Experimental triangulation of coalition signals: varying designs, converging results

Gschwend, Thomas; Meffert, Michael F.

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Experimental Triangulation of Coalition Signals: Varying Designs, Converging Results

Michael F. Meffert
Leiden University
Department of Political Science
Wassenaarseweg 52
2333 AK Leiden
Netherlands
Phone. +31-71-527-3862
Fax. +31-71-527-3815
m.f.meffert@fsw.leidenuniv.nl

Thomas Gschwend
University of Mannheim
Center for Doctoral Studies in the
Economic and Behavioral Sciences
68131 Mannheim
Germany
Phone: +49-621-181-2087
Fax: +49-621-181-3699
gschwend@uni-mannheim.de

1. Experiments as Flexible Tools for Theory Testing

It is probably fair to say that political science has not been a welcoming discipline for experimental research (McDermott 2002). Our discipline has always expressed skepticism about the usefulness and the prospects of experimental designs to address the key research questions we care about. But the more political scientists have started to think carefully about causal relationships and what is required to test them, the more they came (or should come) to realize that our traditional methodologies and research designs are also not sufficient. The latter have serious limitations as well, and some of these limitations can be addressed by experimental methods. Because experimental designs have unique strengths compared to other research designs, it is not surprising that the use of experiments has evolved and increased over time (Morton and Williams 2010). Put simply, experiments are flexible tools for theory testing that allow us to establish causality by clearly separating causes and effects.

In this chapter we will focus on one particular but striking advantage of experiments. When the key explanatory factor lacks variance, that is, when no observable data to test a theory is available, experiments can provide an elegant solution for this problem. Even if they come with their own difficulties and drawbacks, political science can only gain by embracing experimental designs. They not only provide an answer when traditional methods fail but also open up new opportunities and possibilities for political science research.

As an illustrative example throughout this chapter we use the effect of pre-election coalition signals by parties on strategic vote decisions and discuss three different experimental approaches designed to test this effect. Coalition governments are a common outcome in many multiparty systems, and voters might take possible coalitions after the next election into account at the ballot box. During campaigns, parties sometimes signal to voters the desirable and undesirable coalition partners. For instance, German parties often resort to explicit appeals for strategic voting in the form of a 'rental vote' (*Leihstimme*). Supporters of one of the two major parties, Christian Democrats (CDU/CSU) or Social Democrats (SPD), are asked to 'rent out' their vote in favor of the preferred small coalition partner when the latter is in danger of falling short of a minimum vote threshold (e.g. the Free Democrats). In case of such a failure, the major party will likely have no prospects to lead the next government, even when running strong. A more detailed motivation of our substantive research question is provided below. First, however, we elaborate our argument that experiments are a flexible tool for theory testing, discuss some advantages and disadvantages of experimental designs, and introduce the concept of experimental triangulation.

We start with the assumption that researchers want to test a theory. As textbooks instruct us, this requires a careful definition of the theoretical concepts, the derivation and specification of observable implications, and the selection of appropriate cases that allow the measurement of causes and effects (King et al. 1994, Gschwend and Schimmelfennig 2007). When selecting cases, researchers will often face the challenge that appropriate observable data is simply not available to adequately test a theory. Suppose we are interested in the effect of a particular contextual factor on individual behavior such as a coalition signal or other specific campaign messages. If such a message is sufficiently loud and clear, all informed voters will receive it. But how would we be able to determine if it had any effect? If the message was constant throughout the campaign, our key explanatory variable would lack variance. We have only data from this one election, a single case. Thus, all respondents in an election survey will have been exposed to the same message, and no respondent would have received an alternative, counterfactual message. It would not be possible to determine the impact of a constant message with any level of confidence. In fact, any political scientist interested in the effects of institutions and institutional rules on political behavior will almost certainly face a similar challenge.

What can be done in such a situation? If increasing the number of cases is not an option, crafting a clever experimental design can provide a methodological solution. Experiments are ideal for exactly this kind of situation because they enable the researcher to create the necessary variance. Guided by theory, the researcher can operationalize and manipulate the explanatory factor(s) in such a way that meaningful causal tests become possible. Experiments essentially create scenarios that represent different states of the world. By randomly assigning the manipulated explanatory factors to participants, we can make comparisons and estimate the causal effects. The differences (or lack thereof) between treatment and control groups will tell us whether participants react and behave as hypothesized.

Reducing the complexity of the real world to theoretically meaningful but often very narrow differences naturally raises the question of external validity. A simple manipulation does not represent reality as we experience it in everyday life, nor should it do that. The advantage of experiments is to submit hypotheses to causal tests, if necessary by breaking up complex causal chains into smaller steps that can be tested individually. Thus, experiments can systematically address what happens under theoretically relevant circumstances, even if they may not occur this way in the real world (e.g. Mook 1983). After successfully demonstrating the predictive value of a theory, researchers are well advised to address the external validity of their findings. This might require additional experiments or observational data from surveys and similar designs. In fact, the combination of complementary research designs might often be the best strategy.

Experiments are by no means a free lunch. They frequently require tough decisions. There are no cookbook recipes that tell us what to do and how to test a particular theory. Different experimental designs come with different advantages and disadvantages, and a researcher will have to decide what is most appropriate in a given situation. For example, a researcher who wants to rule out all confounding influences on her measures of causes and effects needs to fully control all aspects of the study by creating or inducing all key variables, including the preferences of participants. Any measure that relies on preexisting preferences is not fully controlled by the researcher and might introduce some confounding factor. At other times, however, it might make perfect sense to leverage participants' preexisting preferences, especially in realistic decision contexts. It would be futile to try to directly manipulate a powerful predisposition such as party identification. A simple

party label will automatically elicit strong reactions and beliefs. A smart experimental design will at minimum simply measure and control such powerful reactions but ideally take advantage of them and utilize them within the experimental design.

