Estimating Effect Sizes in Survey Experiments*

Michael Peress†

May 17, 2016

Abstract

Survey experiments often exploit the fact that public is not perfectly informed. Such survey experiments will attempt to estimate the causal effect of a variable X on another variable by informing a randomly selected subset of the survey respondents that X is true. I develop a framework for identifying effect sizes in such survey experiments. These survey experiments are typically analyzed using the difference of means between the treatment and control group. I define two relevant quantities of interest—the *total effect* and the *real world effect*. The difference of means recovers neither of these quantities. I develop experimental designs and estimators for recovering the total effect and the real world effect.

*I would like to thank Jason Barabas, Dan Butler, Kabir Khanna, Paul Testa, and participants of a seminar at the Midwest Political Science Association conference (Chicago, 2016) for helpful comments and suggestions.

†Department of Political Science, SUNY-Stony Brook. `michael.peress@stonybrook.edu`

1 Introduction

Consider the following four research questions:

- Do Supreme Court decisions affect public opinion?
- Do the positions that candidates take affect their electoral prospects?
- Does a political scandal affect a candidate's support?
- Does the use of WMDs affect the willingness of the American public to support intervention in a foreign conflict?

Each of these questions could be potentially studied using a randomized survey experiment where part of a sample of respondents is used as a control and the remaining respondents are informed about a Supreme Court decision, a candidate's position, a political scandal, or the use of WMDs in a conflict. The survey experiment would make use of the fact that not all survey respondents will be aware that (for example) the Supreme Court has ruled in favor of "ObamaCare", that Marco Rubio supports comprehensive immigration reform, that George Allan used a racial slur at a campaign event, or that chemical weapons were used in the Syrian conflict. This data would typically be analyzed using the difference of means between the treatment and control group, with the difference attributed to the respondents gaining the information provided to the treatment group.

This approach has been widely applied, but my focus here is specifically on estimating effect sizes in such survey experiments. Consider the first of the research questions identified above. The Supreme Court decides to uphold ObamaCare. A survey experiment is employed where a random subset of respondents is treated with the information that the Supreme Court made such a decision. A difference of means is calculated between the treatment and control groups. What is desired is an estimate of the difference in public opinion between a state of the world where the Supreme Court rules that ObamaCare is constitutional vs. a state of the world were no such case is considered by the Supreme Court (I term this the *real world effect*).[1] The manipulation we employ is far removed

[1] This is of course not the only possibility—we may desire to measure the difference in public opinion between a state of the world were the Supreme Court rules that ObamaCare is constitutional vs. a state of the world were the Supreme Court rules that ObamaCare is not constitutional. Such a possibility will be considered later in this article.

from this manipulation though—instead, we are comparing public opinion between a state of the world where the public has a natural level of knowledge of the Supreme Court decision to a state of the world were the public is fully educated of the Supreme Court decision. These effects could differ substantially—if the level of knowledge is high in the natural state (if almost everyone is aware of the Supreme Court decision), the survey experiment will produce a small estimate, because there a few individuals to educate. The actual effect could still be quite large, if news of the decision has spread widely.

A case could be made that the problem identified above is of limited concern because the survey experiment will produce an estimate that has the same sign as the real world effect. The second research question identified above produces an instance where calculating the size of the effect is absolutely essential. Consider two positions taken by President Obama during his presidency. Barack Obama supported gay marriage and supported military intervention in the Syrian conflict. Suppose we are interested in determining which of these issues has a larger effect on public opinion. Here, we may mean two different things. We may mean the effect of a position change on each issue on Obama's approval rating (this would be what I termed the real world effect). Alternatively, we may be interested in a manipulation which measures the difference between a state of the world where everybody believes Obama opposes gay marriage (or intervention in Syria) with a state of the world where everybody believes that Obama supports gay marriage (or intervention in Syria). I term this the *total effect*. Employing the difference between the treatment and control means the survey experiment would estimate neither of these effects. In particular, knowledge of Obama's support for intervention in Syria may have been much lower than knowledge of Obama's support for gay marriage, leaving less room for finding a large effect in the survey experiment. We would thus underestimate the total effect because the survey experiment would produce an estimate that would conflate salience with the natural knowledge level. If the real world effect were of interest, the estimate produced by the survey experiment would be similarly problematic.

This paper is devoted to defining the types of interventions which are typically of interest, developing a modeling framework for survey experiments that provide information to respondents, identifying conditions under which meaningful interventions can be identified, and providing experimental designs and estimators that can recover estimates of these interventions.

2 Existing Designs

2.1 Cross Sectional Designs

Considering the four research questions identified above, I first discuss cross-sectional designs and event studies, and identify their potential limitations. The second research question—do the positions that candidates take affect their electoral prospects?—can be studied using a cross-sectional design (Wrone, Brady and Cogan, 2002; Shor and Rogowski, 2010; Hollibaugh, Rothenberg and Rulison, 2013). One can correlate the position of candidates with the preferences of voters among the candidates. This approach is potentially vulnerable to reverse causality—Lenz (2009, 2010), for example, argues that voters will adopt the positions of their favored candidate. One need not take Lenz's position to see a problem though—even if one believes voters have a general ideological orientation and vote based on this orientation, on particularly low salience issues, voters may simply not have enough information to have their own position and may cue off of their preferred candidates. A survey experiment can solve this problem by manipulating respondent's beliefs about a candidate's position and observing the effect of his stated preference between the candidates. A survey experiment can directly address the potential for two-way causality by manipulating respondents knowledge of the candidates positions and observing whether change occurs in the respondents evaluations of the candidates or the respondents own stated positions.

A similar threat to inference is present in the first research question—do Supreme Court decisions affect public opinion? While we may observe a correlation between public opinion and Supreme Court decisions, this correlation may obtain because the Supreme Court affects the public, or because the public affects the Supreme Court. A survey experiment is useful here for the same reason—voter knowledge of the Supreme Court decision can be directly manipulated. A second issue with the cross section design here is that while there are many House, Senate, and gubernatorial candidates taking positions on a given issue, there is only one Supreme Court. A cross-sectional study would have to identify a set of potential Supreme Court cases, identify cases in which the Supreme Court has ruled, and identify a meaningful measure of public opinion on the relevant issue for that ruling. Applying this approach would be far from trivial.

The third and fourth research questions are even more difficult to study using a cross-sectional approach. Scandals are rather unique events and the use of WMDs in a conflict that the U.S. could potentially enter even more so.

2.2 Event Studies

Each of these research questions could potentially be approached using an event study. For example, support for U.S. intervention in Syria could be measured before and after the reported use of chemical weapons. A candidate's level of support could be measured before or after a scandal is reported. The public's support of a policy could be measured before or after a Supreme Court decision is announced. Each study is potentially useful, but also potentially problematic. The persuasiveness of each of these studies is predicated on the fact that there have no other intervening events, between when public opinion is measured. Consider the case of a Supreme Court decision. In order to avoid intervening events, in may be tempting to measure public opinion immediately prior to and immediately after the decision is announced, but this approach is problematic if voters learn of the Supreme Court decision slowly over time. If a vast majority of the public will eventually learn of the Supreme Court decision, is an estimate of the effect of the Supreme Court decision based on a measure one day after the decision, when only a small fraction are aware of the decision, really that relevant? There are similar issues with applying an event study to the other research questions identified.

