Common Knowledge:

Elicited Priors in Political Science

by

Sarah Beth Bouchat

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Political Science)

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

Date of final oral examination: 8 May 2017

The dissertation is approved by the following members of the Final Oral Committee:

Alexander Tahk, Assistant Professor, Political Science (Chair)
Scott Gehlbach, Professor, Political Science
Melanie Manion, Professor, Political Science, Duke University
Rikhil Bhavnani, Assistant Professor, Political Science
Ian Coxhead, Professor, Agricultural and Applied Economics
Edmund Malesky, Associate Professor, Political Science, Duke University

Abstract

This dissertation examines the application of elicited priors in political science through three related essays. The first essay introduces a method for aggregating elicited priors that bridges the qualitative/quantitative divide, while the second and third essays apply this method, illustrating how it can facilitate broadening the notion of "expertise" used in prior elicitation, thereby accessing different formats and venues of elicitation, overcoming practical modeling challenges by increasing prior precision, and expanding the use of prior elicitation.

The first essay examines the use of elicited priors in scientific literature and evaluates their limited social science applications. A core challenge to the adoption of elicited-priors approaches in social science is the divergence of opinions that elicitation and analysis must incorporate. Current methods for aggregating elicited priors from multiple experts—pooling and averaging—do not adequately incorporate diverse expert perspectives. This essay proposes a Dirichlet Process clustering method to overcome this challenge and illustrates its effectiveness using data from Jackman and Western (1994).

The second essay, co-authored with Alexander Tahk, explores and interrogates the notion of "expertise" in prior elicitation. Using the 2016 U.S. national election results for validation, this essay presents evidence that "elite" and representative expert samples can perform comparably in a roulette elicitation framework. Results from both samples, separately and delineated by demographics, show that while more educated and knowledgeable respondents provide more accurate estimates of 2016 vote share, the Dirichlet clustering approach applied in this chapter aggregates divergent assessments to provide an accurate overall estimate.

The final essay applies the Dirichlet clustering technique to legislative participation data from Myanmar. This application addresses a practical problem of

quasi-perfect separation arising from sparse data in a zero-inflated negative binomial model, but also illustrates the novel use of text-based elicitation from newspaper sources. Just as the second essay demonstrated that educational credentials need not imply expertise for elicitation, this essay also broadens the basis for eliciting and applying previous knowledge to current scholarship.

Acknowledgements

This dissertation would not have been possible to complete without significant assistance from advisors, colleagues, and friends. I benefitted from sources of support too numerous to name individually here, but I want to acknowledge several whose contributions throughout my graduate career have been particularly impactful.

First, I have had the good fortune to receive training, advice, support, and mentorship from not one, but three main advisors. Melanie Manion was an excellent academic and personal guide as I entered the PhD program, and has continued her engagement even after leaving the UW. I remain eternally grateful for her insightful critiques, irreverent commentary, and brutal honesty (and also the 5 a.m. deadlines that kept me on track). Scott Gehlbach graciously took me on in Melanie's absence, spending many hours giving advice that ran the gamut from grammar to professionalization, and inviting me to participate in Set Z despite a continued lack of understanding of the post-Soviet sphere and an erstwhile interest in personal finance politics. Melanie and Scott's expertise in authoritarian politics, incisive engagement, attention to detail, and willingness to overcome my interpersonal shortcomings have improved my scholarship as well as my life in general.

Without a doubt, though, surviving these last few years of graduate school and seeing this dissertation project through was only possible because of the encouragement and dedication of time, effort, and understanding from my chair, Alex Tahk. As I transitioned toward a more methodologically focused dissertation project, Alex quickly became my de facto advisor—long before he would ever get credit. In addition to spending multiple hours a week in directed readings indulging my curiosity about everything from Bayes to machine learning and back, Alex spent an incredible amount of time helping me theoretically shape and write

code for this dissertation project. There is a Burmese saying, ချက်ဆို နားခွက်က မီးတောက် : *with a "click," a spark is made.* Alex embodies exactly this idea: just like a flint-lock gun can spark with the squeeze of a trigger, he immediately understands concepts in their entirety with only the very slight suggestion or outline of an idea.[1] Without Alex's insights and support, I could never have transformed a vague notion about elicited priors into a fully fledged dissertation project, and without his dedication of countless office hours, I certainly would never have finished. I owe Alex an immense debt of gratitude and would definitely not be where I am today, personally or professionally, without his mentorship and guidance.

I am also incredibly thankful for thoughtful comments and critiques from the other members of my dissertation committee—Rikhil Bhavnani, Eddy Malesky, and Ian Coxhead—as well as the engagement of faculty and colleagues both within and outside my department, including John Ahlquist, Ryan Bakker, Justin Grimmer, Simon Jackman, Jon Pevehouse, Eleanor Powell, Brandon Stewart, Rocío Titiunik, David Weimer, the 2016 EITM cohort and faculty, as well as participants at APSA and MPSA, the Society for Political Methodology, the New Faces in Political Methodology conference, and the University of Wisconsin Comparative Politics Colloquium and Models and Data group. Even before beginning my tenure at Wisconsin, I benefitted immensely from the help and support of Jim Alt, Jenny Mansbridge (who is responsible for at least one of the three copies of Strunk & White I have been gifted over the years), and Tim McCarthy—without them, I would not have had the opportunity to attend graduate school. Only with the help of such an amazingly intelligent, thoughtful, and talented group of mentors and colleagues was I able to transition from political theory to comparative politics to methodology and appreciate every step.

[1] This is especially lucky for me because sometimes the nexus of math and words is a black hole where ideas go to die.

In addition to fantastic support in political science, I benefitted greatly from the support of those in the Southeast Asia sphere. Financial and logistical support from the University of Wisconsin Southeast Asia Center, and from Michael Cullinane in particular, as well as comments and questions from participants at the 2016 Association for Asian Studies conference, the 2016 International Burma Studies Conference, and the 2015 Burma/Myanmar in Transition Conference at Lingnan University greatly improved my Burma-related research. In addition, I am particularly lucky to have had Burmese language guidance from saya(ma)s Saw Tun, Mya Hla, Than Than Win, Tharaphi Than, Kyaw Win Tun, and May Zin Nyein Ein.

Many friends leant their critical eyes and moral support to me and this dissertation project. At the University of Wisconsin, I am especially grateful to Annie Anderson, Hannah Chapman, Desiree Desierto, José Luis Enríquez Chiñas, Rachel Jacobs, Anna Oltman, Rachel Schwartz, Emily Sellars, Samantha Vortherms, and the members of Set Z and Chinese Politics Workshop. Outside of the UW, Jennie Ikuta has offered equally necessary sage advice and sparkle hearts as a grad school survivor.

In addition to all of these academic sources of assistance, I am lucky to have had several sources of personal support. Beyond their direct support for my education, my parents instilled values of determination and self-discipline from an early age, without which it would have been impossible for me to make it through this many years of graduate school. I am also thankful for the love and assistance from my partner, Vanessa Ng, over the last ten years. Her willingness to listen to nearly endless streams of consciousness, to serve as a test audience for new ideas and less-than-polished writing, and to tolerate long hours never leaving the house in service of a deadline have been invaluable.

Finally, I am particularly grateful for the friendship and unconditional personal and professional support I have received from Nat Olin. From our first meeting at prospective student weekend, through prelims and several years as roommates, Nat has offered innumerable pep talks, honest feedback, around-the-clock debugging assistance, and meaningful companionship even after leaving Madison.

I am indebted to all of you more than I can express.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Hayek describes the main threat to establishing economic order as "the fact that knowledge of the circumstances of which we must make use never exists in concentrated or integrated form, but solely as the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess" (Hayek 1945, 519). Indeed, the diffusion and divergence of knowledge presents a problem not just for economic systems, but for political ones, and for their study in a social scientific framework. Broadly, this dissertation explores ways of collecting, organizing, and aggregating diffuse information to improve the study of political questions. More specifically, it investigates the incorporation of prior information from experts into Bayesian analyses of political phenomena through an *elicited priors* framework.

1.1 Elicited Priors & Bayesian Analysis

This dissertation examines the use of elicited priors across disciplines and proposes methods to increase the applicability of elicited-priors approaches, particularly in social science. In Bayesian analyses, prior probability distributions reflect beliefs about quantities or processes of interest that, together with observed data and the likelihood function, inform posterior estimates of quantities. *Elicited priors* are a subset of prior distributions that specifically seek to include information, substantive knowledge, or area expertise into statistical analyses. These prior probability distributions, rather than reflecting the individual researcher's prior beliefs or reflecting more general prior probability distributions chosen for the convenience of conjugacy, attempt to accurately reflect the state of knowledge on a given quantity of interest. To do so, these prior probability distributions are *elicited* from "experts" either in person through interviews, focus groups, or surveys, or via assessments based on previously published work. Broadly speaking, the elicitation process attempts to instantiate the area knowledge of experts about some quantity or variable about which a researcher has uncertainty in a fully specified probability distribution. The specificity of elicited priors can often address practical estimation challenges while also providing a theoretically satisfying connection between past, current, and future research agendas. More practically, as O'Hagan et al. (2006) note, elicitation is a common practice in engineering and business projects, where the "uniqueness" of a given endeavor requires drawing on expert assessments in order to leverage their analogous expertise and quantify uncertainty (9). Likewise, in addition to Bayesian applications, experimental design is another area in which elicitation from experts in order to improve design (9).

Elicitation targets *epistemic uncertainty*—uncertainty due to imperfect knowledge—rather than *aleatory uncertainty*—uncertainty arising from randomness in a process (10). Unlike frequency probabilities, epistemic uncertainty refers to unrepeatable quantities or events, and therefore requires assessment based on *personal* or *subjective* probability that reflects a "degree of belief" (11). Criticisms of this type of approach tend to take one of two forms, either practically or theoretically based. As a practical matter, psychological literature gives reason for concern about the accuracy of elicitation procedures and the resulting quality of elicited priors. Psychological heuristics such as "anchoring" as well as limitations in the precise expression of probabilities mean that elicitation, as with many measurement exercises, also has some degree of error. As O'Hagan et al. describe, "[it] is also important to recognise [sic] that experts construct probability judgments in response to the stimulus of questioning: their probabilities are not pre-formed values simply waiting to be expressed" (O'Hagan et al. 2006, 20). On the theoretical side, moreover, skepticism of the "subjectivity" of elicited priors threatens their widespread application, at least in part because of the derision with which "subjectivity" is viewed in scientific circles. In Bayesian applications, informative priors, of which elicited priors are a subset, are often seen as overly "influential" with respect to the data under examination, whereas some view uninformative/less informative or reference priors to have a more "objective" characteristic.

1.2 A Theory of Scientific Progress

A Kuhnian framework for scientific progress highlights the potential advantages of elicited-priors approaches. Kuhn's characterization of scientific progress, as well as his questioning of the objectivity of scientific truth, align well with the notion that elicited priors can document and directly incorporate past discoveries into contemporary analyses (Kuhn 1996). In particular, the perspective adopted in this dissertation is one in which elicited-priors approaches should seek not just to serve as an echo-chamber of a few established opinions, but rather to include a variety of diverse perspectives characteristic of social science research—much like the competing paradigms Kuhn describes. Alongside the transparency that elicited-priors approaches afford to the benefit of scientific progress,[1] this dissertation seeks to articulate technologies for aggregating and incorporating diverse perspectives into statistical analyses through the use of priors. Enabling elicited priors to encompass several competing or overlapping perspectives of a single phenomenon, in particular, should aid in the advancement of knowledge accumulation.

Still, competing philosophies of science remain skeptical of approaches like the use of elicited priors, which include beliefs that may not be "falsifiable." As Deborah Mayo and Aris Spanos describe,

> ... for the most part, scientists wish to resist relativistic, fuzzy, or postmodern turns; should they find themselves needing to reflect in a general way on how to distinguish science from pseudoscience, genuine tests from ad hoc methods, or objective from subjective standards in inquiry, they are likely to look to some of the classical philosophical representatives.... Notably, the Popperian requirement that our theories and hypotheses be testable and falsifiable is widely regarded to contain important insights about responsible science and objectivity; indeed, discussions of genuine versus ad hoc methods seem invariably to come back to Popper's requirement.... However, limiting scientific inference to deductive falsification without any positive account for warranting the reliability of data and hypotheses is too distant from day-to-day progress in science. (Mayo and Spanos 2010, 2)

Even "subjective" approaches like the use of elicited priors, then, may aid researchers in overcoming practical challenges in the investigation of social scientific questions.

1.3 Objectivity vs. Subjectivity

A key advantage of Bayesian analysis is the ability to transparently and concretely incorporate prior information into inference, so that the understanding of phenom-

[1] Gelman (2009) argues that Bayesian inference and Bayesian model-checking themselves act as both normal science and scientific revolutions in Kuhn's definition.

ena changes through a process of continually including new information (Moses and Knutsen 2012, 274-275). Precisely how and whether specific prior information is incorporated, however, is a matter of subjective judgment.

At a fundamental level, apart from a frequentist/Bayesian divide or an objective/subjective divide, statistical modeling engages humans in a process of decision-making about what factors are relevant and to what degree, and this process must be, to some extent, subjective. In the realm of Bayesian statistics, furthermore, many varieties of objective and subjective approaches to the use of priors exist. Objective, or minimally informative, priors can take a variety of forms: Jeffreys priors (Jeffreys 1961), reference priors (Bernardo 1979; Berger and Bernardo 1992), maximum entropy priors, minimal description or message length priors (Wallace and Dowe 1999), invariance priors, matching priors (Datta and Mukerjee 2004), admissible priors, etc. (Berger 2004). Indeed, objectivist Bayes is often considered a remaining hope to better integrate Bayesian and frequentist approaches (Bayarri and Berger 2004). Subjective priors, conversely, broadly describe a set of priors incorporating specific beliefs from experts, previous studies, or other sources of information. While Bayesian statistics is broadly considered "subjective" in its interpretation or probabilities as degrees of belief, therefore, significant variation exists along an objective–subjective spectrum in terms of the types of priors one might leverage in a statistical analysis. As a catch-all phrase, *weakly informative* priors describe those priors that simultaneously foreground the data while also placing sufficient restrictions on analysis as to produce reasonable results. By contrast, the intention behind eliciting priors is often to incorporate positive beliefs alongside data with minimal restraint because these beliefs are considered reflective of the state of information during which an analysis is conducted.

Taken to an extreme, a subjective Bayesian approach applied here might suggest that there cannot be a "best" way to proceed in the use of prior elicitation or aggregation. Even without assuming an overly extreme position, however, there are several risks, to adopting a subjective approach. Specifically, as Berger (2004) notes, "[one] only has limited time to elicit models and priors from the experts in a problem, and usually it is most efficient to use the available expert time for modeling, not for prior elicitation" (7). Likewise, there is a danger, he says, that "with the subjective Bayes approach, the already scarce time must be used to train [experts] in elicitation, for otherwise the priors obtained can be quite bad" (7). Moreover, elicitation can include "systemic" bias—"the almost universal tendency to underestimate the actual amount of uncertainty about unknowns"—while only providing a few characterizing moments of a prior distribution, rather than its full specification (8). These criticisms, however, pertain mostly to the practical considerations of eliciting priors in an effective manner, rather than to the theoretical concern for whether elicitation improves or detracts from statistical analyses. In a sense, then, to the extent that this project endeavors to increase the transparency and rigor with which priors are elicited, it can be considered an attempt to generate some degree of "objectivity" in the use of subjective priors (Gelman and Hennig 2017).

1.4 Dissertation Overview

This dissertation evaluates the use of elicited priors, and proposes methods for increasing their applicability, through three essays. Broadly, the dissertation seeks to interrogate how best to apply elicited priors in the social sciences, and, having argued for a method that facilitates their application, to detail new understandings of "expertise" as well as new elicitation techniques that can expand the use of elicited priors. In Chapter 2, I provide a critical examination of the literature currently implementing elicited-priors approaches. The assumption of an elicited-priors approach as applied in the sciences is that an "expert" has specific training and credentials, whereas in social science settings an "expert" may have lived experience or access to preferential information. As a result, priors elicited in these settings may be very diverse. I argue that the currently dominant aggregation method for expert priors—averaging—does a disservice to the diversity of perspectives present in social science research. To ameliorate this problem and improve the prospects of applying elicited-priors approaches in future research, I propose a method for aggregating priors based upon Dirichlet Process clustering. I illustrate this method using data from Jackman and Western (1994) and a series of simulated "divergent" priors regarding real data on unionization and comparative political economy.

Chapter 3, co-authored with Alexander Tahk, proceeds from Chapter 2 to investigate the distinction of "expert" and the nature of "expertise," using the outcome of the 2016 U.S. national election to validate survey-based roulette elicitation from both a "mass" (nationally representative) sample of U.S. individuals and from an "elite" (Ph.D. students in political science) sample. This paper evaluates the tension in psychological and political science literature on forecasting, between the "wisdom of crowds" approach that seeks to capitalize on the accuracy of forecasting given large samples of beliefs, and the push to identify "superforecasters," or individuals or modeling frameworks that provide better-than-average predictions by allocating greater weight to more "accurate" individuals. Applying the clustering technique introduced in Chapter 2, and evaluating it against averaging for both the mass and elite survey samples, Chapter 3 seeks to provide an alternative theoretical and practical framework for approaching "expertise." Rather than the atomistic perspective espoused in the psychological literature, in which individuals with particular skills or credentials have better-than-average predictive capacities, the theoretical basis for this chapter instead considers that knowledge is interdependent across individuals, and that while some individuals may have more information than others, in general a better and fuller perspective of social phenomena is achieved when aggregating the differing perspectives of individuals or groups. Practically, this distinction is implemented using the Dirichlet clustering framework, and considering the mass and elite samples both as separate groups, as a pooled set, and according to demographic subgroups.

In Chapter 4, I apply the clustering technique introduced in Chapter 2 to a substantive question in comparative politics: what explains variation in legislative behavior in autocracies? Specifically, I investigate variation in legislative participation in Myanmar's lower house of parliament, the Pyithu Hluttaw. The paper adopts a zero-inflated negative binomial (ZINB) modeling framework to capture the high incidence of non-participation ("zeros") in Myanmar's parliament, and disaggregates "always zeros" from "sometimes zeros" while measuring differential incentives arising from career objectives and policy preferences. Chapter 4 also builds upon the analysis of expertise in Chapter 3 by introducing a novel method for text-based elicitation and evaluating it against survey-based roulette elicitation from three Myanmar experts. Accessing diverse expertise, particularly in challenging developing or authoritarian contexts, requires expanding the pool of "experts" and leveraging information not just in interview contexts but also through written works as a way to systematically apply previous knowledge to current modeling challenges. Priors are elicited from a variety of newspaper sources covering Myanmar's parliament. These priors are applied within the ZINB modeling strategy in order to add information in a sparse data context, and specifically to overcome challenges related to quasi-perfect separation. Chapter 4 illustrates that not only does cluster-based aggregation aid in incorporating elicited priors into complex modeling designs, text-based elicitation can expand the notion of "expertise" and broaden the set of questions for which elicited priors are a valid approach.

I conclude in Chapter 5 by reviewing the main arguments of the dissertation, and by discussing extensions to each of the three essays that can improve and refine their results. In addition, I discuss several other domains for expansion in the study of elicited priors, and the implications for these additional studies on the use of elicited-priors approaches in future social science research.

Chapter 2

Engaging Experts: Dealing with Divergent Elicited Priors in Political Science

2.1 Introduction & Motivation

Priors over parameters constitute one of the most visible manifestations of the theoretical distinctions between Bayesian and frequentist statistics. The information contained within the specified prior, however, will depend upon the knowledge available to a given researcher. An elicited prior seeks to leverage the substantive knowledge of area experts, whether through interviews or published research, in order to improve the accuracy of posterior estimates. Elicited priors as a concept have been detailed since the early work of Savage (1971), Shuford and Brown (1975), Kadane and Wolfson (1998), and Garthwaite, Kadane, and O'Hagan (2005). Formally, Gill and Walker (2005, 841) discusses the elicited prior "as a means of drawing information from subject-area experts with the goal of constructing a probability structure that reflects their specific qualitative knowledge, and perhaps experiential intuition, about the studied effects." These aims and applications for elicited priors seem both reasonable and desirable, yet the 10 years since the publication of

Gill and Walker's article have seen few instances of researchers, especially in political science, adopting elicited priors as a way to incorporate qualitative knowledge into quantitative research.

This paper argues that the limited application of elicited priors is due to both an unclear prescription for implementing the method and a limited vision for its application. The following sections seek to explore the range of studies from other disciplines that use elicited priors in order to understand both how they are used in practice and how they might be better used to realize the vision of qualitative and quantitative synergy that Gill and Walker present. An empirical section also outlines a specific process for undertaking work using elicited priors. This is illustrated with reference to applications that expand Gill and Walker's scope to encompass data-poor contexts in particular, where the limitations on quantitative data especially benefit from an elicited priors approach.

The concept of elicited priors shows particular promise in combining qualitative and quantitative analytical approaches. A typical modeling approach would rely only on the researcher's assessment of the appropriate array of variables, model type and structure, and method. Eliciting priors from those with expertise in an area offers an opportunity to gather more input for each of these choices from those who have substantive knowledge but are not directly involved in the analysis. This approach has the added benefit, therefore, of maintaining transparency about assumptions in the modeling process. An elicited prior requires documentation of at least the qualifications of the source, whether in-person or published, and can therefore ease the process of replication and progression in similar veins of research.

Despite these advantages, the current work on elicited priors lacks guidance on how to elicit said priors, how to reconcile diverging beliefs, and under what circumstances or for what purposes they would be most useful. These omissions limit the use of elicited priors in the social sciences, and especially in political science. In particular, while the prospect of eliciting priors from experts with both experiential and scholarly insights adds to the appeal of the method, it is exactly the process of identifying "experts," eliciting their views in ways that accurately represent their assessments while being useful in the modeling process, and reconciling the diverging opinions of experts that makes elicited priors difficult to apply. This is especially true in data-poor contexts, where the prior is simultaneously of greater importance to the modeling process and more contentious to construct as a result of the constrained data environment environment. For example, when soliciting input from government-sponsored and opposition-sponsored experts in authoritarian settings, there are no guidelines inherent to using elicited priors as presented by Gill et al. that would allow a researcher to adjudicate between strongly diverging opinions or assess their credibility. Likewise, no guidance is given regarding precisely *how* to elicit a prior—a process that could greatly influence outcomes.

This paper seeks to address these omissions and suggest ways in which elicited priors could be usefully applied. The following section culls techniques from the literature using elicited priors to create suggestions for their broader application in social science. An empirical section follows, discussing strategies for implementing elicited priors and their implications. This section illustrates a new method for aggregating divergent priors using data about unionization, and later demonstrates the promise and limitations of this approach using simulated data. Finally, the conclusions section will offer some next steps for refining these approaches and for broader application of elicited priors to a wider array of social scientific problems.

2.1.1 Priors in Bayesian Analysis

Often, researchers seeking to remain agnostic in statistical analysis employ relatively diffuse priors. These "uninformative" priors are, however, literally informative in at least two senses. First, they directly influence the conclusions to be drawn from the modeling process. This is especially true in cases of low data quality or quantity: diffuse priors limit the ability of the researcher to leverage the model structure in order to draw inferences about empirical phenomena. As Andrew Gelman writes, "[with] well-identified parameters and large sample sizes, reasonable choices of prior distributions will have minor effects on posterior inferences. ... If the sample size is small, or available data provide only indirect information about the parameters of interest, the prior distribution becomes more important" (Gelman 2002, 1634). While other approaches (e.g., clustering and estimating hierarchical models) often seek to address these issues, these statements still justify overarching concern with specifying reasonable priors where possible, with special consideration for instances with poor data. Specifying uninformative priors, furthermore, threatens the trajectory of scientific inquiry as studies build on each other, as Gill and Witko note: "It is also important to observe that the overwhelming proportion of prior distributions specified in published Bayesian social science work still avoids using reasonably informed priors, which unfortunately hurts the steady accumulation and progression of scientific knowledge" (Gill and Witko 2013, 462).

Second, these priors represent a strong positive claim that no useable information about a given parameter θ exists with which to specify a more appropriate or precise prior. More generally, this approach reflects a normative position that favors the omission rather than the careful specification of assumptions throughout the modeling process.

An alternative approach would instead seek to take advantage of the significant knowledge accumulated across fields of study in order to select appropriate, informed priors. From the perspective of the researcher, this approach could represent an alternative agnosticism that does not require reliance on their own knowledge or opinions, but rather documents the acquisition of information from expert sources. This notion effectively underlies the method of "eliciting" priors: whether through interviews with area experts or reference to published works, a researcher can develop appropriate priors for an unknown θ that improve the prospects for modeling in a Bayesian framework.

2.2 Literature Review: Elicited Priors across Disciplines

The notion of elicited priors is not new, yet scholarship examining or employing the technique remains limited. Searching for "elicited prior" in the Web of Science database yields fewer than 50 relevant articles, most of which have substantive foci in biology or psychology rather than political science. In defining the origins of elicited priors, Gill and Walker (2005) reference previous terms such as "community of priors"—which incorporate opinions of both affirming and skeptical experts—and delineate elicited priors into a variety of types, including clinical priors, skeptical and enthusiastic priors, reference priors, etc. (843). Gill and Walker (2005, 844) summarize the three phases of research using elicited priors explained by Spetzler and Staël von Holstein (1975) as the deterministic, probabilistic, and informational phases. The deterministic phase, encompassing variable and expert selection, has costs, but is generally perceived as less challenging than the probabilistic phase, during which priors are actually elicited. Aside from determining data sources, the primary question addressed in the first stage is instead from how many experts one should elicit priors (Gill and Walker 2005, 844).

The informational phase follows, in which elicited responses are tested, evaluated, and scaled for consistency. Gill and Walker (2005) concur with Spetzler and Staël von Holstein (1975) that the methods of eliciting priors ("p-methods," "v-methods," "pv-methods," etc.) present the greatest challenge, but concerns about how best to elicit priors from experts do not explain the very limited uptake of elicited prior methods overall in the social sciences. Rather, I suggest that more careful attention to the selection of "experts" and the calibration of responses opens possibilities for applying this method to new areas of research, specifically in authoritarian and data-poor settings. For example, rather than focusing on the purely statistical problem of how *many* experts should be selected, emphasizing the question of *who* is considered an expert and in what way opens new possibilities for identifying expertise and knowledge from which to generate priors.

2.2.1 Theoretical Contributions to Elicited Priors

While the use of elicited priors in scientific fields far surpasses its use in social science domains, its application remains somewhat limited by remaining skepticism of the "subjectivity" of the Bayesian approach more generally (Lele and Das 2000; Wang and Zhou 2009). Elicited priors, rather than overly diffuse or benchmark "objective" priors, have the potential to provide leverage in low-data circumstances and challenging modeling contexts, but also to substantively impact the quality of estimation in a positive way. In fact, Datta and Ghosh (1991) suggest that the elicited prior represents the "true" prior—presumably that, if sufficient expertise were available and applied, the resulting prior would be accurate. Still, problems remain with processes for engaging even trained individuals in statistical assessments. As Lin, Lin, and Raghubir (2004) note, people are susceptible to biases resulting from self-positivity, controllability of negative events, and especially order of elicitation—where previous questions are used to form assessments for later ones. Likewise, practically speaking, prior elicitation confronts the same issues as typical Bayesian analysis, where, for example, specifying priors for continuous parameters proves difficult since "[doing] so would require infinitely many prior probability judgments" (Dey and Birmiwal 1994).

In a sense, the lack of identifiable expertise in low-data settings, rather than giving rise to the implementation of diffuse priors, should give rise to a concerted effort to model our ignorance—a task Zaffalon (2005) identifies as a persistent challenge in Bayesian analysis. Accounting for prior ignorance and incomplete data, as Zaffalon explains, means that the true model of ignorance requires accounting for "all possible states of knowledge" (1005). One of the key related areas of inquiry in elicited priors is how to handle uncertainty or error within the prior itself. In general, studies posit an elicited prior π_0 in which there may be some "ϵ contamination" or error (Sivaganesan 1993).

Among these scientific papers, however, opinions about how best to elicit priors vary. Lele and Das (2000), for example, argue in favor of directly soliciting guess values:

> It is important to recognize that the concept of a prior probability distribution on the parameters of a statistical model is a statistical construct that is hard for most scientists to visualize. It is more natural for an expert to think in terms of the process under study and not in terms of statistical distributions over a parameter space. The sensible approach, then is to ask the expert to provide guess values for observable data, not a prior probability distribution. (466)

In Lele and Das (2000)'s case, this suggests a hierarchical modeling strategy to combine experts' insights with the observed data under study.

The problem confronted by Lele and Das (2000)—sparse data in a spatial context but the potential for "soft" data in the form of expert opinion—mirrors the difficulty of conducting analyses in authoritarian or low-data situations. Lele and Das (2000) provide an analytical framework for addressing precisely this problem of "misleading" expert priors in their paper examining spatial hierarchical models with elicited priors. In particular, the authors want to account for the dependence between observed data values and elicited prior values when experts are asked for value estimates rather than providing complete prior distributions (468). "This dependence reflects the credibility of the expert," they argue, while noting that "[in] this way, even a misleading expert opinion can be informative" because "if we find that data elicited from an expert are negatively correlated with real data, this is useful information that can be used to suitably adjust our inferences" (468). To do this, Lele and Das imagine that an expert gives an opinion $\bar{E}$ about $\bar{Y}$ drawn from the distribution $f(\bar{y};\theta)$ where the θ parameters are unknown. This opinion $\bar{E}$ is itself distributed $g(\bar{e} \mid y;\eta)$, with η indicating the dependence between $\bar{Y}$ and $\bar{E}$. This leads Lele and Das to call η an "honesty parameter," measuring the "credibility" of an expert. In principle, however, such a parameter captures both the ability of the expert to think and express opinions statistically as well as the potential strategic misrepresentation of information and expertise.

2.2.2 Empirical Applications of Elicited Priors

A variety of disciplines, although primarily within the natural and biological sciences, have utilized elicited priors techniques for research. These studies offer a multitude of insights about how to implement an elicited priors approach, but also highlight the ways in which current practice in other disciplines is not currently well-suited to the study of authoritarian or low-data contexts.

Morris et al. (2013), for example, investigate the effects of vesicoamniotic shunts on lower urinary tract issues for fetuses where they have only 31 female subjects with singleton pregnancies and leverage elicited priors from 52 pediatric nephrologists, pediatric urologists, and fetal medicine specialists (1500). The experts polled in the study largely agreed on the potential effects of treatment, although the authors note that it is "problematic" that these experts' opinions did not align with what is current common clinical practice. Their responses were pooled and averaged to create a prior distribution (1501). This study provides one of several examples in which expert opinions are equally weighted in the implementation of the prior. For social science purposes, however, two issues remain:

(1) How should expert opinions be weighted or aggregated if disagreement occurs?

(2) How should expertise coming from practical experience be reconciled, either with differing experience or with educated or credentialed expertise?

While studies like Morris et al. (2013) illustrate the disjuncture between expert opinion and practice, they do not resolve the concerns facing researchers who wish to use these methods in other social scientific settings.

The procedure for eliciting priors from experts also dialogues with these concerns for how expert priors might be adequately reconciled and combined. As Wheeler et al. (2014) describes, there are many possible ways of "eliciting" prior information. This can be done with previously published work or historical data as well as with interviews of living experts; the primary binding constraint is that the expert has not directly observed the data under current study (678). If the expert has directly observed the data in question, their prior is likely to reproduce these data rather than providing additional information that can aid in inference and analysis. Despite this constraint, there are still many options for how to implement prior elicitation, including eliciting priors about regression coefficients; the distribution of the dependent variable conditioned on fixed values of covariates; quantiles of the predicted dependent variable distribution, etc. (Wheeler et al. 2014, 678).

Some of these methods present greater challenges than others in implementation. For example, eliciting beliefs regarding regression coefficients and their variance can prove difficult even for educated experts (Gill and Witko 2013, 462), as providing accurate estimates of statistical uncertainty of one's beliefs is a particular challenge (Albert et al. 2012, 504). A variety of methods also exist for quantile based direct elicitation of priors (Dey and Liu 2007). Regardless, as Wheeler et al. (2014) notes, "the incorporation of such information into the Bayesian modelling framework aligns with the philosophy of the scientific method, where knowledge that is available before collective data (prior) is used along with the observed current data (likelihood) to inform what we know now (posterior)" (678).

The example offered by Daponte, Kadane, and Wolfson (1997) underscores this point. The authors use elicited priors to project the Iraqi Kurdish population from 1977–1990, and describe three critical advantages to conducting their demographic study in a Bayesian framework. First and foremost, they note that a Bayesian approach can promote communication among demographic researchers: "Making one's beliefs explicit using probability distributions allows other demographers to observe exactly how one views the sources of uncertainty in the phenomenon. Others can then know on what they agree or disagree. The reasons given for particular probability distributions can be an important source of insight" (1256). The authors also note that making these projects explicit in the form of probability distributions enhances their usability in a variety of applications. Finally, they distinguish the Bayesian approach and its advantages from more traditional methods: "Classical models either include or exclude a parameter about which no prior is expressed, which is often equivalent to expressing certainty about its value. Using probability distributions permits one to express states of knowledge in between these two alternatives" (Daponte, Kadane, and Wolfson 1997, 1256). Much like other social scientific applications, as the authors describe, this study of the Iraqi Kurds "lacks high-quality data" and reflects incomplete information —conditions that align it in particular with the study of authoritarian regimes or other low-information environments in political science (1257). Each of these dimensions highlights the critical applicability of an elicited priors approach in an authoritarian or low-data context, and the significance of elicited priors as a source of information to adapt and improve statistical inference. How to integrate the priors offered by experts as part of a study, however, remains unaddressed by this work.

Albert et al. (2012), on the other hand, tackle the issue of how best to aggregate elicited priors directly, proposing a hierarchical modeling approach to account for multiple sources of variation and the potential lack of independence between experts' assessments (503). For illustrative purposes, the authors define a sampling model containing observation $X \sim P_\theta$ where θ is an unknown parameter with prior π that is an "elaboration on a parametric family" such that $\pi \in \{\pi_\gamma, \gamma \in \Gamma\}$. γ is then estimated from the elicited priors (504). It is plausibly the case, as the authors note, that each expert polled provides a different γ, necessitating a procedure to combine these divergent priors. In the authors' review, pooling and averaging predominate as methods for combining these differing priors, where averaging "emphasizes the consensus on elicited quantities" and is advantageous in its "simplicity," but at the same time can "understate variation by ignoring uncertainty" and/or "mis-represent [sic] multiple modes" (Albert et al. 2012, 504). Pooling methods—whether linear or logarithmic—attempt to overcome some of these deficiencies by encompassing all values in a way that can be construed as an additive or multiplicative mixture (504). These divergent priors can then be combined using weights w_ℓ, such as by $\sum_{\ell=1}^{L} w_\ell \pi_{\gamma\ell}$ where ℓ indexes the individual expert whose prior π_γ is being combined with others (504). Devising these weights, however, as the authors report, has primarily been based on "p-values for evaluating how well expert assessments on seed variables align with empirical results" (504). This has the disadvantage, as the authors note, of embracing the "diversity" among the elicited priors without offering a notion of consensus or how individual experts might diverge from that consensus. More directly related to the study of authoritarian regimes and experts, however, this method of weighting expert opinions is intended to resolve discrepancies in experts' ability to offer statistical assessments or measures of uncertainty. In a situation of incomplete information both within the data and among the experts, assessing an individual expert's prior in alignment with the data may less appropriately capture their insights into the underlying data-generating process or motivations of actors captured within the data. Weighting expert opinions in this way, rather than understanding as well their potential incentives to offer misleading information or ability to only offer insights into one part of the "truth," means that existing methods for combining expert opinions are not well suited to use in low-data contexts or the study of authoritarianism in particular.

The method that Albert et al. (2012) propose essentially establishes a hyperprior on the unknown parameter θ. This process treats elicited information as "data" in the construction of a prior for the eventual analysis of interest. The prior probability distribution on θ, given by D_{elicit} gives rise to the following:

$$\pi(\gamma \mid D_{elicit}) \propto f(D_{elicit} \mid \gamma)\pi_0(\gamma)$$

The construction of the distribution of $\pi(\gamma \mid D_{elicit})$ from pooling requires a joint likelihood of expert opinions in order to account for dependence. The resulting model of the expert priors treats prior opinions as being sorted into classes J where members of the same class have the same distribution (Albert et al. 2012, 508). In this the authors acknowledge that smaller groups may represent the divergent opinion of a set of experts who are less represented in the population, for example, or who are less reliable, and the solution they propose is to assume a higher variance parameter for the distribution of a group j with those characteristics. This approach, while including important information or knowledge about the experts into the modeling exercise, relies on assumptions about the validity of the statements made by members of smaller groups of experts, which may not reflect a priori knowledge of experts in authoritarian settings. Likewise, this conception of groups of expert opinions is predicated on an implicit notion that some of these elicitations are accurate whereas others are flawed, rather than a conception that each contains some aspect of the true information.

O'Leary et al. (2009) reinforce that the elicitation method itself should reflect knowledge about the experts who might provide information; that is, it should take into account difficulties in engaging with experts as well as difficulties with experts providing accurate information. "Selection of an elicitation method," they write, "is determined by several issues. These include the expert's knowledge of statistics, their mapping skills, time available, access to experts and funding. The chosen elicitation method should balance the expert's knowledge of statistics and mapping with the output required" (O'Leary et al. 2009, 396). These considerations serve as the basis for three different elicitation methods the article discusses, each with respect to a dataset concerning rock wallaby population estimates. The choice of elicitation tool, the type of elicitation (e.g., indirect P-method, direct, etc.), and the distribution form of the β coefficient primarily distinguish between these methods (388). This distinction is useful in highlighting that many methodological choices already exist in terms of adapting an elicted priors approach to a particular research question and context, but none specifically address the problem of how to aggregate priors across experts beyond the method of weighting with p-values for alignment between the elicited prior and the data that were discussed previously.

These problems notwithstanding, a precedent does already exist within political science for engaging "experts" beyond the credentialed class often sought after by scientists. Kendall, Nannicini, and Trebbi (2015) elicit multivariate distributions from voters reflecting their beliefs about incumbents' valence and ideology. This engagement of "experts" from a more general populous can allow the ultimate prior distribution used in the analysis of interest to reflect the diversity of information about the subject. Likewise, Small (2008) acknowledges that "it may be appropriate in many cases to elicit probabilities from different experts for different parts of the model" (1302). Likewise, Bakker (2009) utilizes elicited priors to enable experts to refine political party placements on a left–right spectrum. The broader application of an elicited priors approach to important problems in political science, however, rides on the ability of the approach to be adapted to other contexts and questions. The next section discusses how in particular this approach should be beneficial for the study of authoritarian political processes, and where its current practice falls short in offering clear signs of applicability.

2.3 Empirical Applications to Authoritarian and Low-Data Contexts

The aforementioned studies often engage expert opinion for priors in order to leverage the greater precision this lends to the analysis. An analogous application in political science relates to questions concerning authoritarian regimes or developing country contexts where low data and data manipulation/falsification are present. Bayesian analysis is often most appropriate for small-n studies, and more informative priors from elicitation can aid in resolving pathological problems arising from small data or complex modeling structures. This area therefore seems like a straightforward one in which to apply an elicited priors approach, but the current literature has several shortcomings that prevent the direct importation of elicited priors into a social science domain—especially where the study of authoritarian regimes or low-data contexts is concerned.

First, while some of the above studies offer more detailed accounts of how to elicit priors for scientific applications, these accounts have not been translated to a social science setting. This would be particularly important when expanding the notion of "expertise." Elicited priors are commonly thought to come from "experts" in a given field. While this may be a sufficient distinction in fields like medicine, where credentials come through schooling and certification, it is less clear for social science applications. When trying to understand regime behavior under authoritarianism, for example, one could conceivably ask professors who study the subject, bureaucrats who work under those conditions, activists who oppose such a system, or average citizens who live under authoritarian rule. Each of these groups has some degree of "expertise" to offer, but each would likely also have varying levels of statistical literacy to align with the protocols developed for eliciting priors in scientific contexts.

A second, and related, problem is the issue of diverging priors. The common procedure in the scientific literature, although not often discussed in detail, is to average elicited priors or to use a "consensus prior" after having elicited information from a series of experts. These aggregation rules intentionally eliminate or deemphasize information that diverges from the most common or popular prior. This may make sense for a panel of credentialed experts in a scientific field: for example, a researcher may be interested in the "state of the field" where the effect of a medication on prognosis is concerned. This rule makes less sense, however, in a circumstance like the one described previously, where the notion of expertise has been expanded to encompass several valid sources of information from a social scientific perspective. Particularly in authoritarian regimes, these different individuals may have widely varying perspectives on the issue at hand. These differing perspectives could be a result of different experiences, different relationships to the ruling regime or party (co-opted or oppositional), or different access to information in what is likely to be an environment rife with incomplete information.

A comprehensive analysis seeking to leverage this expertise nonetheless needs more guidance on how to incorporate and assess these divergent priors. Averaging the priors would discard potentially significant information, and arriving at a "consensus" could prove theoretically problematic. For example, imagine there are two individuals whose priors roughly overlap or correspond, but a third whose prior is significantly distinct. Should the researcher treat the two that correspond as the "consensus" or should the researcher be skeptical that those individuals have reproduced some sort of "party line" that intentionally obscures the underlying statistical relationship, whereas the third individual is telling the "truth"? A more ecumenical approach would seek to incorporate each of these, with the understanding that each potentially contains some part of an overall "truth."

Particularly when discussing social and political phenomena, “experts” may have strongly divergent opinions about the relationship between covariates and outcomes that, unlike in some scientific fields, cannot be resolved immediately by appealing to the data. Some of these experts may be more reliable than others in the sense of being able to provide statistical assessments, but in authoritarian contexts in particular, situations may arise in which experts have differing but not apparently false assessments of statistical likelihood as a function of either their desire to selectively provide (or misrepresent) information or their differing perceptions of truth caused by skewed information environments in authoritarian regimes.

In the following sections, I present an approach for addressing these concerns by providing an aggregation tool that can synthesize divergent expert opinions for Bayesian analysis. This approach does not seek to improve on methods for providing internal, statistical consistency of elicited priors, but rather seeks to deal with the divergence in potential elicited priors and the possible propensity for experts in authoritarian contexts to misrepresent the relationship between covariates and outcomes.

2.3.1 A Dirichlet-Based Method for Elicited Priors

In this section, I propose an approach for resolving the dilemma of prior elicitation and aggregation in social scientific contexts, and especially in studies of authoritarianism. Following this discussion, I provide examples of other uses for Dirichlet approaches as well as an example of an empirical circumstance where the approach could be applied with respect to elicited priors.

The method proposed in this paper adapts a Dirichlet process in order to resolve the issue of aggregating divergent elicited priors before implementing a Bayesian analysis. This Dirichlet-based method allows for multiple possible aggregation rules for elicited priors, enabling each of the priors to be "mixed" in a single distribution, and varying the concentration parameter of the Dirichlet distribution to seek a final prior that emphasizes greater consensus, equally weights all opinions, averages over opinions, etc. As Neal (2000) describes, "[mixtures] with a countably infinite number of components can reasonably be handled in a Bayesian framework by employing a prior distribution for mixing proportions, such as a Dirichlet process, that leads to a few of these components dominating" (249). This flexible framework for establishing prior distributions that carry substantial significance in a low data context is aimed to improve the integration of qualitative knowledge into quantitative assessments of particularly challenging contexts and questions. This approach facilitates a greater number of "new" kinds of experts serving as subjects for prior elicitation. Incorporating these differing kinds of expertise both enables better use of the information and expertise that already exists in authoritarian and other low-data contexts, as well as offers an opportunity for better engagement with qualitative researchers, who have long utilized this type of expertise in their own work.

The Dirichlet distribution is defined as

$$\frac{1}{B(\alpha)\prod_{i=1}^{K} x_i^{\alpha_i - 1}}$$

where α specifies a concentration parameter for the K "clusters" of the distribution. The Dirichlet as a mixture model of mixtures can be thought to play a "partitioning" role for clusters of observations. In the Chinese restaurant process example illustrating a Dirichlet, a customer entering the restaurant can be seated at a table already occupied by some diners, or at a new table. The probability of being seated at a given table corresponds to the distribution of occupants at one table relative to the others.

Applying this process to the elicited priors case, each elicited prior is separated into its constituent components: the proposed value for the dependent variable y as well as the values of independent variables x that corresponded to the y value. These proto "distributions" are treated as data for the purposes of the Dirichlet process. The Dirichlet will dynamically create clusters of observations, essentially finding moments of consensus since the population of these clusters is weighted by the existence of other data to occupy the cluster. Through an updating process to create these clusters, the Dirichlet process will eventually provide posteriors corresponding to each expert from whom a prior was elicited. These posteriors can then be used to create a prior distribution for the eventual analysis of the data of interest.

This framework is flexible, however, in terms of the specification and in terms of how input is weighted. An explicitly symmetric Dirichlet, which can be expressed as:

$$\frac{\Gamma(\alpha K)}{\Gamma(\alpha)^K}\prod_{i=1}^{K} x_i^{\alpha-1}$$

would denote that all components of α are of equal value (equal weighting). Likewise, how the posterior information from the Dirichlet process is transformed for use as a prior in the data analysis leaves room for defining different aggregation rules: averaging, deemphasizing extreme values, deemphasizing consensus, etc. This process of influencing the aggregation itself reflects an implicit hyperprior by the researcher placed on the priors, and the aggregated prior, from the experts used for elicitation, specifying how much weight to place on different experts' opinions or on different components of the aggregation. Note, however, that while this flexibility is an advantage for extending this method, the initial application assumes that the researcher operates from a position of ignorance. That is, this method can work in circumstances when the researcher does not have their own beliefs about the relative credibility of experts' opinions but instead wishes to incorporate a variety of perspectives. The approach as applied in this paper is agnostic with respect to how much weight certain types of prior beliefs should have; the ability to apply a hyperprior would be most useful in circumstances under which a researcher learned new information about the statistical relationship in question after having elicited priors from experts, but did not elicit priors from additional experts that could be incorporated into the Dirichlet clustering process itself. For example, one could suppose that the researcher discovered an older reference text with a compelling argument to be included in the analysis, or a newly published work on the subject under study became available, but its author could not be reached for prior elicitation. The researcher could potentially leverage this new information in the form of a hyperprior on the aggregated elicited priors while documenting their reasons for doing so. At the same time, the assumption that the researcher operates from a position of ignorance in the initial application discussed here bears repeating: without having previously identified what types of sources would be more or less credible, the researcher cannot and does not unfairly "place their finger on the scale" of the analysis by using the proposed Dirichlet clustering method. Rather, the researcher potentially stands to learn from this method as it highlights whether a diverse pool of experts have been targeted for elicitation (i.e., whether a large or small number of clusters exist that reflect differences of opinion). This transparency about the diversity of beliefs present in the analysis following the clustering process is a distinct advantage of this approach to aggregation. Not only does it benefit the current research by illustrating potential biases, it also improves the ability of future researchers to build upon this work by way of identifying additional or more diverse beliefs or sources of data to incorporate.

This proposed method of aggregation adds significant value relative to the existing methods of pooling or averaging. In addition to being better equipped to deal with the potentially divergent priors elicited in political contexts relative to medical or biological fields of study, this method provides a formal and transparent way to incorporate the researchers' beliefs about the credibility of a source, and to weigh the credibility of sources against each other and incorporate them into the final prior. This method also has the distinct potential to engage scholars across the qualitative-quantitative divide, as it does not resolve issues of expert selection but instead requires relying on the experience and knowledge of qualitative scholars who have previous familiarity with the case at hand to identify "expert" individuals and to evaluate their credibility.

2.3.2 Example: Latent Dirichlet Allocation

A common application of Dirichlet processes is in Latent Dirichlet Allocation (LDA). This process provides a model of a corpus, or body of textual works, where each document comprises a random mixture of latent "topics" and topics are distributions of words (Blei, Ng, and Jordan 2003, 996). For the basic LDA model with documents w in corpus D, as described by Blei, Ng, and Jordan, you select $N \sim \mathtt{Poisson}(\xi)$, $\theta \sim \mathtt{Dirichlet}(\alpha)$, and for each N words in w_n, you choose a topic $z_n \sim \mathtt{Multinomial}(\theta)$ and a word w_n from $p(w_n \mid z_n, \beta)$ (996). In this case the mix over latent topics θ is given with a Dirichlet having a concentration parameter of α. The ultimate allocation process seeks to understand how documents, which vary according to their words, can be sorted into a variety of topics. This is analogous to the mixture and allocation process that would occur in the elicited priors application of the Dirichlet process, where latent states correspond to experts' hidden dimensions of consensus. That is, the elicited priors are analogous to documents in the LDA example, which are comprised of components used to elicit them (e.g., across coefficients, according to quantiles, etc.). Experts offering these priors may have been subject to the same kinds of constraints (e.g., career or family concerns) or had access to or limitations from the same kinds of information (e.g., censorship, governing versus governed classes) that would shape the prior they gave. Allocating these into latent classes using the Dirichlet allows the researcher to examine the groupings present in the elicitations in order to better weight and incorporate the contributions of each individual expert.

2.4 Dirichlet Clustering Visualized

One of the key contributions of this method is to allow researchers seeking to utilize expertise from a diverse and perhaps divided set of individuals to be able to aggregate and incorporate that divergent expertise into their prior distributions. The Dirichlet-based approach allows for substantively similar elicited priors to be clustered for more efficient estimation and a more representative final prior. Adequately capturing this diversity within the Dirichlet, however, depends on several parameters that ultimately come from the initial research design, including how many experts to engage and how many questions to ask and/or datapoints to collect for each covariate used in the final analysis.

The figures that follow illustrate how well the Dirichlet approach identifies clusters among priors as a function of the number of covariates and the number of questions asked. As these figures illustrate, the clustering process performs best when receiving enough, but not too much, information about experts' priors. For example, it performs equally well in returning the true clustering with 20 experts versus 10, as long as there were 10 questions for 10 covariates. Slightly greater uncertainty is introduced, however, when a larger number of questions are asked relative to fewer covariates, and even more uncertainty appears when fewer questions are asked relative to the number of covariates. In general, maintaining the shape and scale parameters throughout the sampling process provides approximately 80% accuracy regardless of the number of experts, questions, and covariates input.

Clustering samples are visualized in the following figures (see Appendix for further discussion of these visualizations). The square on the left-hand side of each pair indicates the "true" membership of experts to clusters, while the square on the right-hand side indicates the estimated probability that experts share a cluster. Green squares indicate that experts share a cluster, whereas purple squares indicate that experts do not share a cluster. Shades in between indicate uncertainty. In each square figure, reading from the bottom left corner to the top right corner indicates into how many total clusters the "experts" are allocated. For example, in Figure 2.4, 20 experts are allocated into 4 unique clusters, each of which is a distinct green square as seen in the right-hand side plot. In the right-hand side plot, for example, cell (1,1) is shaded green because expert 1 shares a cluster with herself. Figure 1 illustrates that experts 1 and 20 are each in their own clusters, while experts 2–11 share a cluster and experts 12–19 share a cluster. The same general pattern of clustering is evident in the left-hand side plot of Figure 1, although there is greater uncertainty about which experts share the larger two "true" clusters, while the estimation remains fairly certain about experts 1 and 20.

In Figure 2.4, meanwhile, 10 experts are more neatly allocated into two clusters where a relatively large number of questions are asked about a smaller number of covariates. As before, the plot on the right-hand side indicates that experts are "truly" in two clusters: one containing experts 1–7 and the other containing experts

8–10. The estimated probabilities are less certain: they accurately capture the cluster with experts 8–10, but assign some probability to experts 2 and 7 each having their own clusters, while expert 1 would share a cluster with experts 3, 4, 5, and 6.

Because the Dirichlet Process allocates experts to clusters using a "rich get richer" principle—that each new expert is more likely to be assigned to an extant cluster than to initiate a new cluster—the amount of information distinguishing the opinions of experts plays a critical role in cluster differentiation. Too many questions relative to covariates may generate artificial uncertainty because of the perceived distinction between opinions arising from small differences in answers, while too few questions or covariates may not provide sufficient information to definitively allocate an expert.

Figure 2.1: Cluster Allocation with 20 experts, 10 questions, 10 covariates

Figure 2.2: Cluster Allocation with 10 experts, 20 questions, 5 covariates

2.5 Dirichlet Process Aggregation Applied: Western and Jackman (1994)

In order to illustrate the flexibility of this approach in incorporating multiple divergent priors, I take up the analysis conducted by Bruce Western and Simon Jackman in their 1994 paper, "Bayesian Inference for Comparative Research." In the paper, the authors highlight the many benefits of Bayesian analysis for comparative politics research. In particular, they emphasize that the use of priors to encapsulate differing ideas about the state of the world and combine them with data to draw inferences formalizes a process already undertaken in comparative politics research, albeit less transparently and with less ability to directly adjudicate between competing views. The authors illustrate this argument by drawing on a then-recent debate between Michael Wallerstein and John Stephens concerning the most important factors giving rise to unionization in advanced industrialized democracies.

Western and Jackman's aim is very similar to my project in this paper: namely, to emphasize the importance of priors as instantiations of knowledge, and to highlight the challenge of dealing with differing perspectives even within a prior probability framework. In the context of the Wallerstein/Stephens debate, this concern is applied to the study of union density—union members as a percent of the labor force. While Wallerstein argued for the size of the labor force as the most critical determinant, Stephens favored industrial concentration as an explanation. As Western and Jackman note, these two variables are nearly perfectly negatively correlated (Jackman and Western 1994, 416). Furthermore, the sample of interest —20 industrialized nations in a single year—both suggests against a frequentist approach and increases the challenge of the collinearity between the favored explanatory variables. Drawing on this debate, and with reference to outside sources, Western and Jackman construct plausible prior means and variances for Stephens and Wallerstein for each of three explanatory variables relative to unionization: left government, log labor force size, and economic concentration. Each of these prior means and the corresponding precisions are represented in Figure 2.3 below.

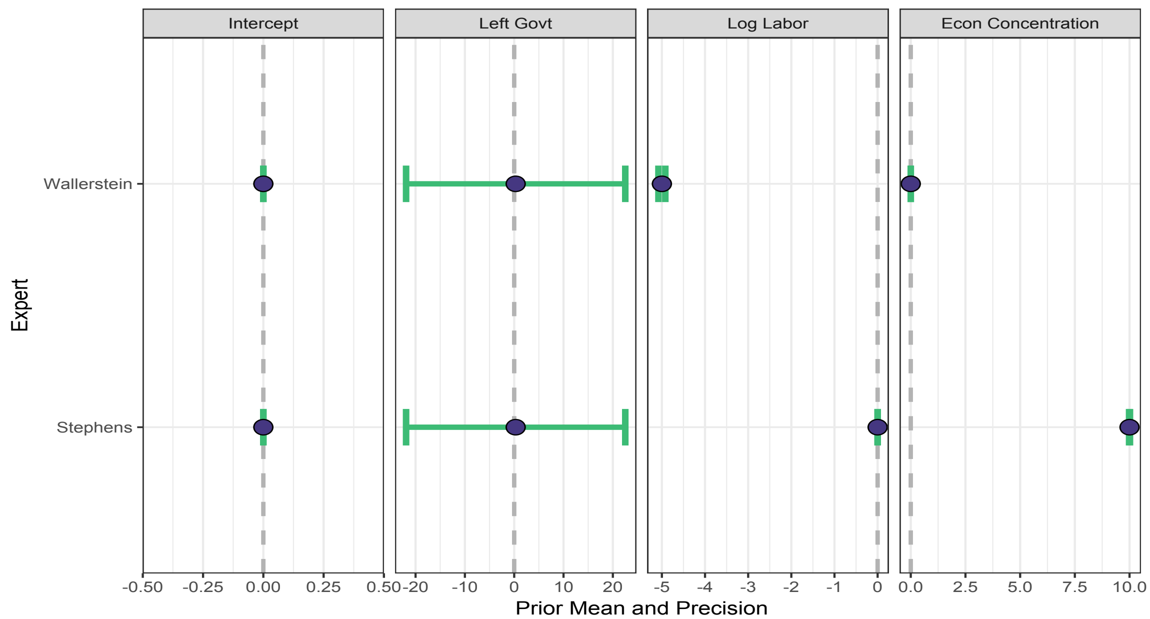


Figure 2.3: Wallerstein and Stephens' Priors and Precisions as shown in Jackman and Western (1994)

The most notable difference, as suggested by their written works, is that Wallerstein believes log labor force should have a strong negative effect on unionization (Stephens believes it has no effect) and that in turn Stephens believes economic concentration has a strong positive effect on unionization, whereas Wallerstein believes it has none.

In the subsequent analysis, Jackman and Western use each of these priors respectively to conduct a simple Bayesian linear regression, where unionization serves as the dependent variable, with left government, log labor force, and economic concentration as predictor variables.

$$\text{Unionization}_i = \alpha + \beta_1 \cdot \text{Leftgovt}_i + \beta_2 \cdot \text{LogLabor}_i + \beta_3 \cdot \text{EconConcent}_i + \epsilon_i$$

Their analysis uses 20 observations of country-level data on unionization and each of the explanatory variables. They first estimate the regression with uninformative priors to find a baseline result, and then apply Wallerstein's prior and Stephens' prior separately in turn to compare the results. Their findings illustrate both the power of priors to shape the conclusions we draw from our data analyses, and the need for better techniques to incorporate what might be divergent priors into analyses. While their paper emphasizes each of these priors separately to illustrate the divergent results, experts such as Wallerstein and Stephens may have equally valid insights into the research problem, or may have pieces of the same picture. How should these differing views then be reconciled?

2.5.1 Extending Western and Jackman

To illustrate the efficacy of the Dirichlet-process-based method I propose for aggregating priors, I reassess the data used by Western and Jackman using a slightly larger body of hypothetical priors. Rather than simply using the two paradigmatic examples of Stephens and Wallerstein, I construct hypothetical priors for "experts" who may favor any of the explanatory variables identified by Western and Jackman, as reasonable scholars of the literature may have a diverse set of beliefs about key explanatory factors. For example, while Western and Jackman treat "left government" as simply a control variable, my analysis assumes that some experts might consider left governments to play a significant role in determining unionization. Fewer studies evaluate unionization as an outcome, but among correlational studies that also investigate union activity in conjunction with economic and political institutions, several discuss the possible influence of political orientation (Hall and Soskice 2001; Borrel 2004; Lipset and Katchanovski, n.d.; Behrens, Hamann, and Hurd 2004), even where these effects may be indirect via regulatory and labor mobility policy (Farber and Western 2001; Lee 2005). To reflect this diversity of possible perspectives, I construct priors for 10 hypothetical experts, including Wallerstein and Stephens. In addition to these two iconic scholars, I suppose that there is someone skeptical of economic explanations—one who believes that government policies solely determine unionization; two different camps of communist views, one governmentalist that believes a liberal state can support unionization and another more classical, believing that only a large labor force could solidify the aims of the proletariat and lead to organization; a neoliberal who believes that the competition arising from economic concentration (industrialization) should decrease unionization; two labor-supporting experts who believe that the labor force is most strongly determining but who disagree about the magnitude; and two "uncertain" experts, who can agree on the same explanatory variables as the other experts but who believe the magnitude of effects is small. These perspectives, while hypothetical and diverse, are reflective of some key disjunctures in the literature on unionization —notably, whether institutional factors influencing labor union organization (e.g., union governance and incentives, networks, firms) or structural factors (e.g., demographics, industry characteristics) are more determinative (Western 1994, 2002). For example, the "uncertain" experts may prefer explanations internal to union organization, which are not reflected in the covariates under examination here, and therefore appear less certain about the effects of our variables of interest. By contrast, the governmentalist, communist, and neoliberal experts fall broadly into the category of those favoring more "structural" explanations for unionization. These positions are, however, abstractions based on work that was published after the original Western and Jackman analysis; as others have noted, yet further research remains to be done on unionization that may identify additional perspectives as well (Ahlquist 2017).

This selection of hypothetical priors for fictitious experts indicates at least one scope condition, and one key distinction, of the Dirichlet process clustering approach to aggregation. The addition of experts to this analysis illustrates that a clustering methodology is less necessary when an analysis includes only a very small number of experts. While a research could change priors governing the concentration parameter of the Dirichlet in order to incorporate their own information about the possibility of a given number of clusters—that is, a higher concentration parameter may be chosen if the researcher believes each expert is likely to represent their own perspective on a topic, while a lower concentration parameter corresponds to more aggregation of experts into clusters—a reasonable assumption, and the one used here, is that experts begin the simulation process effectively in their own clusters, and can be aggregated into a cluster with other experts when sufficient similarity in elicited parameters exists. Adopting this approach means that using Western and Jackman's original example with only two experts—Wallerstein and Stephens—would likely result in each expert being their own cluster, eliminating a benefit to aggregation. More broadly, however, the proposed approach will be less distinguishable from averaging as the concentration parameter becomes higher (that is, as experts are increasingly in their own clusters), or when no apparent underlying pattern of clustering exists. In general, clustering where consensus exists (that is, coherent clusters) will be more informative in the analysis than clusters that are very loosely defined or have high variance; under conditions of loosely defined clustering, then, the advantages relative to averaging also diminish.

Aside from this general scope condition, this vignette with hypothetical experts also demonstrates a critical distinction between this Dirichlet approach and other clustering approaches, notably k-means. Generating hypothetical experts and justifying their priors for the sake of this example means that the possible "schools of thought" at work with respect to this question are to some extent known. This characteristic corresponds more closely to a k-means clustering approach, where the number of clusters is specified in advance of the estimation. As even this example will show, however, having a sense of the ideological positions of a given set of experts need not correctly identify the perhaps more general "schools of thought" to which those experts' positions belong, nor does it suggest a way to aggregate those positions. The Dirichlet-based approach has an advantage, therefore, in not requiring a pre-specification of the number of clusters that exist among elicited priors, and these clusters can also update as more experts' opinions are added to the data. Researchers can therefore use the clustering process to inform them, to some extent, about the type and diversity of perspectives they gather through elicitation by looking at the clusters and cluster members that result from the Dirichlet Process. By the same token, however, this characteristic suggests that care should be taken when interpreting the results of the Dirichlet clustering, as the clusters themselves are artificial categories not necessarily identified with known "schools of thought." That is, much like the analogy to Latent Dirichlet Allocation articulated previously, a researcher may be able to identify a coherent "school of thought" with a given cluster as a benefit of this approach, but the membership of the clusters may also not be easily or directly interpretable and should be approached cautiously.

The prior means and precisions chosen for each of these hypothetical experts is reflected in Figure 2.4 below. The variances selected in the original Jackman and Western (1994) are often quite large, and others are quite small, and I follow this convention in choosing variances corresponding to the priors of the new hypothetical experts. Precisions ($\frac{1}{\sigma^2}$) are shown for visual clarity. As is evidenced in the figure, these prior means and precisions all seem like reasonable positions that experts on the issue of unionization might hold, but at the same time they reflect considerable diversity.

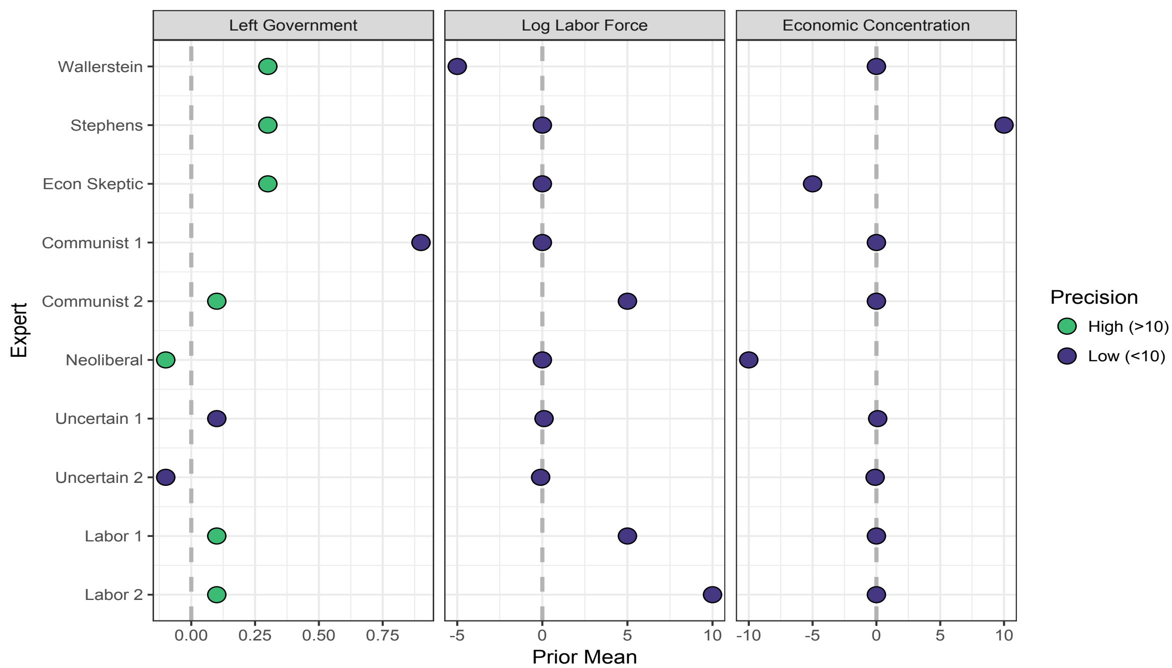


Figure 2.4: Hypothetical Expert Priors and Precisions

One approach to estimating a model with such diverse opinions of experts would generate separate estimates, with each expert's prior used in succession, and then compare results. This framework is reasonable if we believe that while each prior is legitimate, only one provides the "true" answer. As is especially true in studies of authoritarianism, where the information environment is incredibly fragmented, however, one can also imagine a circumstance in which each of these expert's priors contains some element of the truth, but having an aggregate picture would be preferable.

Current research in elicited priors recognizes this tradeoff to some extent. Because expert priors are often elicited in focus group settings, rather than in separate sessions, however, this problem is addressed by attempting to achieve a consensus prior, or trying to pool or average the priors of the attending experts. This approach, I argue, loses considerable information, and a Dirichlet-based method of aggregation should perform better, particularly when expert opinions are divergent. To illustrate this with the Jackman and Western (1994) data, I estimate the model from the original paper, and demonstrate the resulting posterior means and credible intervals when the 10 hypothetical priors are averaged versus when they are clustered in a Dirichlet process. I omit a comparison to pooling because the result would depend upon researcher-chosen weights for expert opinions.

Figure 2.5 illustrates the clustering of these 10 experts, real and hypothetical, through a Dirichlet process. Experts are allocated into six clusters that can be seen in the green squares across the diagonal of the plot, each suggestive of a "school of thought" to which these experts might belong. Wallerstein, the neoliberal expert, and the first uncertain expert each occupy their own clusters. Stephens shares a cluster with the second uncertain expert as well as with the labor-oriented expert whose opinion about the magnitude of the effect of the labor force is more moderate ("Labor 1"). The expert skeptical of economic forces shares a cluster with the second labor expert, whose assessment of the effect of labor is more pronounced. Notably, both labor-oriented experts' assessments of the effect of log labor force have low precision, and the low precision of their opinions as well as those of the uncertain experts should make them more "flexible" in terms of their cluster allocation given a fixed amount of information. The communist experts share a cluster, with slight uncertainty about whether Wallerstein may also be a member of that cluster. This clustering will influence the ultimate regression analysis: because expert priors are considered in clusters, rather than individually, each expert sharing a cluster will have relatively less influence on the results than if expert opinions were simply averaged, taken at equal and face value. Rather, opinions have greater weight when dispersed across clusters, where in this setup each cluster ("school of thought") is given equal weight, but the constitutive priors within each cluster are tempered. Furthermore, where latent "schools of thought" may not be apparent at the elicitation stage, the clustering process itself provides information about how many and what types of different perspectives exist among the experts polled—a characterization that may drastically change depending upon the pool of experts a researcher selects.

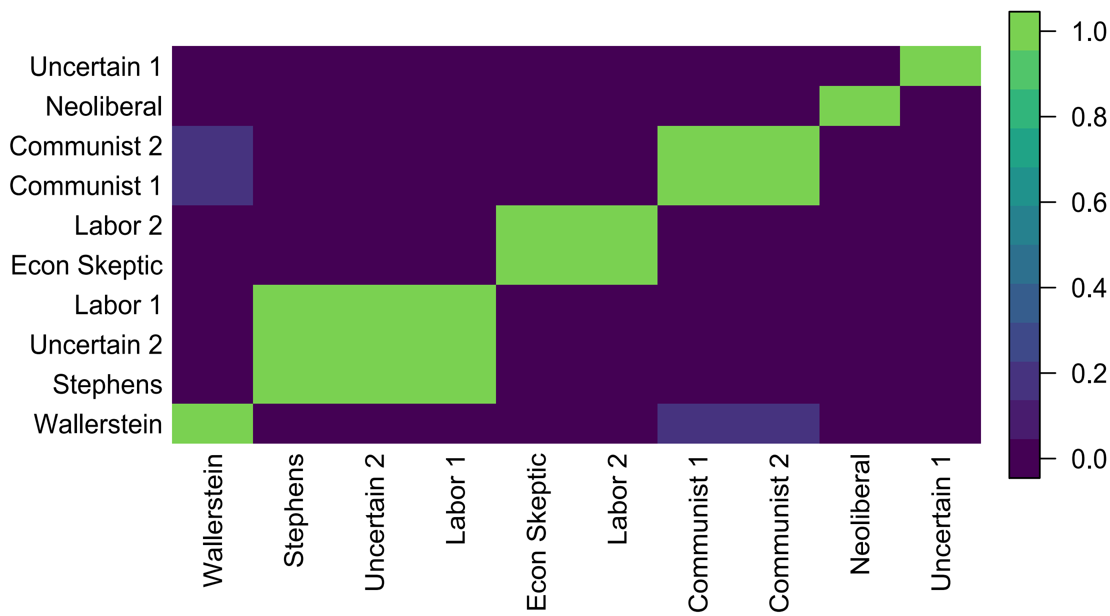


Figure 2.5: Jackman & Western Experts in Clusters

The results in Figure 2.6 demonstrate the ability of the Dirichlet process aggregation to better incorporate a series of divergent priors into a common analysis of unionization, relative to averaging priors. In estimating an effect for left gov-

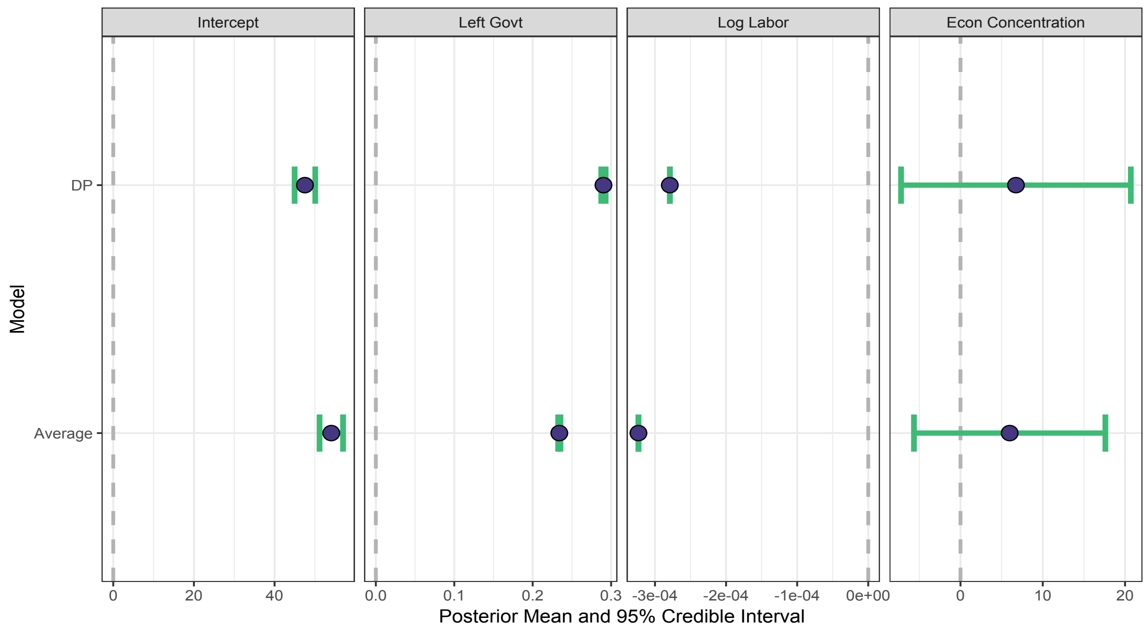


Figure 2.6: Averaging vs. Dirichlet Process Prior Results

ernments, the Dirichlet process model recovers a coefficient similar to the high-precision prior means supplied by Wallerstein, Stephens, and the Economic Skeptic, each of whom occupy separate clusters, whereas the model that averages priors gives slightly more weight to the views of the "Communist 2" and neoliberal experts by equally aggregating across individuals. In particular, because the communist experts share a cluster in the Dirichlet setup, their assessments of left government carry less weight than in the averaging model. A similar dynamic appears in the log labor force coefficients, which in both models reflect the prior means that were near zero in most cases, but which differ in that greater weight is assigned to the "Labor 1," "Labor 2," and "Communist 2" experts in the Dirichlet process, who are spread across clusters.

These results indicate that the Dirichlet process effectively handles the diversity of perspectives that experts may bring to an empirical analysis and reflect in their prior distributions. At the same time, its results appear less volatile than those of the averaging model because the Dirichlet process clustering is able to incorporate extreme views while identifying latent credibility in the data structure that serves to diffuse less informative extreme perspectives. Even so, the Dirichlet framework is very flexible: the standard setup treats clusters as having equal value, but a researcher themselves may have priors about the credibility or reliability of experts. The researcher's priors in turn can be instantiated as a hyperior, providing different weights for the clusters created through the Dirichlet process. In this way, even with as few as 20 observations, the information held by experts in the field, as well as researchers investigating current questions, can be effectively leveraged to conduct data analyses.

2.5.2 Implicit Weighting of Clusters

Researcher hyperpriors to affect the relative "weight" that clusters carry in an analysis should be applied with care and adequately justified given the aims of the research. A researcher may choose, on the basis of new information pertinent to the research question or the experts from whom priors were elicited, to influence the elicited priors with their own, but the assumption throughout this paper is that the researcher operates from a position of ignorance; if not, the researcher need not expend time and energy on complex elicitation. That is, the researcher is assumed not to have a notion of whose beliefs might be "right" or "wrong" at the time of elicitation, and therefore treats all elicited information as having equal value.

Even so, the researcher must understand how the proposed modeling framework handles different distributions of experts across clusters. While as a baseline clusters are "treated equally" in terms of their validity or credibility, the clustering process itself applies an implicit weighting with respect to the size of cluster membership. A greater number of members in a cluster provides more information, implicitly making cluster parameters more precise and lending greater influence over results. Consider the following example for the purposes of illustration. Suppose that only two schools of thought existed with respect to unionization, one reflecting Wallerstein's prior beliefs and the other Stephens'. Suppose then that the researcher elicited priors from a large number of experts—1000 total—across these two schools of thought, yet 990 of those experts had beliefs analogous to Wallerstein while only 10 had beliefs analogous to Stephens. This is simulated here using draws from a normal distribution centered on the prior mean or prior variance values of Wallerstein's and Stephens' priors (with $\sigma = 1$), as in figures 2.7 and 2.8 below.

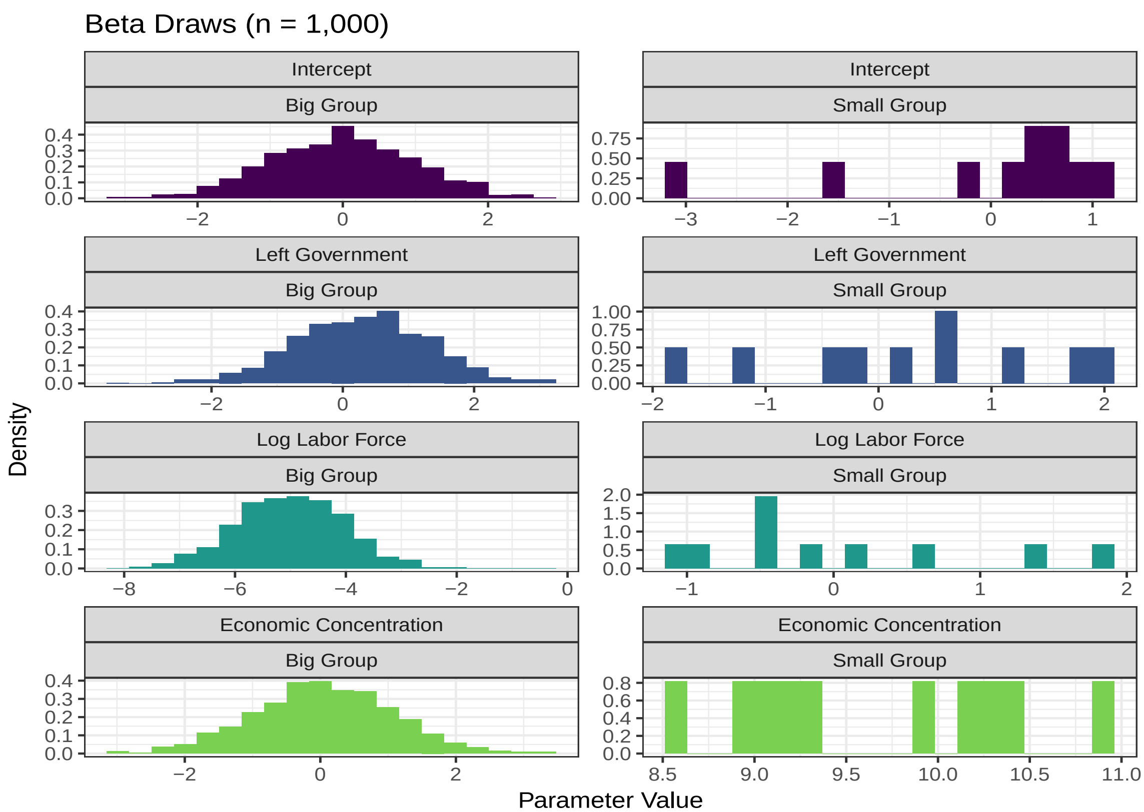


Figure 2.7: Simulated Prior Means for Implicit Weighting

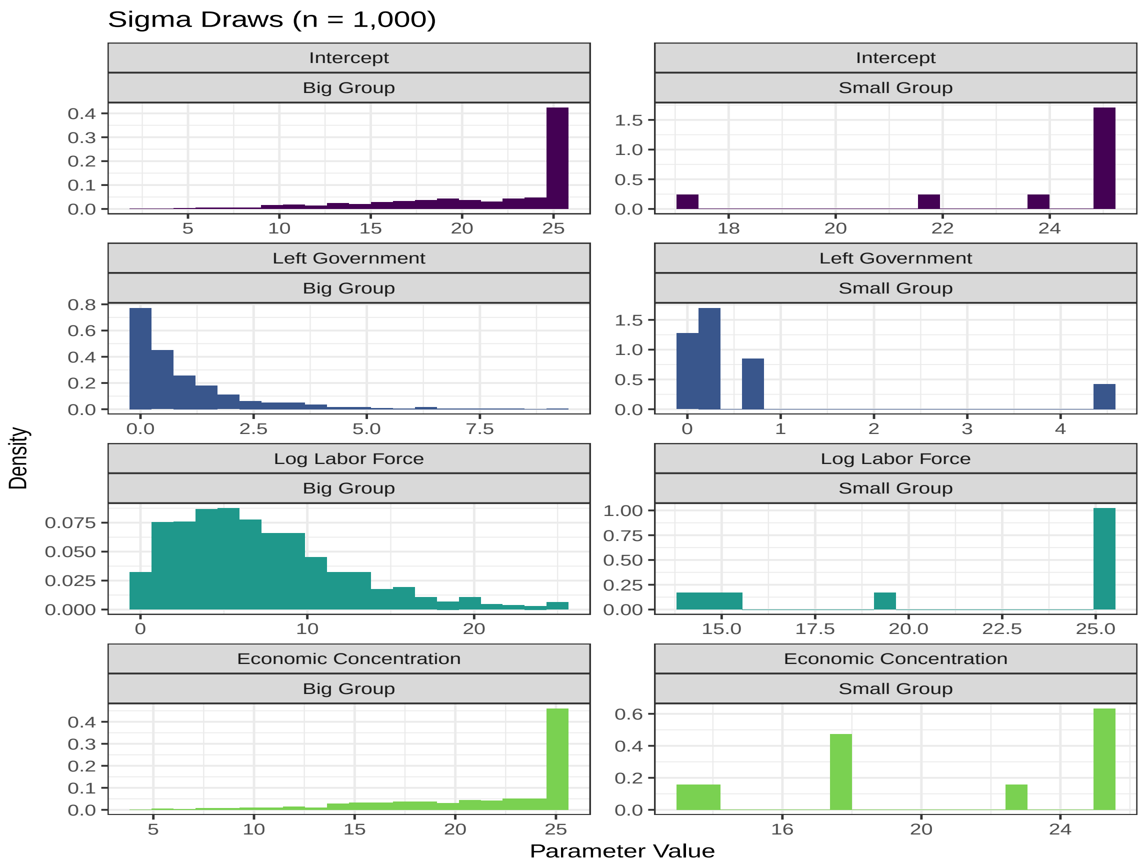


Figure 2.8: Simulated Prior Variances for Implicit Weighting

These two "clusters" of experts are then used to generate prior means and variances in the same modeling framework as the previous section. The results are reflected in Figure 2.9 in comparison to results where these priors are averaged across all 1000 experts—analogous to having elicited priors from experts equally distributed across two clusters. The results demonstrate that the posterior means for the model using clustering reflect greater "weight" toward Wallerstein's priors (refer to Figure 2.3) in comparison to the averaging case as a result of the larger membership of the cluster reflecting Wallerstein's beliefs.

Recognizing this implicit weighting is one plausible reason that a researcher may choose to apply a hyperprior of their own in the course of the analysis. For

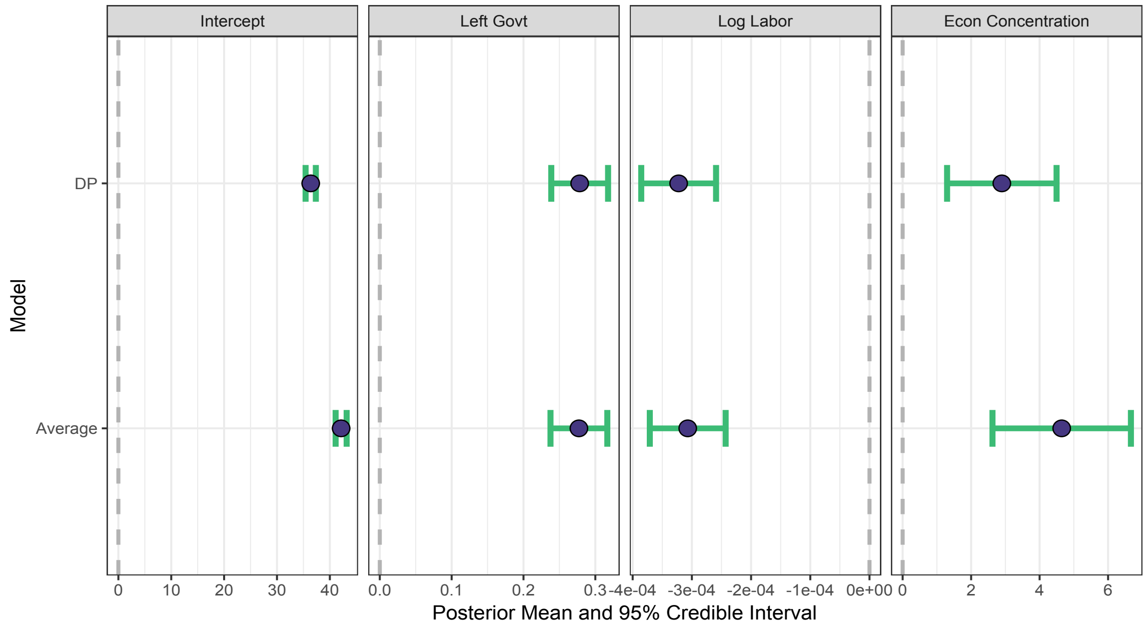


Figure 2.9: Model Comparison: Implicit Weighting of Clusters vs. Averaging

example, particularly in challenging modeling contexts, a researcher may not have equal access to experts across schools of thought, even if the researcher would like to elicit priors from all of them in good faith. In this case, the researcher may find that their clusters are unbalanced, with many more members in one cluster reflecting one school of thought and fewer representing the other. The researcher could, then, attempt to compensate for this implicit weighting via their own hyperprior. While this example provides one reasonable case in which a researcher may choose to modify the clustering framework proposed here, further research is necessary to examine the indications and implications for applying a weighting scheme across clusters, including under what conditions it is most reasonable and according to what protocol.

2.6 Conclusions and Next Steps

An elicited priors approach, as the literature discussed above demonstrates, provides incredibly useful additional insight in Bayesian analysis across a wide array of disciplines. In applying the method within political science, however, examples are much more limited. As I have argued, this is likely due to the fact that the current body of literature does not provide adequate guidance for adapting elicited priors to social science settings. In particular, because "expertise" can be a much broader and more nebulous concept in social scientific contexts, researchers are likely to encounter greater diversity in the priors they elicit. Current techniques do not adequately justify methods for aggregating these divergent opinions, nor do they offer concrete methods for adjudicating the value of some priors versus others beyond the statistical accuracy of the statements.

For researchers in the social sciences, and particularly political science, more work is needed to be able to adapt an elicited priors approach to relevant research questions. This is particularly the case for work that addresses authoritarian regimes or takes place in low-data environments. These settings, often synonymous with small-n work and poor data or information quality, are precisely the types of settings where a Bayesian approach should predominate because of its ability to more easily handle modestly sized data. In these settings, however, "experts" are likely to diverge more significantly in their opinions due to differing access to information and biases resulting from political status. Accounting for these differences in an elicited priors framework requires moving beyond the most common aggregation methods currently available (pooling and averaging).

In this paper, I have proposed a Dirichlet-based framework for addressing these diverging priors, borrowing a technology commonly applied in text-as-data type analyses. This method has greater flexibility and transparency, and can allow the researcher to aggregate priors according to different latent categories of consensus without only relying on the priors' statistical validity in reference to the data. This method stands to be particularly useful in settings where small-n and/or a complex modeling structure require informative priors, and where in principle experts exist to offer prior information but they are an as-yet untapped resource. As the example with the Jackman and Western (1994) data demonstrates, the Dirichlet-based approach competently deals with the diversity of potential expert opinions, and facilitates estimating a model even with sparse data.

The approach suggested here and the emphasis on eliciting priors may have its own downsides. In particular, elicitation of priors and the documentation of the research process in low-data contexts is likely to be difficult and more time consuming than other approaches. The aim, however, is to generate an approach that would both facilitate the use of elicited priors more broadly, while also particularly championing elicited priors as a tool for use in challenging research contexts where the lack of data or absence of identifiable expertise are particularly problematic. This goal is especially well-suited to closing part of the quantitative/qualitative divide that has emerged in some areas of political science, since the elicited priors approach would allow quantitative scholars to still conduct analyses while partnering with qualitative scholars whose deep contextual knowledge could aid in the identification of experts and the elicitation process as well.

2.6.1 Additions and Extensions

To further the analysis included in this paper, I aim to conduct a series of simulations that will illustrate the applications of the proposed method, and its performance relative to current methods such as pooling and averaging. In particular, this will involve simulating both continuous and discrete models/data with a series of generated priors "elicited from experts." The simulation will show how differing distributions of experts' information (e.g., divergent versus convergent, large versus small numbers of experts, large versus small number of data points elicited, etc.) contribute to the estimation process with each of these methods for prior elicitation.

To guide this empirical assessment of the proposed model, however, I also aim to provide a formal model of information elicitation. A general model of information elicitation is useful as a baseline for developing expectations about the type of information we as researchers receive when using an elicited priors approach, especially in an authoritarian setting. In particular, the authoritarian context for elicitation means that the information environment is constrained: individuals have limited access to information and their understanding of situations and consequences may be biased as a result. Whether and what information an "expert" offers during elicitation is a product of this constrained information environment, and this condition in particular should shape how information from different "experts" is and can be combined to better understand social processes at work. Constructing a formal model of why, whether, and what information individuals are likely to offer during elicitation given authoritarian constraints can provide more rigorous expectations for what methodology to employ when eliciting and aggregating these priors.

2.7 Appendix

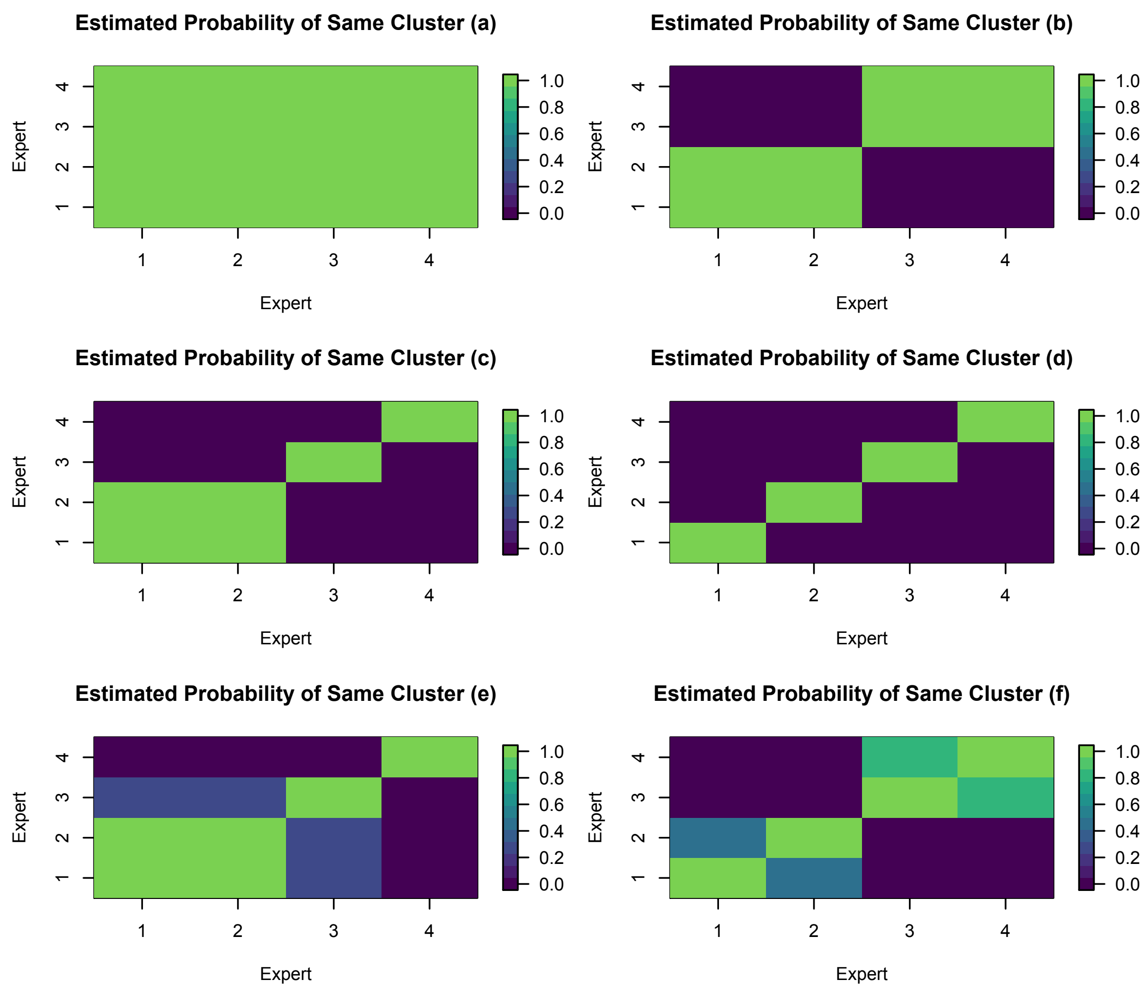


Figure 2.10: Example Clustering Square Plots

Figure 2.10 above illustrates several possible clustering outcomes in a series of square plots. For the sake of simplicity, in each of these example plots, suppose that the researcher has elicited priors from only four experts. Each of these experts is represented by a cell in the square plot, as indicated by the x and y axis labels, which number the experts from 1–4. In plot (a), all four experts are allocated to the same cluster. This is indicated by the fact that each cell in the plot is shaded green, which means that in 100% of samples, each expert shared a cluster with every other expert. The unique number of clusters across a set of samples are given by the number of green square across the anti-diagonal of the square. In plot (a), there is only one green square, corresponding to a single cluster. In plot (b), however, there are two green-shaded squares on the anti-diagonal, indicating that experts have been allocated into two distinct clusters. In plot (b), experts 1 and 2 are allocated into the same cluster, while experts 3 and 4 are allocated into a separate cluster. The green shading in cells (1,1), (1,2), (2,1), and (2,2), where coordinates indicate expert numbers, visualize the fact that expert 1 shares a cluster with themselves, expert 2 shares a cluster with themselves, and experts 1 and 2 share a cluster with each other, because the cells (1,2) and (2,1) are also shaded green. The dark purple shaded squares indicate that the corresponding experts shared a cluster in 0% of samples. Hence, experts 3 and 4 in plot (b) never share a cluster with experts 1 and 2.

The same logic applies to plots (c) and (d), which illustrate the experts being allocated into three and four distinct clusters, respectively. In plot (c), experts 1 and 2 share a cluster while expert 3 is in a cluster alone and expert 4 is also in a cluster alone. In plot (d), by contrast, each expert is allocated to their own cluster. While the elicited priors in plot (a) indicate limited differences among the opinions of experts surveyed, in plot (d), these opinions differ sufficiently that each expert's beliefs constitute a distinct "school of thought."

Plots (e) and (f) illustrate uncertainty in the clustering process. Because clusters are assigned in a sampling procedure, the clustering outcome may reflect uncertainty as to the cluster assignment for a given expert. This uncertainty is represented in the shades between green and purple, where darker shades indicate that an expert was allocated to a cluster in fewer samples, and lighter shades indicate that an expert was allocated to a cluster in more samples. In plot (e), for example, expert 4 is always in a cluster alone, and experts 1 and 2 are always in a cluster together. Expert 3, however, is sometimes assigned to a cluster by themselves, and is sometimes (but more rarely) assigned to share a cluster with experts 1 and 2. Plot (f), by contrast, illustrates differing levels of uncertainty. Every expert, by definition, must share a cluster with themselves, so this plot indicates with certainty that each expert is in a cluster with themselves (green on the anti-diagonal). The shading in plot (f) also indicates that, to the extent that experts are in fewer than four distinct clusters, experts 1 and 2 would share a cluster and experts 3 and 4 would share a separate cluster. Whether there are two, three, or four distinct clusters is a matter of uncertainty, however: experts 1 and 2 more often are in distinct clusters (20% of samples have them sharing a cluster as indicated by the dark blue squares in cells (2,1) and (1,2)), while experts 3 and 4 are more often in the same cluster (80% of samples have them sharing a cluster, as indicated by the darker green in cells (4,3) and (3,4)). When computing the cluster values for priors, cluster assignments are averaged across samples in order to incorporate this uncertainty.

Chapter 3

Collective Wisdom: Rethinking Expertise in Elicited Priors

3.1 Introduction

What is "expertise" and who is an "expert?" Expertise is central for informing priors in Bayesian analysis, whether directly through elicitation or indirectly by influencing the beliefs of researchers. The benefits of eliciting priors diminish as those surveyed are less informed, both because vague priors add less information to the data and because elicitation can be costly to conduct. Effective elicitation, therefore, depends to some extent on identifying which sources have "expertise" on which an analysis can draw, but what type of expertise should inform social scientific research is an open question. Invocation of the term "expert" in the course of elicitation, in fact, serves to instill trust in the research process by indicating that only those with particular knowledge had potential influence over the empirical analysis. Likewise, because technologies for aggregating expert opinions into coherent priors for Bayesian analysis have been fairly limited, keeping the set of sources for elicitation constrained eases practical problems in conducting research with informed priors. Nevertheless, expanding the definition of "expert" is critical to obtaining informed priors that fit a social science research agenda—especially with respect to contexts where differential access to information and/or underdevelopment systematically undermine traditional institutions (e.g., education systems) through which researchers could identify and select experts.

This paper interrogates the distinction between "experts" and "the masses," evaluating and validating what divided "expertise" provides when eliciting priors from groups with differing perspectives. In particular, we leverage the November 2016 Presidential election in the United States to assess the predictive capacities of two groups—a "mass" sample and an "elite" sample—of individuals. Using online surveys conducted in October 2016, we elicited priors on vote share for the presidential election as well as hypothetical U.S. House elections. We implement a Dirichlet Process clustering method to aggregate these priors, allowing us to incorporate very diverse perspectives across both samples of respondents while also grouping respondents' opinions into latent "schools of thought." Using the real election outcomes, then, we are able to validate both (a) what "types" of experts are able to provide more accurate predictions of these electoral outcomes and (b) whether and under what conditions the proposed clustering method improves predictions relative to averaging.

3.1.1 Expanding Expertise

Tom Nichols' recent book, *The Death of Expertise*, decries current trends toward anti-intellectualism and substituting away from education-based knowledge toward social media and information technologies (Nichols 2017). While Nichols views the anti-elitism inherent in skepticism of experts to be a celebration of ignorance (aligned with the Dunning-Kruger effect), the tension between expertise arising from a system of "gatekeepers" to knowledge and experiential expertise is not resolved by these types of polemical claims (Nichols). In particular, perhaps shoring up credentialed expertise is important in scientific contexts like the one Nichols mentions—who do you trust more: a doctor versus your Aunt Ginny—or traditional settings for elicitation, but this is far less clear for social science research (Nichols). First, the broadening or "democratizing" of the information environment likely has differing effects in contexts like the United States relative to developing or authoritarian systems. Differential access to information is not normatively neutral: systems of inequality mean that formal education is often available to only a small subset of society, even in developed contexts. This type of inequity is only implicitly addressed in Nichols' work, but it highlights the fact that identifying an expert pool on the basis of formal educational credentials could generate bias in the information elicited.

Second, and even without a claim to bias, the perspective that expertise resides primarily in those with formal education neglects the kind of experiential expertise that might better inform certain research questions.[1] This discrepancy also highlights at least one way in which limiting the selection of experts . For example, the Women Also Know Stuff project within political science has undertaken the creation of a database of women experts in various subfields of the discipline after recognizing that many media reports and conference panels featured an all-male cast of "experts," supposedly for lack of ability to identify any non-males who worked on a particular topic (Boydstun et al. 2017). This project, and the attendant Twitter hashtag #WomenAlsoKnow, underscore the fact that a given researcher, selecting a small pool of experts for prior elicitation, likely selects for expertise given their own biases about what constitutes knowledge or who qualifies as an "expert." The practical ability to incorporate more and diverse opinions through

[1]Succinctly, in the words of Jethro Tull, "your wise men don't know how it feels/ to be thick as a brick."

the clustering methodology discussed in this paper should lessen this constraint by facilitating the inclusion of more voices in a given research project.

Third, this distinction between types of expertise or knowledge itself undergirds a very all-or-nothing notion of understanding social scientific phenomena. Rather, and more likely, even "experts" on a particular subject of study have only partial information about the underlying variables and mechanisms, and aggregating across differing perspectives would provide a fuller picture of the processes at work than soliciting opinions only from those with the same perspective. Taking this to an extreme, Sloman and Fernbach (2017) argue from a psychological perspective that humans create a "community of knowledge" for themselves that allows for the functioning of society, and that suggesting that knowledge and expertise are atomistic instead reinforces a problem by which individuals are less willing or able to identify the limitations of their own expertise.

Each of these dimensions of the debate over expertise suggests that more work is necessary to interrogate the category of "expert," and in particular to evaluate what constitutes an "expert" when eliciting priors for social science research. In the sections that follow, we provide an overview of current research on the nature of expertise in elicitation, as well as particular examples from election forecasting. Following this discussion of the literature, we describe our survey methodology for eliciting priors related to the 2016 U.S. national elections, as well as the clustering methodology used to aggregate these priors.