What is the best experimental design? The short answer is that it does not exist. Every researcher will have to decide on the most appropriate design to test a certain theory in a given context. If a single experiment cannot give a complete and satisfying answer, as it is frequently the case, more than one experiment might be the solution. We call such a research strategy *experimental triangulation*. Researchers vary the operationalization of key measures or the setup of the experiment in order to test different aspects and mechanisms of the hypothesized cause-and-effect relationship. Taken together, this set of experiments offers a more complete and valid explanation of the social phenomenon of interest.

The term triangulation is borrowed from celestial navigation where it indicates a technique to infer one's geodetic position from the measurement of different sights such as the sun and the horizon (a role taken over by satellites for modern GPS-based navigation). In the social sciences, the concept can be traced back to the idea of improving measurement by using different measures. More specifically, Campbell and Fiske (1959) proposed the multitrait-multimethod matrix to obtain more valid measures of traits. The first explicit reference to triangulation we are aware of was made by Webb et al. (1966):

> Once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. The most persuasive evidence comes through a triangulation of measurement processes. If a proposition can survive the onslaught of a series of imperfect measures, with all their irrelevant error, confidence should be placed in it. Of course, this confidence is increased by minimizing error in each instrument and by a reasonable belief in the different and divergent effects of the source of error. (p. 3).

The concept of triangulation has been extended beyond measurement in several ways. Denzin (1970) outlined various types of triangulation, among them the use of independent data sources (data triangulation), different researchers (investigator triangulation), and different research methodologies (method

triangulation). While not without criticism (e.g. Blaikie 1991), triangulation can be defined as a process in which different measurement strategies or sources of information validate each other and overcome their potential individual weaknesses to enhance the confidence in our conclusions.

Like multiple measures of a single concept, we can talk about triangulation with a multi-method approach when we devise independent tests of the same theory with different methods. If multiple but complementary theory tests come to similar conclusions, we have more confidence in the research findings. But as Mathison (1988) points out, different measures, methods, and sources might not always converge but rather offer inconsistent or even contradictory outcomes. A triangulation strategy consequently can lead to a much more complex and thorough understanding of a social phenomenon.

In this chapter, we elaborate how scholars can, within the same methodological paradigm, creatively leverage different experimental designs to triangulate their findings within the same research program. With multiple experiments, we can use the specific strengths of one particular experimental paradigm to address and compensate for the limitations of another experimental paradigm. The obvious advantage in contrast to, say, a regular multi-method approach, is that through experimental triangulation scholars do not have to compromise the strengths of experimental designs *per se* with the use of other methodologies to triangulate theory tests. While multi-method designs are of course still possible and even desirable, we argue that the particular strength of experimental triangulation is that it facilitates the use of experiments as a flexible tool to devise several independent tests of the same theory. Of course, different designs might sometimes lead to different answers, raising the question about how to evaluate and interpret such divergent results. We will return to this question in our conclusion.

2. Illustrative Example: Coalition Signals and Strategic Voting in Multiparty Systems

In first-past-the-post systems a strategic voter is typically defined as someone who cast his or her vote for a party other than the most preferred party because the former has a better chance of winning (Cox 1997, Fischer 2004). According to the theory of strategic voting, a strategic vote requires an instrumental motivation and

rational expectations about the outcome of the next election. According to this definition, it is necessarily insincere. At first glance, studying strategic voting in multiparty systems might seem to be a hopeless endeavor (but see Cox 1997). But more recently, several studies have offered evidence that strategic voting not only makes sense in PR systems but have offered supporting evidence as well (e.g. Abramson et al. 2010, Blais et al. 2006, Bargsted and Kedar 2009, Meffert and Gschwend 2010). These studies suggest that voters not only defect from marginal parties but have a variety of reasons to cast a strategic vote.

The theory of strategic voting assumes that voters cast their ballot in order to maximize their expected utility based on their party preferences and their expectations about the outcome of the next election (Cox 1997). With coalition governments, strategic voters must not only form expectations about the likelihood that parties win representation in parliament but also consider which coalitions are viable and likely. Based on these expectations, they can decide how to vote in order to best influence government formation, if only to influence the weight of each party in an almost certain coalition (Meffert and Gschwend 2007). Given the complexity of the decision task, it is likely that voters use simple heuristics such as coalition signals by parties to simplify the decision task. Especially coalition signals should help voters to narrow down the large number of theoretically possible coalitions to the relevant few.

At the individual level, strategic voting is typically studied with survey data from particular elections. The challenge to determine the effect of coalition signals on voting behavior is by now a familiar one: a single election usually does not provide much variation in the key independent variables, polls and coalition signals. Both tend to be fairly stable and consistent before elections, and every voter will receive more or less the same information. Consequently, it is not possible to determine with confidence that a strategic voter would have decided differently if the polls had suggested a different election outcome or if parties had offered different coalition signals. In order to overcome this lack of variance we turn to experimental designs. This strategy allows us to create theoretically relevant decision scenarios in and outside the laboratory that should either facilitate or inhibit strategic voting. We use experimental manipulations to create variance in the key explanatory factors, and the comparison of treatment and control groups allows us to directly test our hypotheses about coalition signals.

