phi_1^{BT}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0'\phi_1'\phi_0^{BT\prime}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$
Unblinded lottery yields 1, blinded yields 0	$\frac{\phi_0\phi_1\phi_0^{BT}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0\phi_1'\phi_0^{BT}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0'\phi_1\phi_0^{BT\prime}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \underline{y}$	$\frac{\phi_0'\phi_1'\phi_0^{BT\prime}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$
Unblinded lottery yields 1, blinded yields 1	$\frac{\phi_0\phi_1\phi_0^{BT}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0\phi_1'\phi_0^{BT}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$	$\frac{\phi_0'\phi_1\phi_0^{BT\prime}\phi_1^{BT}}{4}$ $s(UT) = \bar{y}$ $s(BT) = \bar{y}$	$\frac{\phi_0'\phi_1'\phi_0^{BT\prime}\phi_1^{BT\prime}}{4}$ $s(UT) = \underline{y}$ $s(BT) = \underline{y}$

Table G.12: Subjective probabilities assigned to different (non-zero probability) states and consquences of actions in these states (case $3'$).

APPENDIX H

FIGURES

Figure H.1: *Illustration of placebo-effect and blinding bias.*

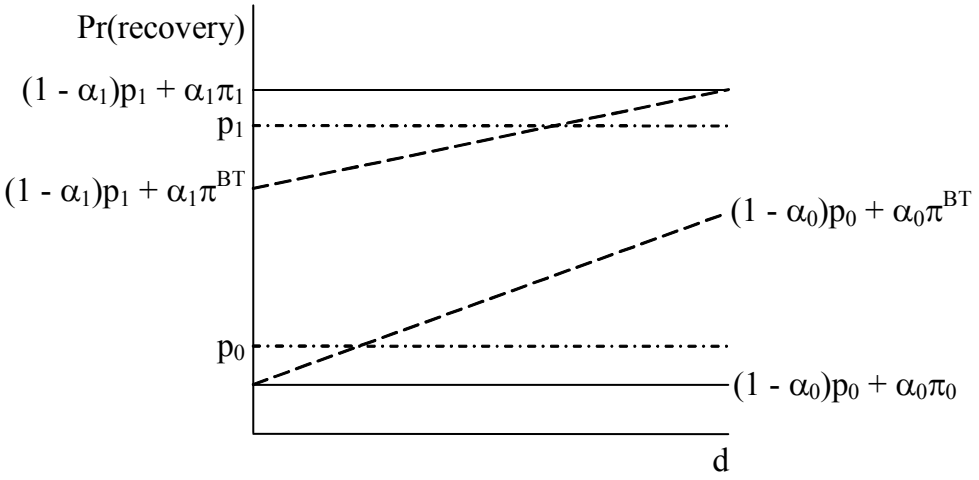


Figure H.2: *Illustration of selection bias.*

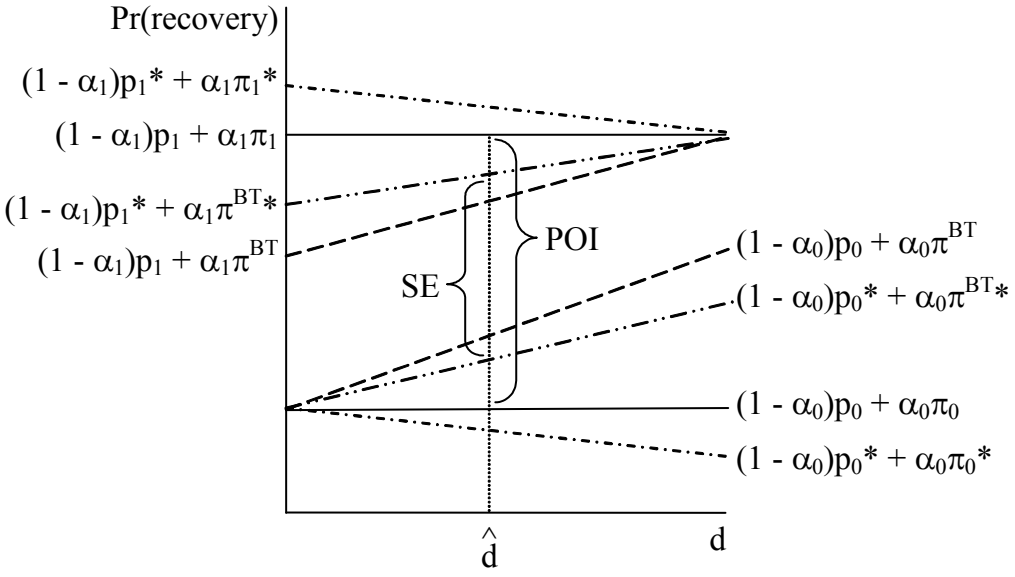


Figure H.3: *Example of proposed design.*

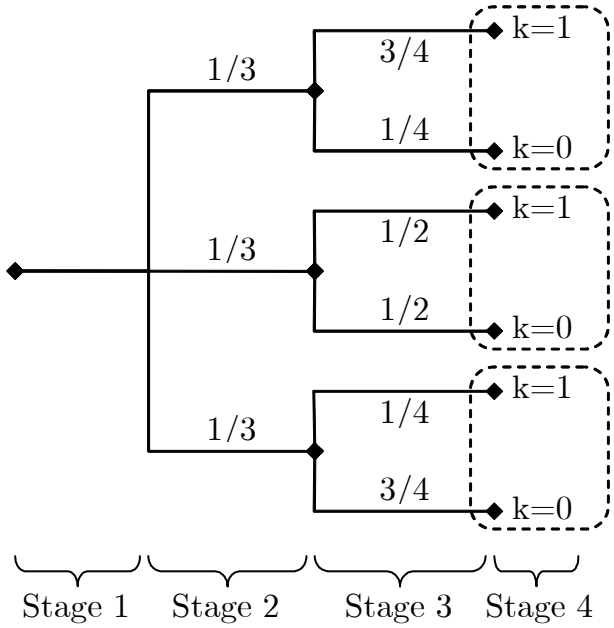


Figure H.4: *$\mathcal{D}$-efficient selection of* $\left\{d_j, n_j\right\}$ *given* (m, n)*, with* $n = 20$ *or* 40.

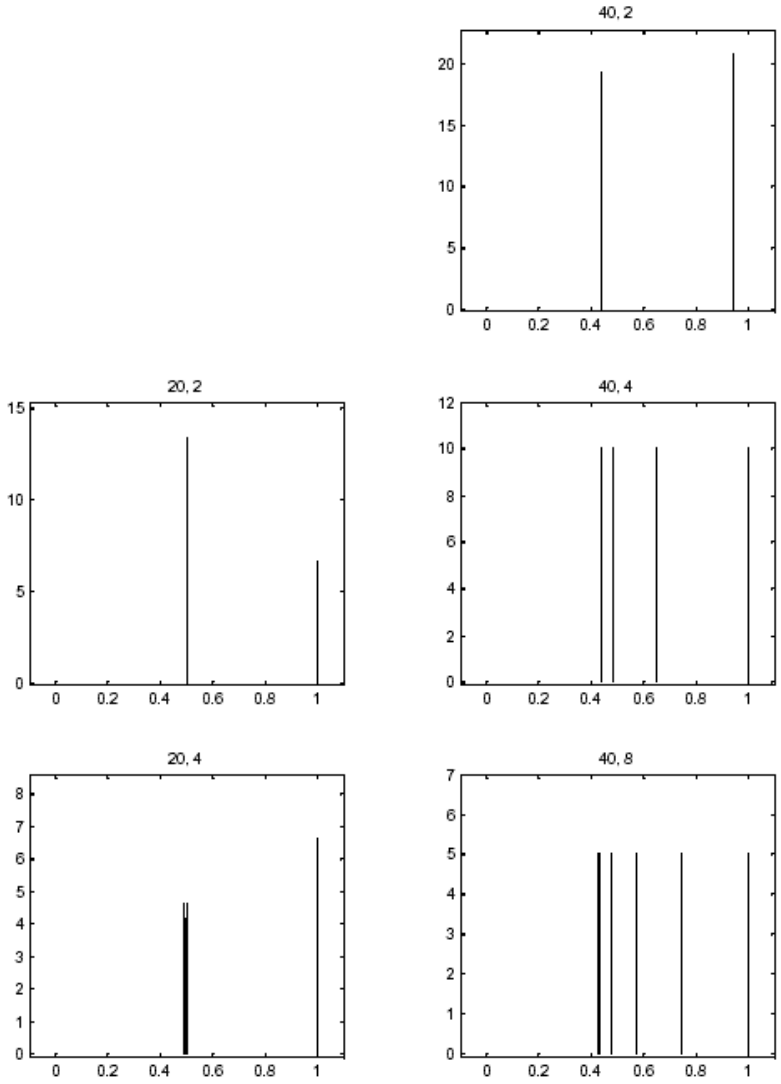


Figure H.5: *$\mathcal{D}$-efficient selection of* $\left\{d_j, n_j\right\}$ *given* (m, n)*, with* $n = 80$ *or* 160*.*