3.2 Literature: Expertise, Elicitation, and Forecasting

3.2.1 Expertise & Good Judgment

While utilizing expert opinion has traditionally formed the basis of elicited-priors approaches, relatively little attention has been paid to defining and refining a notion of "expertise" that best serves Bayesian analysis. Literature in psychology and management science has sought to understand both how best to identify individuals who inherently perform better in statistical tasks, and utilize their knowledge and/or abilities to improve estimates, and how best to leverage the "wisdom of crowds," while being cognisant of the limitations of any single individual within a group.

For example, it has long been recognized—even as early as Aristotle—that an aggregate assessment of probability is preferable to individual assessments alone (termed the "wisdom of crowds"). At the same time, however, researchers have sought to improve upon this aggregate, not least because ensuring a large sample of respondents for a given question can prove practically challenging. Budescu and Chen (2014) specifically designed a model to more heavily weight the contributions of individual forecasters ("judges") who perform better than the group average, and test this model on economic forecasts with respect to the European Central Bank to show its efficacy. Similar approaches have been undertaken in other management and psychology research in an attempt to shift weight toward better-performing forecasters (Karvetski et al. 2013; Satopaa et al. 2014). These approaches are intuitively appealing because they can help to address issues in survey sampling: it is not clear how many individuals one needs to sample in order to encounter diminishing marginal returns in forecast accuracy.

Likewise, from a psychological perspective, not only are some individuals presumed to simply be better in forecasting tasks, but some of the additional accuracy in particular individuals' responses is attributable to abilities to overcome common human fallacies in understanding statistical statements. Specifically, the way in which uncertainty is expressed has a significant impact on individuals' ability to respond with an accurate assessment of their own (Dhami et al. 2015, 754).

These observations from the literature underlie the forecasting process implemented in the Good Judgment Project, founded by Philip Tetlock, wherein individuals provide forecasts for political events of relevance to the intelligence community in a tournament style, and the project identifies "superforecasters" with greater-than-average accuracy to improve future estimates (Tetlock 2017). While in principle this approach seems both sensible and simple, a few challenges arise. First, as Tetlock himself acknowledges, defining and measuring "accuracy" of predictions is fraught (Tetlock and Mellers 2014). Furthermore, the task of identifying "superforecasters" requires the assumption that good performance on one prediction task can translate to good performance on another: "superforecasters" are designated through a process of iterative validation but are not seen as having any particular domain-area expertise that might allow them to be identified prior to undertaking forecasting tasks. Superforecasters in Tetlock's framework are assigned to work in teams to provide forecasts, but posthoc analysis of superforecaster accuracy rendered four seemingly vague distinctions thought to give rise to their success: "(a) cognitive abilities and styles, (b) task-specific skills, (c) motivation and commitment, and (d) enriched environments" (Mellers et al. 2015). Aside from the "enriched environment" condition that arises from team activity, these characteristics might generally describe someone with some higher-than-average level of education who was interested in the underlying question for an elicitation project, since these characteristics might provide the basis for cognitive abilities, skills in statistical reasoning, and motivation.

While the psychological literature pertaining to forecasting recognizes a "wisdom of crowds" principle, then, the countervailing effort to identify "superforecasters" whose predictions surpass those of their peers, in order to improve the general accuracy of forecasting, has currency. This approach faces practical challenges, but also conflicts with a notion of expertise or understanding of social scientific phenomena in which knowledge is interdependent (that is, a fuller picture can be constructed across individuals' understanding) and one of expertise that pertains to a specific domain area per se. Likewise, the process for selection emphasizes characteristics that are challenging to identify prima facie. The clustering approach proposed in this paper, by contrast, acknowledges that different types of information are best drawn from across a broader set of individuals, while also recognizing that some individuals or groups of individuals may perform better in statistical reasoning tasks, or in forecasting specifically.

3.2.2 Forecasting Literature: Representativeness & Aggregation Challenges

"As long as there have been elections," writes Hillygus (2011), "people have tried to predict the outcomes" (964). Beginning with predictions based upon elite observations and the opinions of "knowledgeable observers" or "political insiders," election forecasting has evolved to incorporate increasingly diverse sets of polls and models (964). Now that technologies for estimating elections outcomes have improved significantly (Lewis-Beck and Stegmaier 2014), even greater weight is placed on accurate measurement and inclusion of variables known to matter (e.g., economic conditions and incumbency effects), and on incorporating more data over time (Lewis-Beck and Tien 1996; Evans and Ivaldi 2010; Linzer 2014). Bayesian approaches in particular have explicitly taken into account prior data (e.g., previous electoral outcomes and polls) when generating estimates for national elections (Rigdon et al. 2009; Lock and Gelman 2010; Linzer 2013). Making correct design decisions significantly impacts the accuracy of modern forecasting relative to more traditional expert-based assessments. The type of validation undertaken in this paper mirrors this ongoing, if implicit, concern in the forecasting literature. While some skepticism remains as to the usefulness of generating general evaluations of forecasting algorithms (Campbell 2008), improving forecasting estimates necessarily requires transparency in method, attention to measurement, and accommodations for time-varying effects (Campbell 2014). Specifically, in addition to considering which variables are best suited to forecasting models, researchers investigating forecasting methods contend with

(a) which individuals (via polls) or information sources to draw from, and

(b) how best to aggregate estimates across models or samples.

Greene (1993), writing about the challenge of producing a forecasting model that performs better than "pundits," describes an alternative method as follows:

> ...I organized an election pool in my department, in which contestants had to select the winner of the presidential race, the Democratic percentage of the two-party popular vote, and the winner in each state. Nineteen people entered, mostly graduate students in fields other than U.S. politics. The competitors were certainly not experts in U.S. elections, but could best be described as informed observers. Predictions were submitted in mid-September, admittedly a little later than the models, but still well before the election. The election pool performed as well as any of the models.... Sixteen of the nineteen contestants (84%) picked Clinton as the winner. The mean prediction for the Democratic percentage of the two-party vote was 52.4%, only one point off the 53.4% that Clinton actually received. But most importantly, the 95% confidence interval for the pool's average was only 4%, compared to 5% for Abramowitz' model and 8% for Fair's. (20)

This anecdote highlights a question similar to the one addressed in this paper: under what conditions will informed respondents perform better in predictions than alternative data sources and models? Relatedly, in what ways might "experts" provide more accurate assessments than the masses, or than "pundits"? Literature evaluating later elections details the advantages and disadvantages of some of these alternatives, including futures markets, campaign polls, and regression models (Jones 2008), noting that at least with respect to the U.S. 2004 election, futures markets outperformed other forecasting methods. In general, futures markets are known to have better estimates over longer time horizons (Berg, Nelson, and Rietz 2008; Atanasov et al. 2016). Why exactly futures markets provided better forecasts than other methods is left unexamined, but the underlying logic runs somewhat counter to proposals such as Greene's: with a large enough group of participants (and with market incentives in place), predictions increase in accuracy.

By the same token, the theory of opinion polling largely rests on representative sampling methods, yet the choices of polling firms when conducting these surveys—choices related not just to the sampling frame but also to wording and measurement—has significant potential to induce non-random error in measurements (Campbell and Lewis-Beck 2008, 191). As Gelman and King (1993) note, polls vary significantly and do not reflect "rational" responses over the entire course of the campaign, even when arriving at relatively accurate conclusions. Furthermore, particularly with decreasing rates of telephone survey response, and although alternatives such as mail surveys have been considered to increase response rates (Visser et al. 1996), Wang et al. (2015) show that even an unrepresentative sample of individuals (in this case polled through the Xbox gaming platform) can provide accurate election forecasts. That said, the accuracy of these non-representative sample results depends to some extent on correct weighting; other attempts to use nontraditional venues for eliciting forecasting information, such as Twitter, have been less successful (Huberty 2015).

Micro approaches to forecasting featuring individual vote expectations also show promise in accurate forecasting ability, irrespective of reported political involvement or campaign interest (Lewis-Beck and Tien 1999). Rather, in addition to complications such as the closeness of the political contest and the timing of the survey, individual variables such as education play a meaningful role in the accuracy of individual forecasts (181). That is, while in general large groups of individuals can provide accurate election forecasts, with a micro-level approach, some individuals will provide more accurate forecasts than others, and overall accuracy is improved when weighting the responses of more "competent" respondents more heavily (Andreas Erwin Murr 2011; Andreas E. Murr 2015).

Furthermore, concerns for aggregation across models are also prevalent in the forecasting literature. In general, research demonstrates that combining (usually averaging) across forecasts produces more accurate results than any single component forecast (Graefe et al. 2014; Graefe 2015). Attempts to further refine from averaging, using Ensemble Bayesian Model Averaging, for example, have performed less well (Graefe et al. 2015). In part, this discrepancy reflects the fact that alternatives to averaging require information about which components may be more accurate in order to increase the overall accuracy of the combined estimate. In that sense, then, a tension exists whereby, at both the individual and the aggregate forecasting level, aggregation with attention toward more accurate micro-level predictions improves accuracy, but determining what conditions correlate with improved accuracy at the micro-level presents challenges.

Taken together, these observations from the forecasting literature provide a few questions that this paper further investigates:

(1) Can individual-level predictions from a representative sample provide more accurate estimates than those from an "elite" sample?

(2) Can aggregating elicited predictions from "experts," whether mass or elite, in a Dirichlet clustering process provide more accurate estimates than averaging?

(3) What subgroup conditions correlate with greater accuracy in predictions, and how can these be favored in the aggregation process?

3.3 Survey Methodology

For this project, we utilize web-based surveys to access a broader set of "experts" than would otherwise be accessible via traditional focus group or interview methods of elicitation. In order to elicit prior probability distributions, we conducted two separate surveys: one for a nationally representative "mass" sample of individuals, and one for an "elite" sample composed of Ph.D. students in political science at top 20 graduate programs in the United States. The mass survey sample consists of 408 American individuals identified and solicited for participation by Qualtrics. Participants in the elite survey, also conducted within Qualtrics, were identified for participation via their public profiles and email addresses on graduate program websites, and received emails asking for their participation. The elite sample consists of 57 completed responses of 205 who began the survey.[2] The mass survey participants were compensated for their participation, while the elite sample participants were not. Both surveys collected demographic data, including race/

[2]One completed response was dropped from this set, for a final total of 56 responses, due to inappropriate responses. This individual completed the survey in 5.28 minutes (faster than any other completed survey response in the elite set), answered every question with 10 chips in a single bin irrespective of covariate values, and provided insincere answers to demographic questions. These factors in conjunction suggest that the responses were not reflective of true beliefs.

ethnicity, sex/gender, highest level of education, region of residence, and age. Figures 3.1, 3.2, 3.3, 3.4, and 3.5 below summarize demographic characteristics for each sample, as well as party identification in Figure 3.6.

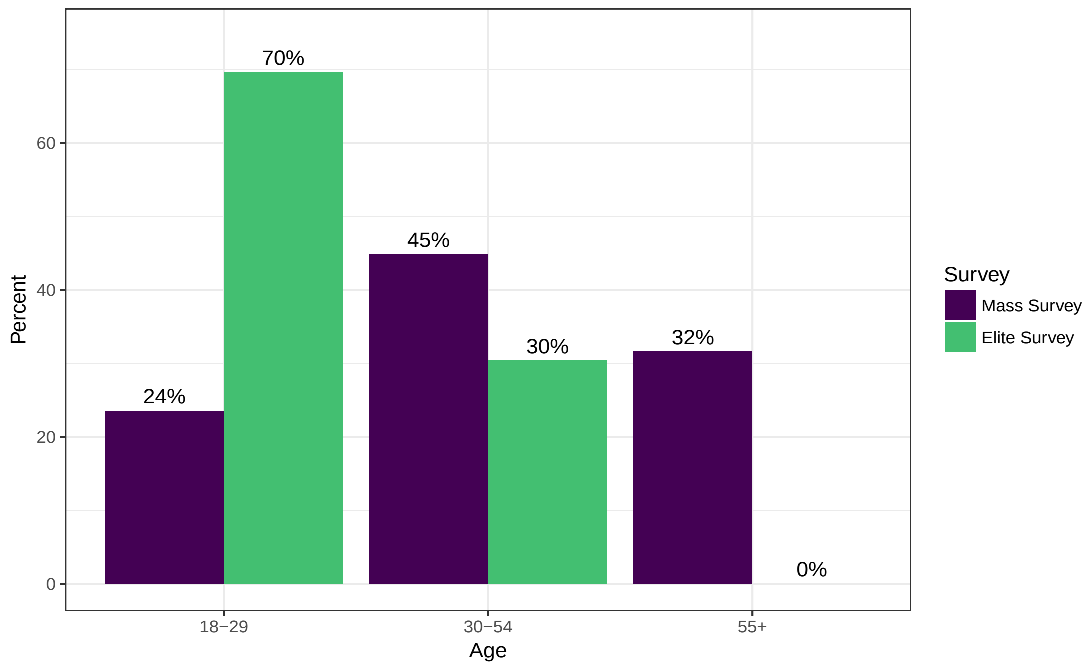


Figure 3.1: Age of Survey Samples

While the mass sample were selected through Qualtrics to provide a nationally representative sample on the basis of key demographic characteristics (age, sex, race, region, etc.), the elite sample is more skewed in several characteristics. Notably, the elite sample consists of a much larger set of individuals who identify as "Strong Democrat," and many more males than non-males. Education differences are expected and intended. In particular, the "elite" sample reflects what is more typically sought in a standard elicitation routine: individuals with significant academic training and credentials related to the subject matter. While Ph.D. students were targeted irrespective of subfield, their academic training could provide more insights into the ways that our chosen covariates might influence electoral

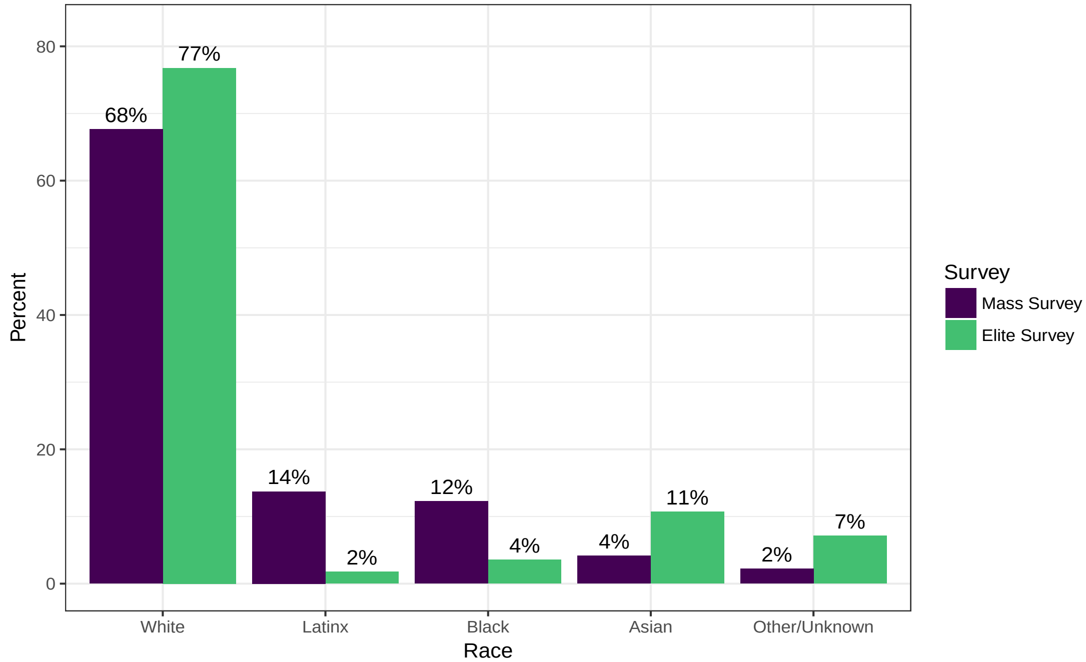


Figure 3.2: Race/Ethnicity of Survey Samples

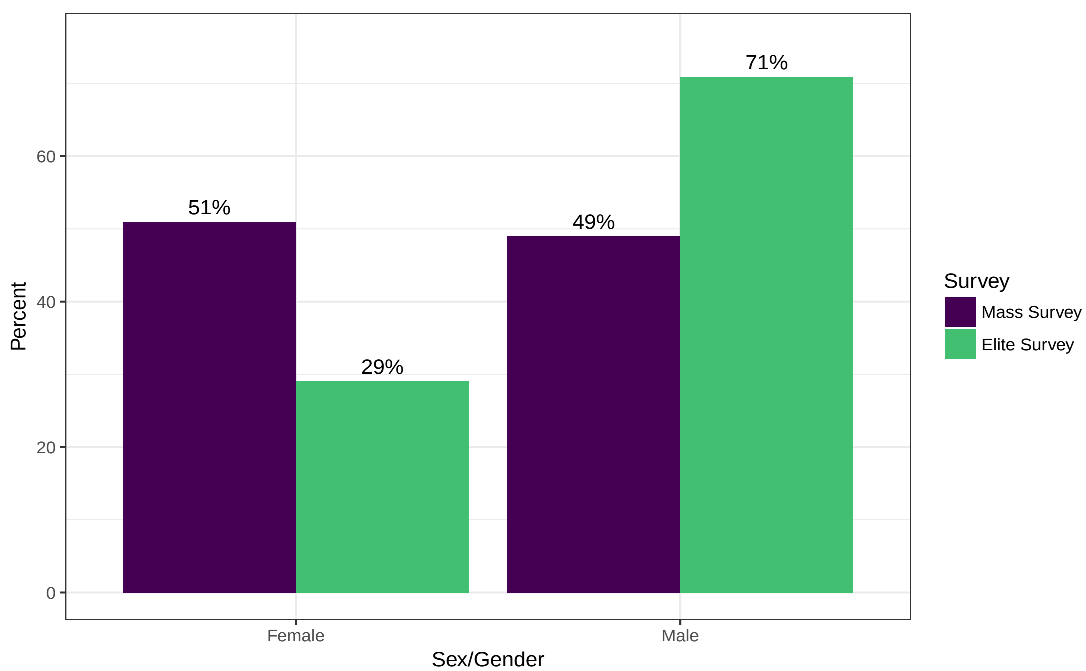


Figure 3.3: Sex/Gender of Survey Samples

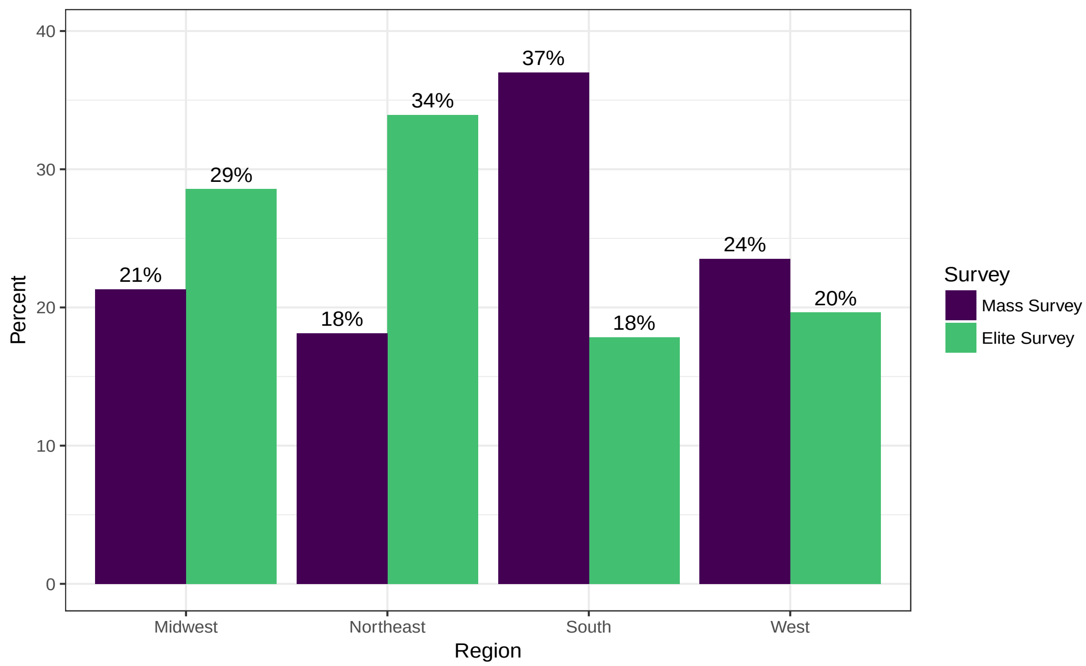


Figure 3.4: Region of Residence of Survey Samples

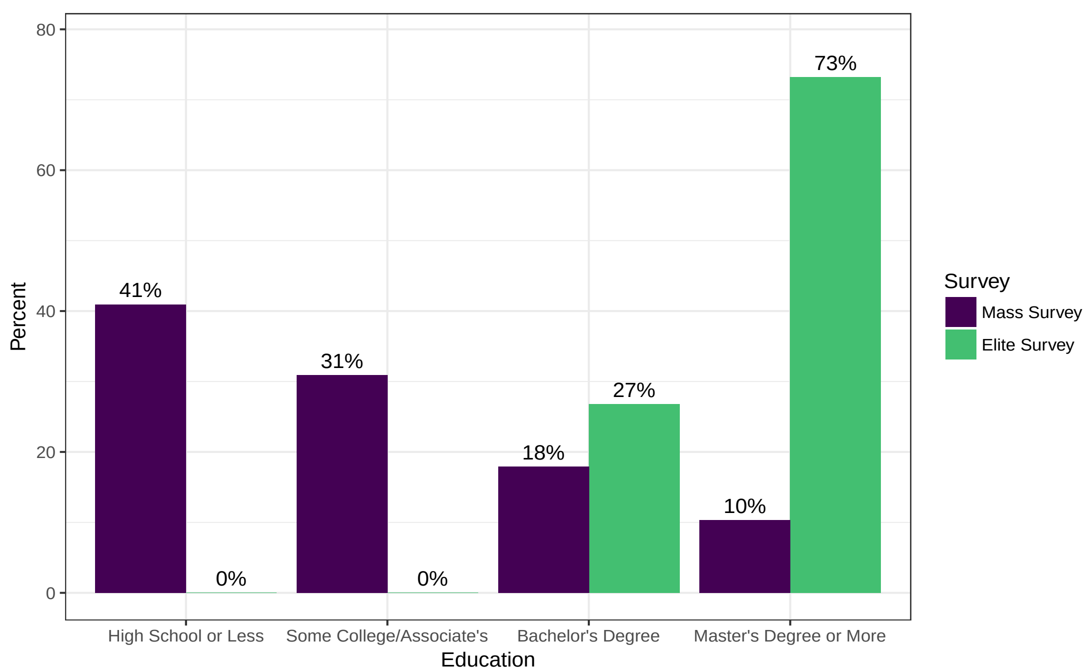


Figure 3.5: Education of Survey Samples

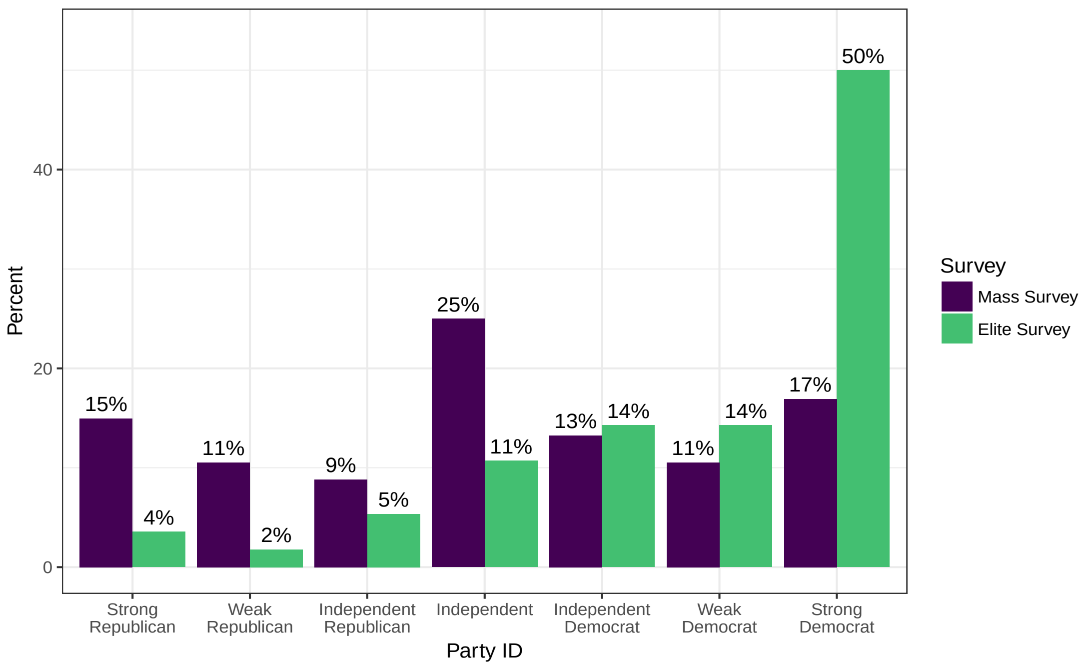


Figure 3.6: Party Identification of Survey Samples

outcomes, and they may provide more specific (less vague) priors than those in the mass sample. Likewise, while the mass sample received compensation for their participation, the elite sample would self-select into participation and completion, which should bias the sample toward individuals who at least believe themselves to be more knowledgeable about the subject matter.

Respondents in both surveys were randomized to receive questions that addressed either Hillary Clinton or Donald Trump's vote share in the November 2016 Presidential election. Subjects answered questions in a "roulette" elicitation format as articulated in the Sheffield Elicitation Framework (SHELF) (O'Hagan and Oakley 2016). Subjects received 10 "chips" for each question, which they could use to "bet" in each bin of vote share probabilities, forming a probability density. Respondents were presented with 6 bins per question that encompassed vote ranges between 40 and 60% (i.e., $< 40\%$, $40 - 45\%$, etc.). With 10 chips, each chip represented a 10% probability. That is, if a respondent believed that Hillary Clinton had a 50% chance of receiving a vote share in the $51-55\%$ range, that respondent would place 5 chips in that bin. Question validation ensured that each respondent used all 10 chips for each question. An example question is shown in Figure 3.7 below.

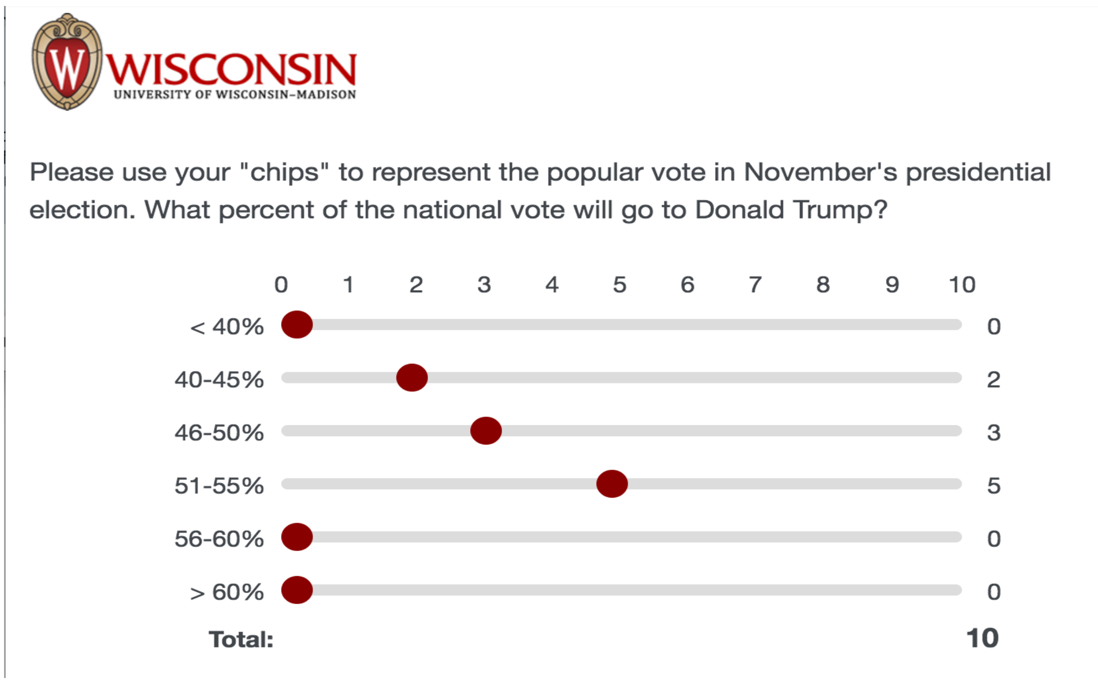


Figure 3.7: US Elicitation Survey Question Example

Respondents first answered what proportion of the national vote would go to a candidate (either Clinton or Trump), then what proportion that candidate would receive in each of three swing states: Florida, Ohio, and North Carolina. Following these raw predictions, respondents were presented with a random selection of questions for their given candidate across five swing states: Florida, North Carolina, Ohio, Pennsylvania, and Wisconsin. For each of these state questions, prompts varied to reflect how much considerations such as the economy, national security, and campaign advertising might impact vote share. In every question, the state unemployment rate varied (either 4% or 6%) in comparison to the national unemployment rate, set at 5%. Questions also specified whether, globally, there had been 1 terrorist attack or 3 terrorist attacks in the previous month. Finally, each question specified whether the given candidate (Clinton or Trump) had equally as many, or 20% more, campaign ads as the other. The mass sample survey included a timed validation, where subjects whose survey durations lasted less than 170 seconds were dropped from the final sample in order to ensure sufficient consideration for each answer.

In addition to the questions in the mass survey, respondents in the elite survey also received a selection of questions regarding 2016 U.S. House of Representatives races in hypothetical districts, asking them to use their chips to evaluate the proportion of the vote an incumbent would receive. Each question characterized the district by the vote share Obama received in the 2012 election (45%, 50%, or 55%). The question specified the incumbent's NOMINATE score in relation to others within their party (15th or 30th most conservative of 46 Republicans, or 19th or 38th most liberal of 57 Democrats) (Carroll et al. 2015), reflecting the first and second tertiles for NOMINATE scores. Questions also indicated whether the challenger had ever held elected office (yes/no).[3] Each question also included the state unemployment rate, varying between 4% and 6% relative to the national rate of 5%, as in the general survey questions. In total, mass survey respondents answered 35 questions, and elite survey respondents answered 47 questions.

Beyond demographic data and substantive elicitation questions, subjects answered five political knowledge questions adapted from the American National Election Survey in order to assess their general familiarity with American politics (ANES 2015). The questions respondents answered were:

[3]As a result of a technical error, one blank prompt appeared in a random selection of surveys. Respondents were instructed to answer this question with their chips in order to satisfy validation and complete the survey, but responses for this question were dropped from the final sample.

1. Which party has a majority in the U.S. House of Representatives, as of August 2016?
2. Do you happen to know how many times an individual can be elected President of the United States under current laws?
3. For how many years is a United States Senator elected—that is, how many years are there in one full term of office for a U.S. Senator?
4. Is the U.S. federal budget deficit—the amount by which the government's spending exceeds the amount of money it collects—now bigger, about the same, or smaller than it was during most of the 1990s?
5. On which of the following does the U.S. federal government currently spend the least? (Foreign Aid, Medicare, National Defense, Social Security)

The relative rates of correct answers to these political knowledge questions are visualized in Figure 3.8 below. These results illustrate that, as expected, the elite sample has, in general, more correct answers to these measures of political knowledge.

Our preliminary analysis includes responses irrespective of respondents' scored political knowledge. The answers to these questions can, however, serve two potential purposes. First, responses could be used to discount the contributions to cluster estimates of those whose political knowledge is less than a threshold level (e.g., the sample average). Second, by contrast, these questions may serve to highlight the disjuncture between political knowledge as typically measured and "expertise" regarding potential outcomes for the 2016 election. That is, perhaps some individuals have less knowledge of the formal structure of American politics but

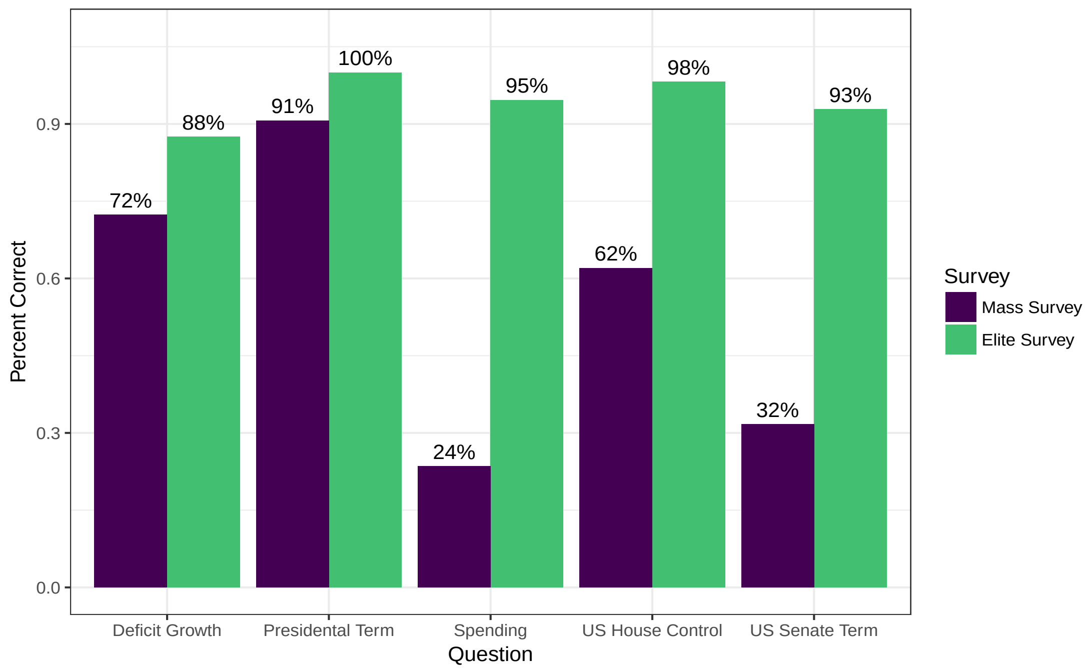


Figure 3.8: Political Knowledge: Mass vs. Elite Answers

are more aligned with sources of information that correlate strongly with the election outcome.

3.4 Clustering Methodology

The survey responses for both samples of experts provide bin counts of chips, which can be used to characterize each respondent's prior distribution for the influence of a particular covariate (economy, security, campaign advertising, state) on the ultimate election outcome in terms of vote share. While one could average over these elicited distributions in order to obtain a single, unified prior distribution (either for each sample or across samples of respondents), as is standard practice in focus-group elicitation settings, averaging would allow outlying responses to inordinately influence the ultimate distribution. This paper instead uses a Dirichlet Process clustering framework in order to identify latent "schools of thought" within the elicited prior distributions and aggregate the clusters of "schools of thought" rather than the individual distributions themselves.

Undertaking this clustering process, however, requires care to specify the sources and types of uncertainty present in the data collection process itself. Not only is the phenomenon under study uncertain, but the process of elicitation itself introduces measurement uncertainty with respect to experts' beliefs about the phenomenon under study. In order to simplify the elicitation procedure and conduct it in a survey framework, rather than in a traditional focus group setting with significant feedback from a facilitator, the analysis must account for the uncertainty experts experience in placing chips into bins to reflect their beliefs. Likewise, because of the desire to aggregate these elicited beliefs across clusters, uncertainty in terms of the clustering process must be accounted for. Uncertainty in this analysis thus arises in each of four levels across the elicitation:

1. **Underlying probabilities:** Some underlying probability determines the vote share of either the Presidential or Congressional candidates in the elicitation survey. Elicitation attempts to provide an estimate of these probabilities, subject to question structure and the correct specification of relevant covariates.

2. **Cluster probabilities:** The elicitation framework adopted here assumes that there are several possible "schools of thought" governing which variables impact vote share and to what degree. These clusters arise from the underlying probability governing the phenomenon itself. Implementing a basic Dirichlet Process to allocate experts to clusters with a prior on the concentration parameter suggesting that each expert may form their own cluster produces a large number of clusters, implicitly suggesting that the chip allocation for a given expert must be *exactly* the same in order for experts to share a cluster.

This assumption is implausible for reasons discussed below related to chip allocation. Instead, we relax this assumption to allow more uncertainty in cluster parameters. This in turn induces uncertainty in expert allocation to clusters.

3. **Expert probabilities:** Just as the clusters reflect differing perspectives on the underlying distribution, experts are considered "draws" from given clusters. The definition of clusters is a result of an iterative sampling process and is thus probabilistic; that is, an expert may be in one "school of thought" 80% of the time and in another 20% of the time depending on their chip allocation relative to others in the sample.

4. **Chip probabilities:** Likewise, chips effectively represent "draws" from the prior distribution each expert has in mind to describe the underlying probability process. Chips placed in the roulette bins in our survey are therefore also subject to uncertainty. We assumes chips are drawn from a multinomial distribution according to the bin segmentation we defined in the survey. That is, the bins provide arbitrary cutpoints in a continuous distribution between 0% and 100% vote share, and the probability of a chip being allocated to a particular bin reflects some uncertainty about the probability mass. For example, a given respondent may have believed that the vote share Hillary Clinton would receive in Florida was closest to 50%. Because our bin labels are 46–50% and 51–55%, however, this respondent may feel that allocating chips to the higher (51–55%) bin is more reflective of their assessment of 50% because they believe 46% is unlikely. This type of logic suggests a correlation between neighboring bins that we later account for using a Gaussian Process on a logistic scale. This approach has several advantages in addition to main-

taining conjugacy. While a Dirichlet Process would generate a "clumpy" distribution of bin allocations, the Gaussian Process provides some smoothing that can reflect correlation between neighboring bins, and the logistic scale accommodates the fact that the Gaussian is technically continuous, whereas our probabilities cannot be negative.

Procedurally, then, while the distribution of clusters is modeled with a Dirichlet, we model the difference between cluster distributions and expert distributions as Gaussian, and the difference between expert distributions and chip distributions as Gaussian as well in order to take advantage of conjugacy. The likelihood of a given number of chips in a bin is considered proportional to a multinomial distribution. Including a normalizing constant, if the underlying distribution of probabilities for the bins is given by π, and clusters are given by ρ, then

$$\rho_{ikq} = \frac{\pi_{ikq} e^{\epsilon_{ikq}}}{\sum_j \pi_{jkq} e^{\epsilon_{jkq}}}$$

where i indexes bins (for j in 1 to J, where J is total bins), k indexes clusters, and ℓ indexes experts. $g(\ell)$ in turn provides the group (cluster) assignment for expert ℓ, and errors $\epsilon_i \sim N(0, \Sigma)$.

Experts, in turn, are given by ν so that

$$\nu_{i\ell q} = \frac{\rho_{ig(\ell)q} e^{\xi_{i\ell q}}}{\sum_j \rho_{jg(\ell)q} e^{\xi_{j\ell q}}}$$

where errors $\xi_i \sim N(0, \mathcal{T})$. In this initial analysis Σ and $\mathcal{T}$ are treated as non-correlated diagonal matrices, where $\mathcal{T}$ is fixed so that experts irrespective of cluster give consistent responses. Inducing correlation between Σ and $\mathcal{T}$ later will generate more correlation in the errors across bins. In addition, while these values are currently considered fixed, later extensions will update them dynamically in the sampling process to provide a more realistic outcome. Still, this framework allows cluster variance to differ from error variance, and allows both to vary by cluster as well as question in the analysis. That is, cluster variance captures the differences in priors across clusters, while error variance captures the differences or inconsistencies in answers within clusters. The separate cluster variance and error variance in the model in future analyses can provide useful heuristics for refining aggregation and validating cluster predictions, as discussed in the conclusions.

Implementing this approach requires modeling the multinomial probabilities in our data with a Pólya-gamma sampler as described in Polson, Scott, and Windle (n.d.). A Dirichlet-multinomial approach does not model dependency between draws, as we have in our data, whereas a Pólya-gamma augmentation reframes the multinomial distribution in terms of random auxiliary variables whose likelihoods are Gaussian (Linderman, Johnson, and Adams 2015), and this allows us to maintain conjugacy.[4]

To incorporate covariate conditions, furthermore, constants are added to the estimation process to distinguish the underlying distribution for each individual (one that does not depend on covariates), as well as the individual and cluster variances, before including covariate conditions.

3.5 Results

In order to validate the clustering process for elicited priors across both sets of experts, prior bin counts from each survey are incorporated into the Gaussian and Dirichlet processes. The output of this process includes two parameters for a beta distribution corresponding to each question asked in the survey. The preliminary

[4] Additional description of this method to be included in a future extended Appendix.

analysis, reported here includes the first eight survey questions asking about electoral outcomes at the national and state level for 3 swing states (Florida, Ohio, and North Carolina), as well as a battery of questions including hypothetical covariate values (state, unemployment, campaign ads, terrorist attacks, etc.). Parameters are averaged over 4000 samples each. Future analysis will include not only questions with these covariates applied to the presidential election across both the mass and elite surveys, but also the U.S. House hypothetical conditions in the elite sample.

The figures below show the clustered priors for the mass and elite survey samples in comparison, for questions with and without covariates. These priors are compared to the 2016 electoral outcomes at the national level and by state, including both the true vote proportion as well as the two-party vote proportion (New York Times 2017; Wasserman). As the figures illustrate, in general both the mass and elite samples, aggregated with this clustering process, perform fairly well.

3.5.1 Overall Mass vs. Elite Results

Figure 3.9 provides an overall comparison of the mass and elite samples in state and national contexts without additional covariates (that is, for questions where terrorism, campaign ads, and unemployment levels are not specified). The point for each mass or elite line segment represents the mean of 4000 samples, while the line segments represent the 95% bounds across samples. This figure provides initial evidence that the elite sample need not outperform even a small representative population sample, potentially indicating that "expertise" need not perfectly correlate with educational credentials. Furthermore, that the mean prediction across samples aligns well with eventual vote share suggests three promising conclusions: that elicitation has value for recovering accurate information about quantities of interest; that the clustering method proposed in Chapter 2 aggregates sometimes disparate individual assessments from survey respondents into a reasonable prediction of vote share; and that the elicitation protocol (roulette, via online survey) does not significantly compromise estimates.

Figure 3.10 refines this preliminary view, providing more nuance to distinguish between samples by including covariate values. For these questions in the survey, respondents in each sample were presented with hypothetical conditions that might influence vote share, such as the number of terror attacks occurring globally in the month before the election; the state unemployment rate; and the advertising margin of the candidate in the question. These questions provide implicit information that cues respondents to weigh the relative impact of security, economic policy, and campaign strategy in the eventual vote margin of the candidate. One immediately evident trend in these results is that the elite sample outperforms mass respondents with respect to Donald Trump's vote share, while the samples perform comparably in their assessments of Clinton's prospects.

Figures 3.11 through 3.15, furthermore, demonstrate that the elite sample consistently has lower variance in their estimates than the mass sample, with the range much more narrow around the true and two-party electoral outcomes. This trend is more evident in the prior distribution plots in the next section and the appendix. Even so, results from Ohio and Wisconsin, with covariates, illustrate circumstances under which the mass sample can either outperform or achieve parity with the elite sample, particularly in evaluating Clinton's potential vote share. In the case of Ohio, the mass sample mean is often more accurate even as the range is wider. In the case of Wisconsin, some distinctions across covariate conditions suggest how mass and elite samples may have weighed economic and security conditions differently. For example, the October 2016 unemployment rate was closest to the 4% covariate condition (Bureau of Labor Statistics 2016), and Clinton had a consistently greater ad margin than Trump throughout the entire campaign (Associated Press 2016). At least in the case of Clinton's vote share, the mass sample mean is consistently closer to a true value under the 4% unemployment covariate condition, while the elite sample means do not differ very significantly across all covariate conditions. The unemployment condition plausibly reflects a lived experience from which experts may be able to offer more accurate assessments, whereas the terrorism levels and ad margin may have been interpreted more hypothetically, or have the potential to distinguish respondents on the basis of their access to or consumption of news media. Despite this fungibility in how respondents might interpret covariate conditions, the more implausible a condition or set of conditions, the greater the variation in responses to elicitation one should expect. This uncertainty is more evident in the figures in the next two sections.

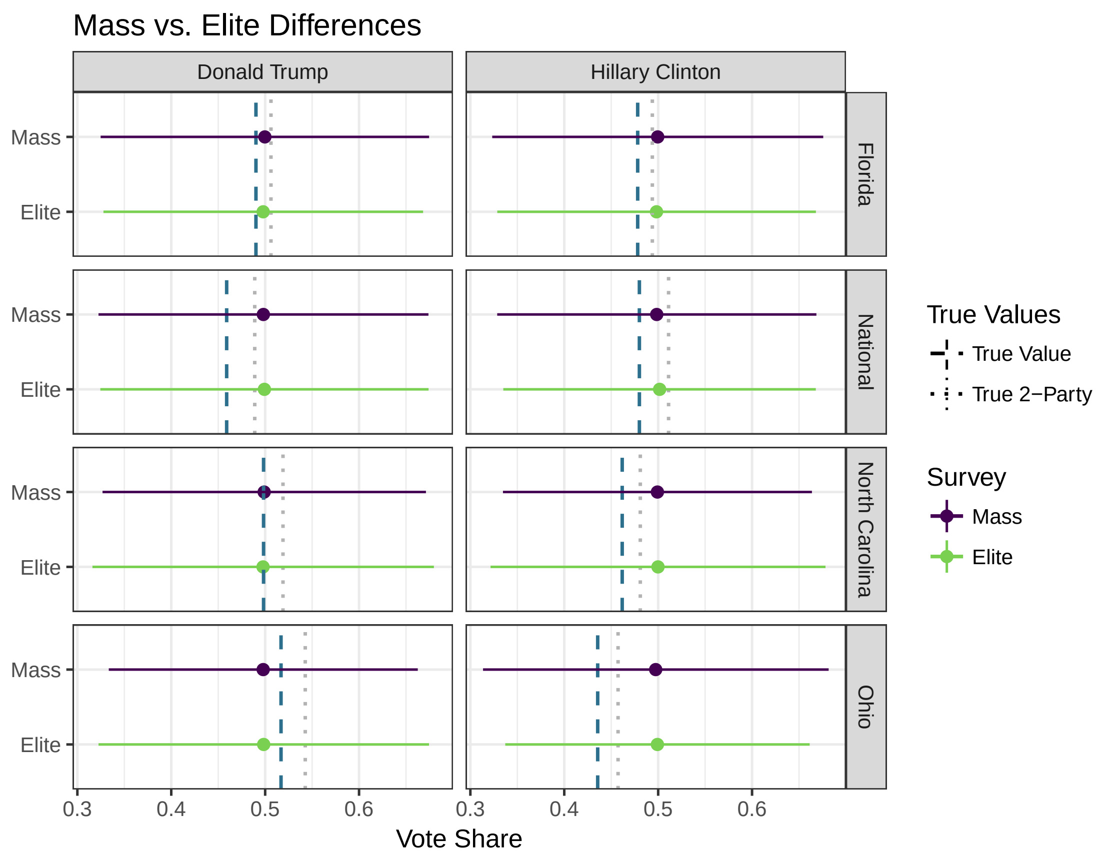


Figure 3.9: Mass vs. Elite Survey without Covariates

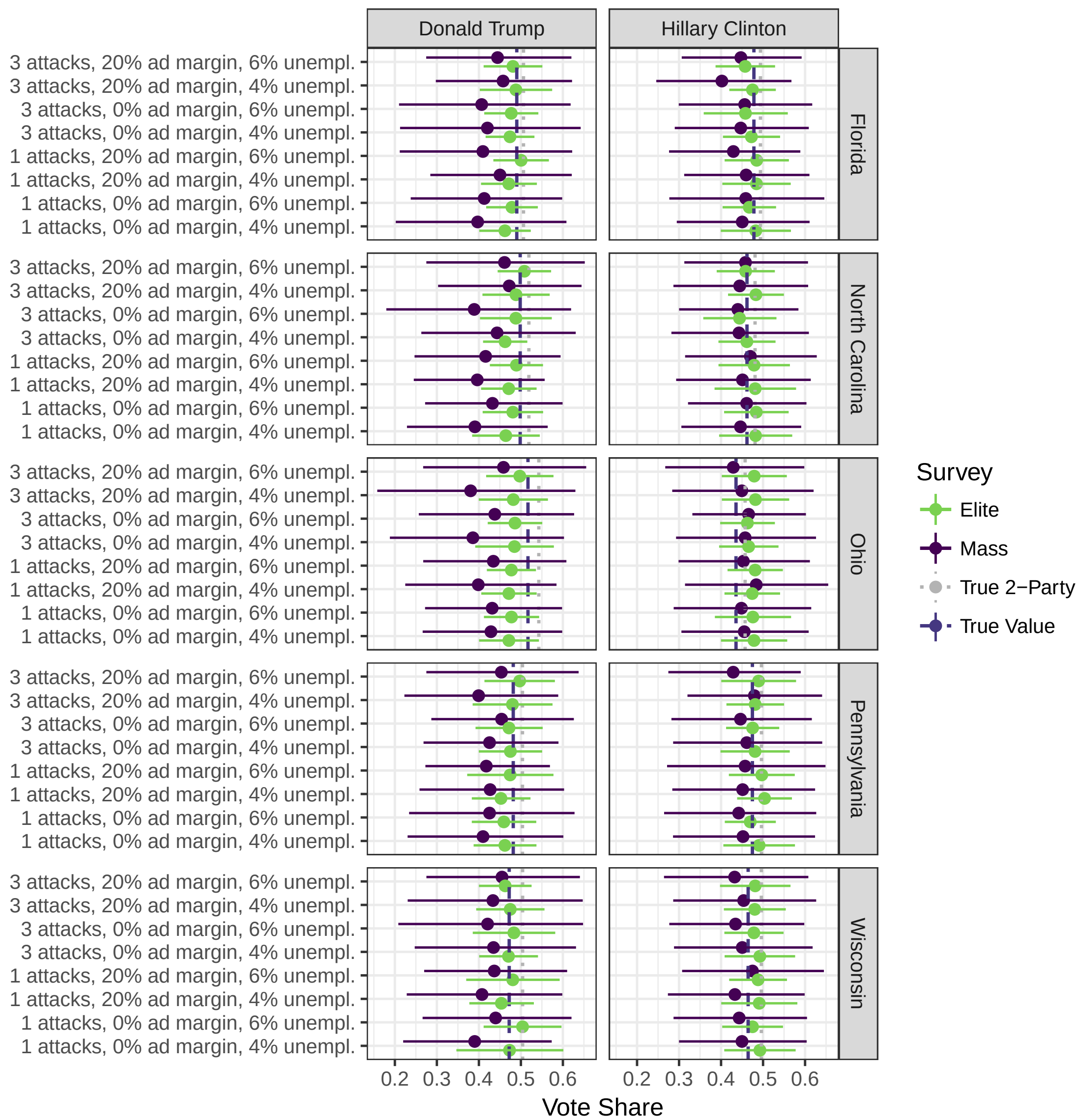


Figure 3.10: Mass vs. Elite Survey with Covariates

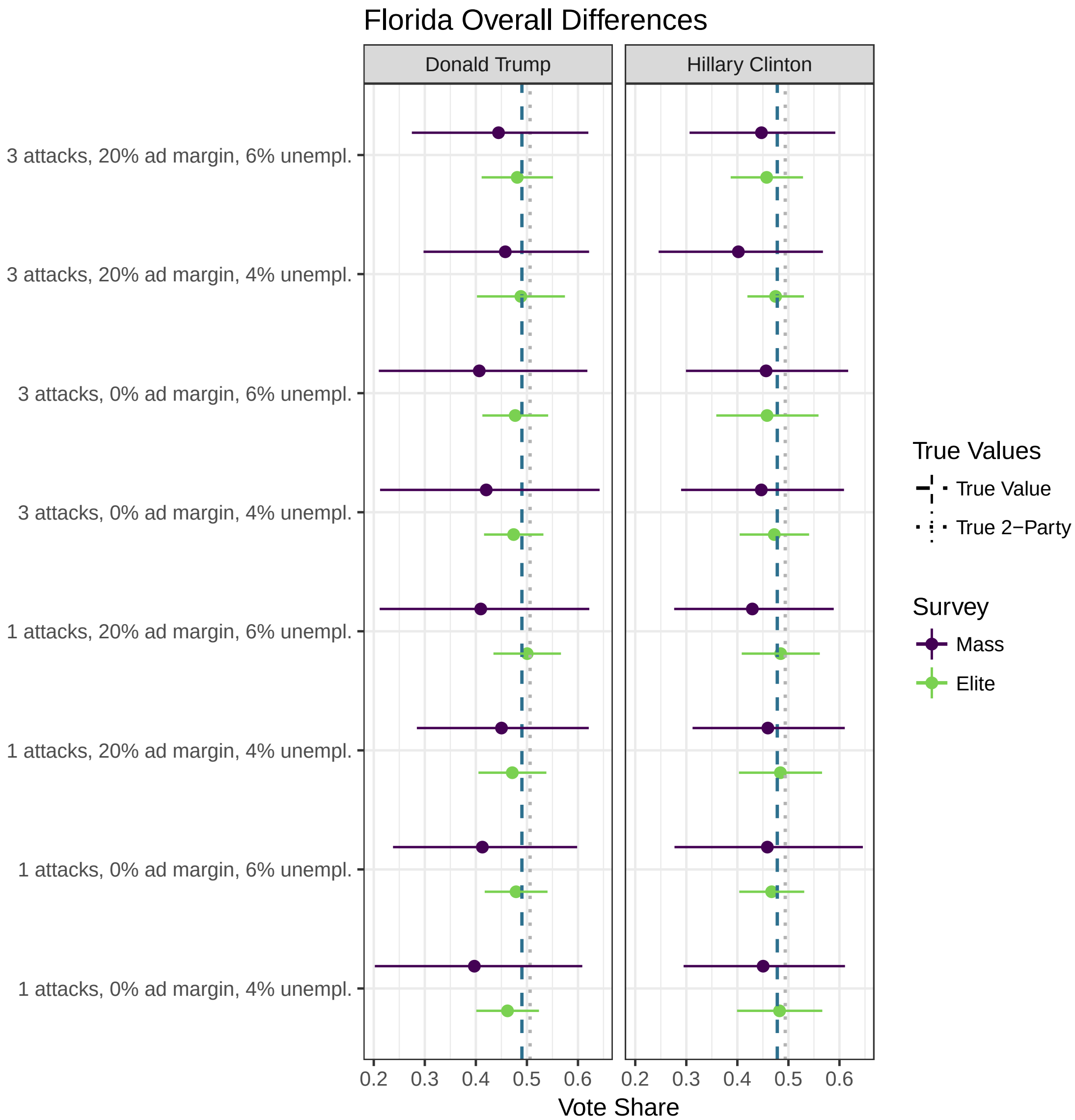


Figure 3.11: Mass vs. Elite Survey FL with Covariates

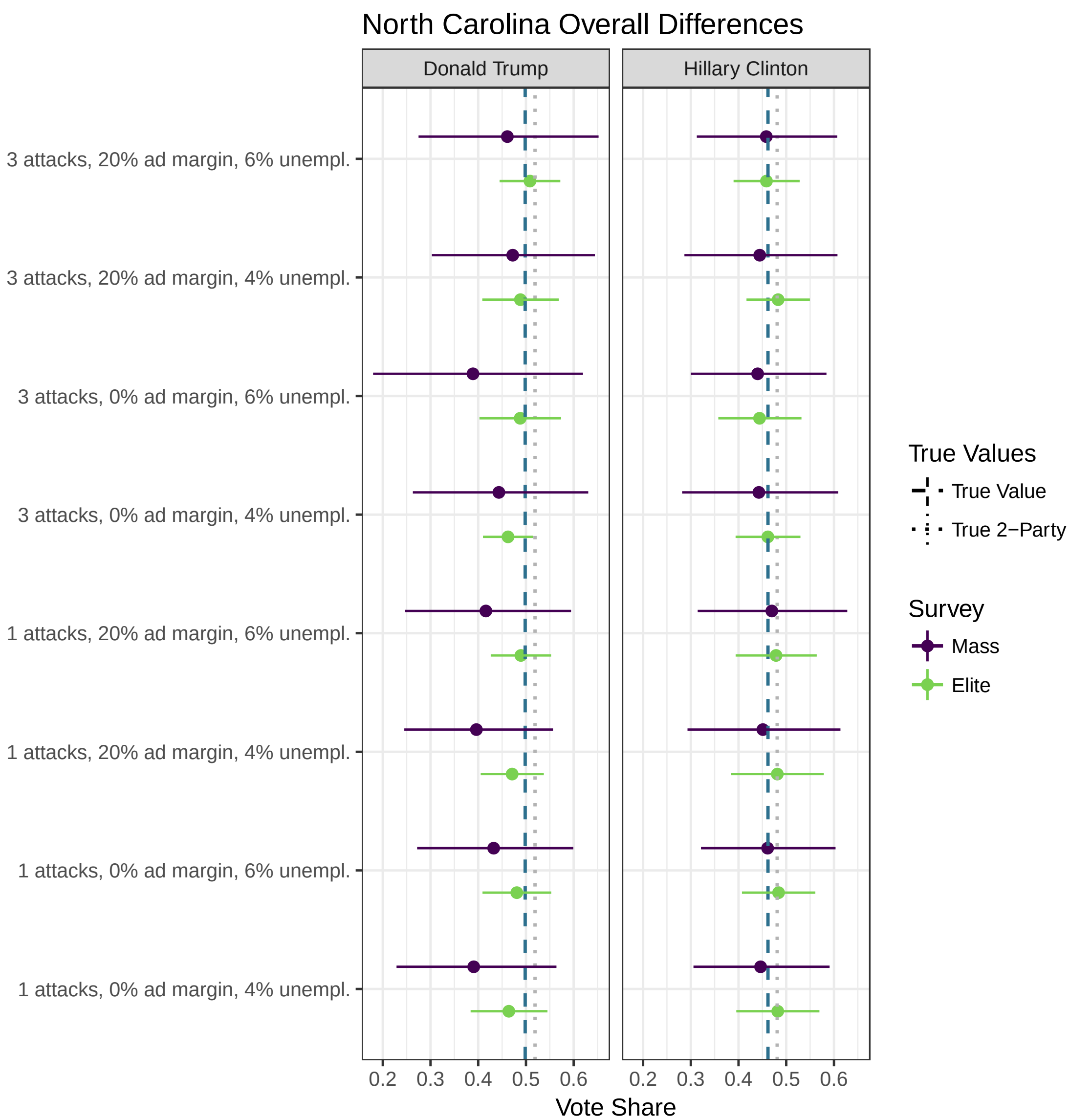


Figure 3.12: Mass vs. Elite Survey NC with Covariates

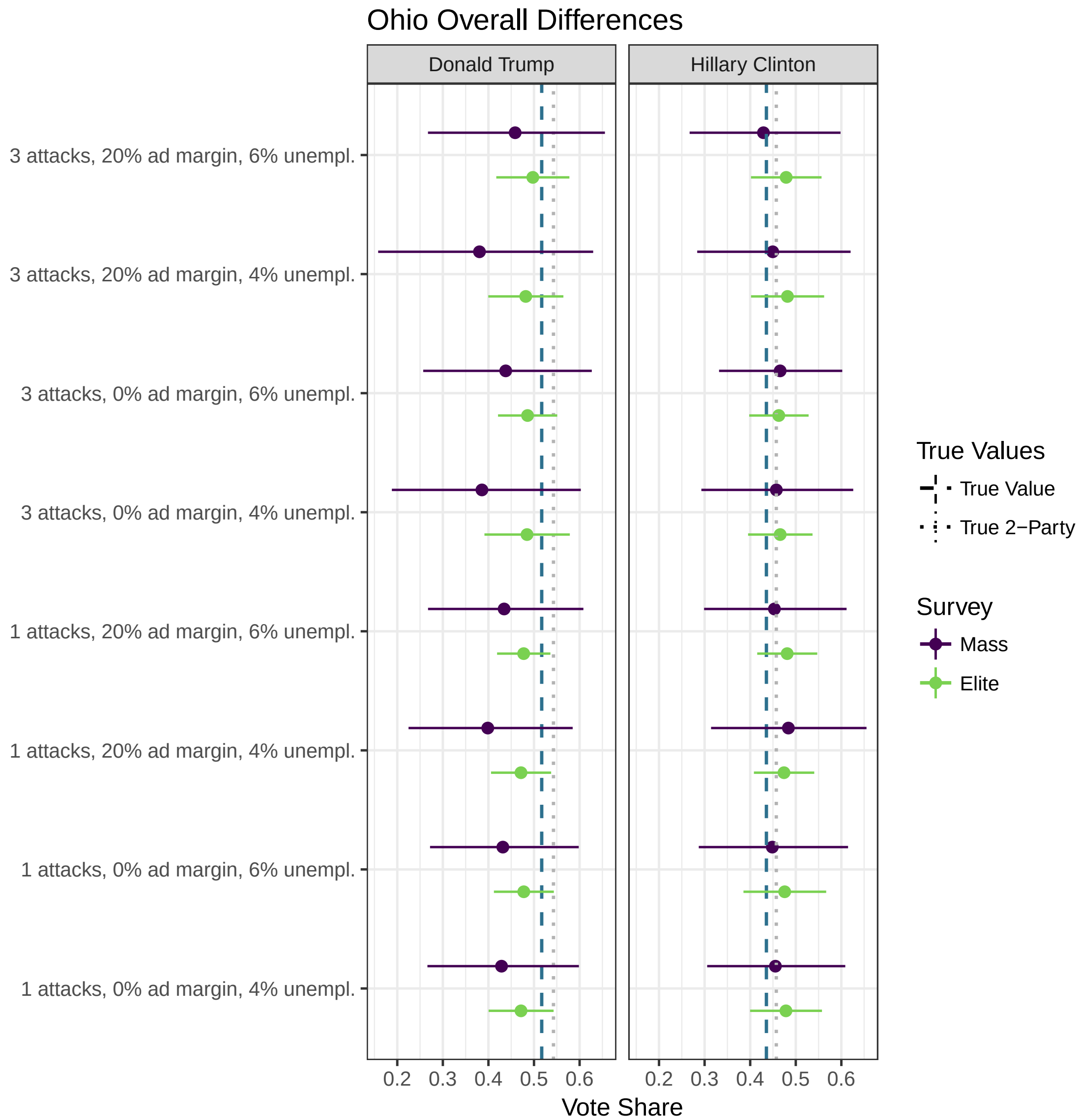


Figure 3.13: Mass vs. Elite Survey OH with Covariates

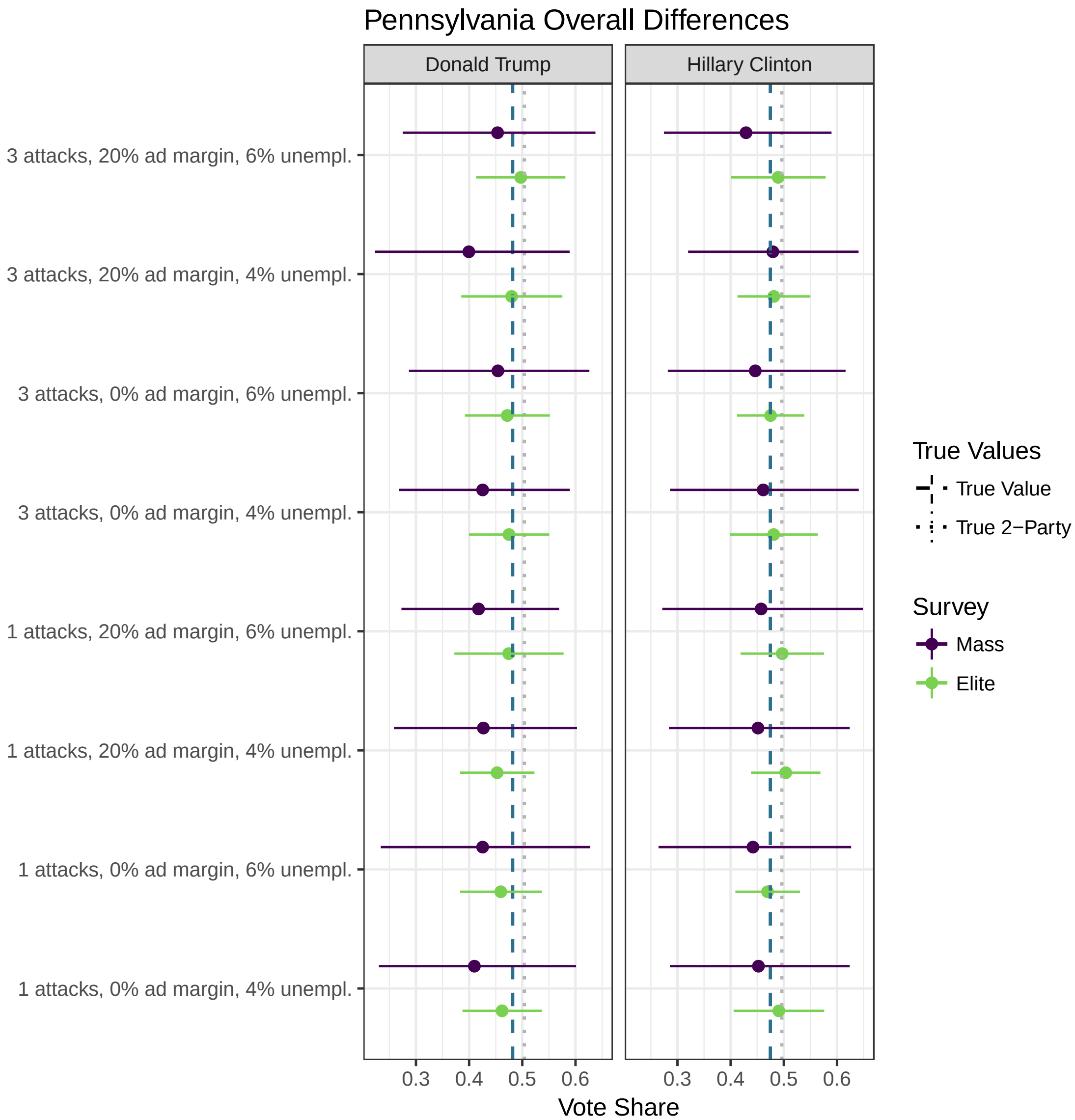


Figure 3.14: Mass vs. Elite Survey PA with Covariates

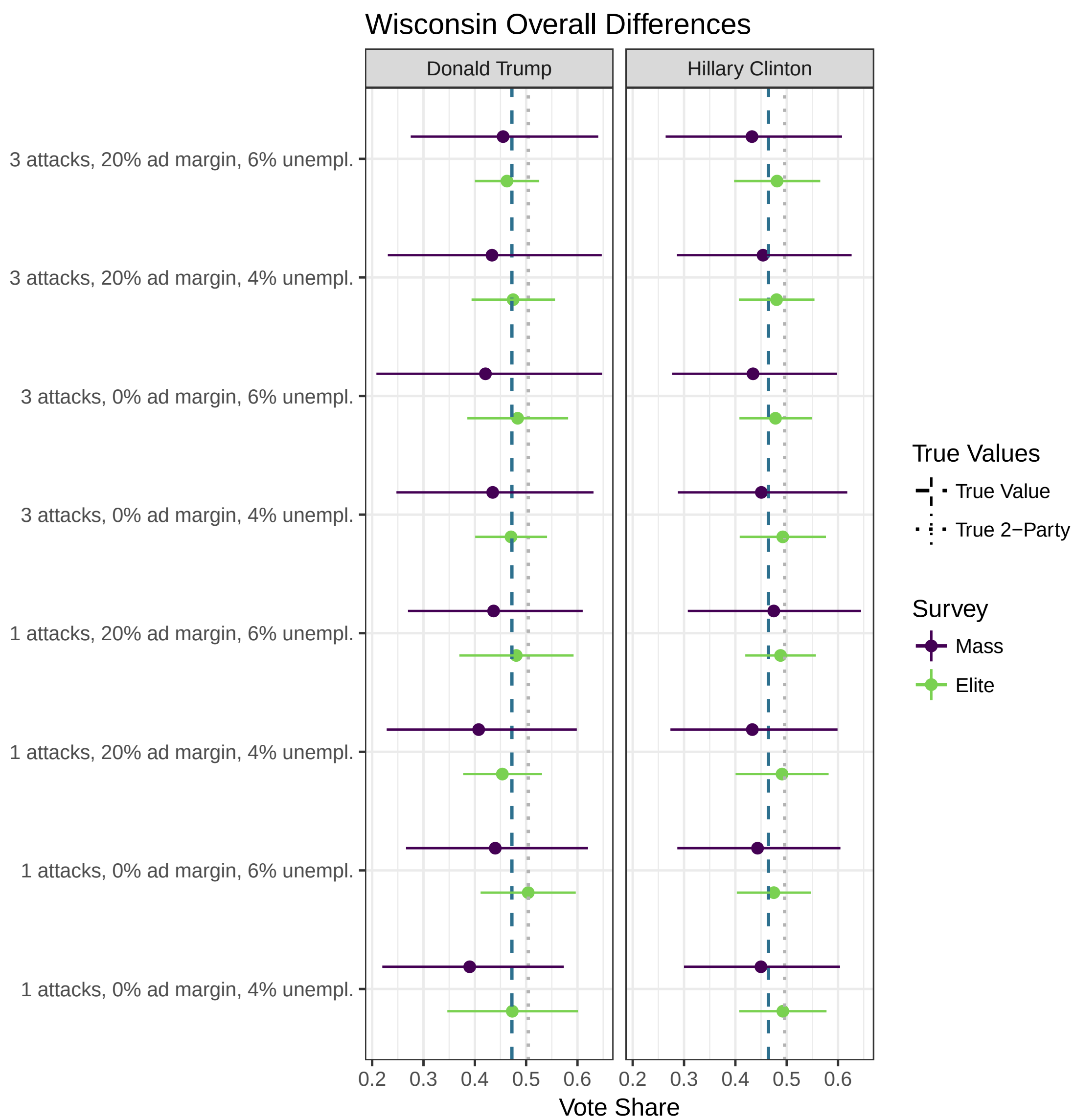


Figure 3.15: Mass vs. Elite Survey WI with Covariates

3.5.2 Overall Results: No Covariates

The range of responses and uncertainty in estimates across mass and elite samples is better illustrated through the plots of the clustered prior distributions in the figures that follow. As with the difference plots previously discussed, the results without covariates demonstrate few differences between the mass and elite samples across candidates and states; in fact, both samples appear to reflect the same biases with respect to the true vote share outcomes.

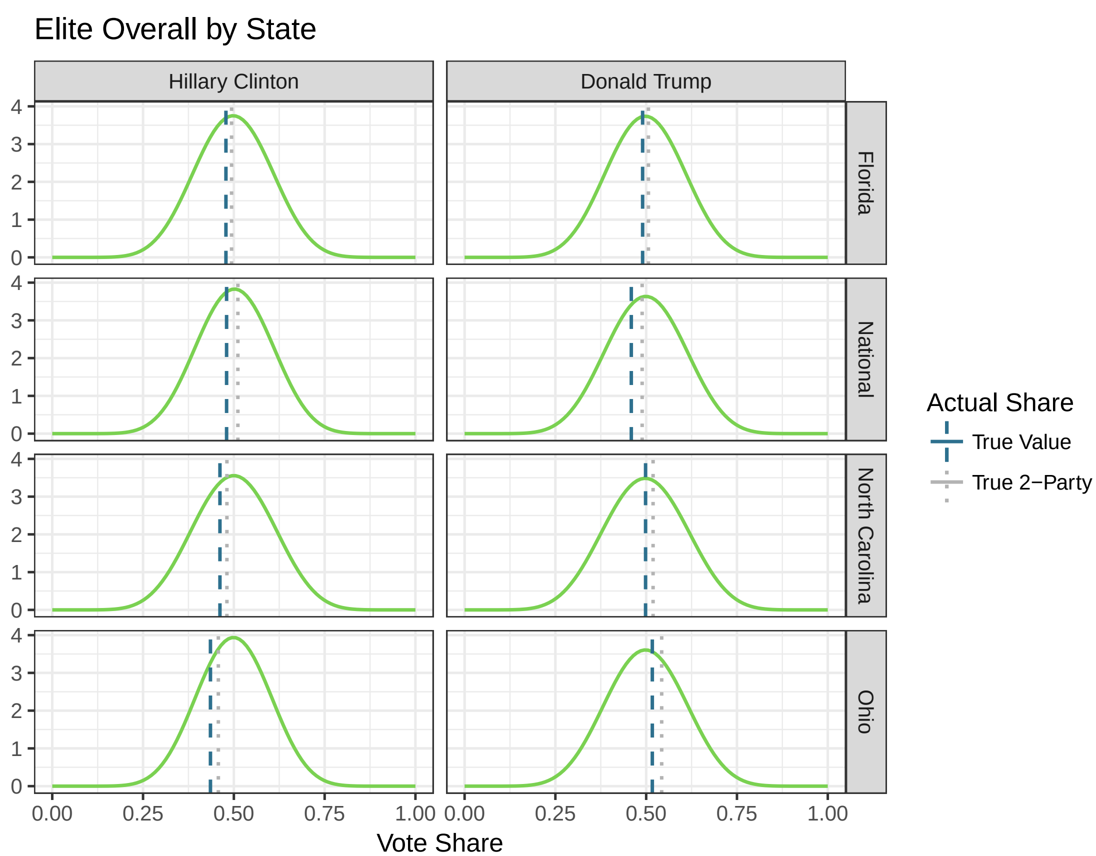


Figure 3.16: Overall: Elite overall without Covariates

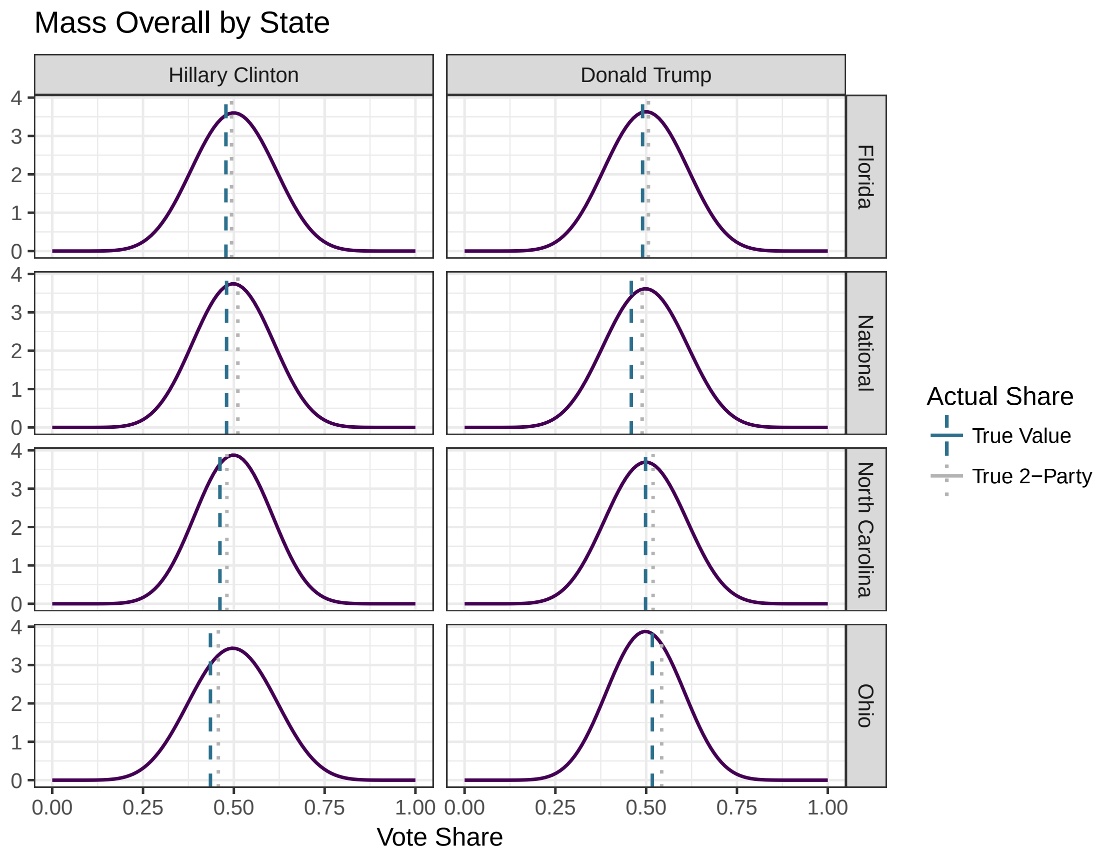


Figure 3.17: Overall: Mass overall without Covariates

3.5.3 Overall Results: With Covariates

In the following figures reflecting prior distributions under covariate conditions, however, the relative certainty of the elite versus mass samples becomes clearer. Irrespective of state or covariate conditions, the elite sample reflects greater certainty, with a narrower distribution of estimates, than the mass sample. Notably, the mass sample appears further from the true values in their assessments of Trump across states (always underestimating his vote share), although skew in the distribution means that the mean across samples is not as inaccurate as the density suggests. Despite having more diffuse priors than the elite sample, the mass and elite priors regarding Clinton often coincide with the true values, as previously discussed. At the same time, the difference in dispersion between the overall mass and elite priors is only truly evident when introducing covariate conditions, which potentially arises from greater difficulty or discomfort with hypothetical conditions or probabilistic assessments among respondents in the mass sample, particularly because questions with covariates followed more generic questions without. This supposition does not, however, explain the greater uncertainty with respect to Trump's electoral performance. Moreover, this explanation suggests that more educated members of the mass subset, or perhaps those with greater political knowledge, should perform better and/or have greater certainty in their expected outcomes. The plots in the following section provide evidence to evaluate this claim, delineating each sample according to demographic characteristics.