3. Coalition Signals in Three Experimental Designs

Testing the causal effects of coalition signals requires that coalition signals vary, either in terms of their presence or absence or in terms of their nature (valence), advocating (positive) or ruling out (negative) a specific coalition. The basic design and operationalization can follow a simple logic. By randomly assigning different versions of the coalition signal to participants, it is possible to determine whether signals have the hypothesized impact by comparing the key outcome variable for the different experimental groups. Experiments allow the systematic variation of coalition signals and measure their effect on randomly assigned groups.

Experiments can take many different shapes and forms. The settings can range from a tightly controlled lab environment over a real world field setting to (often representative) surveys, and methodological rules and standards differ by tradition (Morton and Williams 2010). Experiments in the economic tradition tend to confront participants with abstract, context-free, and transparent decision scenarios. The information available to participants might be incomplete, creating uncertainty, but it should never be deceptive or false. In order to rule out external and potentially confounding influences, preferences are induced and assigned by the experimenter and not based on existing preferences of participants. This gives the experimenter in economic experiments a very high degree of control. The abstract nature of these experiments and the induced preferences make it possible to assess the quality of decision making in a straightforward manner. Because the correct decision is known to the experimenter, it is very easy to determine good and bad or optimal and wrong decisions. Participants experience success and failure as monetary gains and losses.

Following these basic principles, we designed an economic experiment that presented participants with an abstract game with fictitious parties and induced party preferences in a laboratory setting (Meffert and Gschwend 2007). The coalition signals were operationalized as salient information but associated with high ambiguity and uncertainty. The quality of the decision was determined as a monetary payoff. Table 1 presents an overview of the key characteristics of our studies.

[Table 1 about here]

Psychological experiments, on the other hand, try to create realistic decision scenarios, not in terms of mundane realism, but in the sense that they rely on pre-existing individual preferences and differences and try to pose decision scenarios that

capture the attention and involvement of the participants (McDermott 2002). A key difference to economic experiments is the frequent use of concealment and deception for experimental manipulations. The information given to participants is optimized to create a convincing manipulation, not to provide objective and verifiable facts. From an ethical perspective, the use of concealment and deception makes it mandatory that participants are debriefed at the end of the study. Any misrepresentation of facts needs to be corrected.

Psychological experiments of electoral decision making rely frequently on fictitious scenarios in order to control the amount and content of information available to participants. However, it is very common to use existing parties and existing party preferences, relinquishing much more control than economic experiments. The psychological experiment described below went one step further by embedding it in two ongoing state election campaigns in Germany (Meffert and Gschwend, forthcoming). The decision scenario presented to participants was thus highly realistic, and most information provided to participants was taken from the actual party platforms. However, the experiment still took place in a laboratory setting with a convenience sample of student participants. The experiment used deception to operationalize and manipulate coalition signals and poll results. The manipulated information was embedded in a subtle and unobtrusive way in other campaign information. The goal was to create theoretically relevant decision scenarios that should (not) induce strategic voting. The key dependent variable was a hypothetical vote decision in the state election, but not tied to any monetary payoff or incentive (though participants received a fixed participation fee).

Laboratory experiments usually use convenience samples that pose a challenge to external validity and the generalization of the study results to the world outside. In this respect, cross-sectional surveys with a general population sample have a clear advantage over laboratory experiments, even if they fall short when assessing causal relationships. That said, it is sometimes possible to combine the advantages of randomized manipulations and control of laboratory experiments with the representative nature of general population surveys. If a manipulation can be included in a survey questionnaire, the combination of a randomized experimental manipulation with a representative population sample is a near perfect solution.

In the survey experiment described below, respondents were interviewed in a pre-election survey and confronted with four scenarios in the form of short vignettes,

in a randomized order. The vignettes presented respondents with different coalition signals and asked for any (hypothetical) changes in vote intentions. Needless to say, these respondents did not receive a financial incentive for participation or 'optimal' answers.

3.1 Economic Experiment

The main purpose of the economic experiment was a causal test of strategic voting in multiparty systems with proportional representation, minimum vote thresholds, and coalition governments under ideal conditions—all participations had an induced monetary incentive for strategic voting and no incentive for expressive or habitual voting (for details, see Meffert and Gschwend 2007). An important initial design decision was to use a decision scenario with four parties because three parties allow only for a trivial number of coalitions while five parties already lead to an (exponentially increasing) explosive number of coalitions and highly complex decision scenarios. The election scenario consisted of four parties (A to D) competing for the votes of 15 voters, distributed randomly in a two-dimensional space. Voters could maximize their expected utility by moving the location of next government as close as possible to their own location, compared or relative to the government location after a sincere vote for the preferred (closest) party. The reduced distance constituted the monetary payoff, while wrong decisions that moved the government further away from the voter location constituted a monetary loss. A voter decision (or government) is called optimal if no other party choice (government) leads to a higher payoff. The participant was the only swing voter while the other 14 simulated voters always supported their preferred party.

[Figure 1 about here]

The critical component of the decision scenarios was how coalition governments would be formed after an election. The procedure followed four sequential rules. The first and very obvious criterion was an *absolute majority* of seats in parliament for a single party. If no party had the support of a majority, a coalition government became necessary. The key rule was the *minimum distance* of two (or three) parties in the political space that reached an absolute majority. The following two rules were used to break any ties that might exist after the second rule. First, a two-party coalition would beat a three-party coalition (*minimum number of parties*),

and if this still could not resolve the tie, the coalition with the lower vote share would be formed (*minimum vote share*). If all four rules failed to produce a government, the election ended in a stalemate without any payoff. It is important to note that government formation was explicitly and entirely based on the electoral strength and proximity of the parties. Pre-electoral coalition signals played no part in government formation, and thus should *not* have played any role for voters. Participants were fully informed and familiar with the rules of government formation.