2.3 Survey Experiments

The goal of the previous subsection was not to convince the reader that cross-sectional studies and event studies should never be applied to studying the four research questions identified above, but instead to provide the motivation for survey experiments as one approach to answering such research questions. Survey experiments have been widely applied because they have high internal validity—we can be fairly confident that changes we observe between the treatment and control group can be attributed to the treatment. Survey experiments can however suffer from poor external validity (Gaines, Kuklinski and Quirk, 2009; Barabas and Jerit, 2010). If fact, the type of problems we identify in this paper of estimating substantive effect sizes could be viewed as one such threat to external validity—the manipulation we quantify in the survey experiment is different from the real world effect and total effect we are often interested in estimating.

Of course, survey experiments are not the only type of experimental approach that can be applied. In principle, the problem of external validity could be circumvented using field experiments. Consider, however, the problem of estimating the effect of issue positioning on a candidate's vote share. A field experiment would require the researcher to randomly assign the position of a cross-section of candidates for office. This is clearly not a feasible research design in most situations.[2] Here, a survey experiment is a viable alternative—although we cannot manipulate the positions of candidates, we can manipulate the beliefs of survey respondents of the positions of candidates. The infeasibility of directly manipulating candidate positions makes survey experiments an attractive experimental approach for studying candidate positioning. Similarly, we cannot directly manipulate whether the Supreme Court makes a decision or whether chemical weapons are used in a foreign conflict, but survey experiments provides a second best alternative that sacrifices external validity for more wide applicability.

My focus in this article is on survey experiments that manipulate the information available to respondents. Such survey experiments can be characterized as those that provide respondents with hypothetical information, those that provide respondents with true information, and those that provide respondents with false information. The advantage of experiments that provide hypothetical information is that the researcher is much less constrained—for example, Doherty, Dowling and Miller (2011) are able to compare the impact of sex scandals and financial scandals on election outcomes by asking survey respondents to evaluate hypothetical candidates. Similarly, Hainmeuller and Hopkins (forthcoming) apply conjoint analysis to study immigration attitudes, by providing respondents with information about hypothetical immigrants. By contrast, consider Egan and Citrin (2009), who inform survey respondents about six decisions made by the Supreme Court and

[2] Wantchekon's (2003) study is a rare exception.

measure the respondents' opinions on those six issues. Berinsky et al. (2011) used the third variety of survey experiment, treating some of the subjects with misleading information.

Designing a study were respondents are informed of real information is clearly both more limiting and more difficult to implement. The advantage comes in terms of external validity. When studies present hypothetical options to survey respondents, respondents may cue off of whatever limited information is provided to them and effect sizes may be exaggerated. Experiments that provide respondents with truthful knowledge and ask respondents to evaluate real candidates or policies can potentially produce estimates with greater external validity, but this advantage is not fully realized in existing work. Specifically, existing work may estimate manipulations that are not of direct relevance and may conflate knowledge with salience. The framework presented here can potentially alleviate these problems.

The techniques can be best applied to survey experiments (and occasionally experiments that do not involve a survey) that provide treated units with knowledge of true facts. Beyond Egan and Citrin (2009), a number of recent studies meet this criteria. Humphreys and Weinstein (2012) inform survey respondents in Africa about the actual records of candidates for office and measure electoral behavior in the treatment and control groups. Lupu (2013) provides survey respondents with information about inter-party alliances and party switching and measures support for political parties. de Figueiredo, Hidalgo and Kasahara (2013) inform potential voters about the actual corruptions records of challengers in a Brazilian election and measure turnout, spoiled ballots, and candidate choice. Butler and Nickerson (2011) provide state legislators information about the policy positions of their constituents on two issues and measure their voting behavior on related roll call votes. Broockman and Butler (forthcoming) provide survey respondents with the positions of state legislators and measure respondents' support for the same positions.

In each of the six studies identified above, the effect identified is informative, but we could learn more by applying the techniques developed in this paper. For example, in Humphreys and Weinstein (2012) we identify the magnitude of the effect of informing voters of the candidates records. Using my techniques, we could also identify the magnitude of the effect of all voters being informed vs. no voters being informed. In Butler and Nickerson (2011), we identify the magnitude of the effect of informing state legislators about the opinions of their constituents. Using my techniques, we could also identify the magnitude of the effect of all state legislators being informed vs. no state legislators being informed and the magnitude of the effect of a 1 percentage point change in support for a policy on the probability that a state legislator votes in favor of that policy.

3 Modeling Framework

Consider the following setup. We would like to know the effect of the binary variable X (e.g. the position that a politician takes on an issue) on the outcome variable Y_n (e.g. the survey respondent's approval of the politician). Individuals do not know the value of X, but instead have beliefs about the probability that $X = 1$. We let P_n denote the beliefs of respondent n about the probability that $X = 1$. We let $Y_n(x, p)$ denote the potential outcome of Y_n as a function of the value of X, x, and the value P_n, p. We let $P_n(x)$ denote the potential outcome P_n as a function of the value of X. We assume that $P_n \in \{0, 1\}$, or that respondents either believe that $X = 1$ or $X = 0$.[3] We assume a linear model for the conditional expectation of the outcome variable,

$$E[Y_n(x, p)] = \alpha + \beta p \tag{1}$$

In the framework, the outcome variable Y_n, typically a measure of a survey respondent's opinion, is directly affected by the individual's belief that $X = 1$ and only indirectly affected by X itself, where the indirect effect is captured by the dependence of the belief $P_n(x)$ on the event $X = x$. In other words, we assume that a survey respondent does not react to what he does not know. Once we assume that $E[Y_n(x, p)]$ does not depend on x and $P_n \in \{0, 1\}$, the linearity assumption here is without loss of generality. In addition, we assume homogeneous effects here for simplicity, but this assumption can to a degree be relaxed. We assume that $X = 1$ is true and that the survey experiment informs individuals of this truth in the treatment group.

Consider a conventional survey experiment. In the data, we observe a sample of individuals of N individuals. We let Y_n denote the outcome variable and we let T_n denote treatment status where

[3] We make this assumption because in a survey, it is easier to measure discrete beliefs. We can allow individuals to have beliefs in the form of probabilities if we assume that a respondent will report $X = 1$ with probability P_n.

$T_n = 1$ for individuals in the treatment group and $T_n = 0$ for individuals in the control group. We have that,[4]

$$E[Y_n|T_n = 0] = E[Y_n(P_n(1), 1)] = \alpha + \beta E[P_n(1)] \tag{2}$$

$$E[Y_n|T_n = 1] = E[Y_n(1, 1)] = \alpha + \beta \tag{3}$$

Here, $E[Y_n|T_n = 0] = E[Y_n(P_n(1), 1)]$ because in the control group, the outcome variable will be governed by the natural level of knowledge, $P_n(1)$, and $E[Y_n|T_n = 1] = E[Y_n(1, 1)]$ because in the treatment group, all individuals will be assumed to be informed of the true value, $X = 1$. In large samples,

$$\bar{Y}_0 = \frac{\sum_{n=1}^{N}(1 - T_n)Y_n}{\sum_{n=1}^{N}(1 - T_n)} \xrightarrow{prob.} E[Y_n|T_n = 0] = \alpha + \beta E[P_n(1)] \tag{4}$$

$$\bar{Y}_1 = \frac{\sum_{n=1}^{N} T_n Y_n}{\sum_{n=1}^{N} T_n} \xrightarrow{prob.} E[Y_n|T_n = 1] = \alpha + \beta \tag{5}$$