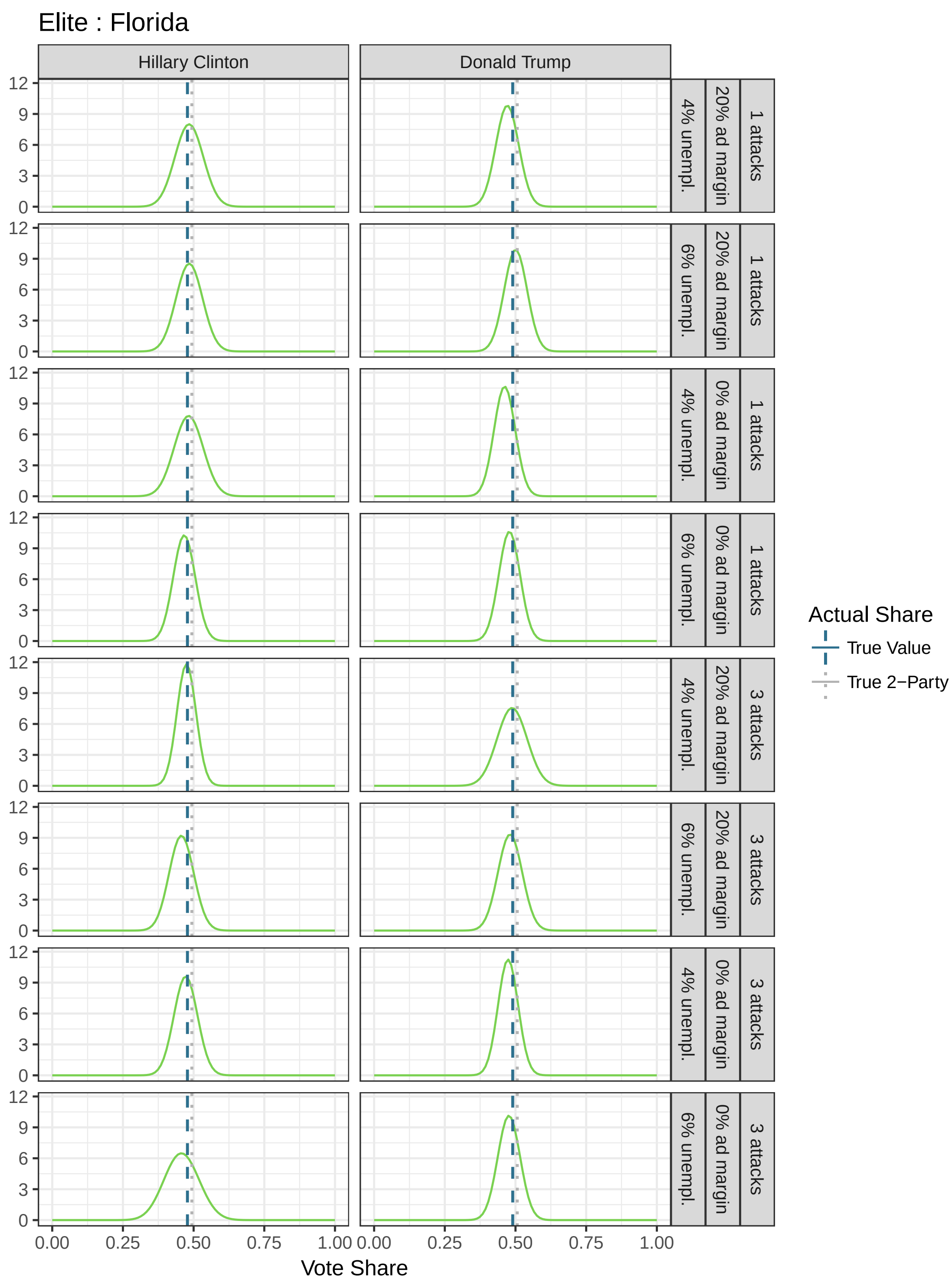


Figure 3.18: Overall: Elite with Covariates FL

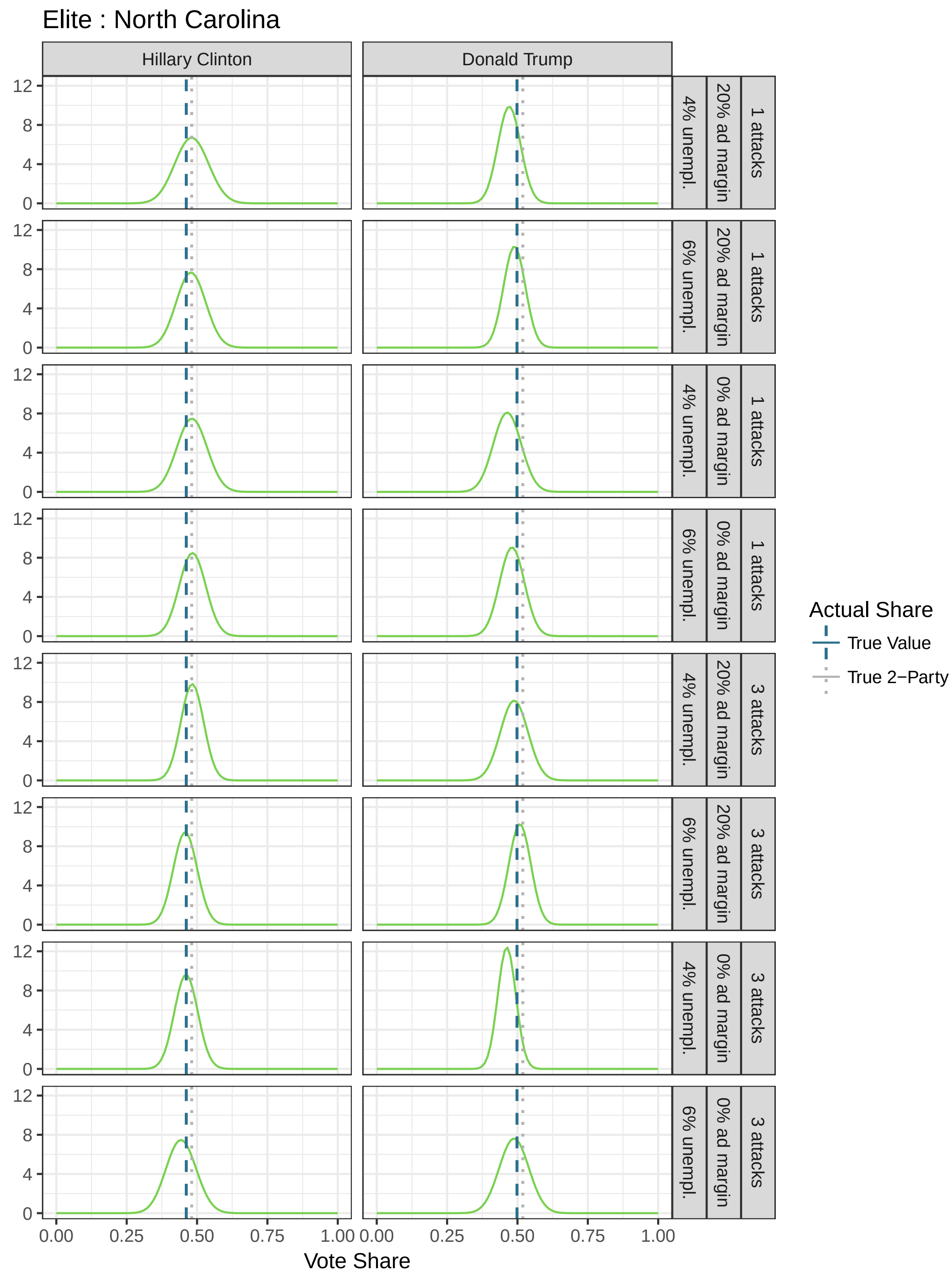


Figure 3.19: Overall: Elite with Covariates NC

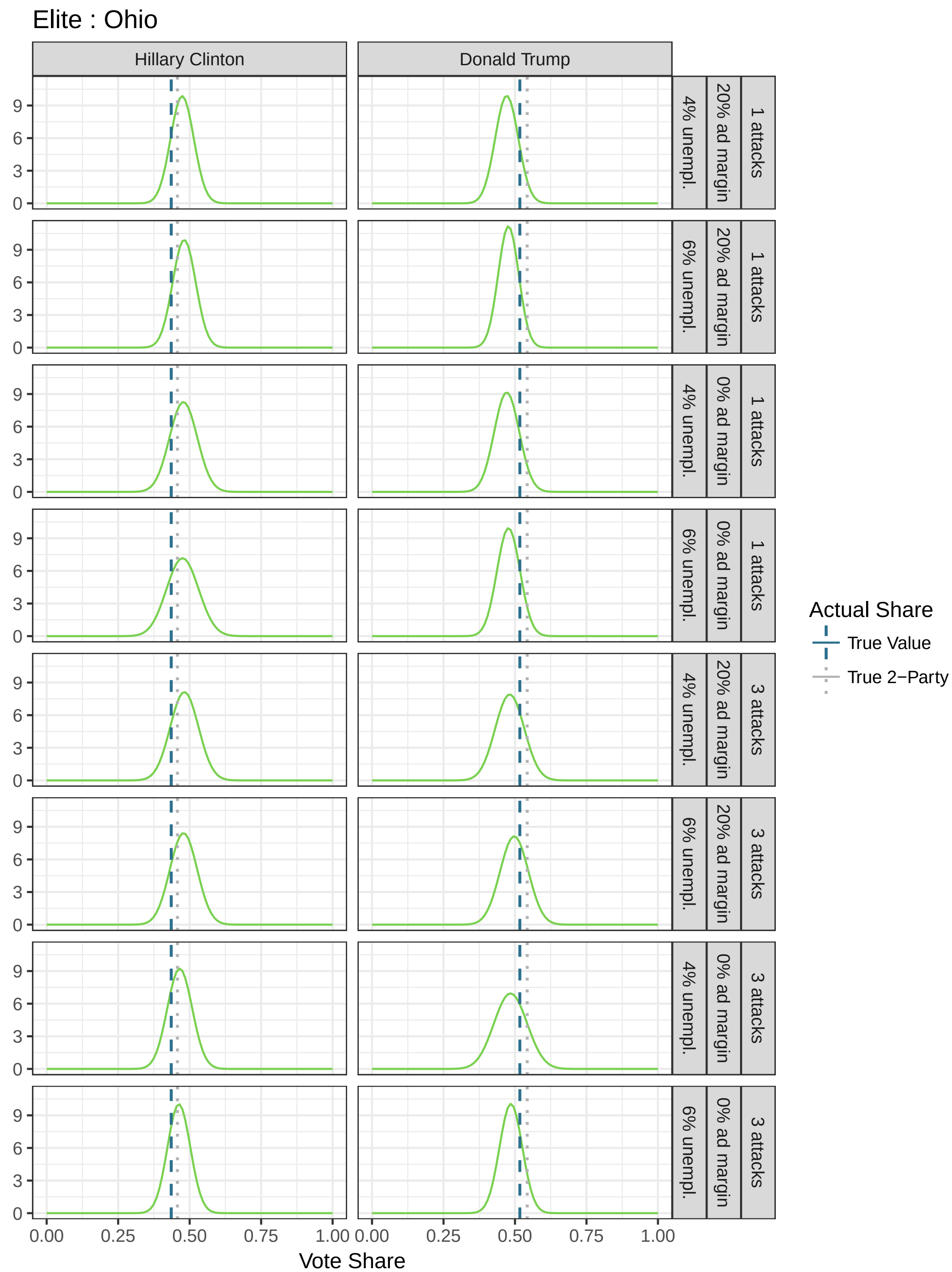


Figure 3.20: Overall: Elite with Covariates OH

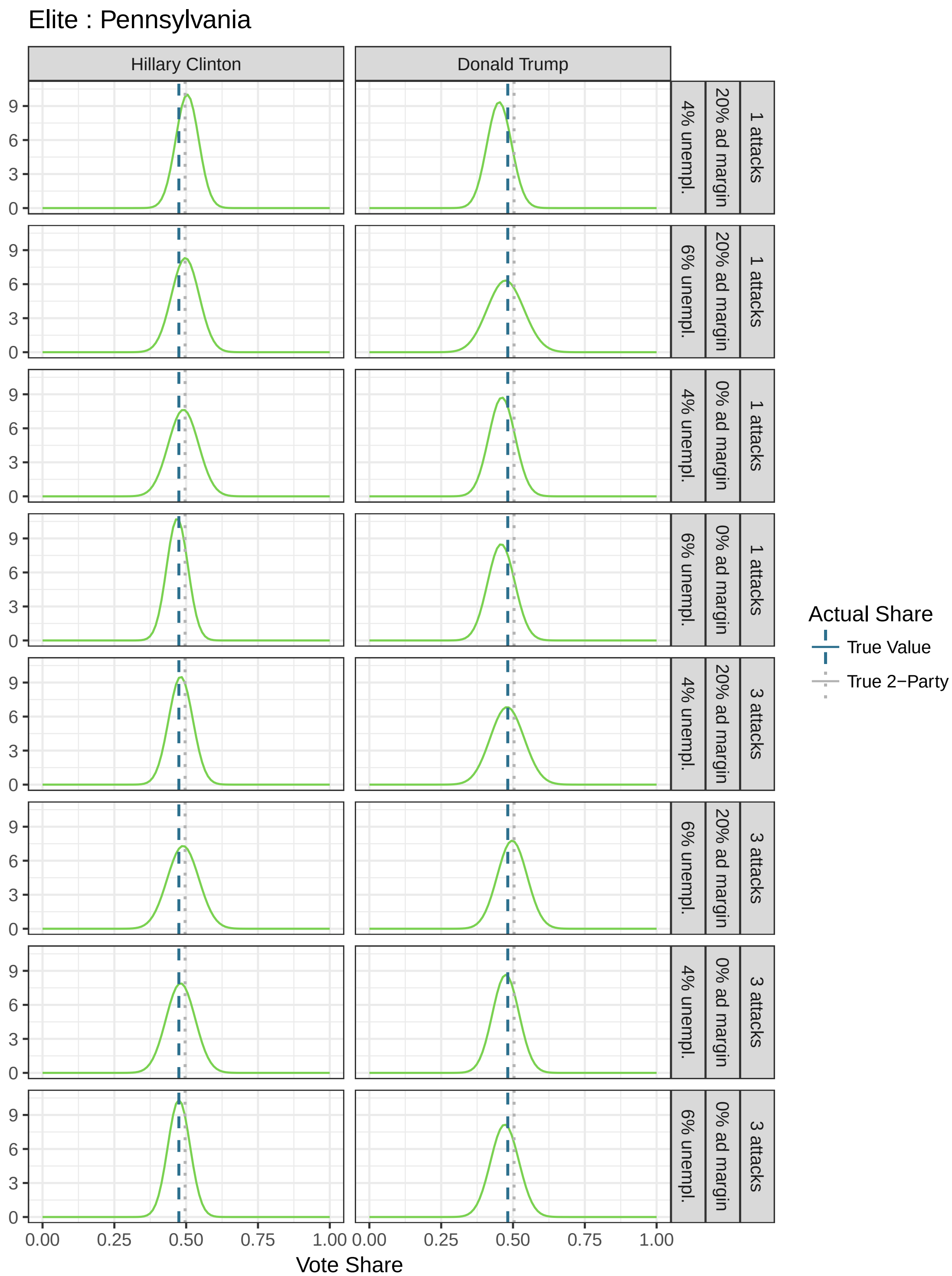


Figure 3.21: Overall: Elite with Covariates PA

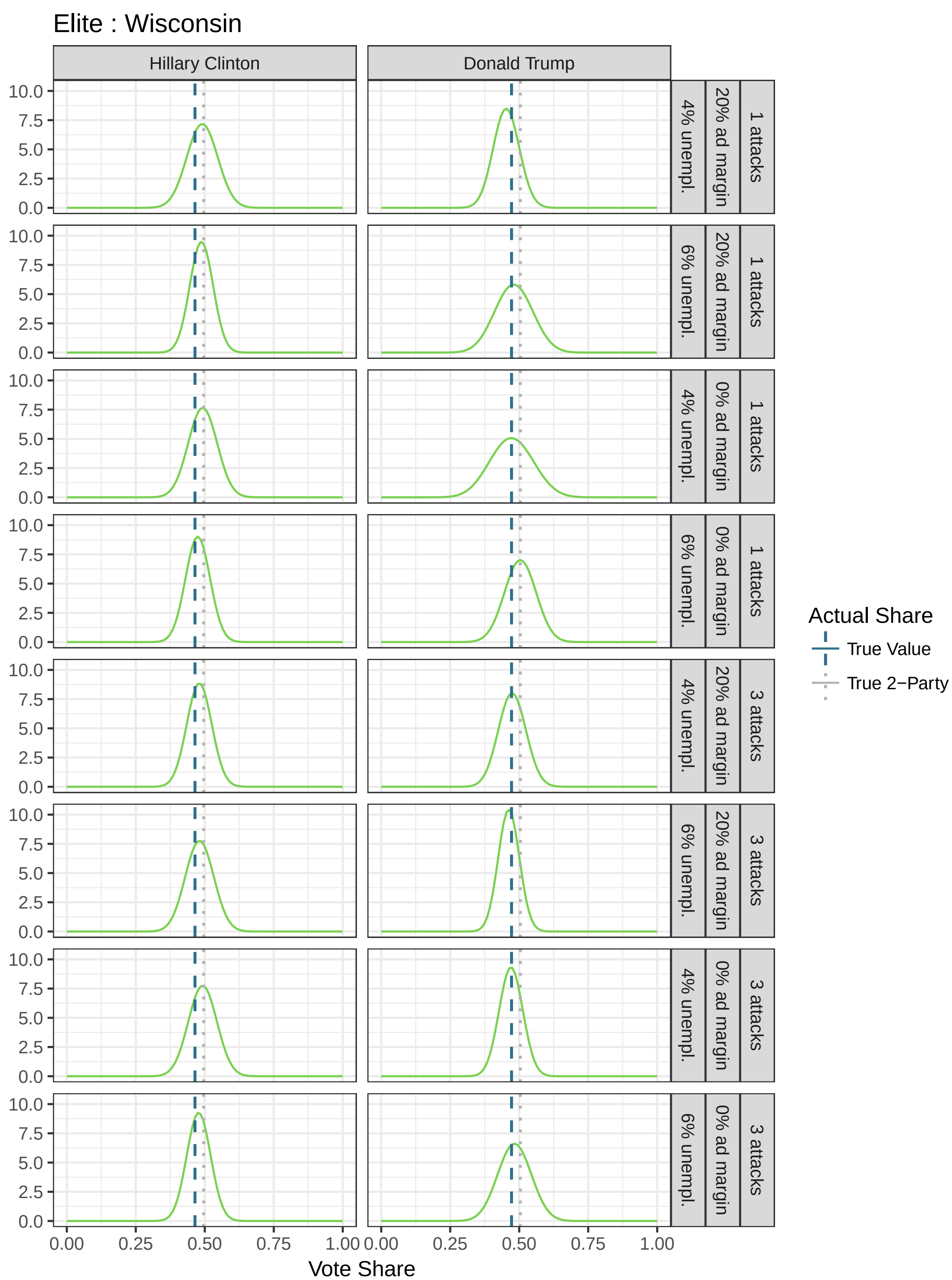


Figure 3.22: Overall: Elite with Covariates WI

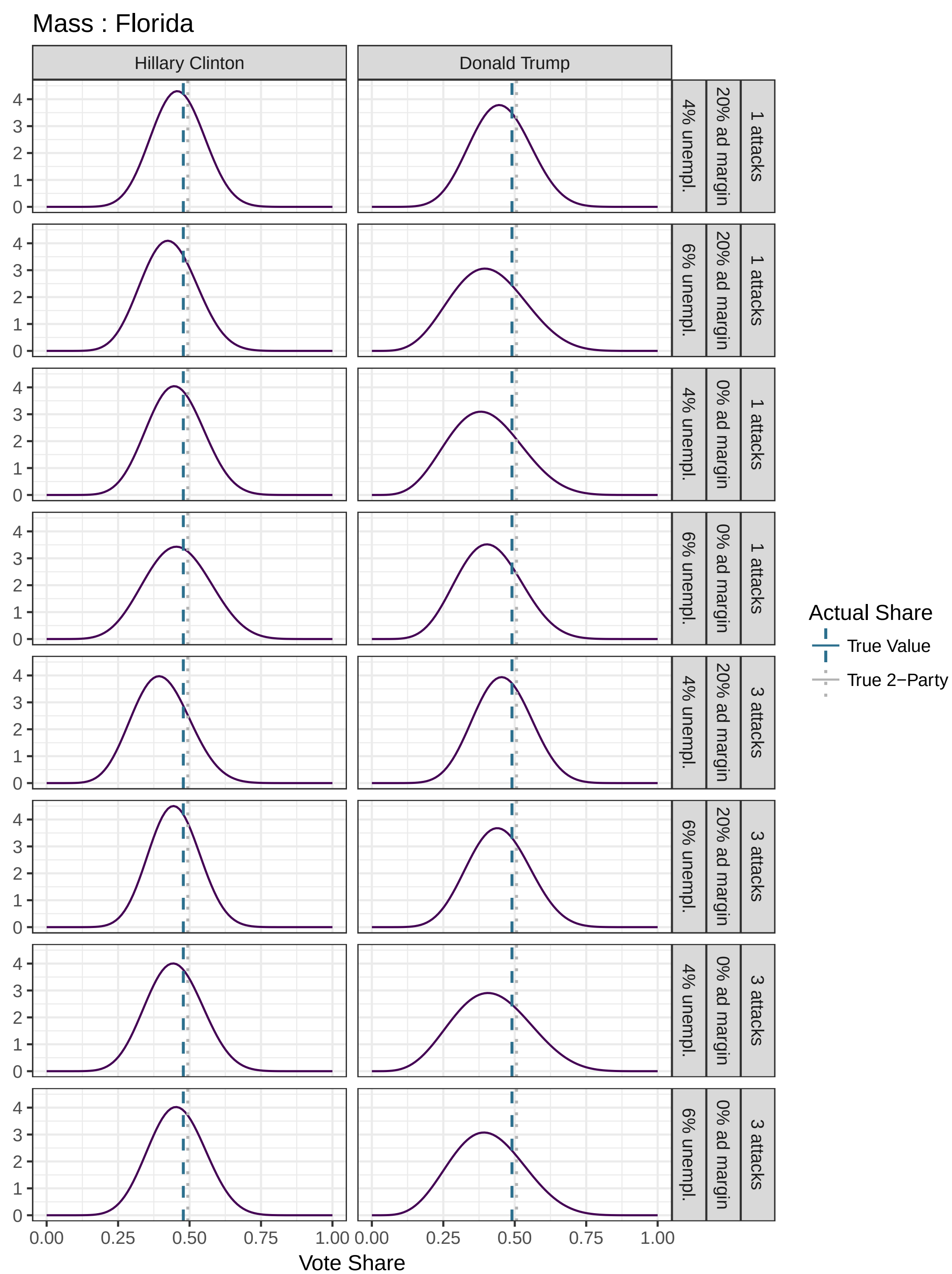


Figure 3.23: Overall: Mass with Covariates FL

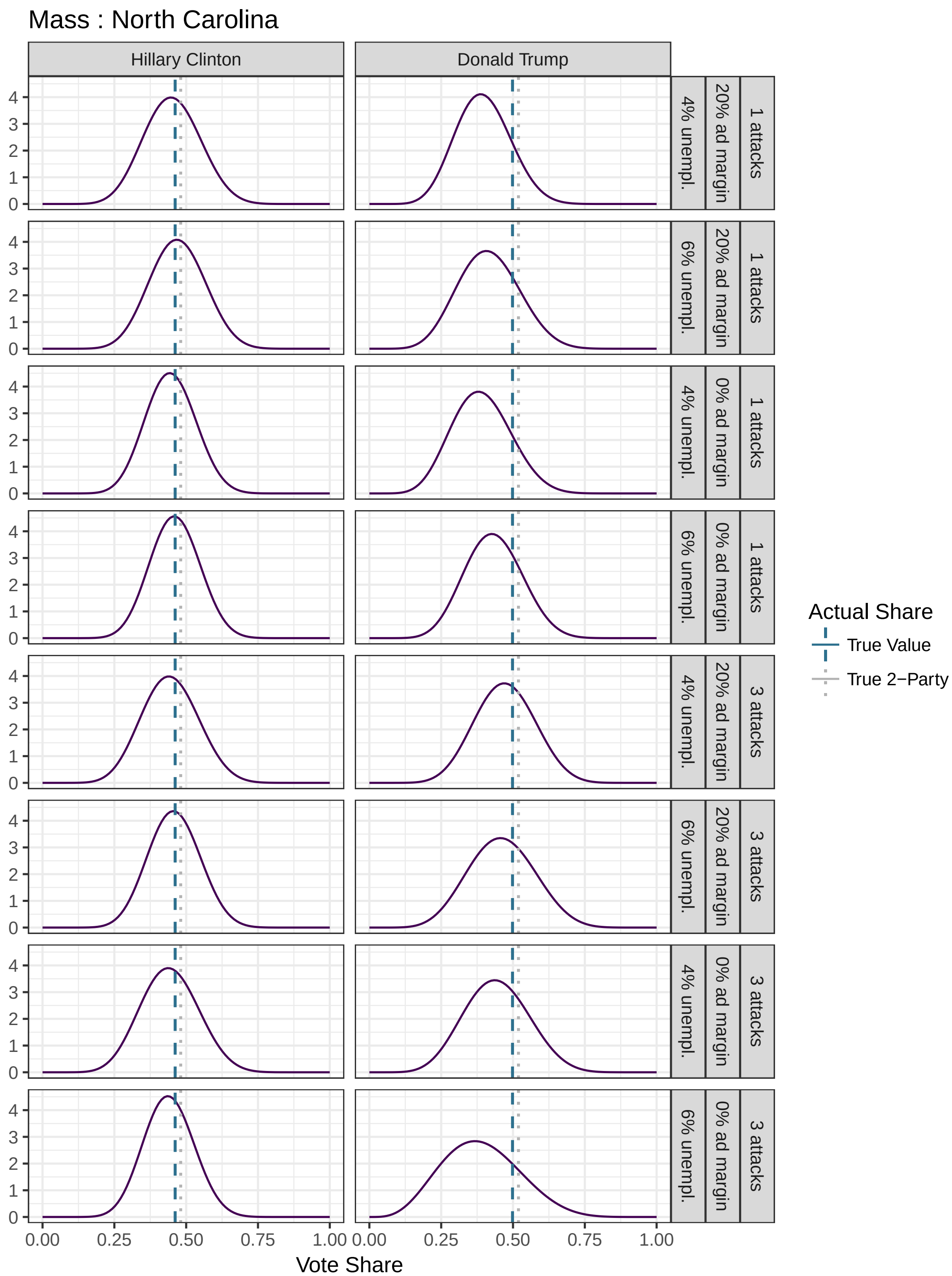


Figure 3.24: Overall: Mass with Covariates NC

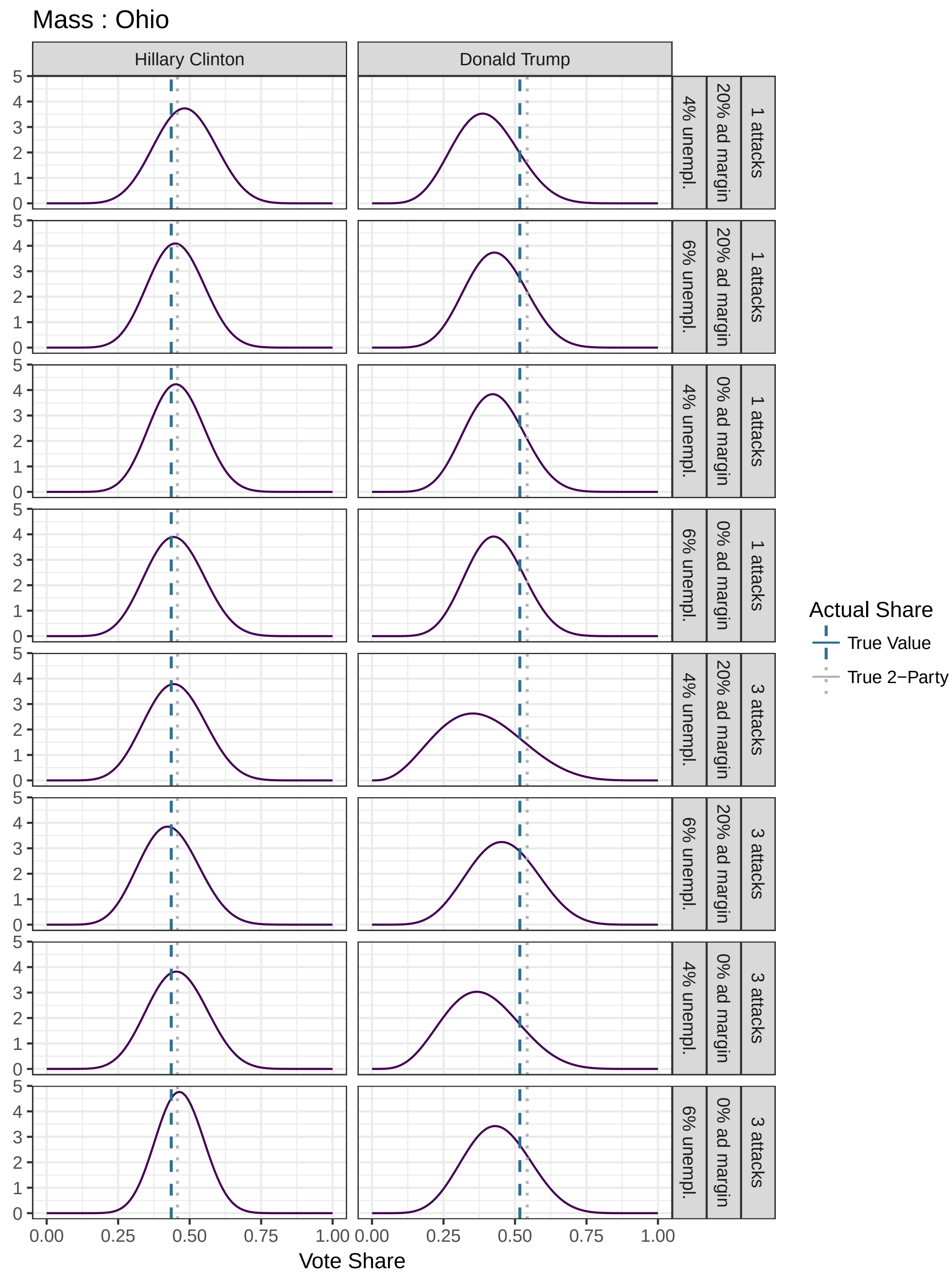


Figure 3.25: Overall: Mass with Covariates OH

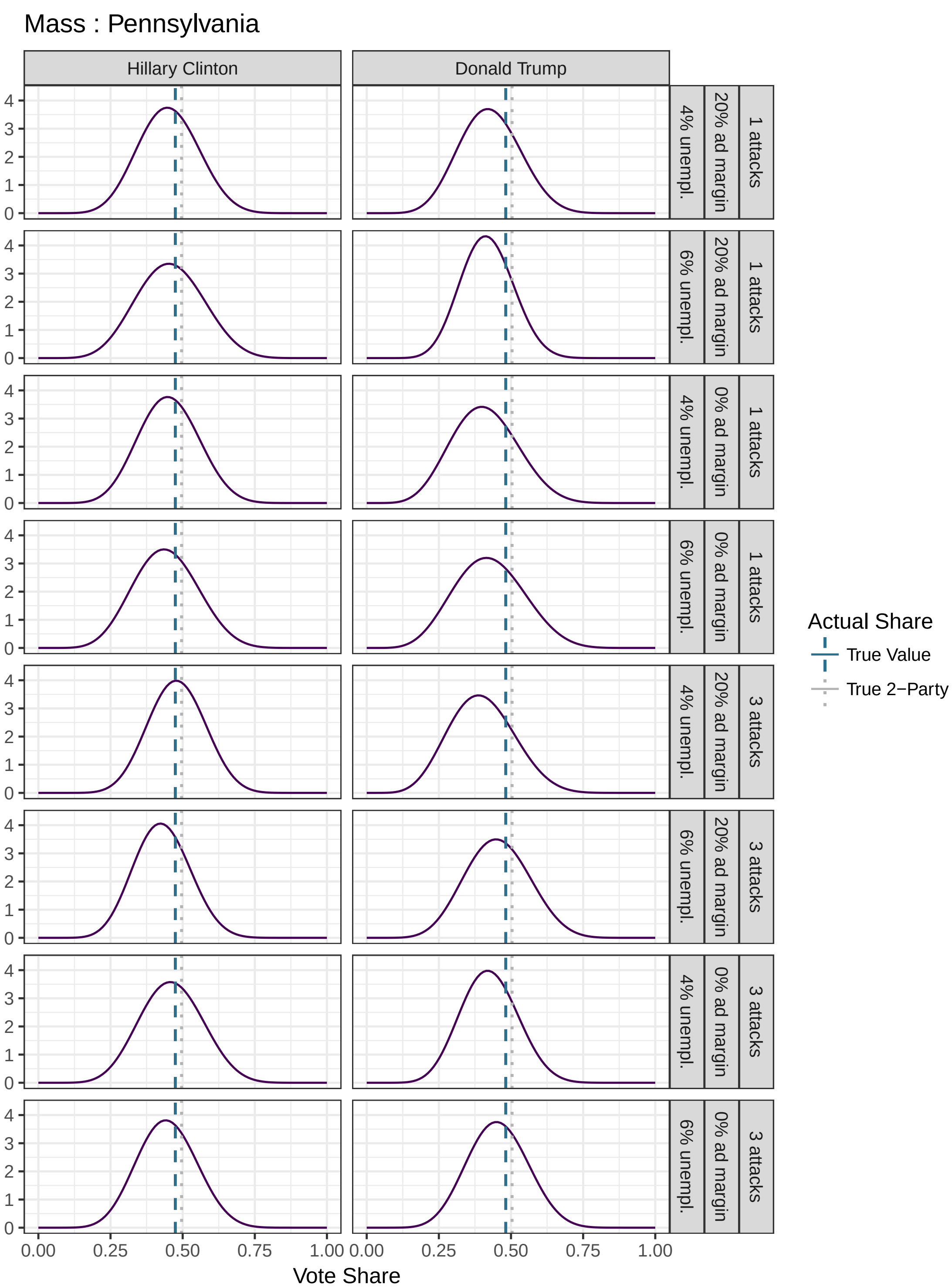


Figure 3.26: Overall: Mass with Covariates PA

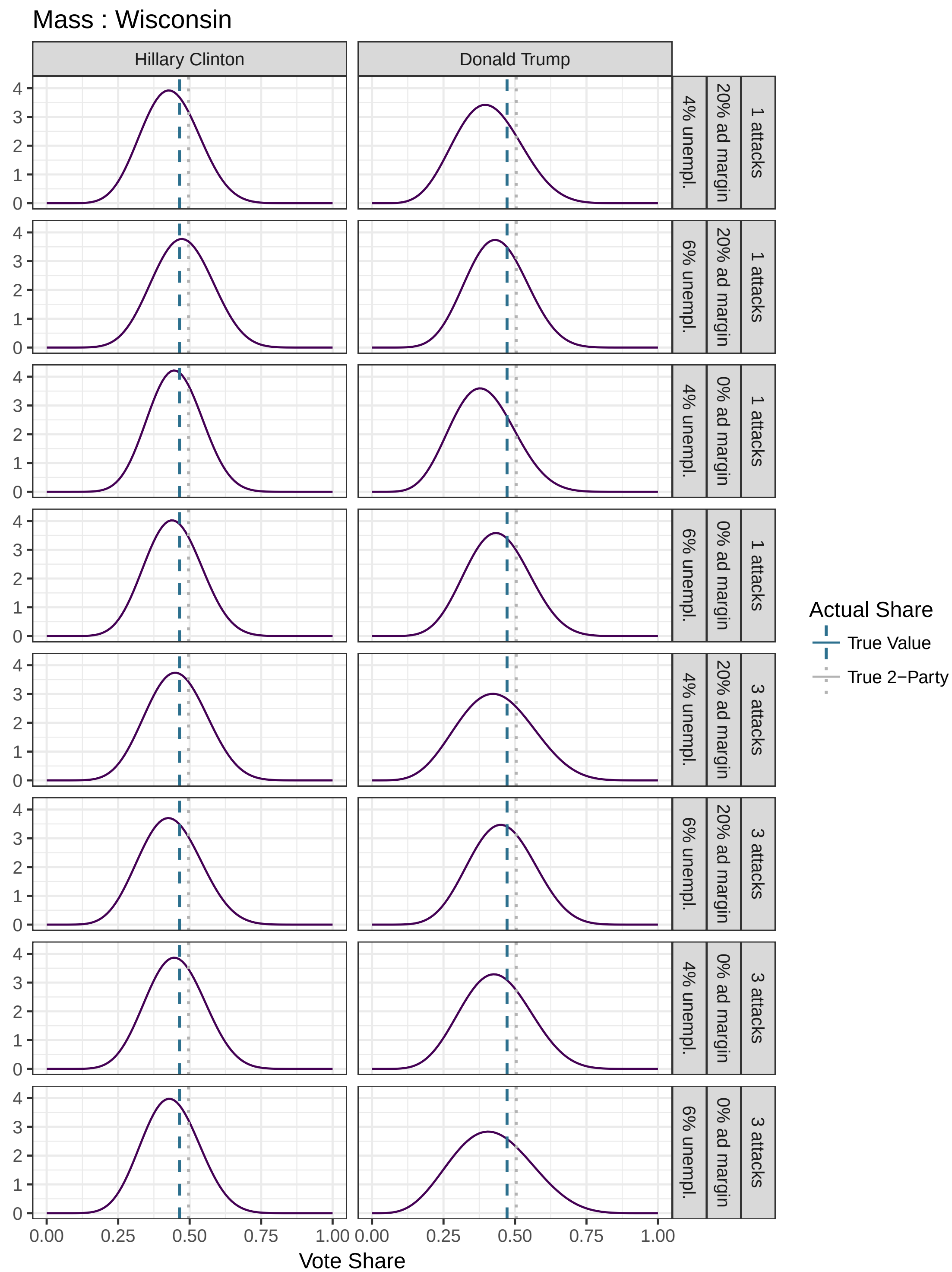


Figure 3.27: Overall: Mass with Covariates WI

3.5.4 Difference Plots: Without Covariates

Implicit in the notion that elicitation should target "experts" is the assumption that education or experience will confer particular area knowledge. This knowledge, in turn, should improve estimates in empirical analyses in ways that justify the costs of elicitation. Targeting distinct mass and elite samples for this project engages with this assumption, but does not directly demonstrate the mechanism that either education or pertinent knowledge translate to better elicited estimates; after all, PhD students differ from the general population on a variety of dimensions, including age, race, and party identification. The following difference plots compare estimates across the mass and elite samples broken down according to their demographic characteristics, both with and without covariates. While some categories do not allow for comparison (e.g., there are no members of the elite sample who are over age 55 or who answered 0 political knowledge questions correctly, and there are no members of the mass sample with a Ph.D.), these figures allow for direct comparison along dimensions such as education and knowledge that are often thought to correlate with expertise. Throughout these plots by state, mass and elite samples have extremely similar estimates within demographic categories.

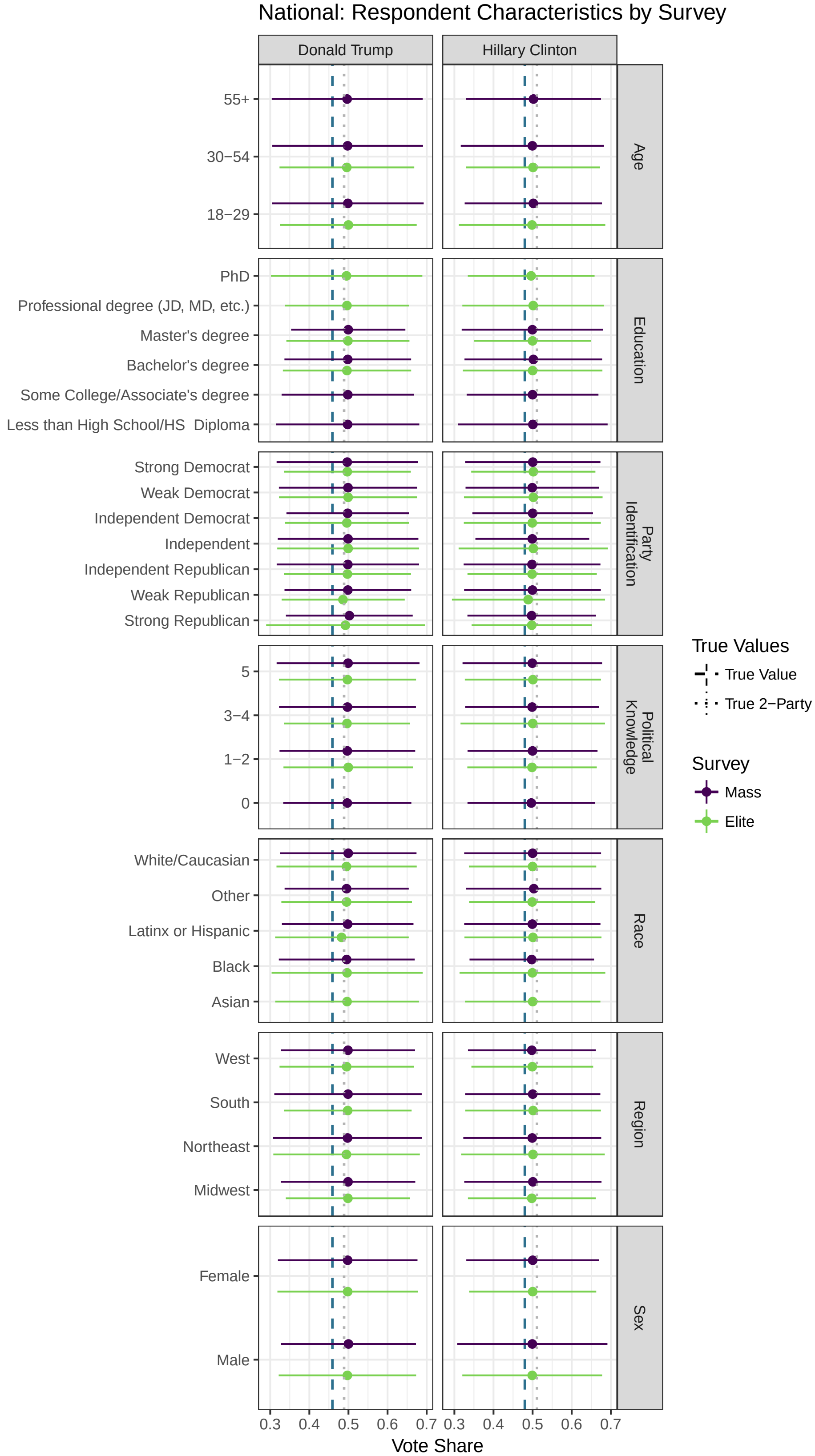


Figure 3.28: Differences without covariates: National

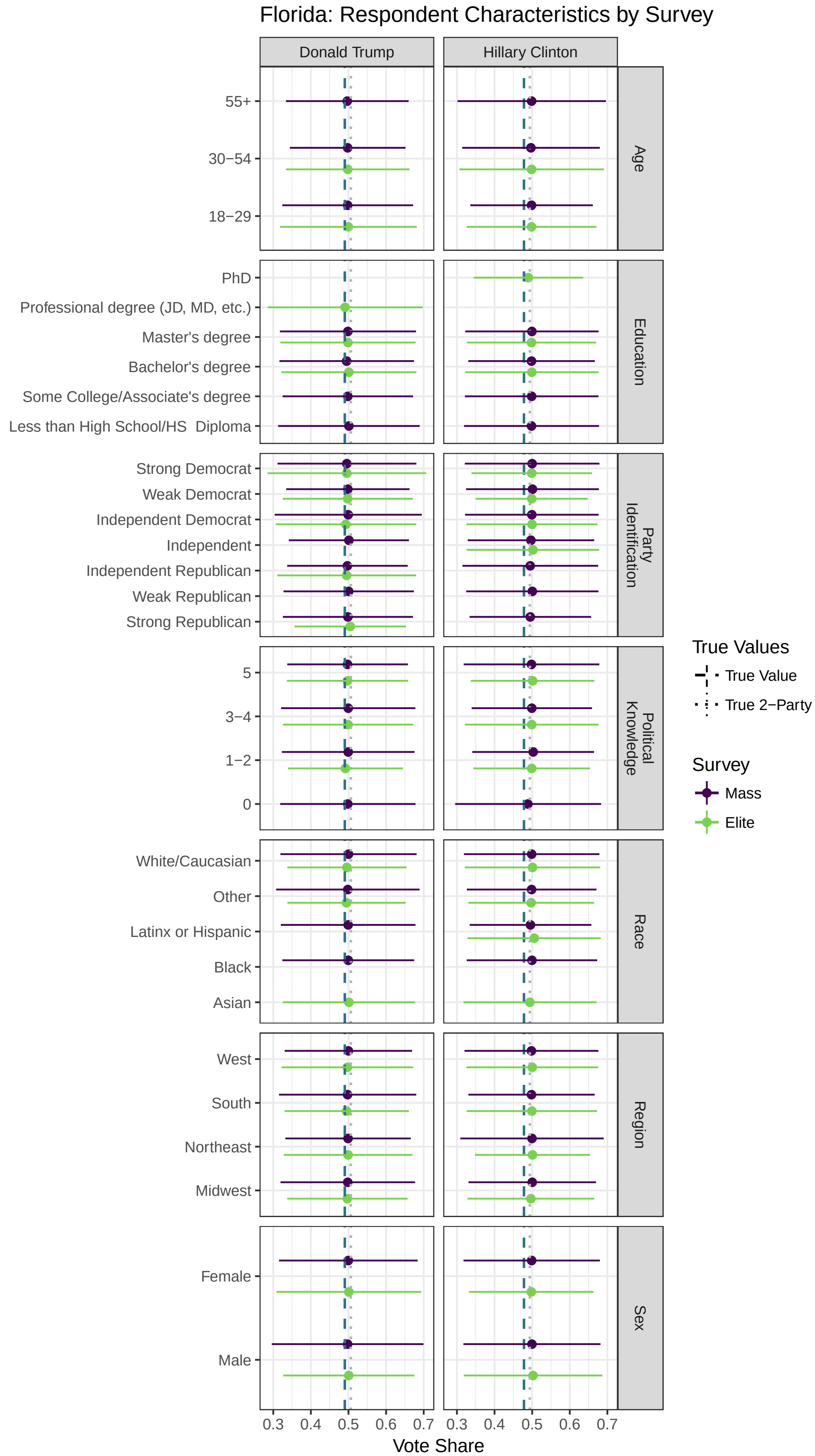


Figure 3.29: Differences without covariates: Florida

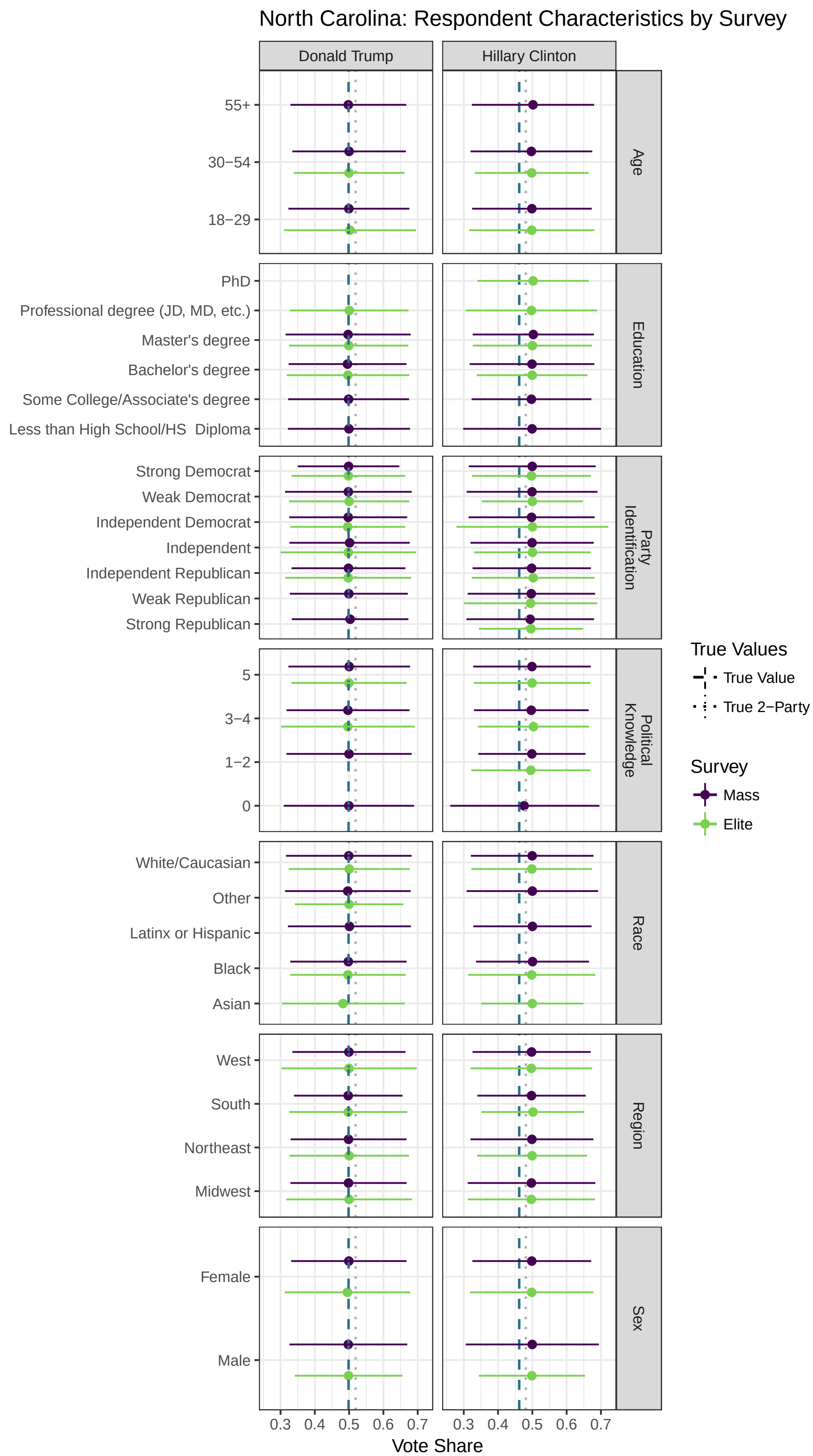


Figure 3.30: Differences without covariates: North Carolina

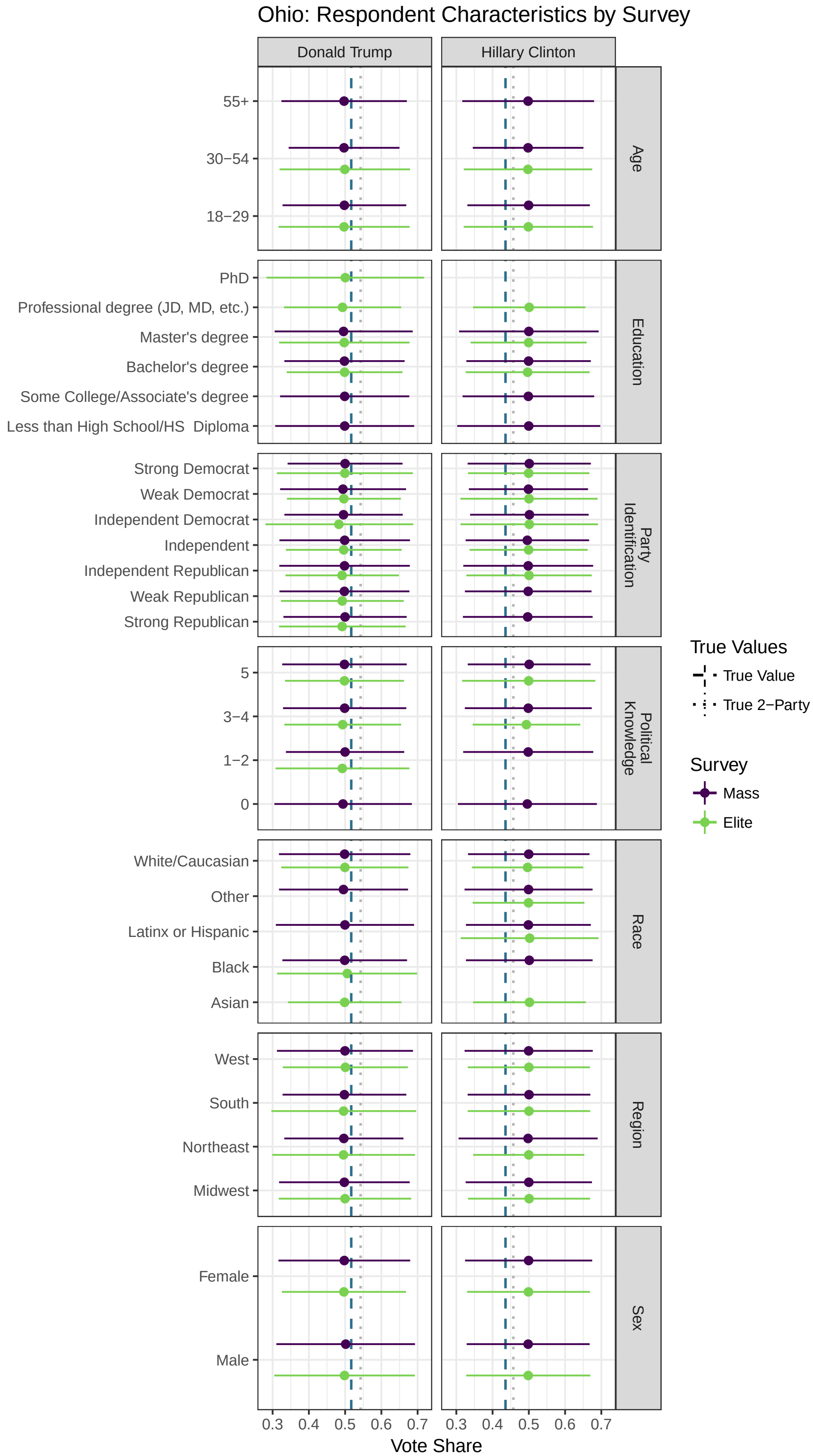


Figure 3.31: Differences without covariates: Ohio

3.5.5 Difference Plots: With Covariates

As with the pooled results for mass and elite samples, the performance of each group differs more when including covariates. Elite subsets have much less variation across samples in their estimates, although means are not always nearer to true outcome values. Dispersion in the full prior distributions of elite and mass subgroups are visualized in figures in the Appendix. Within the difference plots in this section, however, the aim is to evaluate whether particular demographic subsets provide more accurate predictions, and under which covariate conditions they do so. In general, these results can provide a preliminary indication of what kinds of individuals might be more likely to offer accurate assessments through elicitation. In particular, if experts with more educational credentials or a higher score on political knowledge questions provide mean estimates nearer to the true values, these results would serve as evidence that "expertise" does indeed reside with "experts" as conventionally defined.

As before, missing estimates across these plots suggest that no respondents of that demographic characteristic received a particular question (e.g., no individuals in the mass sample evaluating Clinton's performance in Florida also, following the survey, answered 0 political knowledge questions correctly). First and foremost, the results according to education and, to a lesser extent, political knowledge, demonstrate some preliminary evidence that expertise can or does reside with "experts" as traditionally defined. That said, it is important to note that no direct comparison between those in the political science PhD programs and those in the general population who may have a PhD either in political science or another field is possible. Current PhD students may benefit from being directly and contemporaneously engaged in educational endeavors related to politics in ways that those with just a general higher level of education may not. The set of experts in the elite sample identifying as having a PhD level of education is also small, lending to greater precision in their estimates.

Likewise, while elite responses are often more precise, and often more accurate, they are not always closer to true values even when they are more certain of their answers. That is, higher levels of education or political knowledge need not imply the absence of bias. Subsetting the covariate results by party identification demonstrates how strong party affiliation can and often does correlate with providing a biased assessment of each candidate's potential vote share. Biases appear in the expected directions, with "Strong Republican" respondents in either sample overestimating Trump's vote share and underestimating Clinton's, while "Strong Democrat" respondents tend to overestimate Clinton's vote share (although not always to the same extent, particularly in the Florida results) while underestimating Trump's.

These results offer evidence in support of elicitation along several dimensions. First, when not using covariate combinations in elicitation questions, mass and elite samples perform relatively similarly and answers are usually centered near or around true values. This indicates that elicitation may not need to be so costly as is often assumed, depending on research aims. If the goal of eliciting priors is to provide bounds for extreme values or address concerns of quasi-perfect separation (as will be further discussed in the next chapter), for example, elicitation from either a small convenience sample of identified "experts" or a relatively small representative sample can perform similarly, and elicitation perhaps need not require extremely time-intensive elicitation frameworks. Where more precision is desirable, a non-random "elite" sample can still be relatively modest in size and perform admirably. In circumstances where identifying individuals or sources with the correct "expertise" is challenging or impossible, however, a representative population sample could be used instead. Conversely, in particularly challenging contexts such as authoritarian or developing country settings where representative sampling is difficult or impossible, or for very specific research questions, a small, non-random set of "experts" may be sufficient. This observation has the immediate implication that restricting the definition of "expert" to those with more education or professional expertise may not be preferred; on the other hand, a modestly sized representative sample may be more costly to contact, and these results suggest that a smaller, non-random sample may perform just as well. By the same token, however, these results are conditional on the elicitation framework used. That is, with an elicitation tool other than roulette, or with a survey or focus group structure other than the online model used in this study, differences resulting from education or underlying knowledge could be more evident.

Second, as the bias in the responses delineated by party identification demonstrate, even assessments from modestly sized samples of respondents, and even where those samples, as in the elite case, are relatively homogenous on other dimensions (e.g., race and education), can have significant variation. Visualizing vote share estimates by party identification appears to reify a concern of scholars skeptical of "subjective" research designs involving elicitation, that eliciting priors even from so-called "experts" has the potential to introduce significant bias. Attempting to identify and recruit "experts" without bias is likely an impossible goal, but the overall group responses of both the mass and the elite samples, whether with or without covariates, demonstrate that the Dirichlet-based clustering proposed in the previous chapter and applied here provides a solution. Clustering can help to incorporate divergent assessments, and the overall estimates provided by each sample provide consistently accurate results.