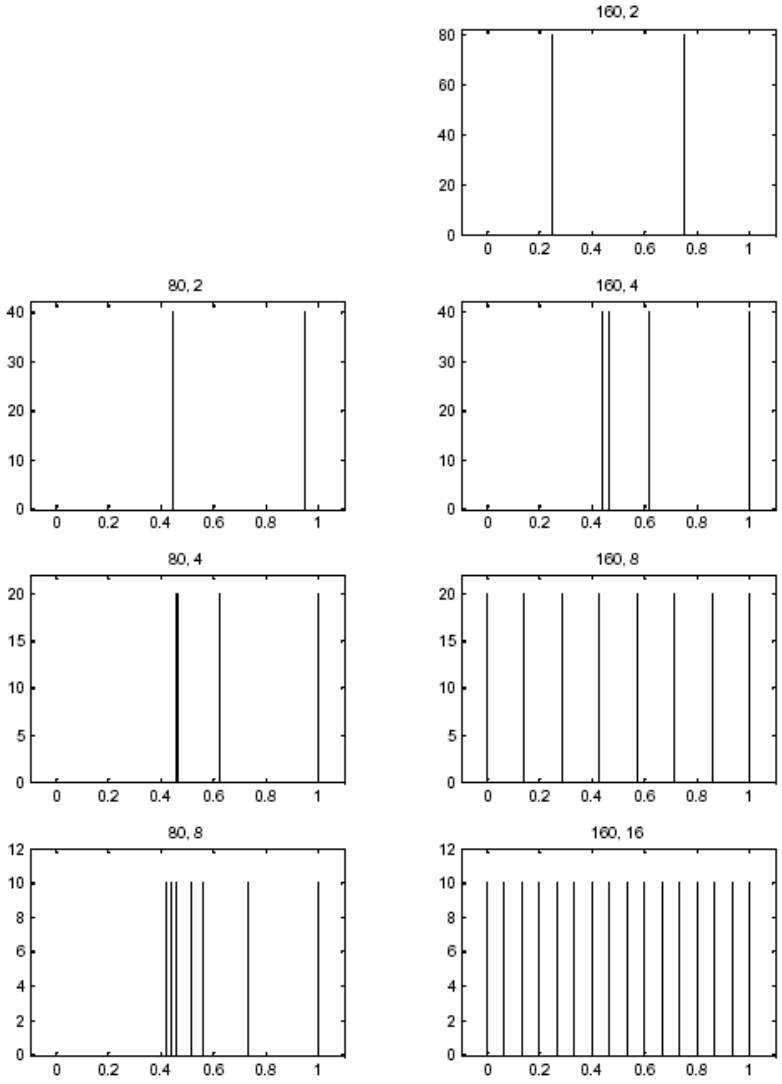
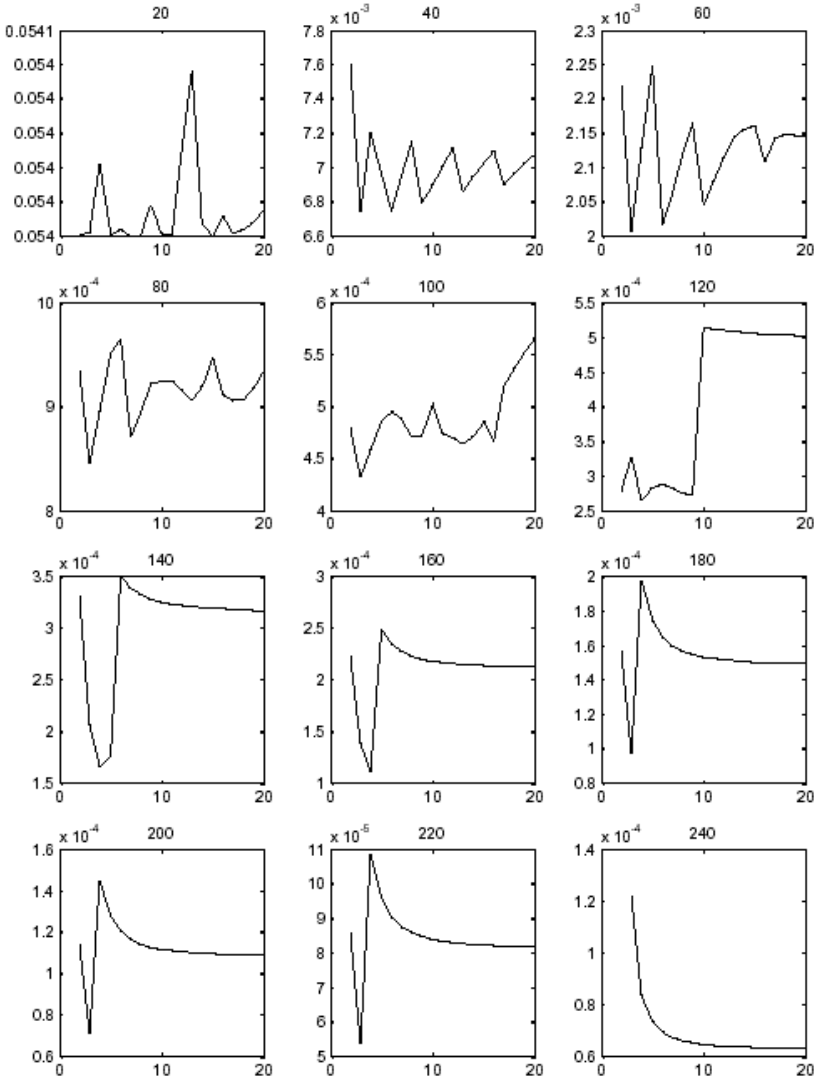