This implies that we can identify $\alpha + \beta E[P_n(1)]$ and $\alpha + \beta$ from a conventional survey experiment. The conventional estimator, the difference of means $\bar{Y}_1 - \bar{Y}_0$ will identify,

$$\theta_{se} = \beta(1 - E[P_n(1)]) \tag{6}$$

This is not the quantity we are ultimately interested in however. One quantify we might be interested in is,

$$\theta_{tot} = E[Y_n(x, 1)] - E[Y_n(x, 0)] = \alpha + \beta - \alpha = \beta \tag{7}$$

This quantity, which we call the total effect, quantifies the effect of moving the entire population

[4] Note that here and elsewhere, we interpret expectations as integrating over all random variables. For example, since $E[Y_n(p, x)] = \alpha + \beta p$, we have $E_Y[Y_n(P_n(1), 1)] = \alpha + \beta P_n(1)$. Taking expectations of both sides once more, we have $E[Y_n(P_n(1), 1)] = E_P[E_Y[Y_n(P_n(1), 1)]] = \alpha + \beta E[P_n(1)]$. To simplify the notation, we drop the subscripts denoting which variable we are integrating over through the paper.

from certainty that $X = 0$ to certainty that $X = 1$. A second quantity we might be interested in is,

$$\theta_{real} = E[Y_n(1, P_n(1))] - E[Y_n(0, P_n(0))] \tag{8}$$

$$= \alpha + \beta E[P_n(1)] - \alpha - \beta E[P_n(0)] = \beta E[P_n(1) - P_n(0)]$$

This quantity, which we call the real-world effect, quantifies the effect of an intervention which changes $X = 0$ to $X = 1$. When such a change is made, not all individuals will change their beliefs about X, and individuals who do not change their beliefs (e.g. because they are unaware that a change has been made) will not change their values for the response variable.

The effect we identify in the survey experiment, θ_{se}, has the same sign as both θ_{tot} and θ_{real} (provided that $E[P_n(1)] \neq E[P_n(0)]$ and $\mathrm{E}[P_n(1)] \neq 1$), but is equal to neither of them. We have $\theta_{se} = \theta_{tot}$ only if $E[P_n(1)] = 0$ (i.e. if everybody believed that $X = 0$ before the manipulation). We have that $\theta_{se} = \theta_{real}$ only if $E[P_n(1)] = \frac{1}{2} + \frac{1}{2}E[P_n(0)]$, a condition which will typically not hold. For example, if 25% of the population believes that $X = 1$ when in fact $X = 0$, then we would require that 62.5% of the population believe $X = 1$ when in fact $X = 1$, a rather silly requirement to assume in any particular application.

If in general, the survey experiment identifies neither the total effect nor the real world effect, does it identify anything manipulation of interest? The answer to this is yes—consider a manipulation where individuals are educated about the true value of X,

$$E[Y_n(1, 1)] - E[Y_n(1, P_n(1))] = \theta_{inf} \tag{9}$$

We term this the information effect. It is clear that $\theta_{se} = \theta_{inf}$, so the substantive effect typically estimated in the literature captures a meaningful manipulation. For example, we might be interested in whether Obama's vote share would increase if the public were informed that he was indeed a U.S. citizen. In this case, the conventional survey experiment will identify the quantity of interest. However, in other cases, the information effect will not be relevant. A number of studies consider whether the Supreme Court affects public opinion, recovering $\theta_{se} = \theta_{inf}$, but in this case, θ_{real} would be more relevant—$\theta_{se} = \theta_{inf}$ is a measure of the change in public opinion that would be observed if voters were fully educated about a particular Supreme Court decision, it does not identify the affect of a ruling by the Supreme Court on public opinion.

The treatment effect we identify in the survey experiment depends on $E[P_n(1)]$ and the real world effect depends on both $E[P_n(1)]$ and $E[P_n(0)]$. Suppose that we vary $p = E[P_n(1)] = 1 - E[P_n(0)]$. In this case, we are varying the probability that a respondent knows the true state of X and we assume that the likelihood of a mistake is the same when $X = 1$ and $X = 0$. We illustrate the treatment effect relative the total effect and real world effect in Figure 1. The total effect is constant (in the plot, we normalize it's value to 1). The real world effect is zero when $p = \frac{1}{2}$—in this case, X has no effect on the outcome Y because it changing X does not change the respondent's beliefs about X. If $p = 1$, changing X changes all the respondents beliefs about X, so the real world effect is equal to the total effect. The more likely respondents are to notice the change in X, the bigger the effect of changing X will have on the outcome. The treatment effect identified in the survey experiment behaves quite differently—higher values of p mean that fewer respondents can be affected by the treatment, and will lead to a smaller value for the treatment effect.

To see how this might effect our conclusions, imagine we are interested in the effect of different issue positions on approval of President Obama. Imagine for a moment that the total effect is the same for the President's position on deporting undocumented workers and gay marriage. If a larger share of respondents are correctly informed about Obama's position on deporting undocument workers, the treatment effect will be smaller even though the total effect is the same and the real world effect is larger. Further complicating things is that the respondents are more likely to be informed about an issue that is salient–that is, p is likely to be high when the total effect is large. In this case, the treatment effect will be highest on exactly those issues where the total effect and real world effect are smallest. Obama's support for abortion, for example, will have a very small treatment effect as almost all survey respondents will already be aware of Obama's position, despite the fact that the abortion is a salient issue (meaning that it will have a large total effect) and that