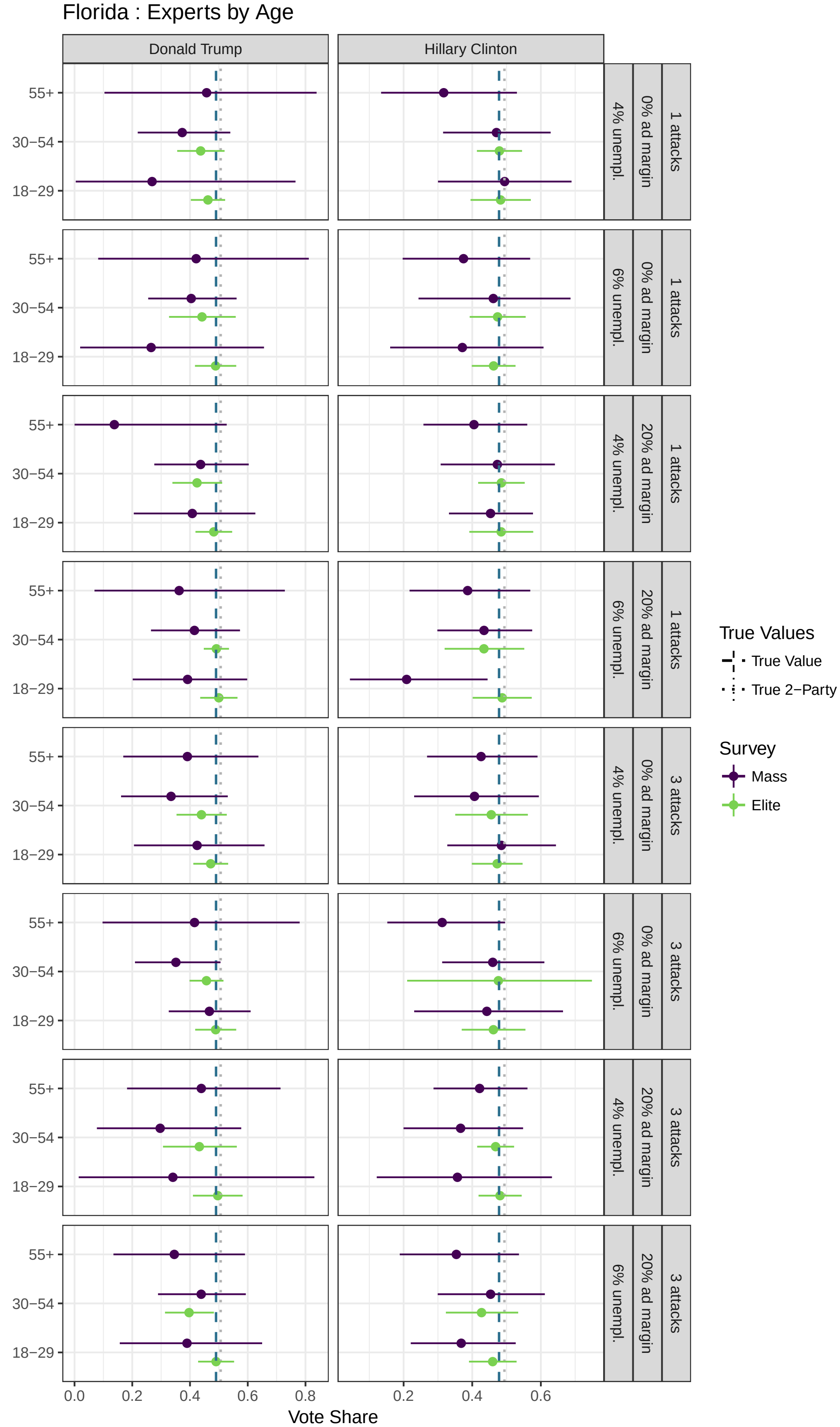


Figure 3.32: Differences with covariates: Florida Age

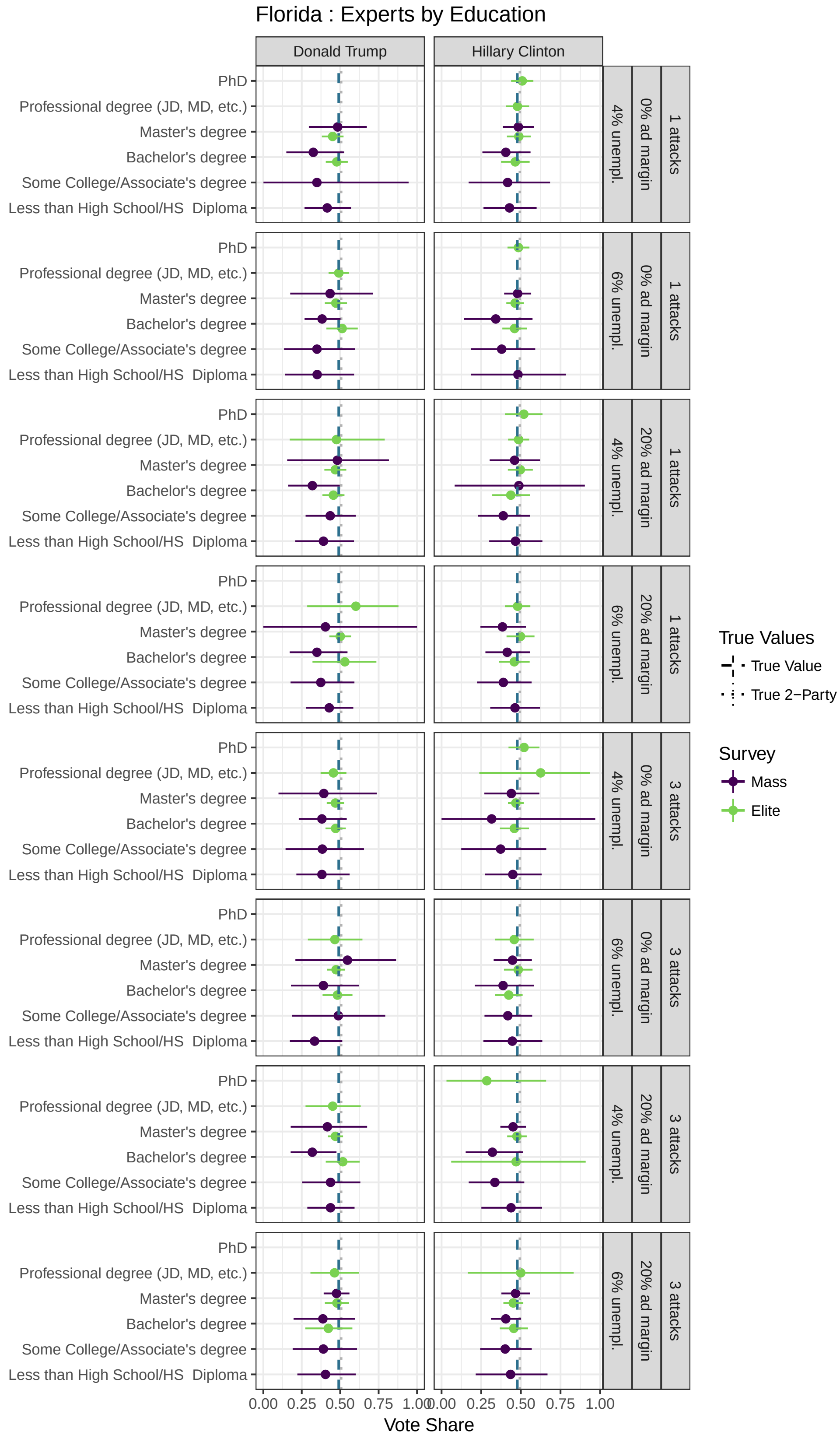


Figure 3.33: Differences with covariates: Florida Education

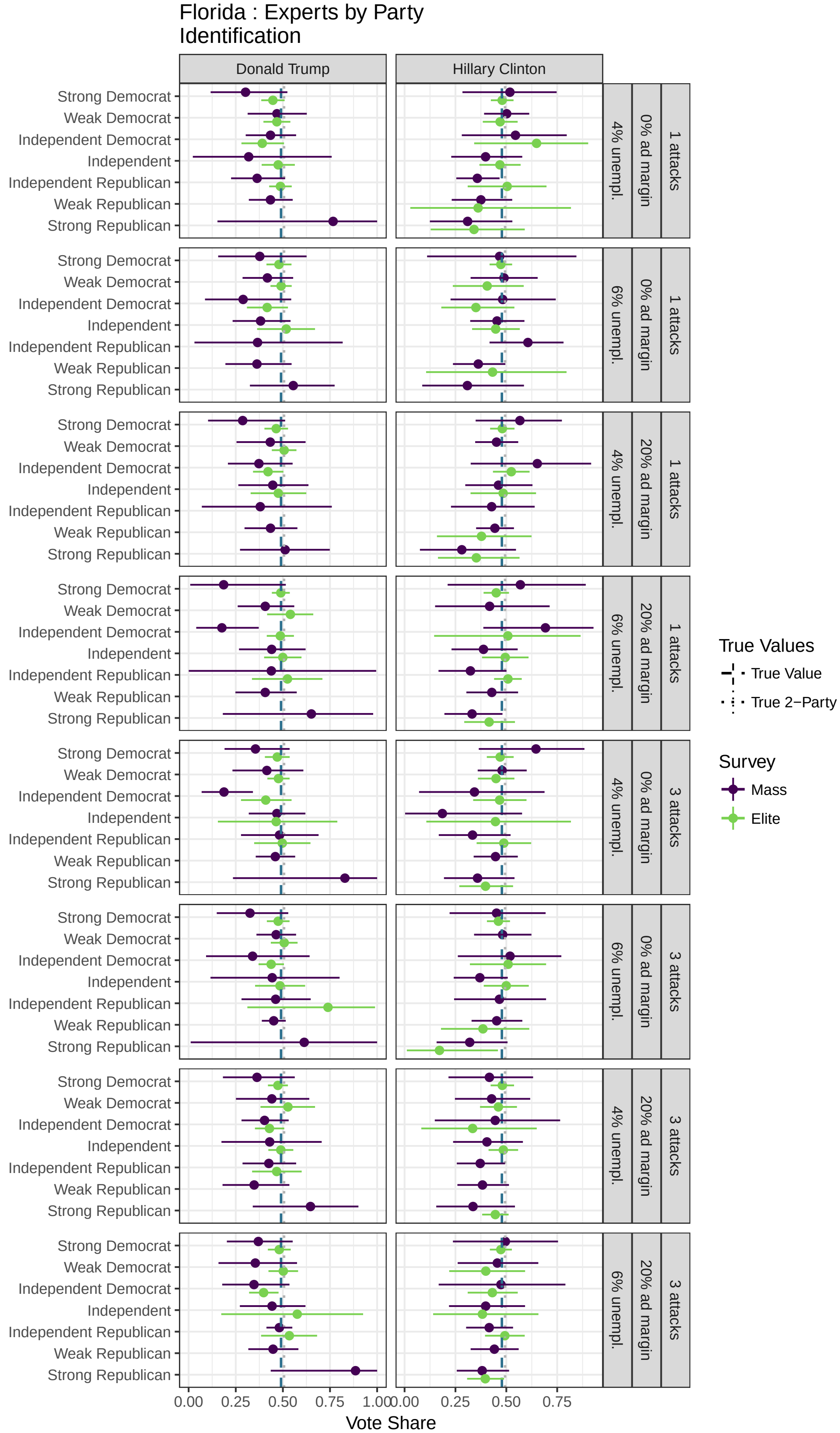


Figure 3.34: Differences with covariates: Florida Party Identification

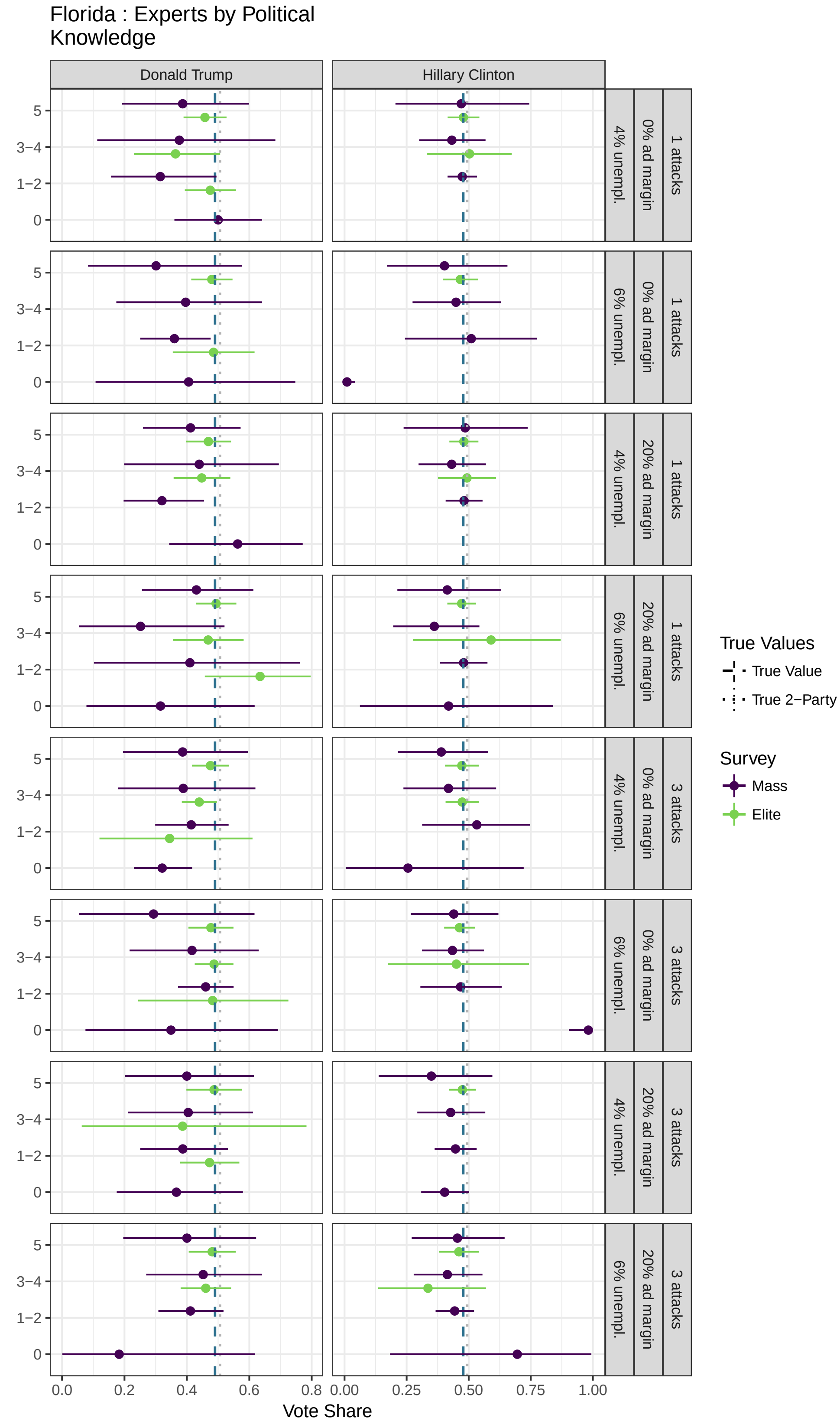


Figure 3.35: Differences with covariates: Florida Political Knowledge

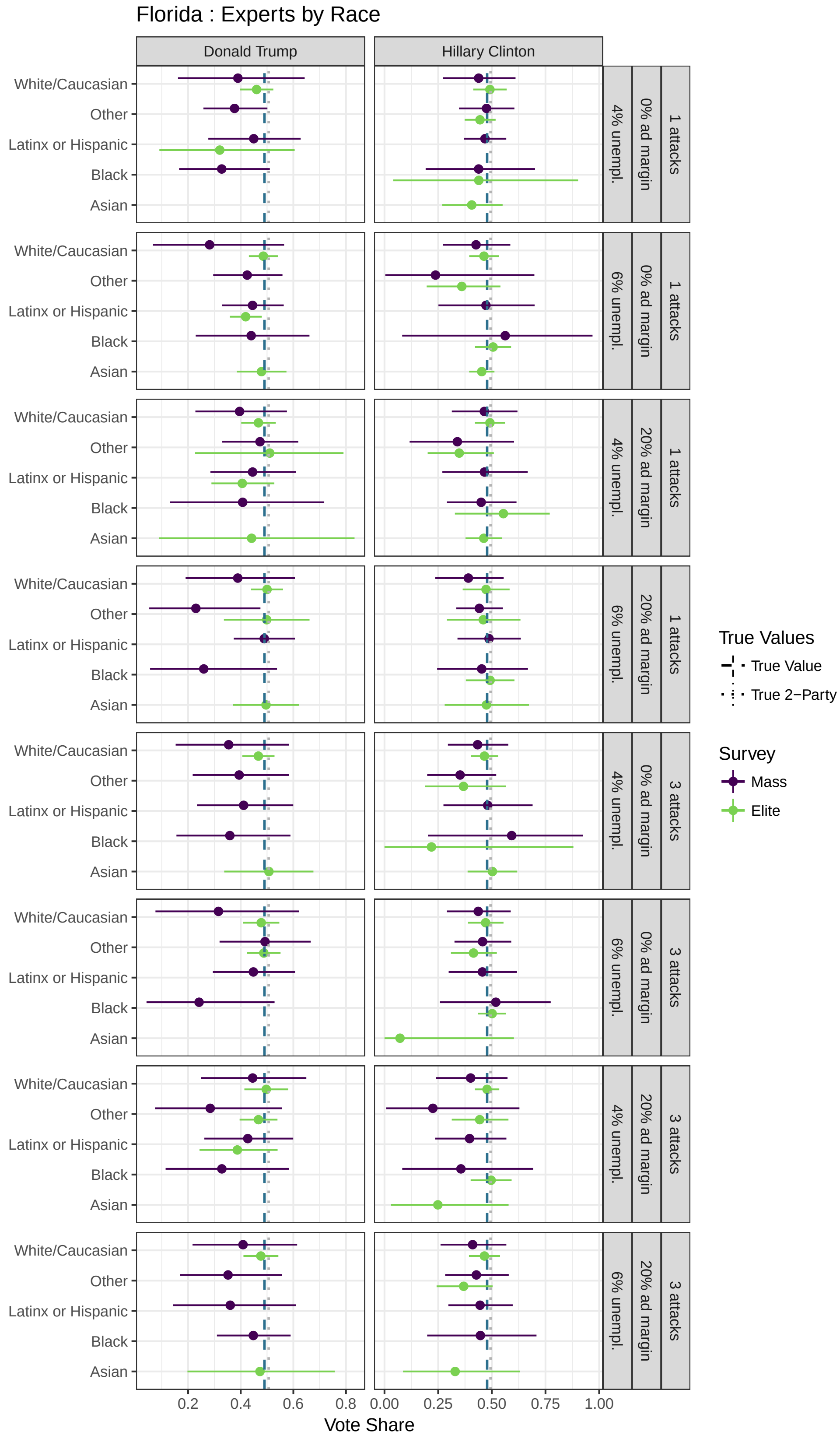


Figure 3.36: Differences with covariates: Florida Race

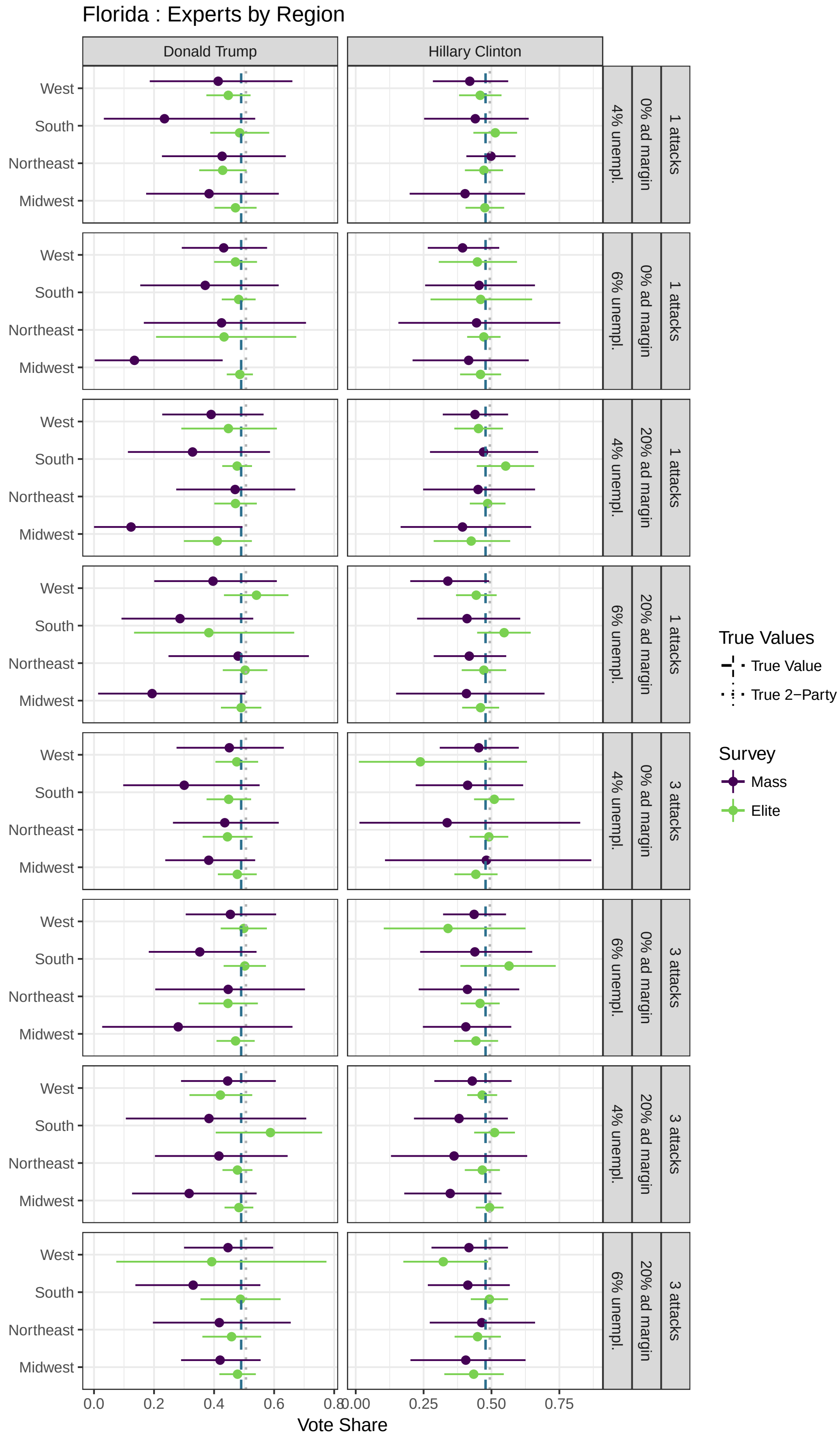


Figure 3.37: Differences with covariates: Florida Region

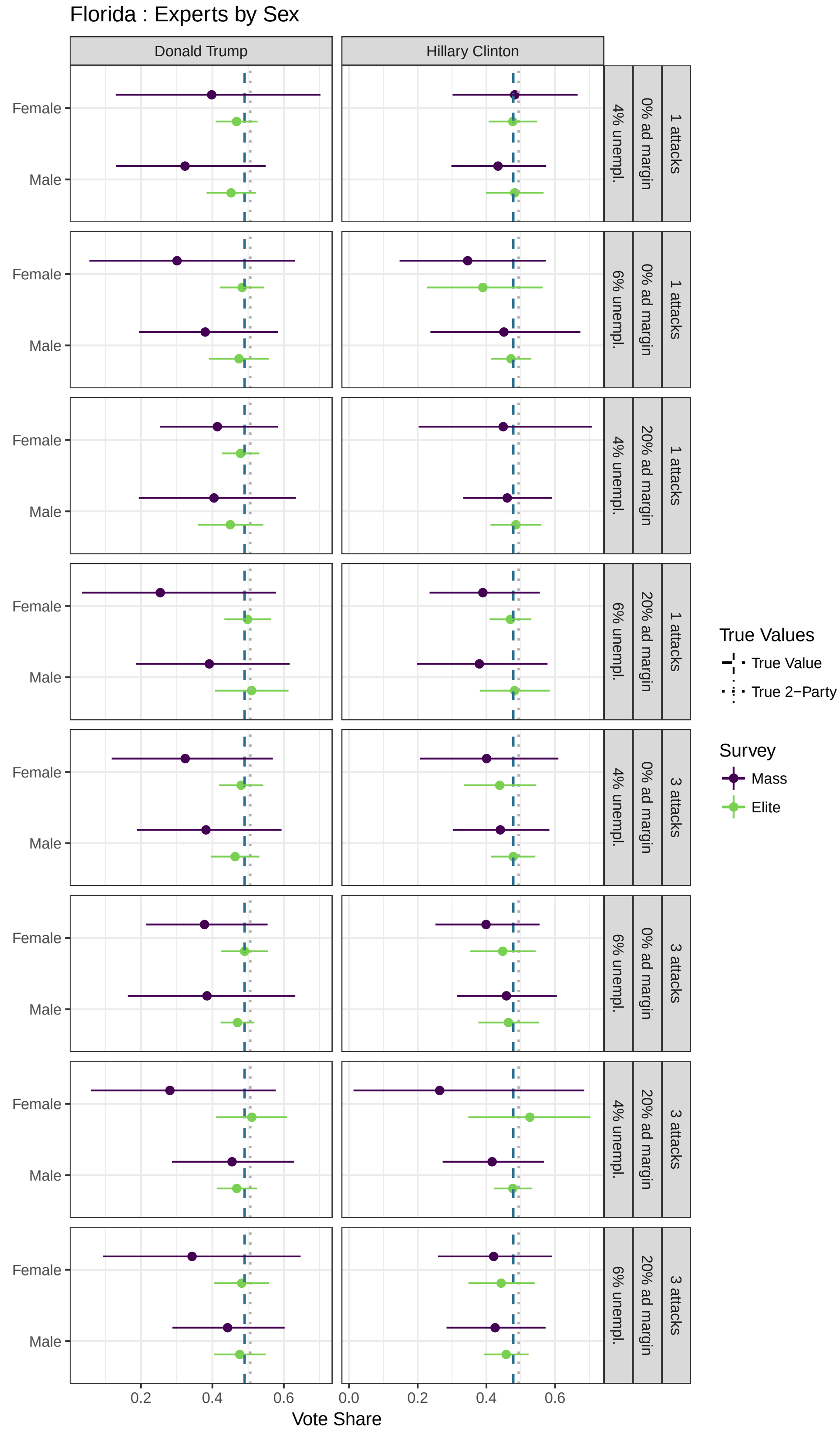


Figure 3.38: Differences with covariates: Florida Sex

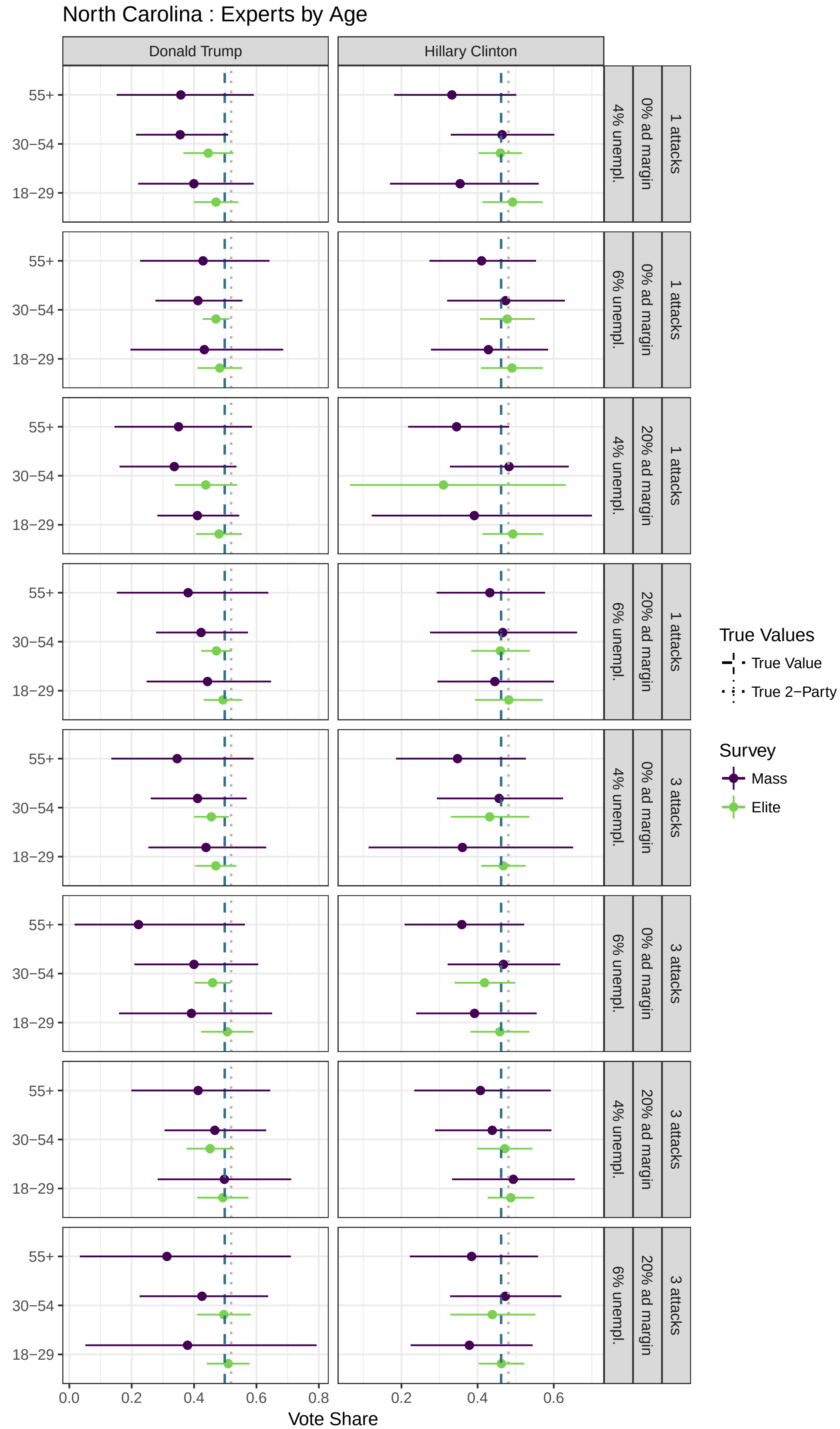


Figure 3.39: Differences with covariates: North Carolina Age

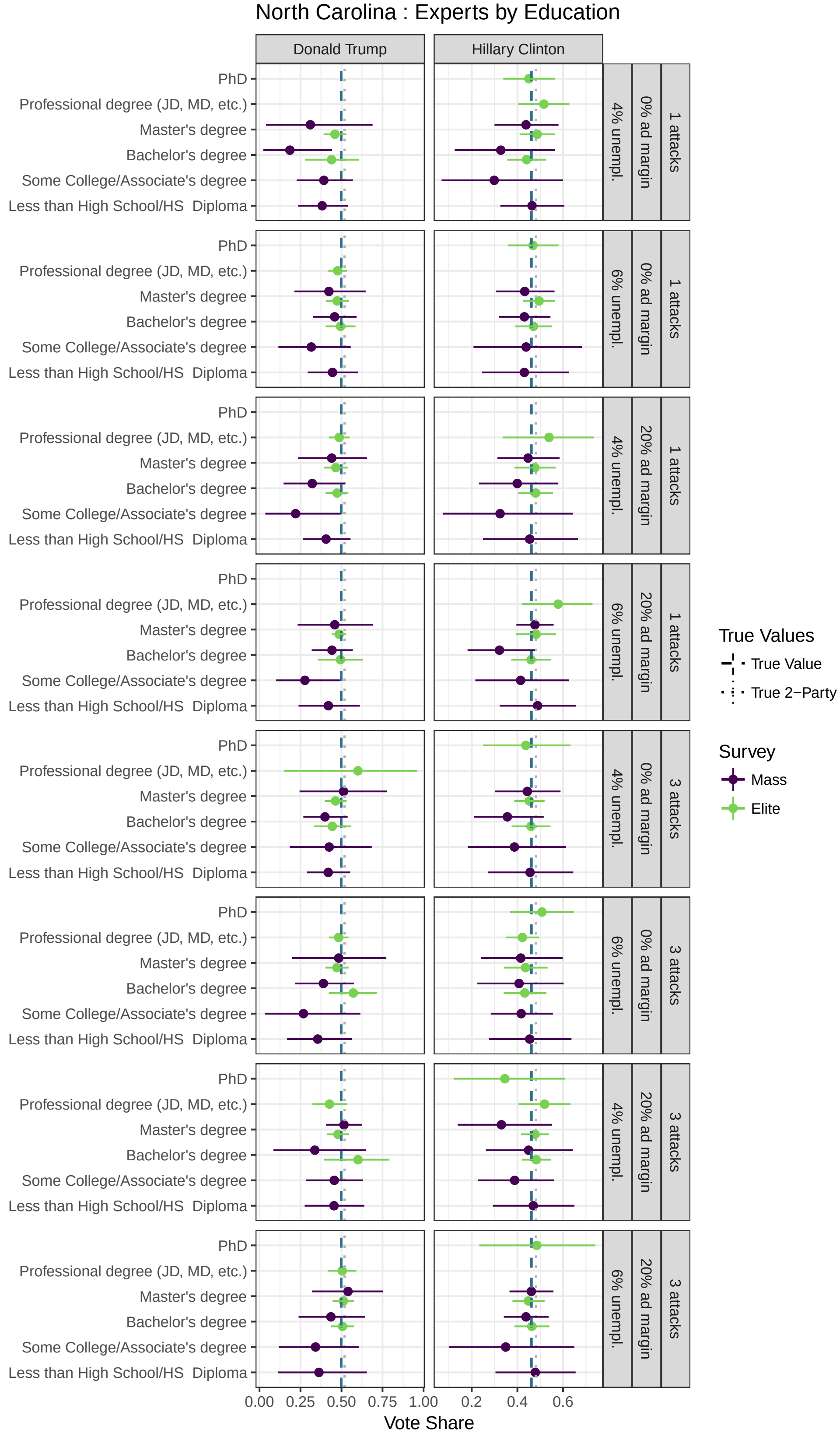


Figure 3.40: Differences with covariates: North Carolina Education

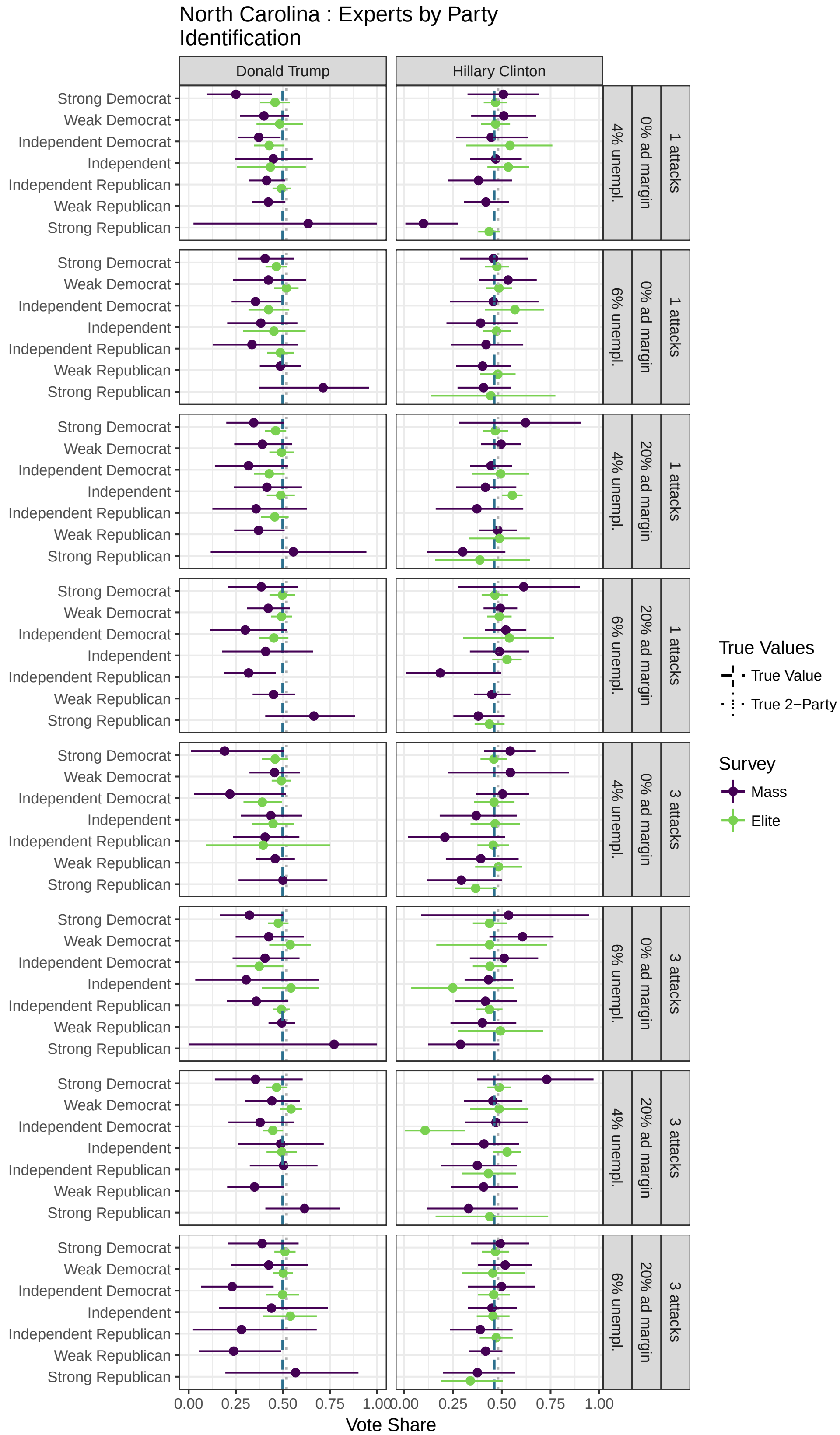


Figure 3.41: Differences with covariates: North Carolina Party Identification

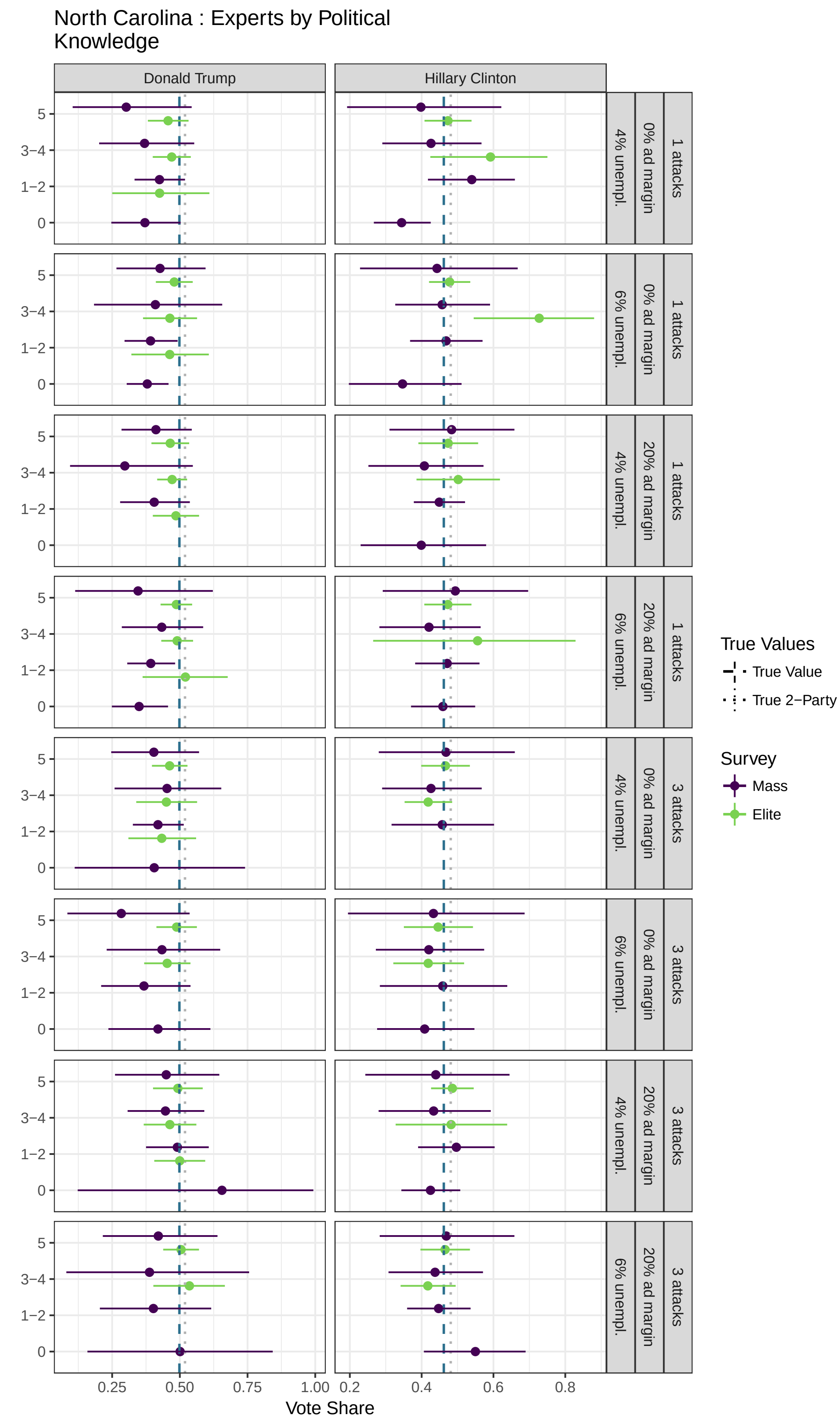


Figure 3.42: Differences with covariates: North Carolina Political Knowledge

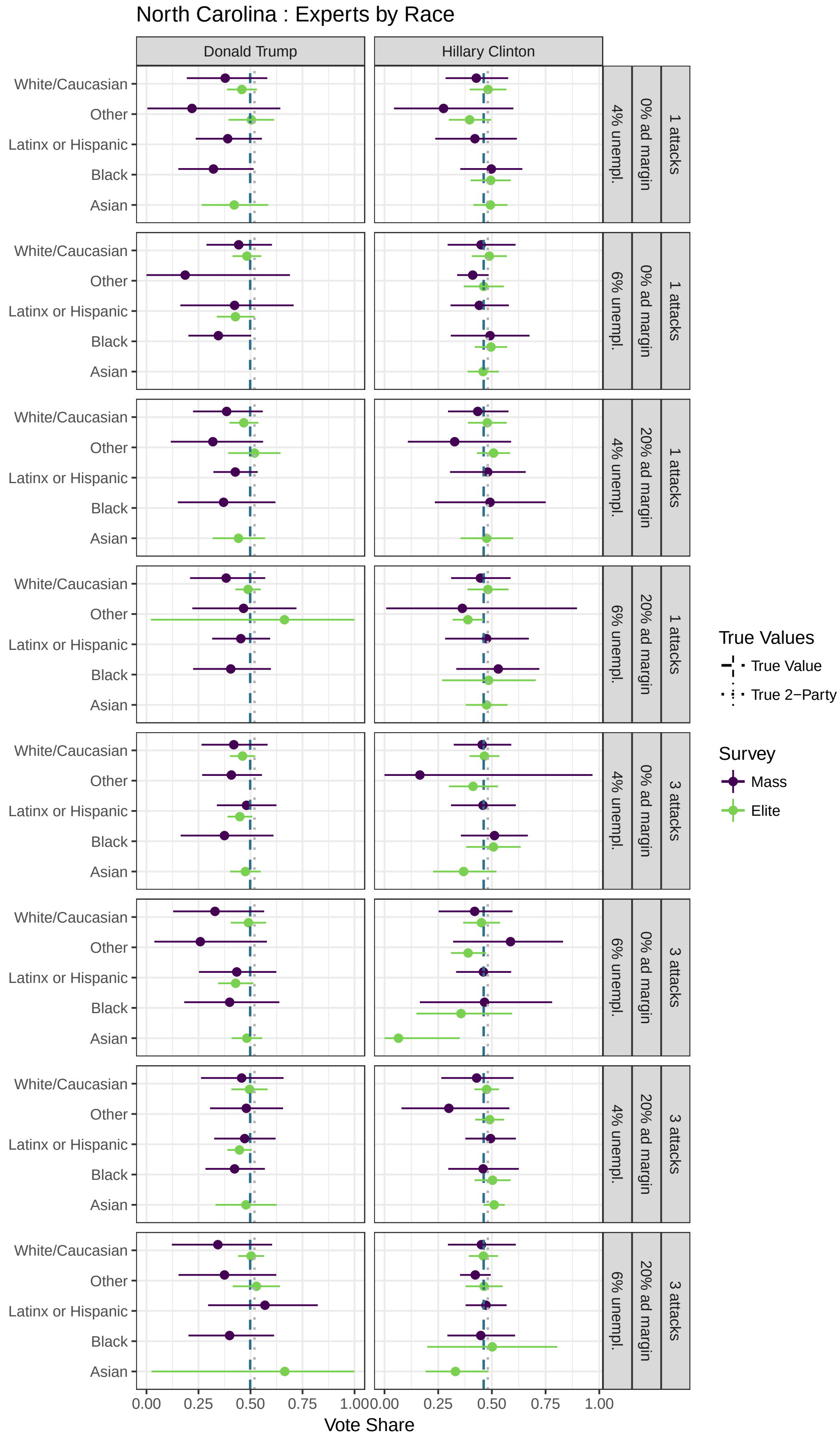


Figure 3.43: Differences with covariates: North Carolina Race

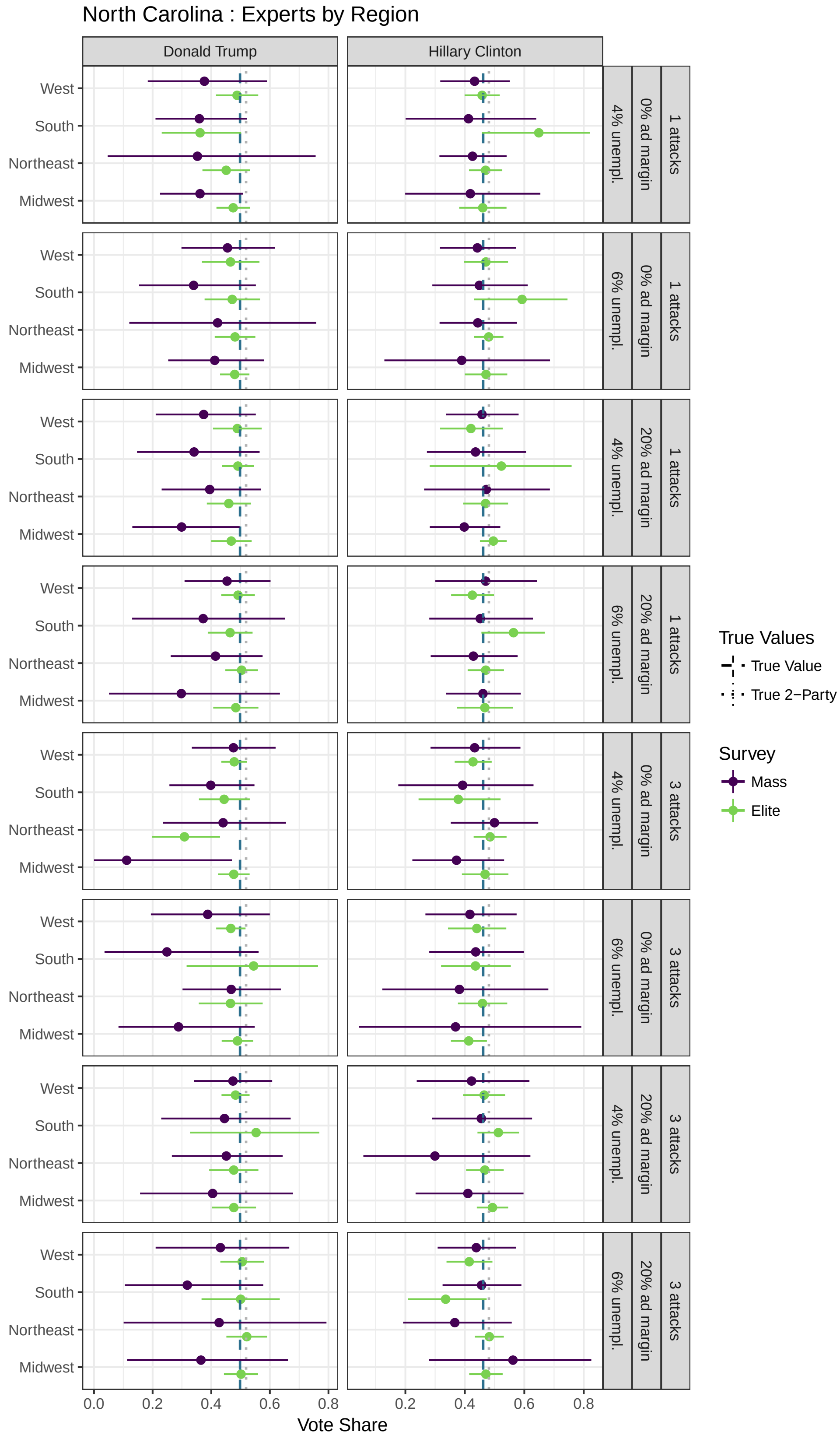


Figure 3.44: Differences with covariates: North Carolina Region

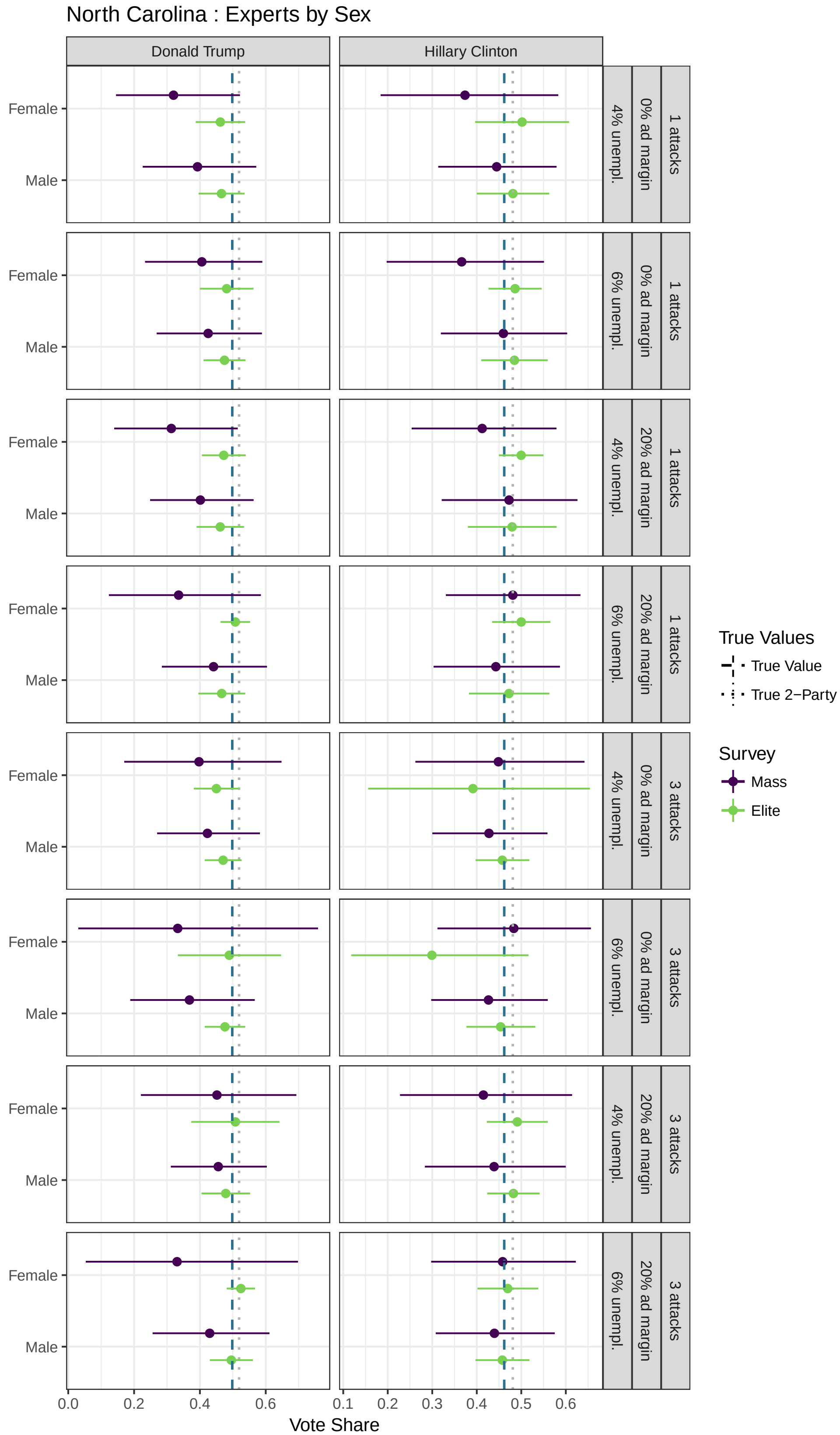


Figure 3.45: Differences with covariates: North Carolina Sex

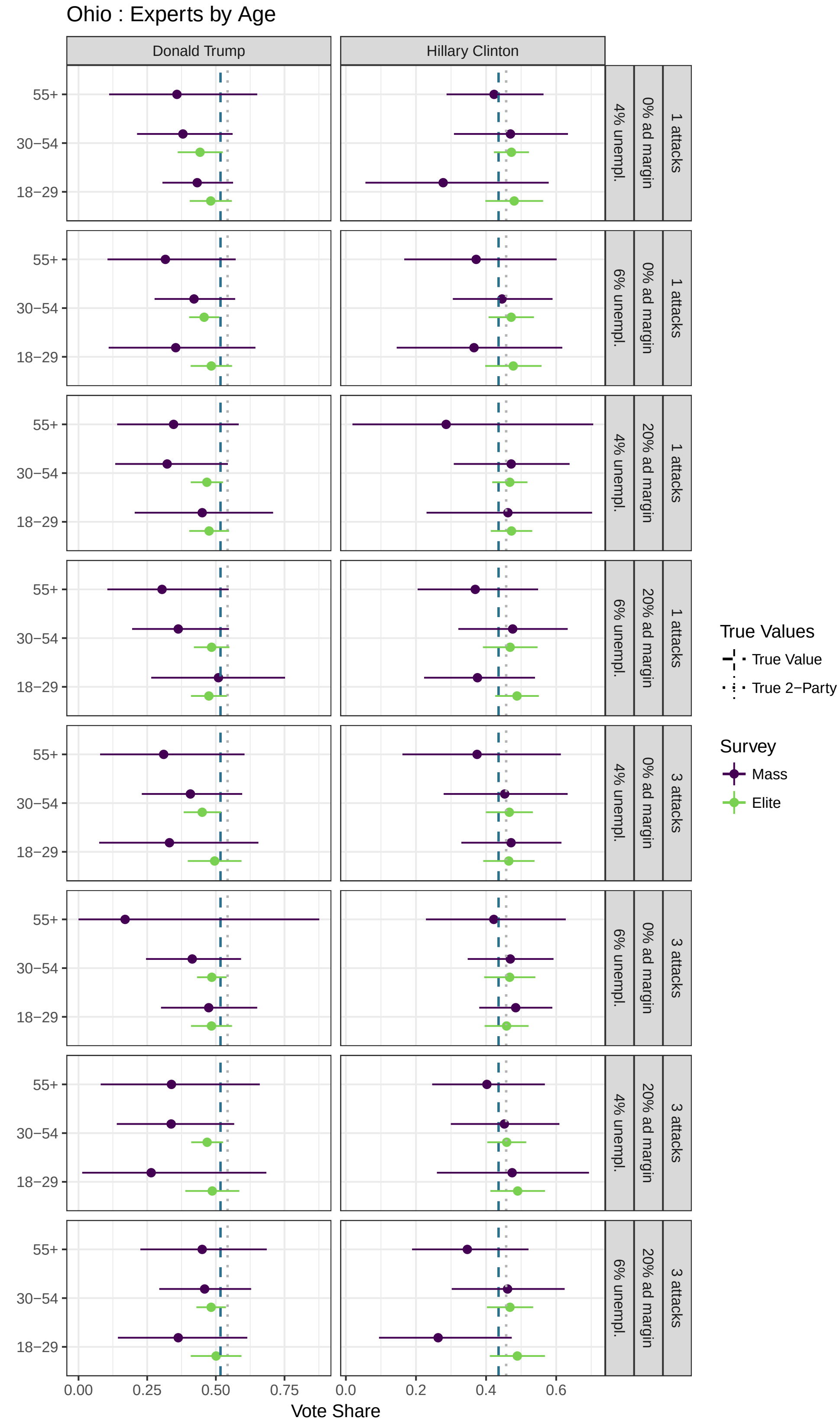


Figure 3.46: Differences with covariates: Ohio Age

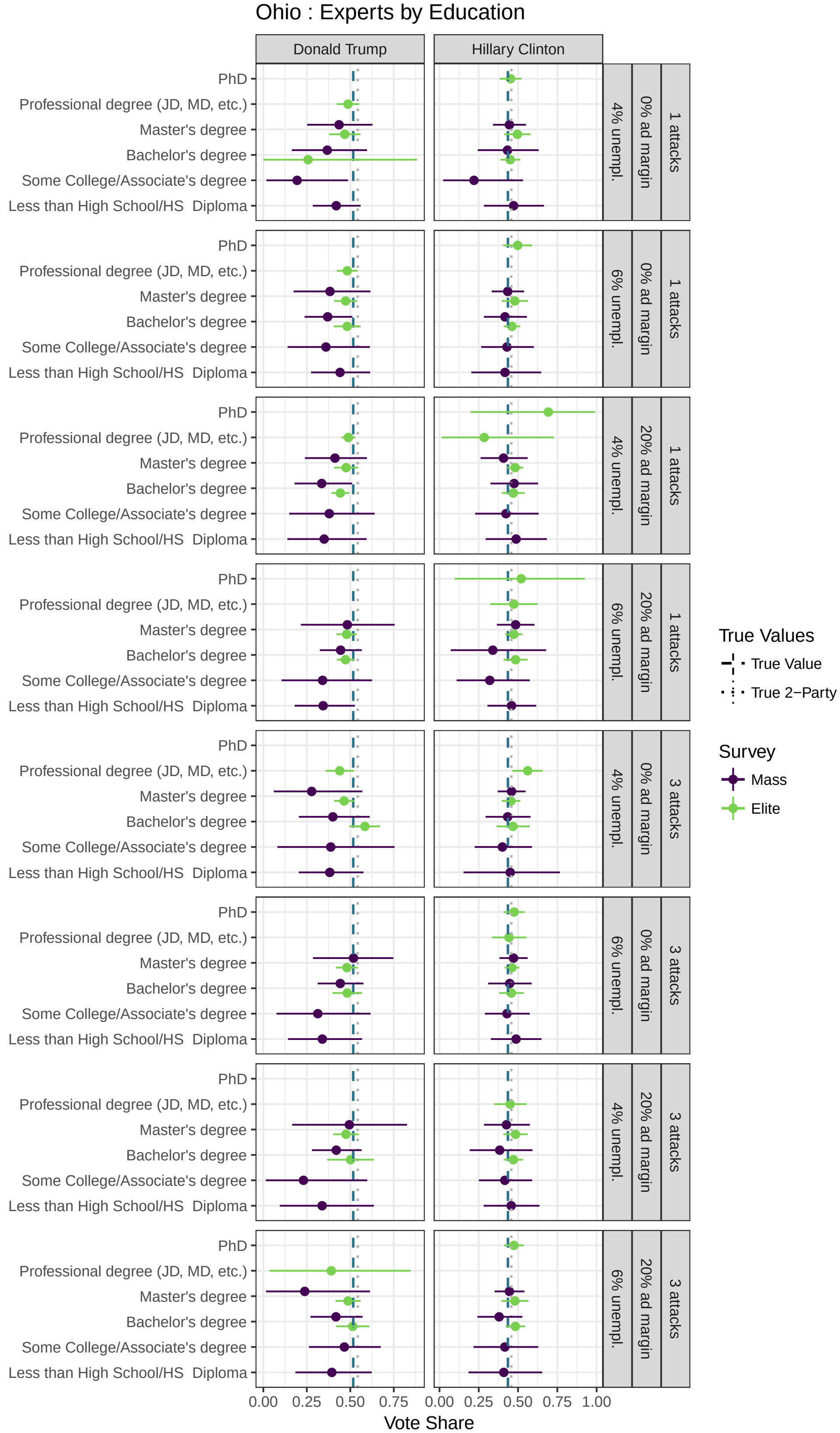


Figure 3.47: Differences with covariates: Ohio Education

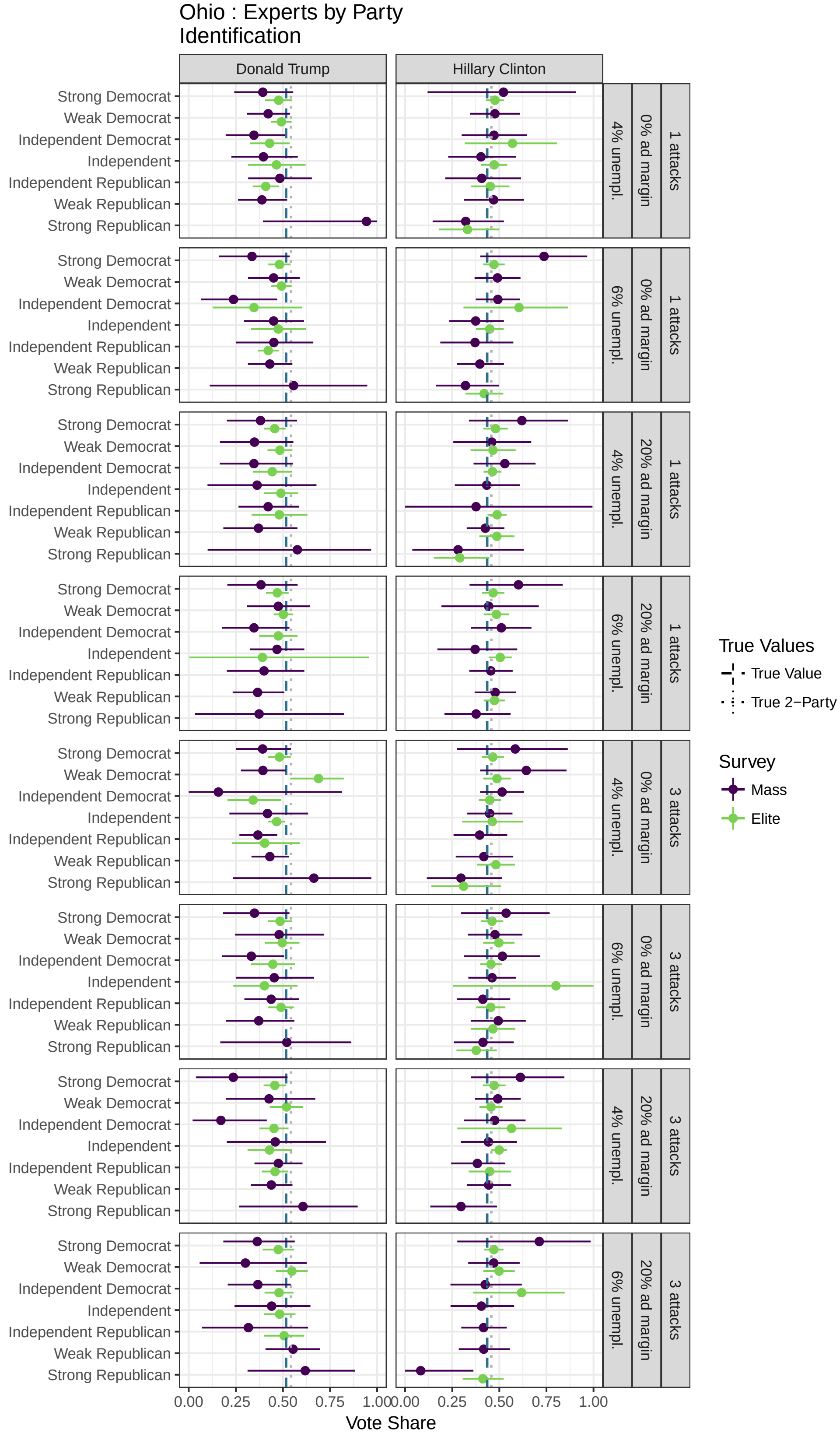


Figure 3.48: Differences with covariates: Ohio Party Identification

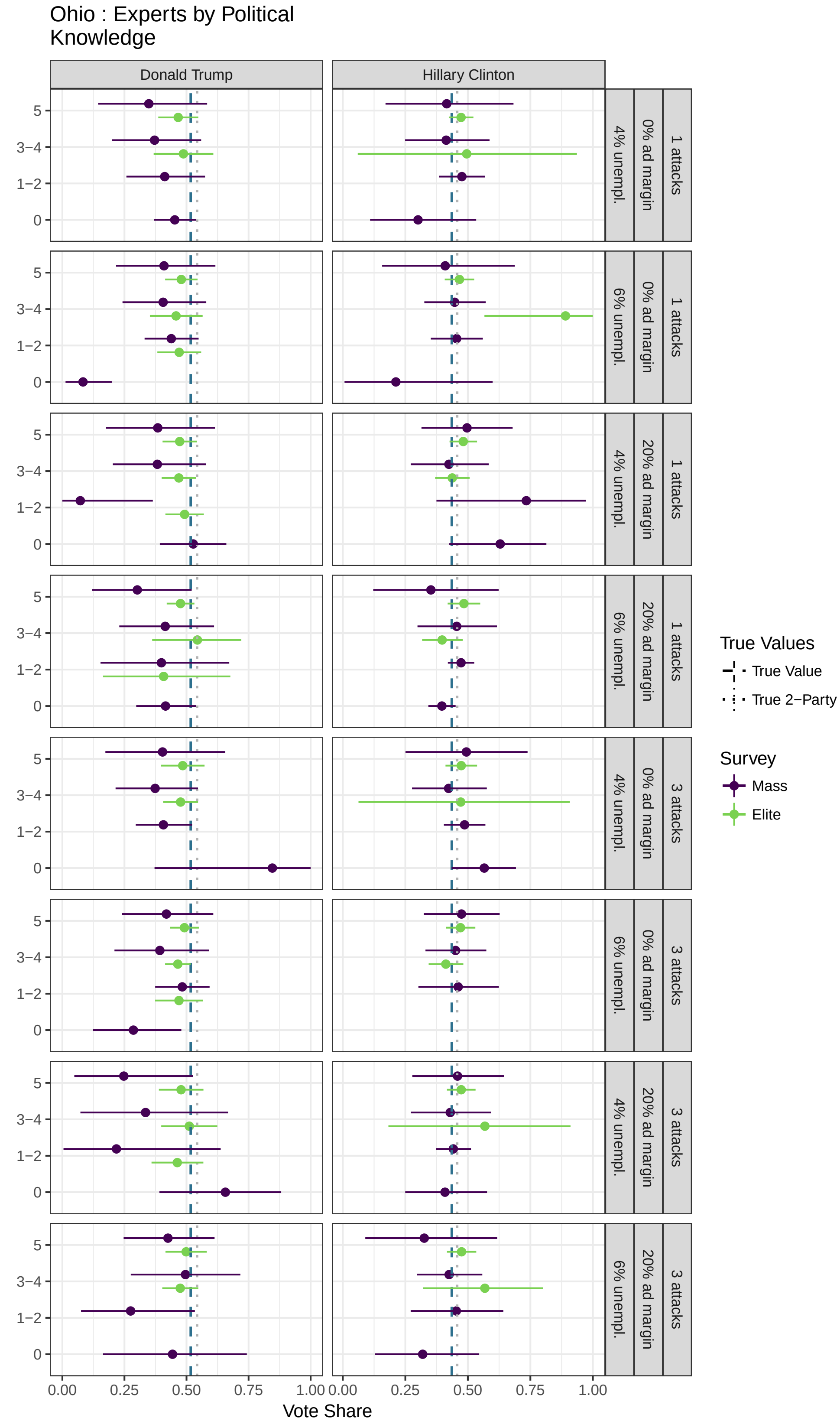


Figure 3.49: Differences with covariates: Ohio Political Knowledge

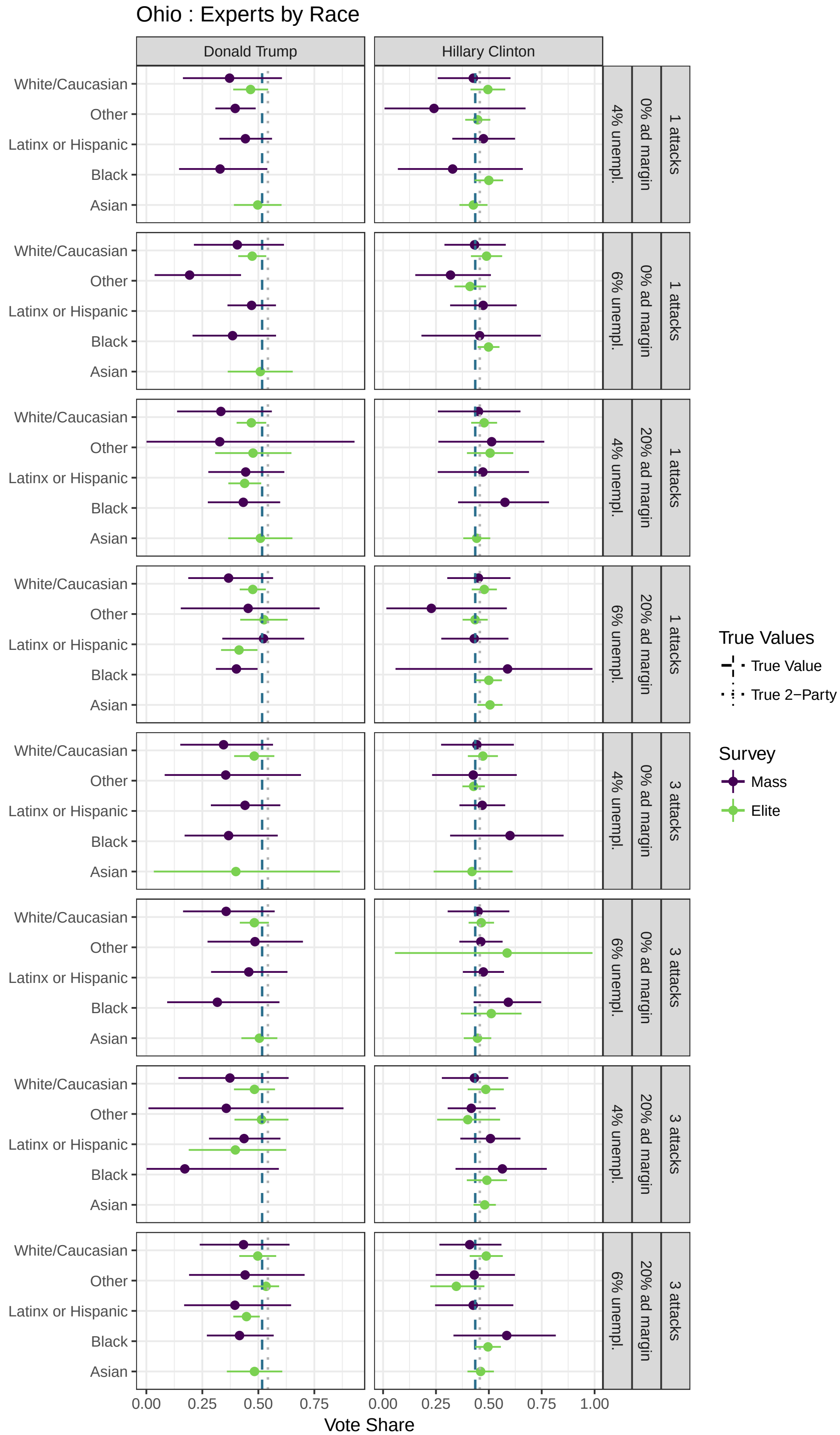


Figure 3.50: Differences with covariates: Ohio Race

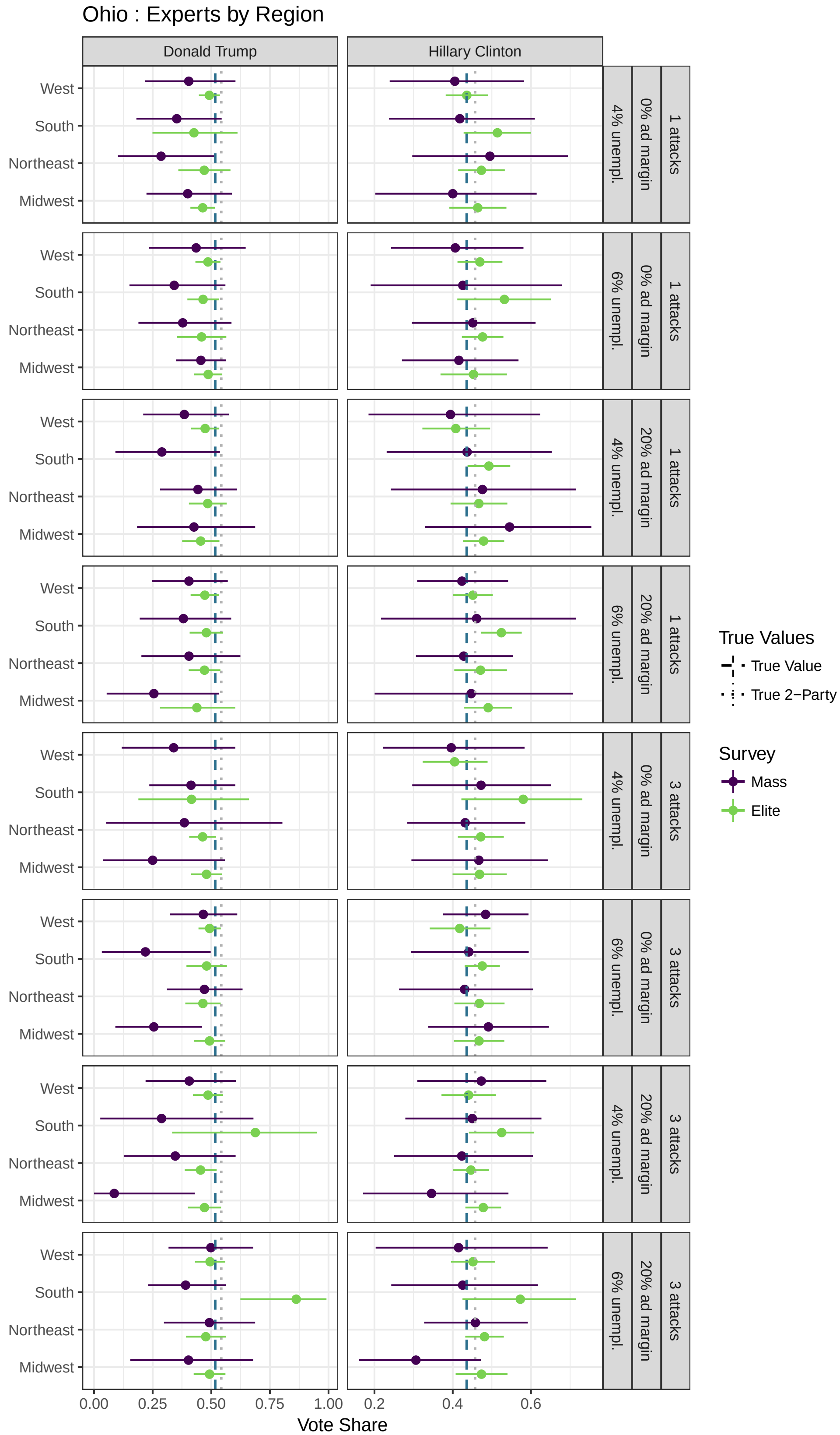


Figure 3.51: Differences with covariates: Ohio Region

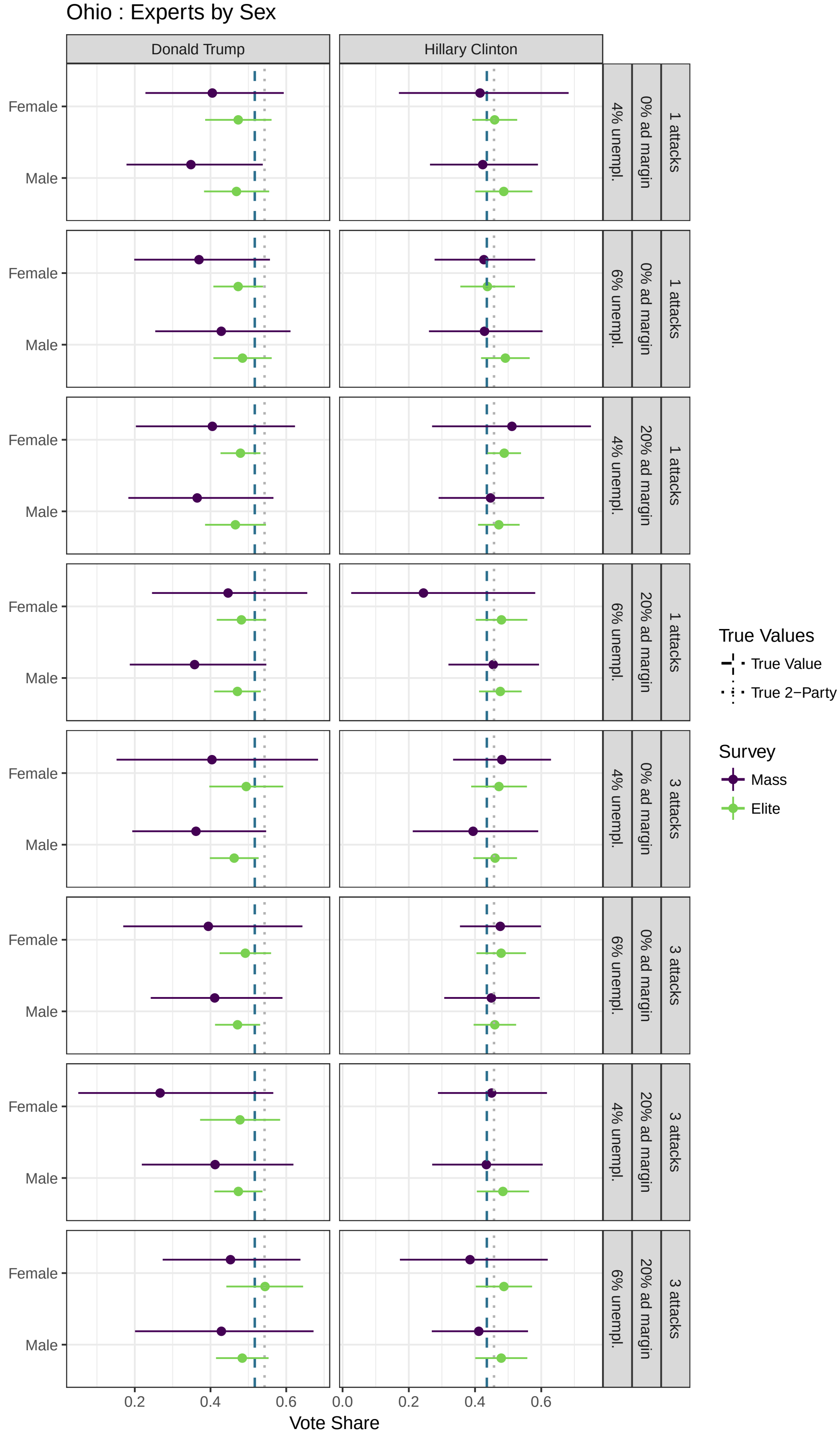


Figure 3.52: Differences with covariates: Ohio Sex

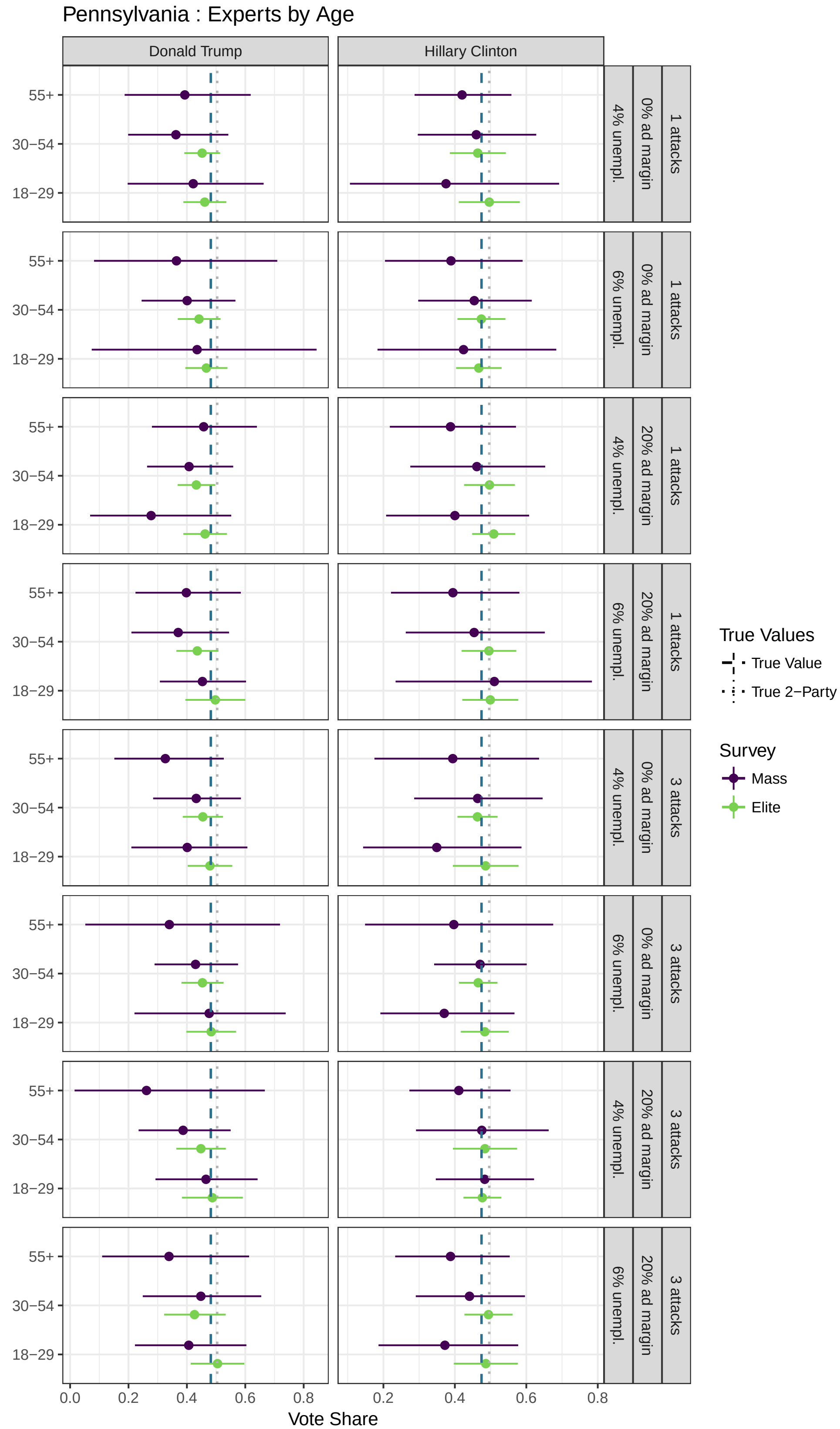


Figure 3.53: Differences with covariates: Pennsylvania Age

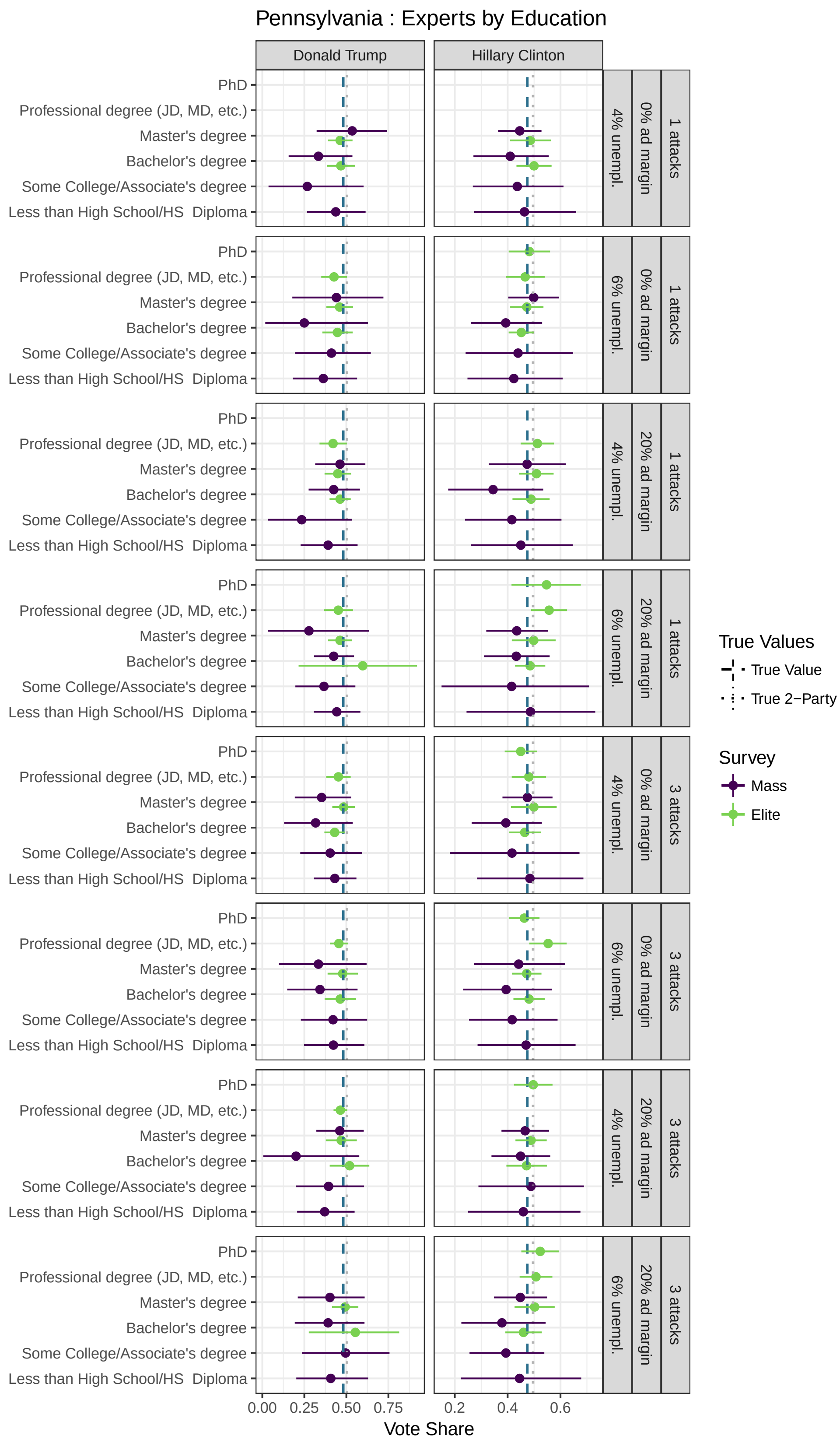


Figure 3.54: Differences with covariates: Pennsylvania Education

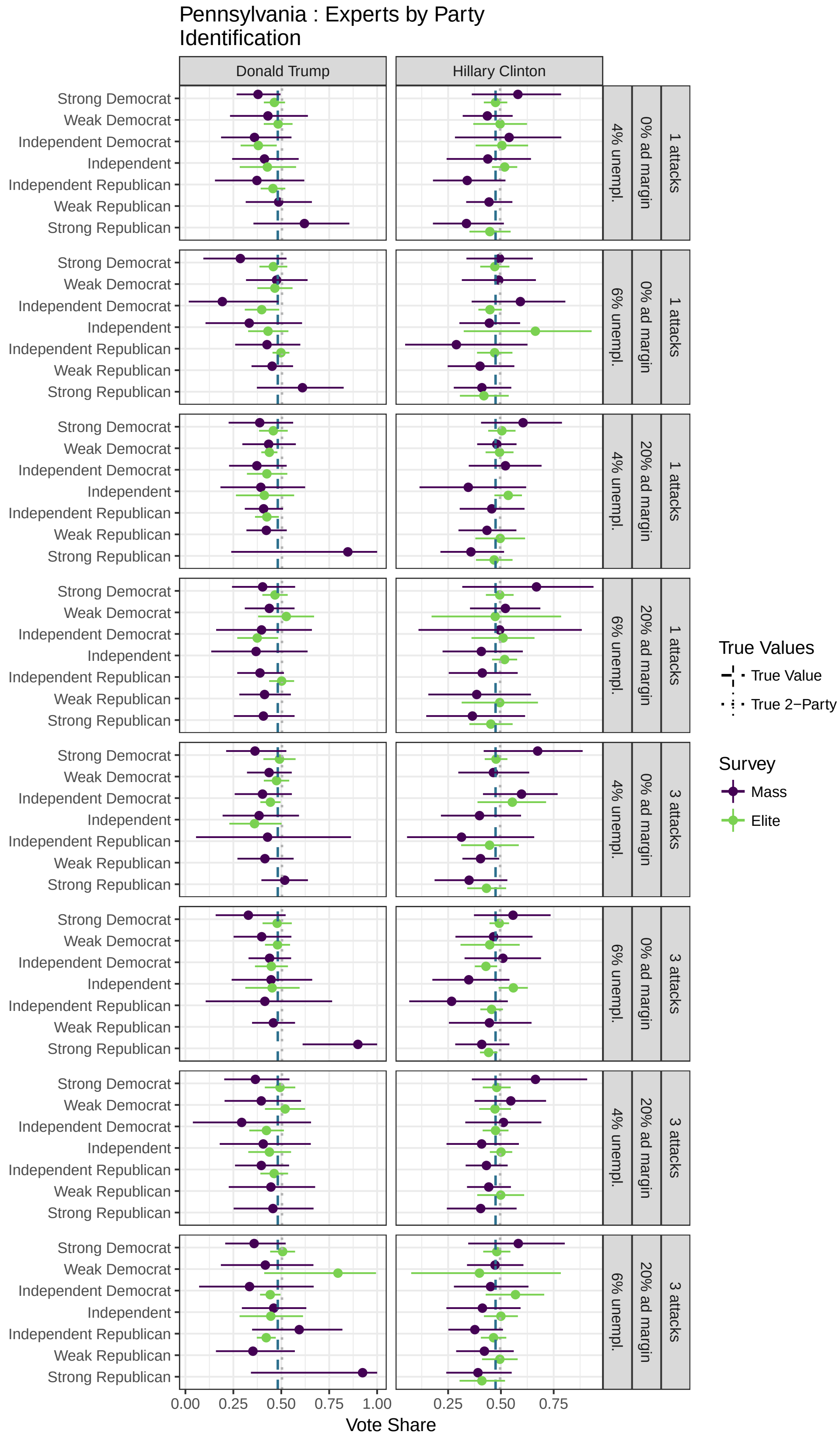


Figure 3.55: Differences with covariates: Pennsylvania Party Identification

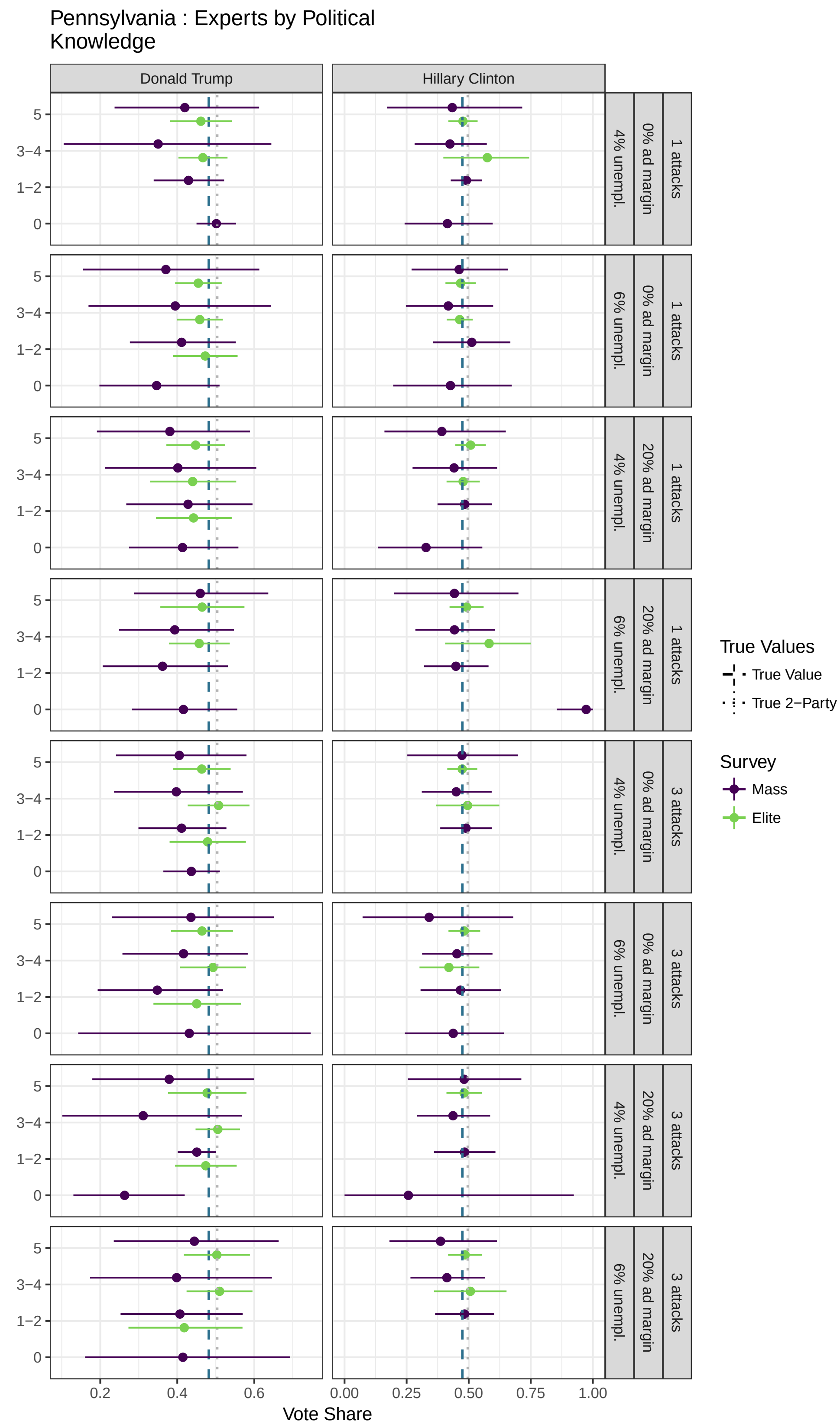


Figure 3.56: Differences with covariates: Pennsylvania Political Knowledge

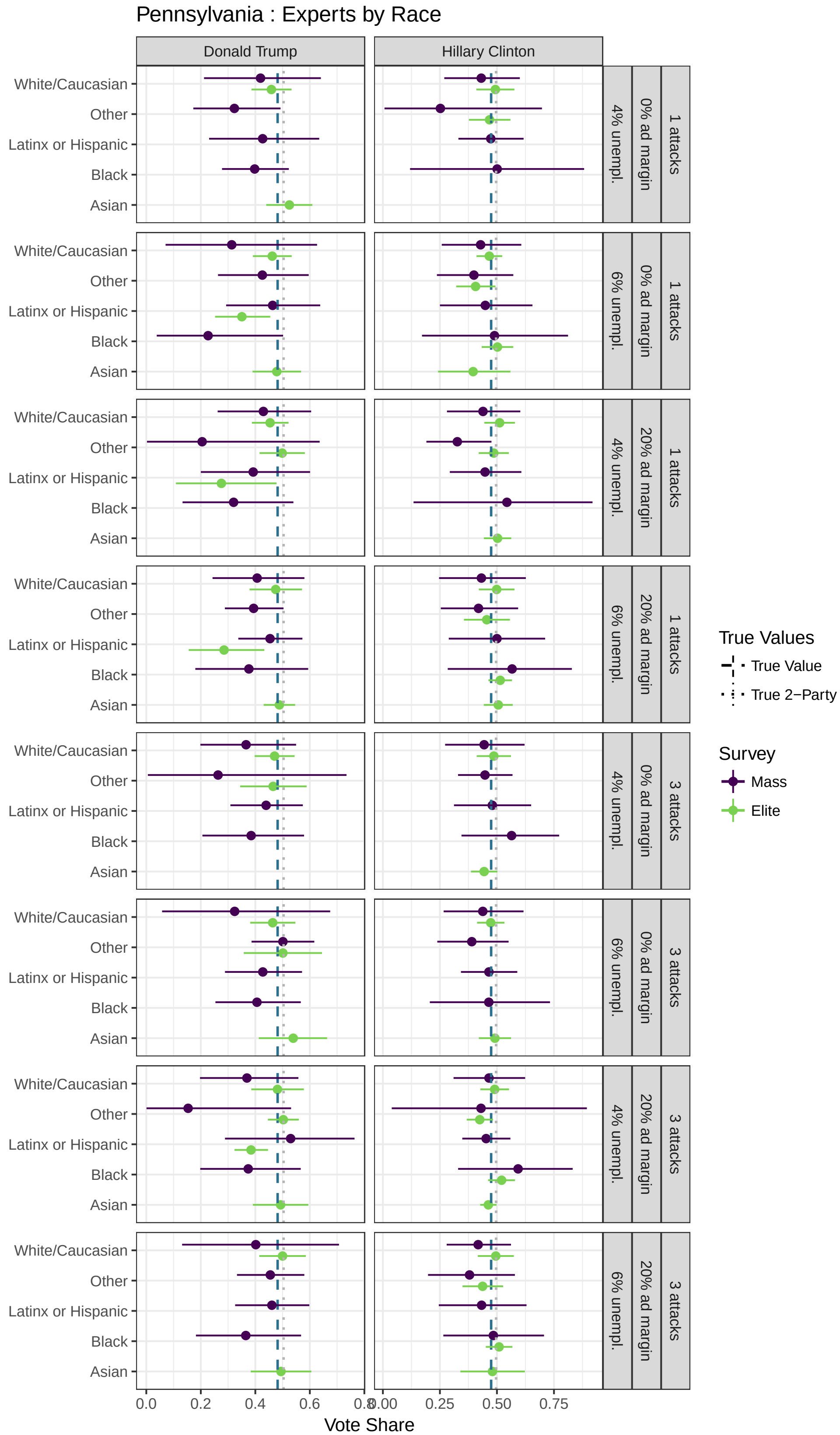


Figure 3.57: Differences with covariates: Pennsylvania Race

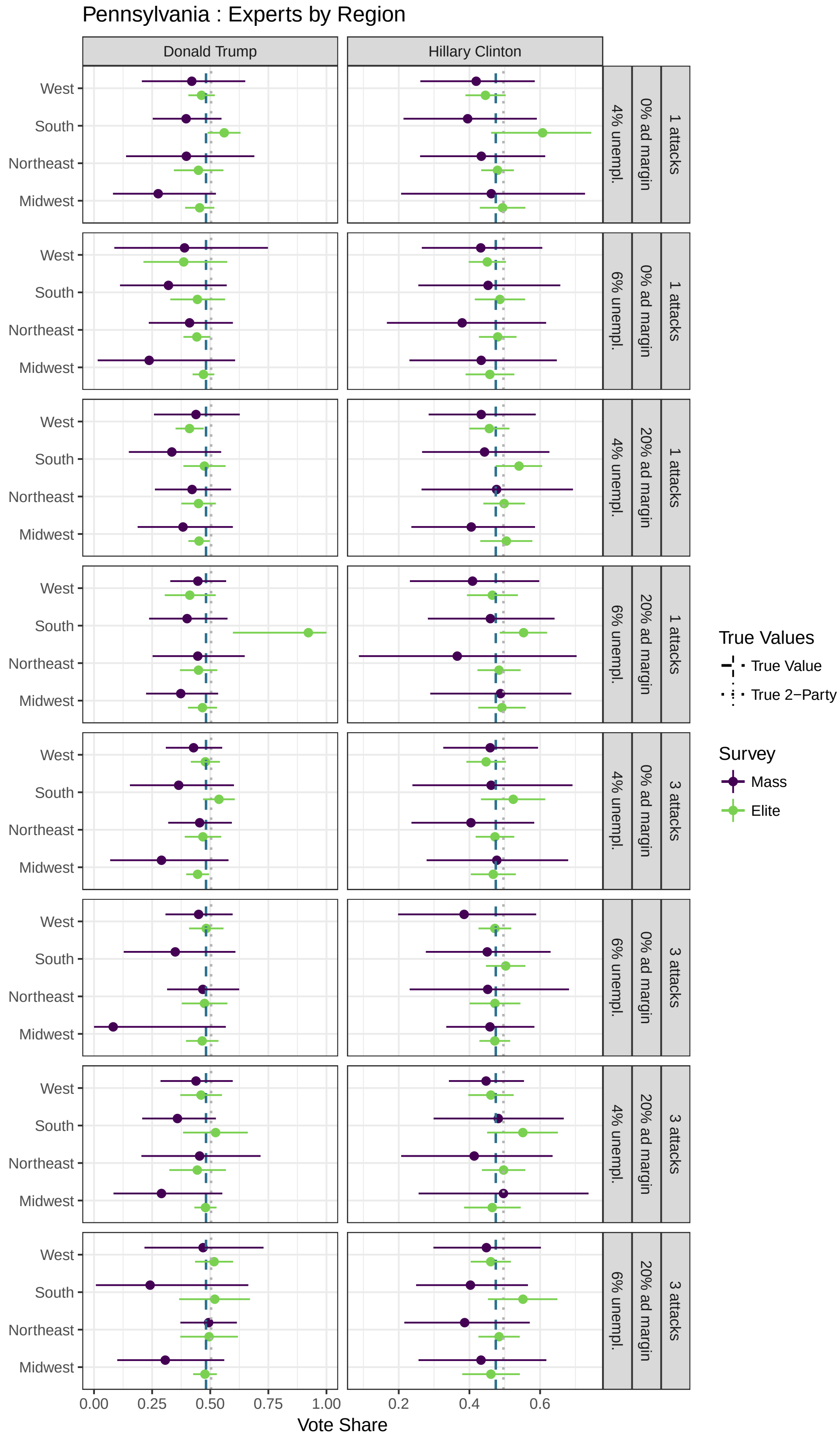


Figure 3.58: Differences with covariates: Pennsylvania Region

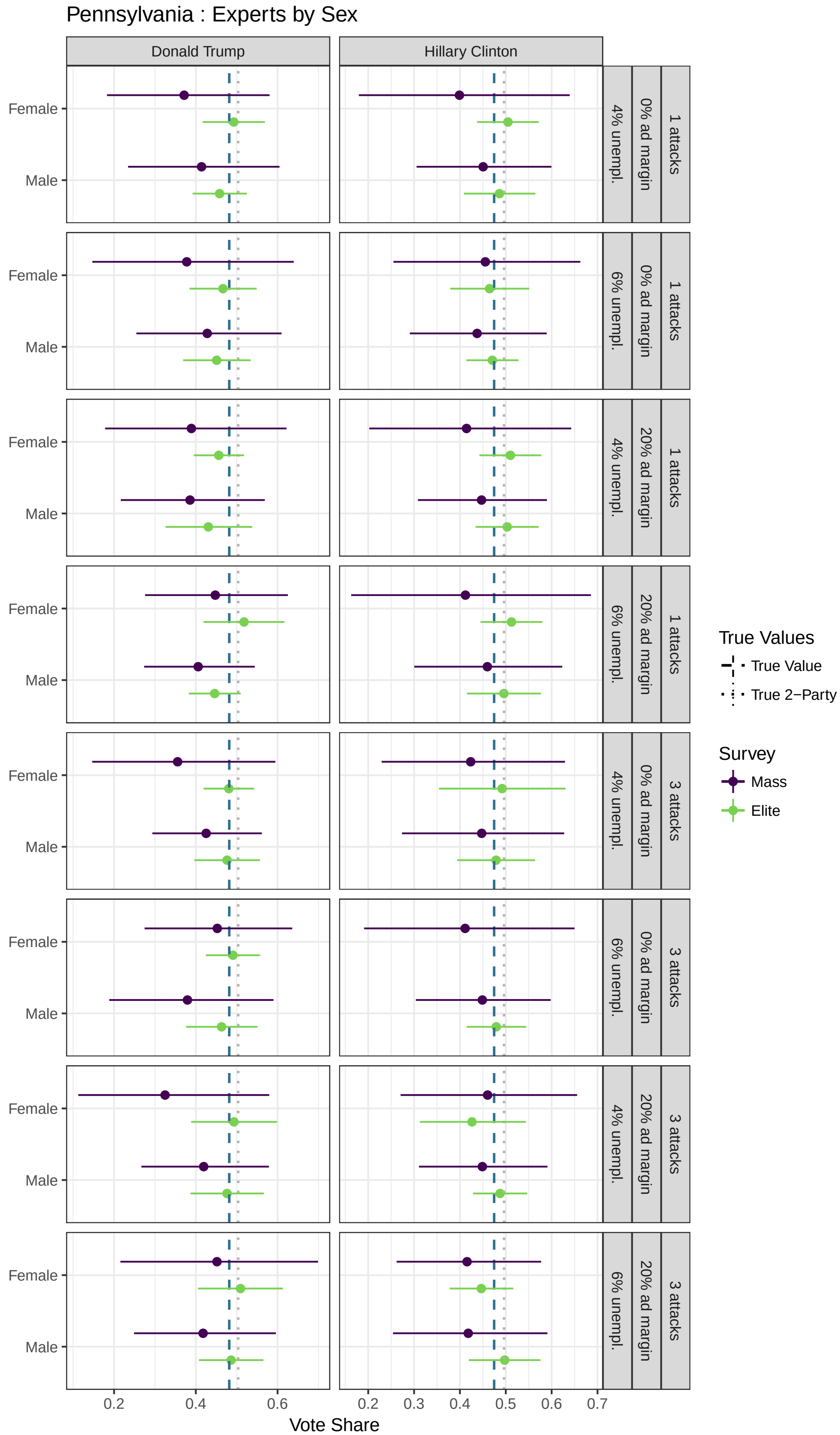


Figure 3.59: Differences with covariates: Pennsylvania Sex

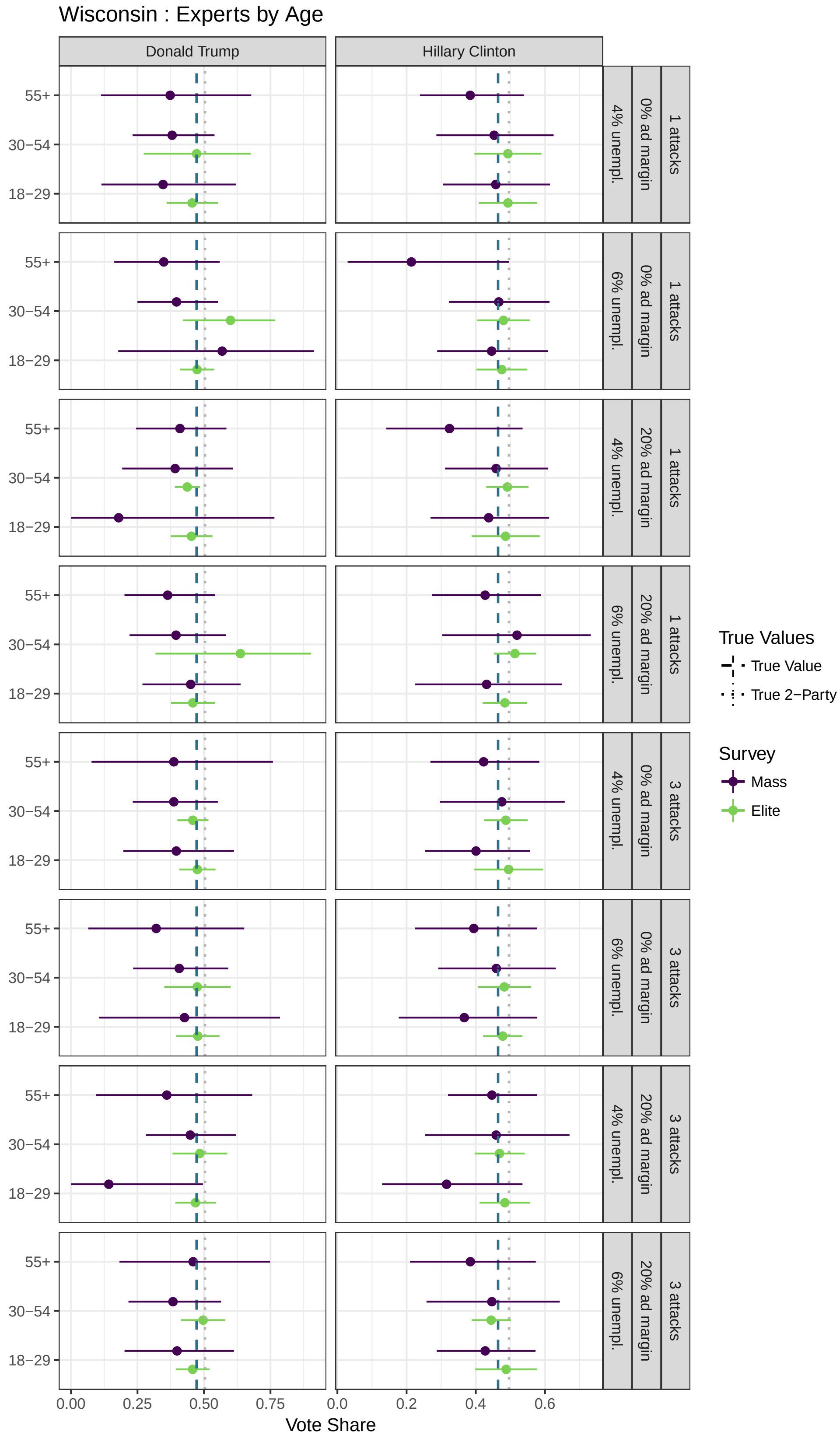


Figure 3.60: Differences with covariates: Wisconsin Age

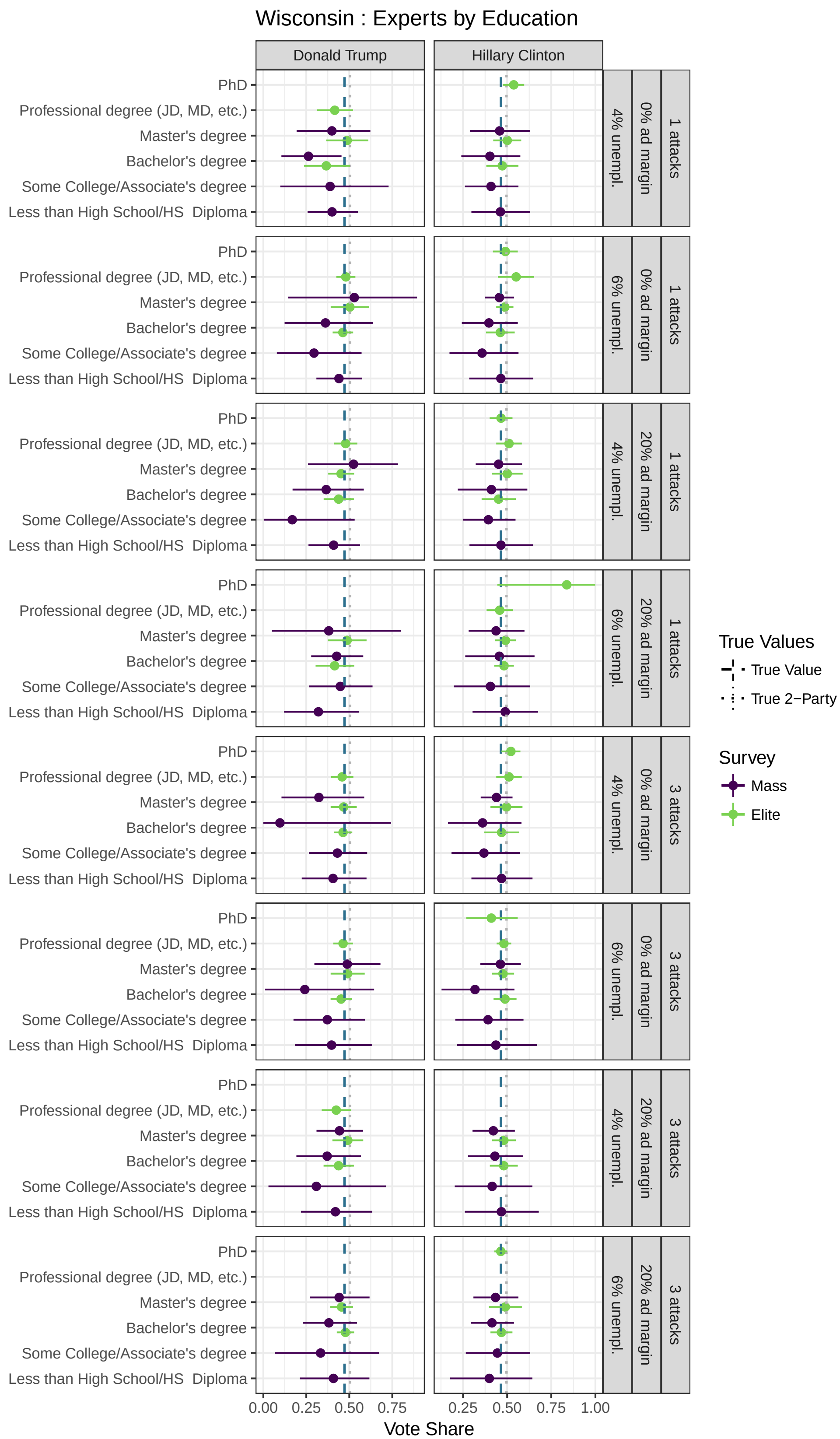


Figure 3.61: Differences with covariates: Wisconsin Education

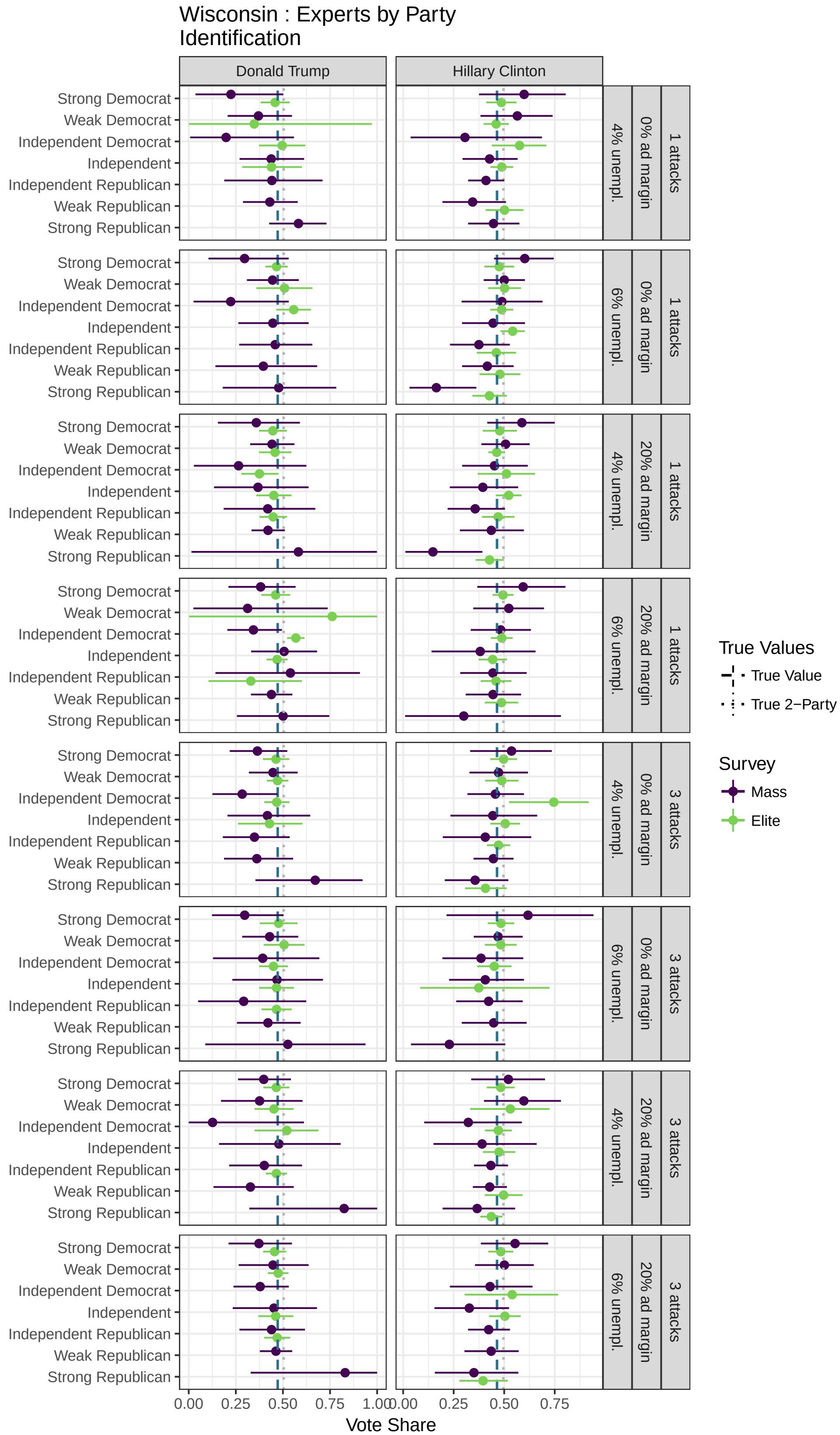


Figure 3.62: Differences with covariates: Wisconsin Party Identification

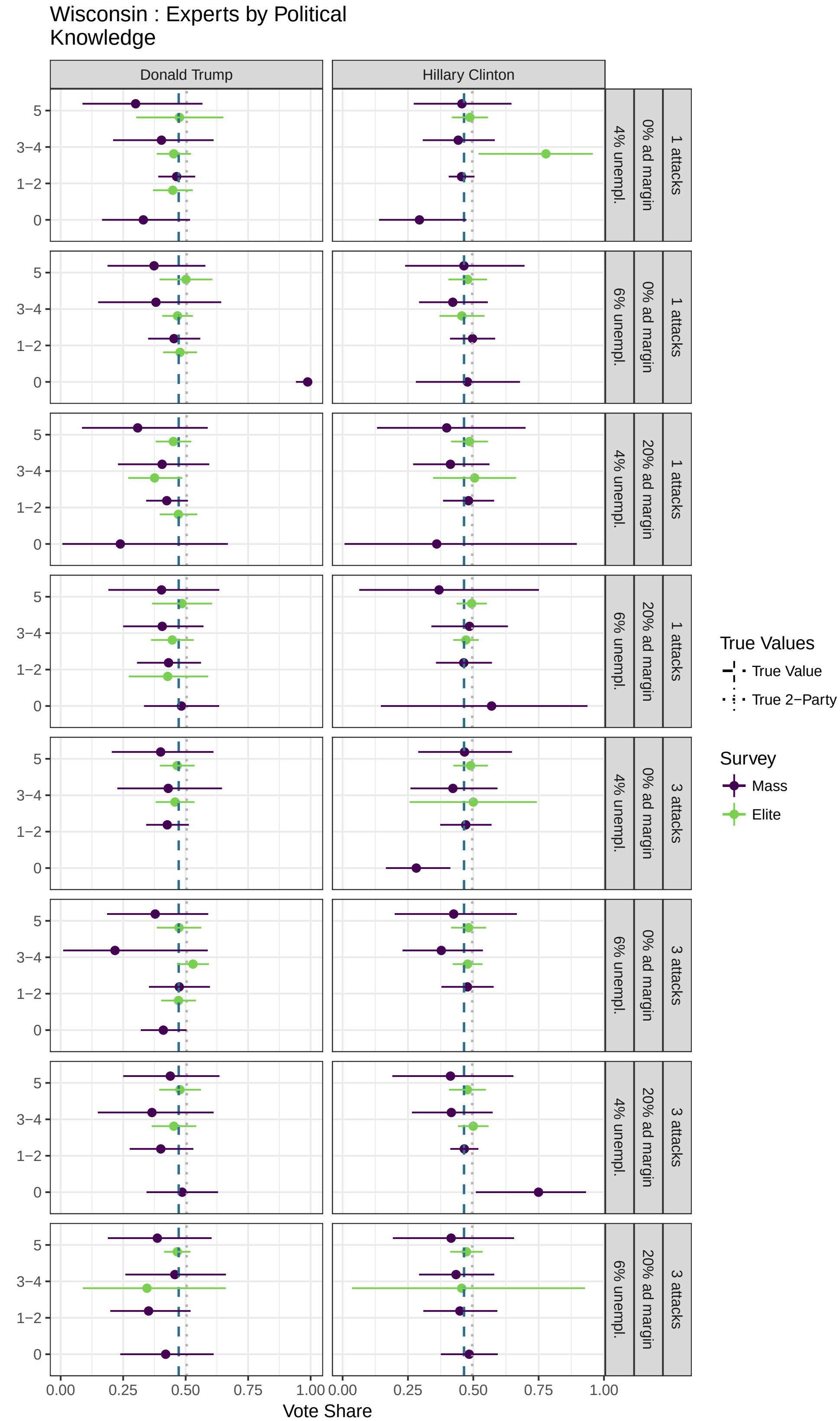


Figure 3.63: Differences with covariates: Wisconsin Political Knowledge

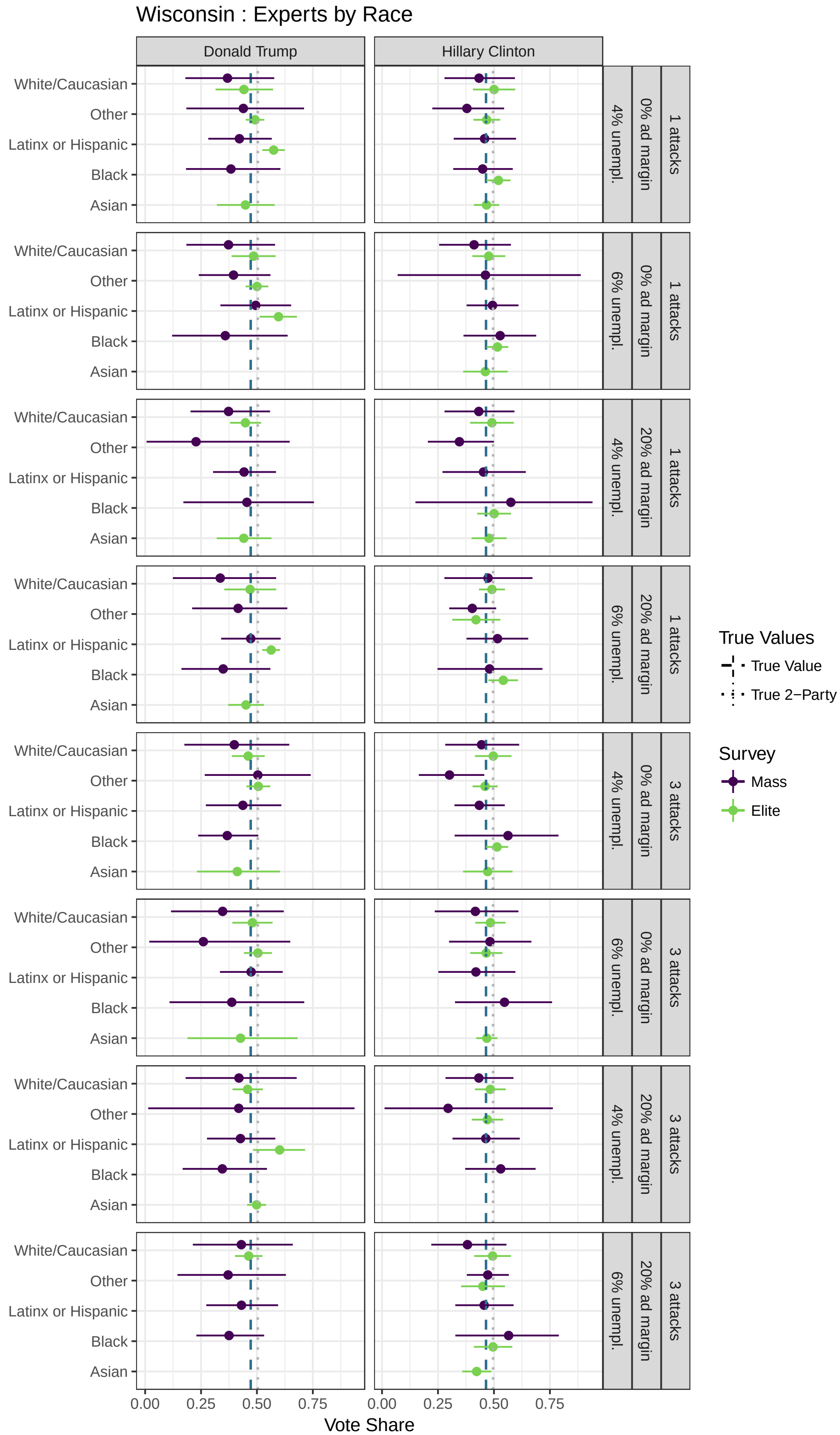


Figure 3.64: Differences with covariates: Wisconsin Race

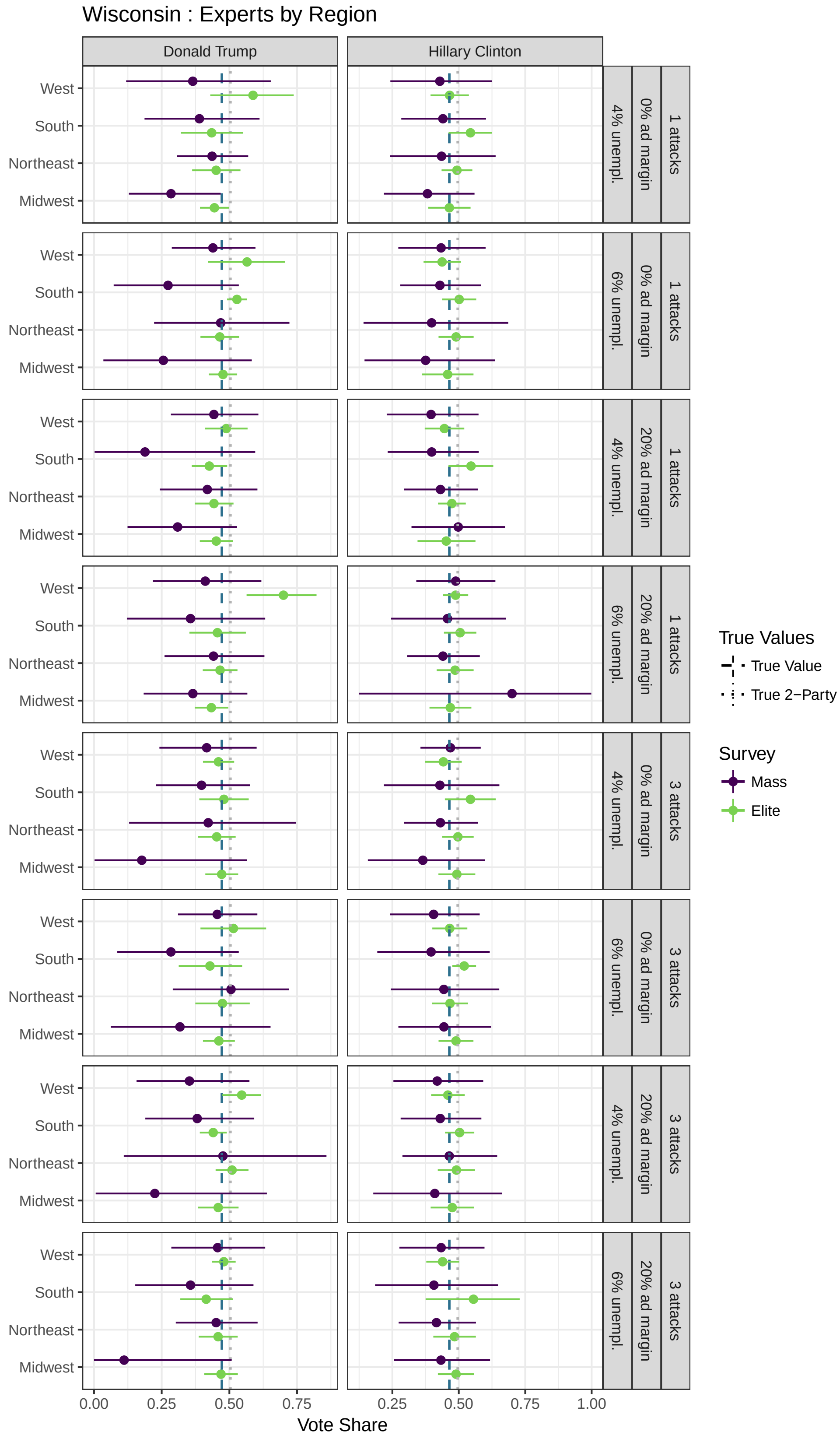


Figure 3.65: Differences with covariates: Wisconsin Region

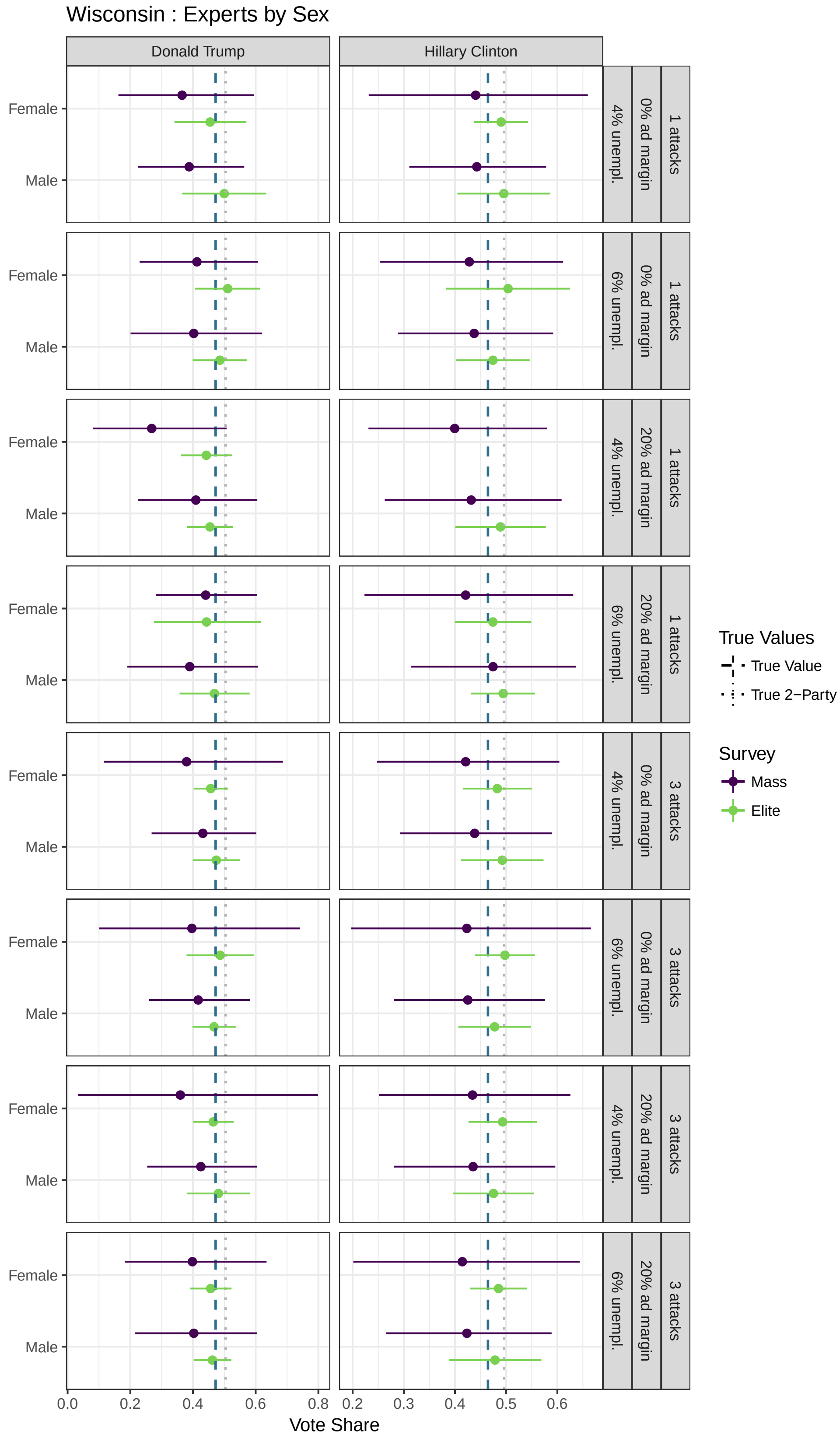


Figure 3.66: Differences with covariates: Wisconsin Sex

3.6 Conclusion & Next Steps

The preceding analysis examines elicited priors concerning outcomes from the 2016 U.S. national election, evaluating the impact of "expertise" across mass and elite samples as well as the role of clustering aggregation in improving predictions. The results evaluated here reflect not only respondents' raw predictions for the national election and returns in Florida, North Carolina, and Ohio, but also incorporate assessments of these returns, in addition to those in Pennsylvania and Wisconsin, in conjunction with relevant covariates proxying for economic and security conditions as well as campaign characteristics. The results incorporating covariate conditions, and those across demographic characteristics of respondents in each of the mass and elite samples, suggest that while education and knowledge may confer the type of "expertise" often assumed in elicitation exercises, this need not be the case. In particular, the modestly sized representative "mass" sample performs comparably to the elite sample in tasks without covariates, and both are relatively accurate, suggesting that "expertise" per se may not overly influence estimates in relatively simple elicitation designs. Likewise, these results demonstrate that the Dirichlet-based clustering proposed in the previous chapter adequately incorporates sometimes disparate estimates to provide an overall accurate evaluation of potential vote share. These results in combination indicate that the selection of "experts" may be flexible to the circumstances; that recruiting experts on the basis of educational credentials may not be necessary where it proves too difficult, but also that a small, non-random "elite" sample of experts can still provide accurate estimates.

A more general question, of course, is the extent to which the abilities of experts in one task, such as this one, correlate with their ability to accurately provide priors related to other questions. Specifically related to election forecasting, for example, research has demonstrated that economic voting may not be as prevalent in contexts outside the United States (Lewis-Beck 2010; Jackman and Marks 1994). This paper seeks to demonstrate that (a) "expertise" for elicitation can reside in those beyond credentialed "experts" and (b) Dirichlet-based clustering of elicited priors is an effective method for aggregating diverse priors to provide useful estimates. Even so, the scope of contexts to which this method most usefully applies requires further definition and specification.

3.6.1 Additional Analysis

The results in this paper will be complemented with several extensions to further interrogate the "expertise" of respondents in each sample. In particular, the following extensions will provide additional insights about how best to identify and construct a pool of "experts" for elicitation:

(1) In addition to including the clustering and analysis for the national election results for the elite sample, the outcomes for U.S. House of Representatives questions for the elite data will be matched against clustered estimates on the basis of covariates (i.e., NOMINATE score, incumbency, unemployment, Obama's vote share).

(2) Further analysis will present results both with the mass and elite samples clustered separately and while combining both samples into a single clustering process.

(3) Clustering results will be evaluated against simple averages for both samples.

(4) The mass survey was conducted from 24–28 October, 2016. The elite survey, by contrast, remained open from 28 October through election day. Because the elite sample had the ability to answer questions through election night, a

further analysis will include groupings by time-stamp to assess whether predictions improve as the election is closer. Literature demonstrates that pre-election polls conducted closer to the election produce more accurate results (Linzer 2013).

(5) Beyond these additions, two extensions to the current analysis can provide a basis for validation of the clustering process itself.

   (a) Both across-cluster variance and error variance can vary by question and can update in the course of the sampling process. Cluster-specific (rather than question-specific) variance can provide some insight into which clusters have more consensus, whether or not this correlates with other measures (e.g., demographic characteristics or political knowledge), and less coherent clusters could be down-weighted.

   (b) In order to implement analyses incorporating the error variance, a prior needs to be applied in order to prevent extreme values (potentially in a hierarchical framework to use a common distribution).

3.6.2 A Note About Expertise & Self-Selection

The elite survey in this project highlights two interesting aspects of expertise-based research that future studies should consider and further interrogate. First, although the elite sample survey required approximately 10–20 minutes to answer, the participation rate remained at 27% of those solicited. Why or under what conditions individuals are willing to engage in research leveraging their expertise requires future investigation in order to understand self-selection into "expert" pools. This observation intersects with the second interesting outcome of the elite survey, which is that although the sample of students solicited was roughly equal in terms of sex, the final set of respondents self-identified as 71% men and only 29% women. This imbalance suggests a difference in what types of individuals might self-identify as "experts" in a given domain, while also reflecting potential differences in opportunity costs for solicited experts that future research should compensate. Distinct from the concern that researchers conducting elicited-priors analyses may select unbalanced samples of experts, this outcome indicates that even a good-faith effort to generate an inclusive sample may not always yield ideal results. Rather, the clustering process articulated in this paper contributes to ameliorating these types of discrepancies in at least two possible ways. First, the clustering process itself may make transparent whether an analysis incorporates only an inordinately small number of "schools of thought"—suggesting that further elicitation may be necessary. Second, to the extent that women or minorities' perspectives diverge from majority perspectives in ways that distinguish them in separate clusters, the clustering process provides greater weight to their opinions and can compensate for under-sampling.

3.7 Appendix

3.7.1 Prior Plots: Without Covariates

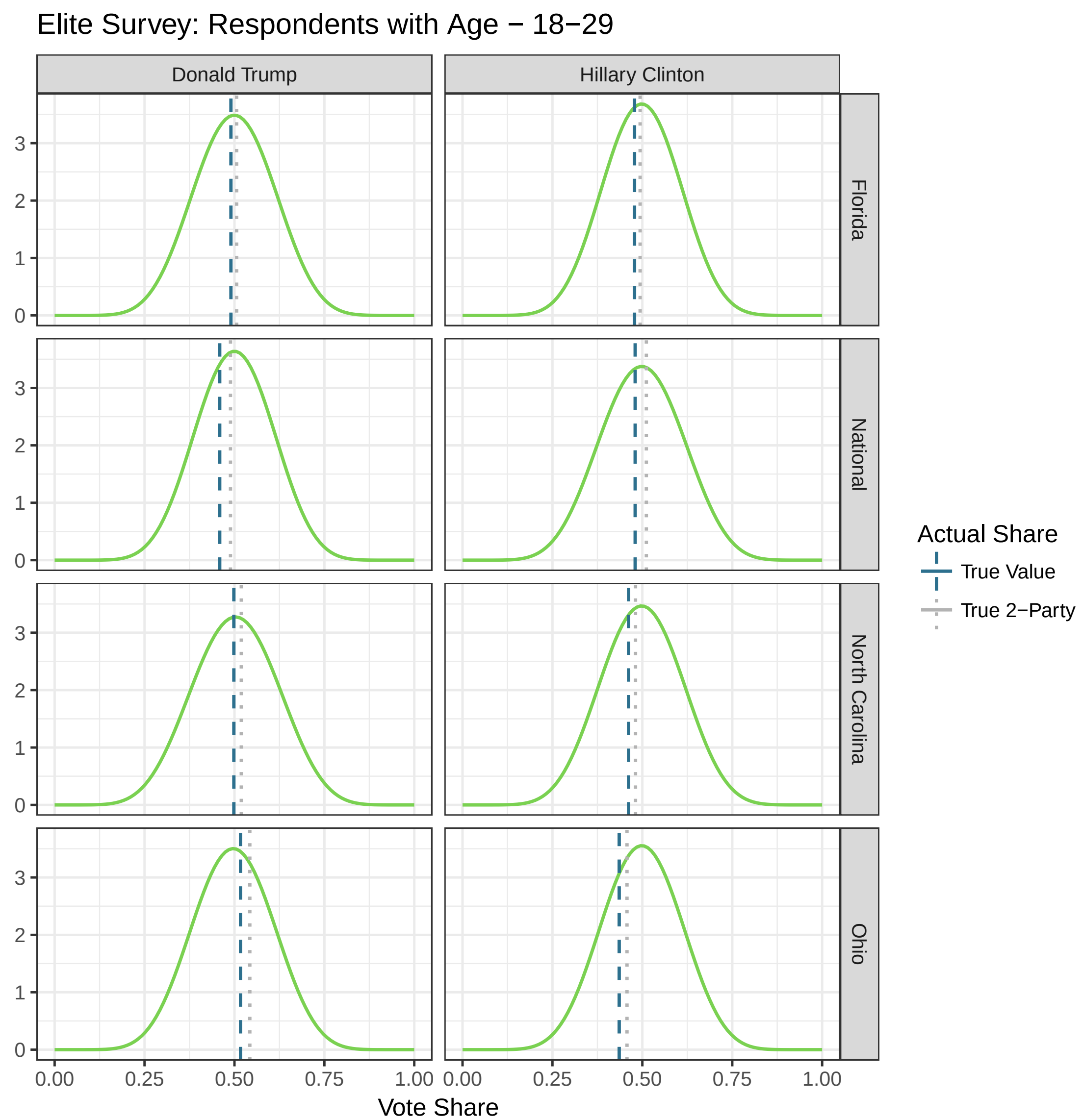


Figure 3.67: Priors without covariates: Elite Age 18-29

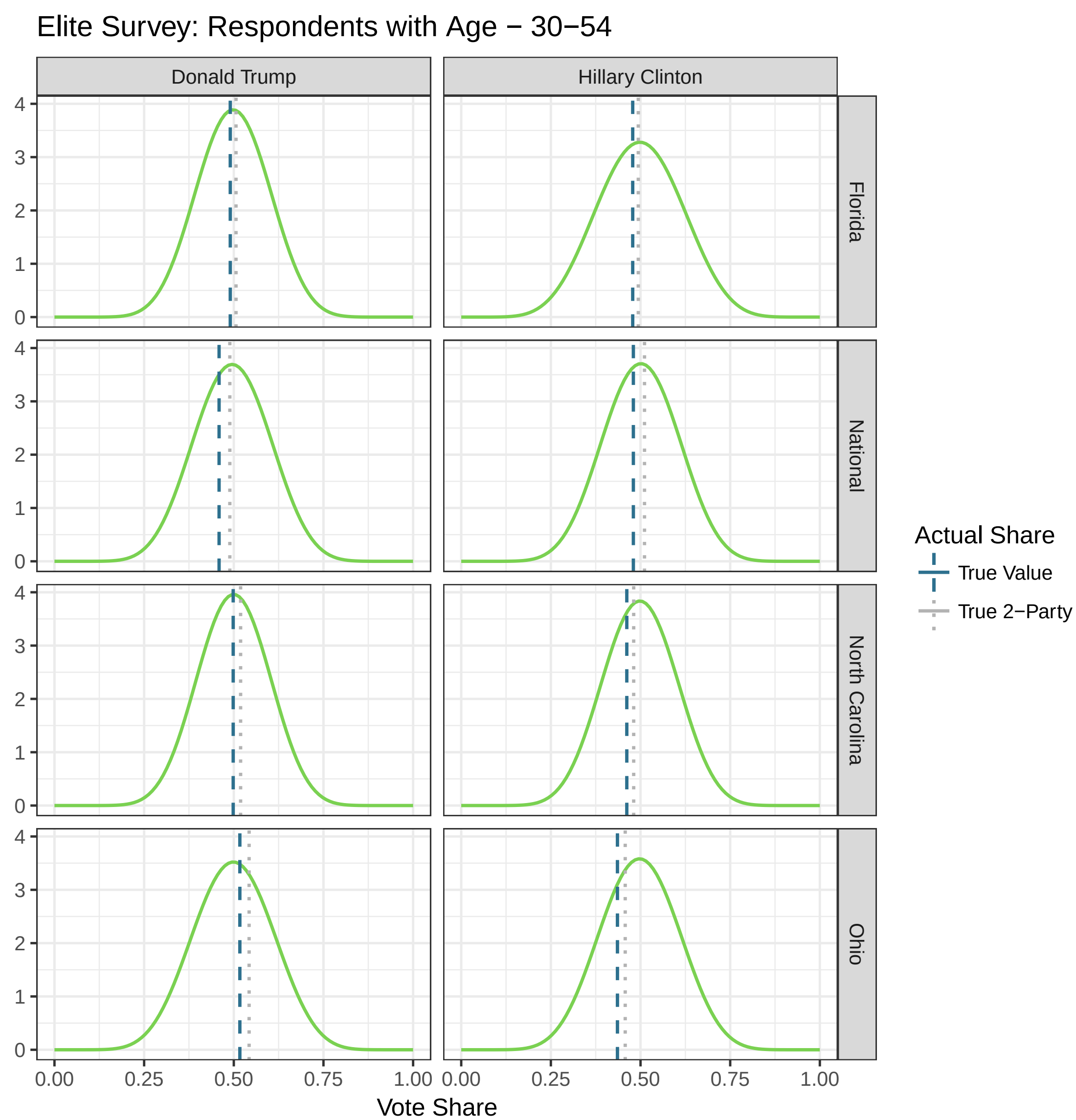


Figure 3.68: Priors without covariates: Elite Age 30-54

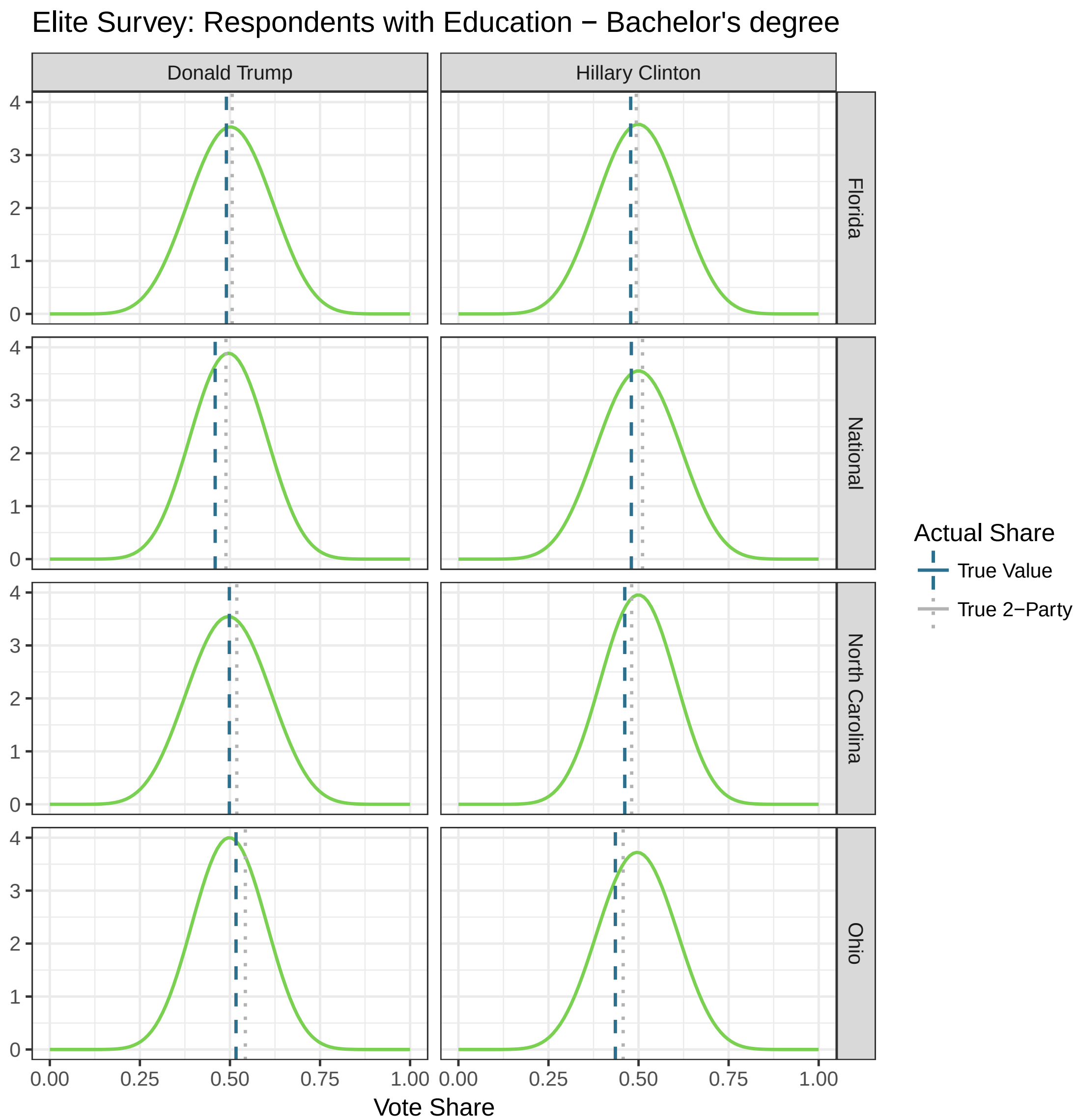


Figure 3.69: Priors without covariates: Elite Education Bachelor's degree

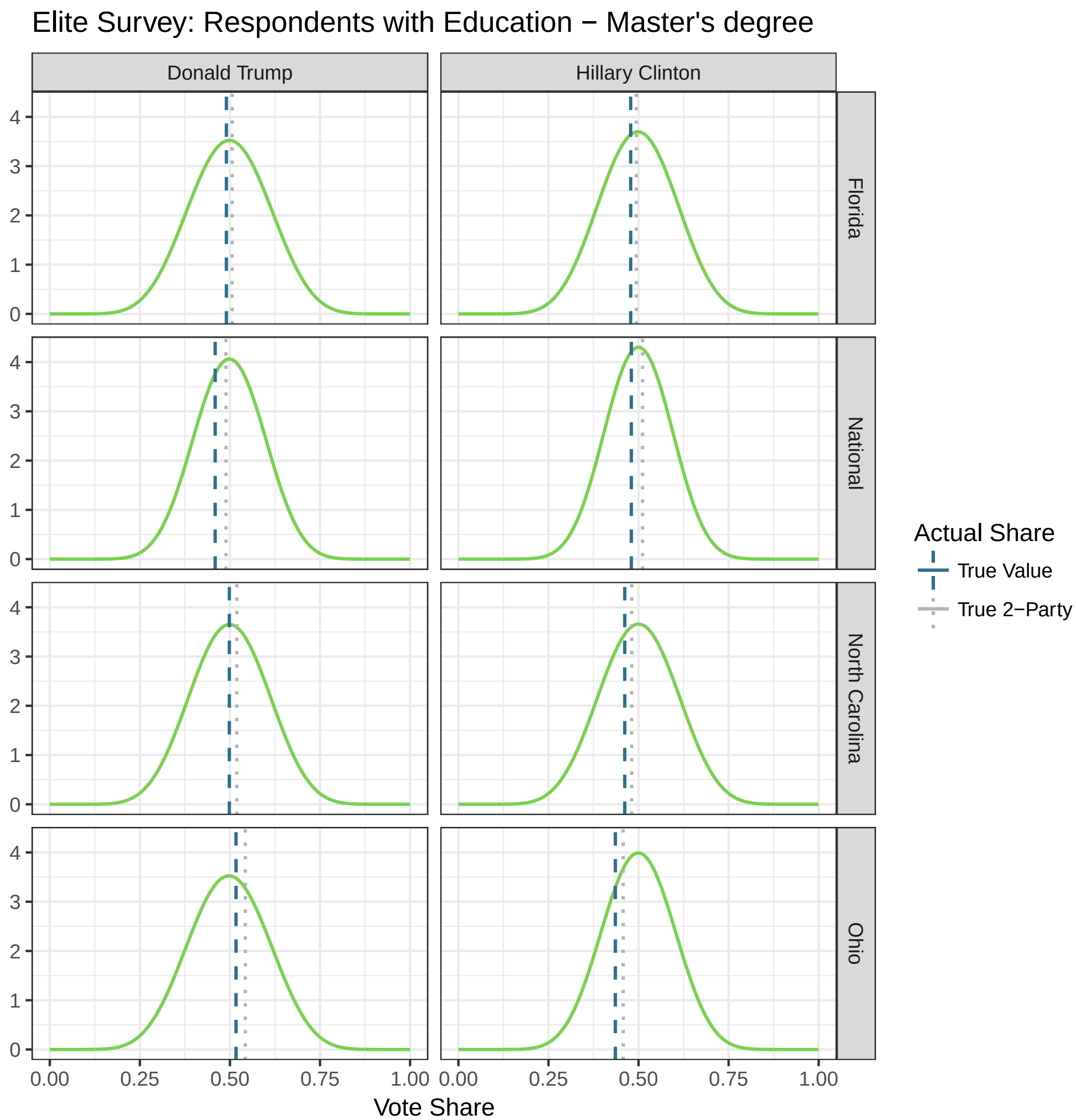


Figure 3.70: Priors without covariates: Elite Education Master's degree

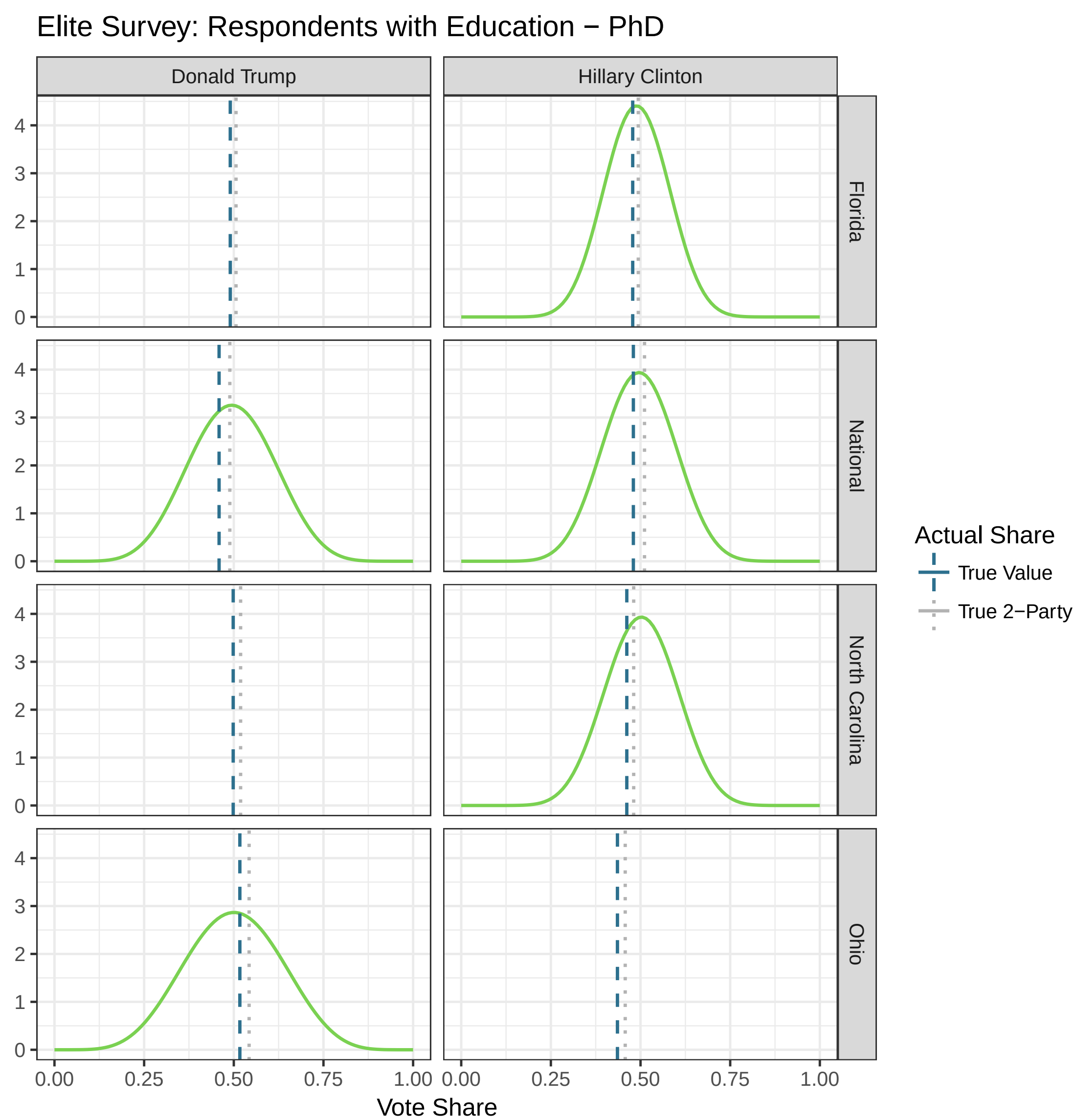


Figure 3.71: Priors without covariates: Elite Education PhD

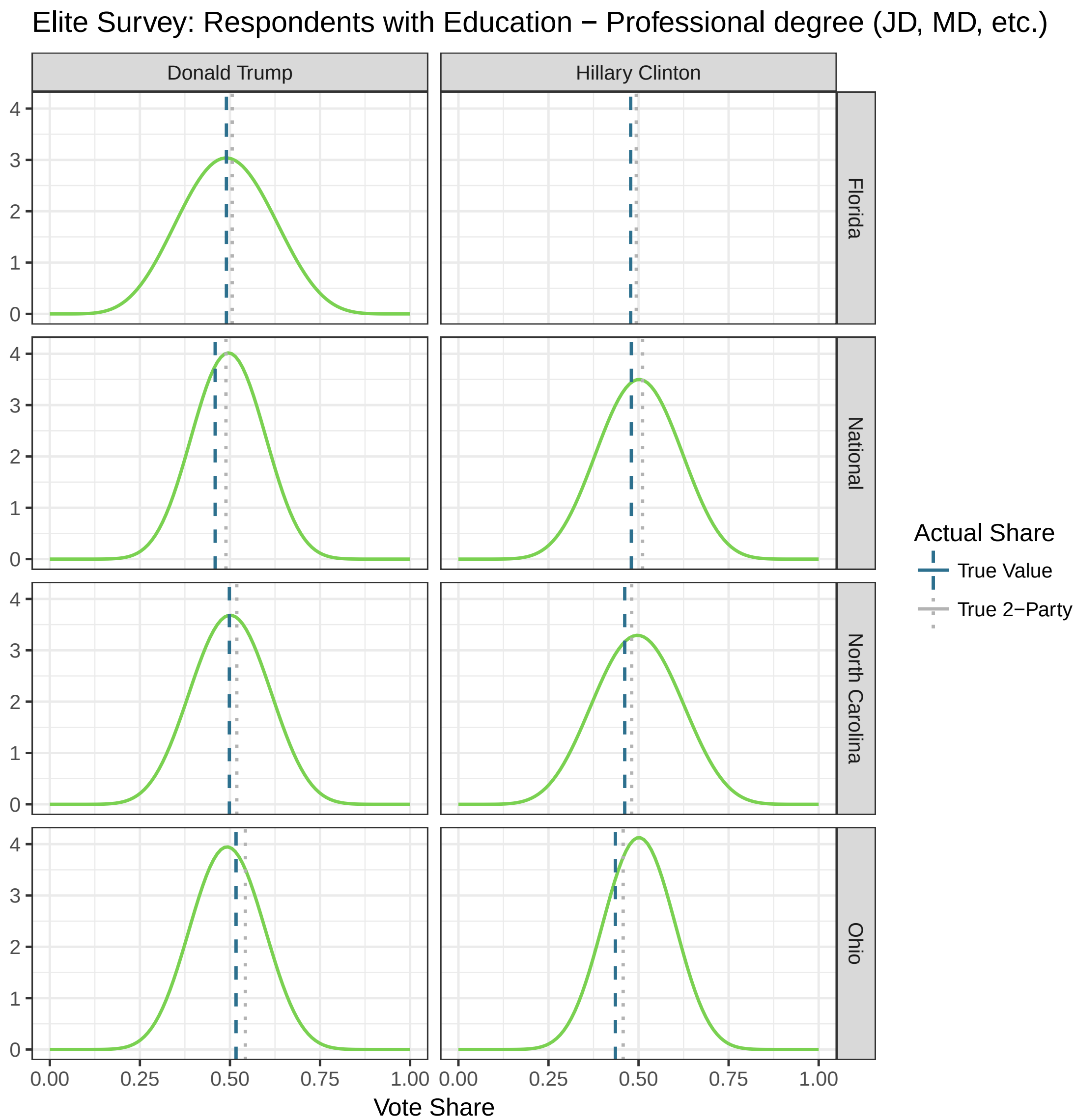


Figure 3.72: Priors without covariates: Elite Education Professional degree JD MD etc

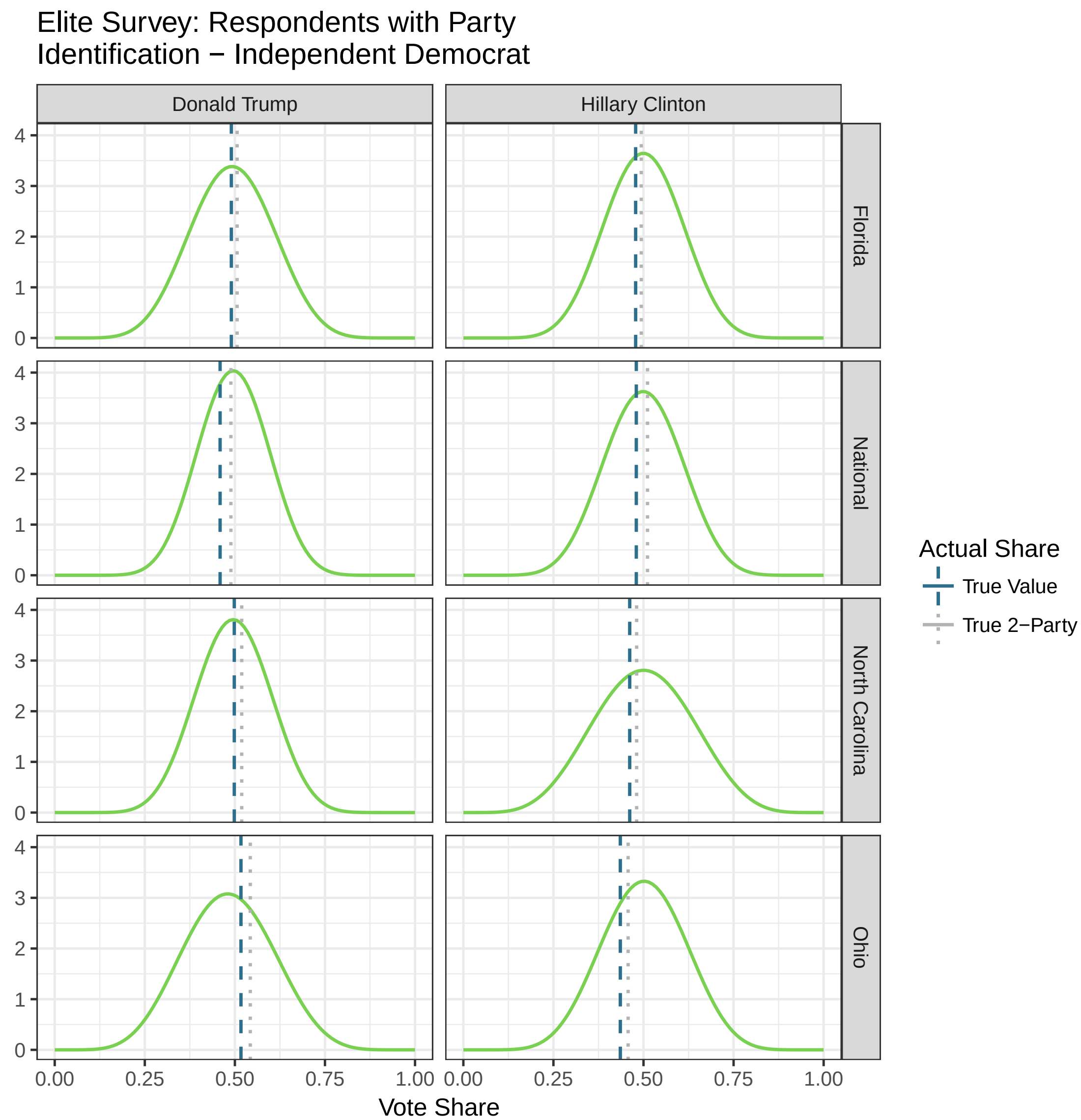


Figure 3.73: Priors without covariates: Elite Party Identification Independent Democrat