Figure H.6: *$\mathcal{D}$-efficient selection of* $\left\{d_j, n_j\right\}$ *and* m *given* n*, for* $n = 20, ..., 240$*.*



REFERENCES

[1] M. van Stuijvenberg, M. H. Suur, S. de Vos, G. C. Tjiang, E. W. Steyerberg, G. Derksen-Lubsen, and H. A. Moll, Archives of Disease in Childhood. **79**, 120 (1998).

[2] A. Hrobjartsson and P. C. Gotzsche, New England Journal of Medicine **344**, 1594 (2001).

[3] I. Kirsch and G. Sapirstein, Prevention and Treatment **1**, http://www.journals.apa.org/prevention/volume1/pre0010002a.html (1998).

[4] T. J. M. A. S. Kirsch, Irving and S. S. Nicholls, Prevention and Treatment **5**, http://www.journals.apa.org/prevention/volume5/pre0050023a.html (2002).

[5] A. Hrobjartsson, European Journal of Clinical Pharmacology **50**, 345 (1996).

[6] S. Penick and L. E. Hinkle, Psychosomatic Medicine **26**, 369 (1964).

[7] S. B. Penick and S. Fisher, Psychosomatic Medicine **27**, 177 (1965).

[8] G. A. Marlatt and D. J. Rohsenow, in *Advances in Substance Abuse*, edited by N. Mello (JAI Press, Greenwich, Conn., 1980).

[9] I. Kirsch and L. J. Weixel, Behavioral Neuroscience **102**, 319 (1988).

[10] M. P. Jensen and P. Karoly, Psychosomatic Medicine **53**, 144 (1991).

[11] M. A. A. A. C. C. G. M. Pollo, Antonella and F. Benedetti, Pain **93**, 77 (2001).

[12] L. S. M. I. K. A. D. G. H. M. Price, Donald D. and S. S. Nicholls, Pain **83**, 147 (1999).

[13] T. Philipson and J. DeSimone, Biometrika **84**, 221 (1997).

[14] Y. M. M. A. Berthelot, Jean-Marie and A. Prost, Joint Bone Spine **68**, 65 (2001).

[15] C. L. P. Voudouris, Nicholas J. and G. Coleman, Pain **43**, 121 (1990).

[16] C. L. P. Voudouris, Nicholas J. and G. Coleman, Pain **38**, 109 (1989).