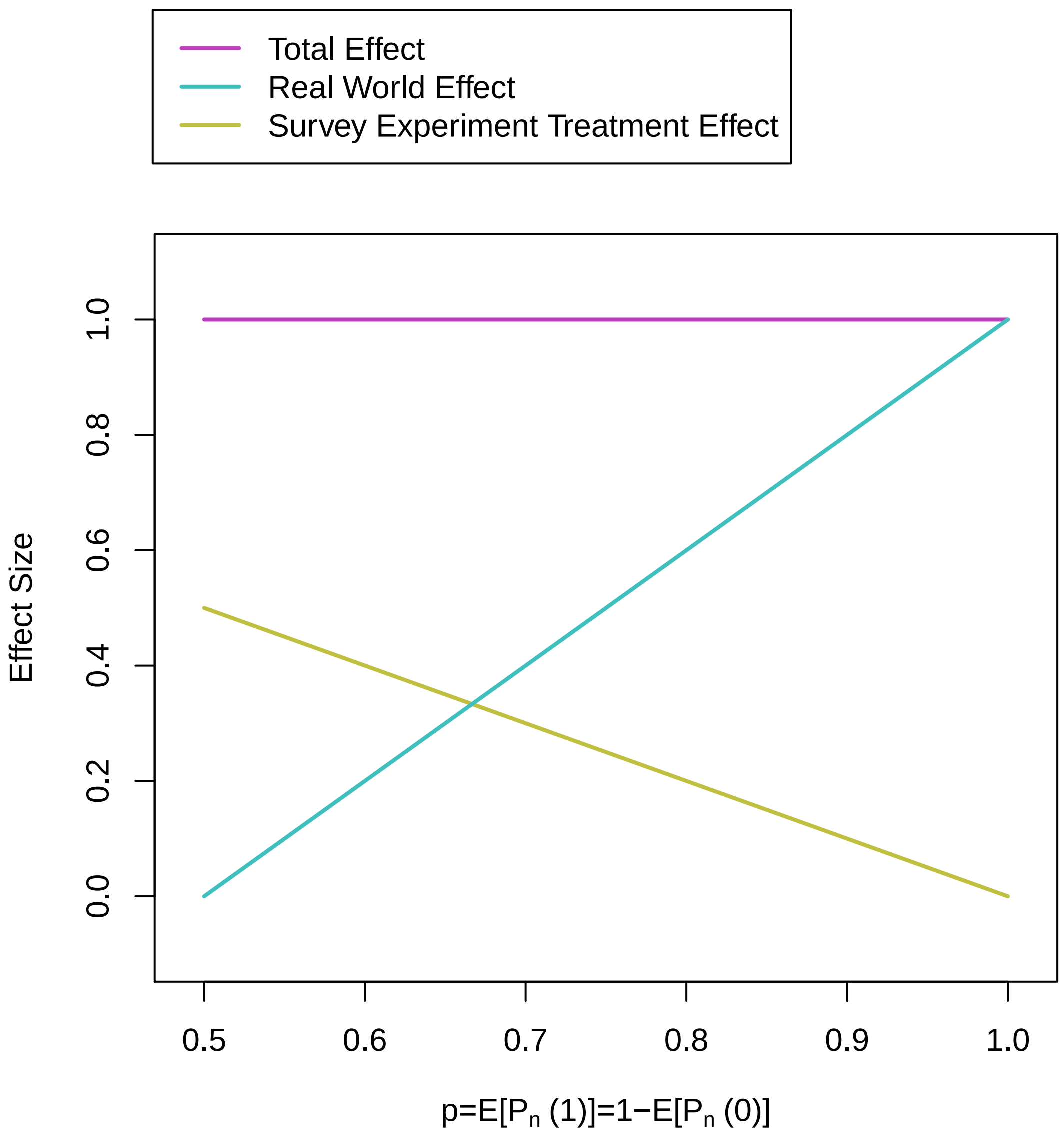

Figure 1: The Total Effect, the Real World Effect, and the Treatment Effect – The size of the total effect is normalize to 1.

a change in Obama's position is very likely to be noticed (meaning that it will have a large real world effect). For these reasons, it is crucial to develop research designs for identifying the total and real world effects.

Under some conditions, we can identify both θ_{tot} and θ_{real}. If we can measure $E[P_n(1)]$, then we observe $\alpha+\beta$, $\alpha+\beta E[P_n(1)]$, and $E[P_n(1)]$, and can back out an estimate for $\theta_{tot} = \beta$. Identifying θ_{real} requires us to have measures of both $E[P_n(1)]$ and $E[P_n(0)]$, or alternatively to measure $E[P_n(1)]$ and assume that $E[P_n(0)] = 0$. The assumption that $E[P_n(0)] = 0$ may indeed be valid in situations and we elaborate on this in the examples below.

An alternative approach for identifying the total effect is to provide respondents with false information that $X = 0$. In this case, we could observe $E[Y_n(0,0)] = \alpha$ in a second treatment group. In this case, we could estimate θ_{tot} using the two treatment groups alone (we would not have to obtain a separate estimate of $E[P_n(0)]$). Whether this is an attractive alternative depends—compliance may be a much more serious challenge as individuals who know that in fact $X = 1$ may not comply with the manipulation, in the sense that they may not believe that $X = 0$.

We can thus summarize our results as follows—if we can assume that $E[P_n(0)] = 0$ or if we can measure $E[P_n(0)]$, then we can identify both the total and real word effect of X on Y_n using a survey experiment, but in addition to the usual design, we also must measure the current (pre-manipulation) beliefs of the population. If we cannot assume that $E[P_n(0)] = 0$ and cannot measure it, then a survey experiment can help us learn about the total effect provided that we also measure current beliefs, but the conventional survey experiment provides no help in learning about the real world effect of the manipulation (though it does allow us to identify the sign of the effect). If $\beta \geq 0$, then we can determine that $0 \leq \theta_{real} \leq \beta E[P_n(1)]$ and if $\beta \leq 0$, we can determine that $\beta E[P_n(1)] \leq \theta_{real} \leq 0$. Identifying the total effect requires adding a measure of beliefs in the control group to the conventional survey experiment. Identifying the real world effect will require further modification of the conventional design, which we consider in the examples below.

4 Examples and Experimental Design

4.1 The Supreme Court Rules that "Obamacare" is Constitutional

We next motivate a number of experimental designs through some examples. The Supreme Court ruled that ObamaCare is constitutional. Supreme Court rulings potentially affect public opinion. However, even if the Supreme Court rules in a certain way, not all individuals will be aware that the Supreme Court has ruled. Let $X = 1$ denote a ruling that ObamaCare is constitutional and let $X = 0$ denote the absence of a Supreme Court case.

The total effect represents the difference between everybody being aware of the Supreme Court decision and nobody believing the Supreme Court has made such a decision. This quantity can be estimated if we know the percentage who approve of ObamaCare, the percentage that believe that the Supreme Court has ruled, and the percentage that would support ObamaCare if they knew the Supreme Court ruled.

To compute the total effect, in addition to the information that would be available from a conventional survey experiment, we would need to obtain an estimate of $E[P_n(1)]$, the expected proportion of individuals who would be aware if the Supreme Court made a decision. This can be most easily be accomplished by adding an item to the existing survey experiment. Specifically, we can ask respondents to report whether they are aware that the Supreme Court has ruled ObamaCare constitutional. Asking this question of the treated respondents may not be desirable. Clearly, we could not ask this question after respondents are treated. Asking the respondents before they are treated may be less problematic, but a safe option is to ask this question to the control respondents only. Note that the recommendation we make in terms of research design for estimating the total effect is a trivial addition to the typical survey experiment—one need only add one item to the survey.

We may also be interested in the real world effect—the impact of the Supreme Court ruling on public support for ObamaCare. In order to estimate this quantity, we would need to either assume that $E[P_n(0)] = 0$, or to estimate $E[P_n(0)]$ in some way. In this example, it may be reasonable to assume that nobody believes that Supreme Court has ruled when it has not ruled.

In this example, should we be more interested in the total effect or the real world effect? Well, that depends. If we put ourselves in the shoes of the Supreme Court justices contemplating the effect of their decisions on public opinion, we would most likely like to know the real world affect. However, if we are studying public opinion on the Supreme Court, and would like to compare (for example) the salience of the ObamaCare ruling to some other rulings, we would most likely like to know the total effect. Considering either θ_{real} or θ_{inf} would conflate the salience of case with the level of knowledge that the Supreme Court decided the case. θ_{tot} is more relevant since it is not dependent on levels of information.

4.2 Obama's Support for Gay Marriage

Barack Obama supports gay marriage. Obama's support for gay marriage affects whether an individual approves of Barack Obama. However, even though Barrack Obama supports gay marriage, not all individuals are aware of this. Let $X = 1$ denote Obama taking a position in favor of gay marriage and $X = 0$ denote Obama taking a position against gay marriage. The total effect represents the difference between everybody believing Barack Obama supports gay marriage and everybody believing Barack Obama opposes gay marriage. This quantity can be estimated if we know the percentage who approve of Obama, the percentage that believe Barack Obama supports gay marriage, and the percentage that would support Barack Obama if they knew he supports gay marriage.