The experiment tested the influence of two critical information sources, polls and coalition signals, by manipulating their availability to voters. Polls were based on the actual distribution of the party preferences in a given election scenario and available with an 80 per cent probability. But even if not available, voters would still receive information about the relative size of the parties, whether it was a major (>25 per cent) or a minor (<25 per cent) party. The operationalization of the coalition signal was more difficult to implement. Ideally, an experimental manipulation is fully randomized and independent from other manipulated factors such as, in this case, the strength and location (or proximity) of the parties. If implemented this way, the signal would show two (or three) randomly chosen parties. However, a signal generated this way would frequently be meaningless, for example by displaying two small parties or two parties at the opposite ends of the political space. It would have no meaning and participants would not take it seriously. The signal had to be both plausible but uncertain, that is, sometimes providing ‘good’ information and sometimes ‘bad’ information—good in the sense that the coalition in the signal would indeed lead to a successful, optimal outcome while a bad signal would indicate a coalition that represents an unsuccessful election outcome. Consequently, the coalition signal was based on a simple decision rule: it showed the two parties closest to each other but required that at least one of the two parties was a major party. This rule essentially represents a simple heuristic for government formation that might or might not be successful. It is also based on information that was always available to participants: the distance of the parties in the political space and the approximate size of the parties (that is, at least one major party).

In about half of the randomly generated decision scenarios selected for the experiment, the signal showed the coalition that represented the optimal government for the voter. In the other half, it displayed a suboptimal government. Note that even if the signal shows the optimal government, these parties do not necessarily include the

party that the participant should vote for in order to produce this government. While the parties in the signal were thus determined by a simple rule, the visibility of the signal to participants was randomized with equal probability. Participants were only told that the signal shows parties that wish to form a coalition, not how the signal was generated. Because the coalition signal played no role in actual government formation, it should be irrelevant information for participants.

The results of the experiment, however, show that the signal did influence the decisions of the participants. Table 2 distinguishes between easy elections with an optimal coalition signal and difficult elections with a suboptimal signal as well as the availability of poll and signal information. If we take the decision scenarios without polls and signals as the baseline, participants were able to make optimal decisions in 51 per cent of the easy elections and 31 per cent of the difficult elections. The availability of polls increases the proportion of optimal decisions to 64 and 41 per cent, respectively. The impact of the signal, when no poll was available, is equally strong, but conditional on the quality. Good signals in easy elections increase the share of optimal decisions to 65 per cent while bad signals in difficult elections lower the share of optimal decisions to 22 per cent. If both the poll and signal are available, the share of optimal decisions in easy elections increased further, but only slightly, to 68 per cent. In difficult elections, the availability of a poll appears to have helped voters to counteract the bad signal. They made optimal decisions in 38 per cent of the elections. The results of the economic experiment suggest that even voters with a strategic (monetary) incentive tend to rely on coalition signals as a heuristic. If the signal is accurate, it can very well substitute for a poll, but if it is bad, voters who follow it tend to make the wrong decisions.

[Table 2 about here]

3.2 Psychological Experiment

The psychological experiment operationalized coalition signals in a highly realistic way. As before, the experiment focused on strategic voting and was conducted in a laboratory setting. However, it was embedded in two real, contemporaneous German state election campaigns in January 2006. The general design and procedure of the study involved exposure to campaign information about the five major parties, with information taken from actual election platforms of the

parties. Participants played the role of a voter and were instructed to inform themselves before the upcoming election. The information was presented on a computer-based information board that always showed six newspaper-style headlines with information. Clicking on a headline opened another window with the associated short article (see Meffert and Gschwend, forthcoming, for details).

The main purpose of the experiment was to test a specific version of strategic voting in PR systems with minimum vote thresholds, threshold insurance. Supporters of a major party might vote for the preferred junior coalition partner if the latter is in danger of falling short of the threshold. Previous research has shown mixed support in favor of such rental votes or *Leihstimmen* (e.g., Gschwend 2007, Pappi and Thurner 2002). At the same time, supporters of small parties that are fairly certain to fall short of the threshold should defect from their party and rather vote for another party that will affect government formation in a beneficial way. In order to test these assumptions and the role of polls and explicit coalition signals by parties, the study used the actual party preferences of the participants.

The manipulation of polls and coalition signals targeted specifically the preferred parties of each participant. At the beginning of an experimental session, participants indicated their party preferences by ranking the five most relevant parties, the two major parties Christian Democrats (CDU) and Social Democrats (SPD) and the three minor parties Free Democrats (FDP), the Greens (Die Grünen), and the Left Party (Die Linke/WASG). This ranking determined which parties were used for the subsequent manipulations. First, the highest ranked major party determined the assignment of a participant to one of two states, CDU supporters to Baden-Württemberg and SPD supporters to Rhineland-Palatinate. These parties were the respective incumbent parties in each state and both were expected to be re-elected by large margins. In other words, the expected winner in each election was held constant for all study participants. It should be noted that the study was conducted in the city of Mannheim, located right on the border between these two states, allowing for a fairly seamless assignment of participants to these different states.

Next, the most preferred small party was used for the poll and coalition signal manipulation. The poll manipulation varied the expected performance of the small party above and below the minimum vote threshold. The signal manipulation used the preferred major and minor party to either explicitly mention this coalition or avoid any reference to it. In short, the two most preferred parties of each participant were

used for manipulations in order to create standardized election scenarios, but the manipulations themselves, the closeness to the threshold and an explicit coalition signal, were randomized.