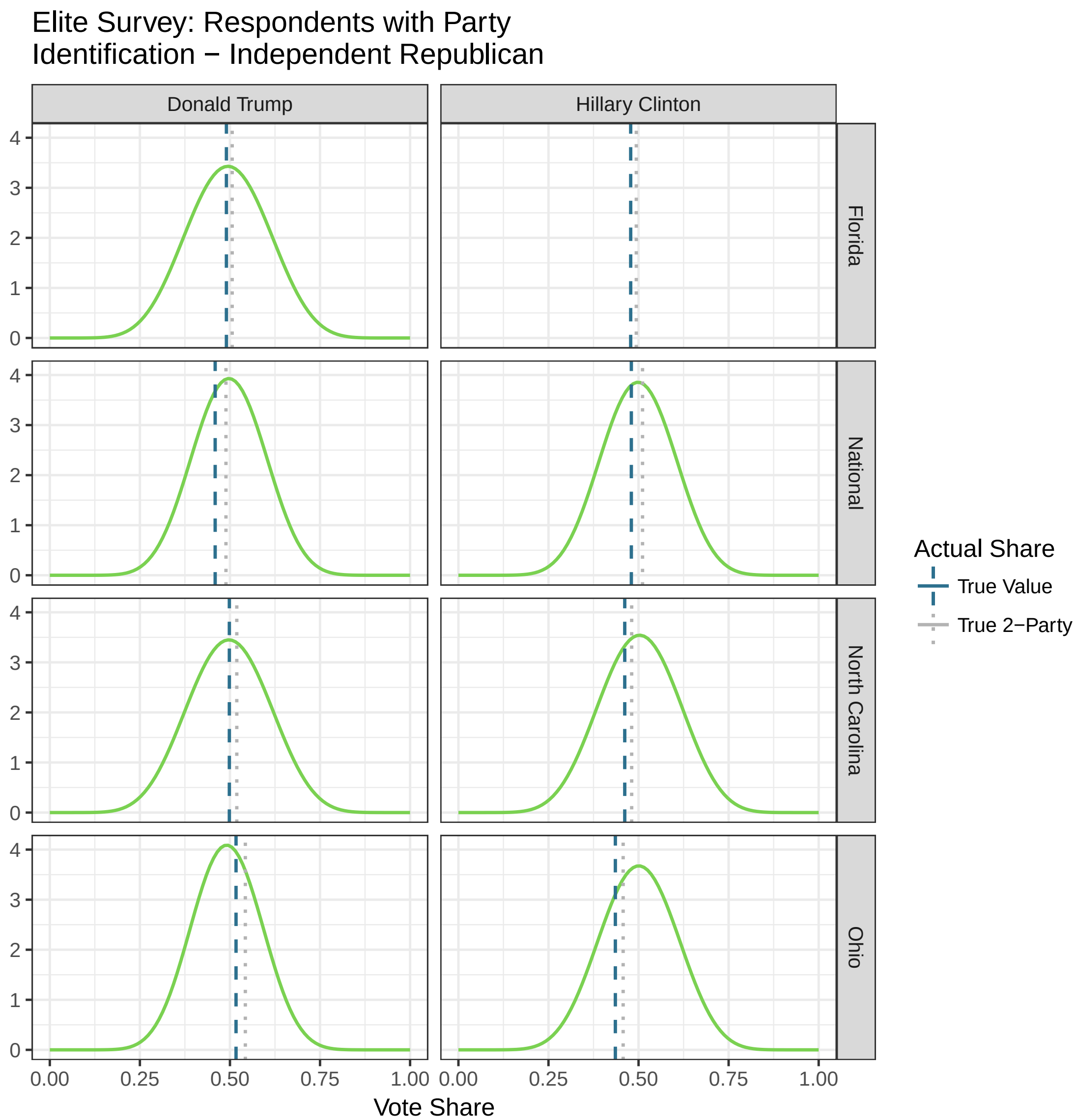


Figure 3.74: Priors without covariates: Elite Party Identification Independent Republican

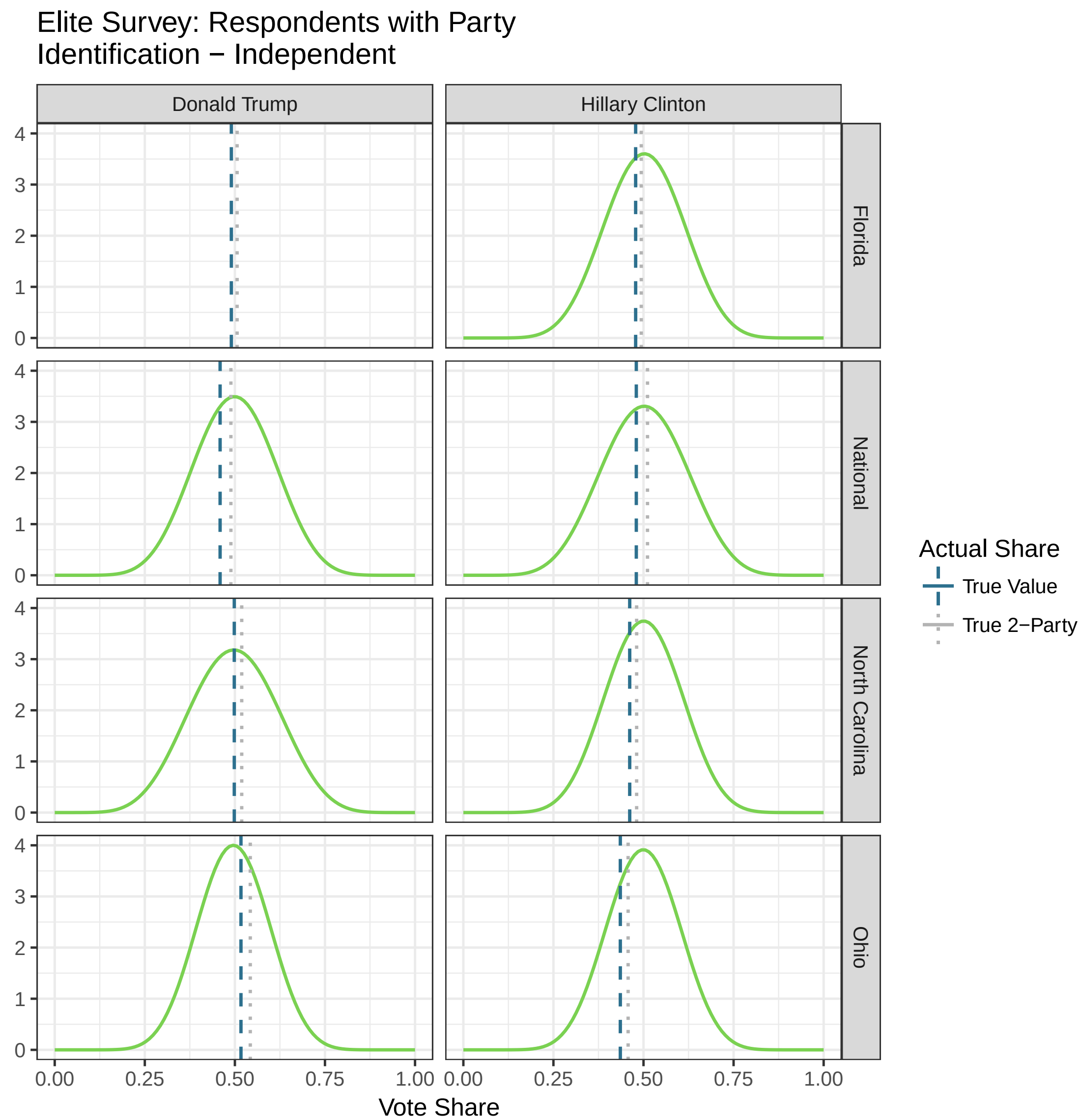


Figure 3.75: Priors without covariates: Elite Party Identification Independent

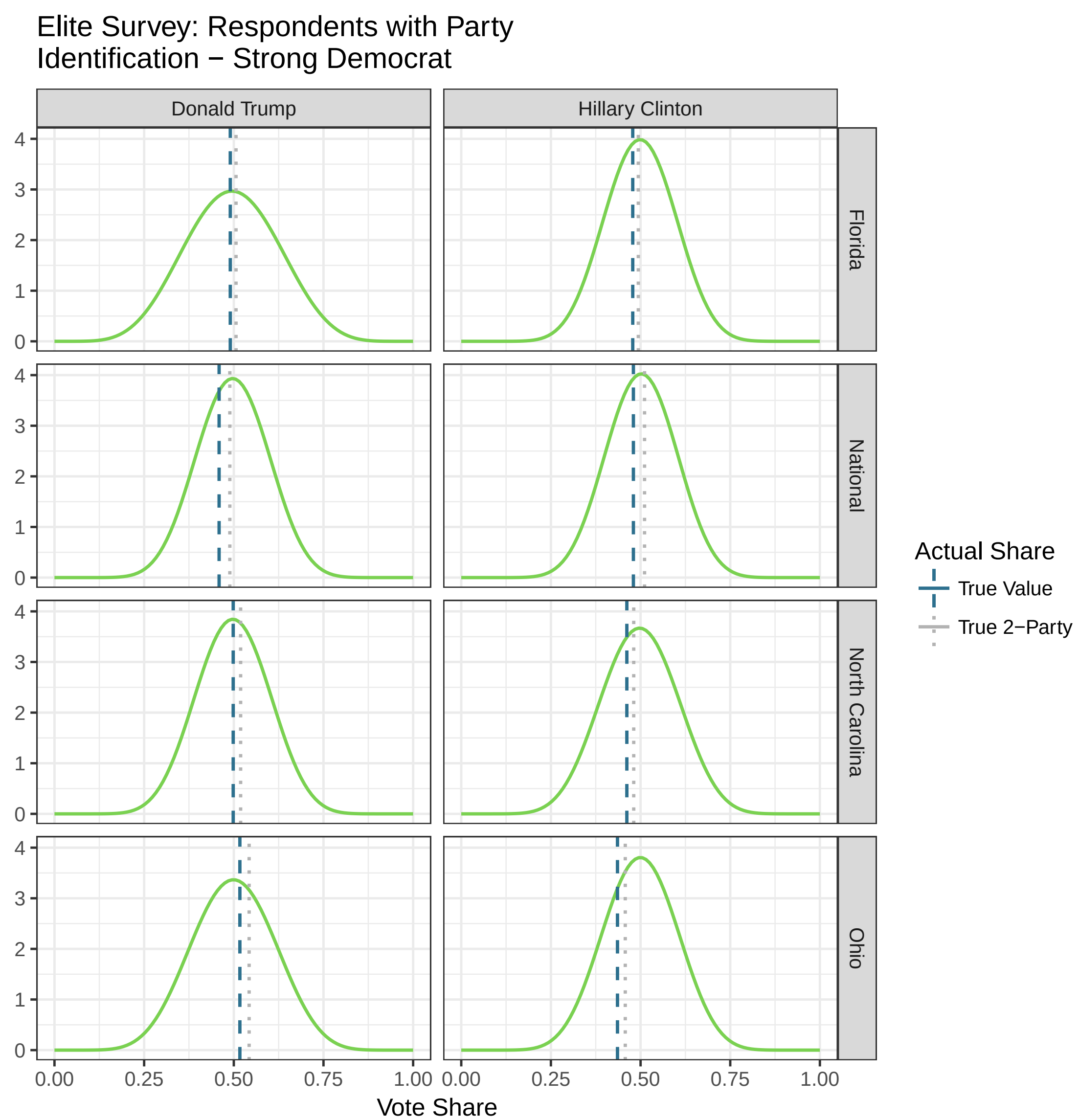


Figure 3.76: Priors without covariates: Elite Party Identification Strong Democrat

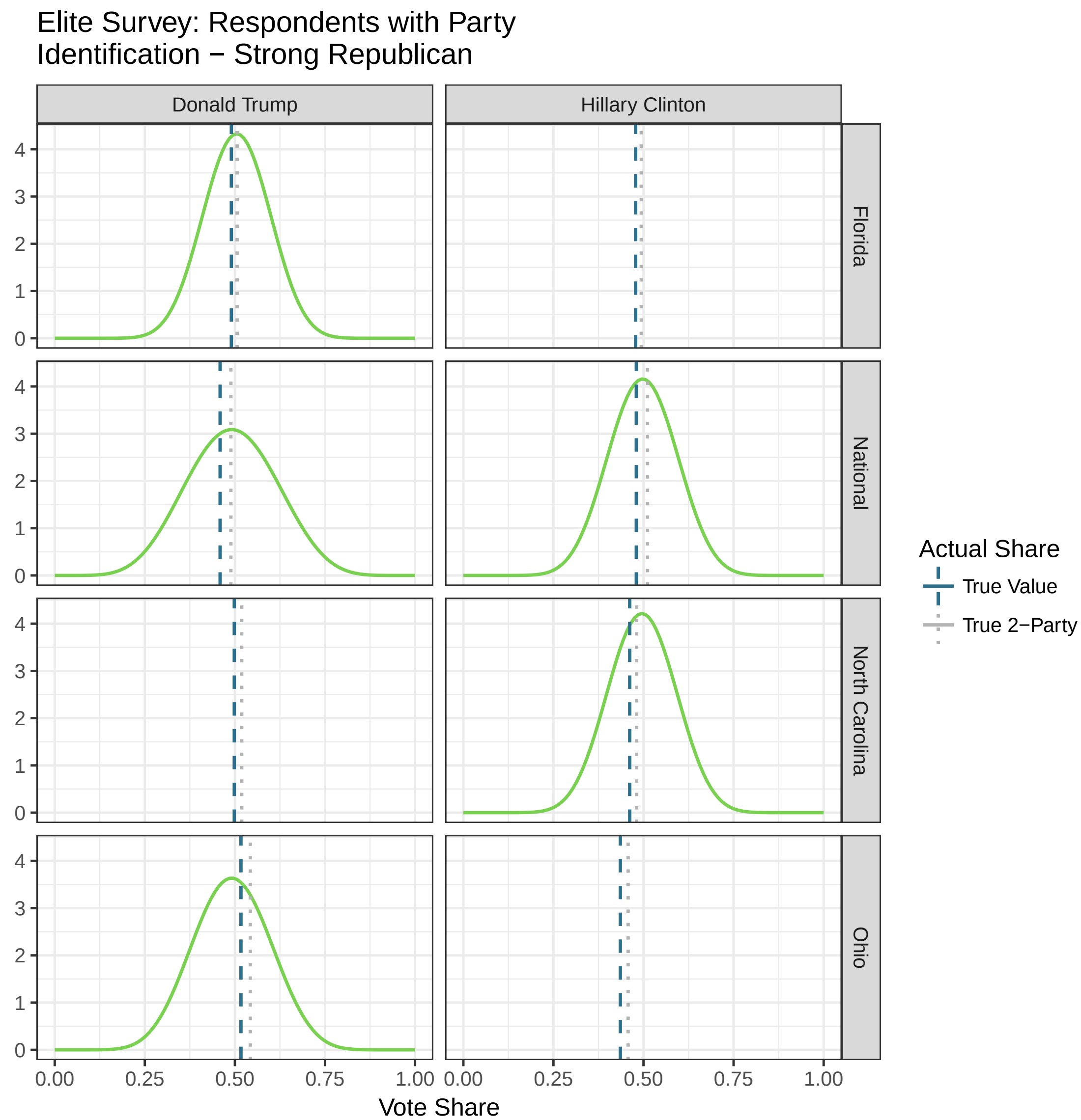


Figure 3.77: Priors without covariates: Elite Party Identification Strong Republican

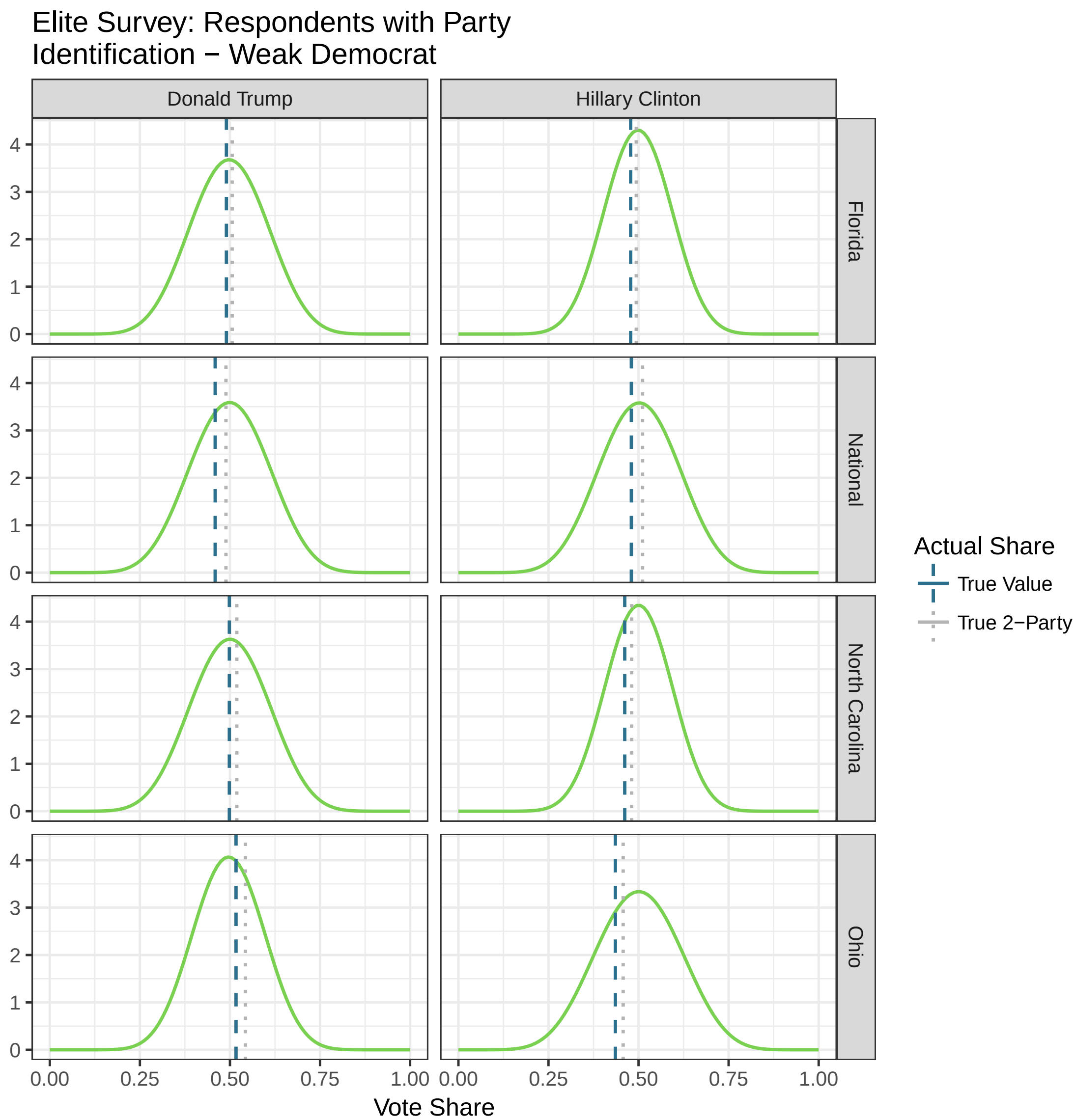


Figure 3.78: Priors without covariates: Elite Party Identification Weak Democrat

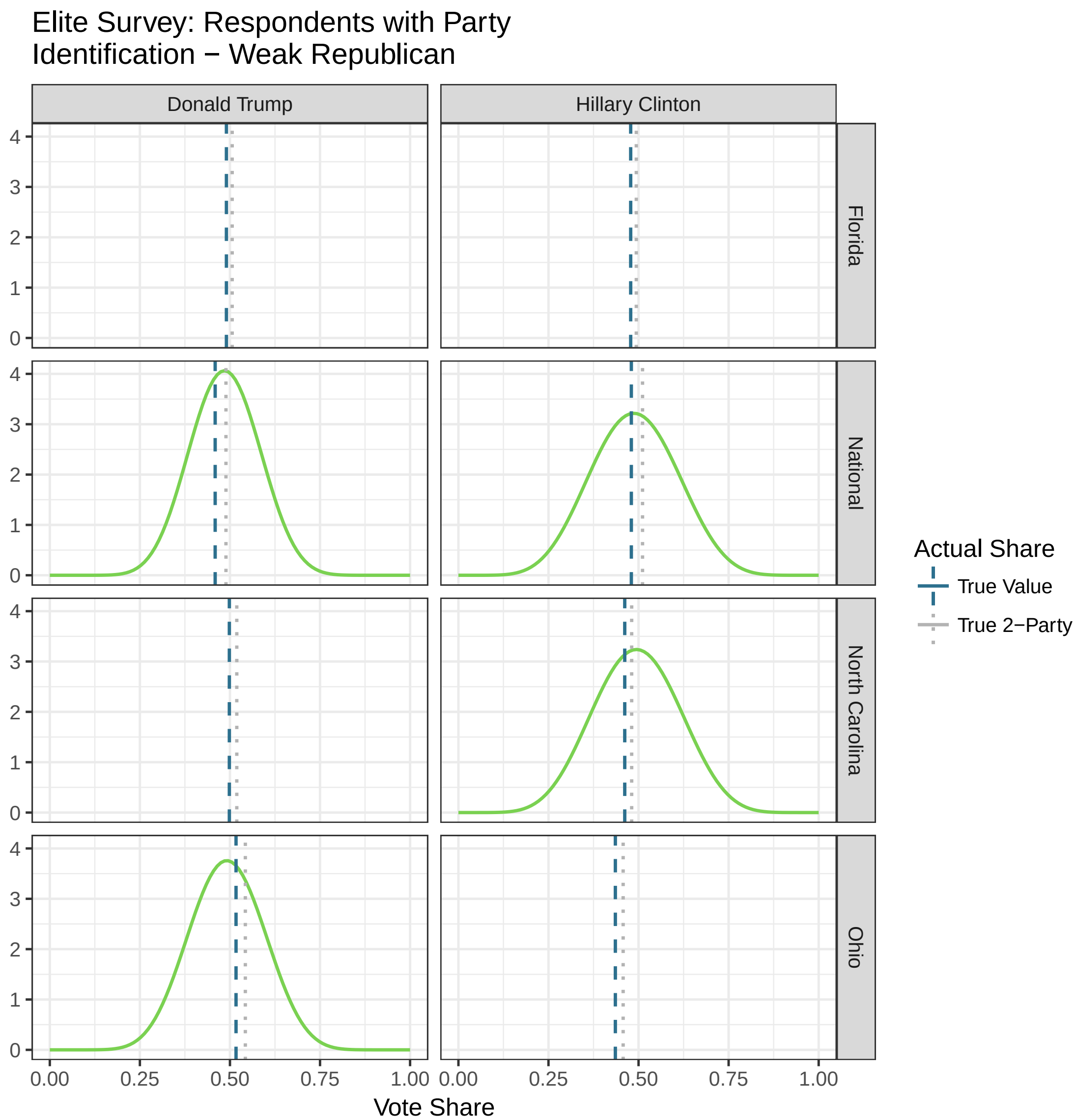


Figure 3.79: Priors without covariates: Elite Party Identification Weak Republican

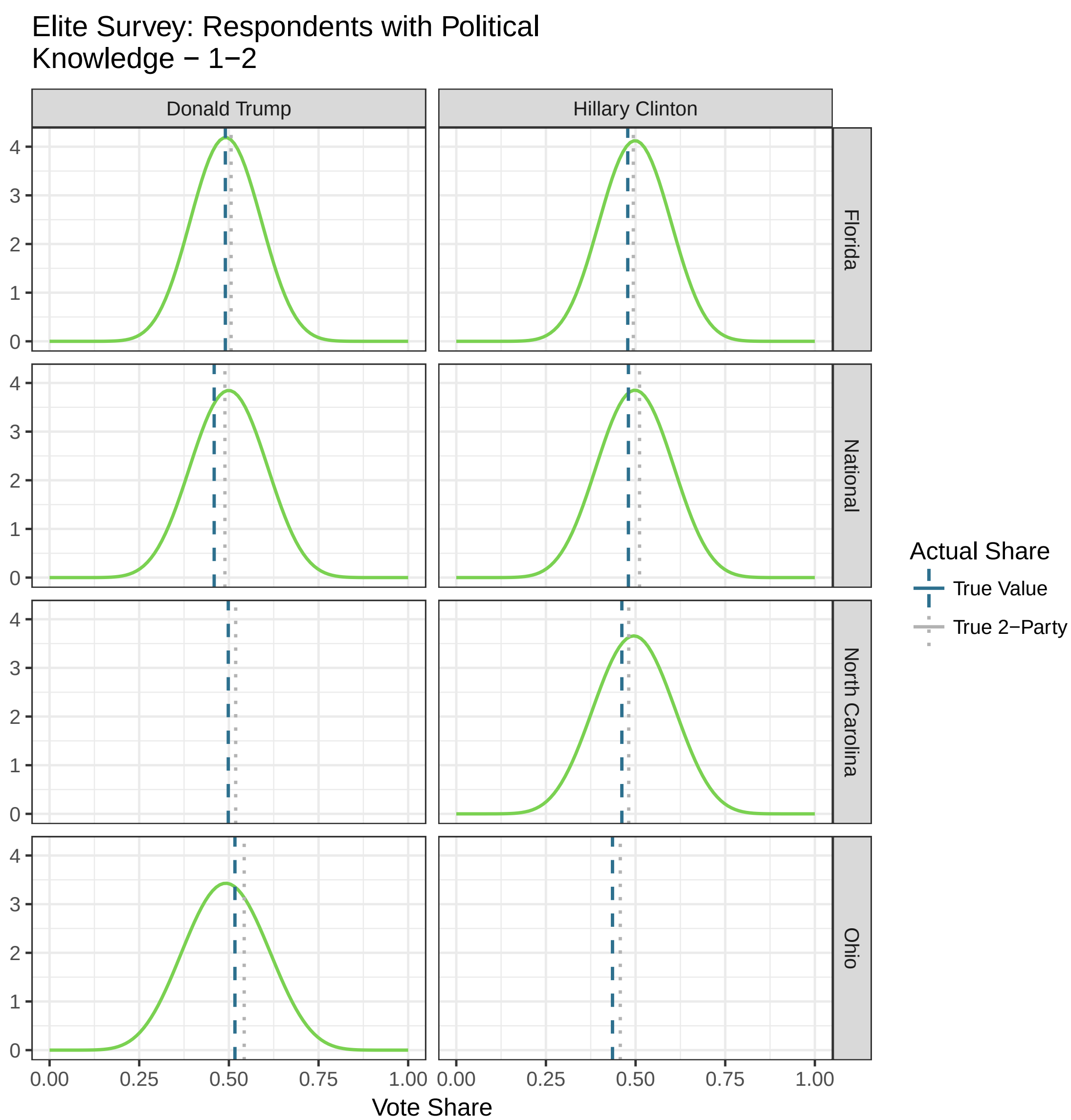


Figure 3.80: Priors without covariates: Elite Political Knowledge 1-2

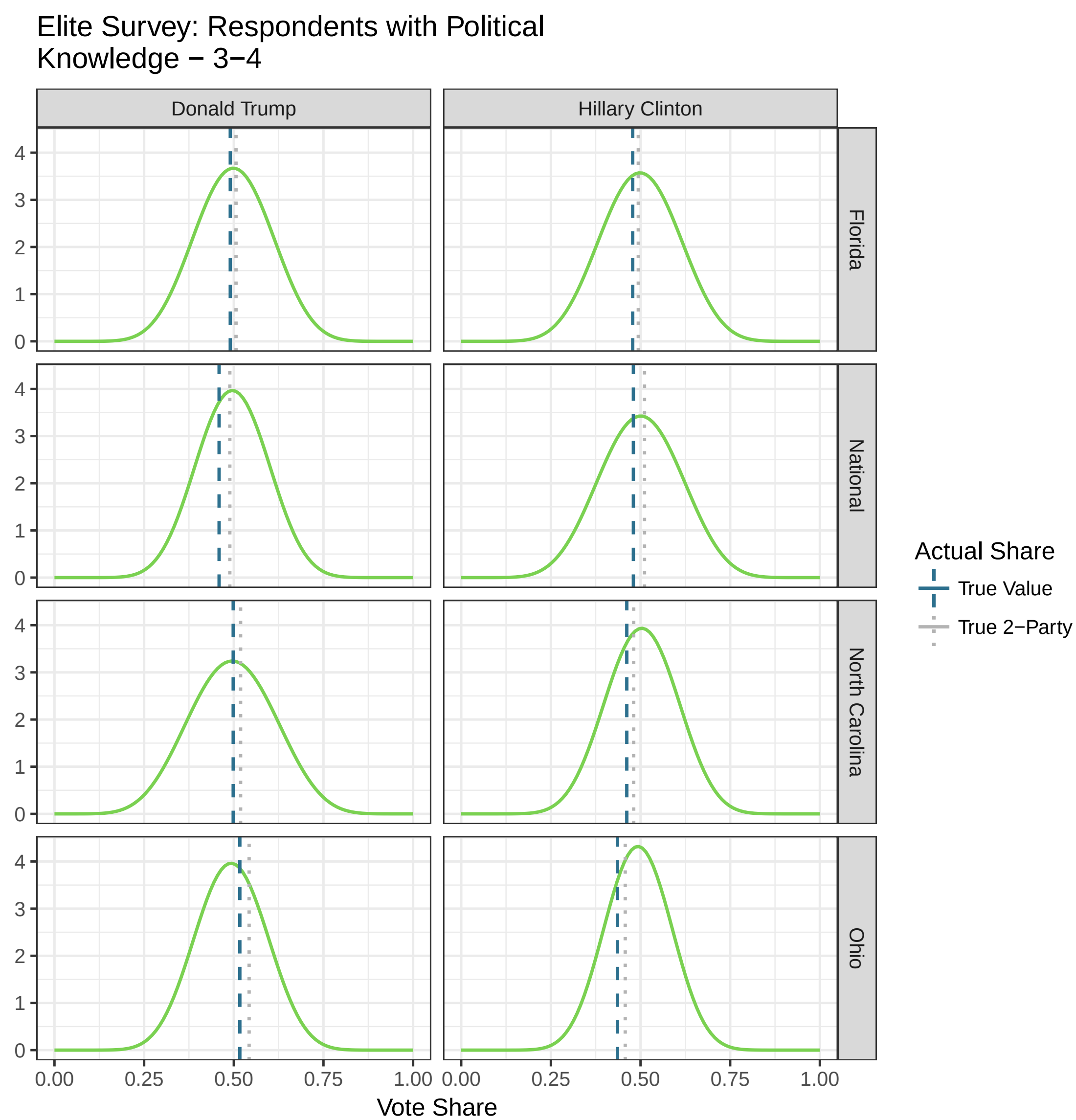


Figure 3.81: Priors without covariates: Elite Political Knowledge 3-4

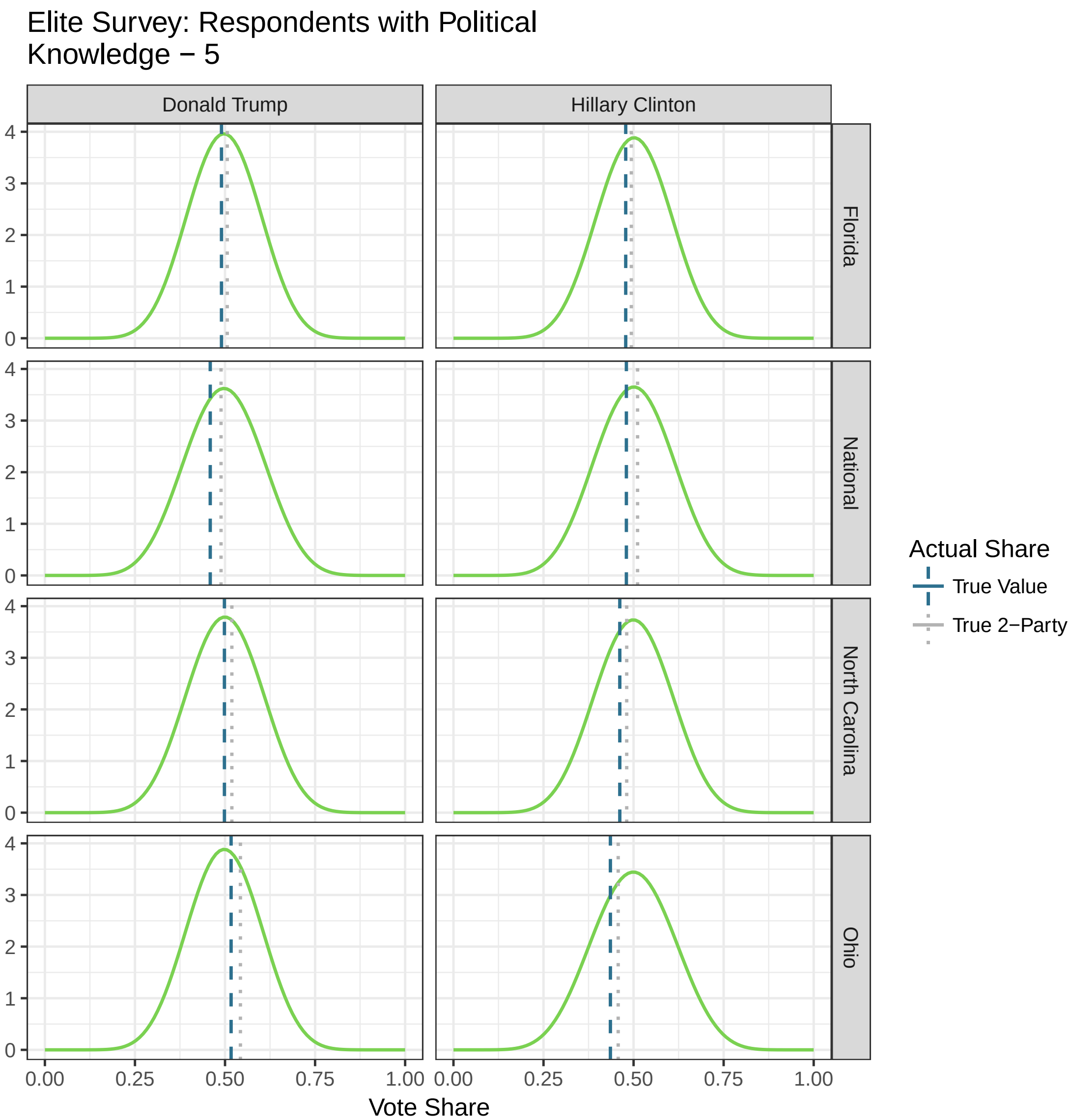


Figure 3.82: Priors without covariates: Elite Political Knowledge 5

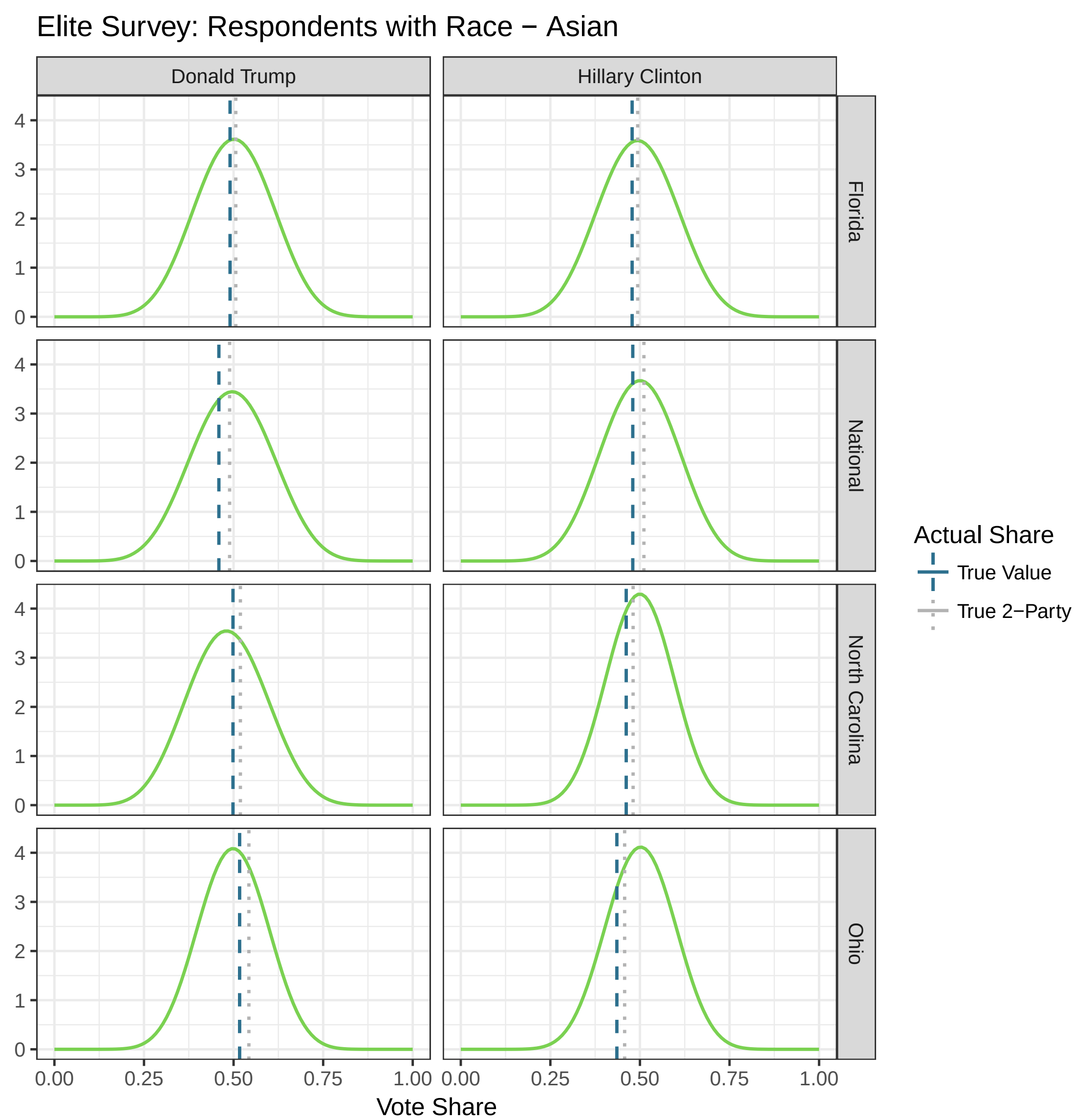


Figure 3.83: Priors without covariates: Elite Race Asian

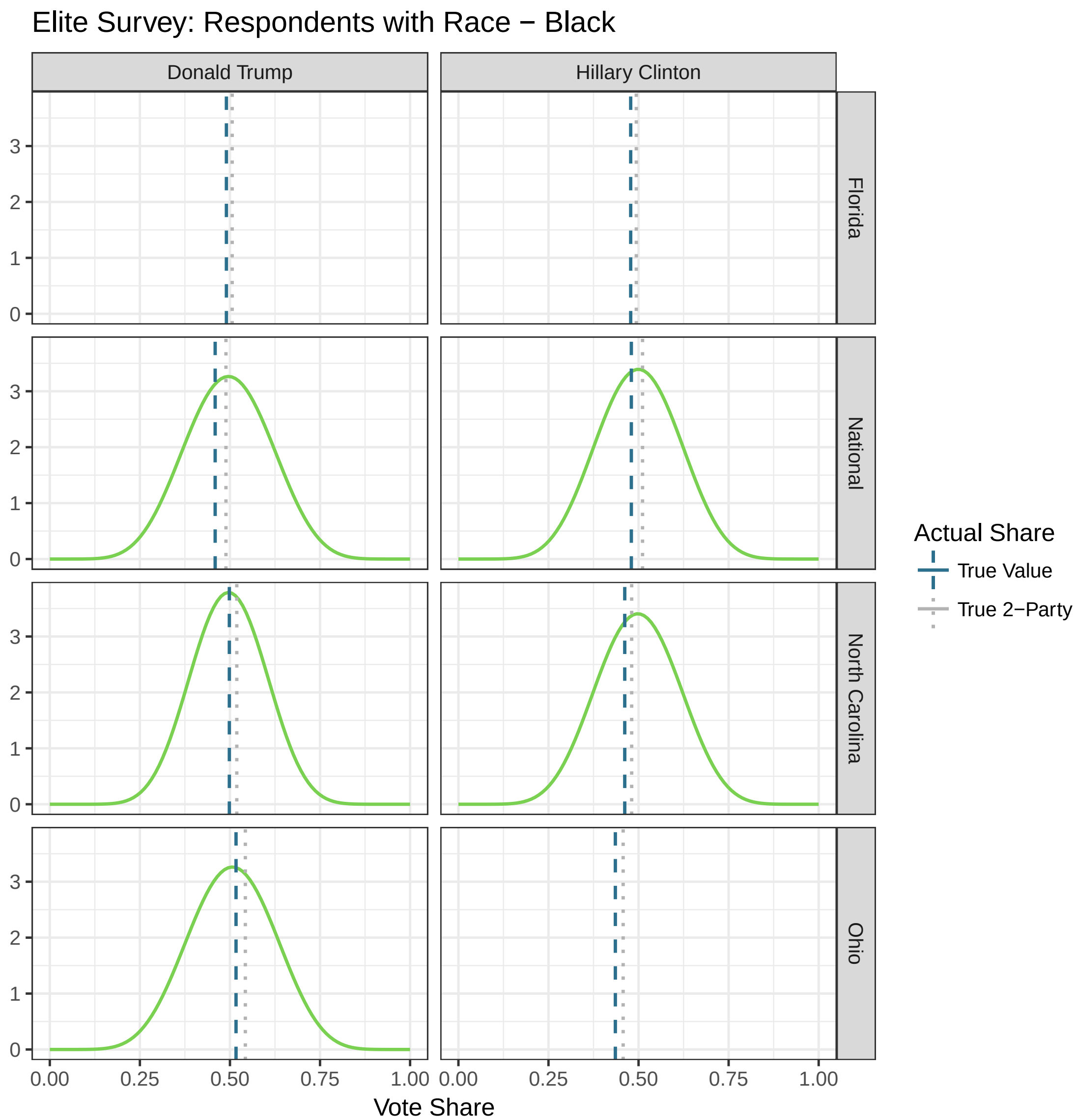


Figure 3.84: Priors without covariates: Elite Race Black

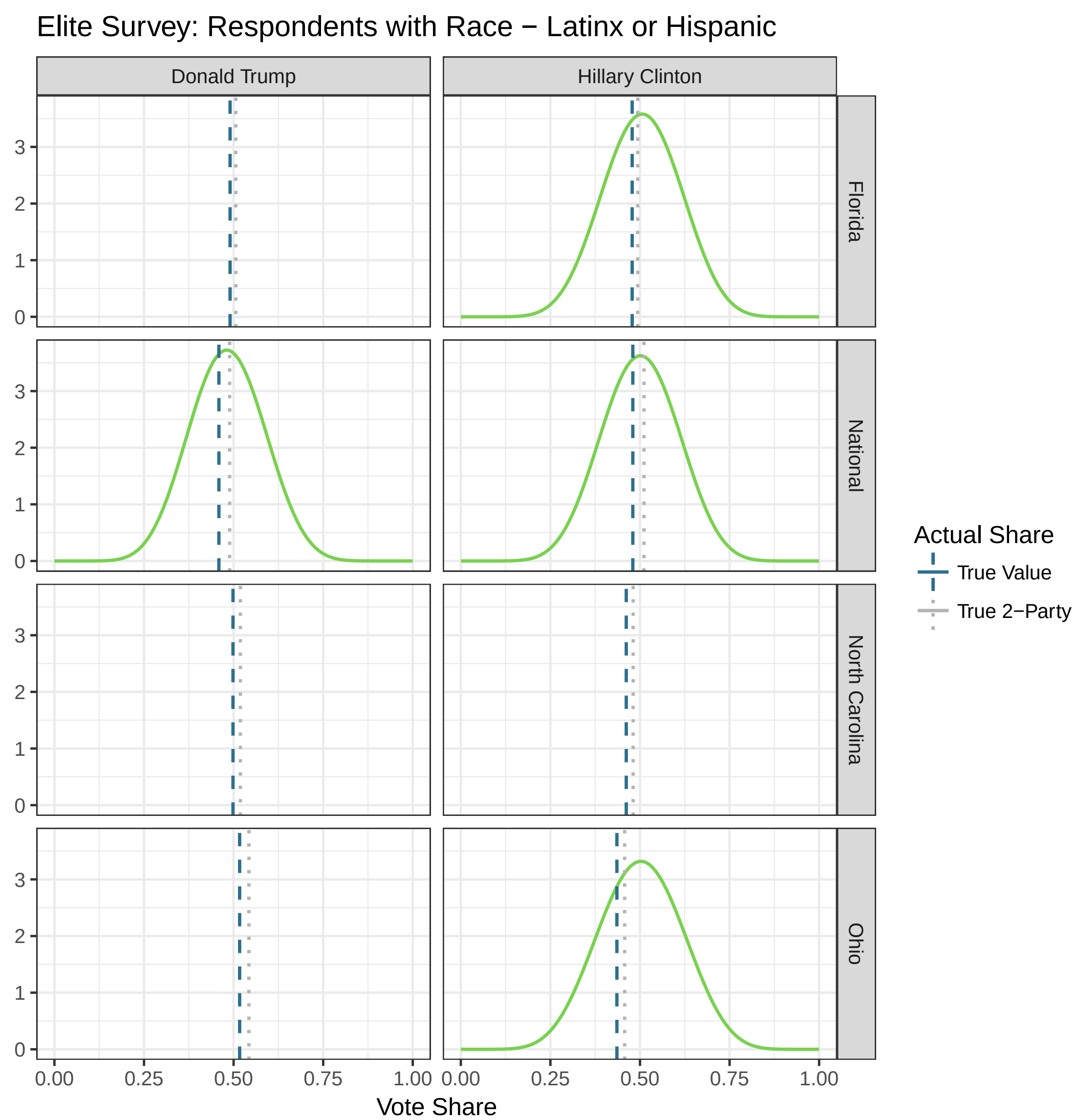


Figure 3.85: Priors without covariates: Elite Race Latinx or Hispanic

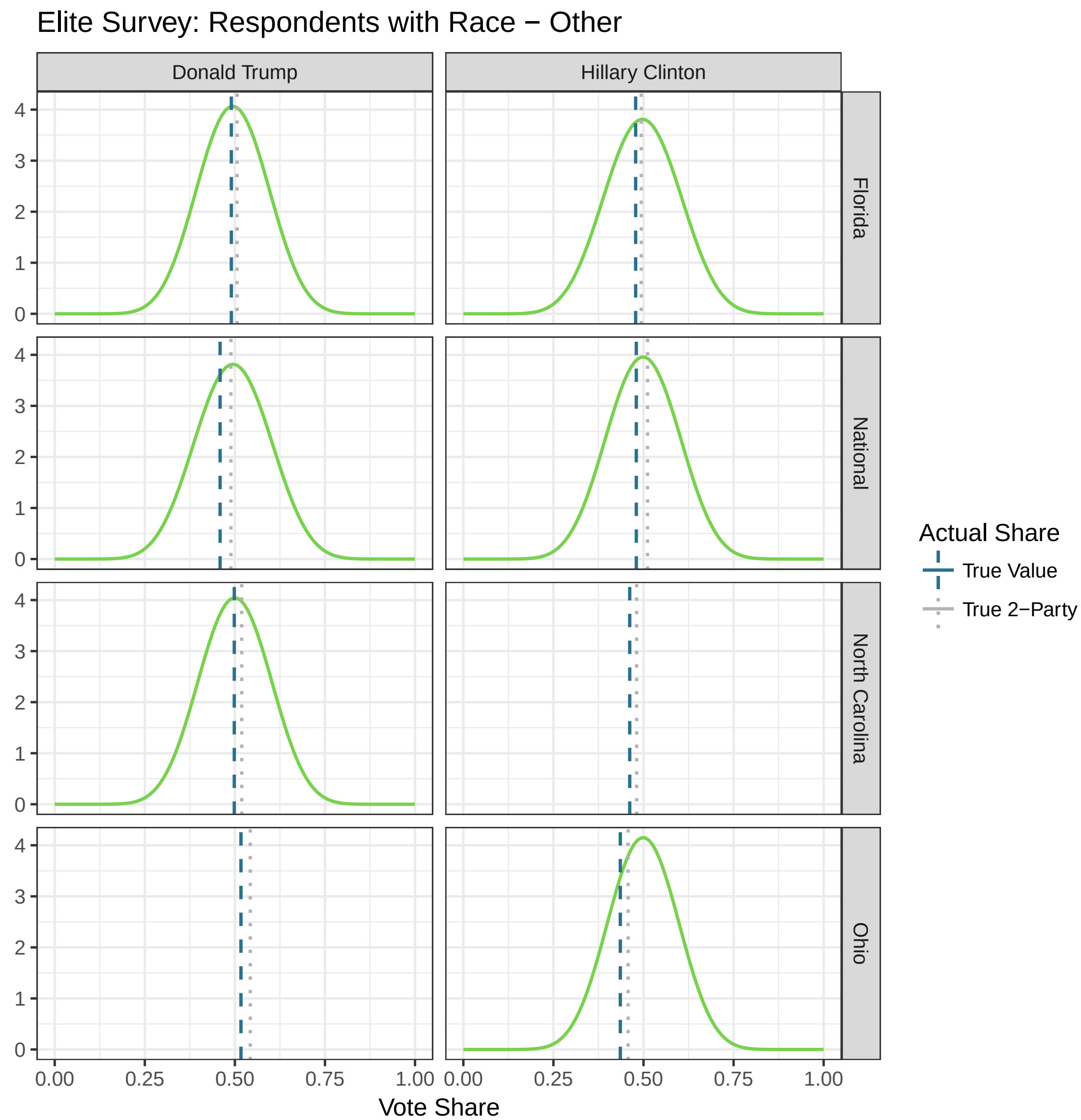


Figure 3.86: Priors without covariates: Elite Race Other

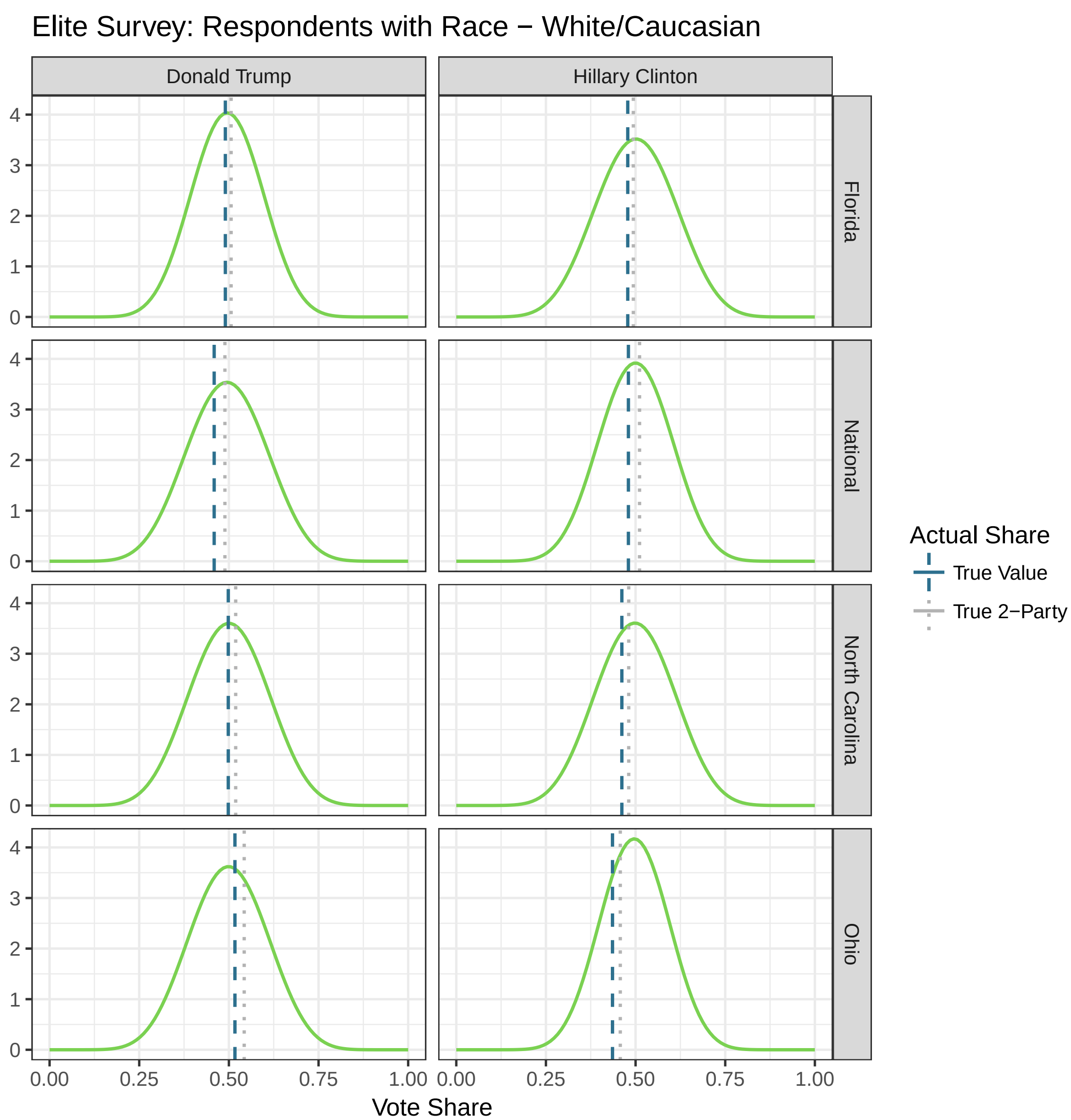


Figure 3.87: Priors without covariates: Elite Race White Caucasian

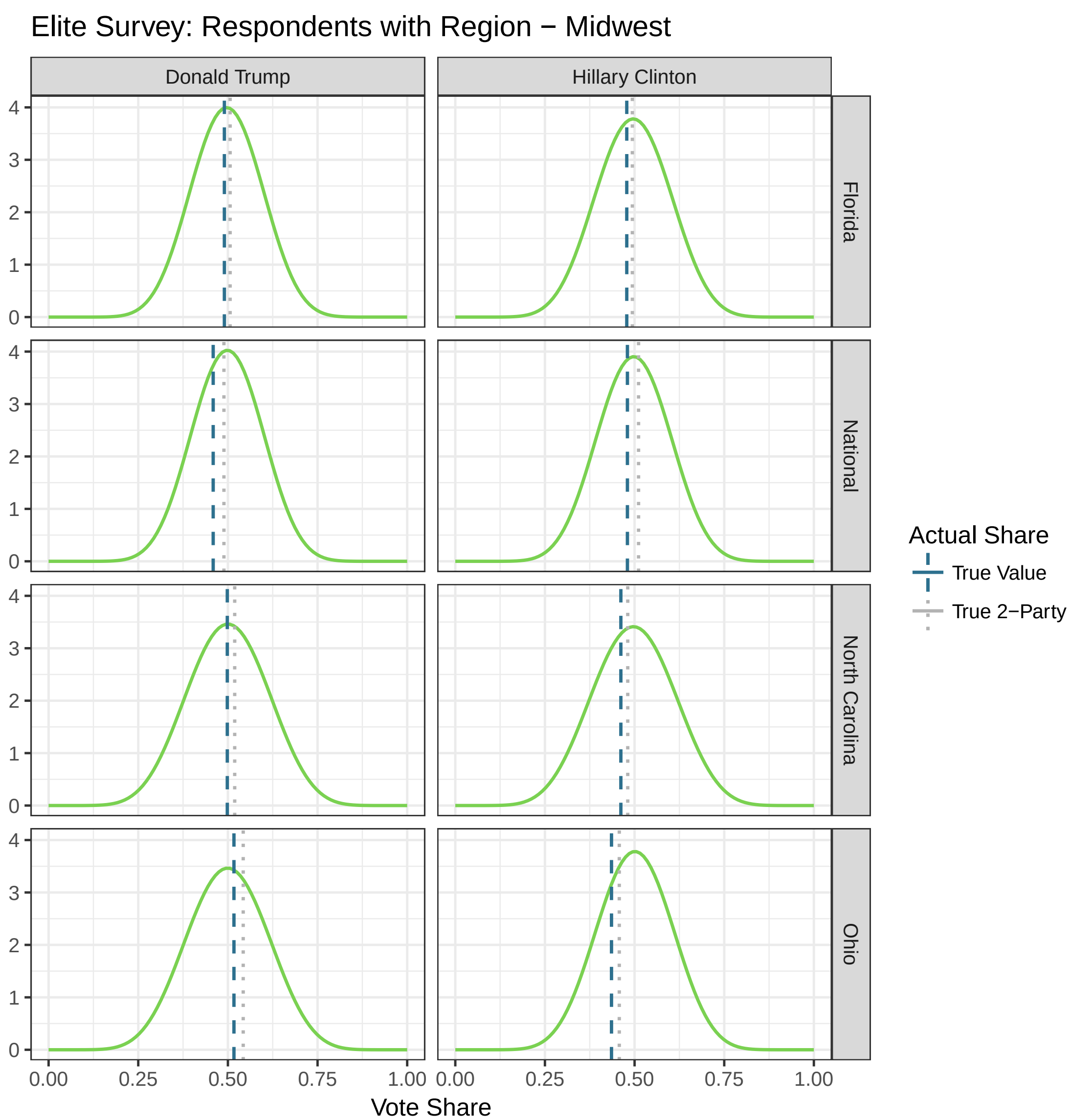


Figure 3.88: Priors without covariates: Elite Region Midwest

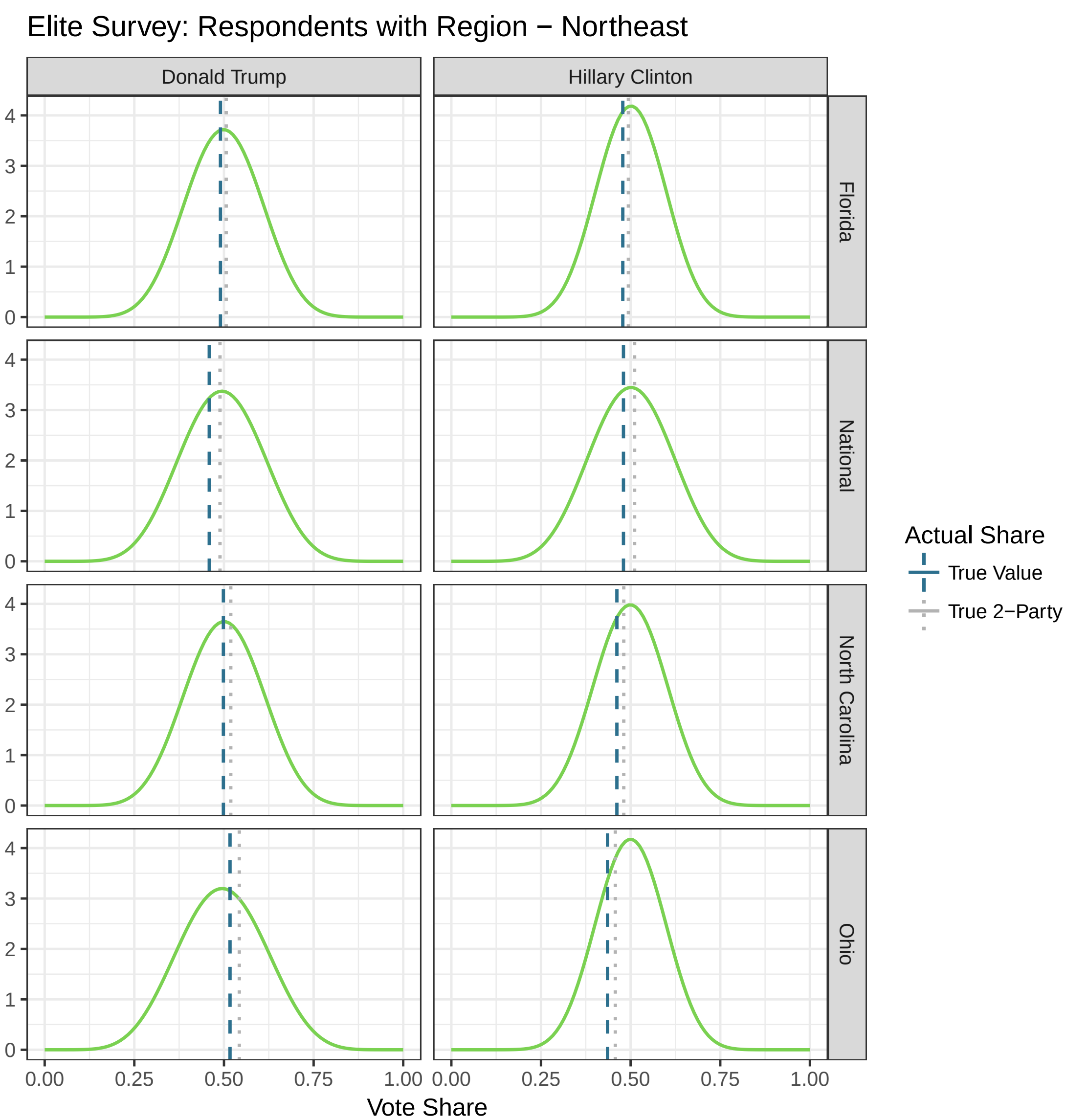


Figure 3.89: Priors without covariates: Elite Region Northeast

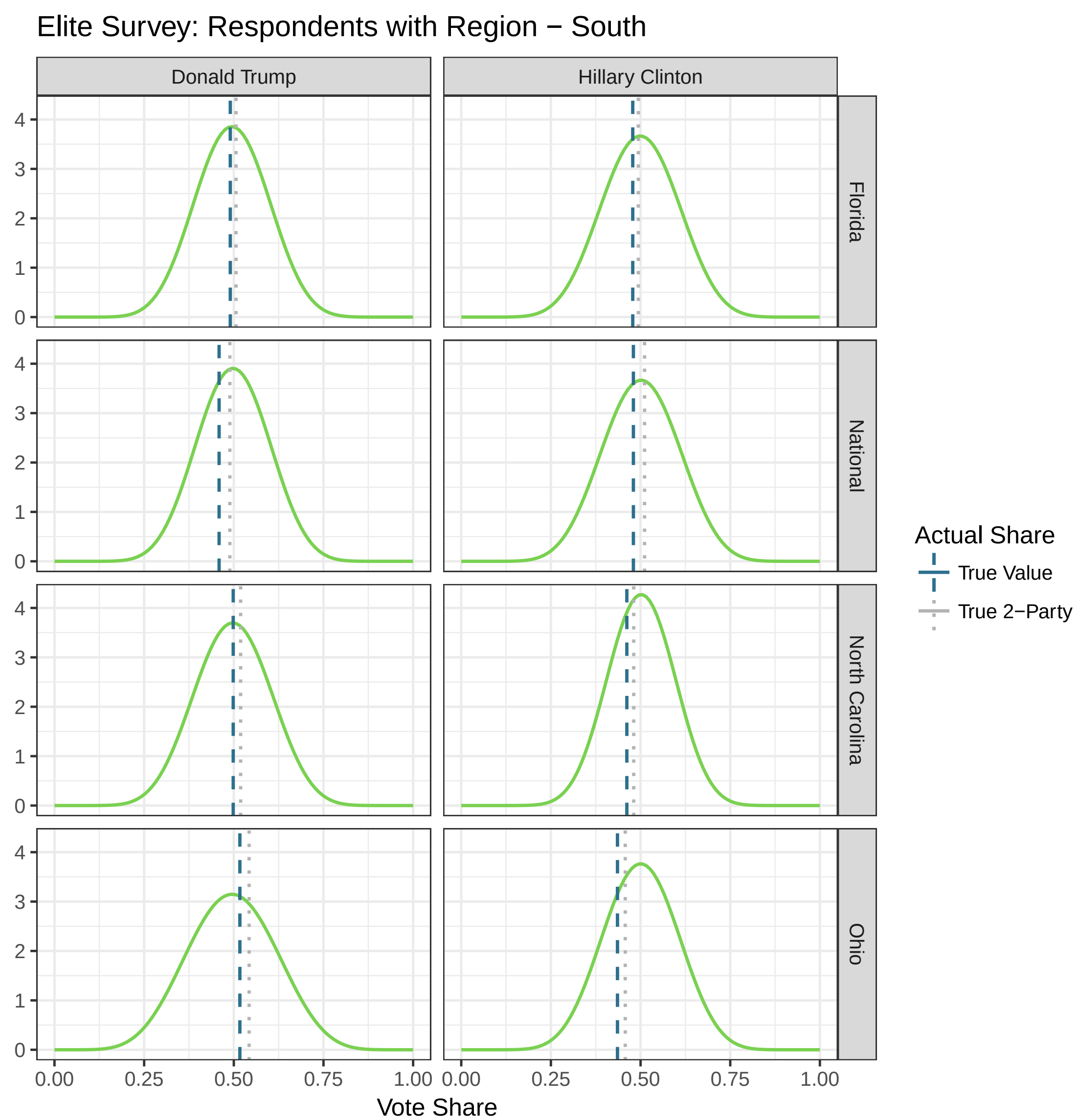


Figure 3.90: Priors without covariates: Elite Region South

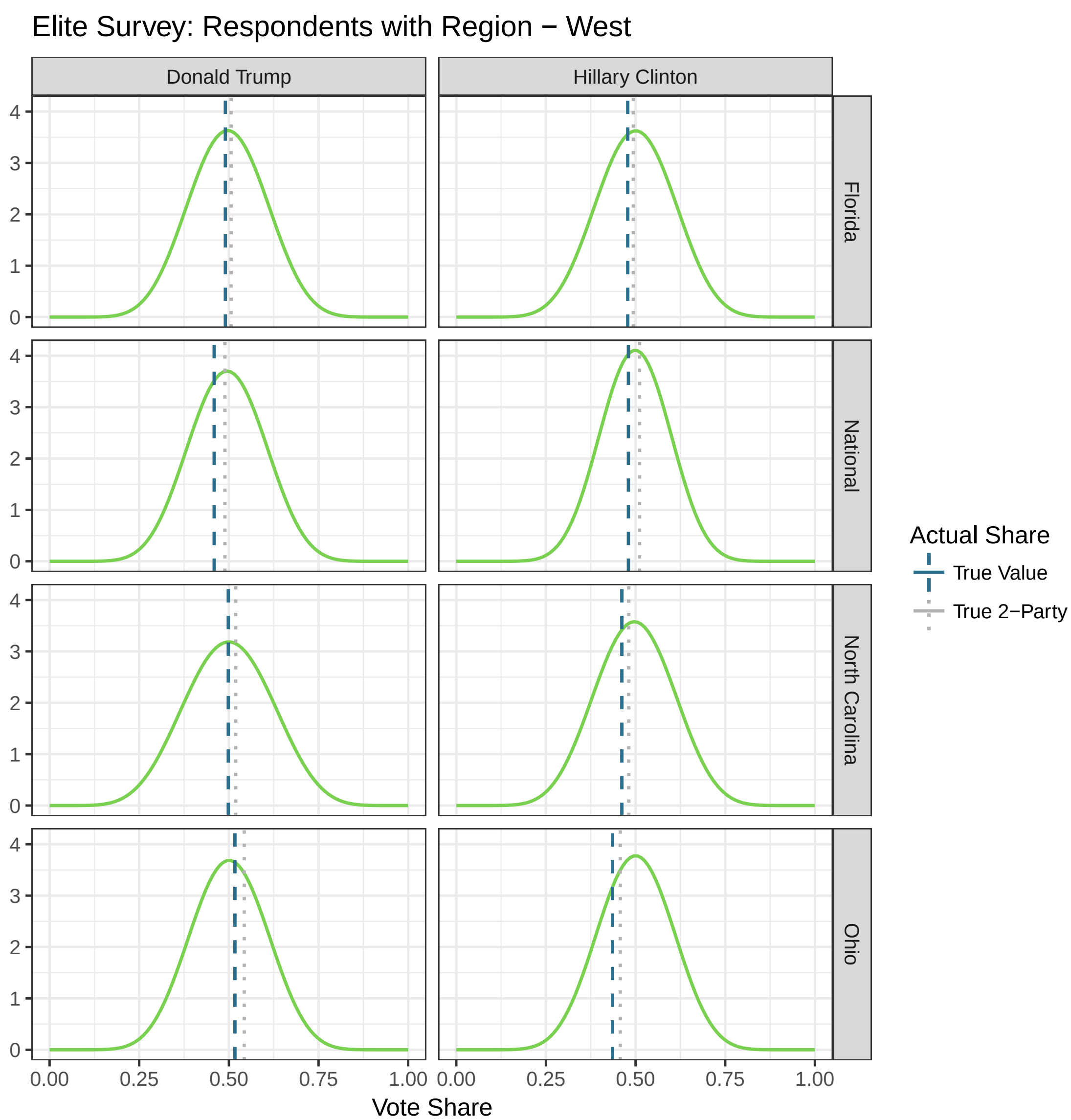


Figure 3.91: Priors without covariates: Elite Region West

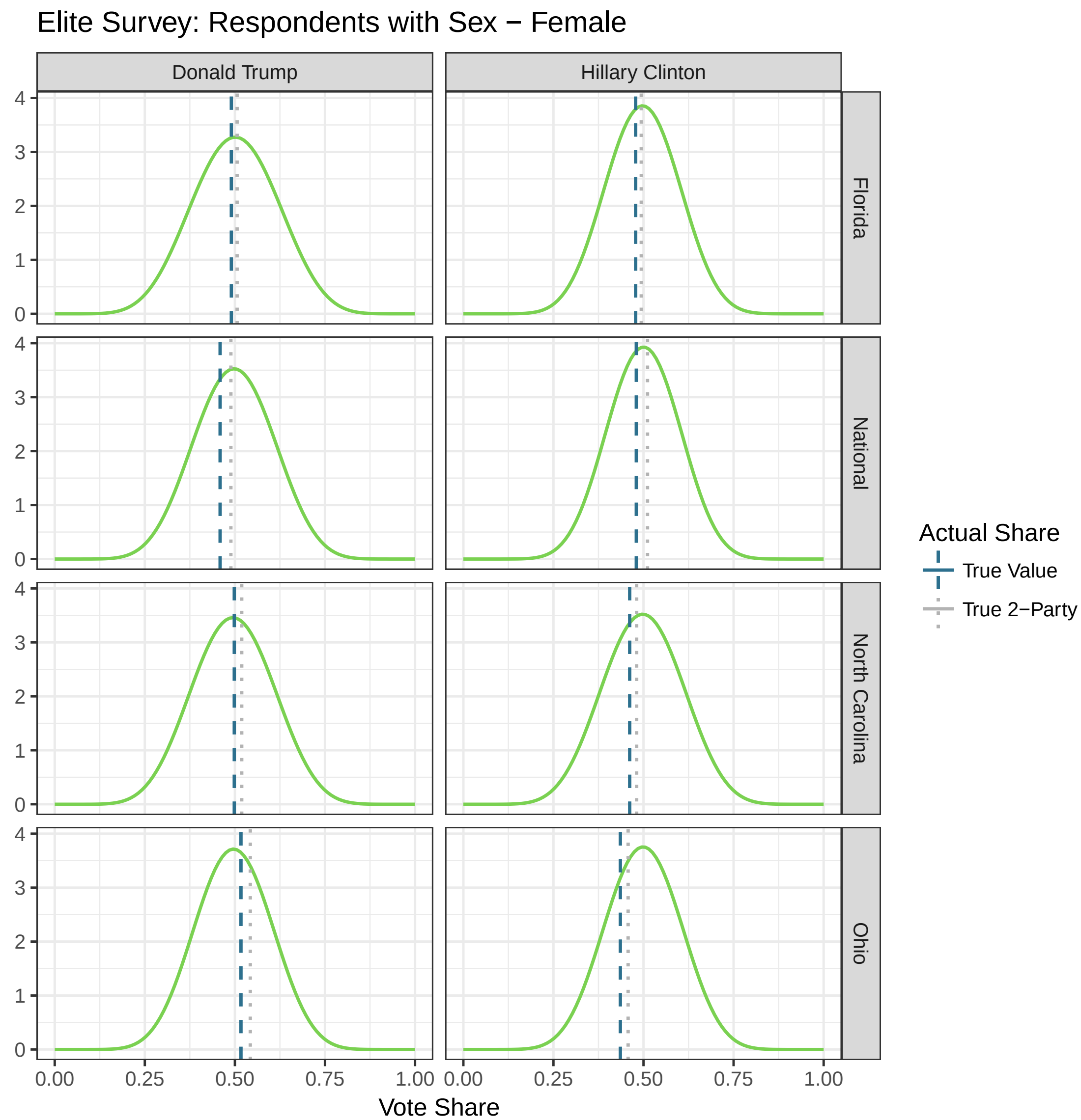


Figure 3.92: Priors without covariates: Elite Sex Female

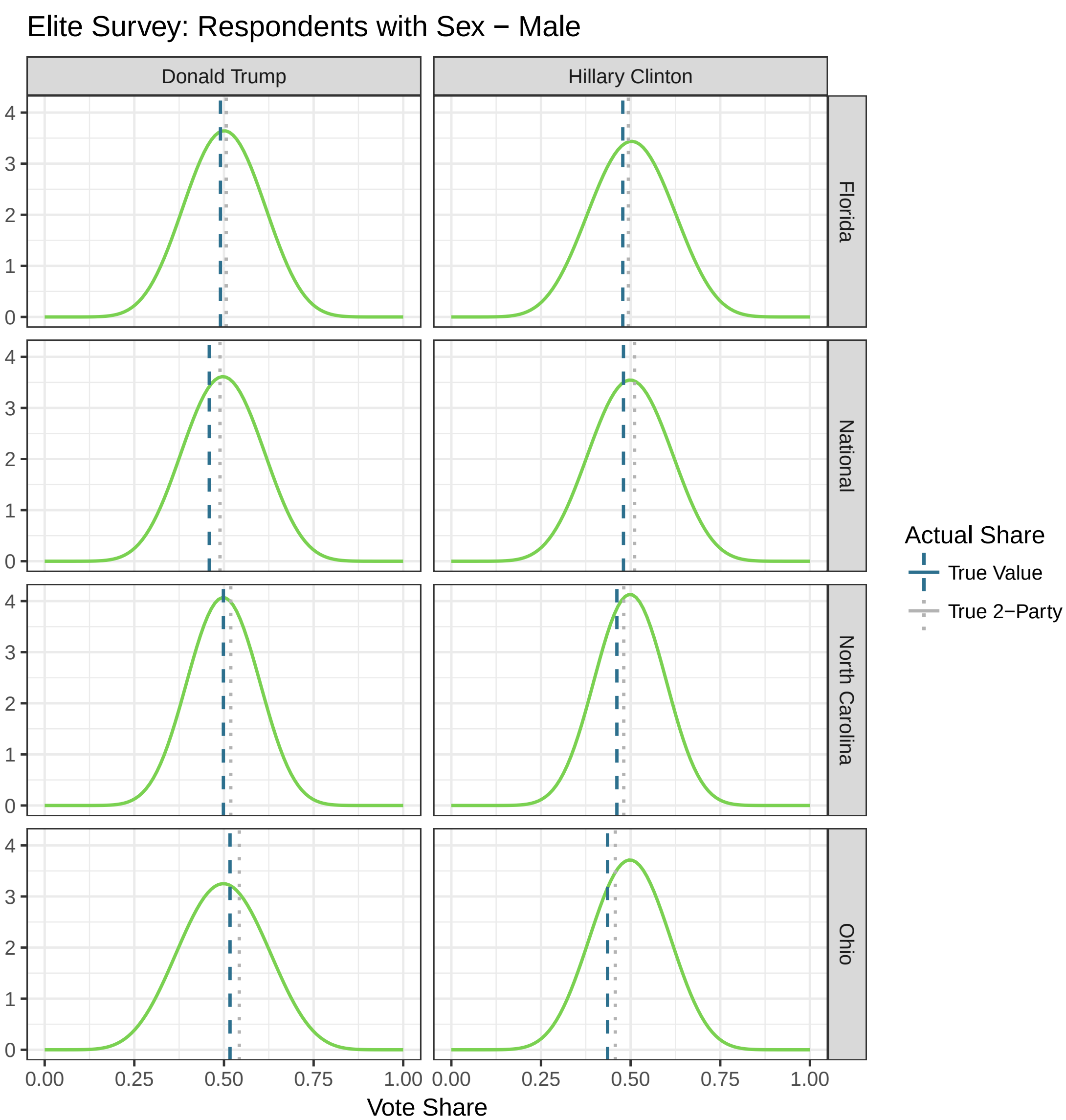


Figure 3.93: Priors without covariates: Elite Sex Male

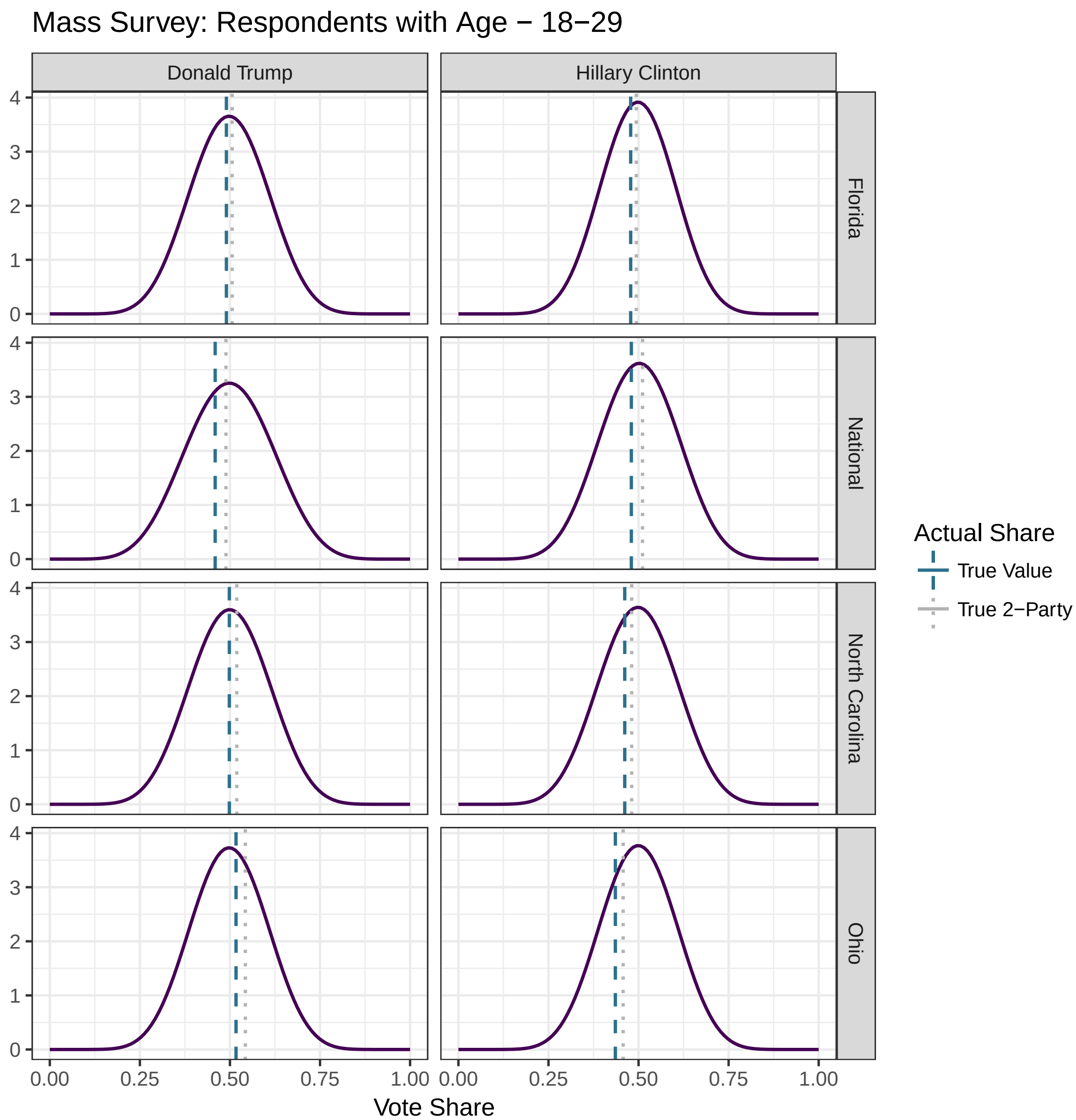


Figure 3.94: Priors without covariates: Mass Age 18-29

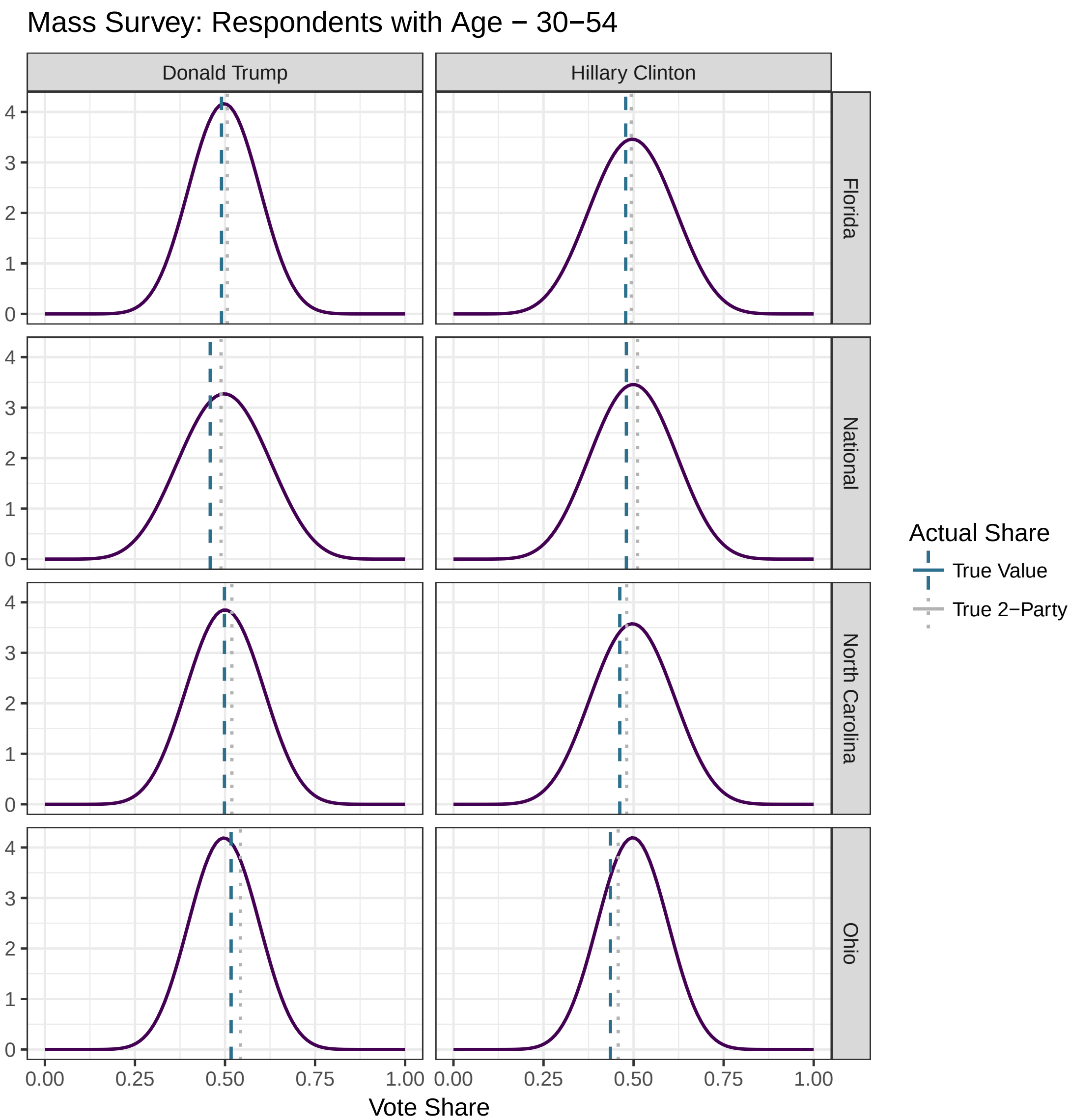


Figure 3.95: Priors without covariates: Mass Age 30-54

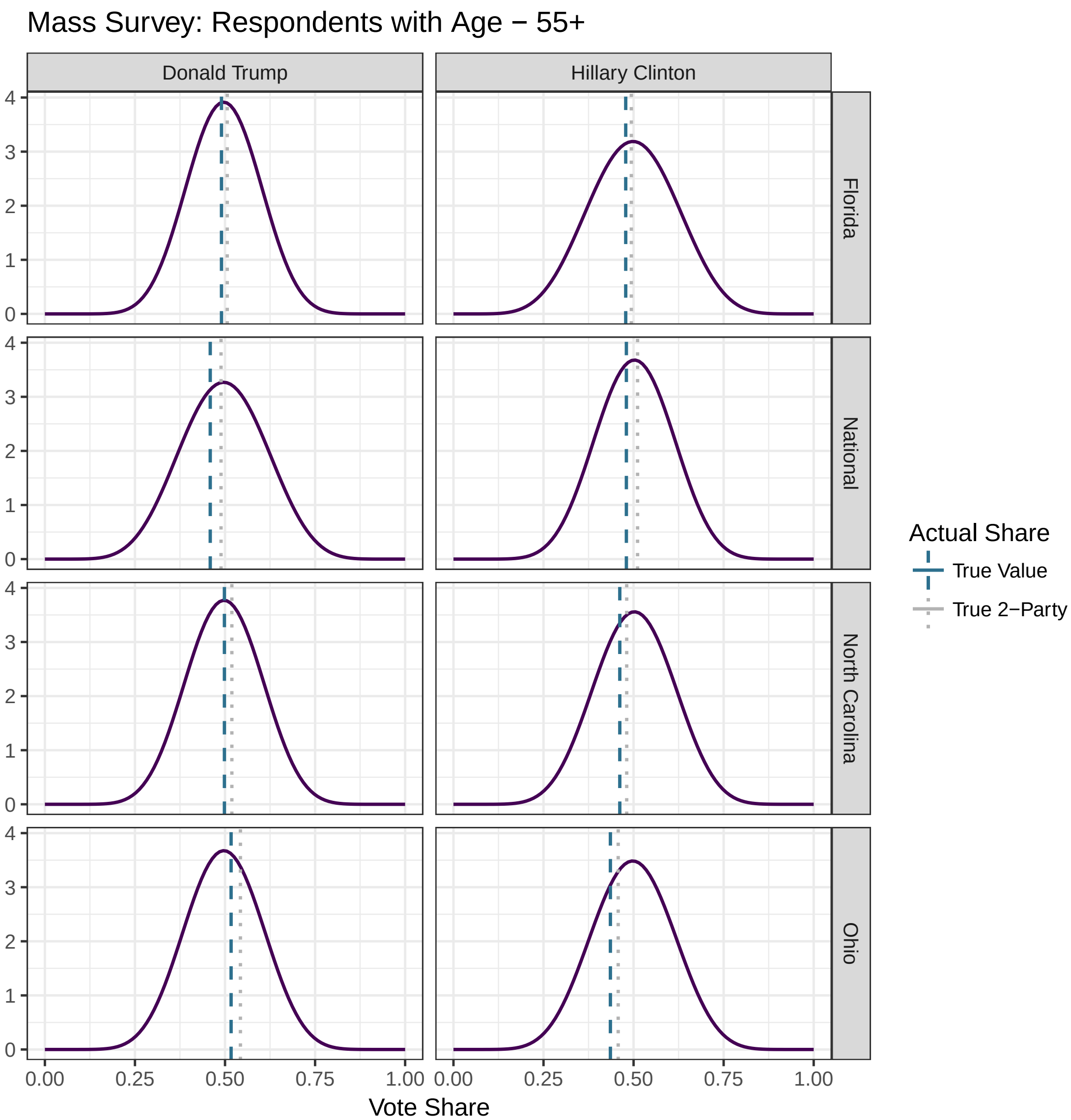


Figure 3.96: Priors without covariates: Mass Age 55+

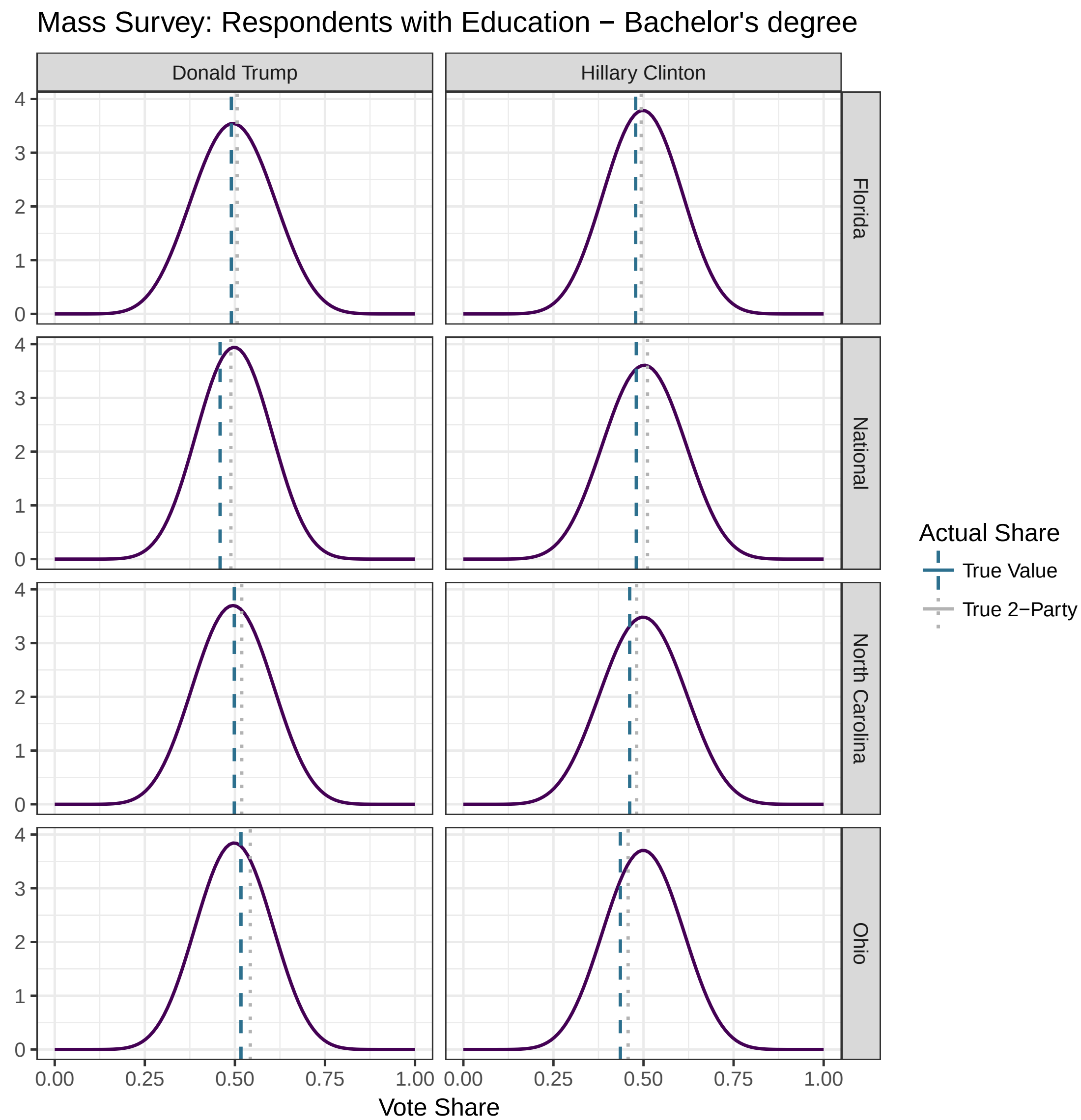


Figure 3.97: Priors without covariates: Mass Education Bachelor's degree

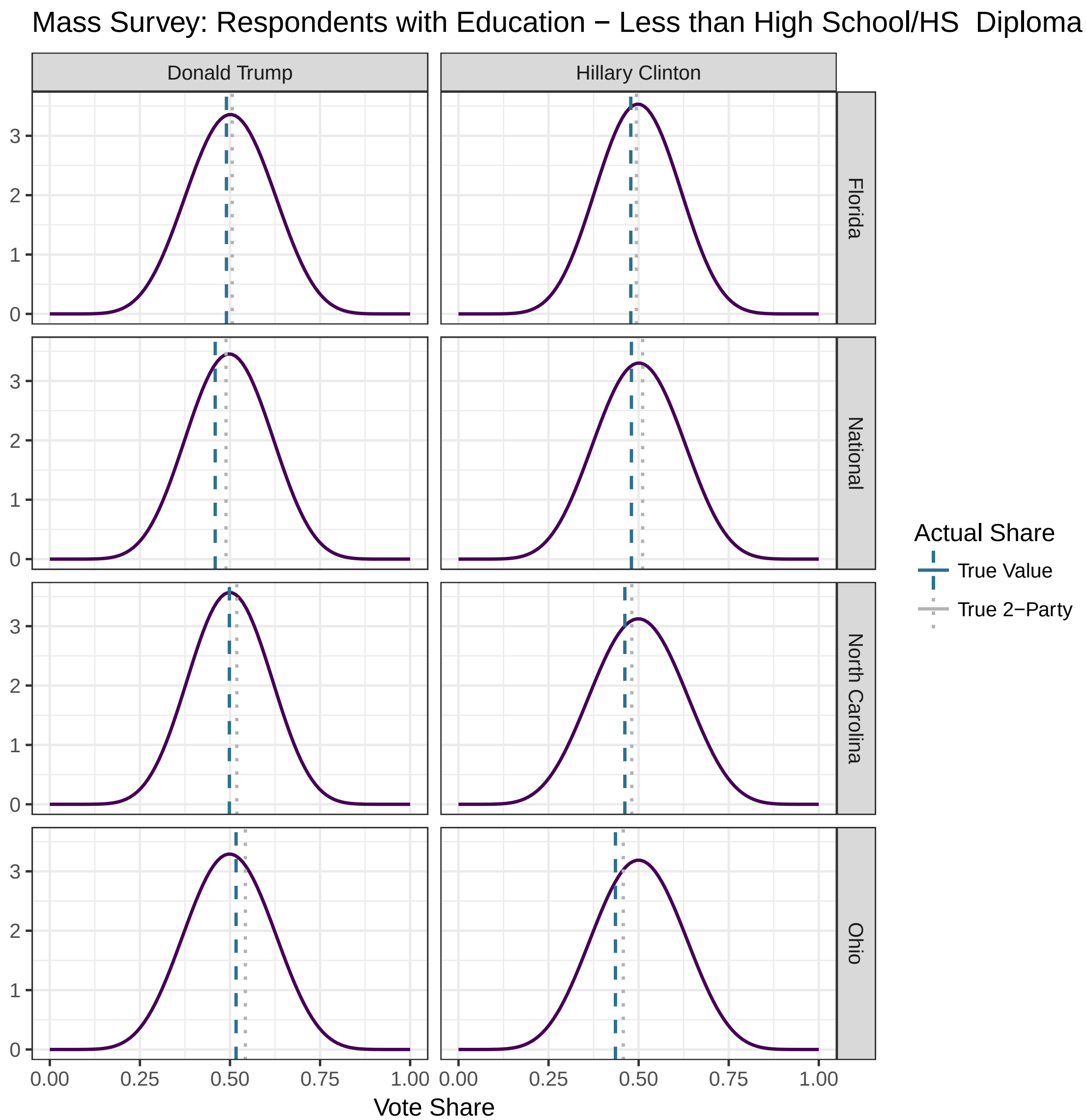


Figure 3.98: Priors without covariates: Mass Education Less than High School HS Diploma