As in the previous application, computing the total effect requires an estimate of $E[P_n(1)]$, which in this case is the proportion of respondents who will come to believe that Obama supports gay marriage if he takes such a position. This can again be accomplished by adding an item to the existing survey experiment asking respondents to report whether they believe Barack Obama support's gay marriage. To compute the real world effect—the impact on Obama's approval rating of taking a position supporting gay marriage—we would need to either assume that $E[P_n(0)] = 0$—that nobody would believe that Obama supports gay marriage if he were to take a position against gay marriage—or to estimate $E[P_n(0)]$ in some way. In this case, symmetry suggests that the assumption that $E[P_n(0)] = 0$ is not reasonable—clearly it does not make sense to assume that

$E[P_n(0)] = 0$ without also assuming that $E[P_n(1)] = 1$. But if we assume that both $E[P_n(0)] = 0$ and $E[P_n(1)] = 1$, then both the control and the treatment will identify $\alpha + \beta$ (i.e. if everybody knows the truth, the manipulation in the survey experiment cannot be effective).

So it remains to develop research designs for identifying θ_{real}. Recall that the missing quantity is $E[P_n(0)]$. Consider first the event study design. We have $X_t = 0$ for $t < t^*$ and $X_t = 1$ for $t \geq t^*$, where t^* denotes the point in time at which Obama announces a change in his position on gay marriage. Let $Y_{nt} = 1$ denote approval of Obama by respondent n at time t. We would like to observe the effect to X_t on $E[Y_{nt}]$—the effect of Obama announcing his support of gay marriage on his approval rating. We could compare the pre and post survey and estimate $E[Y_{nt}] - E[Y_{ns}]$ with $s < t$, but this may not isolate Obama's announcement for other factors. By selecting t and s very close to t^*, we may make this problem go away, but if information is the mechanism, it may take a while for news of Obama's change in position to transmit, so estimating $E[Y_{nt}] - E[Y_{ns}]$ will only tell us the immediate effect of a change in position which may not be representative of the long run effect.

Consider instead the following hybrid design. We estimate $E[P_n(0)]$—the percentage that believed that Obama was against gay marriage when he was against gay marriage—using a survey before Obama announces his position change on gay marriage. We then perform the same survey experiment after the intervention, far enough away from t^* that the information has had a chance to transmit. For this hybrid design to be applicable, we would have to assume that $E[P_{nt}]$ is not confounded in the same way that $\mathrm{E}[Y_{nt}]$ is (i.e. that there are not intervening factors that affect beliefs), but this assumption may be quite reasonable. The other issue is that we might not know about a change from $X_s = 0$ to $X_t = 1$ until it is too late to field a survey. One way around this problem is luck—there may have been a relevant survey conducted beforehand. Alternatively, we could consider the change from $X_s = 0$ to $X_t = 1$ as representing a formal announcement of a position (i.e. an announcement in the state of the union address), in which case we may know the announcement is coming before it happens. While it may seem that such a hybrid study would be difficult to conduct, Barabas and Jerit (2009, 2010) were able to successfully conduct a number of such studies.

In the case of Obama's support of gay marriage, are we more interested in the total effect or the real world effect? Well, that again depends. If we put ourselves in the shoes of Obama's political advisors, we would most likely like to know the real world affect. However, if we are studying voting behavior, and would like to compare the salience of gay marriage with the salience of Obama's support for intervention in Syria, we would most likely like to know the total effect. As before, the total effect isolates the effects of the intervention from the information processes under which people learn about the intervention. From the perspective of a president considering changing his position, the total effect will be quite misleading. A large literature has a argued that Americans posses little detailed knowledge of policy issues so that computing the total effect may over-represent the impact a position taking by the president. A scholar studying voting behavior would be best served by estimates of both the total and real world effects, so as to understand both the importance of issues to voters as well as the process by which voters learn about politicians' positions on issues.

4.3 Congressional Voting on Immigration Reform

There is another type of hybrid study we could consider, and I motivate this by a third example. Specifically, consider a hybrid between a survey experiment and a cross-section study. I depart from the previous assumption that, in fact, $X = 1$, and suppose instead that $X_i \in \{0, 1\}$. Here, i denotes a particular congressman, $X_i = 1$ denotes a conservative position on immigration reform, and $X_i = 0$ denotes a liberal position. P_{ni} denotes whether a respondent represented by congressman i believes that congressman n took the conservative position.

To estimate the relationship between P_{ni} (the proportion that believe the congressman took the conservative position) and X_i (whether the congressman actually took the conservative position), we could estimate a regression of the form,

$$P_{ni} = a + bX_i + cZ_{ni} + \varepsilon_{ni} \tag{10}$$

where Z_{ni} are a set of control variables and ε_{ni} is an error term. Again, we could compare this to the following cross-section regression,

$$Y_{ni} = a + bX_i + cZ_{ni} + \varepsilon_{ni} \tag{11}$$

but we may worry the this equation is confounded because in districts where support for the Republican is high, voters may adopt the conservative position at greater rates. The question then is whether $P_{ni} = a + bX_i + cZ_{ni} + \varepsilon_{ni}$ is similarly confounded—do the proportion of individuals who would believe that the congressman holds a conservative position affect whether the politician holds the conservative position? (after controlling for covariates Z_{ni}). The assumption of un-confounding seems more appropriate in this regression because it would be very difficult for a politician to know the proportion of individuals who would be unaware of a position change and hence react to it by changing his position, particularly once a measure of the congressman's general ideology is controlled for.

Let $Y_{ni} = 1$ denote that respondent n approves of Congressman i. Let $X_i = 1$ if Congressman i voted in favor of an immigration reform bill and let $X_i = 0$ if the Congressman voted against the immigration reform bill. We write $E[Y_{ni}(p, x)] = \alpha_i + \beta_i p$ where p denotes whether the respondent believes Congressman took a conservative position. We also write $P_{ni}(x)$ for the potential outcome that the respondent believes that the congressman took a conservative position. In the survey experiment, we observe $E[Y_{ni}(P_{ni}(X_i), X_i)]$ in the control group and $E[Y_{ni}(X_i, X_i)]$ in the treatment group. The total effect is $\theta_{tot} = E[Y_{ni}(1, 1)] - E[Y_{ni}(0, 0)] = \beta$ and the real world effect is $E[Y_{ni}(P_{ni}(1), 1)] - E[Y_{ni}(P_{ni}(0), 0)] = \beta(E[P_{ni}(1)] - E[P_{ni}(0)])$. In the sample, we observe $P_{ni}(1)$ for i such that $X_i = 1$ and $P_{ni}(0)$ for i such that $X_i = 0$. However, consider the regression, $P_{ni} = a + bX_i + cZ_{ni} + \varepsilon_{ni}$. According to this regression, we have $E[P_{ni}(1)] = a + b + cE[Z_{ni}]$ and $E[P_{ni}(0)] = a + cE[Z_{ni}]$ which implies that $E[P_{ni}(1)] - E[P_{ni}(0)] = b$ so that $\theta^i_{real} = \beta_i b = \theta^i_{tot} b$. This immediately suggests an estimator for the real world effect—we simply multiply the total effect in the district by the slope of X_i in the regression of beliefs on positions.