Participants were exposed to manipulated polls and signals in two ways during the ‘campaign.’ Participants were exposed to always six headlines on the information board that changed in a fixed interval of 45 seconds, whether or not participants clicked on and read any articles. Five headlines on each screen always represented the issue positions or candidates of the five parties (one headline for each party). The sixth headline covered either polls or other, fairly generic state information. In total, the 90 headlines and articles available to respondents covered 13 issues and two candidates for each party as well as five manipulated polls, five generic polls, and five state-specific but generic topics. The order was randomized.

After two screens with headlines, the campaign was interrupted for a pre-election poll that first asked participants to indicate their vote intention at that point, followed by a screen ostensibly showing the results of an actual state election poll (Figure 2). Participants saw a table with the manipulated numerical poll results on the left and a verbal summary (for numerically challenged participants) on the right. At the bottom were two statements attributed to the two preferred parties of each participant. Phrased in the style of newspaper headlines, they either mentioned a coalition or just stated typical campaign statements in response to the poll. Using the parties CDU and FDP as examples, the statements without signal read:

CDU: Poll confirms we are on the right track; Will fight for every vote
FDP: Campaign will be tough; Need to better motivate supporters

In the version with a coalition signal (as shown in Figure 2), the statement read instead:

CDU: Poll confirms we are on the right track; Hope for coalition with FDP
FDP: Campaign will be tough; Appeal for ‘rental votes’ by CDU supporters

Note that the first part of these statements was always identical and only the second part changed. All participants saw this screen and thus were guaranteed to be exposed to the signal manipulation.

[Figure 2 about here]

The second opportunity to encounter the manipulated information was as part of the headlines and articles on the information board. However, participants had to actively choose and read these five articles with manipulated poll and signal information. It does provide a hard behavioral measure of interest in and exposure to poll information. The five articles repeated the same poll and signal information from the pre-election poll discussed above. Each article focused on a different aspect but basically restated the same information. As a rule of thumb, one or two paragraphs restated the poll results and one paragraph discussed coalitions, either mentioning the explicit coalition signal or at a fairly unspecific level. In each experimental condition, every participant was exposed to the same information or content. Only the names of the parties changed according to the individual party preferences of each participant. In terms of programming, the party names were ‘variables’ in a text mask (which also included all the verbs associated with the parties because, grammatically, the Greens are a plural noun and require a different verb form than the singular nouns FDP and WASG).

The operationalization of polls and coalition signals in this experiment has the clear advantage of tapping the actual party preferences of the participants and using a real election campaign as decision scenario and backdrop. This clearly improves the external validity of the study but also imposes certain limitations. First, reality constrains the manipulation of polls and signals to a plausible range. For the polls, the winning major party in each state could not be changed, only the forecasts for the small parties could plausibly range from 4 to 10 per cent (with a minimum vote threshold of 5 per cent). The WASG was running for the first time in both states, creating some uncertainty about its strength. The only baseline salient to participants could have been the results of the previous general election several months earlier in which the three minor parties reached fairly similar and strong results (FDP: 9.8 per cent, Green Party: 8.1 per cent, Left Party/PDS/WASG: 8.7 per cent).

The coalition signal posed a bigger challenge. In both states, the FDP was the junior partner in the incumbent coalition and thus the designated coalition partner after the next election. In both states, however, the situation was more fluid and alternative coalitions could not be ruled out. In both states, the Greens were a plausible alternative coalition partner while the WASG was more or less ruled out by both major parties. Because the signal manipulation automatically used the preferred

parties of each participant, the signal could have shown fairly absurd combinations, in particular a coalition of the conservative CDU with the far-left WASG. This was judged to be an acceptable risk, correctly as it turned out, because such a party preference ranking was highly unlikely. Less serious but more difficult to solve was the fact that some signals would show the incumbent coalition while others would propose a new coalition. Thus, the coalition signals had to be phrased very carefully. They were attributed, for example, to ‘different politicians in both parties’ to make them plausible for any coalition, incumbent or not. The phrases used typical, sometimes off-the-record statements by politicians during real campaigns. Given this complexity, the whole experimental design was tested first in a large pilot study. This test was successful, but as a result, it became necessary to include another poll condition in the main experiment. The manipulation checks of poll and signal manipulations were successful as well, and post-study comments and feedback by participants indicated hardly any suspicion of the manipulated polls and coalition signals.

The results, however, brought some surprises. Only ten participants (or 7.5 per cent of participants in the close-poll conditions that were expected to induce strategic voting) could be classified as strategic voters, pre-empting a meaningful analysis of the effect of coalition signals on strategic voting. However, about a quarter of the participants did defect from their top-ranked party and voted ‘insincerely’ for some other party, independent of the poll manipulation. In a multivariate model predicting insincere voting, coalition signals have a modest positive impact, again suggesting that coalition signals do play a role in vote decisions. But compared to the strong signal effect in the economic experiment, the realistic but fairly subtle signal in the psychological experiment appears to have only a minor impact.

The small number of strategic voters can in part be explained by one of the key and necessary design features. Because the preferred large party was always the certain winner and never faced real competition or even trailed the opponent, this party was essentially removed from strategic considerations that might exist otherwise. Only a replication in other contexts would allow a test of strategic voting under such circumstances. Last but not least, the manipulation of coalition signals during a real election campaign carries a significant risk because real parties might make an announcement during data collection that might undermine the study purpose. In our case, this did not happen.