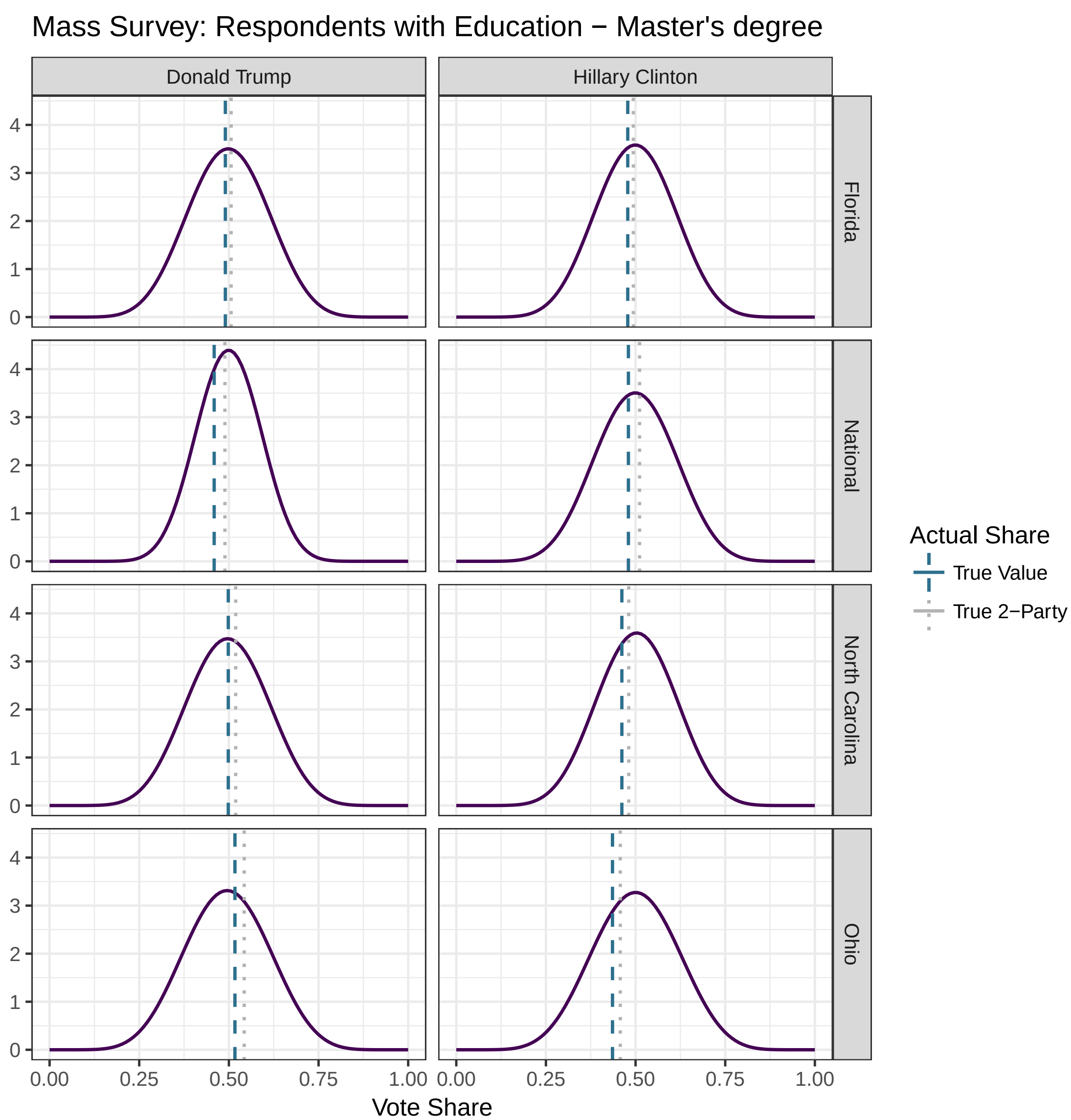


Figure 3.99: Priors without covariates: Mass Education Master's degree

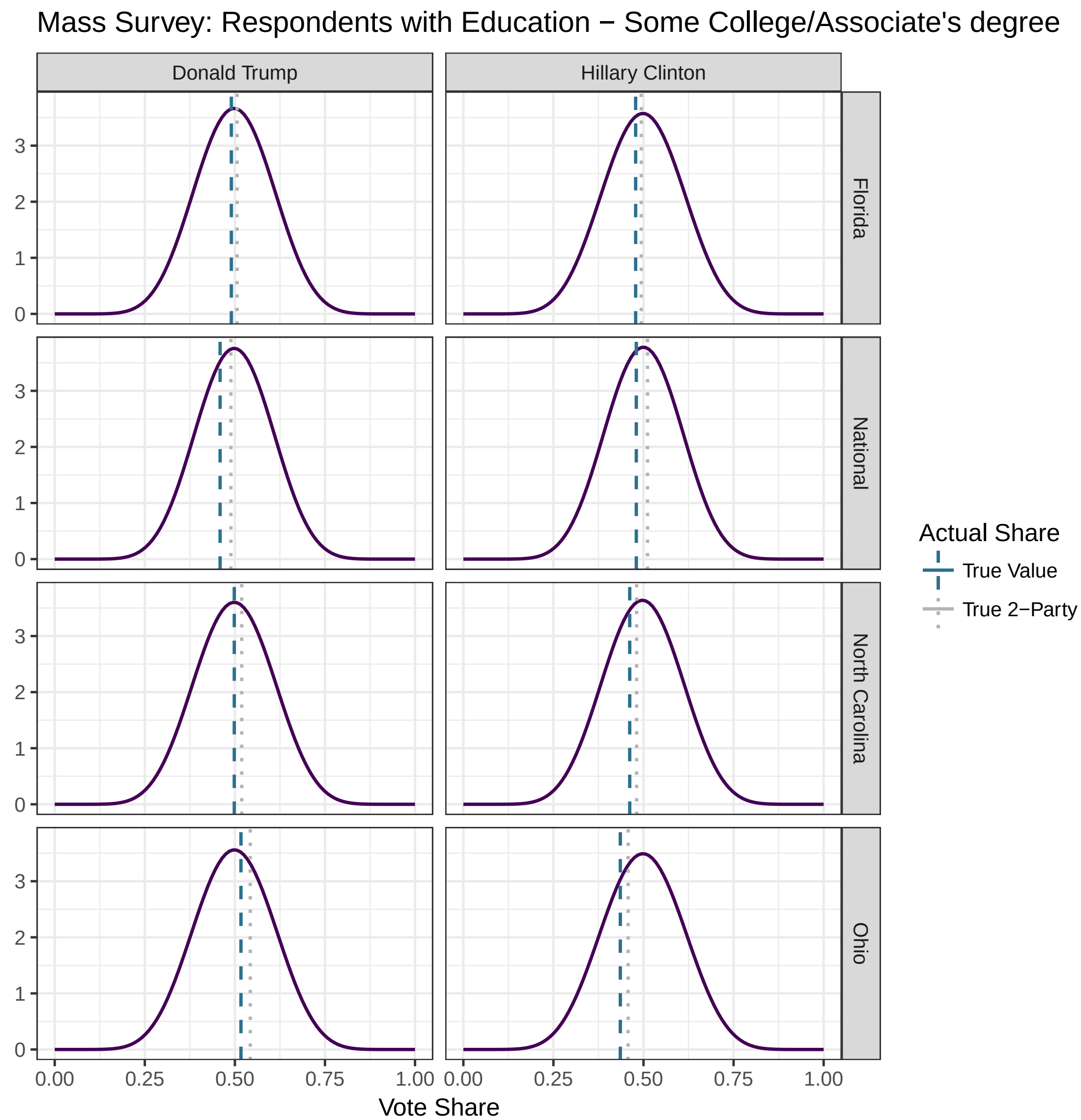


Figure 3.100: Priors without covariates: Mass Education Some College Associate's degree

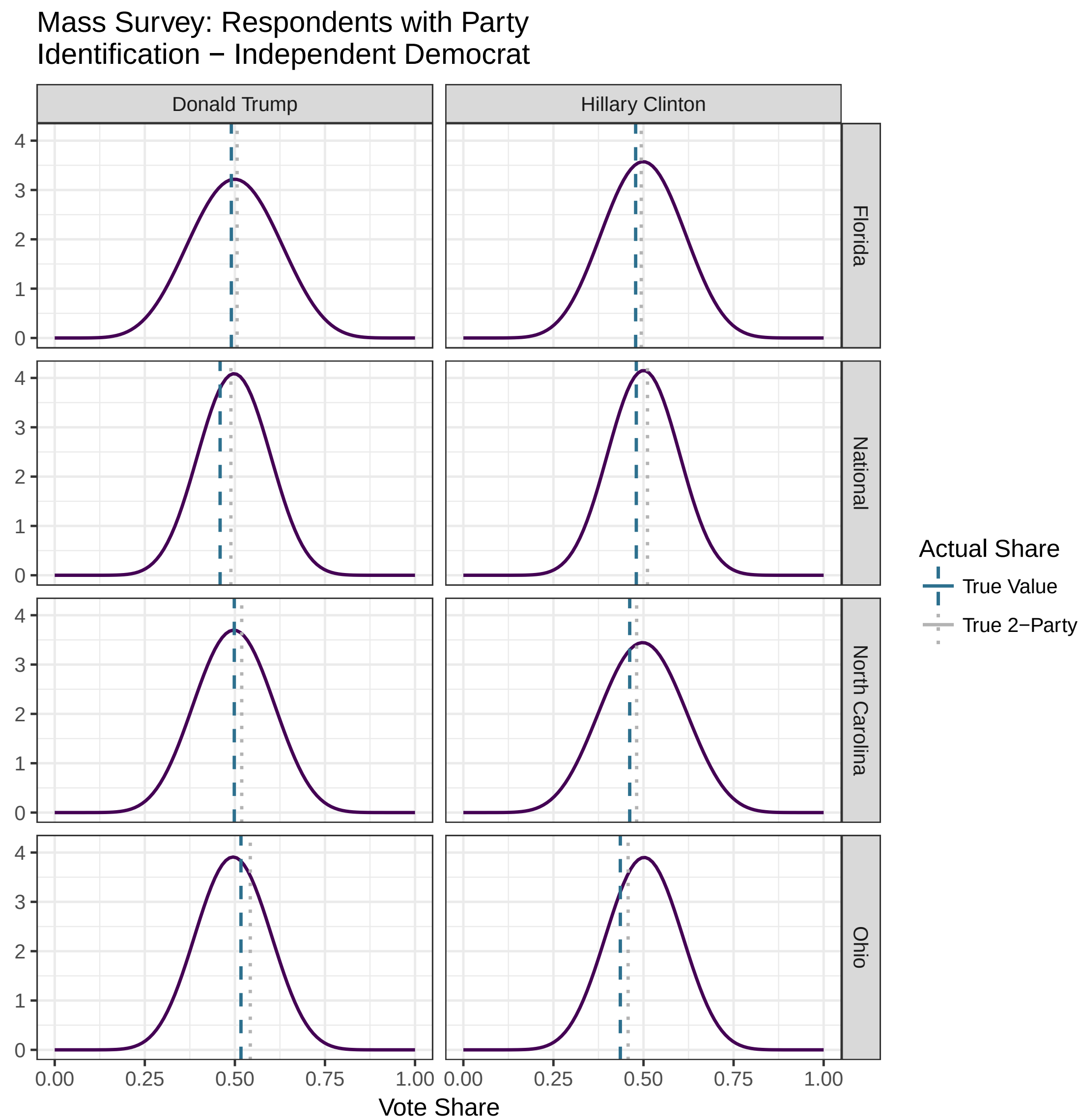


Figure 3.101: Priors without covariates: Mass Party Identification Independent Democrat

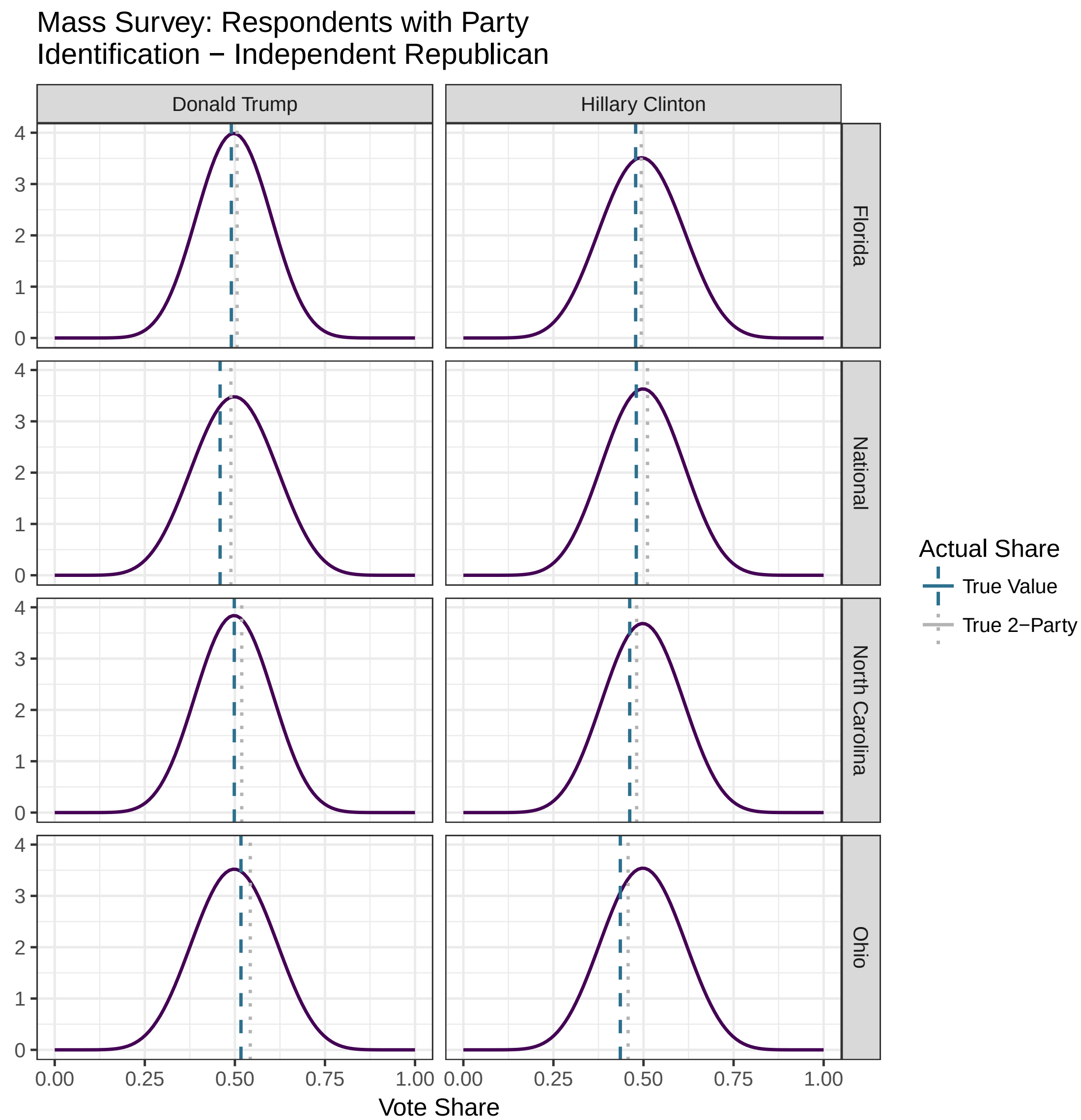


Figure 3.102: Priors without covariates: Mass Party Identification Independent Republican

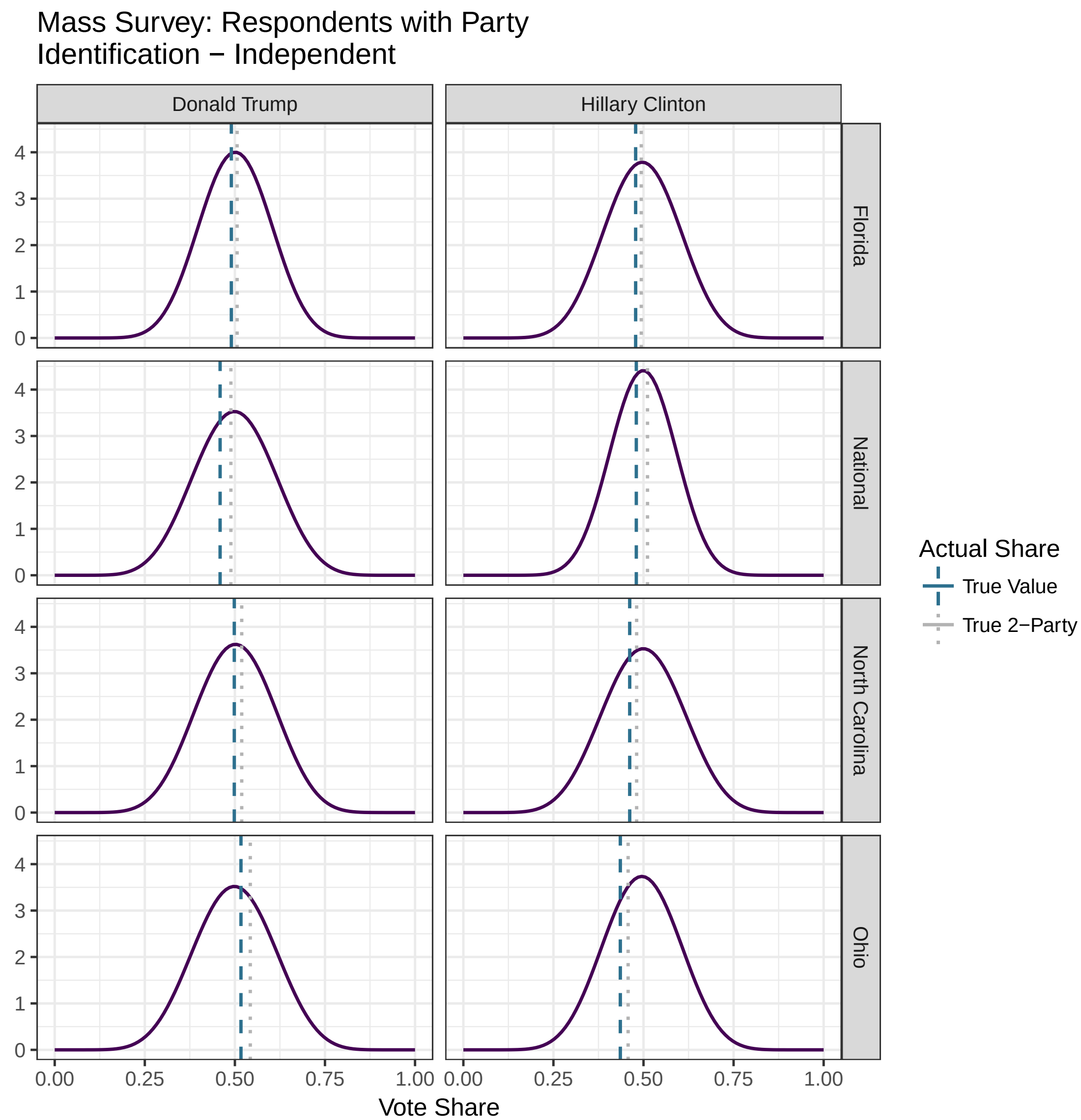


Figure 3.103: Priors without covariates: Mass Party Identification Independent

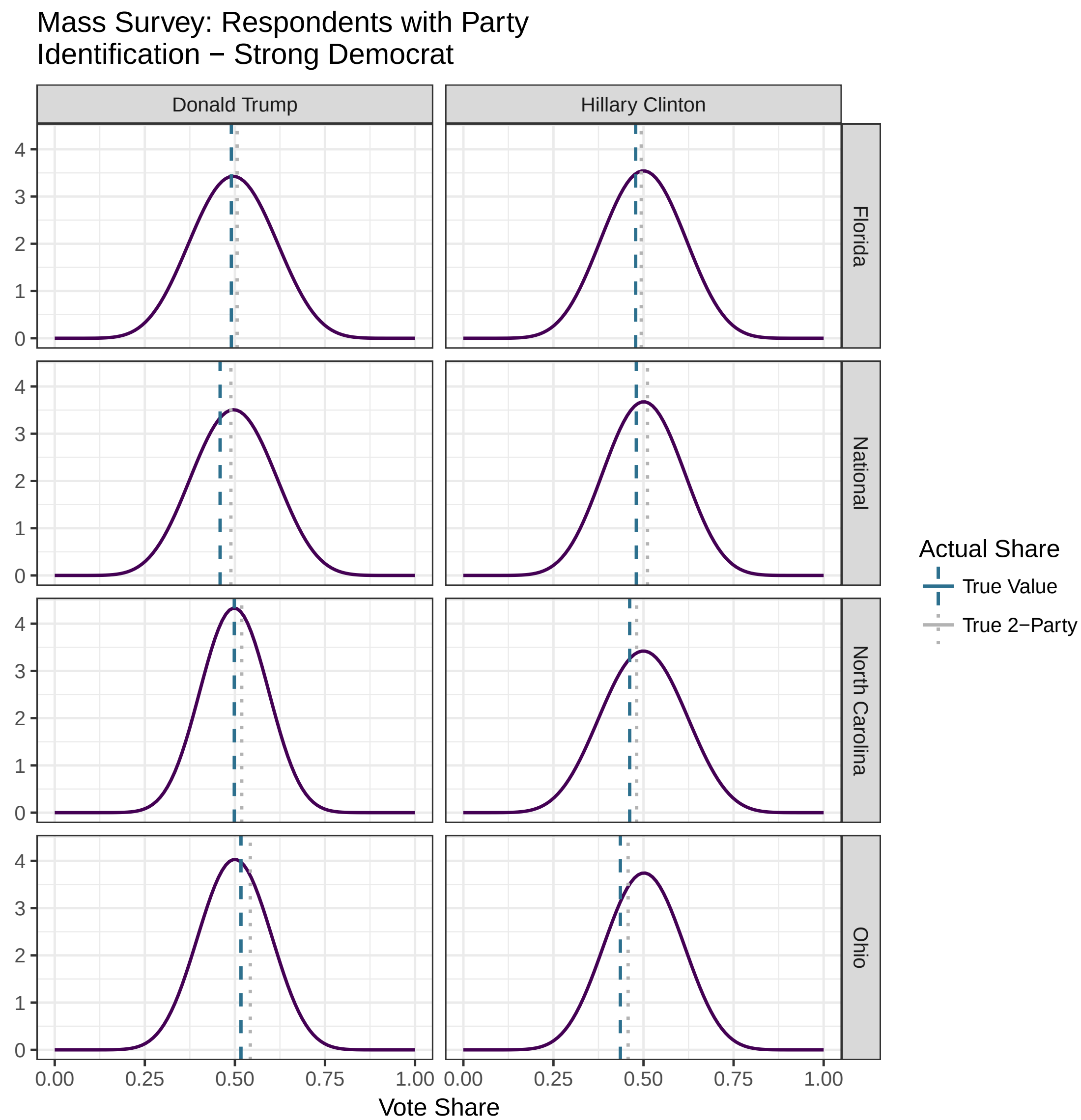


Figure 3.104: Priors without covariates: Mass Party Identification Strong Democrat

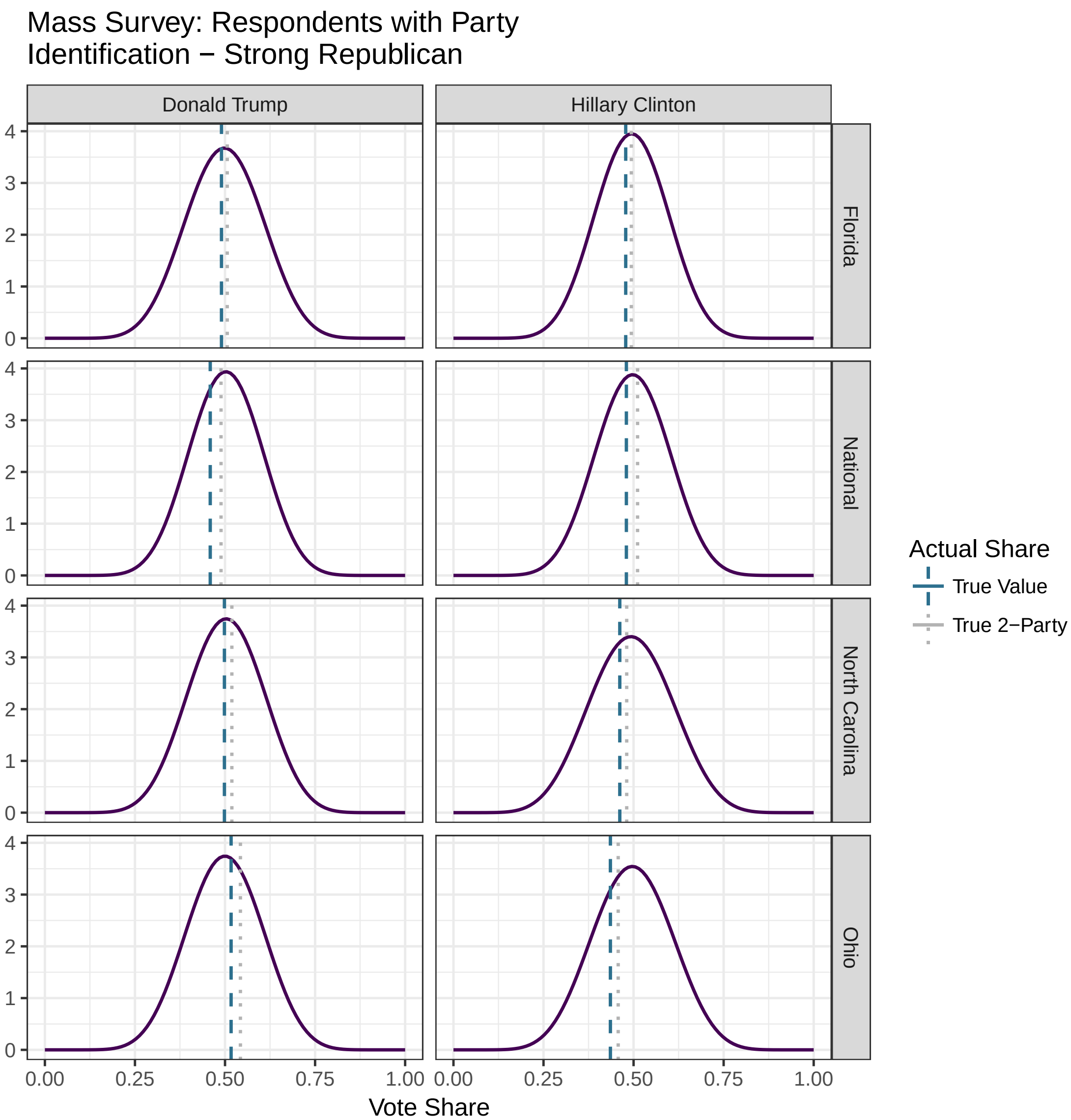


Figure 3.105: Priors without covariates: Mass Party Identification Strong Republican

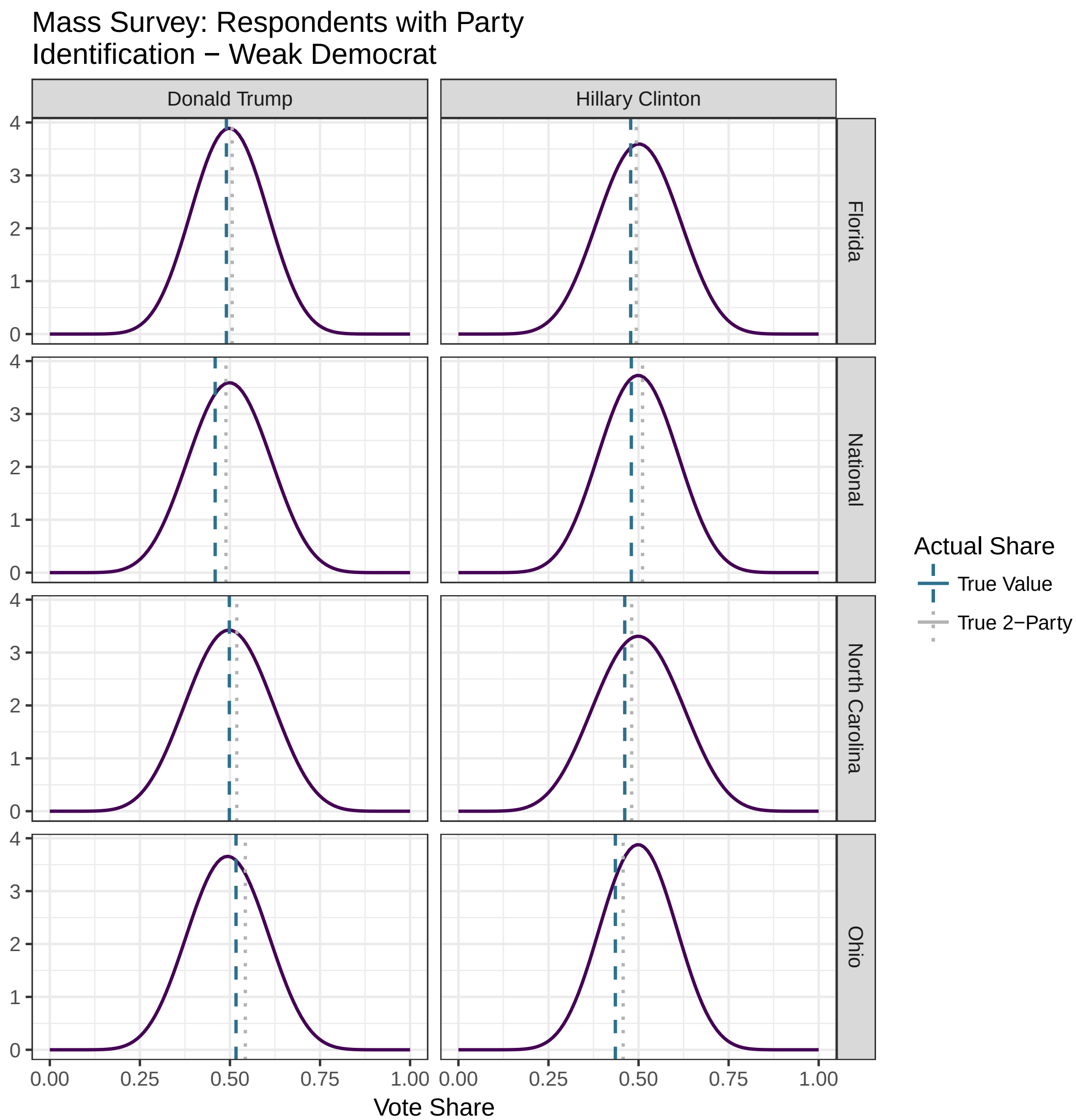


Figure 3.106: Priors without covariates: Mass Party Identification Weak Democrat

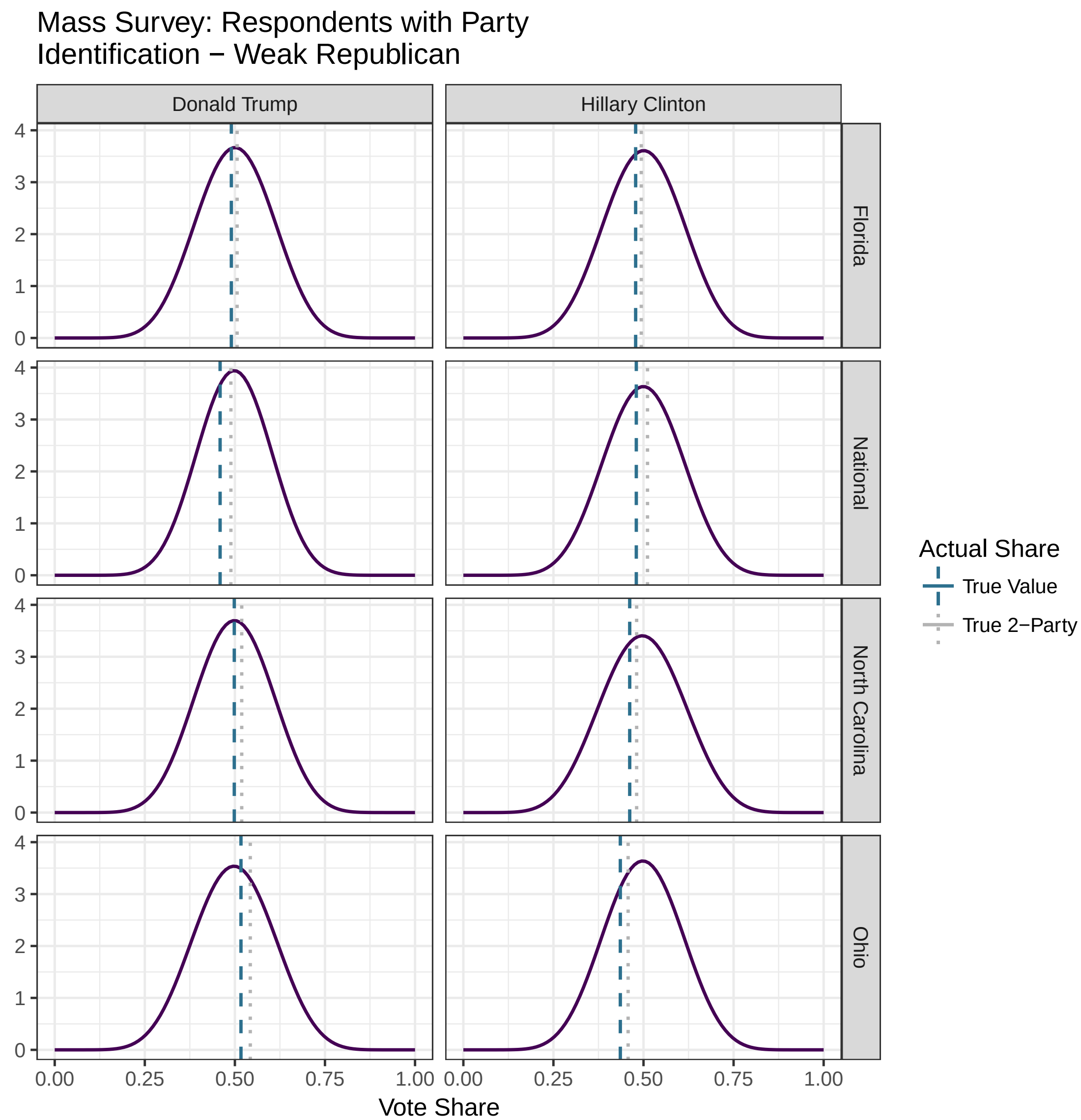


Figure 3.107: Priors without covariates: Mass Party Identification Weak Republican

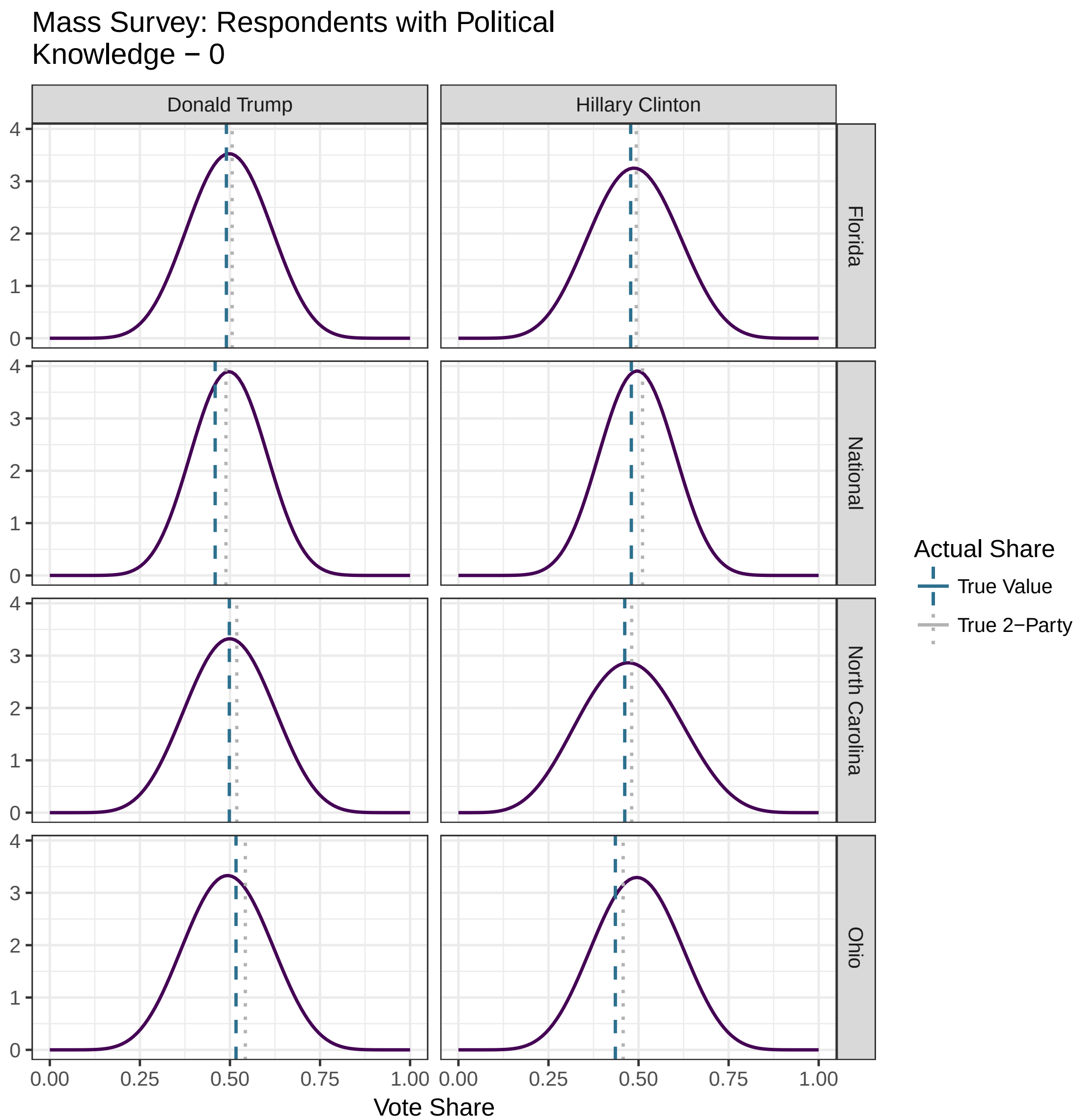


Figure 3.108: Priors without covariates: Mass Political Knowledge 0

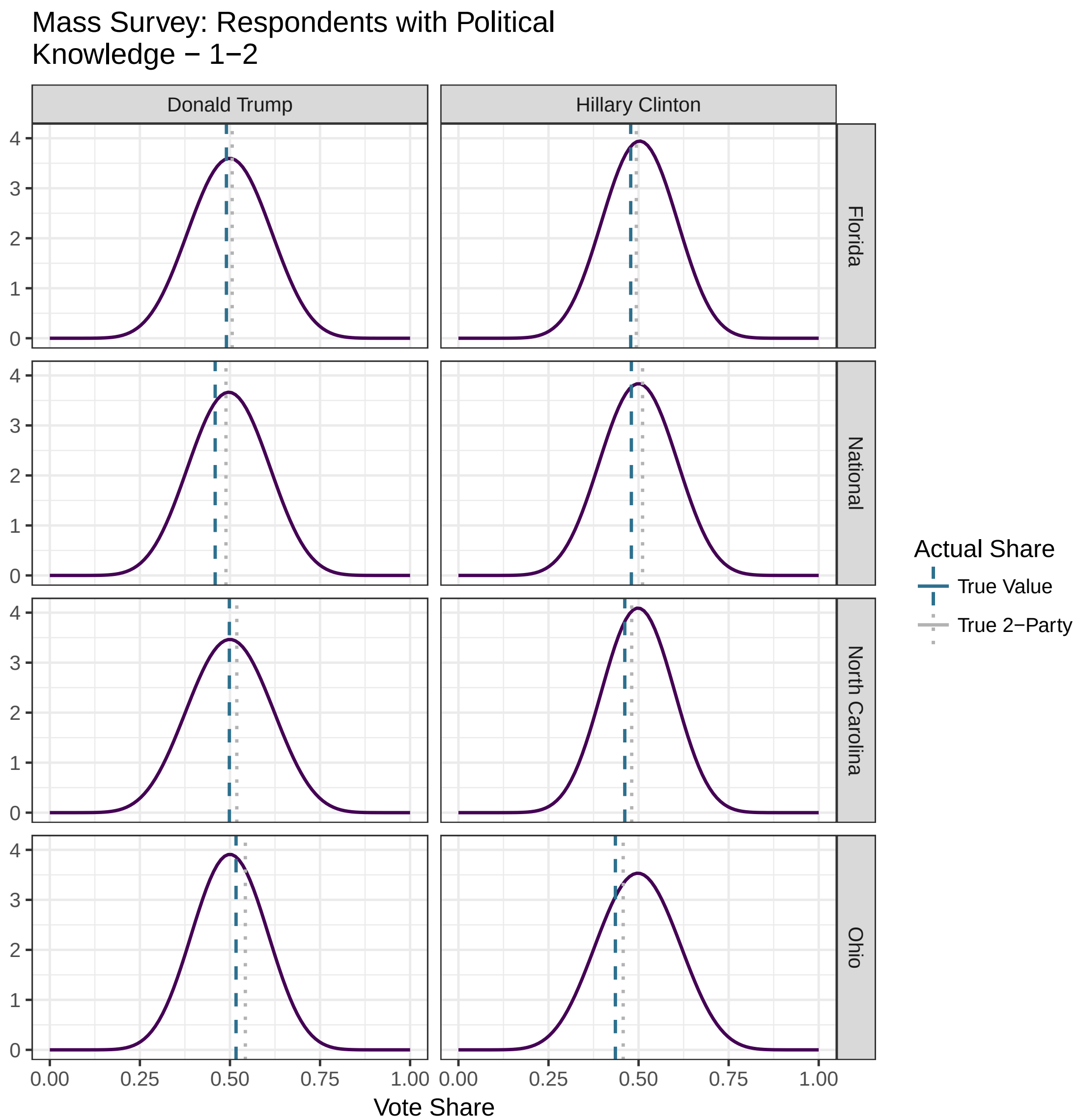


Figure 3.109: Priors without covariates: Mass Political Knowledge 1-2

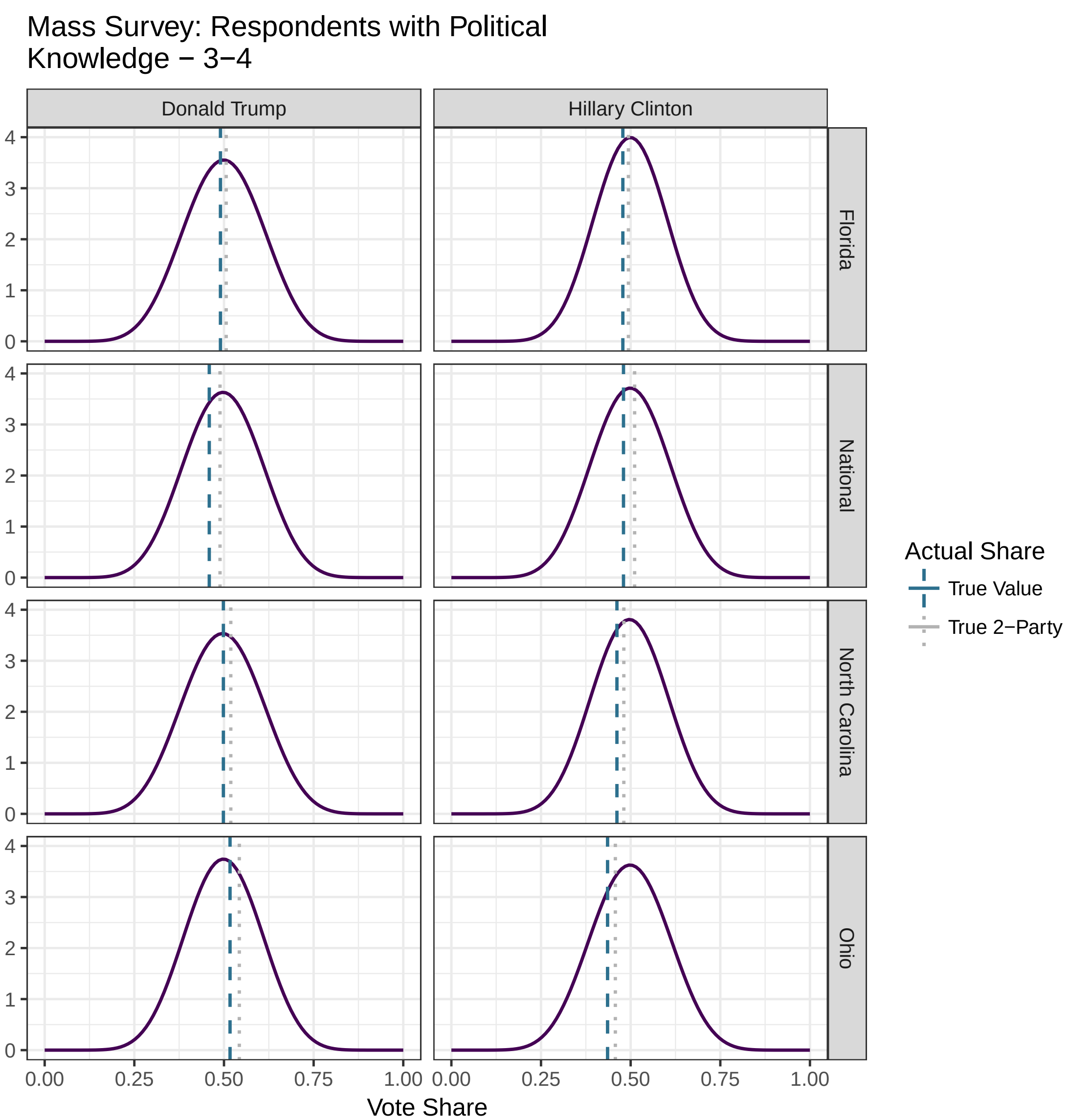


Figure 3.110: Priors without covariates: Mass Political Knowledge 3-4

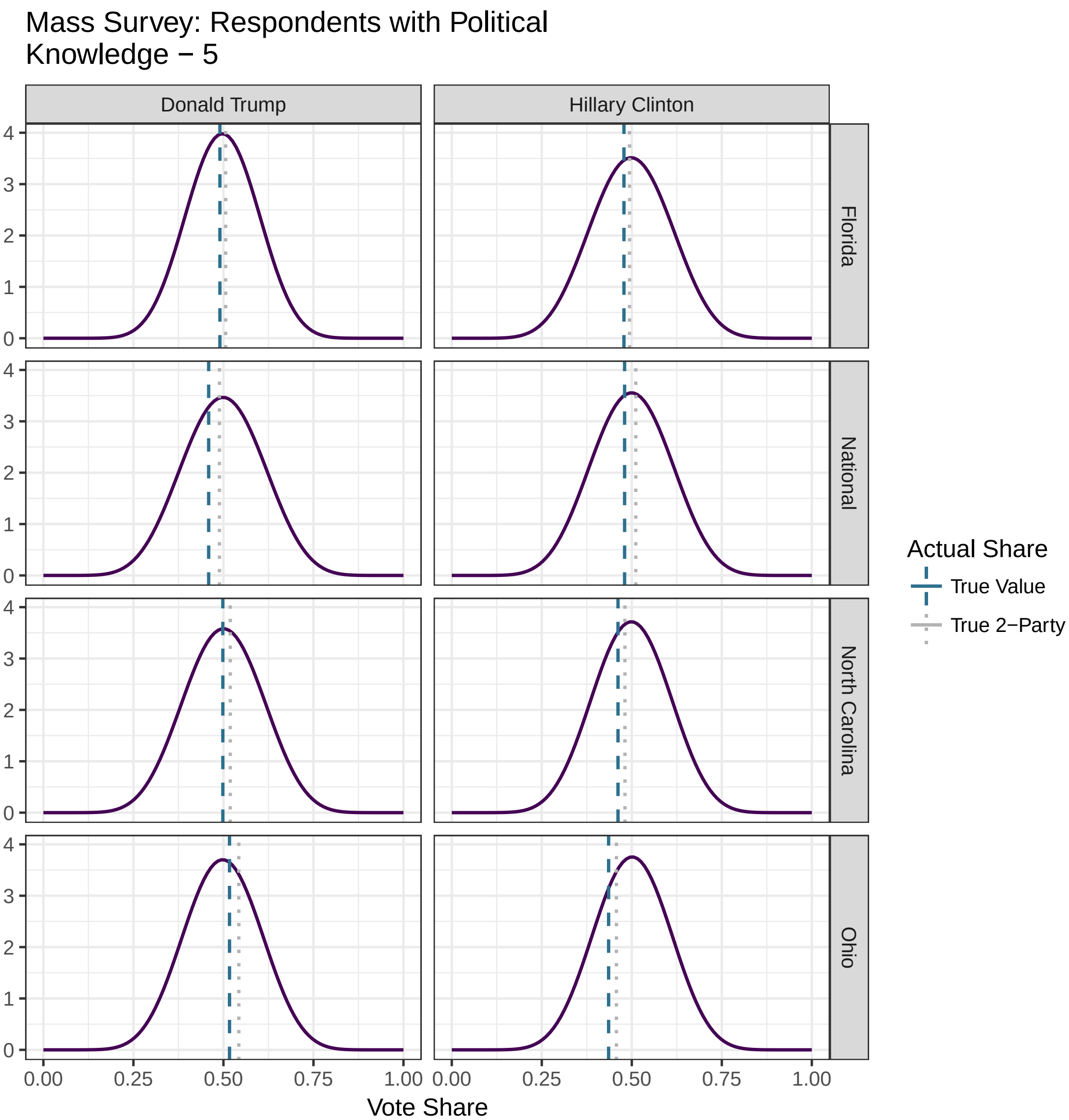


Figure 3.111: Priors without covariates: Mass Political Knowledge 5

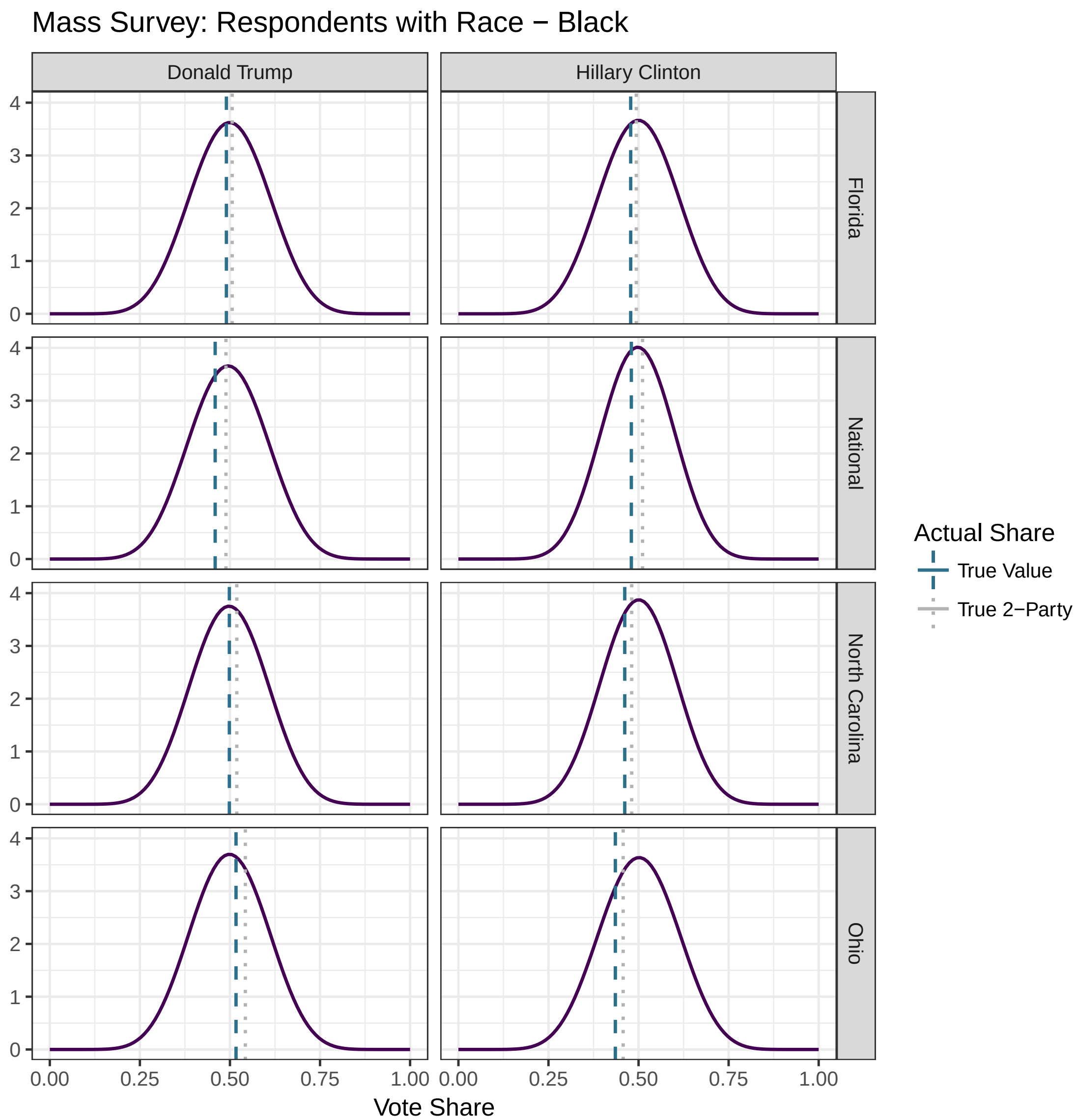


Figure 3.112: Priors without covariates: Mass Race Black

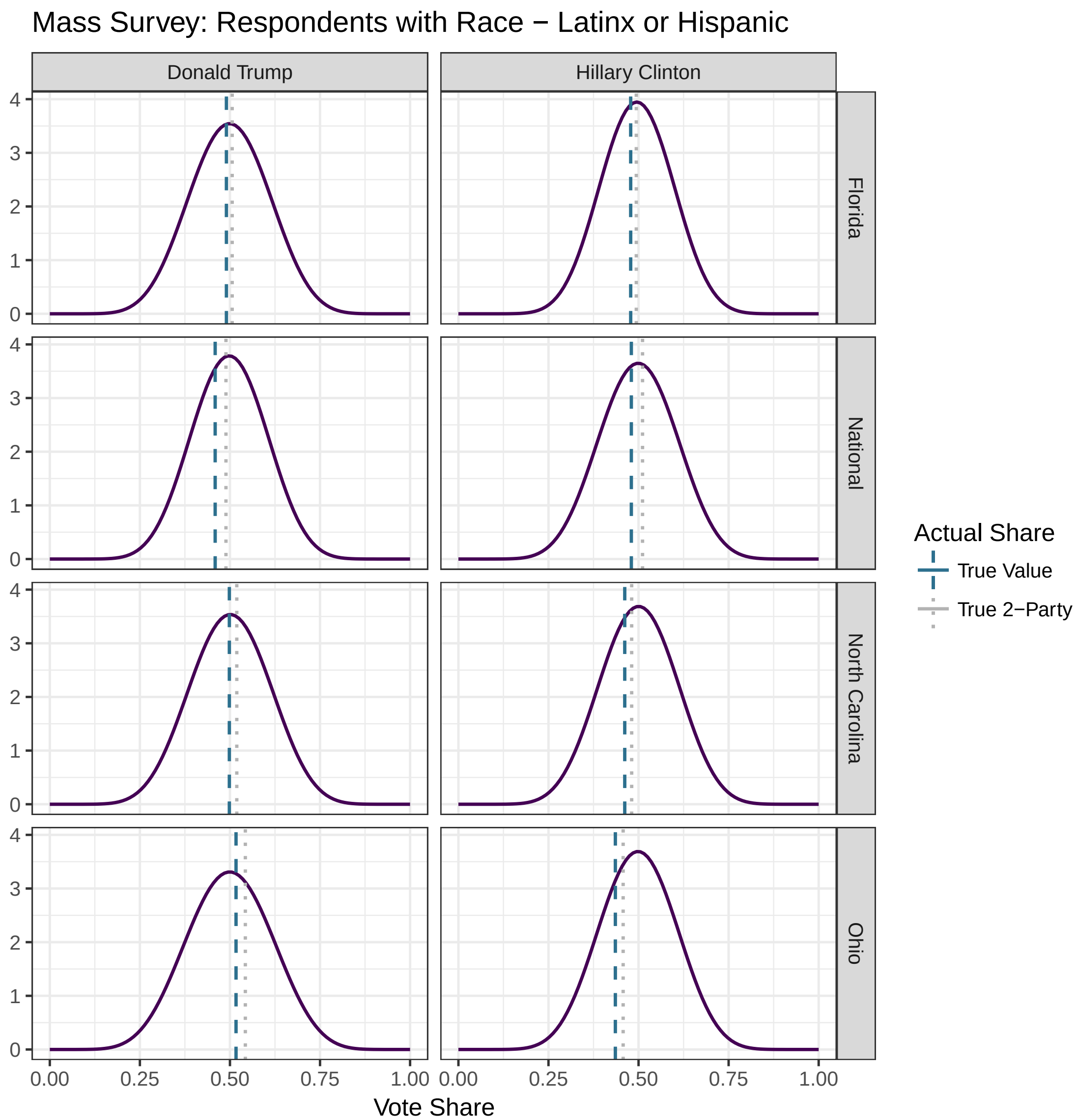


Figure 3.113: Priors without covariates: Mass Race Latinx or Hispanic

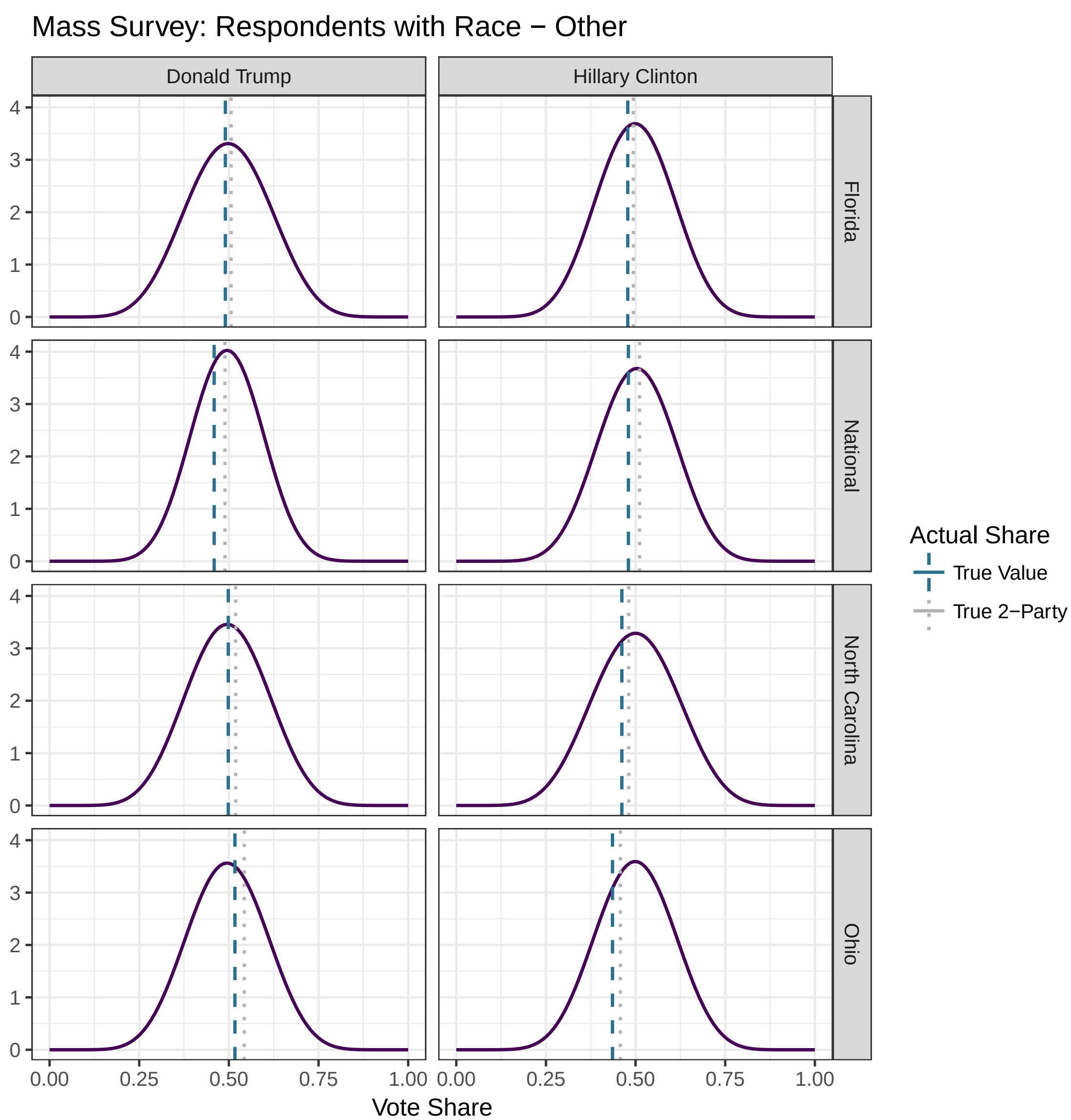


Figure 3.114: Priors without covariates: Mass Race Other

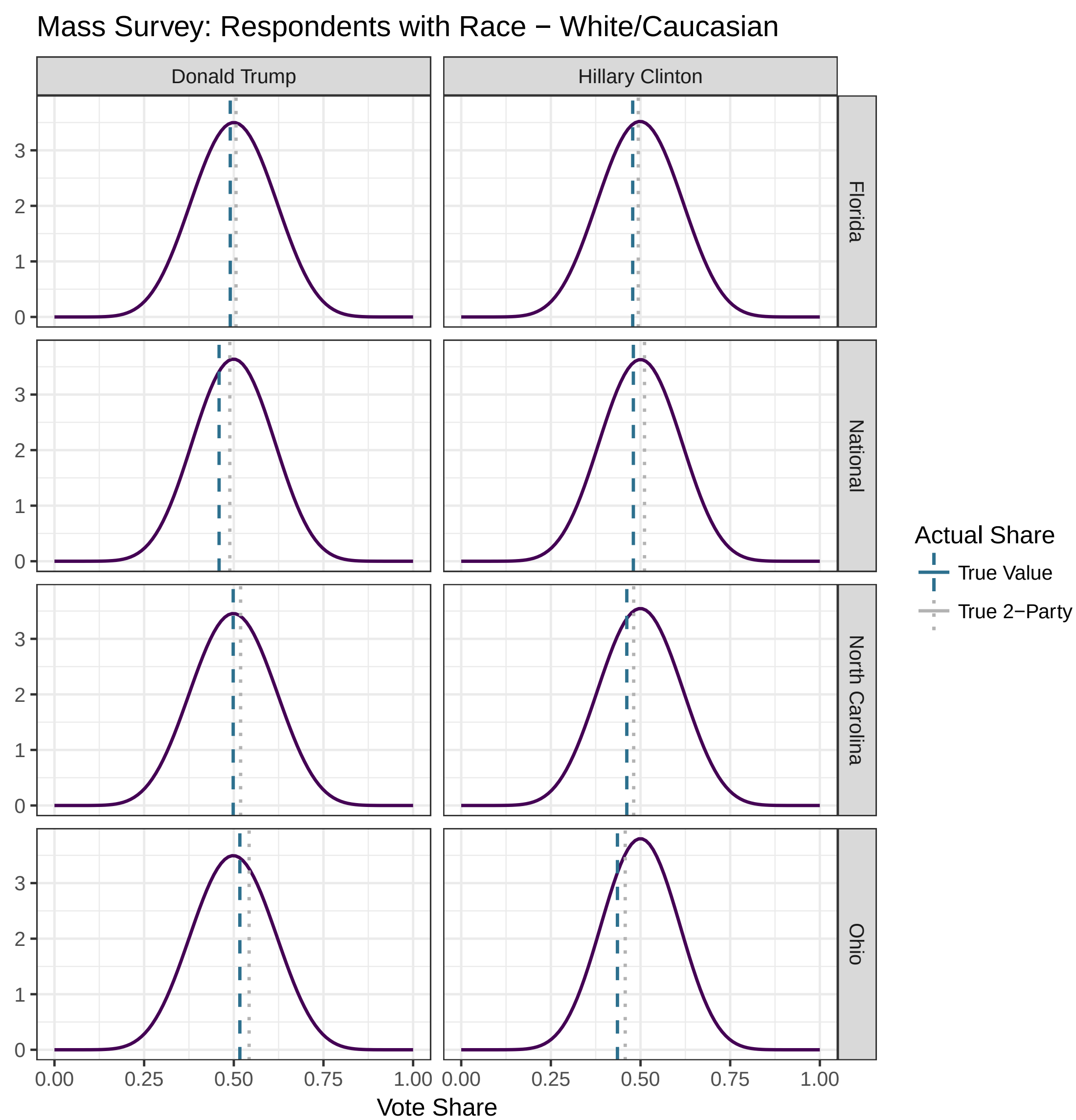


Figure 3.115: Priors without covariates: Mass Race White Caucasian

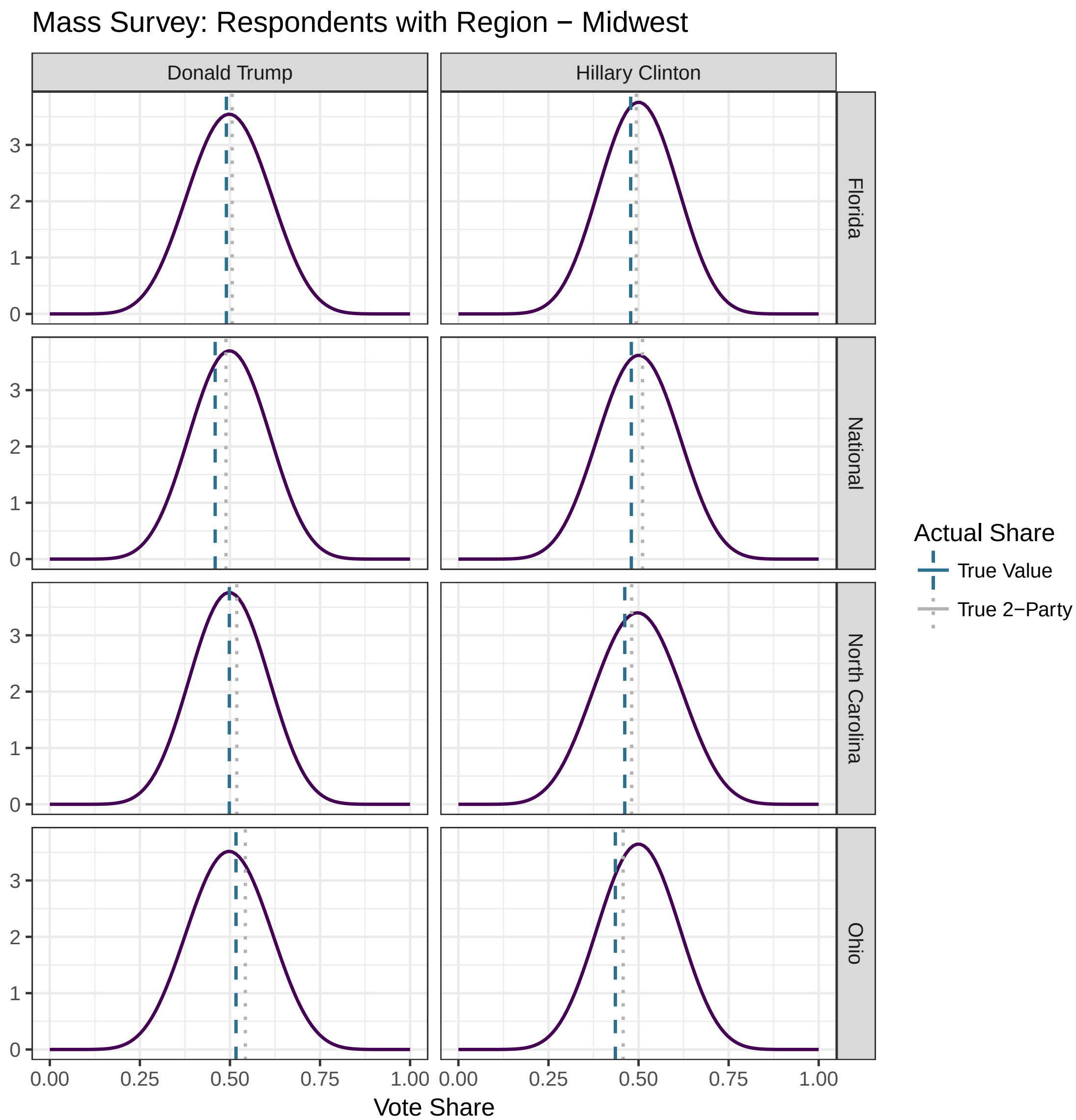


Figure 3.116: Priors without covariates: Mass Region Midwest

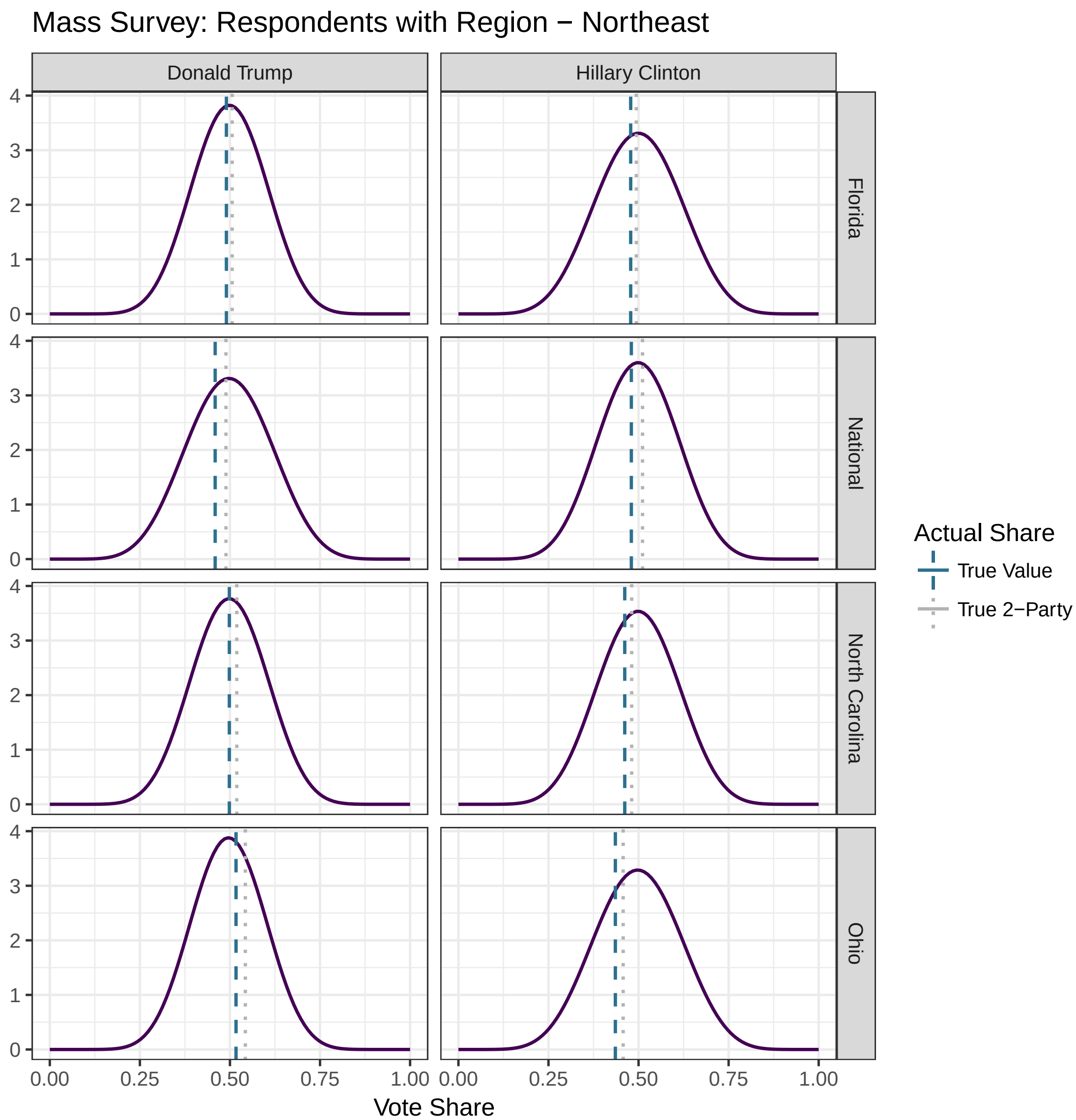


Figure 3.117: Priors without covariates: Mass Region Northeast

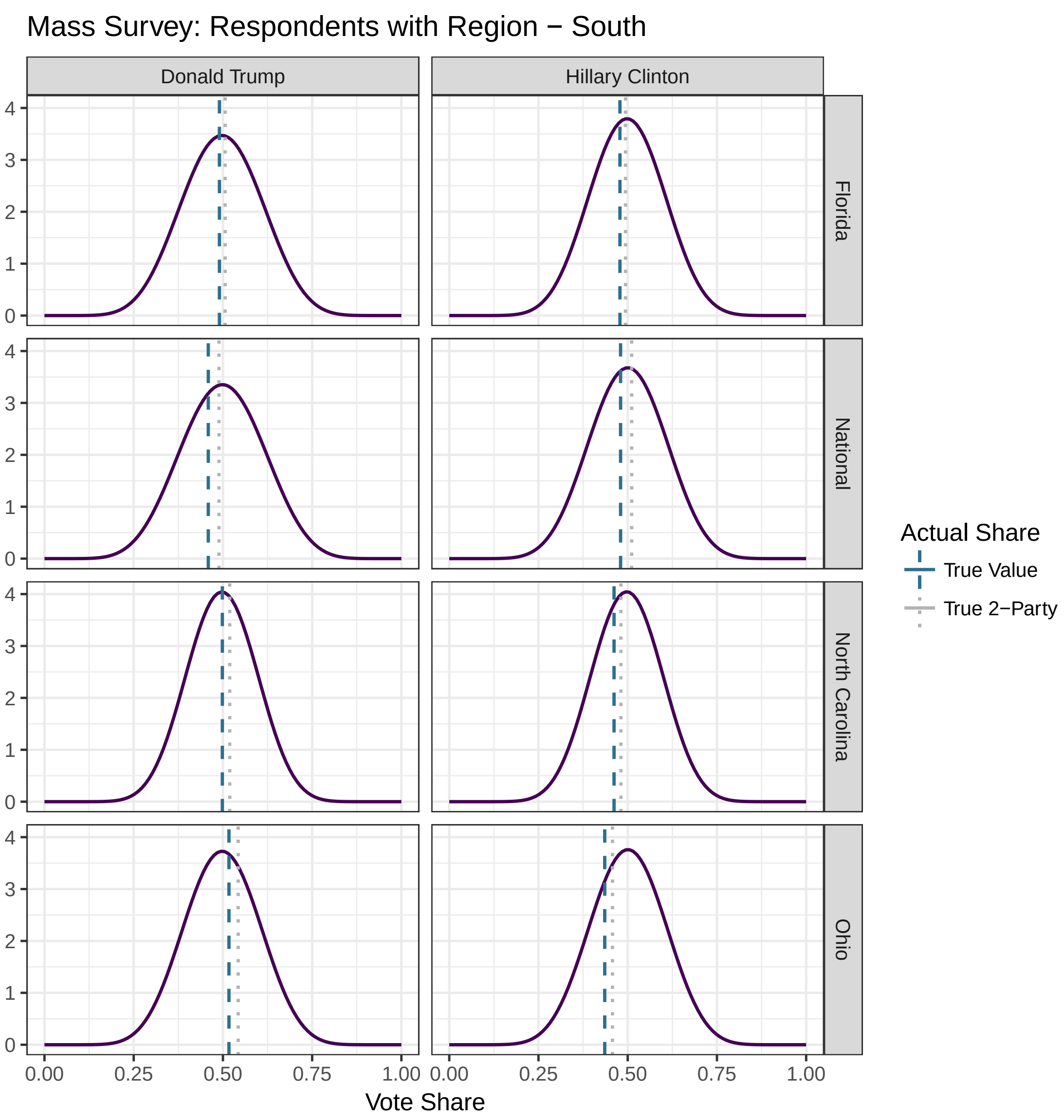


Figure 3.118: Priors without covariates: Mass Region South

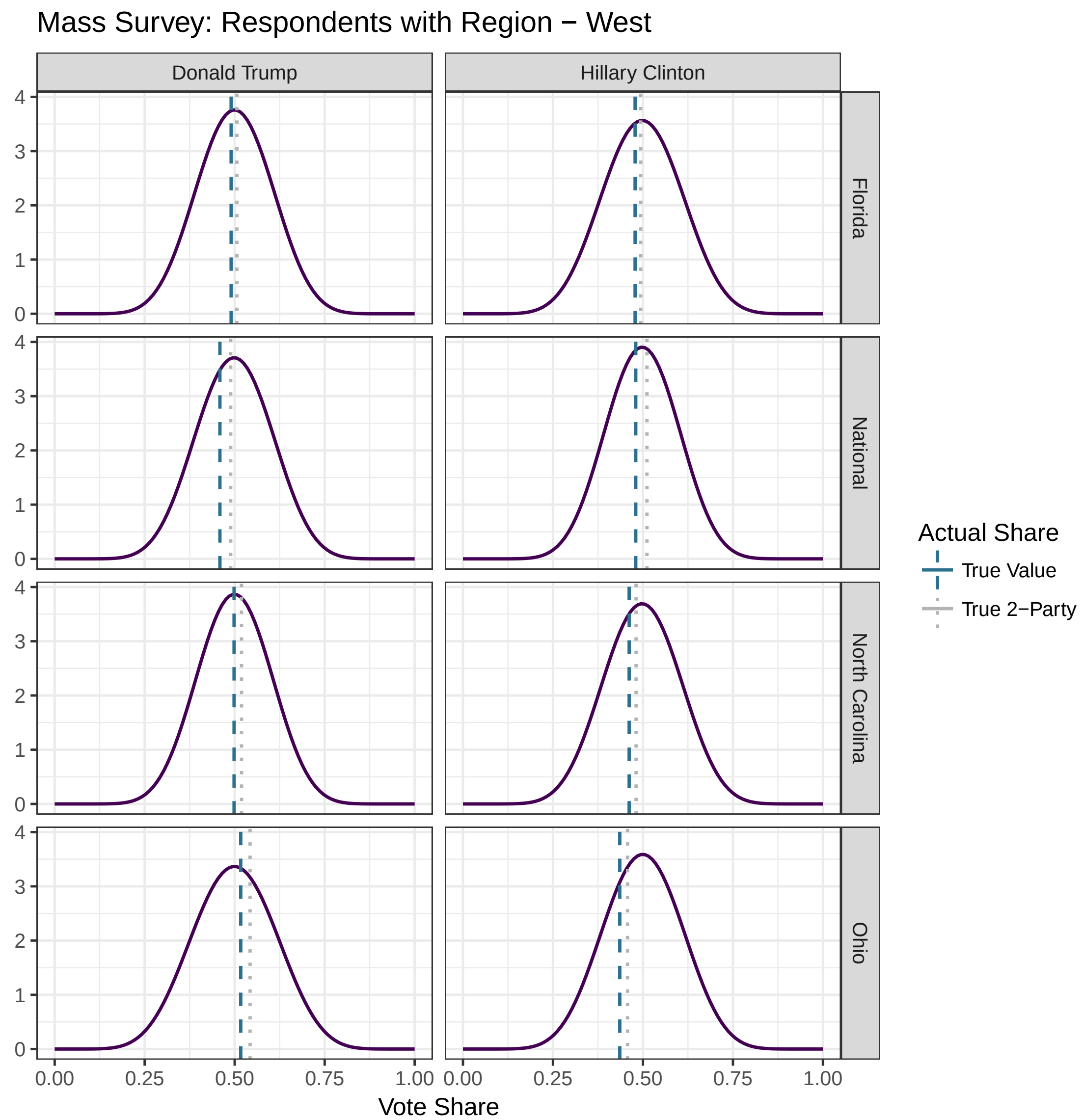


Figure 3.119: Priors without covariates: Mass Region West

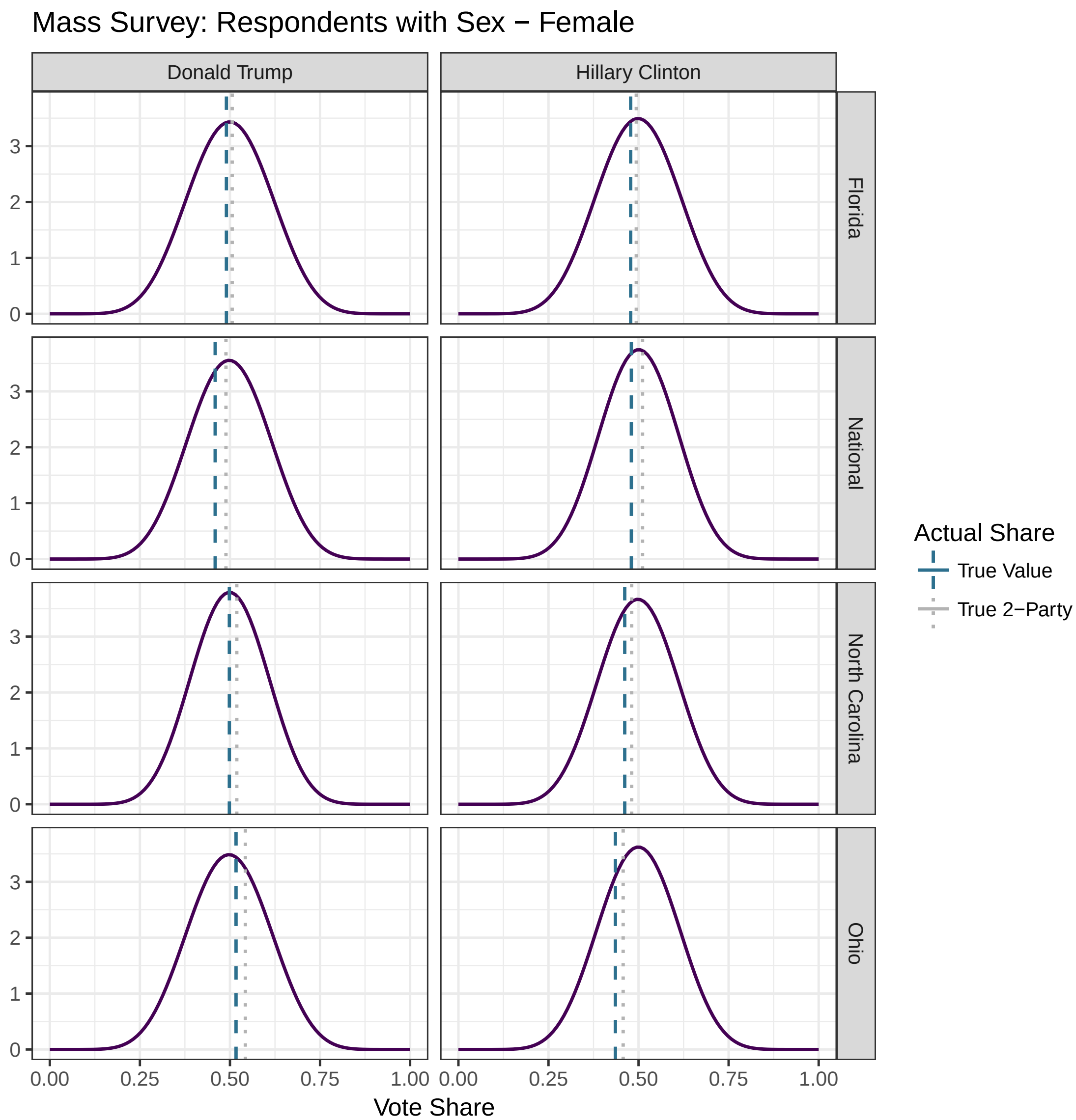


Figure 3.120: Priors without covariates: Mass Sex Female

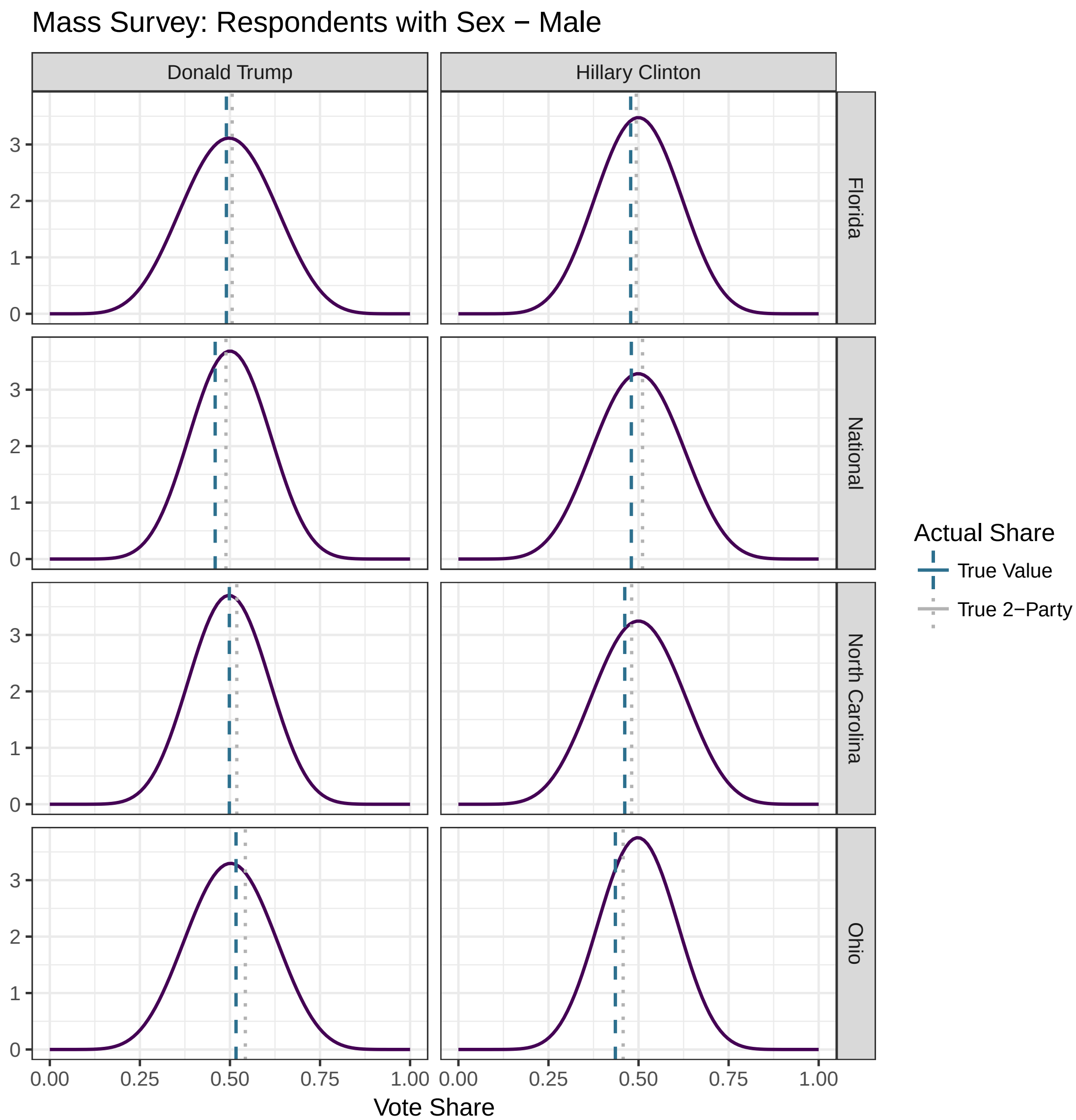


Figure 3.121: Priors without covariates: Mass Sex Male

3.7.2 Prior Plots: With Covariates

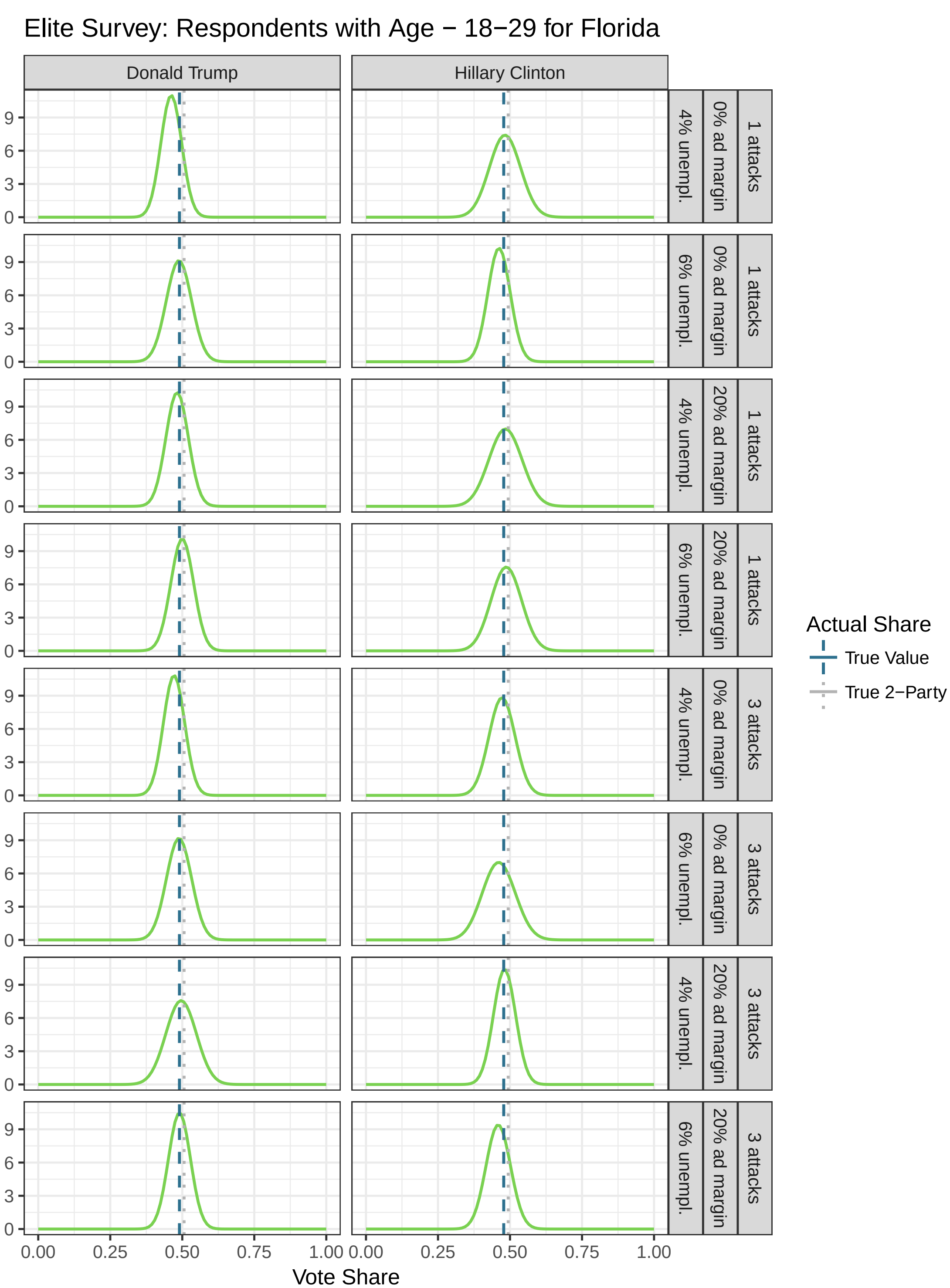


Figure 3.122: Priors with covariates: Elite Florida Age 18-29

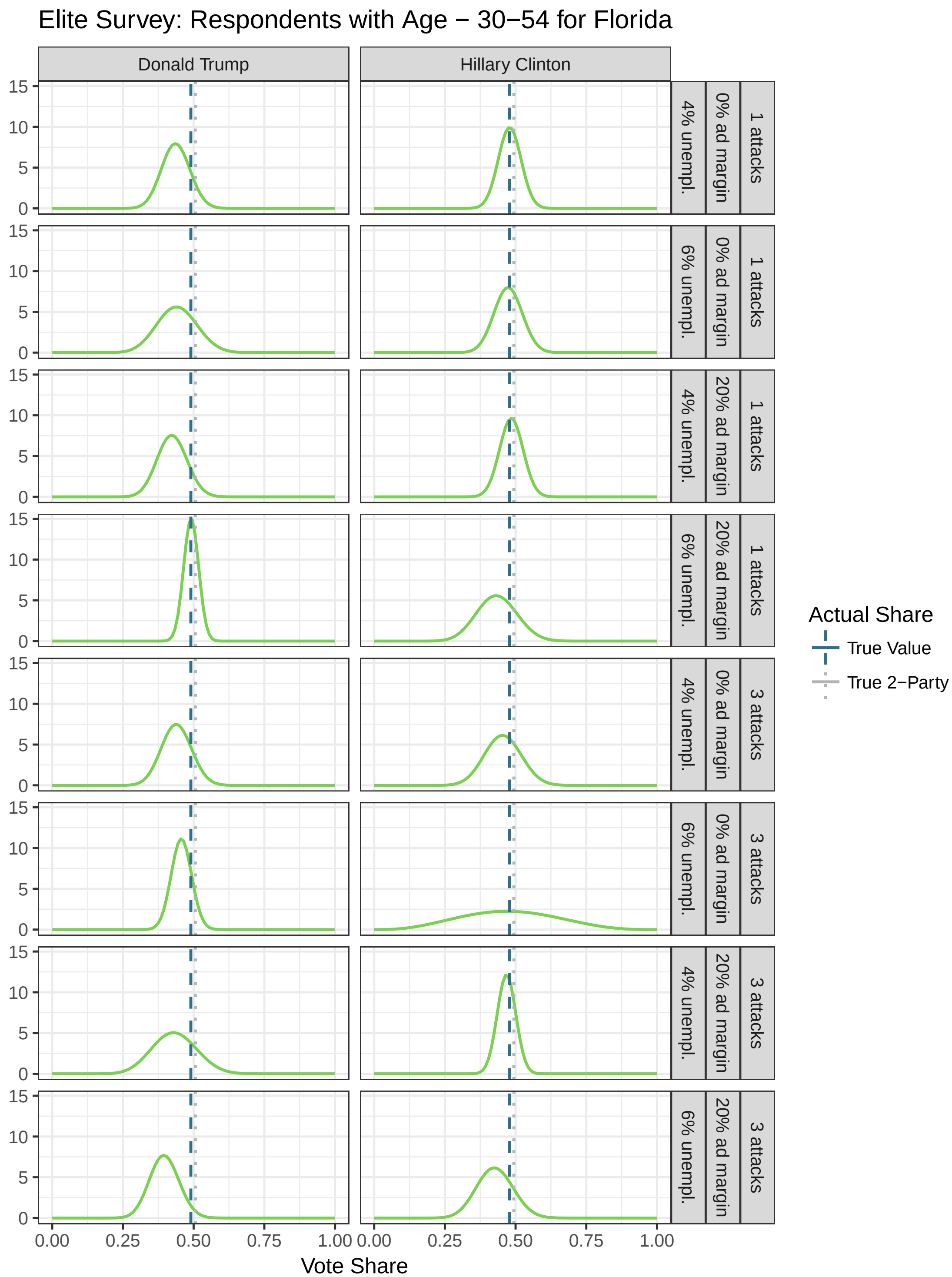


Figure 3.123: Priors with covariates: Elite Florida Age 30-54

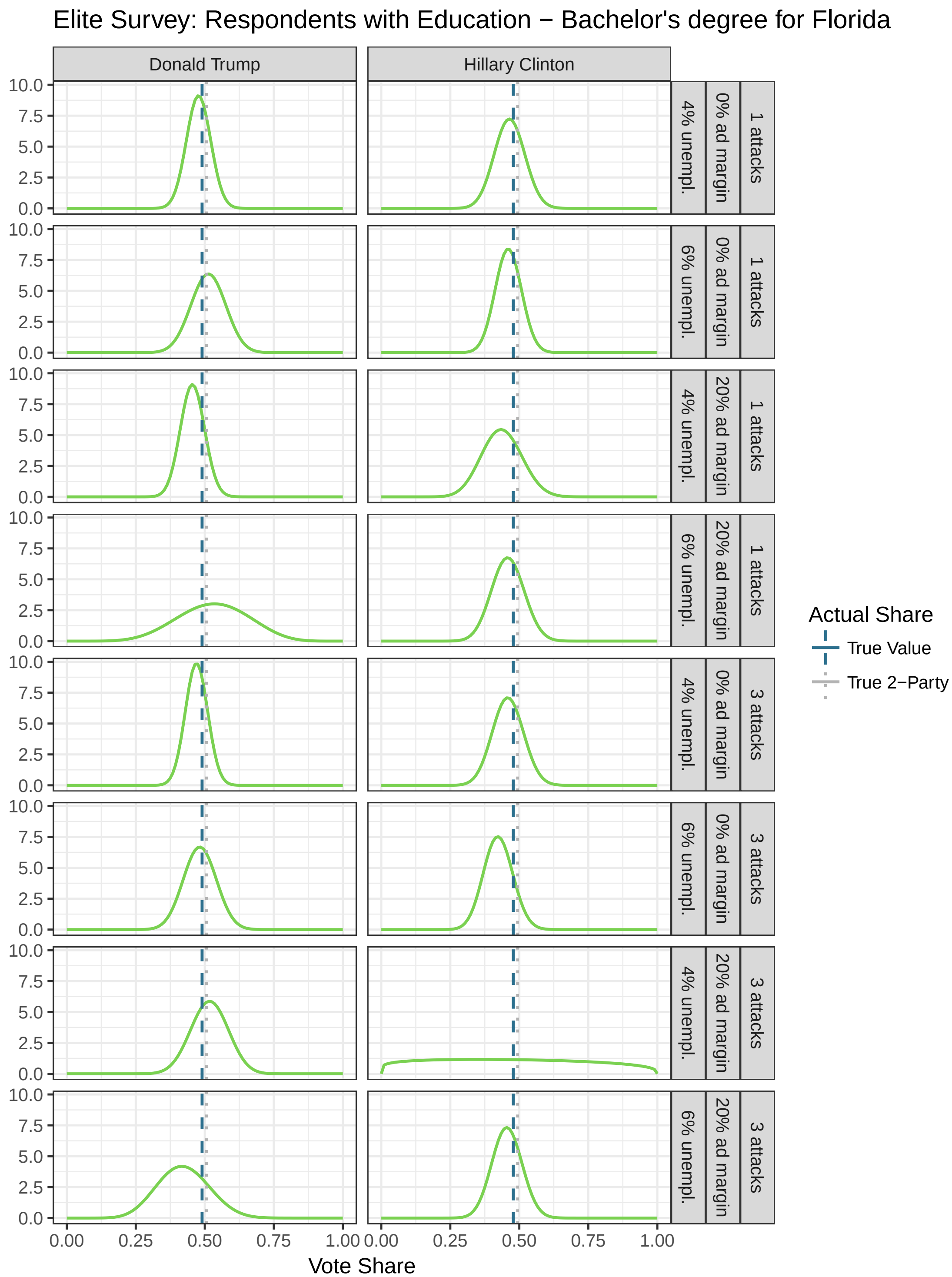


Figure 3.124: Priors with covariates: Elite Florida Education Bachelor's degree

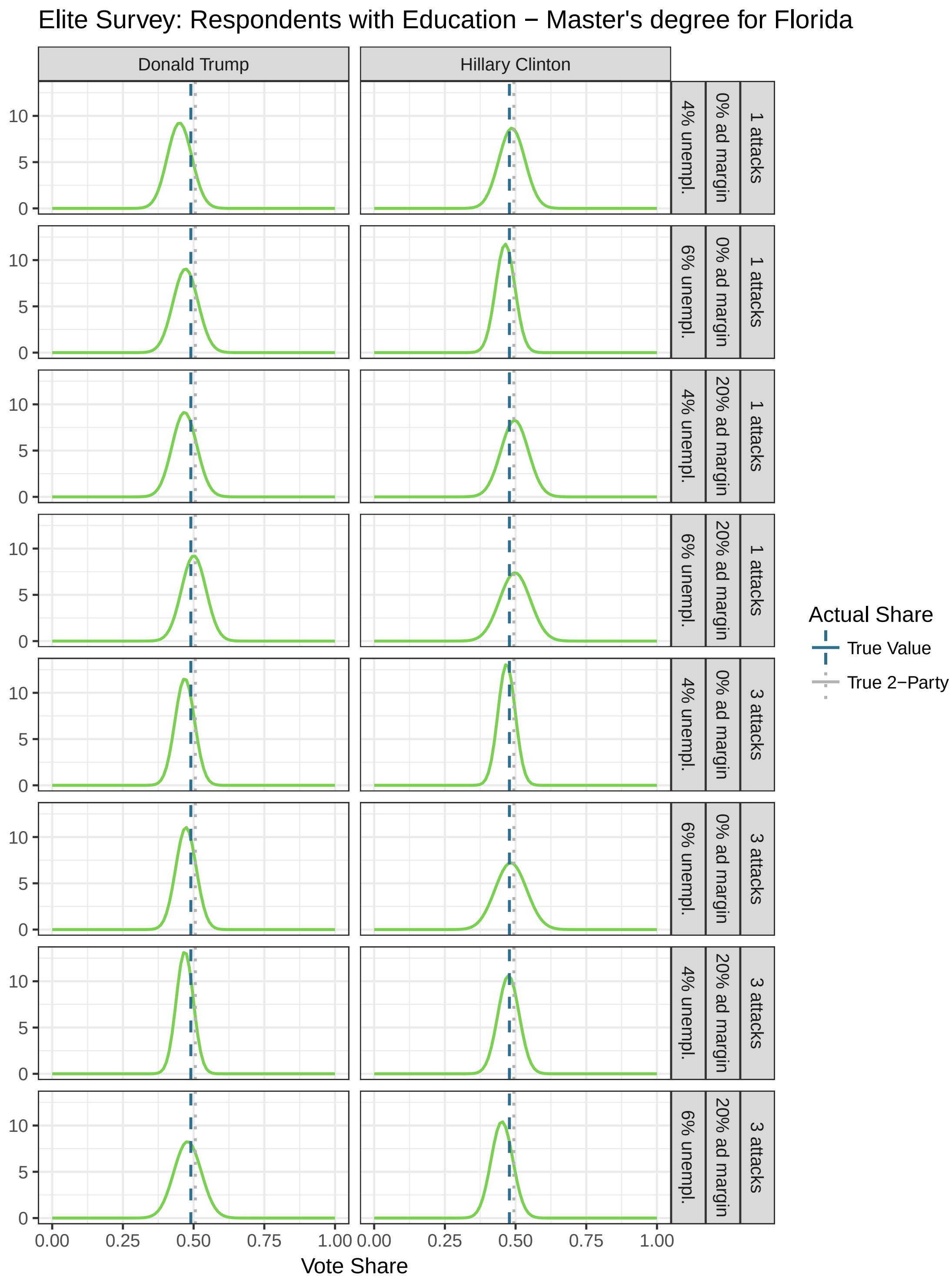


Figure 3.125: Priors with covariates: Elite Florida Education Master's degree

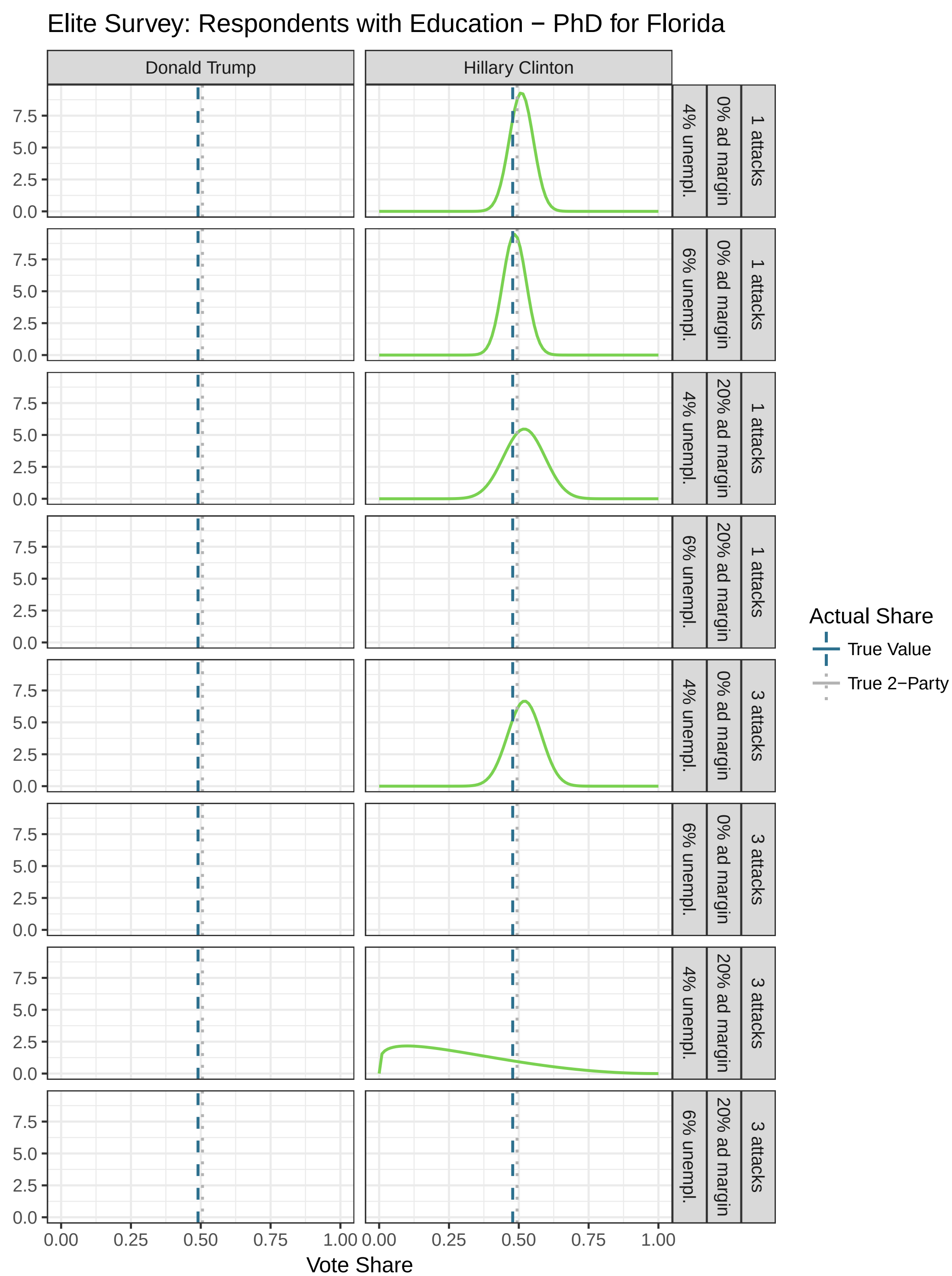


Figure 3.126: Priors with covariates: Elite Florida Education PhD

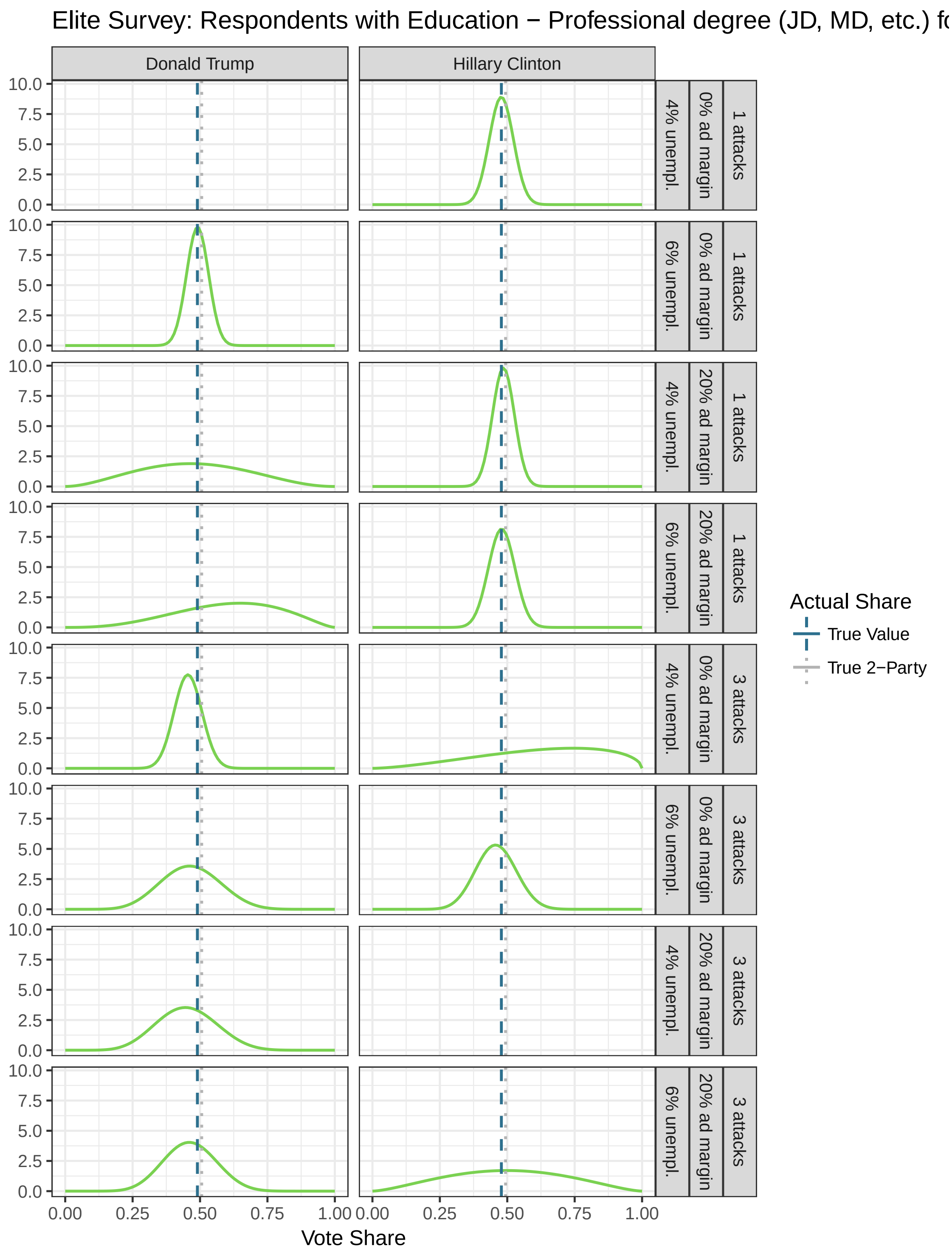


Figure 3.127: Priors with covariates: Elite Florida Education Professional degree JD MD etc