5 Estimation and Inference

5.1 The Total Effect

The identification results lead to the following estimators of θ_{tot} and θ_{real}. Consider a sample of N individuals, denote the outcome by Y_n, and denote the treatment by T_n. Assume that P_n represents belief that $X = 1$, which is measured in the control group. We can estimate,

$$\bar{Y}_0 = \frac{\sum_{n=1}^{N}(1 - T_n)Y_n}{\sum_{n=1}^{N}(1 - T_n)} \tag{12}$$

$$\bar{Y}_1 = \frac{\sum_{n=1}^{N} T_n Y_n}{\sum_{n=1}^{N} T_n} \tag{13}$$

$$\bar{P}_0 = \frac{\sum_{n=1}^{N}(1 - T_n)P_n}{\sum_{n=1}^{N}(1 - T_n)} \tag{14}$$

Note that,

$$E[\bar{Y}_1] = \alpha + \beta \tag{15}$$

$$E[\bar{Y}_0] = \alpha + \beta E[P_n(1)] \tag{16}$$

$$E[\bar{P}_0] = E[P_n(1)] \tag{17}$$

This suggests that we estimate,

$$\hat{\theta}_{tot} = \frac{\bar{Y}_1 - \bar{Y}_0}{1 - \bar{P}_0} \tag{18}$$

Since $\bar{Y}_1 - \bar{Y}_0 \xrightarrow{prob.} \beta(1 - E[P_n(1)])$ and $1 - \bar{P}_0 \xrightarrow{prob.} 1 - E[P_n(1)]$, we have that $\hat{\theta}_{tot} \xrightarrow{prob.} \beta = \theta_{tot}$.

We can derive the standard error of $\hat{\theta}_{tot}$ using the delta method,

$$se(\hat{\theta}_{tot}) = \sqrt{\frac{\hat{\sigma}^2_{Y_0} + \hat{\sigma}^2_{Y_1} + \hat{\theta}^2_{tot}\sigma^2_{P_0} - 2\hat{\theta}_{tot}\hat{\sigma}_{Y_0,P_0}}{1 - \bar{P}^2_0}} \tag{19}$$

where,

$$\hat{\sigma}^2_{Y_0} = \frac{\sum_{n=1}^{N}(1 - T_n)(Y_n - \bar{Y}_0)^2}{\sum_{n=1}^{N}(1 - T_n)} \tag{20}$$

$$\hat{\sigma}^2_{Y_1} = \frac{\sum_{n=1}^{N} T_n(Y_n - \bar{Y}_1)^2}{\sum_{n=1}^{N} T_n} \tag{21}$$

$$\hat{\sigma}^2_{P_0} = \frac{\sum_{n=1}^{N}(1 - T_n)(P_n - \bar{P}_0)^2}{\sum_{n=1}^{N}(1 - T_n)} \tag{22}$$

$$\hat{\sigma}_{Y_0,P_0} = \frac{\sum_{n=1}^{N}(1 - T_n)Y_nP_n}{\sum_{n=1}^{N}(1 - T_n)} \tag{23}$$

We note that if instead of $\bar{P}_0$, we estimate $E[P_n(0)]$ using a third sample (if contamination is a worry) then we simply replace $\bar{P}_0$ with $\bar{P}_2$ (the mean information level in the third sample) and $\hat{\sigma}_{Y_0,P_0}$ with 0 in the above formula for the standard error.

5.2 The Real World Effect

We assume that we have a different sample where we can estimate $E[P_n(0)]$. Typically, $E[P_n(0)]$ would be estimated from the before component of a survey experiment event study hybrid, where we assume that the sample obtained before the event is independent of the sample obtained after the event.

We index these additional respondents by $N+1$ through $N+M$ and we let $\bar{P}_2 = \frac{1}{M}\sum_{n=N+1}^{N+M} P_n$ denote this estimate, where $E[\bar{P}_2] = E[P_n(0)]$. We consider the estimator,

$$\hat{\theta}_{real} = \frac{\bar{Y}_1 - \bar{Y}_0}{1 - \bar{P}_0}(\bar{P}_0 - \bar{P}_2) \tag{24}$$

We have that $\hat{\theta}_{real} \xrightarrow{prob.} \beta(E[P_n(1)] - E[P_n(0)]) = \theta_{real}$. To derive the standard error of $\hat{\theta}_{real}$, we again apply the delta method,

$$se(\hat{\theta}_{real}) \tag{25}$$

$$= \sqrt{\frac{(\bar{P}_0 - \bar{P}_2)^2(\hat{\sigma}^2_{Y_0} + \hat{\sigma}^2_{Y_1}) + \frac{(\bar{Y}_0 - \bar{Y}_1)^2(1-\bar{P}_2)^2}{(1-\bar{P}_0)^2}\hat{\sigma}^2_{P_0} + (\bar{Y}_0 - \bar{Y}_1)^2\hat{\sigma}^2_{P_2} + 2\frac{(\bar{P}_0-\bar{P}_2)(\bar{Y}_0-\bar{Y}_1)(1-\bar{P}_2)}{1-\bar{P}_0}\hat{\sigma}_{Y_0,P_0}}{(1-\bar{P}_0)^2}}$$

where,

$$\hat{\sigma}^2_{P_2} = \tfrac{1}{M}\sum_{n=N+1}^{M}(P_n - \bar{P}_2)^2 \tag{26}$$

In the special case where we are willing to assume that $E[P_n(0)] = 0$, we can simply set $\bar{P}_2 = 0$ and $\hat{\sigma}^2_{P_2} = 0$ in the above formula above.

5.3 The Real World Effect in a Cross Section Hybrid Study

Using the results from Section 5.1, we can estimate θ^i_{tot} using $\hat{\theta}^i_{tot} = \frac{\bar{Y}_{1,i} - \bar{Y}_{0,i}}{1-\bar{P}_{0,i}}$ and we can estimate b using a linear regression with P_{ni} as the dependent variable. If we assume that I (the number of groups) is large, then $\hat{b}$ and $\hat{\theta}^i_{tot}$ will be independent. Since $\theta^i_{real} = \theta^i_{tot}b$, we can estimate $\hat{\theta}^i_{real} = \hat{\theta}^i_{tot}\hat{b}$ and using the delta method, we obtain,

$$se(\hat{\theta}^i_{real}) = \sqrt{\hat{b}^2 se(\hat{\theta}^i_{tot})^2 + (\hat{\theta}^i_{tot})^2 se(\hat{b})^2} \tag{27}$$

where $se(\hat{\theta}^i_{tot})$ can be calculated according to equation (19) and $se(\hat{b})$ is the standard error of $\hat{b}$ that we obtain from the linear regression of P_{ni} on X_i and the control variables.

5.4 Relation to the Wald Estimator

The estimators I propose for the total and real world effects share some similarities to the Wald estimator used to estimate the effect of a treatment with two-sided non-compliance. In this case, we have a group that was assigned treatment ($Z = 1$) and a group that was not assigned treatment

($Z = 0$). The treatment is denoted as $P = 1$ and the control is denoted as $P = 0$. The treatment is the belief that $X = 1$ and Z can be thought of as an instrument or an encouragement to believe $X = 1$. We observe the mean value of the dependent variable, Y, in the group that was assigned treatment ($\bar{Y}_1$) and in the group that was assigned no treatment ($\bar{Y}_0$). The Wald estimator estimates the treatment effect as,

$$\hat{\theta} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{P}_1 - \bar{P}_0} \tag{28}$$

If we assume that there is full compliance in the treated group ($\bar{P}_1 = 1$), we will recover our estimator for the total effect. In many applications of the Wald estimator, the level of compliance in each group ($\bar{P}_1$ and $\bar{P}_0$) will be observed automatically. For example, if the treatment is being mobilized to vote by telephone and the encouragement is being assigned to receive such a telephone call, we will automatically observe the fraction that were successfully treated. In my framework, we do not automatically observe ($\bar{P}_0$) and must measure it within the survey.