3.3 Survey Experiment

Experimental manipulations can also be included in representative population surveys, though with less control and the need for fairly obvious manipulations. Coalition signals are very well suited for this purpose because they merely require that survey respondents are exposed to them before the relevant questions. Thus, our third implementation of coalition signals is fairly straightforward. As part of a representative pre-election survey before the 2006 Austrian General Election, participants were exposed to four different vignettes of hypothetical coalition announcements by Austrian parties. As in the psychological experiment, a real election campaign as decision context and background always poses the acute risk that real events might interfere with the manipulations, such as a party making an unexpected coalition announcement. Unlike laboratory experiments with fictitious decision tasks, a survey that is several weeks in the field offers hardly any control over contextual factors and the study setting that might undermine the manipulated messages. Consequently, the coalition signals had to be phrased explicitly and transparently as hypothetical statements in order to work even in a changed setting.

In order to both avoid such surprises and to create sharply contrasting vignettes, the hypothetical coalitions always mixed and matched one of the two major Austrian parties, the conservative People's Party (ÖVP) or the Social Democrats (SPÖ), with one of the two smaller parties that were expected to perform very well in the election, the moderate but left-of-center Greens (Die Grünen) and the far-right and populist FPÖ (which incidentally was fairly explicit in ruling out any participation in government).

These vignettes were presented shortly after asking the standard question about vote intention. They were introduced by the statement that '[m]ost parties have not made a clear announcement about possible coalitions after the election' and followed by four vignettes, in randomized order:

> 'For which party would you vote if the Greens would clearly reject a coalition with the SPÖ and announce the intention to form a coalition with the ÖVP?'

> ‘For which party would you vote if the Greens would clearly reject a coalition with the ÖVP and announce the intention to form a coalition with the SPÖ?’
>
> ‘For which party would you vote if the FPÖ would drop its intention to not participate in any coalition and rather announce the intention to form a coalition with the ÖVP?’
>
> ‘For which party would you vote if the FPÖ would drop its intention to not participate in any coalition and rather announce the intention to form a coalition with the SPÖ?’

The response to each vignette was recorded with the same party list that was used for the standard vote intention question. This allows within-respondent comparisons of changes in (hypothetical) vote intentions.

Because the vignettes focus on specific parties but were given to all respondents, it is reasonable to expect effects primarily on those respondents who are directly affected by these coalition signals, in particular supporters of the Greens and the FPÖ. Table 3 gives a short illustration how respondents reacted to the vignettes. Among supporters of the Green Party, a signal in favor of the ÖVP and against the SPÖ led to a considerable drop of support while a signal in favor of the SPÖ did not change the support at all. The latter was the preferred coalition of a large majority of Green Party supporters. Among FPÖ supporters, however, any departure from the declared governmental abstinence, whether in favor of the ÖVP or the SPÖ, led to a drop of support for the FPÖ. In both cases, coalition signals affect the vote intentions of supporters. For the Greens, the coalition partner matters and the SPÖ is the clear favorite. For FPÖ supporters, government participation in itself leads to a drop of support, suggesting that at least some supporters see their vote as a protest vote against the mainstream parties. Even if the effects are again more limited, the third study once more supports the notion that coalition signals matter, in a real election and with a representative sample of voters.

[Table 3 about here]

4. Conclusion: Comparing and Evaluating the Results of Different Experiments

How do we know that coalition signals actually have an effect on voters' decision making? If we had merely used observational data, we would almost certainly have faced the problem that our key independent variable, coalition signals, would not have varied much in each of the election campaigns. We simply would not have the necessary variance for a meaningful test of our hypothesis. Instead, we used different experimental designs that allowed us to 'inject' variance by manipulating coalition signals in theoretically meaningful ways. This approach makes a test of the causal hypothesis possible and suggests that coalition signals matter, and not only for strategic voting.

The different operationalizations of coalition signals demonstrate that experiments are flexible tools to test causal relationships even if there is not enough variance in the key explanatory variable. Given that the lack of variance is a frequent problem for research questions in political science, researchers would be well advised to consider and adopt experimental strategies as well. It can not only overcome the limits of other designs but provide the opportunity to address new and seemingly intractable questions. And using an experimental triangulation strategy by employing different types of experiments can further enhance and strengthen our confidence in the findings. In our case, three experimental designs from different experimental traditions—economic, psychological, and survey research—have given us a mostly converging, sometimes inconsistent, but never contradictory pattern of results. Table 4 provides a brief summary.

[Table 4 about here]

The unambiguously good news is that no matter the type of experiment, coalition signals matter! We saw in the economic experiment (that deliberately induced in all participants a strategic mindset) that the manipulation of coalition signals was highly effective. Coalition signals facilitated strategic voting and emerged as a useful heuristic that simplified participants' decision task. But it is a risky heuristic because a given coalition signal might involve parties that are not the optimal vote choice for a participant. Thus, coalition signals can help but also lead voters astray if they trust them blindly.

In the psychological experiment (in which voters could follow either strategic or expressive motivations in a real election context), we found merely a marginal impact of coalition signals on participants' vote choice. Voters were more likely to defect from their top-ranked party and voted for some other party when coalition

signals were present. The fact that participants deserted their preferred party even if the polls did not indicate any instrumental benefit suggests that coalition signals affect not just strategic voting but have a more complex impact. It suggests that coalition signals are a simple heuristic for both strategic and merely insincere voters but might even elicit the expression of a genuine coalition preference. The experiment suggests that the investigation of coalition signals requires a closer look at coalition preferences as well.