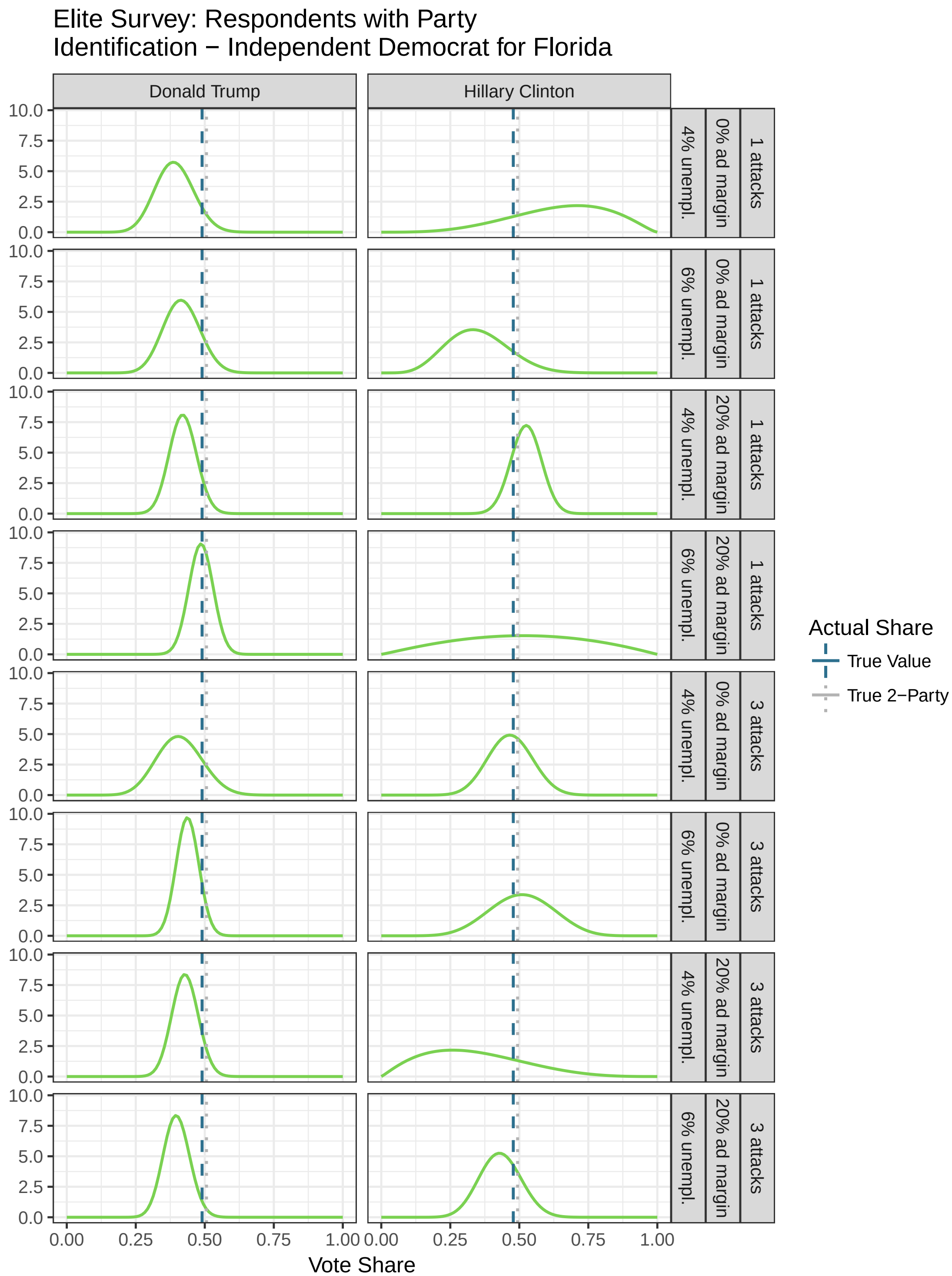


Figure 3.128: Priors with covariates: Elite Florida Party Identification Independent Democrat

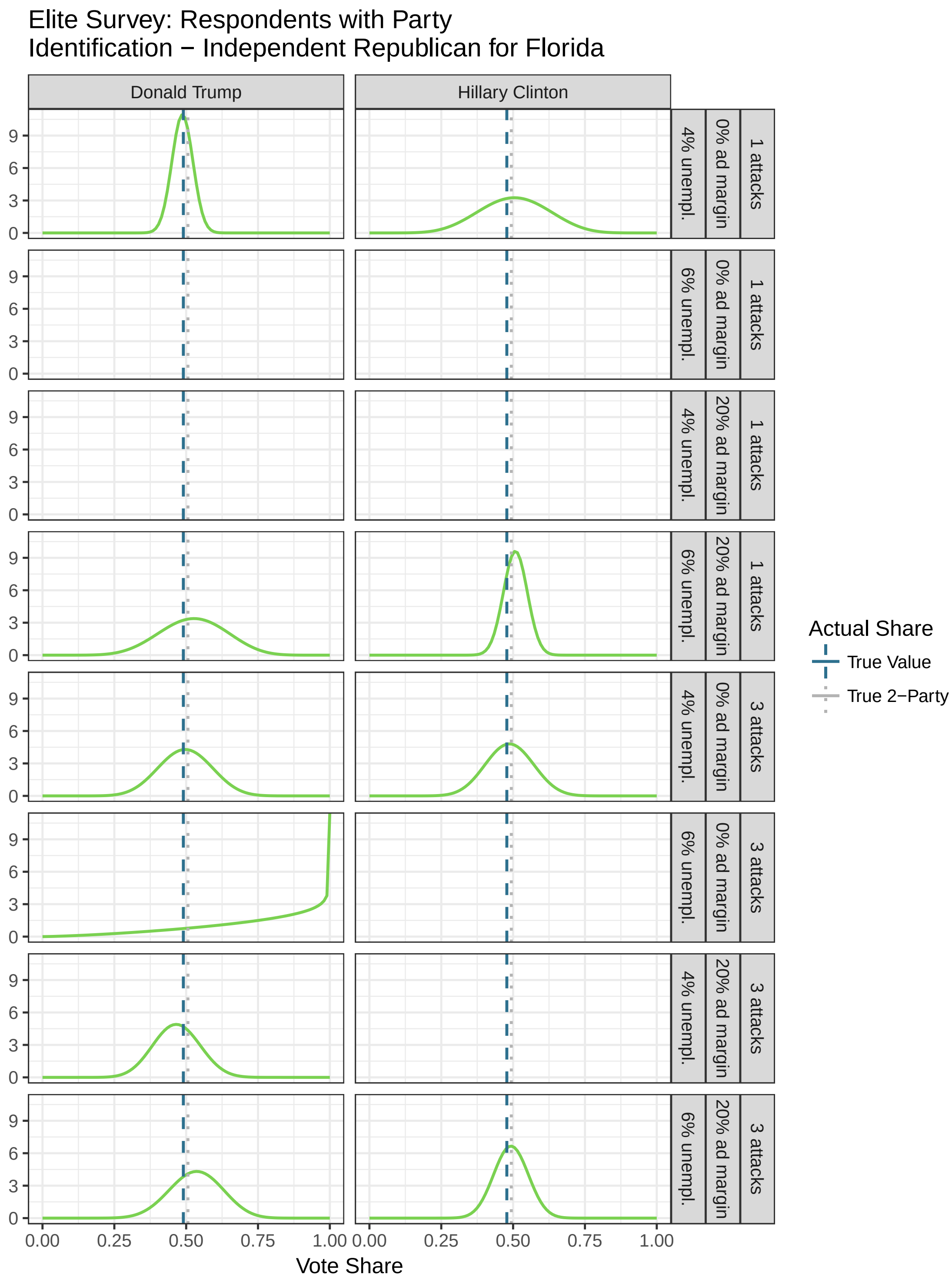


Figure 3.129: Priors with covariates: Elite Florida Party Identification Independent Republican

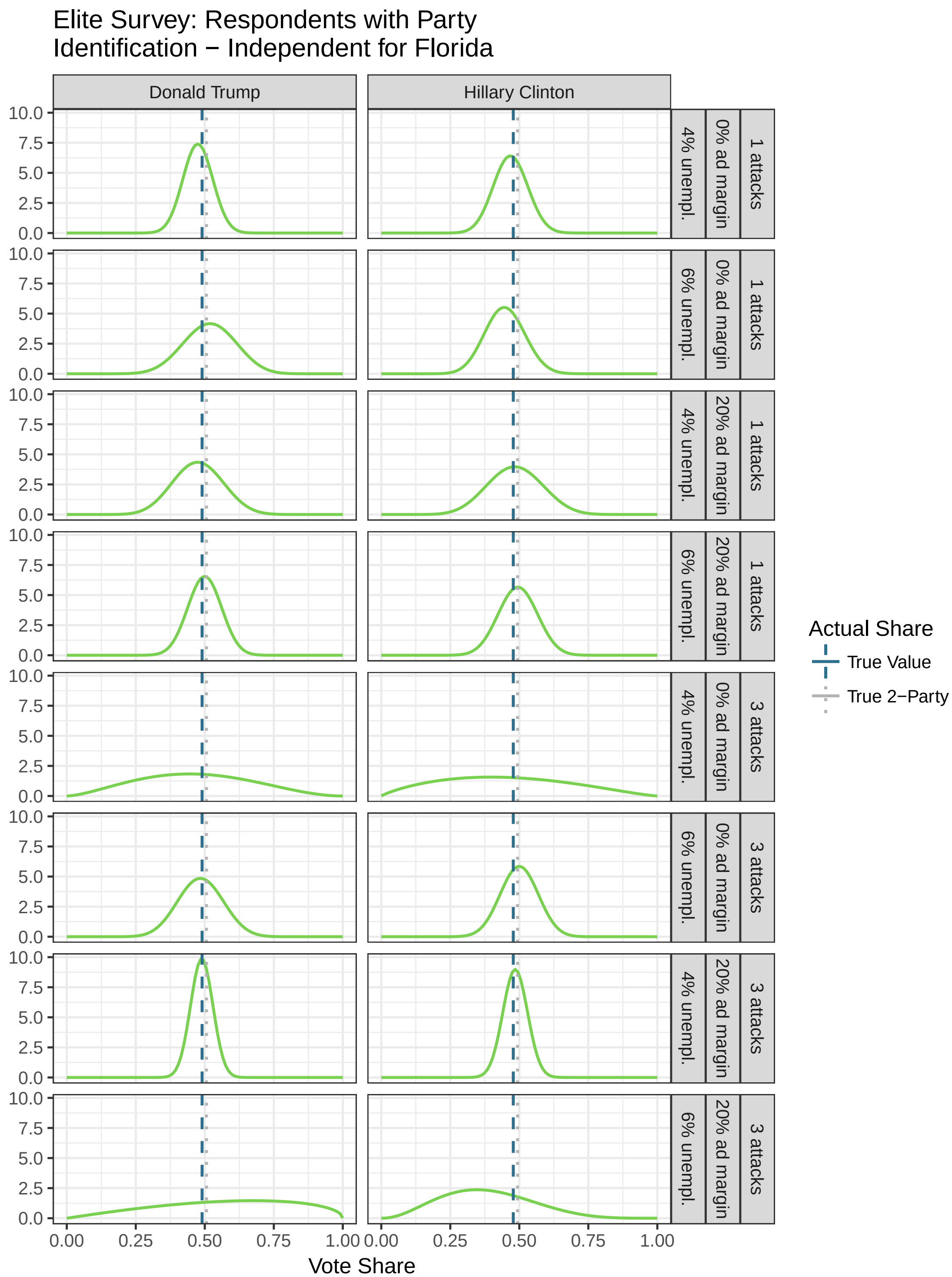


Figure 3.130: Priors with covariates: Elite Florida Party Identification Independent

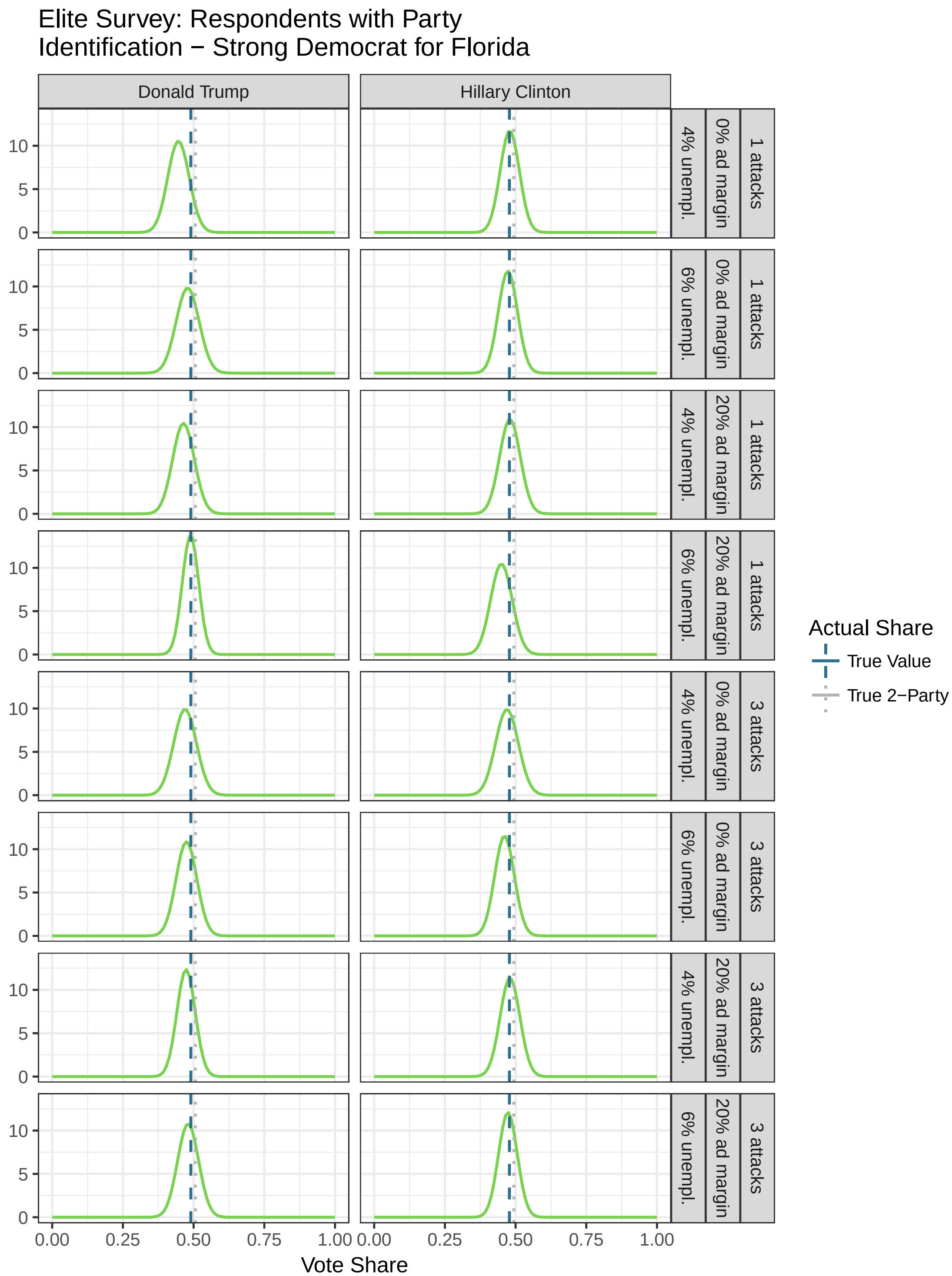


Figure 3.131: Priors with covariates: Elite Florida Party Identification Strong Democrat

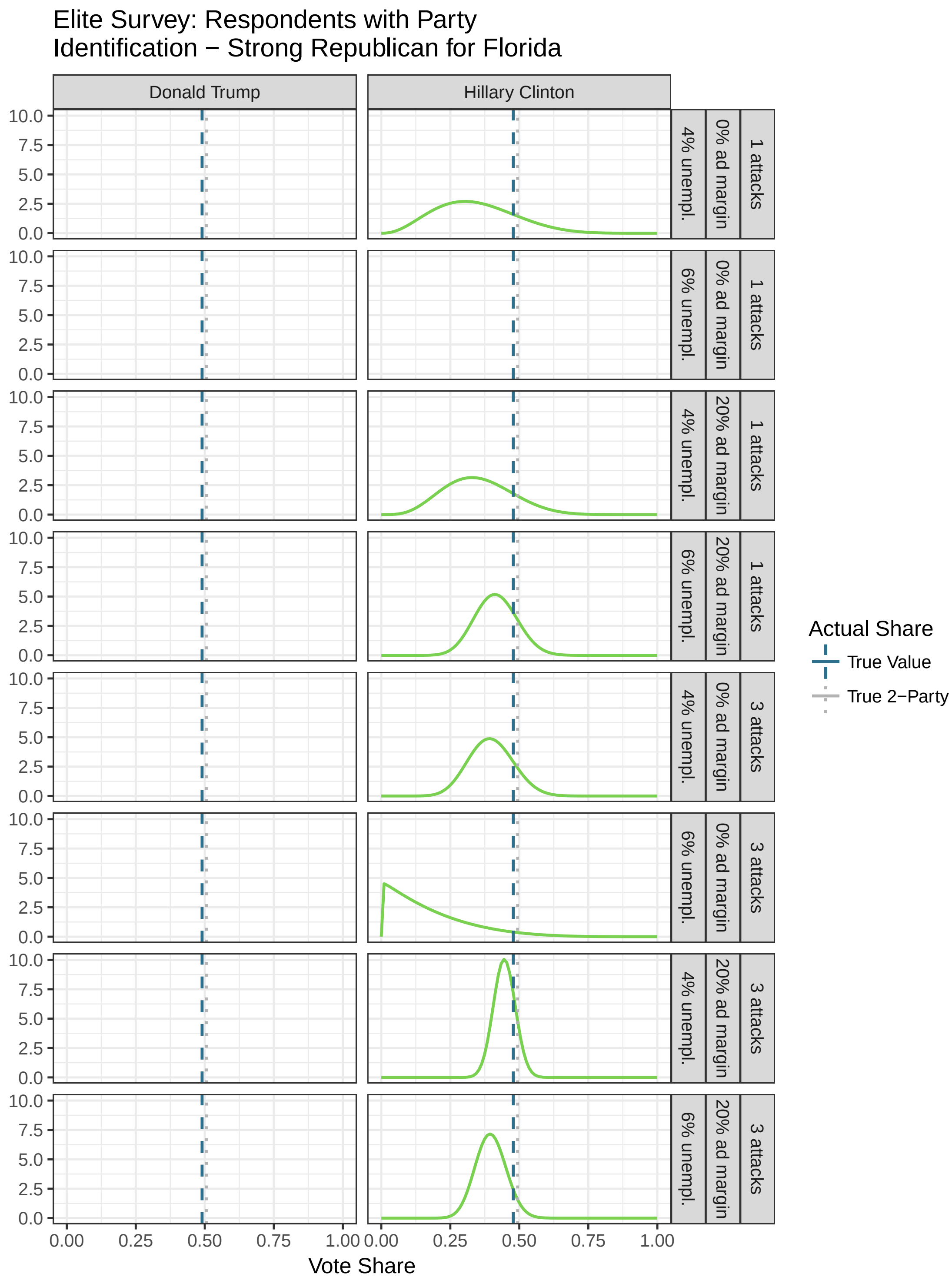


Figure 3.132: Priors with covariates: Elite Florida Party Identification Strong Republican

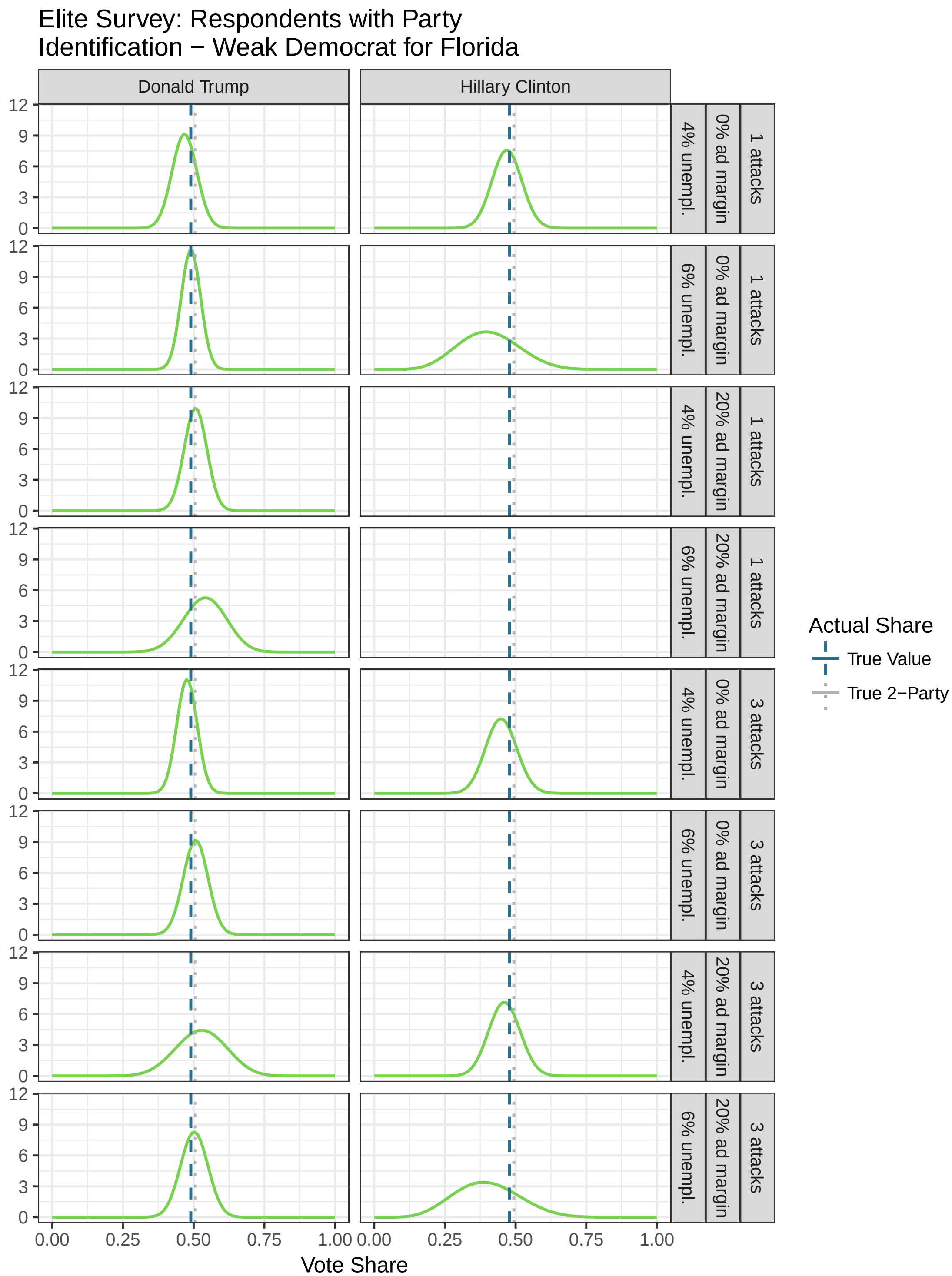


Figure 3.133: Priors with covariates: Elite Florida Party Identification Weak Democrat

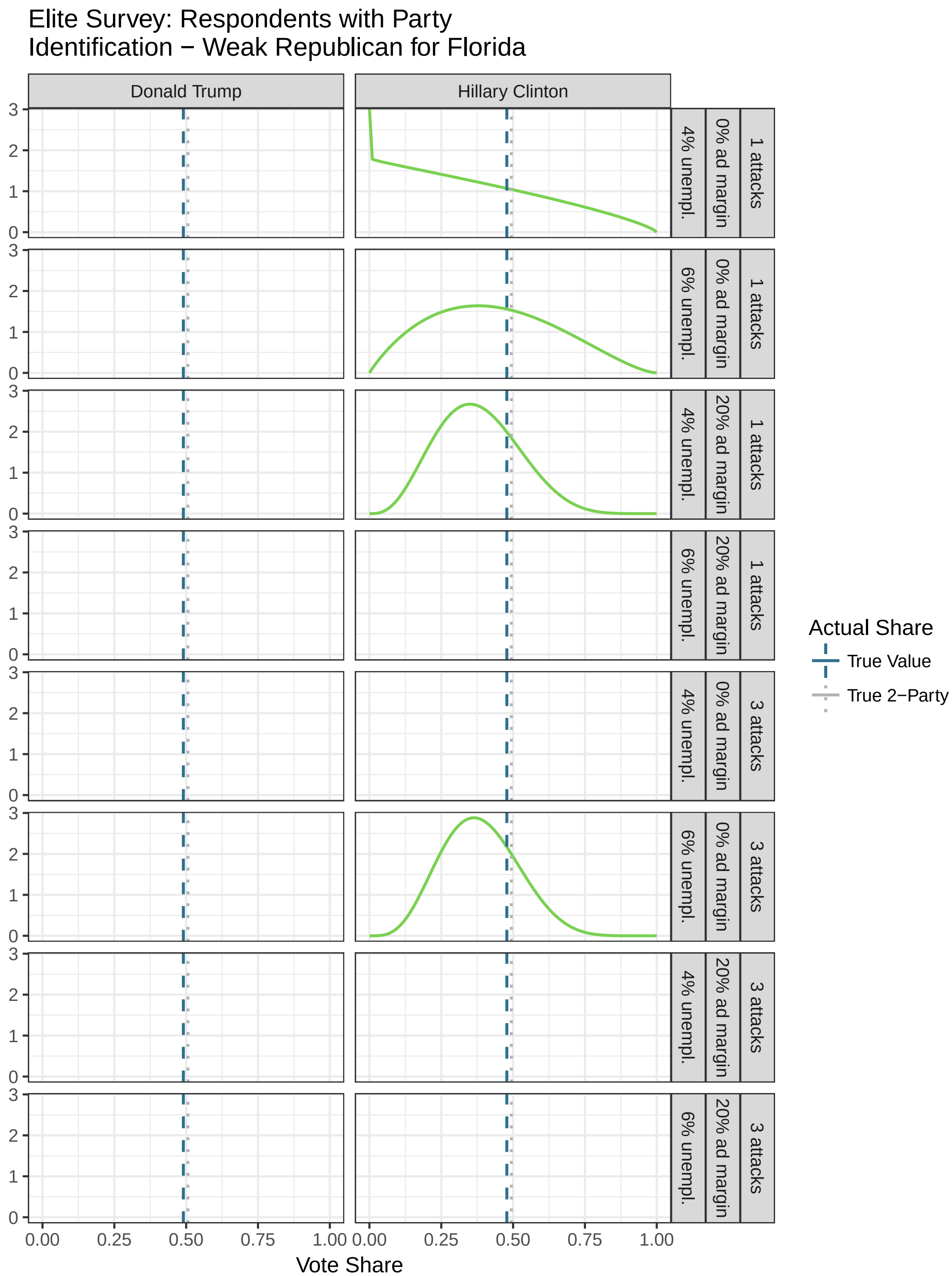


Figure 3.134: Priors with covariates: Elite Florida Party Identification Weak Republican

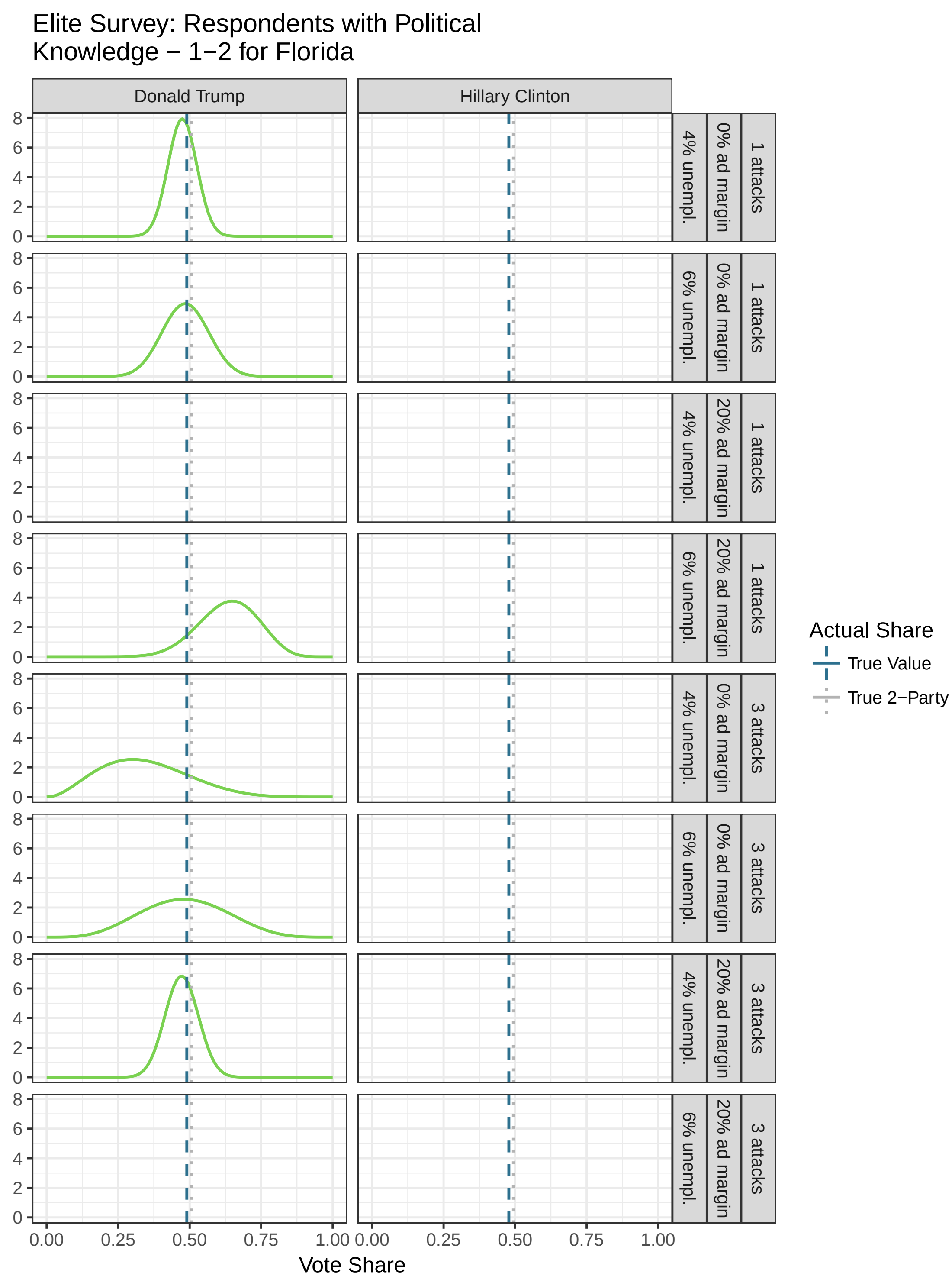


Figure 3.135: Priors with covariates: Elite Florida Political Knowledge 1-2

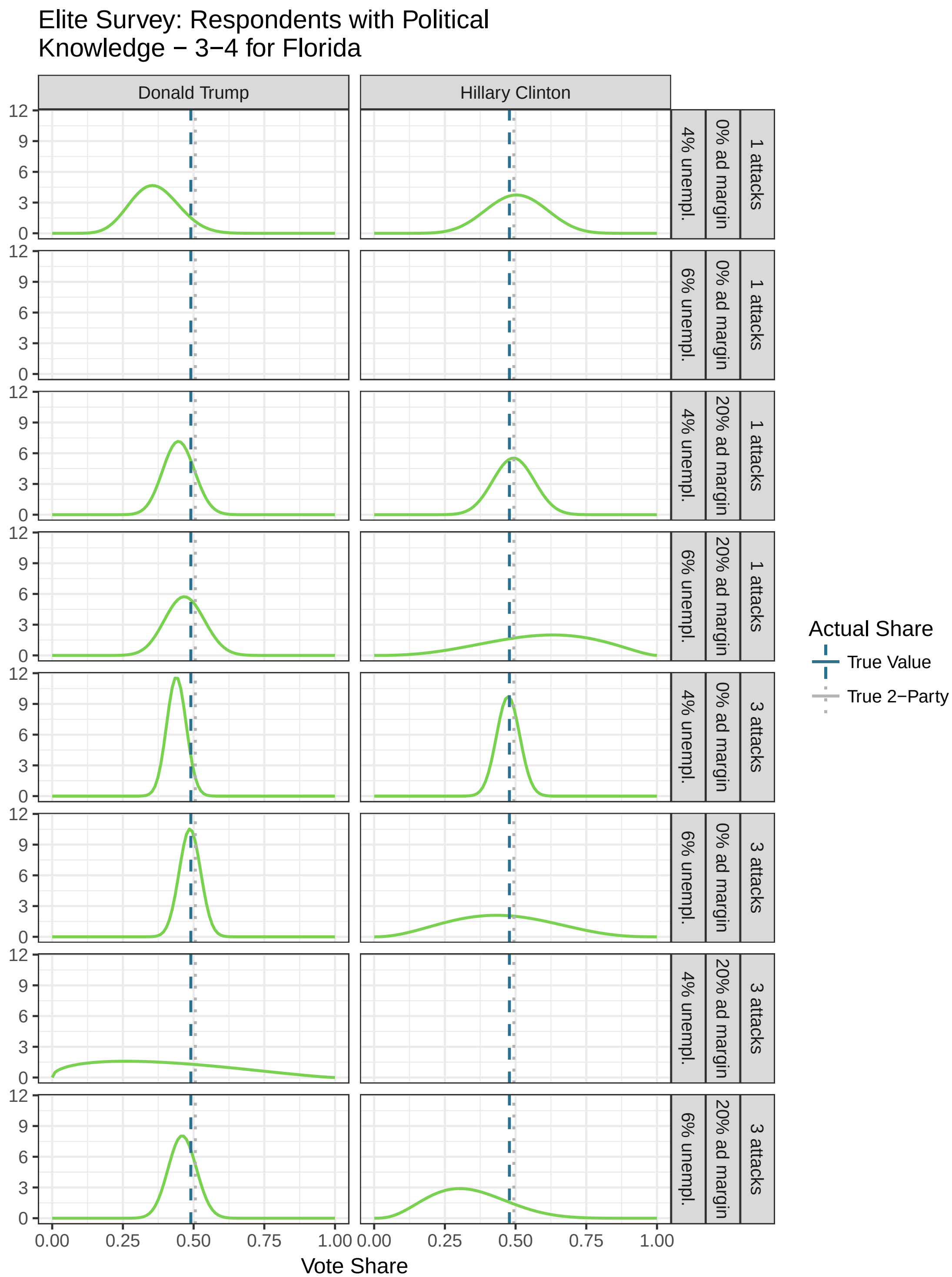


Figure 3.136: Priors with covariates: Elite Florida Political Knowledge 3-4

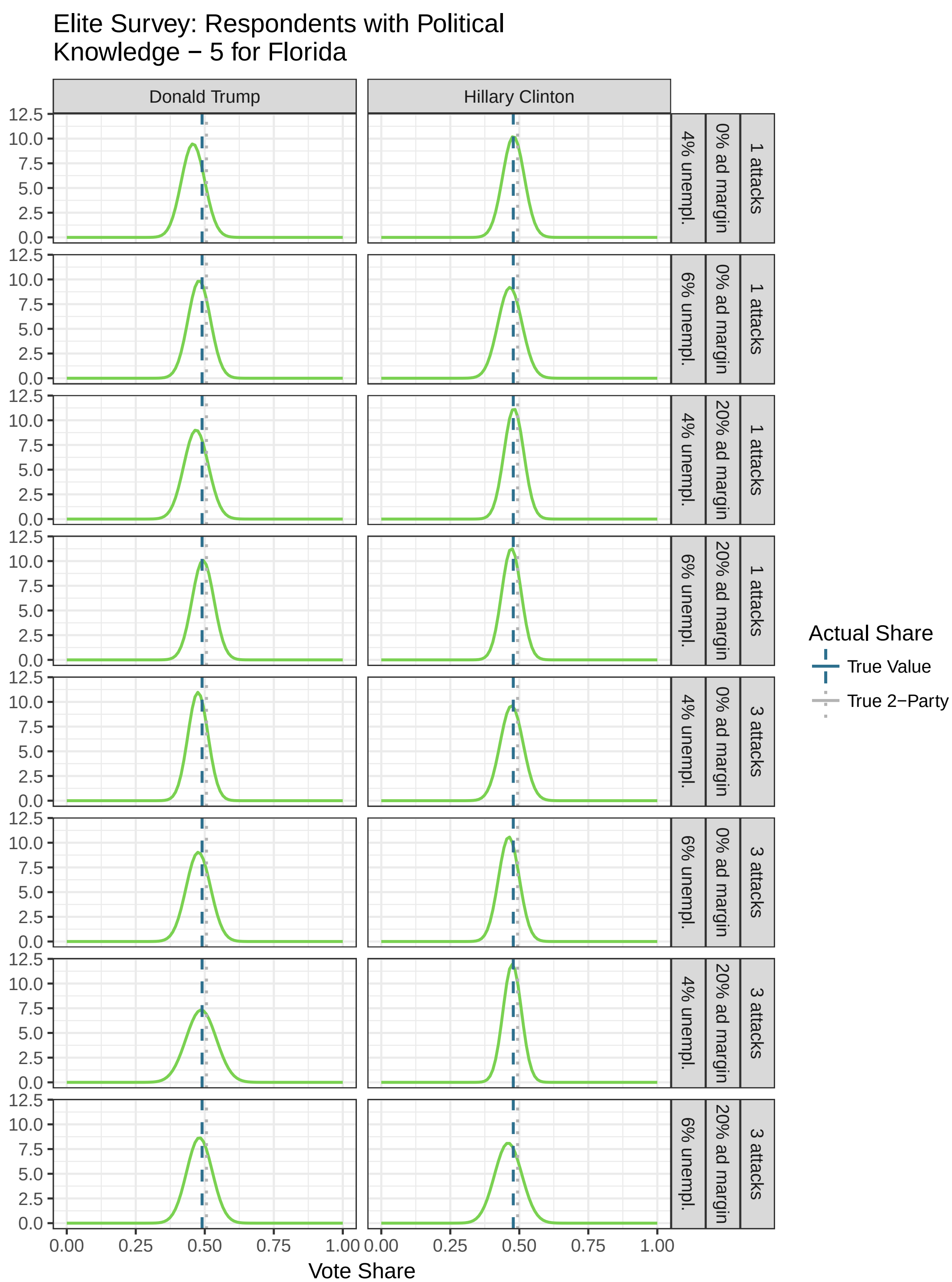


Figure 3.137: Priors with covariates: Elite Florida Political Knowledge 5

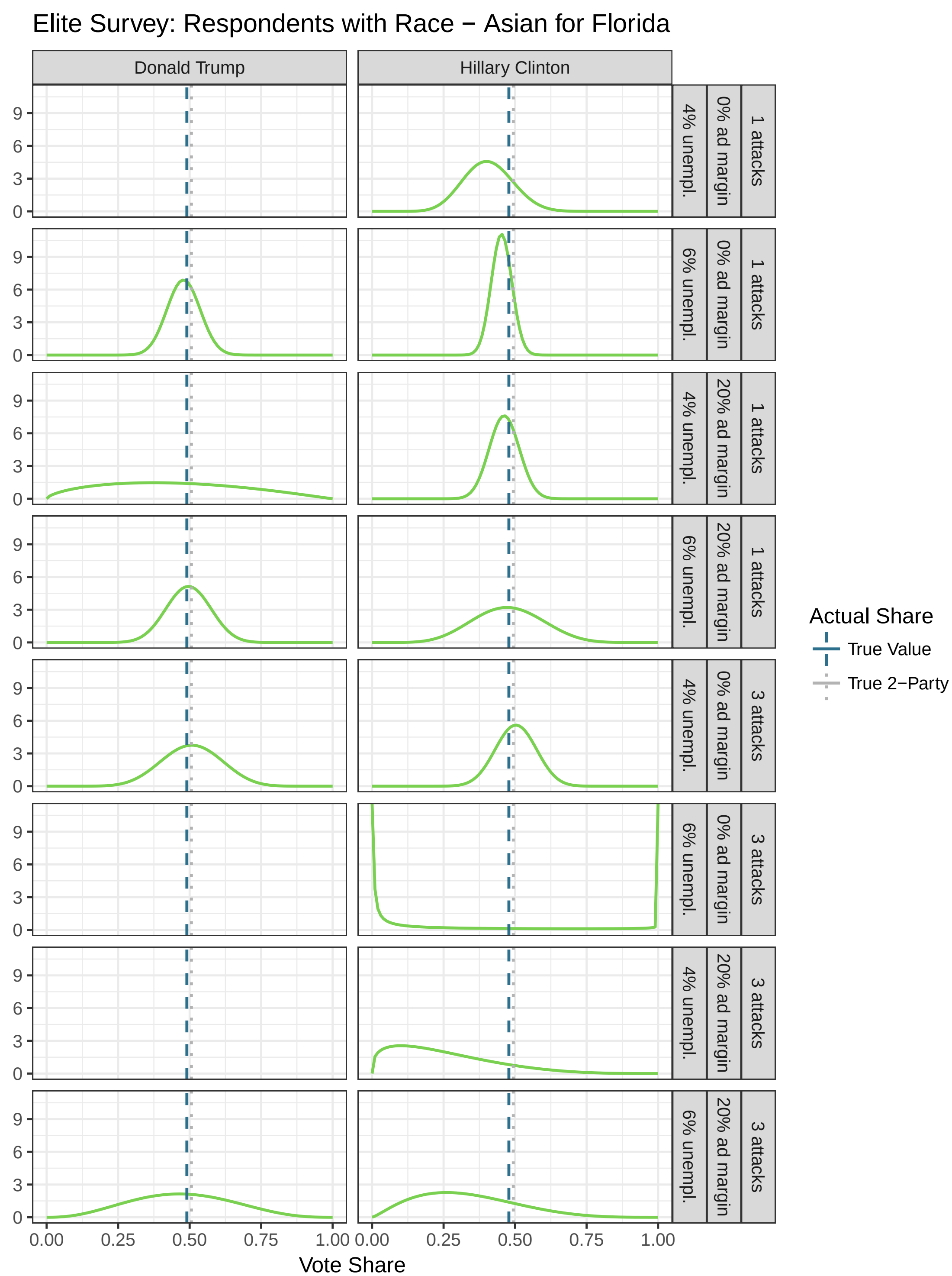


Figure 3.138: Priors with covariates: Elite Florida Race Asian

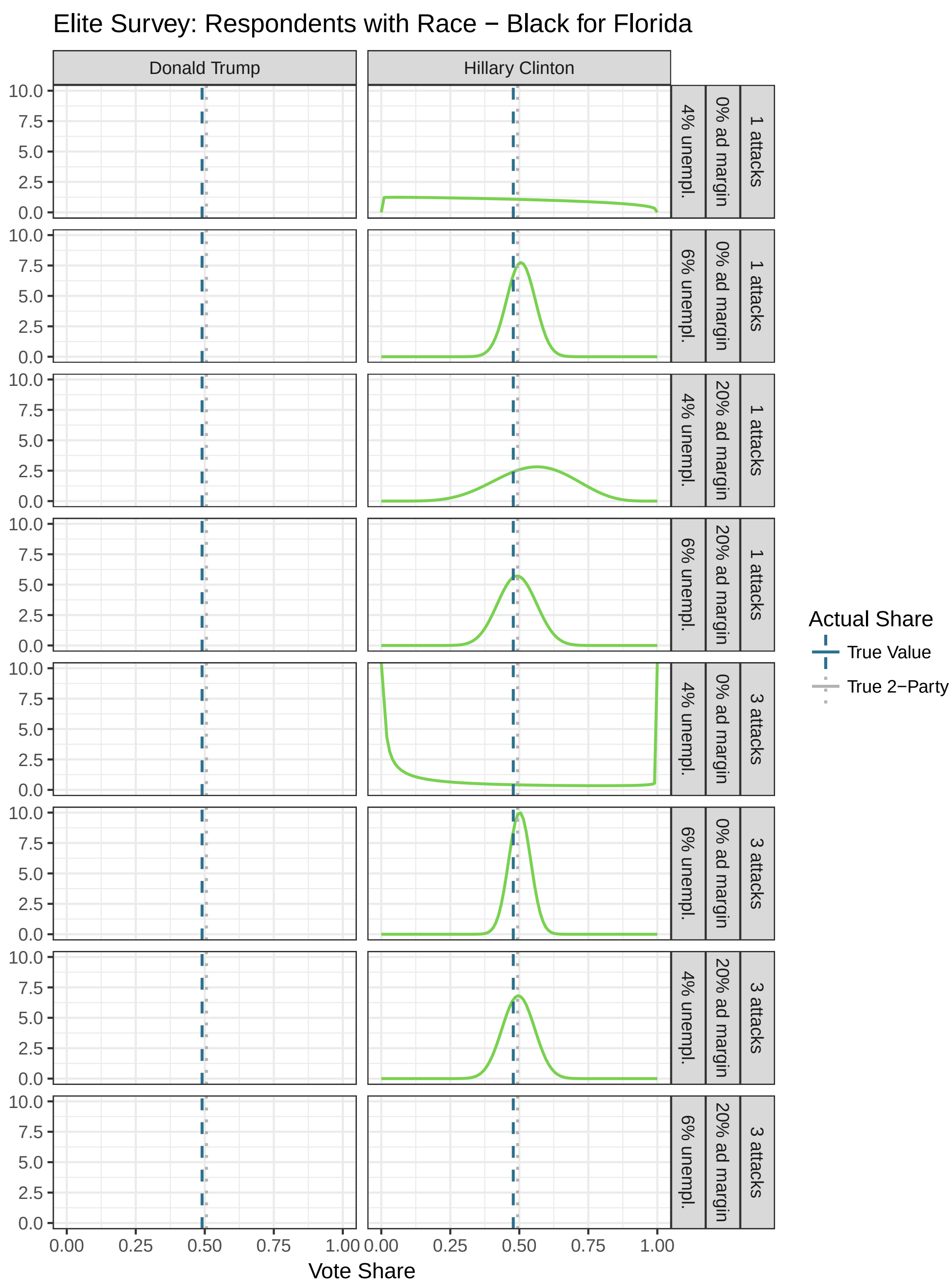


Figure 3.139: Priors with covariates: Elite Florida Race Black

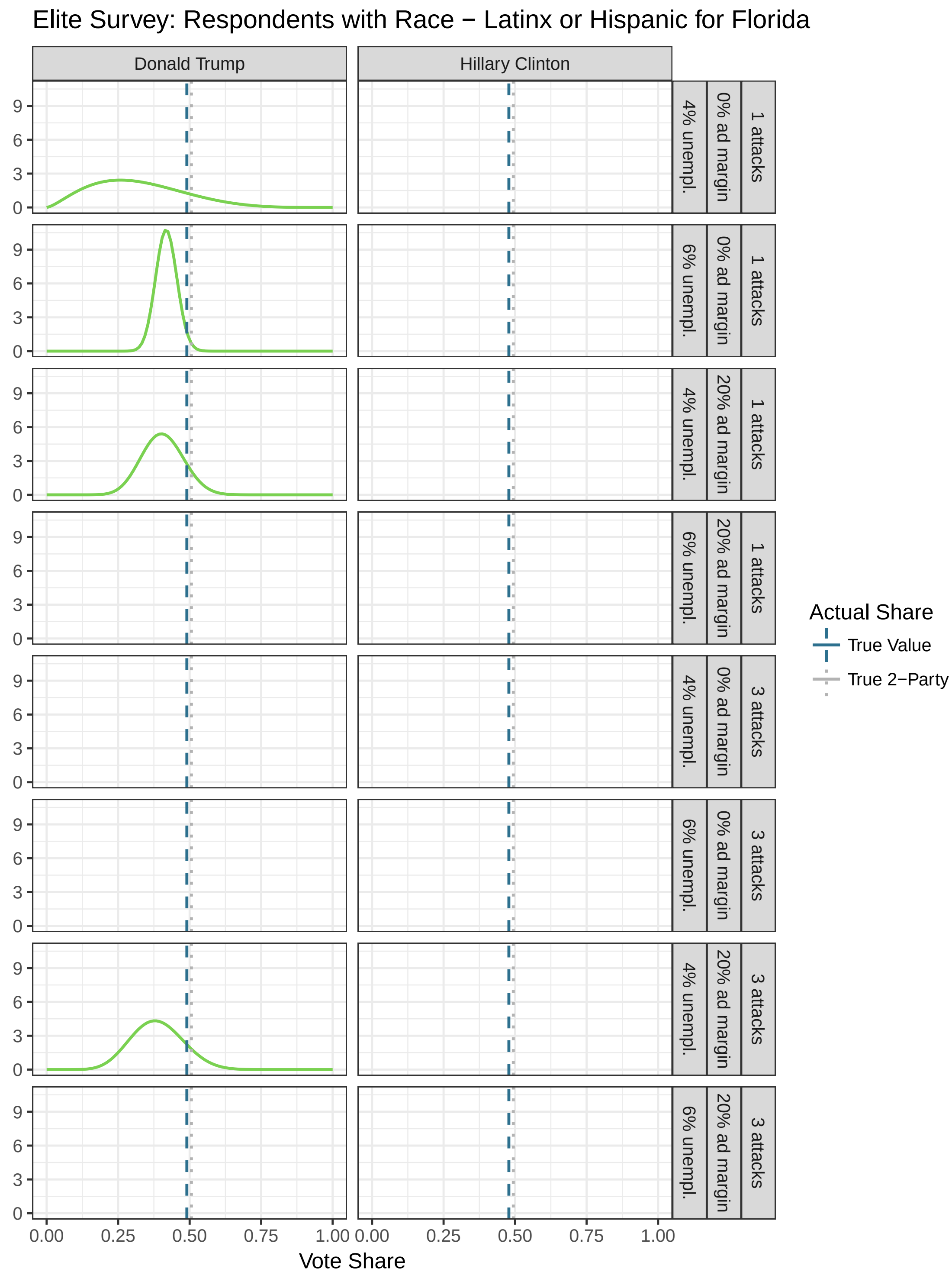


Figure 3.140: Priors with covariates: Elite Florida Race Latinx or Hispanic

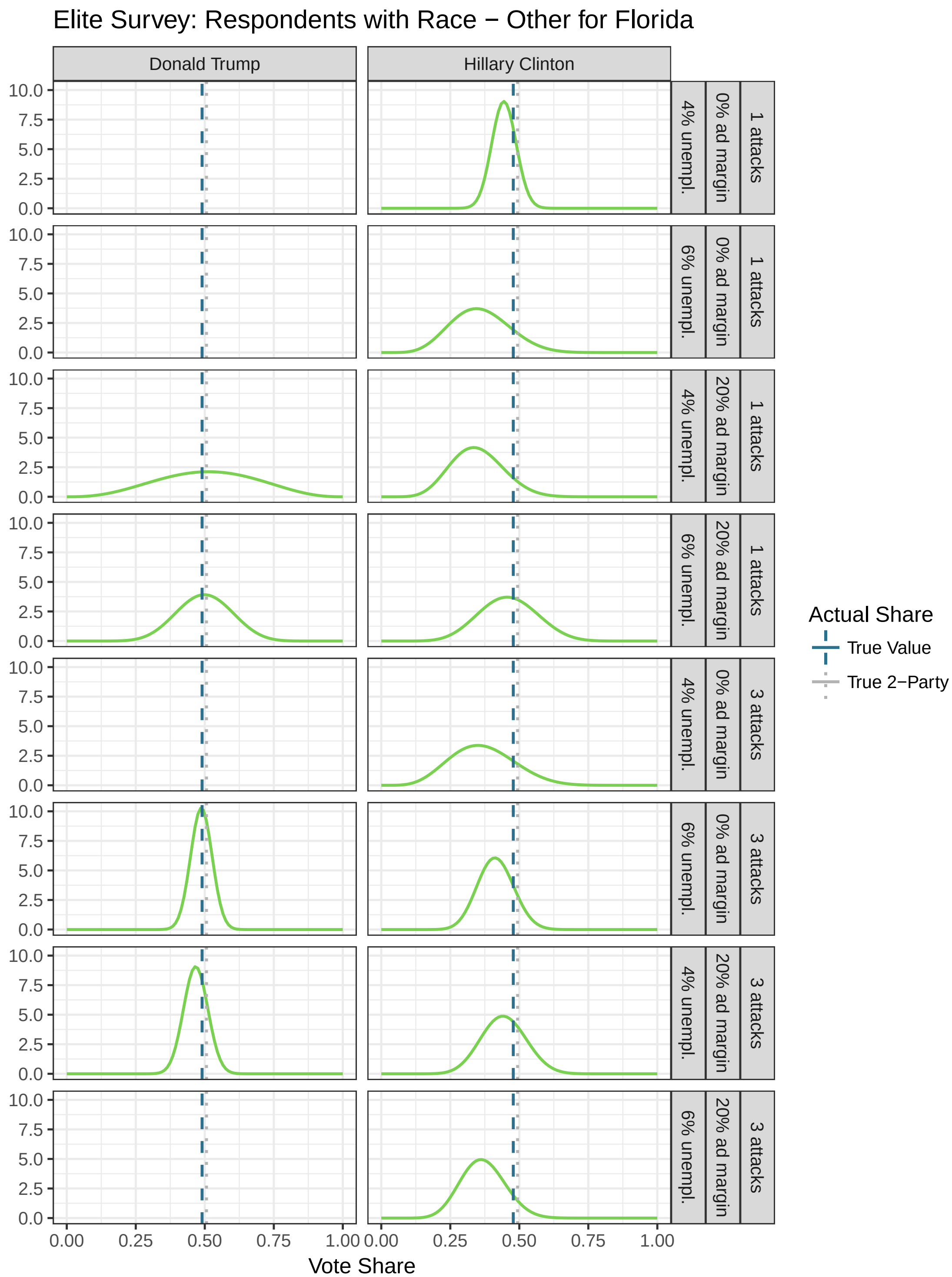


Figure 3.141: Priors with covariates: Elite Florida Race Other

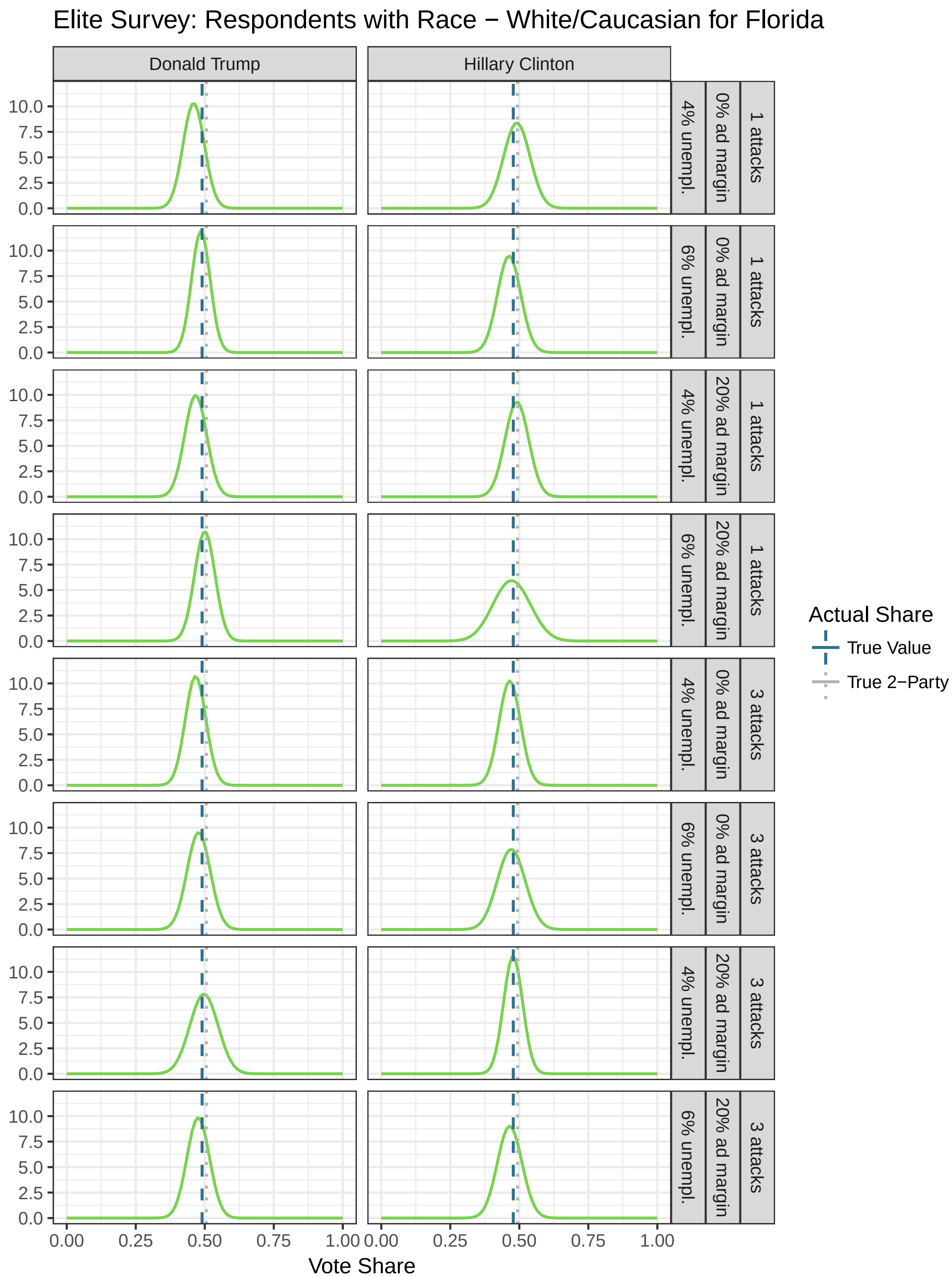


Figure 3.142: Priors with covariates: Elite Florida Race White Caucasian

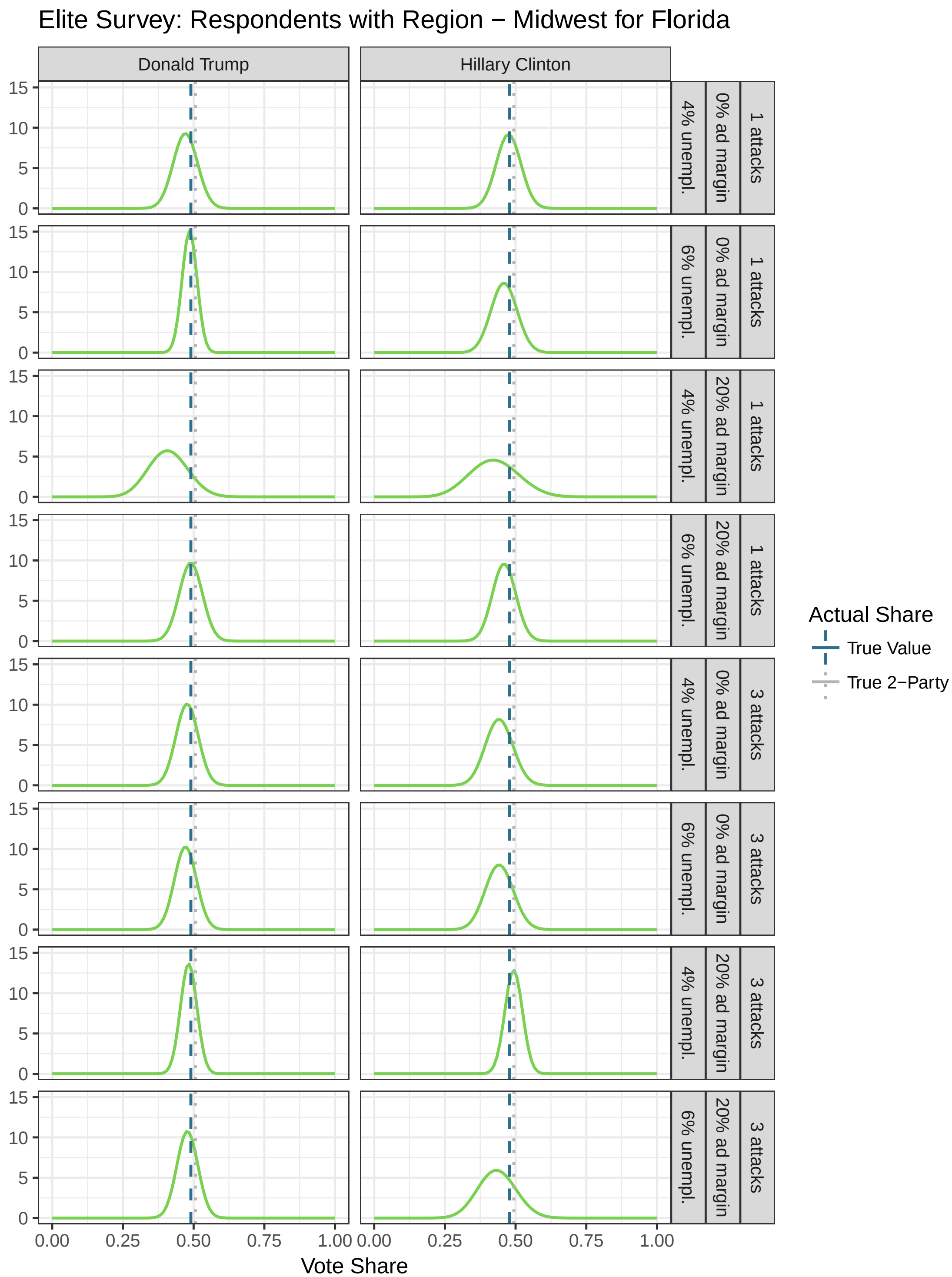


Figure 3.143: Priors with covariates: Elite Florida Region Midwest

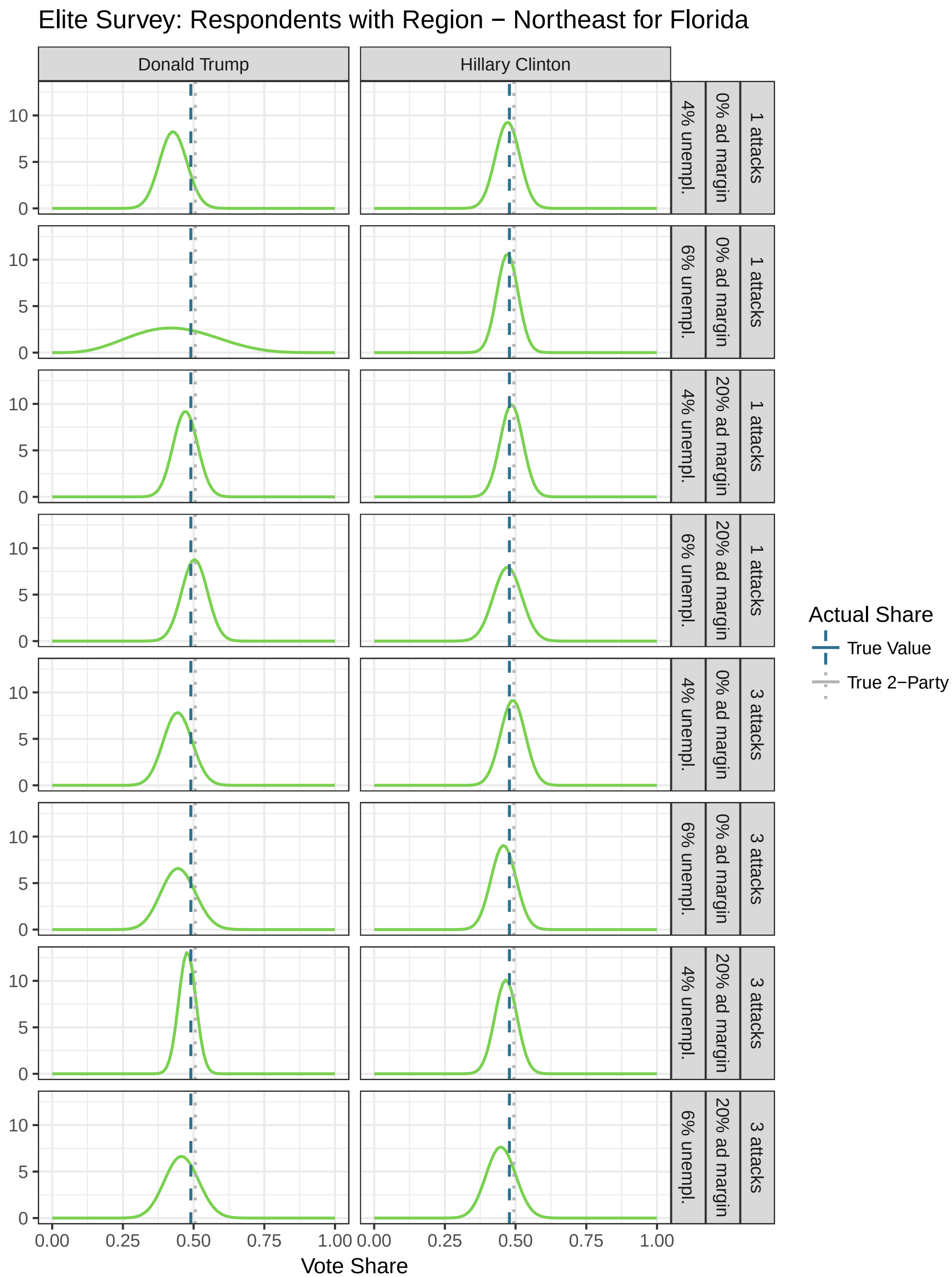


Figure 3.144: Priors with covariates: Elite Florida Region Northeast

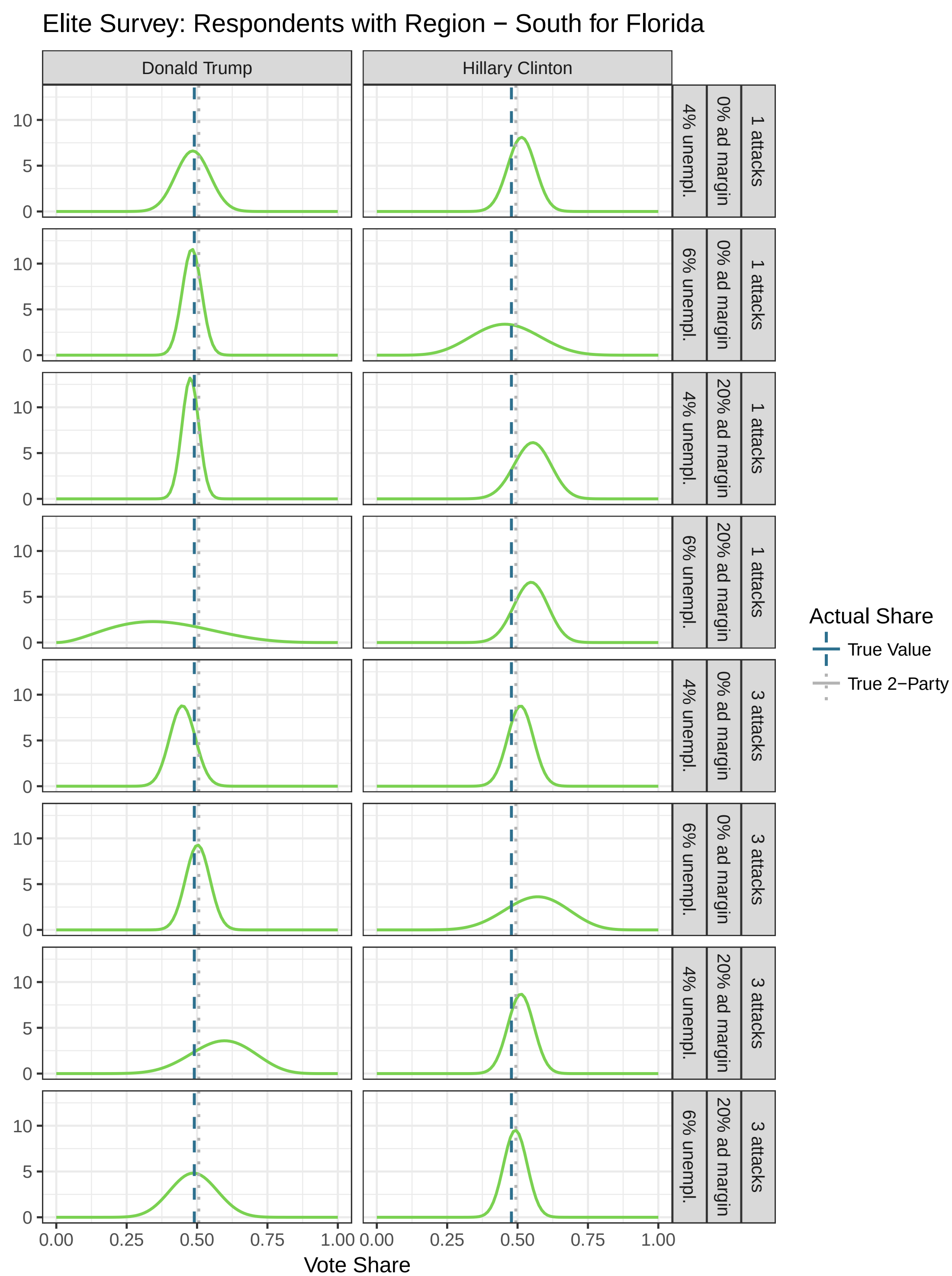


Figure 3.145: Priors with covariates: Elite Florida Region South

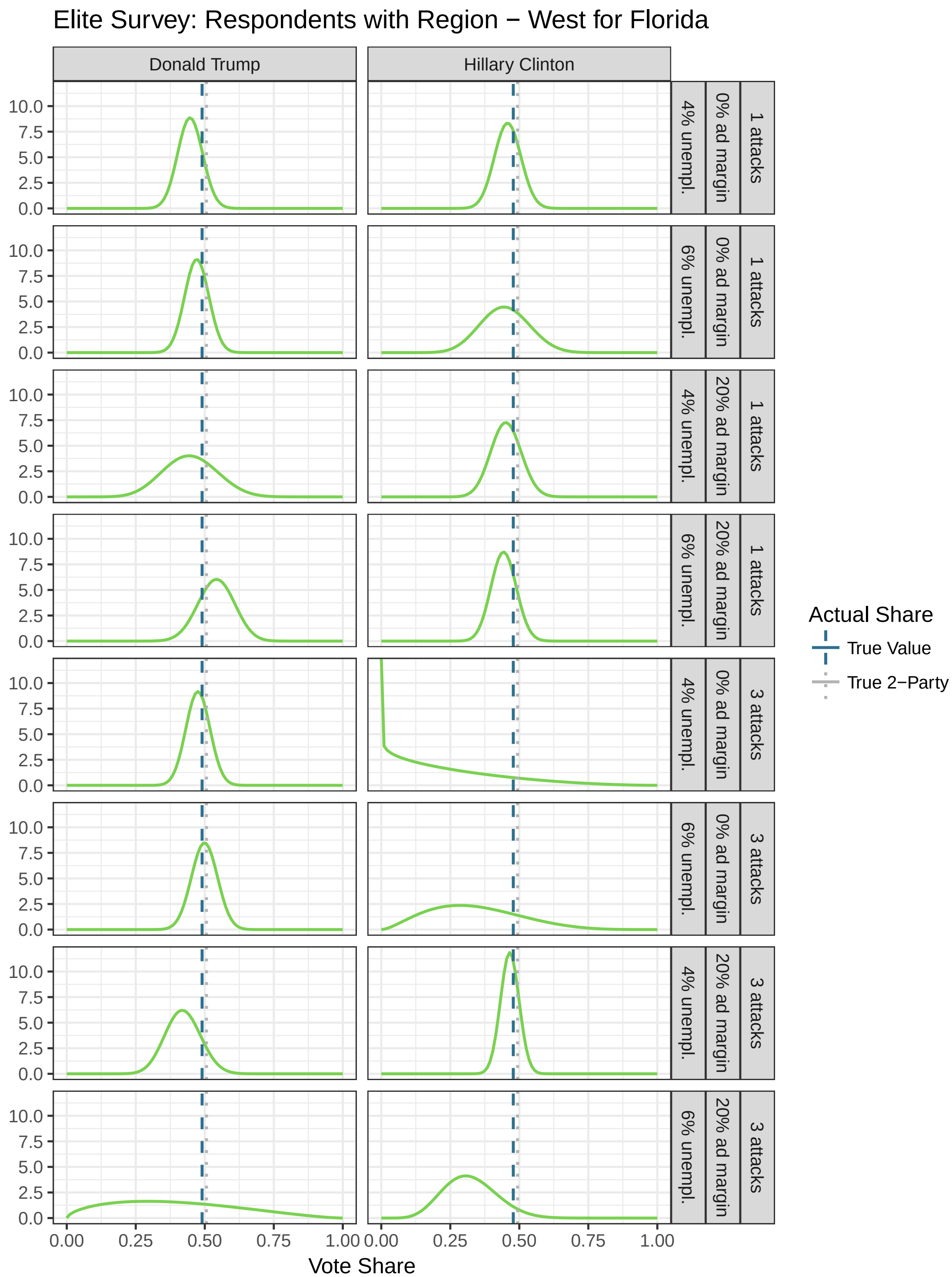


Figure 3.146: Priors with covariates: Elite Florida Region West

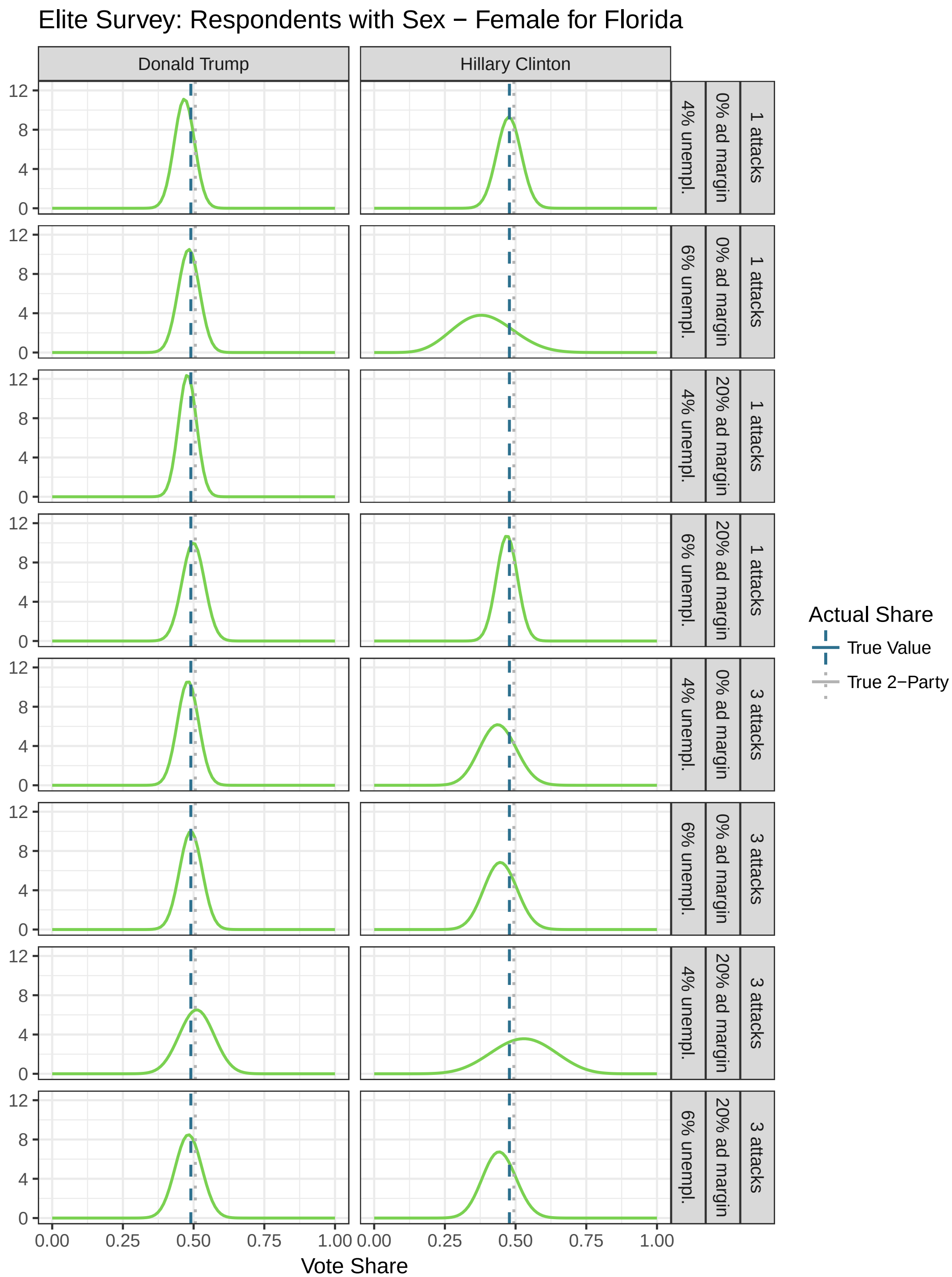


Figure 3.147: Priors with covariates: Elite Florida Sex Female

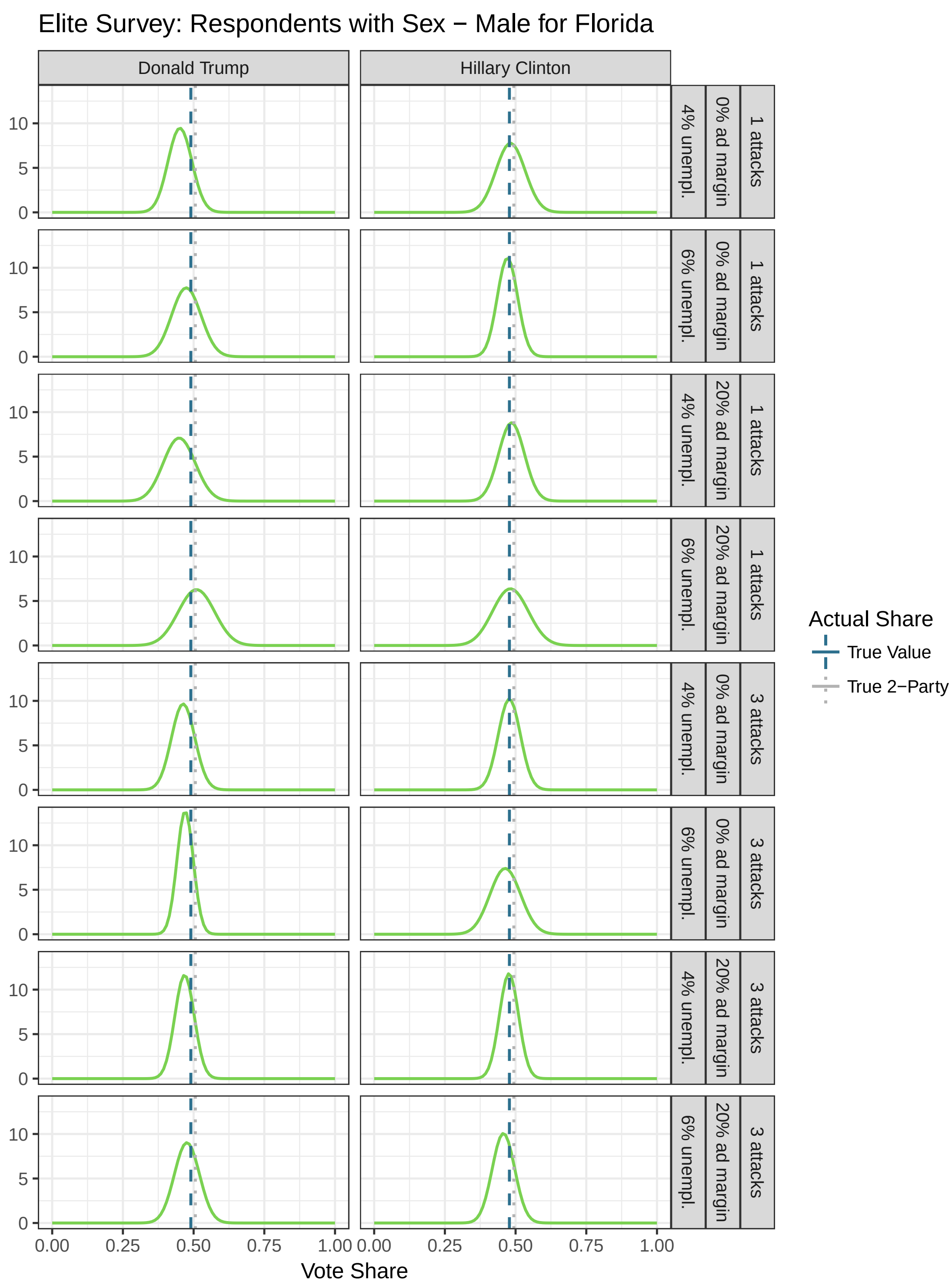


Figure 3.148: Priors with covariates: Elite Florida Sex Male

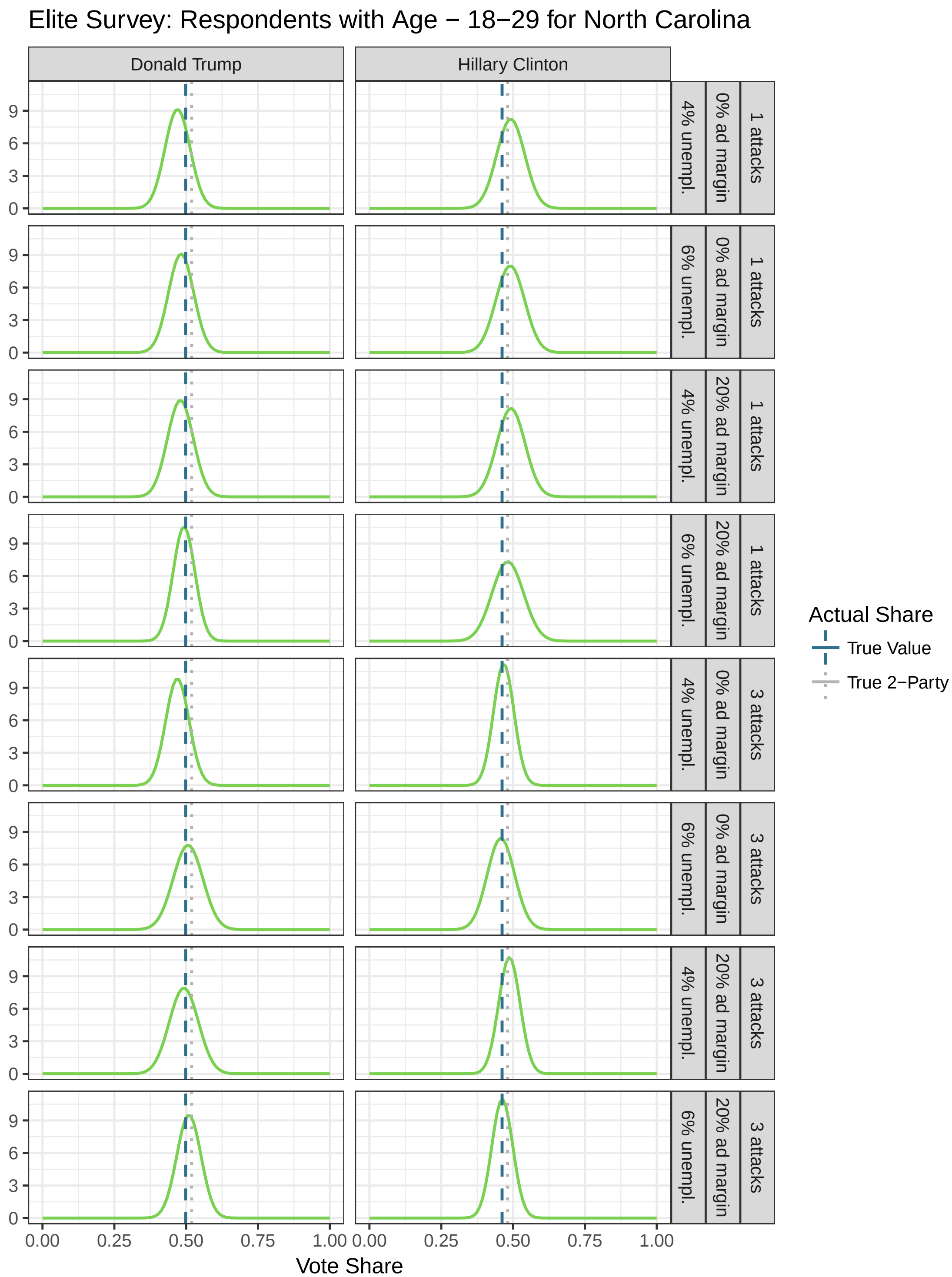


Figure 3.149: Priors with covariates: Elite North Carolina Age 18-29

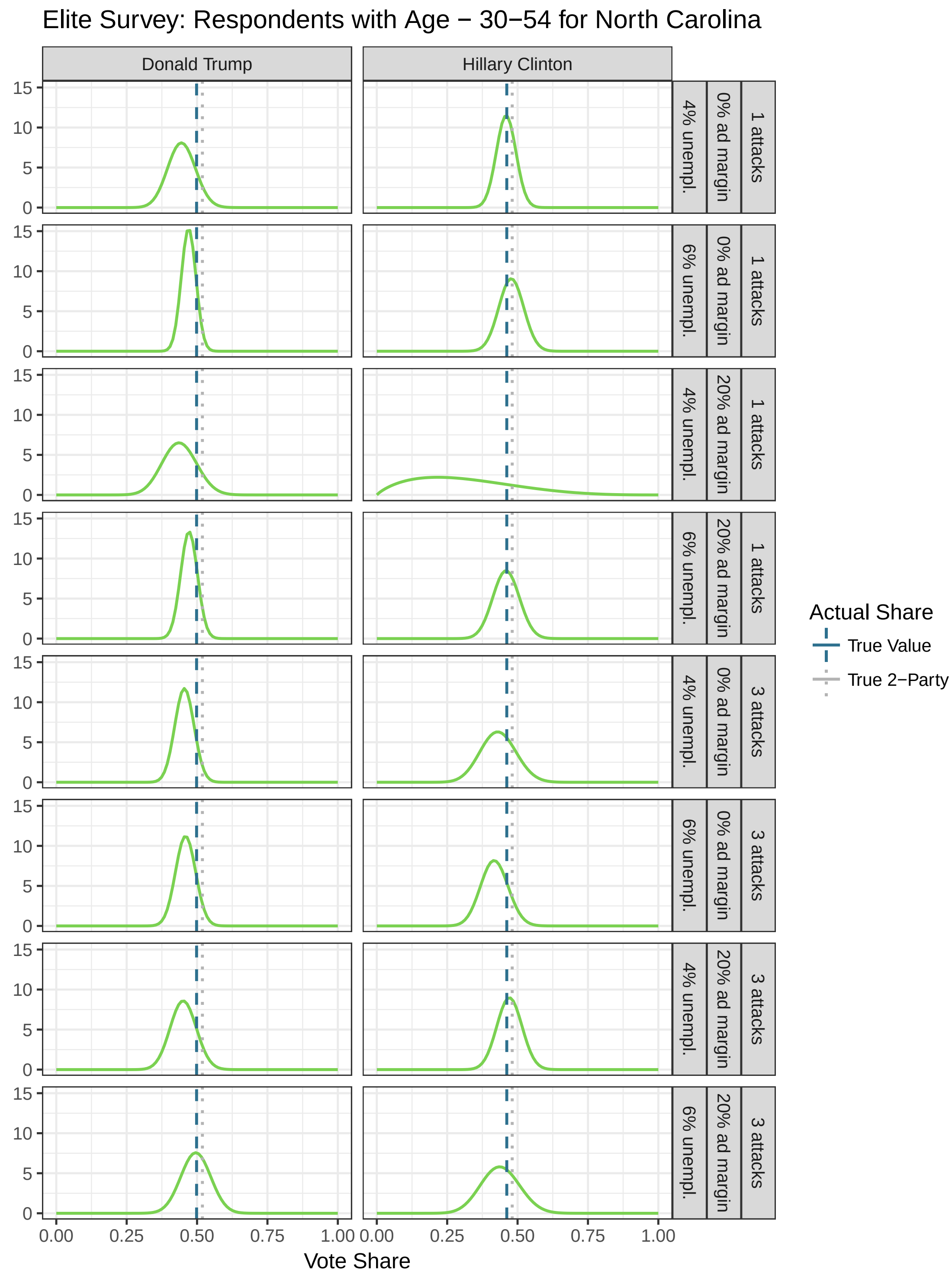


Figure 3.150: Priors with covariates: Elite North Carolina Age 30-54

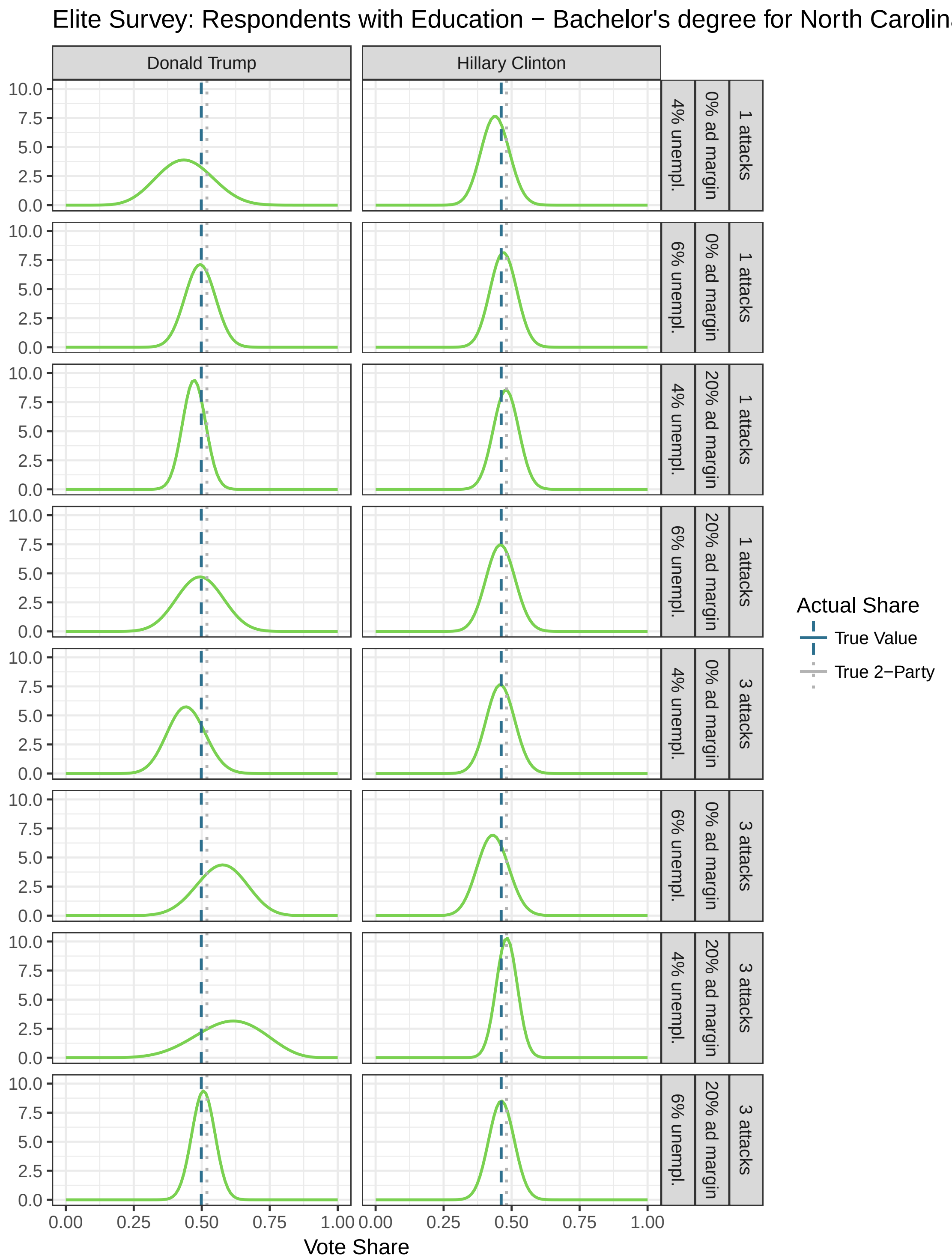


Figure 3.151: Priors with covariates: Elite North Carolina Education Bachelor's degree