The relationship is weaker for the real world effect. We have a second instrument X, which also instruments for P (the belief that $X = 1$). The real world effect is the intent to treat effect of X on Y, so our goal is to estimate the intent to treat effect of an instrument X (taking the action $X = 1$ vs. taking the action $X = 0$) using a different instrument Z (inducing a belief that $X = 1$).

6 Limitations and Extensions

Here, I describe some objections to the basic framework I introduced and discuss ways of extending the basic framework to address these challenges.

6.1 Truthful Reporting

One issue has to do with obtaining truthful answers for estimating $E[P_n(0)]$. A number of studies have grappled with the problem of measuring respondents' knowledge of a true fact (Franklin, Kosaki and Kritzer, 1993; Alvarez and Gronke, 1996; Ansolabehere and Jones, 2010). While directly asking respondents will often be effective, in some cases, respondents may misrepresent their level of knowledge or guess a correct answer. For example, if respondents are asked to report whether they are aware that the Supreme Court decided on ObamaCare, they may falsely report that they are aware. Similarly, if respondents are asked whether the Supreme Court decided for or against ObamaCare, respondents may try to guess the correct answer. Instead, respondents could be asked whether one four items is true, with the true item being that the Supreme Court ruled in favor of ObamaCare paired with three incorrect items. Corrections for guessing from the testing literature could be applied as well (Lord, 1975). In other cases, assuming truthful reporting may not be problematic.

6.2 Noncompliance

A second issue is that in the basic framework, I make the assumption that in the survey experiment, the mean in the treatment group measures the expectation $E[Y_n(1, 1)]$, or the expected outcome when everybody believes that $X = 1$. This can be viewed as a type of full compliance—all individuals believe what they are told to believe. Lack of full compliance may be particulary problematic in cases where the study provides the respondents with false information (and provides one reason to avoid such studies), but perhaps occurs even when the information is truthful. Modifications of the basic framework are possible.

Let $t = 1$ denote treatment status. Let $E[Y_n(x, p, t)]$ denote the potential outcome as a function of the manipulation x, beliefs about the manipulation p, and treatment status t. Let $P_n(x, t)$ denote potential beliefs. We can assume that $E[Y_n(x, p, t)] = \alpha + \beta p$. In the control group, we will measure $E[Y_n(1, P_n(1, 0), 0)] = \alpha + \beta E[P_n(1, 0)]$ and in the treatment group, we will measure $E[Y_n(1, P_n(1, 1), 1)] = \alpha + \beta E[P_n(1, 1)]$. As before, we can measure $E[P_n(1, 0)]$ in the control group. To identify $\theta_{tot} = \beta$, we would have to estimate $E[P_n(1, 1)]$—the proportion that believe that $X = 1$ when the treatment is administered—in the treatment group. Simply asking the respondents to report whether they believe the Supreme Court ruled ObamaCare constitutional (for example) immediately after informing respondents of this fact may not be effective. Instead, the treatment could be administered earlier in the survey and both Y_n as well as beliefs about $X = 1$ could be measured later in the survey. To estimate θ_{real}, we would have to in addition use one of the hybrid research designs considered in Section 4.

6.3 Binary Intervention

A third issue is the focus on a binary item X. Consider the Supreme Court example again. In Section 4.1, we focused on the Supreme Court deciding that ObamaCare is constitutional against the alternative that the Supreme Court did not consider the constitutionality of ObamaCare. Ideally, we would have also considered the counterfactual relative to the Supreme Court deciding against ObamaCare.

The framework can be extended to handle this in the following way. Let $X \in \{-1, 0, 1\}$ where $X = -1$ denotes a decision against ObamaCare. Assume that $E[Y_n(x, p)] = \alpha + \beta 1\{p = 1\} + \gamma 1\{p = -1\}$. Let $P_n(x)$ denote the potential belief that X, $P_n(x) = 1$ denote a belief that the Supreme Court decided that ObamaCare is constitutional, and $P_n(x) = -1$ denote a belief that the Supreme Court decided that ObamaCare is unconstitutional. Here, $\beta - \gamma$ would denote the total effect of a pro-ObamaCare decision relative to an anti-ObamaCare decision, β would denote the total effect of a pro-ObamaCare decision relative to no decision, and γ would denote the total effect of an anti-ObamaCare decision relative to no decision. Assuming a control group and a treatment group that is informed of the truth, we would be able to identify $\alpha + \beta Pr(P_n(1) = 1) + \gamma Pr(P_n(1) = -1)$ and $\alpha + \beta$. We could estimate both $Pr(P_n(1) = 1)$ and $Pr(P_n(1) = -1)$ using a survey item in the control group. This would allow us to recover estimates of β and γ, which then allow us to recover all three total effects.

The real world effect of a pro-ObamaCare ruling relative to no ruling would be,

$$\theta_{real}^{1,0} = E[Y_n(1, P_n(1))] - E[Y_n(0, P_n(0))] \tag{29}$$

$$= \beta(Pr(P_n(1) = 1) - Pr(P_n(1) = 0) + \gamma(Pr(P_n(1) = -1) - Pr(P_n(0) = -1))$$

the real world effect of a pro-ObamaCare ruling relative to an anti-ObamaCare ruling would be,

$$\theta_{real}^{1,-1} = E[Y_n(1, P_n(1))] - E[Y_n(-1, P_n(-1))] \tag{30}$$

$$= \beta(Pr(P_n(1) = 1) - Pr(P_n(-1) = 1)) + \gamma(Pr(P_n(1) = -1) - Pr(P_n(-1) = -1))$$

and the real world effect of a anti-ObamaCare ruling relative to no ruling would be,

$$\theta_{real}^{-1,0} = E[Y_n(-1, P_n(-1))] - E[Y_n(0, P_n(0))] \tag{31}$$

$$= \beta(Pr(P_n(-1) = 1) - Pr(P_n(0) = 1)) + \gamma(Pr(P_n(-1) = -1) - Pr(P_n(0) = -1))$$

To estimate the real world effects, we could assume that $Pr(P_n(0) = 1) = Pr(P_n(0) = -1) = 0$ and estimate $Pr(P_n(-1) = 1)$ and $Pr(P_n(-1) = -1)$ using a hybrid event study survey experiment.

6.4 Homogeneous Effects

Consider the hybrid cross-section survey experiment I introduced to estimate the effect of congressional position taking. I purposely allowed the effect to differ over congressman i. The effect may be different for different congressmen because some individuals are more sensitive to a particular issue, but the effect may also differ because a different proportion of respondents support immigration reform in each district. Allowing β_i to differ over i may recover the correct average effects of certain interventions, but also misses some important features of mass ideology.