The results of the survey experiment replicate and complement the results of the two previous studies. Coalition signals changed respondents vote intention systematically in our representative sample of Austrian voters. We find evidence for those effects not for all signals and on all respondents but primarily on those who are directly affected by the signals. In contrast to the two other studies where coalition signals were an unobtrusive facet of the information environment, the vignettes in the survey experiment explicitly linked the coalition signals with the vote intention. Thus, respondents could not even process this information heuristically. They were rather forced to explicitly and deliberately think about the consequences of different signals on their vote decision, leading to clear and observable shifts in vote intentions.

Our triangulation strategy with different types of experiments leverages the strength of each design to address the limitations of the others. For example, the economic experiment gave us full control over participants' preferences, the signal manipulation, and any contextual influences. In the survey experiment, our control was very weak because we had no influence over what happened in the actual campaign. On the other hand, the survey experiment used a real election and was based on a representative sample, giving it much higher generalizability than the convenience sample in our laboratory experiments.

In terms of internal validity, the psychological experiment falls somewhere in middle. The standardized decision scenarios and randomized manipulations certainly provide a high degree of internal validity, but it is rather difficult to find the hypothesized effects. Subtle manipulations met real and strong political preferences, severely limiting our ability to 'push' participants around.

For a pure theory test, our concern is more with internal than external validity. The fact that we can replicate the strong effects of coalition signals in the abstract economic experiment in weaker form with both a laboratory experiment and a survey experiment during real election campaigns gives us the confidence to conclude that

coalition signals are an important factor that requires more attention in future research.

How can we best assess and compare the different impact of coalition signals across very different experiments? We have two answers. First, it remains puzzling for us how to *directly* compare the size of the effects, and in fact it might even be a futile endeavor. These differences might merely be random, but it is a priori more likely that different types of experiments exert their own 'design effects' similar to so-called 'house effects' of different survey institutes that often produce different numbers even when surveying the same population at the same time. A third possibility is that the differences vary systematically with the different contexts in which they were conducted. Only replications with similar experiments in different contexts will allow us to answer this question. On a more positive note, the second answer is that the findings of all three experiments support and complement each other while indicating stronger and weaker effects under different conditions. This, after all, is the ultimate purpose of experimental triangulation.

To sum up, we argued that experiments are flexible tools for theory testing. Our results indicate that experiments are particularly useful in situations when key explanatory factors lack variation. This is a challenge we often face when designing a study. We have shown some of the strengths and weaknesses of different experimental designs, and the benefits of using an experimental triangulation strategy to both conduct conclusive causal tests of our theories and to generate a complementary and more generalizable pattern of findings. Our hope is that we have convinced our readers that despite all the difficulties and drawbacks, well-designed experiments offer new possibilities for interesting research in political science.

References

P.R. Abramson, J. H. Aldrich, A. Blais, M. Diamond, A. Diskin, I. H. Indridason, D. J. Lee and R. Levine (2010) 'Comparing Strategic Voting Under FPTP and PR', *Comparative Political Studies,* 43(1), 61-90.

M.A. Bargsted and O. Kedar (2009) 'Coalition-Targeted Duvergerian Voting: How Expectations Affect Voter Choice under Proportional Representation', *American Journal of Political Science,* 53 (2), 307-23.

N.W.H. Blaikie (1991) 'A Critique of the Use of Triangulation in Social Research', *Quality & Quantity,* 25, 115-36.

A. Blais, J. H. Aldrich, I. H. Indridason and R. Levine (2006) 'Do Voters Vote for Government Coalitions? Testing Downs' Pessimistic Conclusion', *Party Politics,* 12 (6), 691-705.

G.W. Cox (1997) *Making Votes Count. Strategic Coordination in the World's Electoral Systems* (Cambridge: Cambridge University Press).

D.T. Campbell and D. W. Fiske (1959) 'Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix', *Psychological Bulletin,* 56(2), 81-105.

N.K. Denzin (1970) *The Research Act in Sociology: A Theoretical Introduction to Sociological Methods* (Chicago: Aldine).

A. Downs (1957) *An Economic Theory of Democracy (*New York: Harper Collins).

S.D. Fisher (2004) 'Definition and measurement of tactical voting: the role of rational choice', *British Journal of Political Science,* 34 (1), 152-66.

T. Gschwend (2007) 'Ticket-Splitting and Strategic Voting under Mixed Electoral Rules: Evidence from Germany', *European Journal of Political Research,* 46 (1), 1-23.

T. Gschwend and F. Schimmelfennig (2007) *Research Design in Political Science. How to Practice What They Preach?* (Houndmills, Basingstoke: Palgrave).

G. King, R.O. Keohane and S. Verba (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, NJ: Princeton University Press).

S. Mathison (1988) 'Why Triangulate?' *Educational Researcher,* 17(2), 13-17.

R. McDermott (2002) 'Experimental Methodology in Political Science', *Political Analysis,* 10, 325-42.

M.F. Meffert and T. Gschwend (2007) *Strategic Voting under Proportional Representation and Coalition Governments: A Simulation and Laboratory Experiment* (SFB 504 Working Paper No. 07-55, University of Mannheim).

M.F. Meffert and T. Gschwend (2010) ‘Strategic Coalition Voting: Evidence from Austria’, *Electoral Studies*, 29(3), 339-49.

M.F. Meffert and T. Gschwend (forthcoming) ‘Polls, Coalition Signals, and Strategic Voting: An Experimental Investigation of Perceptions and Effects’, *European Journal of Political Research*.

D.G. Mook (1983) ‘In Defense of External Invalidity’, *American Psychologist*, 38, 379-87.

R.B. Morton and K.C. Williams (2010) *Experimental Political Science and the Study of Causality: From Nature to the Lab* (Cambridge: Cambridge University Press).