[17] H. Benson and R. Friedman, Annual Review of Medicine **47**, 193 (1996).

[18] M. Y. An, centre for Labour Market and Social Research, University of Aarhus, Working Paper 96-01 (unpublished).

[19] D. E. M. S. H. H. G. N. T. J. G. T. de Craen, Anton J. M. and J. Kleijnen, British Journal of Clinical Pharmacology **48**, 853 (1999).

[20] D. E. Moerman, Medical Anthropology Quarterly **14**, 51 (2000).

[21] T. Amemiya and F. Nold, Review of Economics and Statistics **57**, 255 (1975).

[22] G. Maddala, *Limited Dependent and Qualitative Variables in Econometrics*, *Econometric Society Monographs* (Cambridge University Press, New York, 1983).

[23] D. Pregibon, Applied Statistics **29**, 15 (1980).

[24] I. Kirsch, Biological Psychiatry **47**, 733 (2000).

[25] T. R. Stiger and E. F. Pun, in *Encyclopedia of Biopharmaceutical Statistics*, edited by S.-C. Chow (Marcel Dekker, ADDRESS, 2000), pp. 402–407.

[26] T. J. Kaptchuk, Lancet **351**, 1722 (1998).

[27] G. Dobrilla and C. Scarpignato, Digestive Diseases **12**, 368 (1994).

[28] R. J. Levine, Clinical Pharmacology and Therapeutics **42**, 247 (1987).

[29] T. Philipson, External Treatment Effects and Program Implementation Bias, NBER Working Paper No. 0250 (2000).

[30] I. C. R. J. H. Schulz, Kenneth F. and D. G. Altman, Journal of the American Medical Association **273**, 408 (1995).

[31] P. C. H. S. Chalmers, T.C. and J. H. Smith, New England Journal of Medicine **309**, 1358 (1983).

[32] O. C. J. G. P. D. J. A. K. J.-M. S. C. C. Bergmann, Jean-Francois and R. Dahan, Clinical Trials and Meta-Analysis **29**, 41 (1994).

[33] P. Swamy, *Statistical Inference in Random Coefficient Regression Models* (Springer-Verlag, New York, 1971).

[34] G. W. Dehlert, *A First Course in Design and Analysis of Experiments* (W.H. Freeman and Company, New York, 2000).

[35] A. Hofer, M. Hummer, R. Huber, M. Kurz, T. Walch, and W. W. Fleischhacker, Journal of Clinical Psychopharmacology **20**, 699 (2000).

[36] N. Sharpe, Cardiovascular Drugs and Therapy **16**, 75 (2002).

[37] W. A. Kennedy, C. Laurier, J. L. Malo, H. Ghezzo, J. L'Archeveque, and A. P. Contandriopoulos, International Journal of Technology Assessment in Health Care **19**, 8 (2003).

[38] G. Wiedemann, S. Klingberg, A. Wittorf, A. Fischer, and G. Buchkremer, European Psychiatry **15**, 387S (2000).

[39] L. J. Savage, *The Foundations of Statistics* (John Wiley and Sons, New York, 1954).

[40] D. M. Kreps, *Notes on the Theory of Choice*, *Underground Classics in Economics* (Westview Press, Boulder, CO, 1988).

[41] B. M. M. Manning, Willard G. Jr. and J. P. Acton, Journal of Econometrics **11**, 131 (1979).

[42] C. Morris, Journal of Econometrics **11**, 43 (1979).

[43] J. Conlisk and H. Watts, Journal of Econometrics **11**, 27 (1979).

[44] D. J. Aigner, Journal of Econometrics **11**, 7 (1979).

[45] J. M. Lucas, Technometrics **16**, 561 (1974).

[46] M. Bagnoli and T. Bergstrom, university of Michigan (unpublished).

[47] J. Heckman and B. E. Honore, Econometrica **58**, 1121 (1990).

[48] A. Prekopa, Acta Scientiarium Mathematicarum **33**, 335 (1973).

[49] E. Skovlund, Lancet **337**, 1041 (1991).

[50] G. F. H. L. Skovlund, E. and B.-I. Nesheim, European Journal of Clinical Pharmacology **40**, 343 (1991).

[51] G. F. H. L. Skovlund, E. and B.-I. Nesheim, European Journal of Clinical Pharmacology **40**, 539 (1991).