To address this issue, I assume that voters differ in their position V_n with $V_n = 1$ denoting a conservative position and $V_n = 0$ denoting a liberal position. I consider the following model for Y_n,

$$E[Y_n(x, p, v)] = \alpha + \beta 1\{p = v\} \tag{32}$$

where v indexes the voter's position. Here, respondent n is more likely to approve of his congressman if he shares the same position on immigration reform, with the magnitude of this effect given by β. I let $V_n(x,p)$ denote the potential outcome for the voter's position as a function of the candidates position x and the voters beliefs about the candidate's position p. I assume that,

$$E[V_n(x,p)] = \phi + \delta p \tag{33}$$

Here, we are incorporating the possibility that the candidates position causes the voters position. In the control group, we observe,

$$E[Y_n(1, P_n(1), V_n(1, P_n(1))] = E[\alpha + \beta 1\{P_n(1) = V_n(1, P_n(1))\}] \tag{34}$$

$$= \alpha + \beta(E[P_n(1)](\phi + \delta) + (1 - E[P_n(1)])(1 - \phi))$$

$$E[V_n(1, P_n(1))] = \phi + \delta E[P_n(1)] \tag{35}$$

In the treatment group, we observe,

$$E[Y_n(1, 1, V_n(1, 1)] = \alpha + \beta(\phi + \delta) \tag{36}$$

$$E[V_n(1, 1)] = \phi + \delta \tag{37}$$

We can define the total effects as follows,

$$\theta^Y_{tot} = E[Y_n(1, 1, V_n(1, 1))] - E[Y_n(0, 0, V_n(0, 0)] = \alpha + \beta(\phi + \delta) - \alpha - \beta\phi = \beta\delta \tag{38}$$

$$\theta^V_{tot} = E[V_n(1, 1)] - E[V_n(0, 0)] = \delta \tag{39}$$

Provided that we can estimate $E[P_n(1)]$, we can recover the total effects parameters β and δ, where

δ represents the total effect of the congressman's position on the voter's position and β represents the total effect of congruence on the respondent's approval of their congressman.

We can define the real world effects as follows,

$$\theta^{Y}_{real} = E[Y_n(1, P_n(1), V_n(1, P_n(1))] - E[Y_n(0, P_n(0), V_n(0, P_n(0))] \tag{40}$$

$$= \beta(2\phi + \delta - 1)(E[P_n(1) - P_n(0)])$$

$$\theta^{V}_{real} = E[V_n(1, P_n(1))] - E[V_n(0, P_n(0))] = \phi + \delta E[P_n(1) - P_n(0)] \tag{41}$$

which require us to estimate $E[P_n(0)]$ in addition to the quantities required for estimating total effects. Again, $E[P_n(0)]$ could be estimated using the techniques we developed earlier. This exercise here shows how my framework can be adapted to study particular types of heterogeneity.

7 Conclusions

This article focused on estimating effect sizes in survey experiments and focused on two particular effects—the total effect and the real world effect. In most applications, one of these two effects will be of interest and the conventional estimator from survey experiments does not recover either of these quantities of interest. The total effect can be estimated using a design that slightly modifies existing designs—it is necessary to know the percent of the sample that had knowledge of the intervention. This quantity can typically be measured in the control group.

The real world effect is more difficult to recover. In some cases, it may be possible to assume that the nobody believes an intervention happened when it has not happened. For the situations in which we cannot assume this, I proposed two research designs for estimating these counterfactual beliefs—a hybrid event study survey experiment (like the studies preformed in Barabas and Jerit (2010)) and a hybrid cross-section survey experiment.

Most broadly, the recommendations of this article are that those employing survey experiments that manipulate information should adapt their designs to allow measurement of the information processes. The basic framework presented in this paper provides a way of thinking about these information processes and the extensions considered briefly provide some clues as to how the basic framework can be adapted to deal with more complicated problems.

References

Alvarez, R. Michael and Paul Gronke. 1996. “Constituents and Legislators: Learning about the Persian Gulf War Resolution.” *Legislative Studies Quarterly* 21:105–127.

Ansolabehere, Stephen and Phillip E. Jones. 2010. “Constituents Responses to Congressional Roll-Call Voting.” *American Journal of Political Science* 54:583–597.

Barabas, Jason and Jennifer Jerit. 2009. “Estimating the Causal Effects of Media Coverage on Policy-Specific Knowledge.” *American Journal of Political Science* 53:73–89.

Barabas, Jason and Jennifer Jerit. 2010. “Are Survey Experiments Externally Valid?” *American Political Science Review* 104:226–242.

Berinsky, Adam, Vincent Hutchings, Tali Mendelberg, Lee Shaker and Nicholas Valentino. 2011. “Sex and Race: Are Black Candidates More Likely to be Disadvantaged by Sex Scandals?” *Politicial Behavior* 33:179–202.

Broockman, David and Daniel M. Butler. forthcoming. “The Causal Effects of Elite Position-Taking on Voter Attitudes: Field Experiments with Elite Communication.” *American Journal of Political Science* .

Butler, Daniel M. and David W. Nickerson. 2011. “Can Learning Constituency Opinion Affect How Legislators Vote? Results from a Field Experiment.” *Quarterly Journal of Political Science* 6:55–83.

de Figueiredo, Miguel F.P., F. Daniel Hidalgo and Yuri Kasahara. 2013. “When Do Voters Punish Corrupt Politicians?” Working Paper.

Doherty, David, Conor M. Dowling and Michael Miller. 2011. "Are Financial or Moral Scandals Worse? It Depends." *PS: Political Science & Politics* 44:749–757.

Egan, Patrick J. and Jack Citrin. 2009. "Opinion Leadership, Backlash, and Delegitimation: Supreme Court Rulings and Public Opinion." Working Paper.

Franklin, Charles H., Liane Kosaki and Herbert M. Kritzer. 1993. "The Salience of U.S. Supreme Court Decisions." Working Paper.

Gaines, Brian J., James H. Kuklinski and Paul J. Quirk. 2009. "The Logic of the Survey Experiment Reexamined." *Political Analysis* 15:1–20.

Hainmeuller, Jens and Daniel Hopkins. forthcoming. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes Toward Immigrants." *American Journal of Political Science* .

Hollibaugh, Gary E., Lawrence S. Rothenberg and Kristin K. Rulison. 2013. "Does It Really Hurt to Be Out of Step?" *Political Research Quarterly* 66:856–867.

Humphreys, Macartan and Jeremy M. Weinstein. 2012. "Policing Politicians: Citizen Empowerment and Political Accountability in Uganda." Working Paper.

Lenz, Gabriel S. 2009. "Learning and Opinion Change, Not Priming: Reconsidering the Priming Hypothesis." *American Journal of Political Science* 53:821–837.

Lenz, Gabriel S. 2010. *Follow the Leader? How Voters Respond to Politicians Performance and Policies*. Chicago: University of Chicago Press.

Lord, Frederick M. 1975. "Formula Scoring and Number-Right Scoring." *Journal of Educational Measurement* 12:7–11.

Lupu, Noan. 2013. "Party Brands and Partisanship: Theory with Evidence from a Survey Experiment in Argentina." *American Journal of Political Science* pp. 49–64.

Shor, Boris and Jon C. Rogowski. 2010. "Congressional Voting by Spatial Reasoning." Working Paper.

Wantchekon, Leonard. 2003. "Clientelism and Voting Behavior: Evidence from a Field Experiment in Benin." *World Politics* 55:399–422.

Wrone, Brandice C., David W. Brady and John F. Cogan. 2002. "Out of Step, Out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review* 96(1):127–40.