F.U. Pappi and P. W. Thurner (2002) ‘Electoral Behaviour in a Two-Vote System: Incentives for Ticket Splitting in German Bundestag Elections’, *European Journal of Political Research*, 41 (1), 207-32.

E.J. Webb, D. T. Campbell, R. D. Schwartz and L. Sechrest (1966) *Unobtrusive Measures: Nonreactive Research in the Social Sciences* (Chicago: Rand McNally).

Table 1: Key Characteristics of Studies

Key Aspect	**Economic Lab Experiment**	**Psychological Lab Experiment**	**Survey Experiment**
Context	Abstract Game	Real Campaign	Real Campaign
Party Preferences	Induced (no ties)	Measured (ties possible)	Measured (ties possible)
Coalition Signals	Salient & Transparent (uncertain)	Subtle & Unobtrusive (realistic)	Salient & Transparent (hypothetical)
Vote Decisions	Monetary Payoff (optimal)	Hypothetical Vote Decision	Hypothetical Vote Decision
Sample	Convenience Sample (Students)	Convenience Sample (Students)	Representative Sample

Table 2: Share of Optimal Decisions by Election Difficulty and Available Information

Scenarios	**No Info** % (BSE) N	**Poll Only** % (BSE) N	**Signal Only** % (BSE) N	**Poll & Signal** % (BSE) N
Easy Elections	51.7 (3.1) 269	64.2 (1.5) 1097	64.8 (2.8) 301	67.7 (1.4) 1123
Difficult Elections	30.9 (2.3) 408	40.8 (1.2) 1651	21.8 (2.1) 427	37.6 (1.2) 1699

Note: Entries are proportions, with bootstrapped standard errors in parentheses and the number of decisions in each cell. The number of decisions varies due to the random assignment of poll and signal manipulation, the former with unequal probability.

Table 3: Vote Intention for Preferred Party of Green Party and FPÖ Supporters

	Vote Intention for Preferred Party (PP)		
Preferred Party	**Initial** % (SE)	**Vignette with ÖVP-PP Signal** % (SE)	**Vignette with SPÖ-PP Signal** % (SE)
Greens (n = 308)	65.9 (2.7)	53.6 (2.8)	65.6 (2.7)
FPÖ (n = 86)	62.7 (5.2)	51.2 (5.4)	51.2 (5.4)

Note: Entries represent the proportions of Green Party or FPÖ supporters who intend to vote for their preferred party in each condition, with standard errors in parentheses. The preferred party is defined as the party rated highest among all parties.

Table 4: Key Results of Studies

Key Aspects	**Economic Lab Experiment**	**Psychological Lab Experiment**	**Survey Experiment**
Coalition Signals	Highly Effective	Marginal Effect	Conditional Effect
Interpretation	Useful but Risky Heuristic	Non-Strategic Heuristic	Deliberate Decision

Figure 1: Game Screen of Economic Experiment

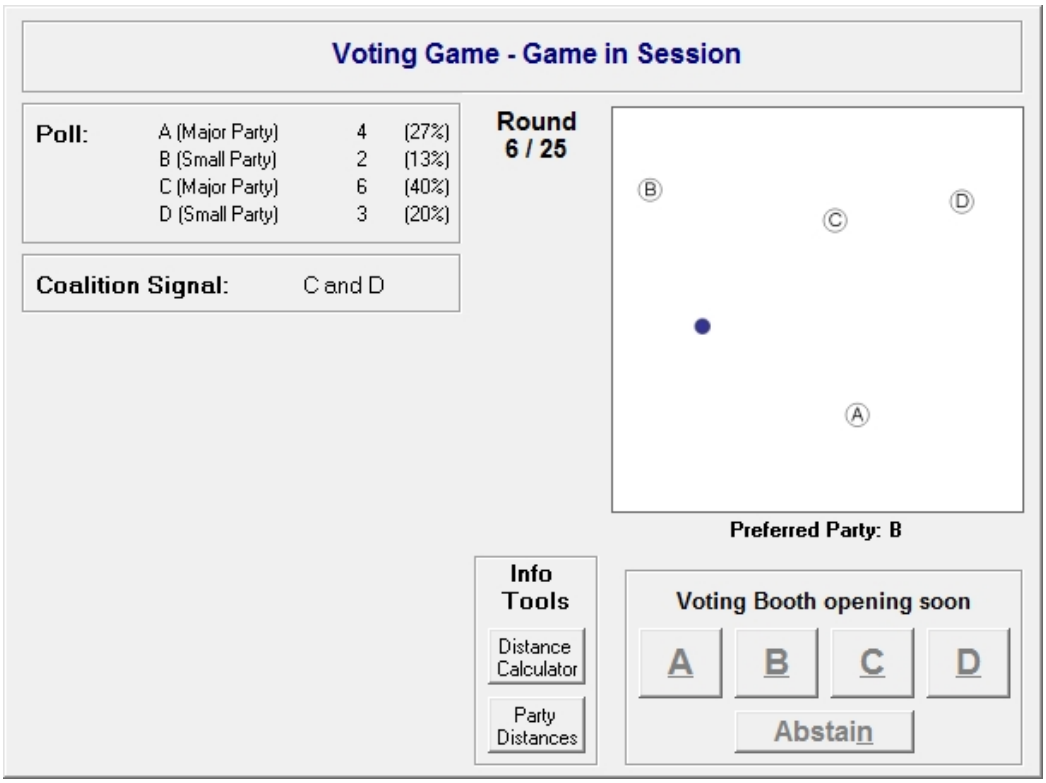

Figure 2: Poll Results Screen of Psychological Experiment