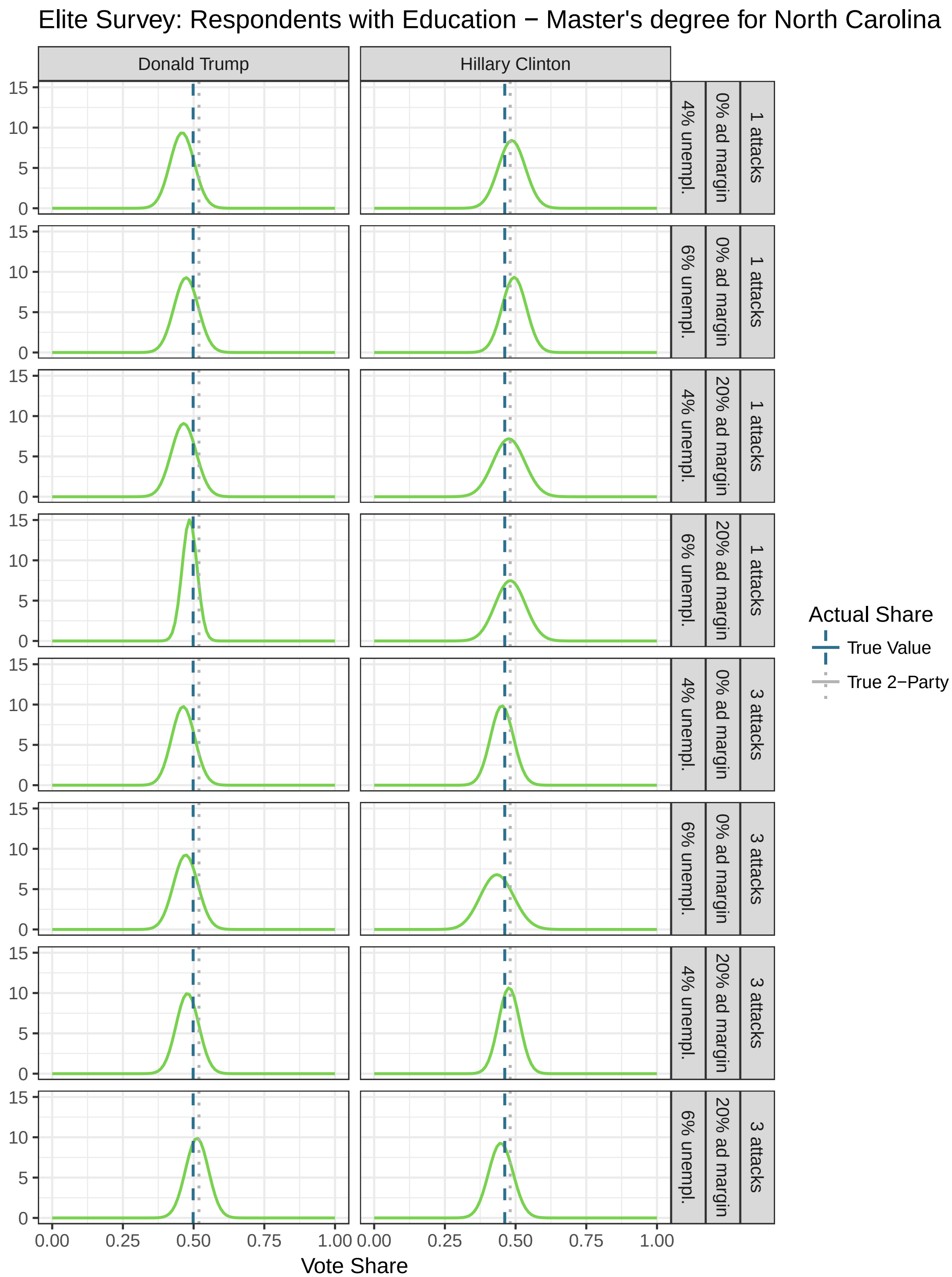


Figure 3.152: Priors with covariates: Elite North Carolina Education Master's degree

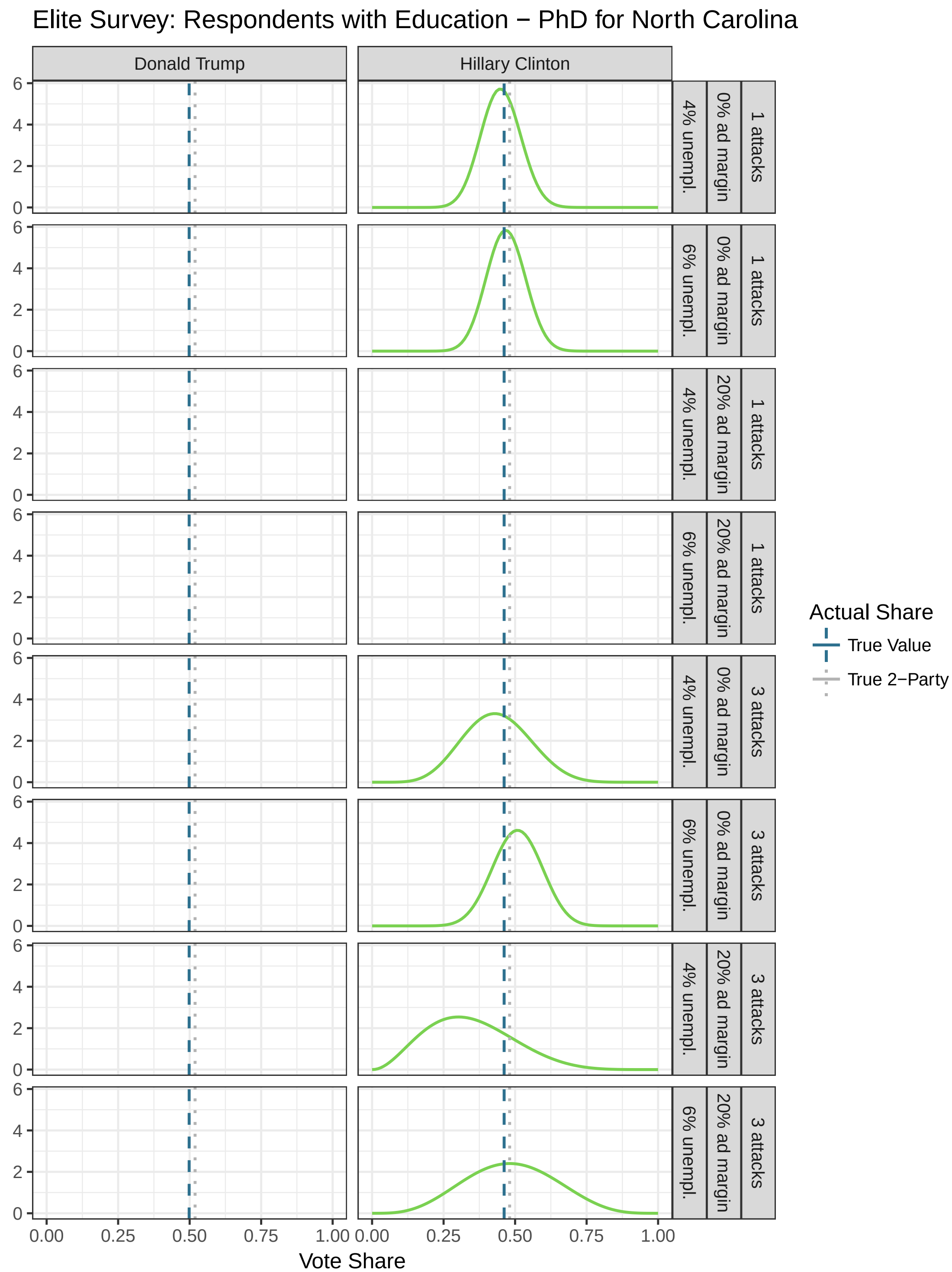


Figure 3.153: Priors with covariates: Elite North Carolina Education PhD

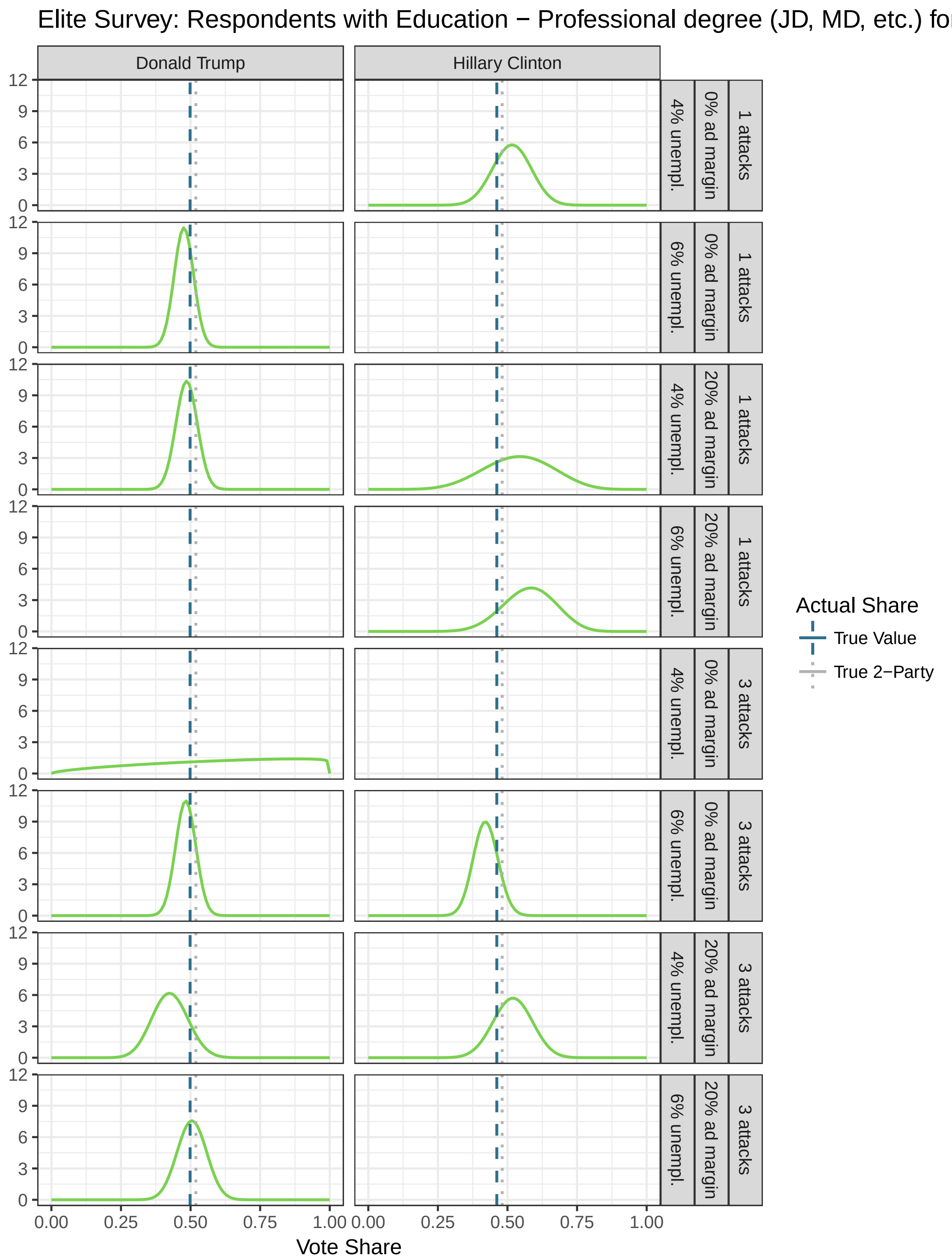


Figure 3.154: Priors with covariates: Elite North Carolina Education Professional degree JD MD etc

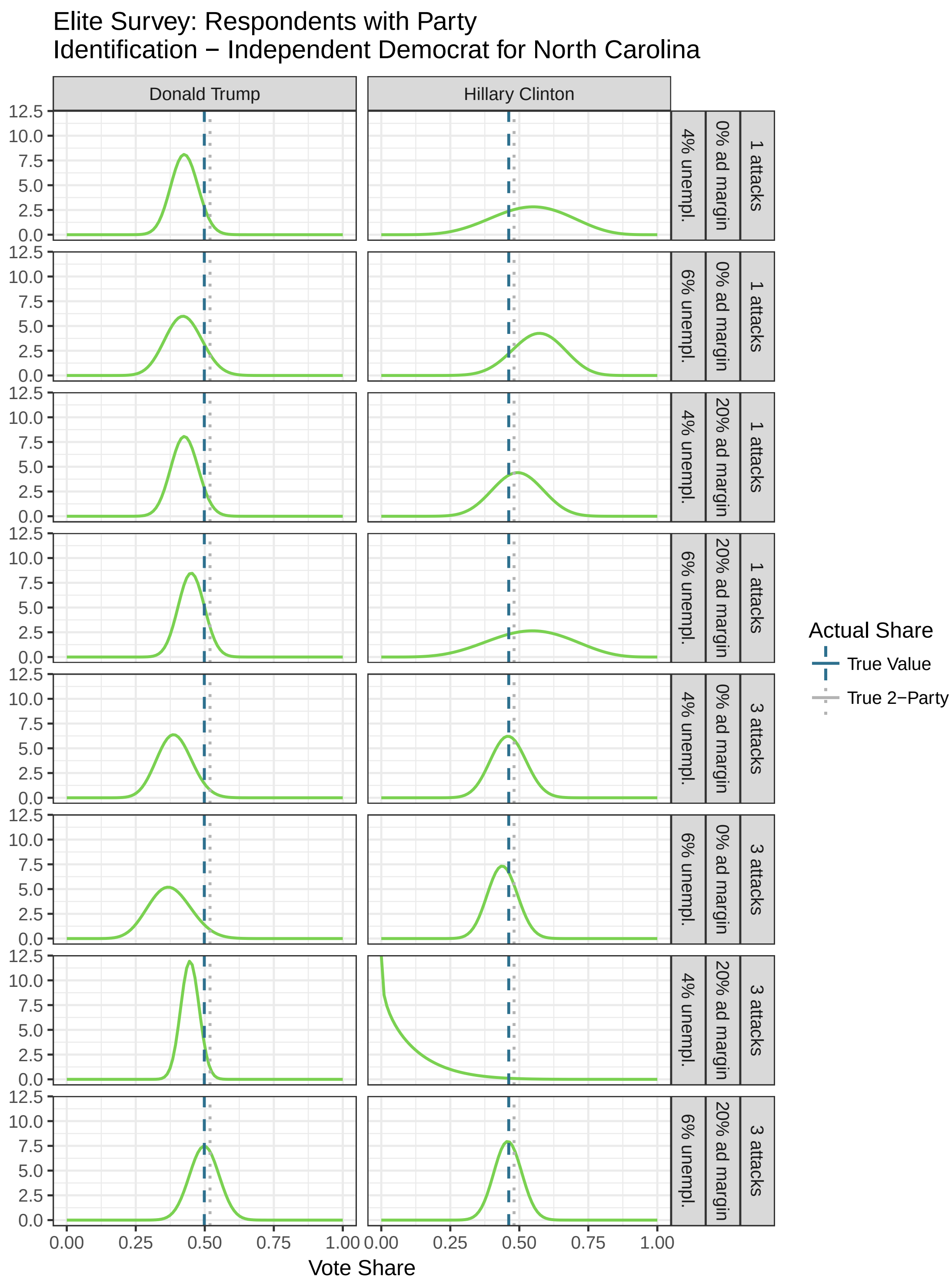


Figure 3.155: Priors with covariates: Elite North Carolina Party Identification Independent Democrat

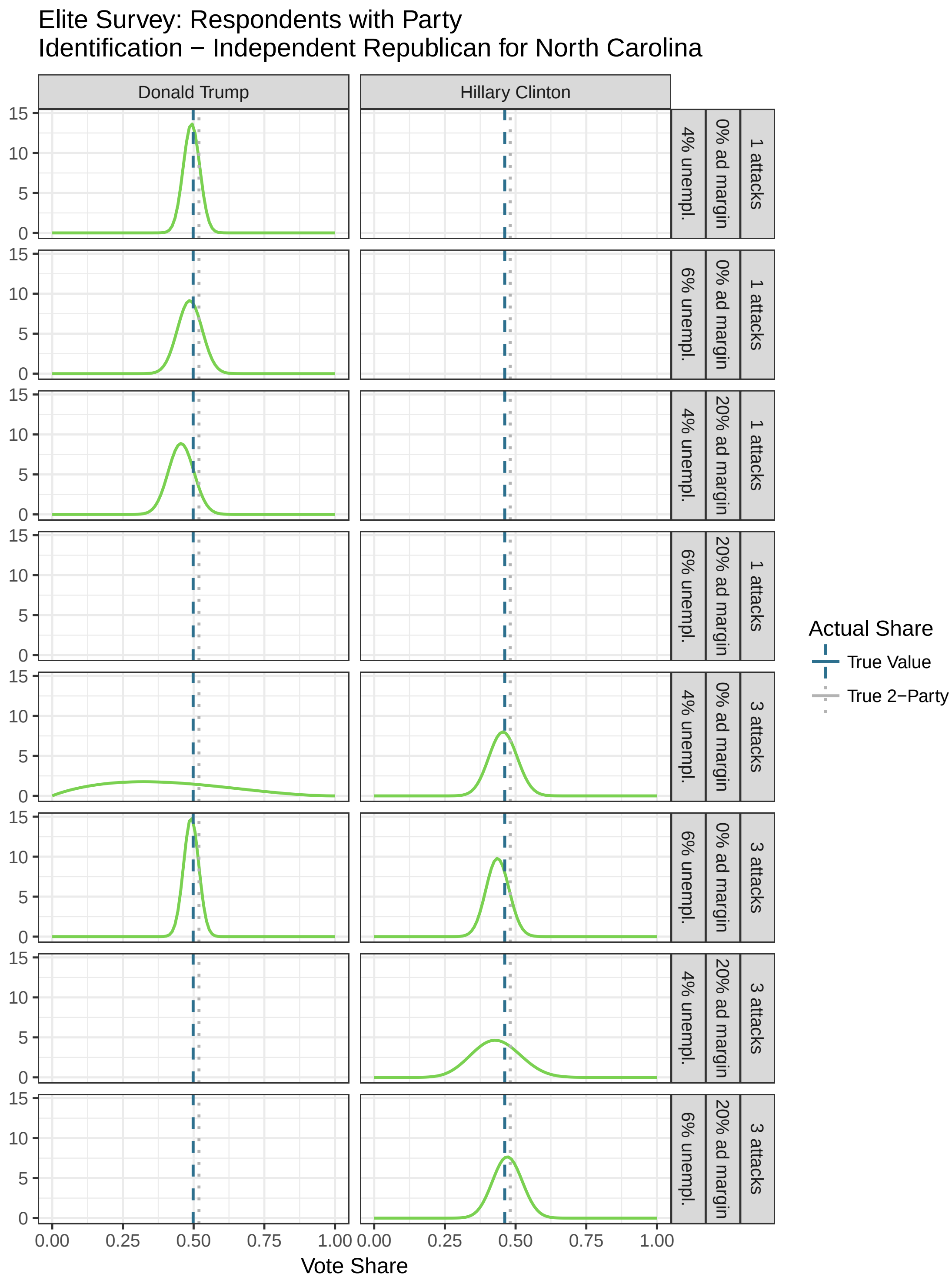


Figure 3.156: Priors with covariates: Elite North Carolina Party Identification Independent Republican

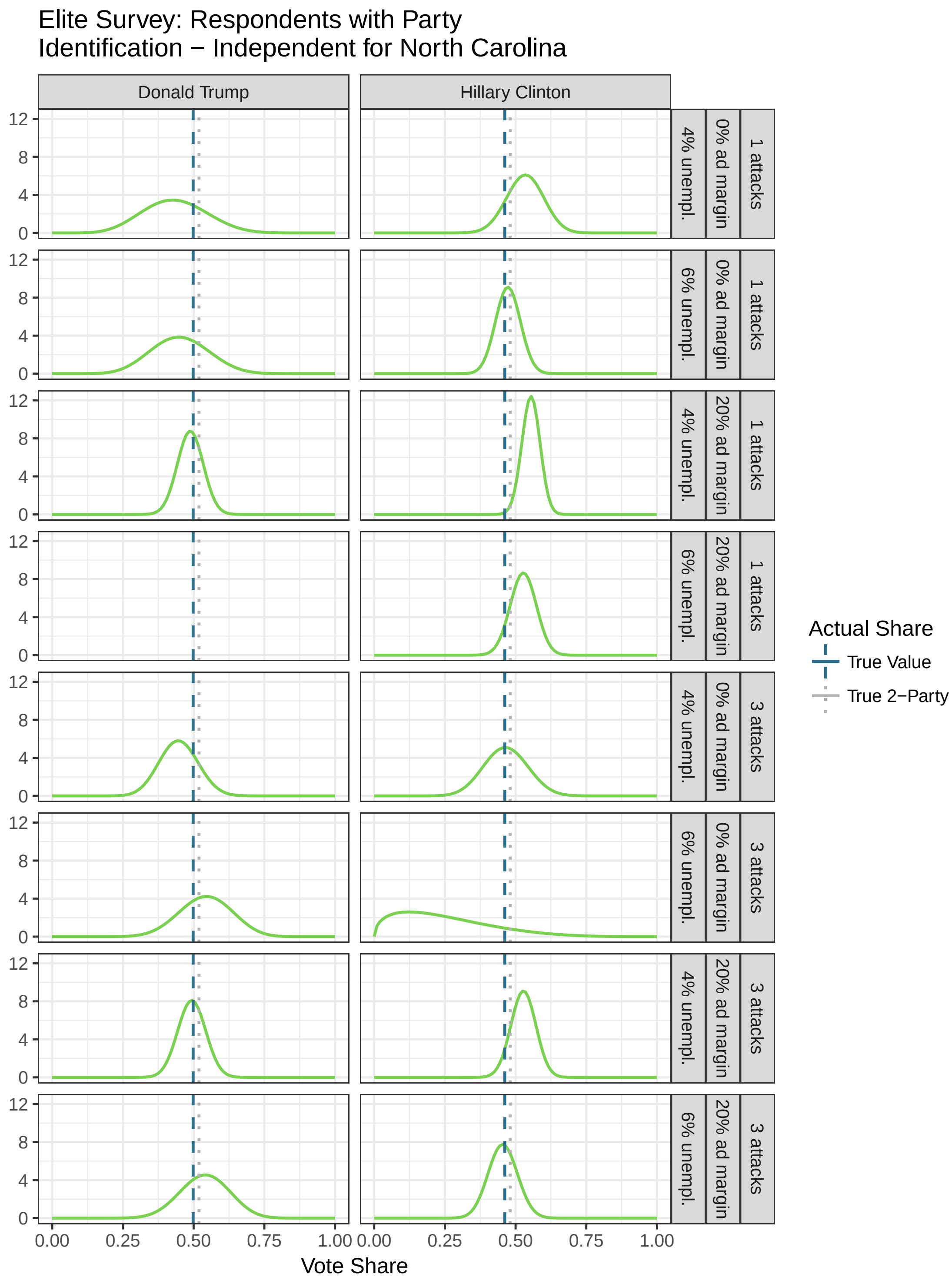


Figure 3.157: Priors with covariates: Elite North Carolina Party Identification Independent

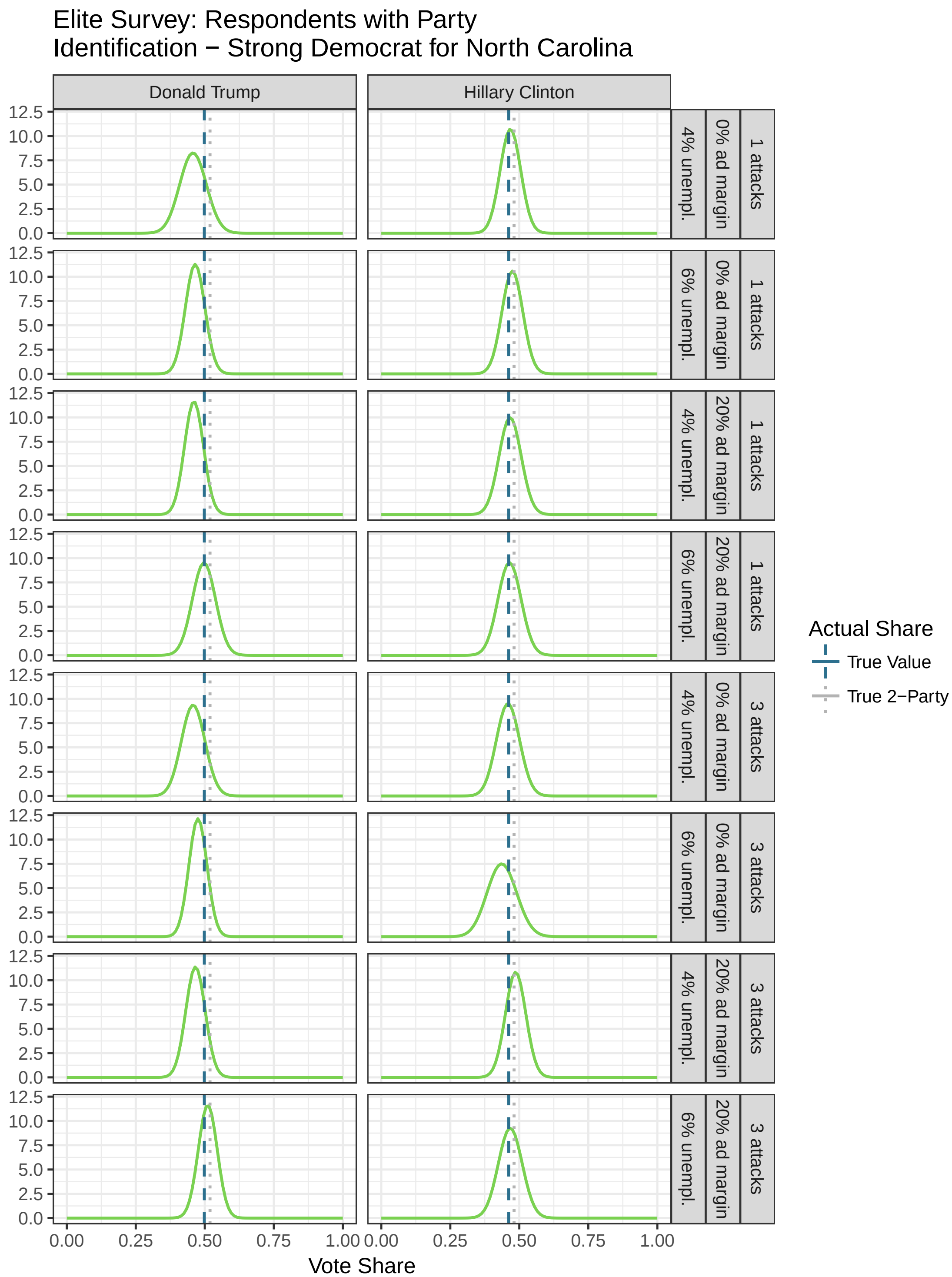


Figure 3.158: Priors with covariates: Elite North Carolina Party Identification Strong Democrat

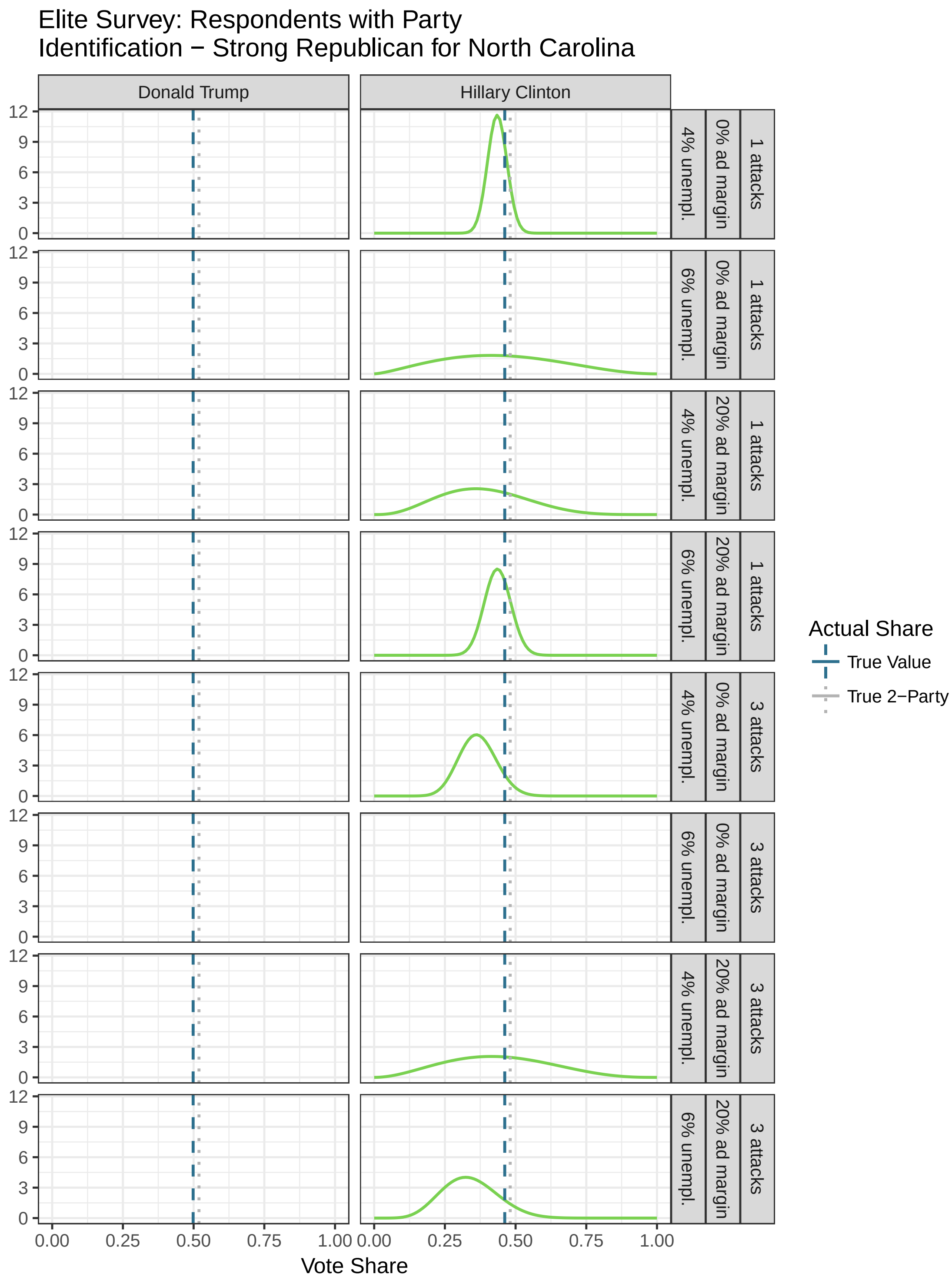


Figure 3.159: Priors with covariates: Elite North Carolina Party Identification Strong Republican

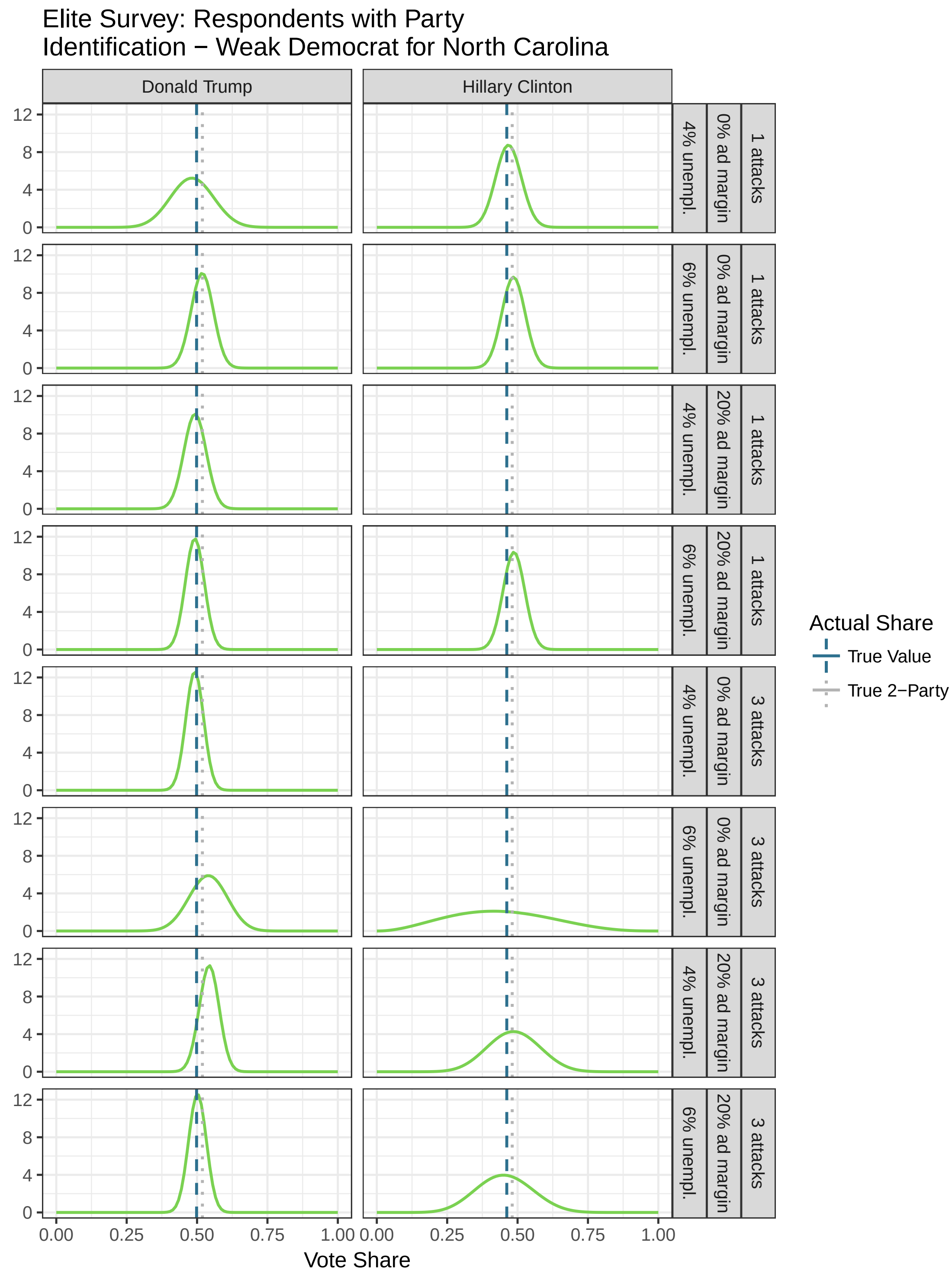


Figure 3.160: Priors with covariates: Elite North Carolina Party Identification Weak Democrat

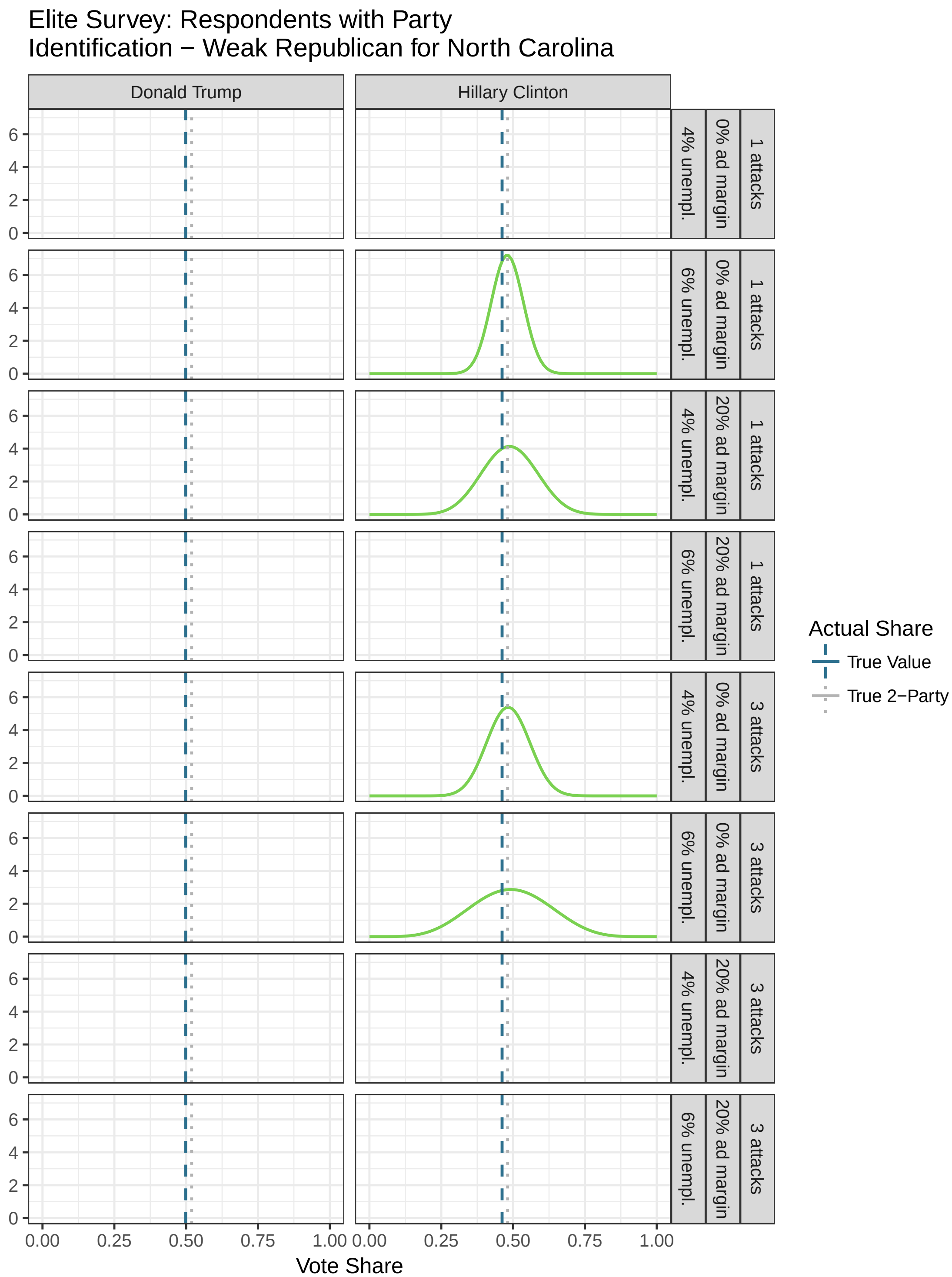


Figure 3.161: Priors with covariates: Elite North Carolina Party Identification Weak Republican

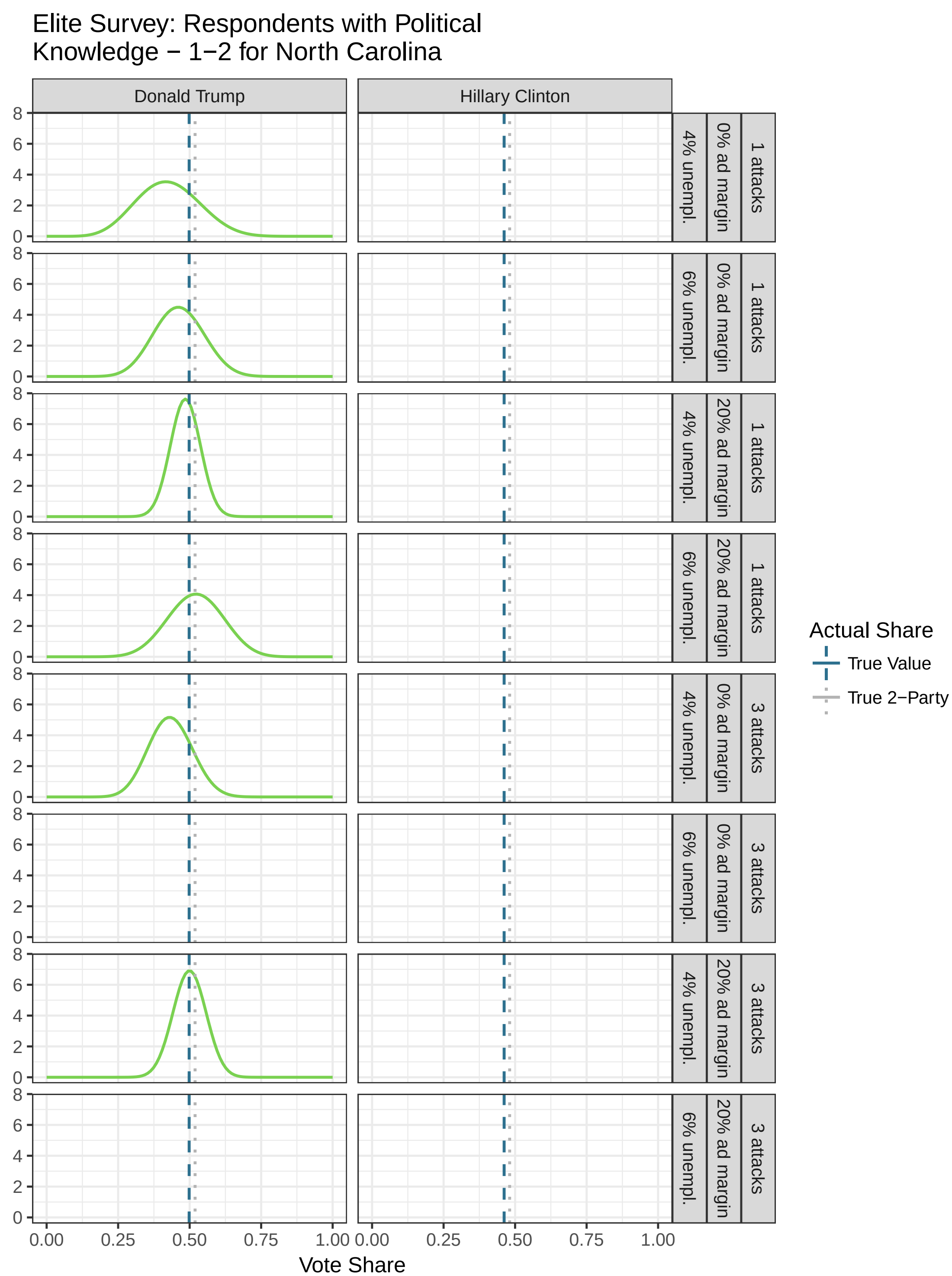


Figure 3.162: Priors with covariates: Elite North Carolina Political Knowledge 1-2

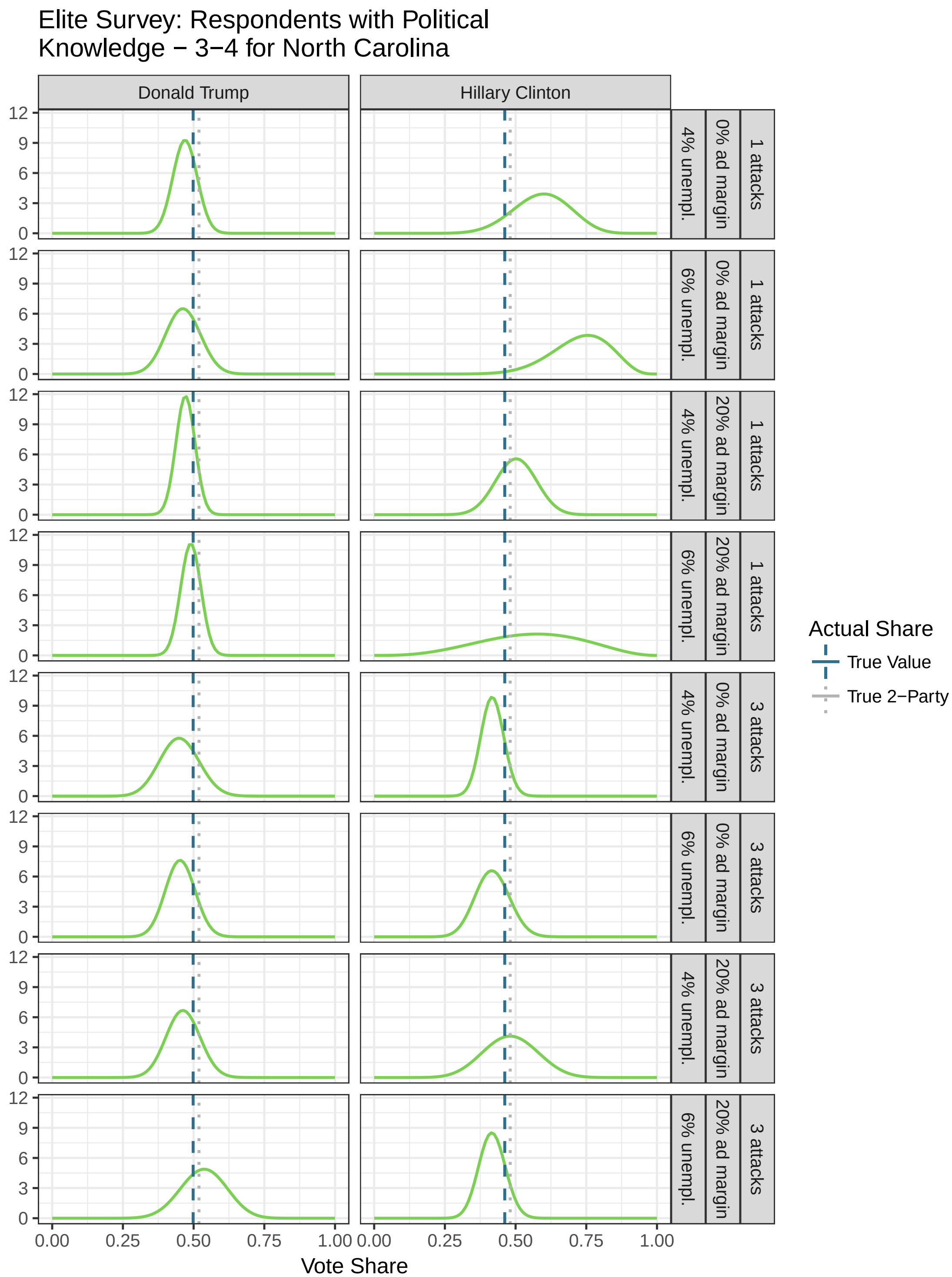


Figure 3.163: Priors with covariates: Elite North Carolina Political Knowledge 3-4

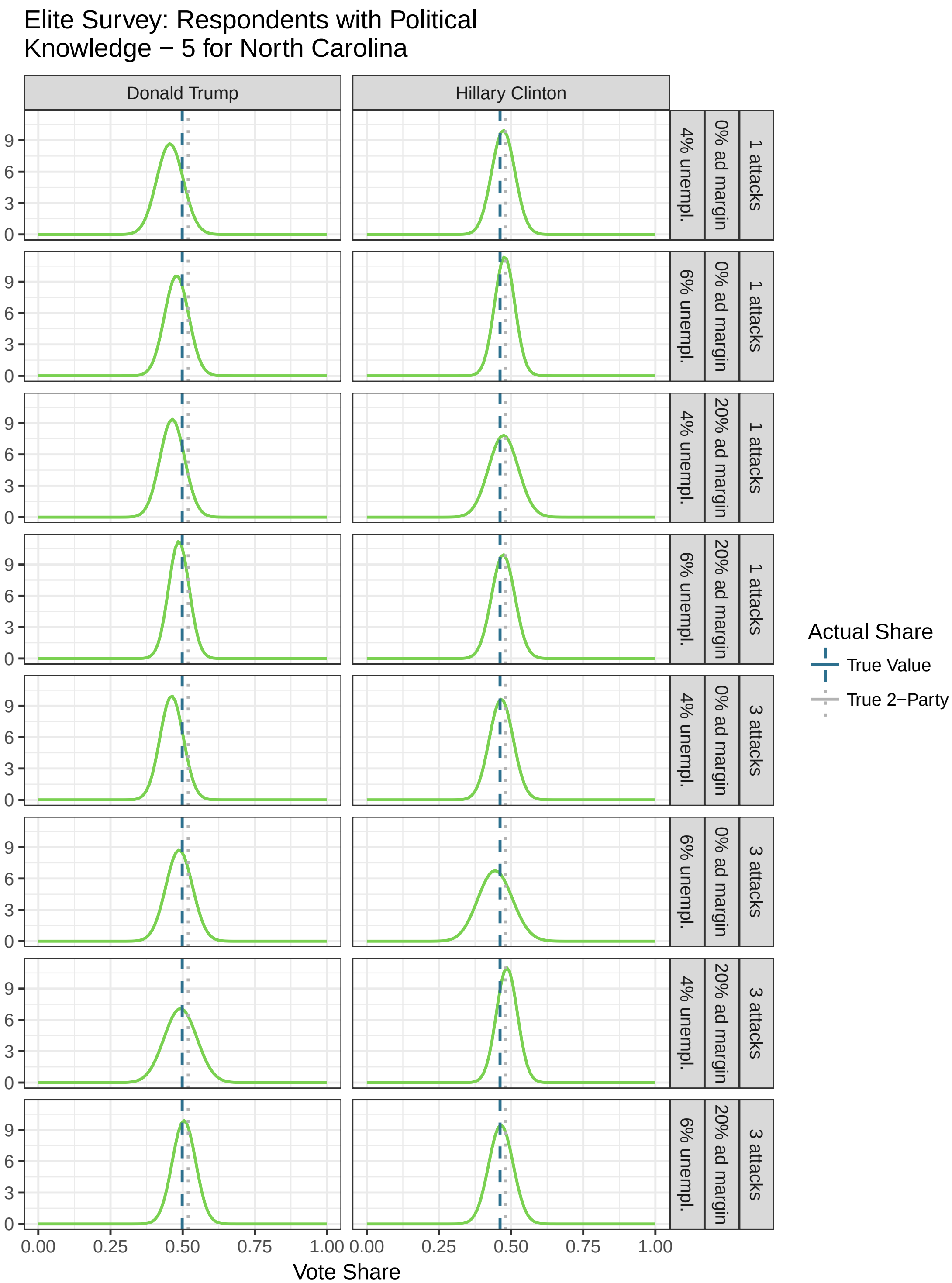


Figure 3.164: Priors with covariates: Elite North Carolina Political Knowledge 5

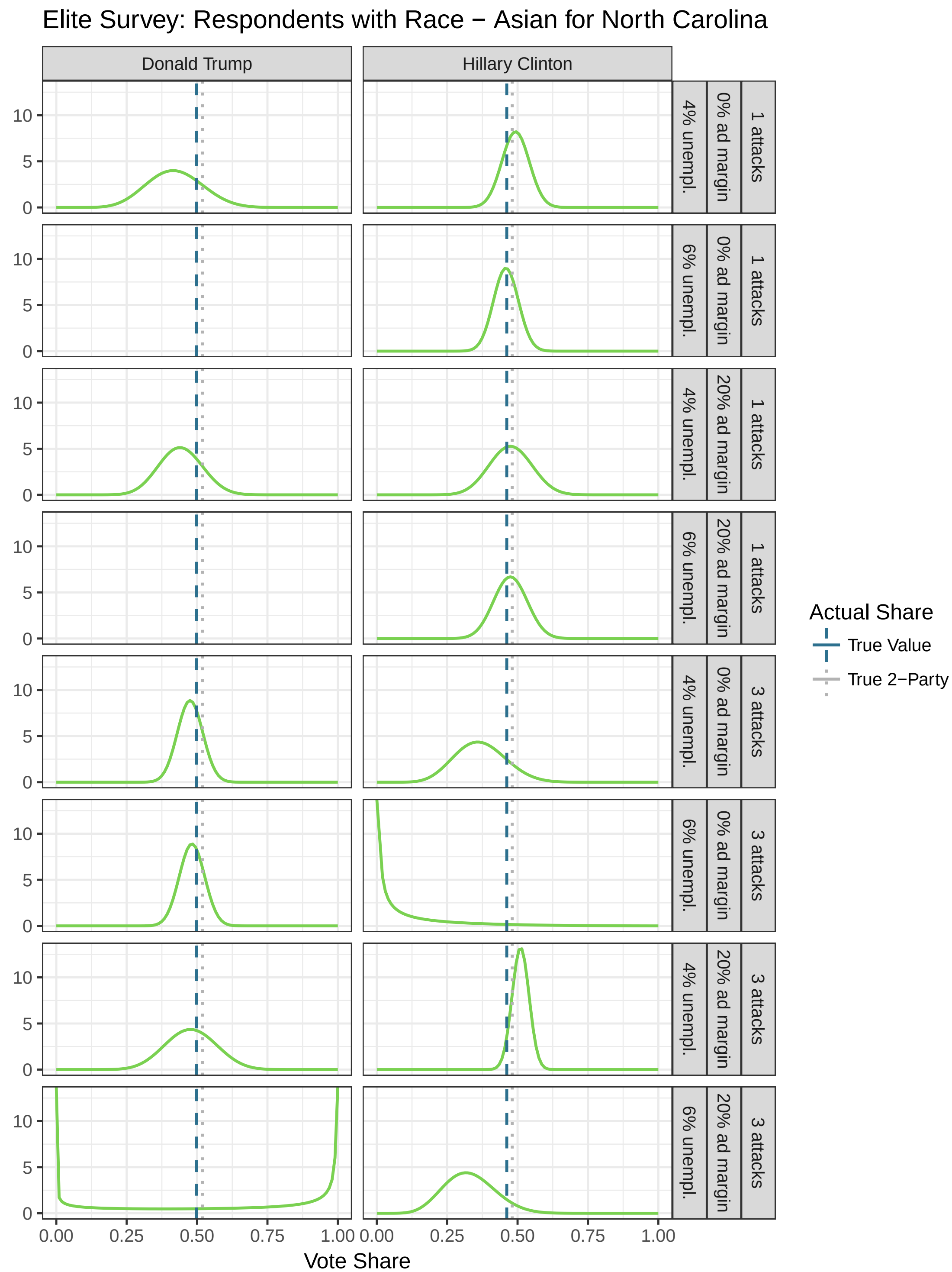


Figure 3.165: Priors with covariates: Elite North Carolina Race Asian

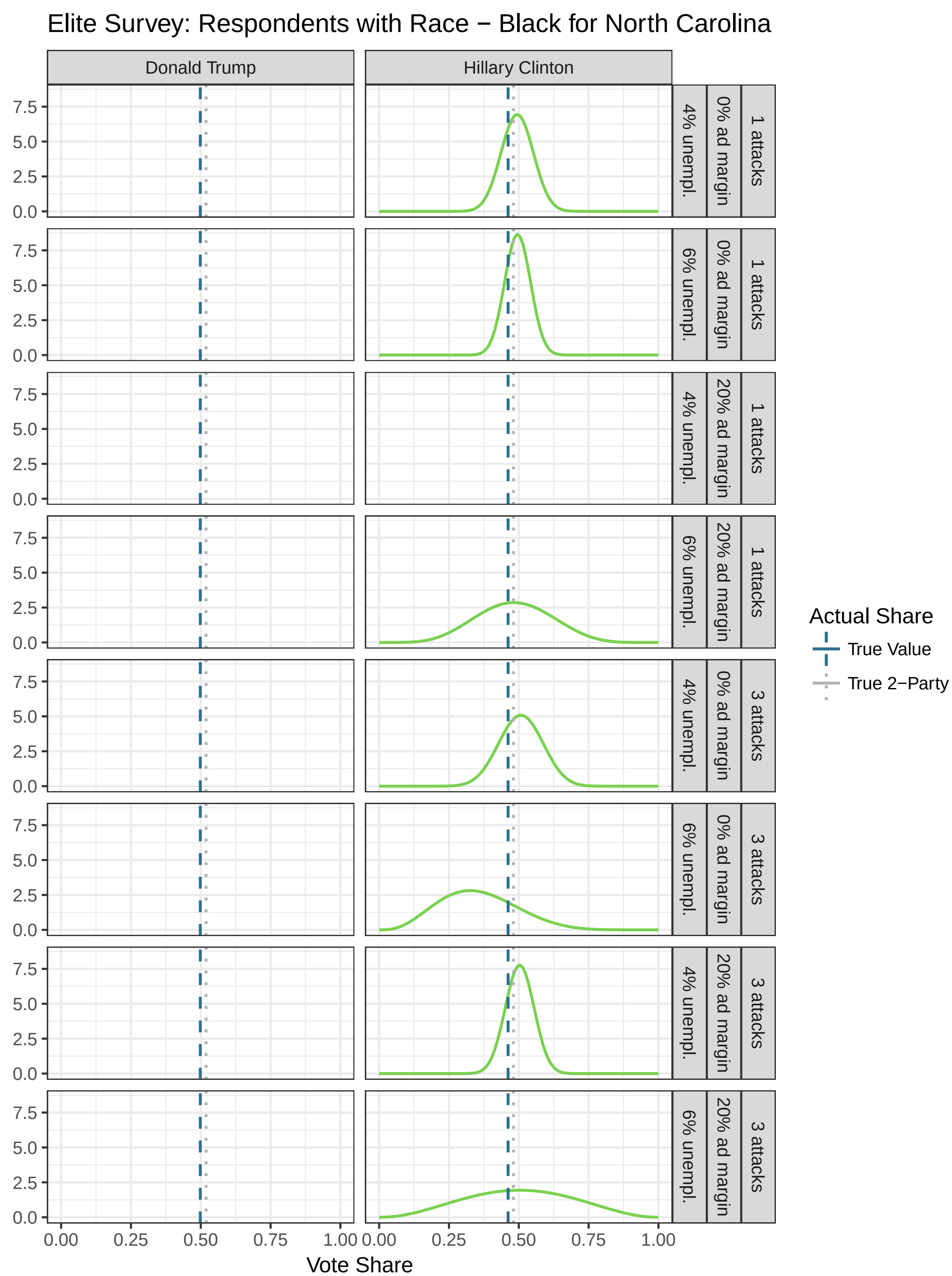


Figure 3.166: Priors with covariates: Elite North Carolina Race Black

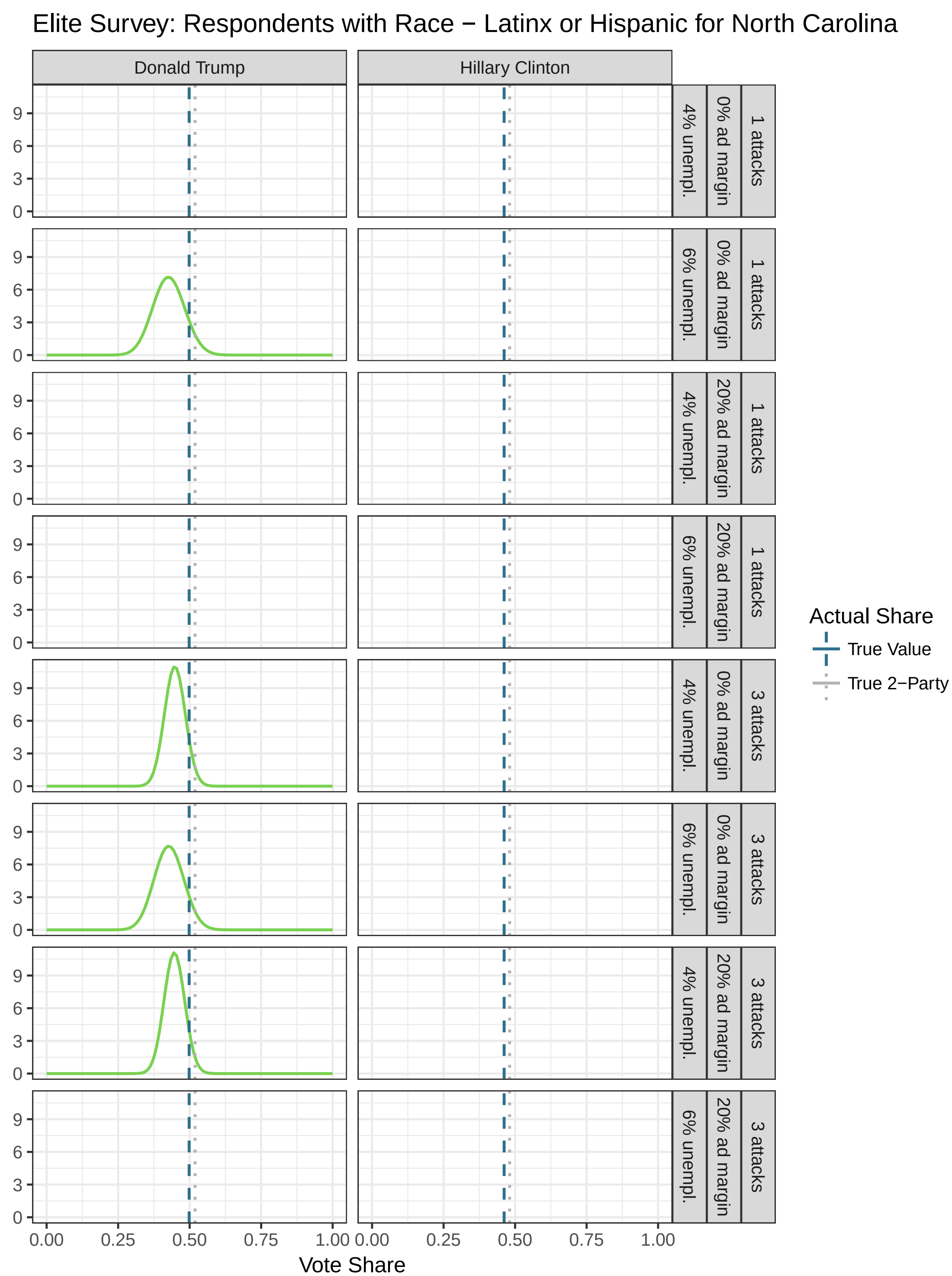


Figure 3.167: Priors with covariates: Elite North Carolina Race Latinx or Hispanic

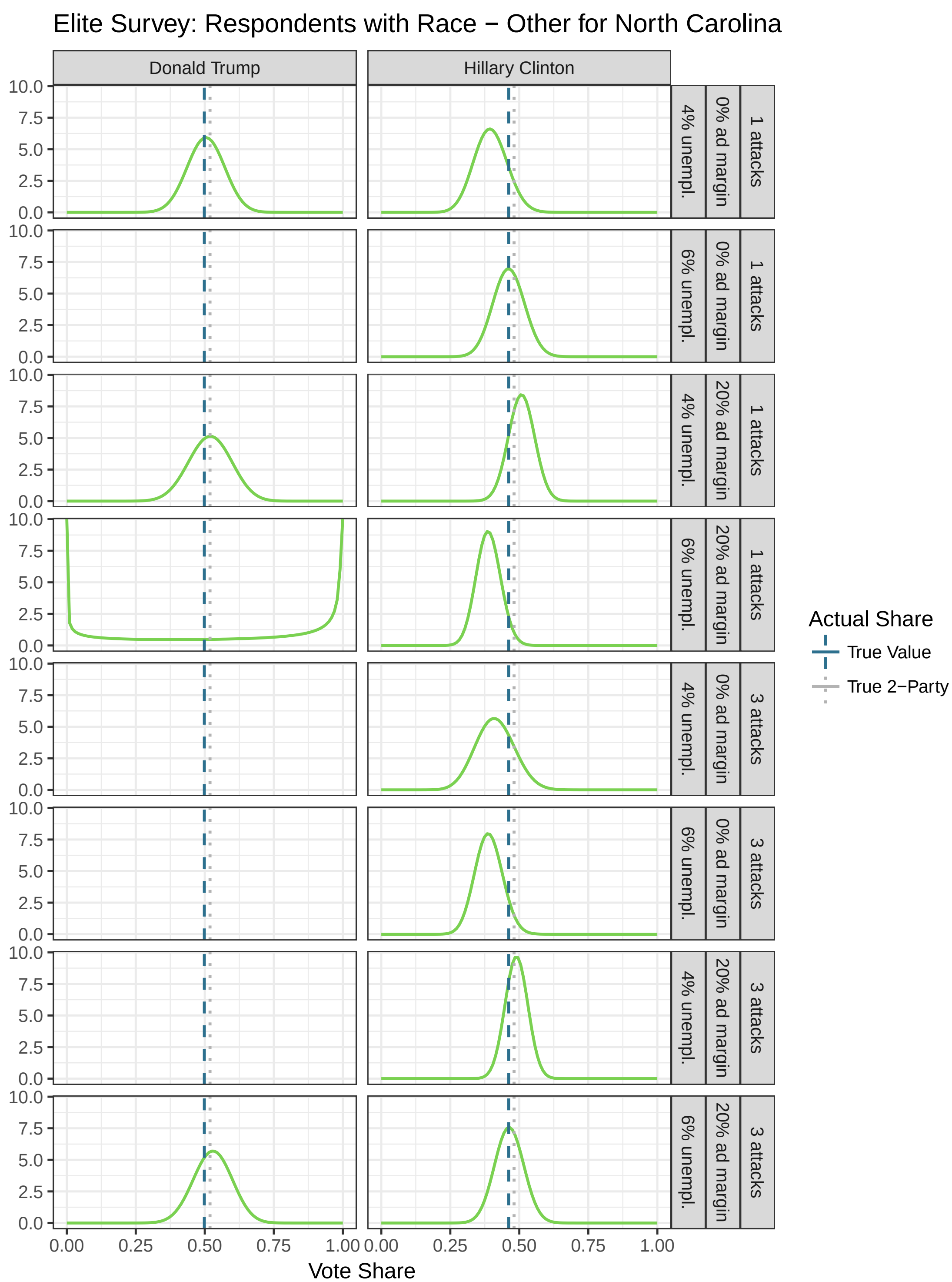


Figure 3.168: Priors with covariates: Elite North Carolina Race Other

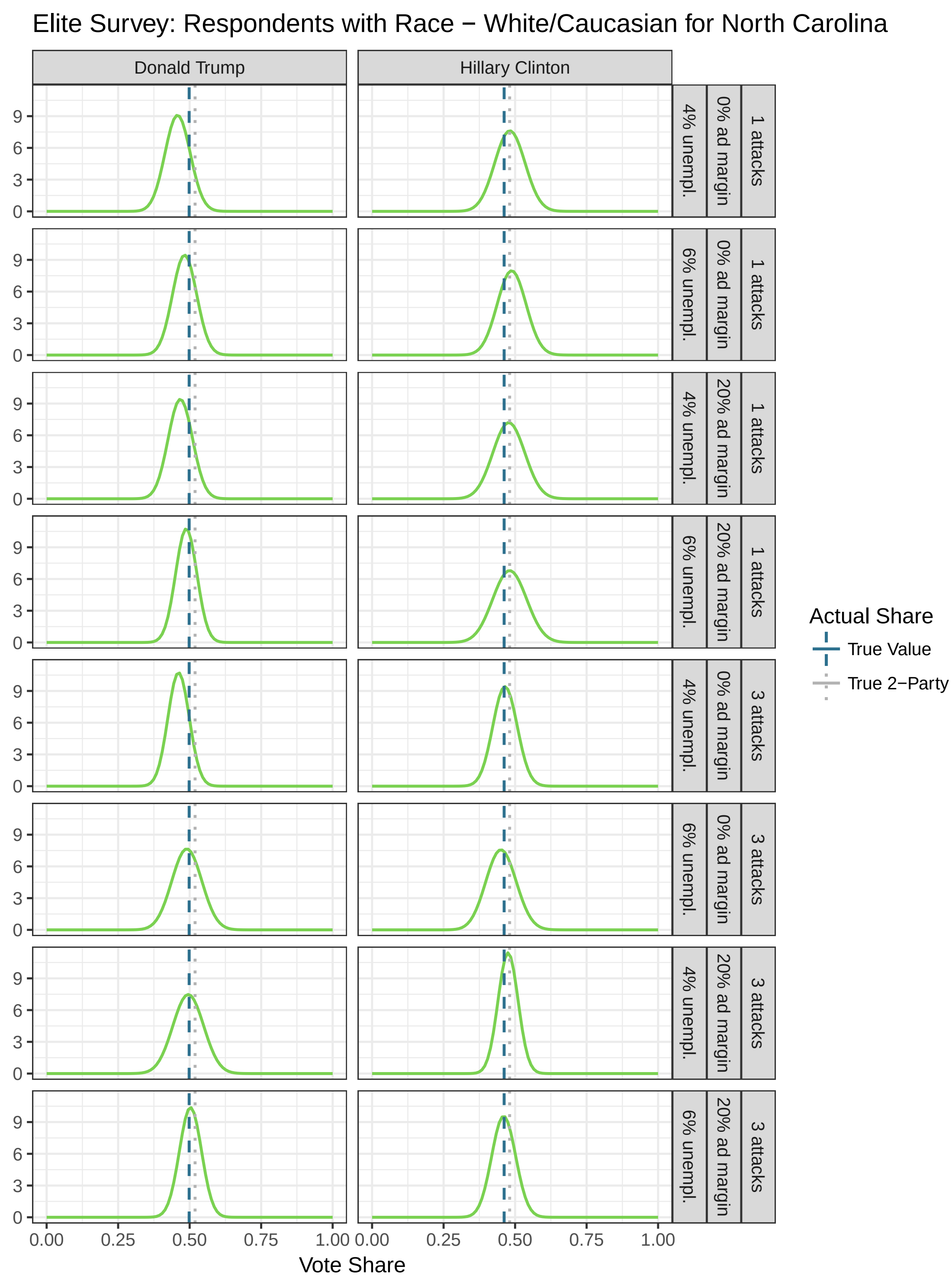


Figure 3.169: Priors with covariates: Elite North Carolina Race White Caucasian

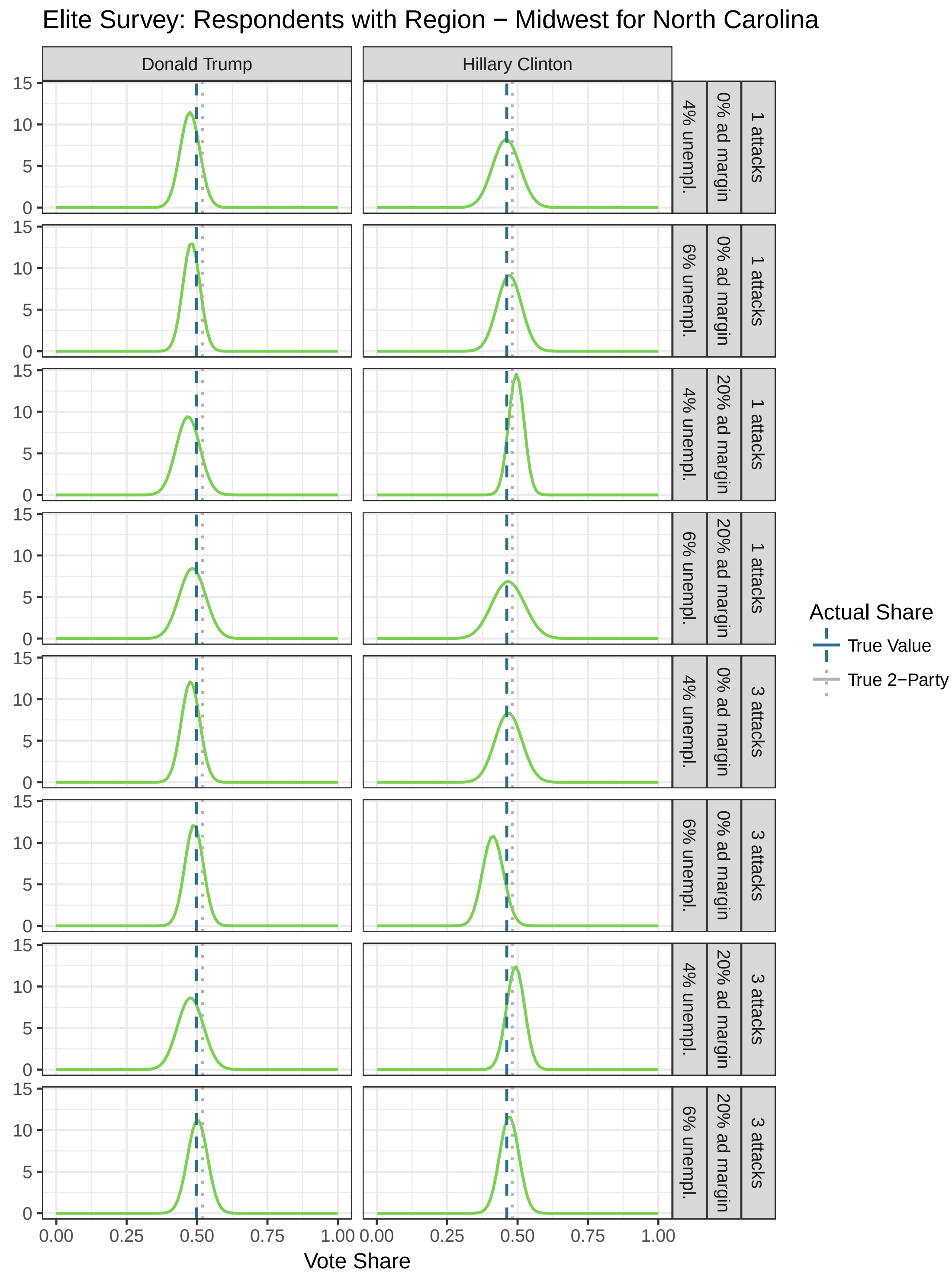


Figure 3.170: Priors with covariates: Elite North Carolina Region Midwest

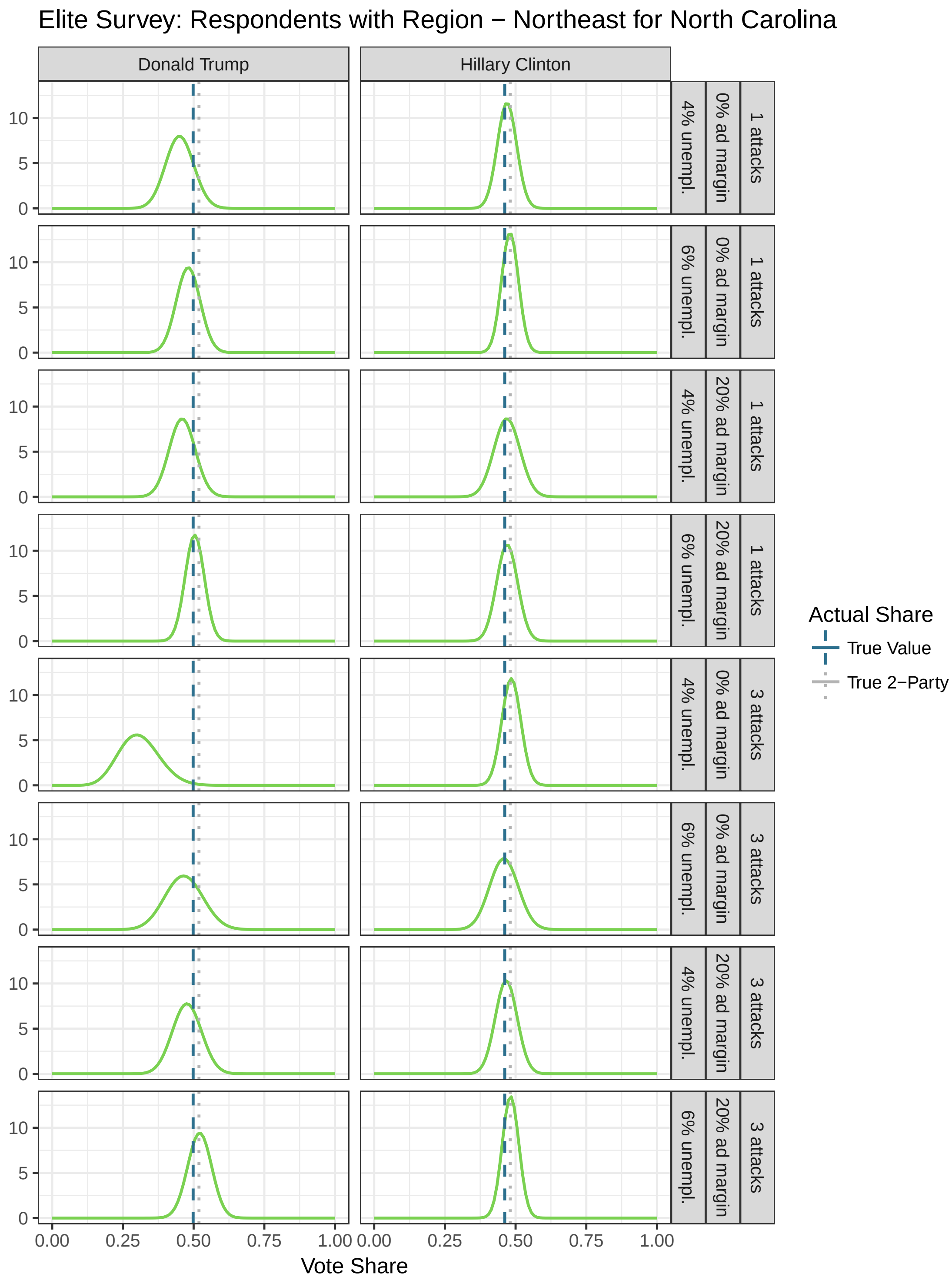


Figure 3.171: Priors with covariates: Elite North Carolina Region Northeast

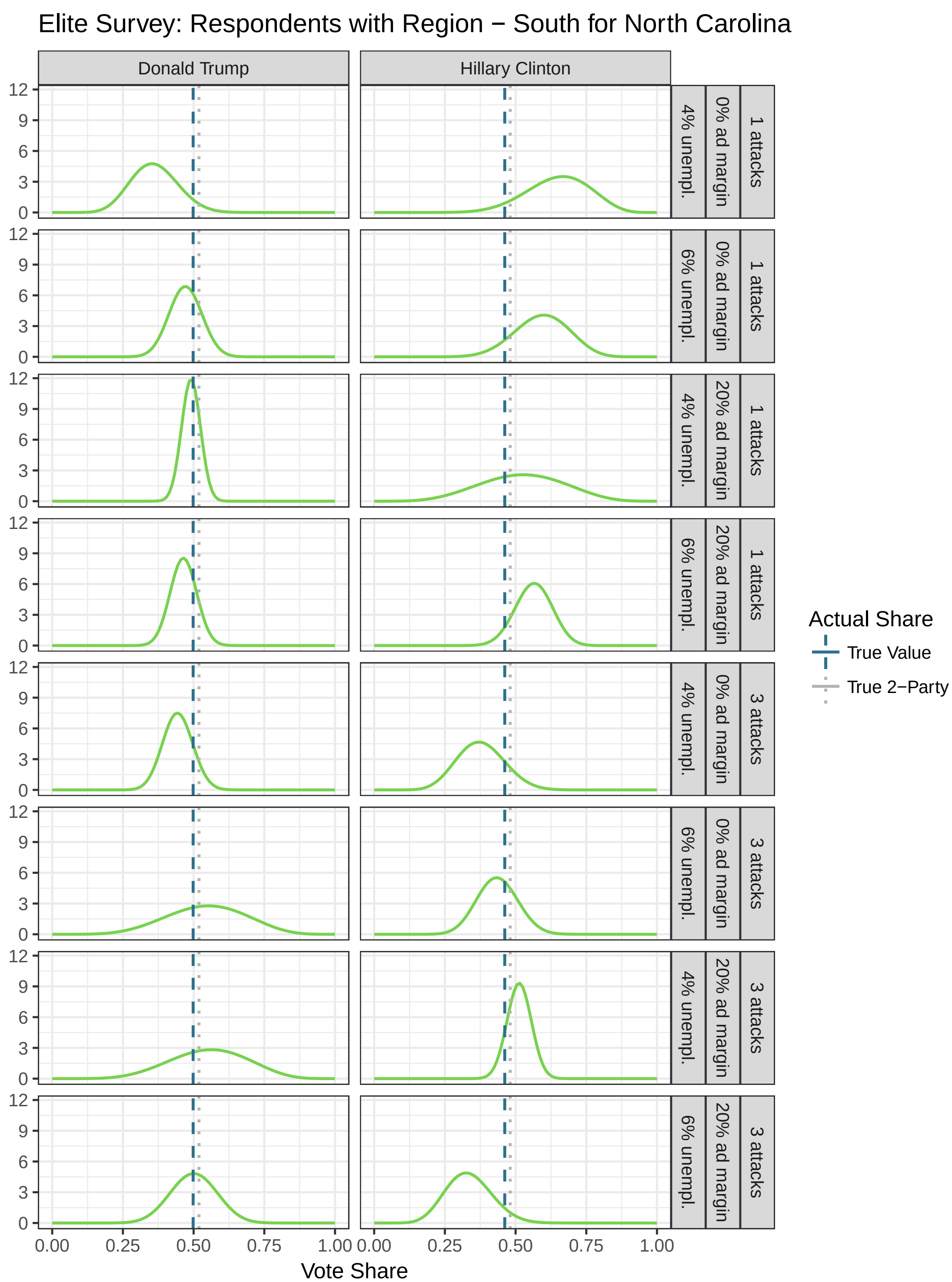


Figure 3.172: Priors with covariates: Elite North Carolina Region South

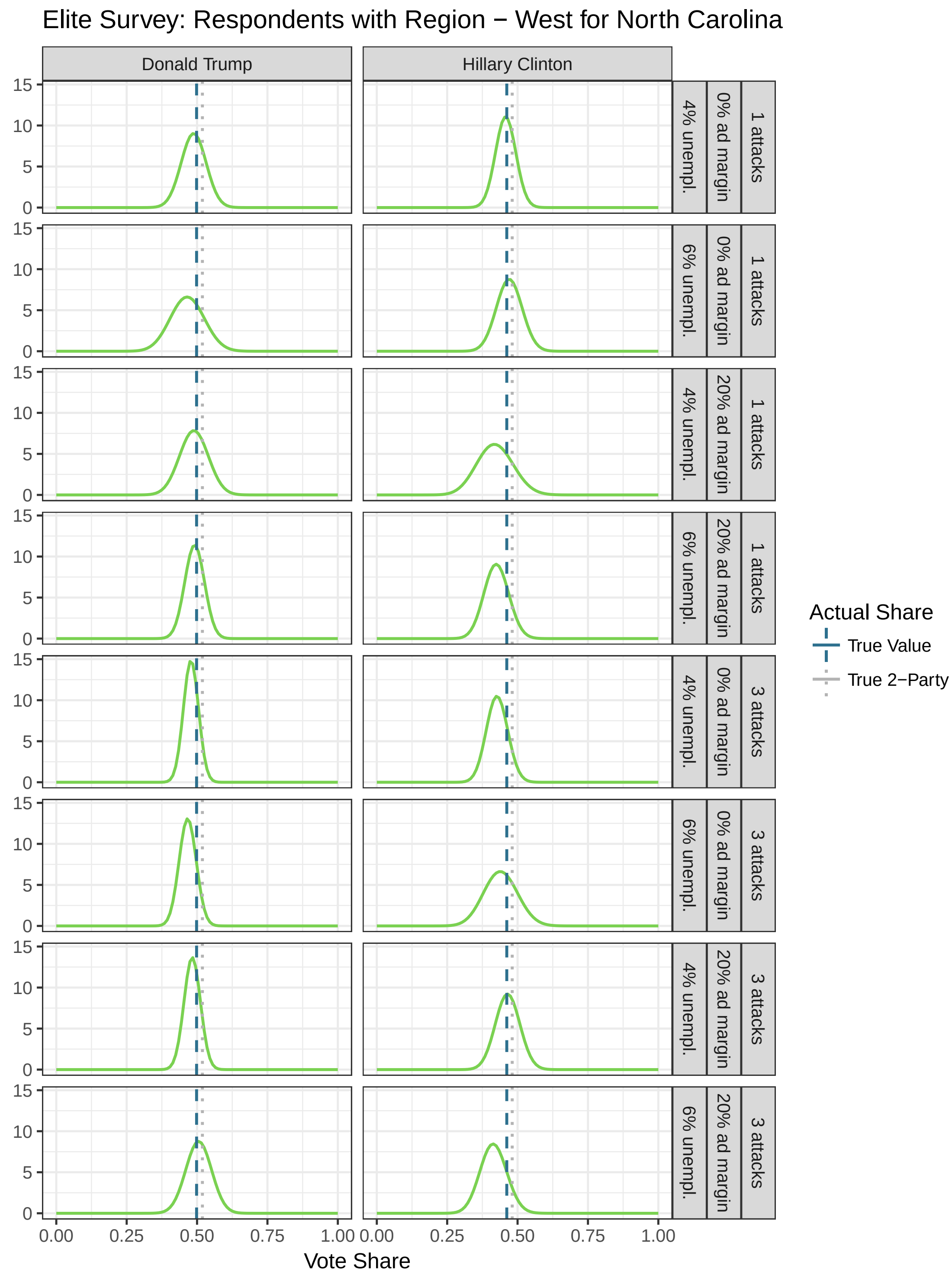


Figure 3.173: Priors with covariates: Elite North Carolina Region West

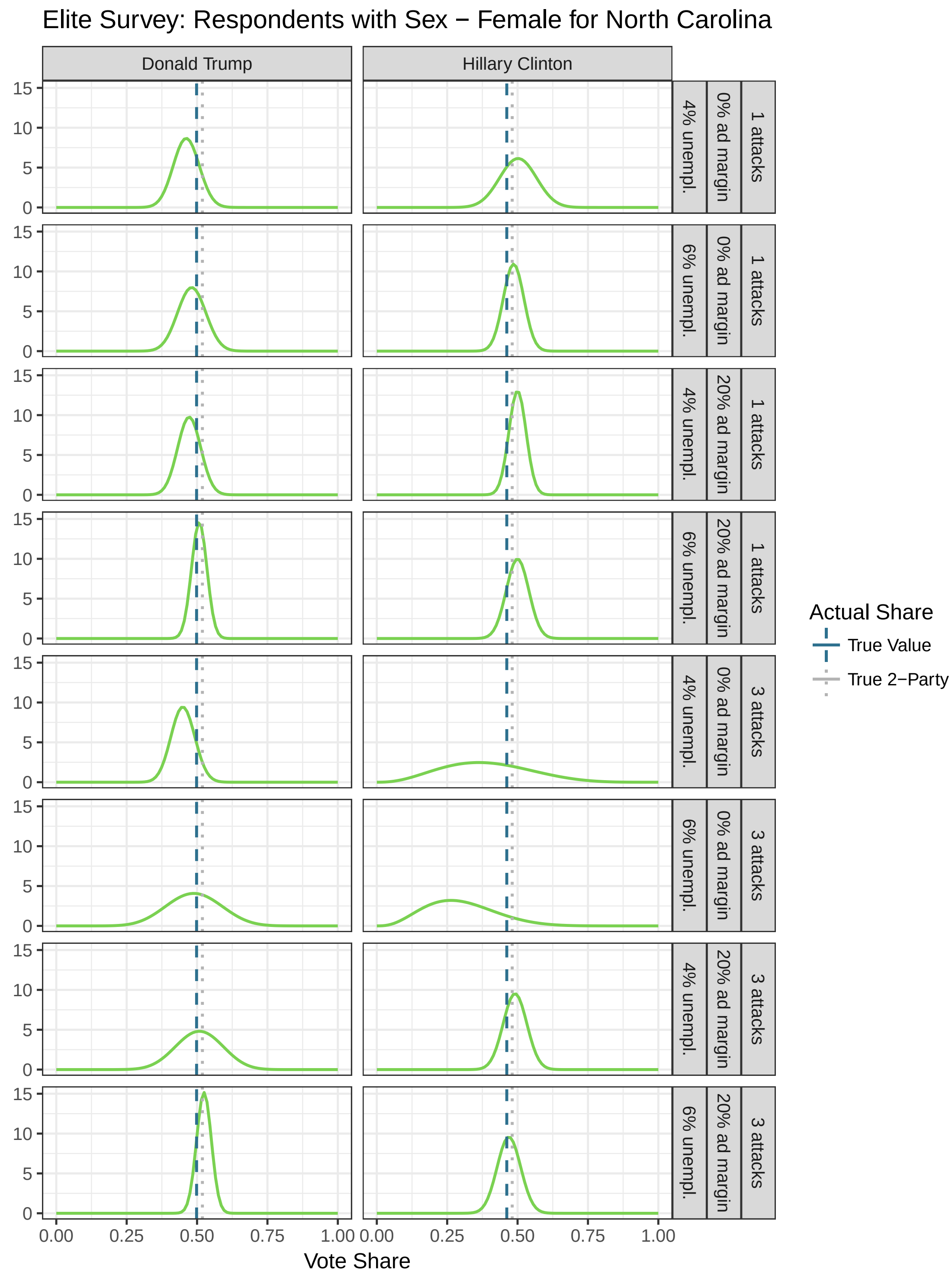


Figure 3.174: Priors with covariates: Elite North Carolina Sex Female

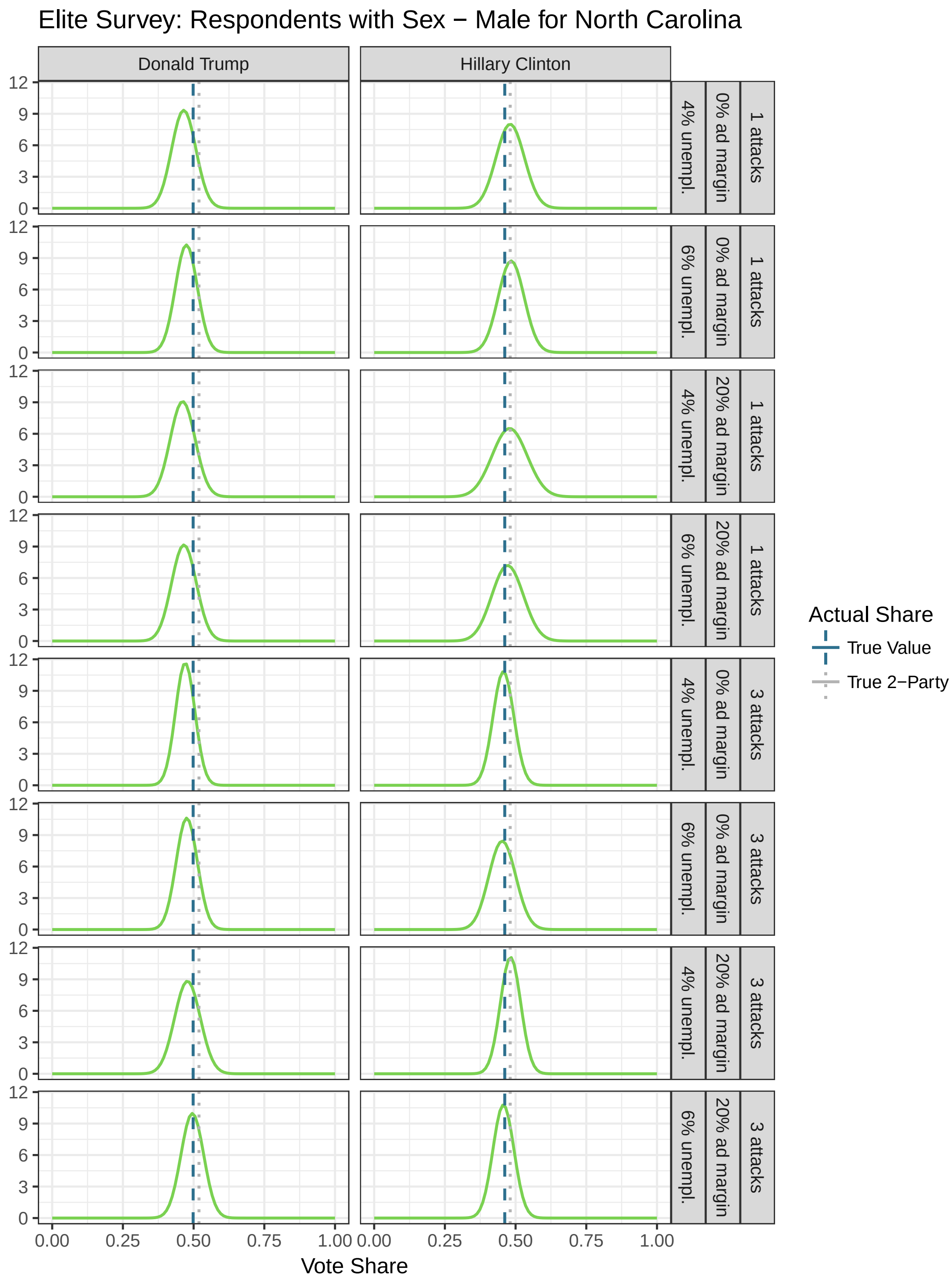


Figure 3.175: Priors with covariates: Elite North Carolina Sex Male

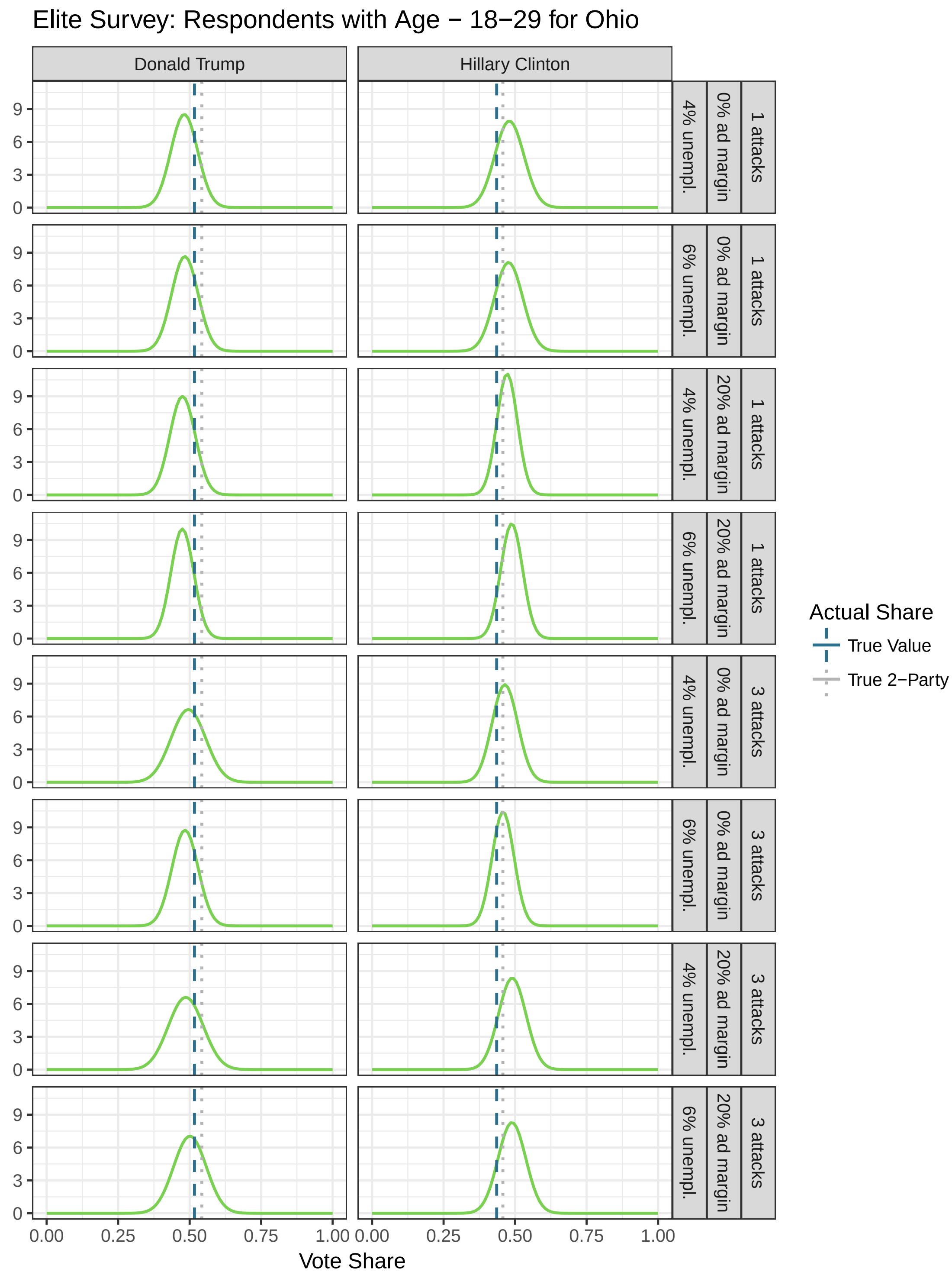


Figure 3.176: Priors with covariates: Elite Ohio Age 18-29

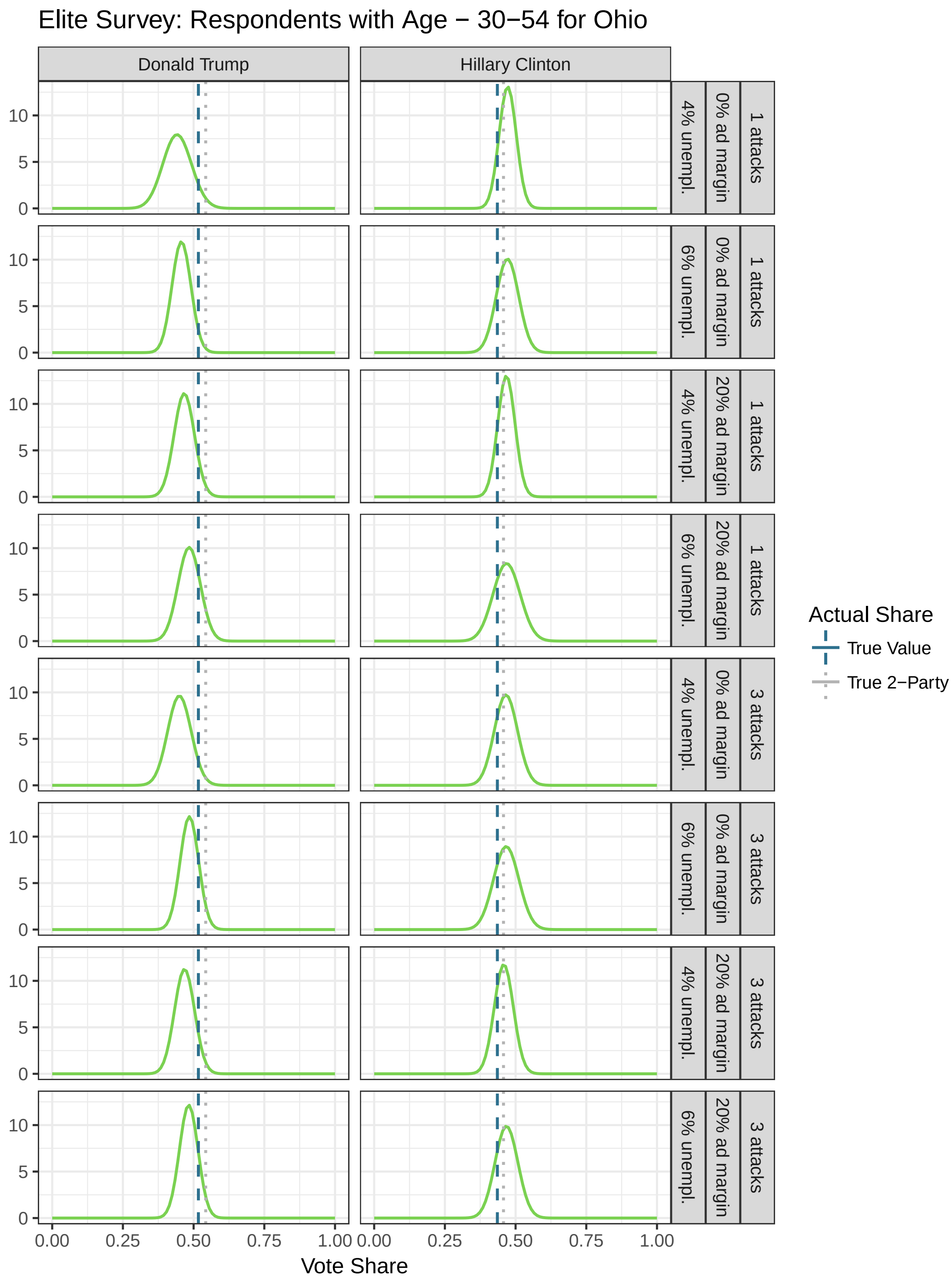


Figure 3.177: Priors with covariates: Elite Ohio Age 30-54

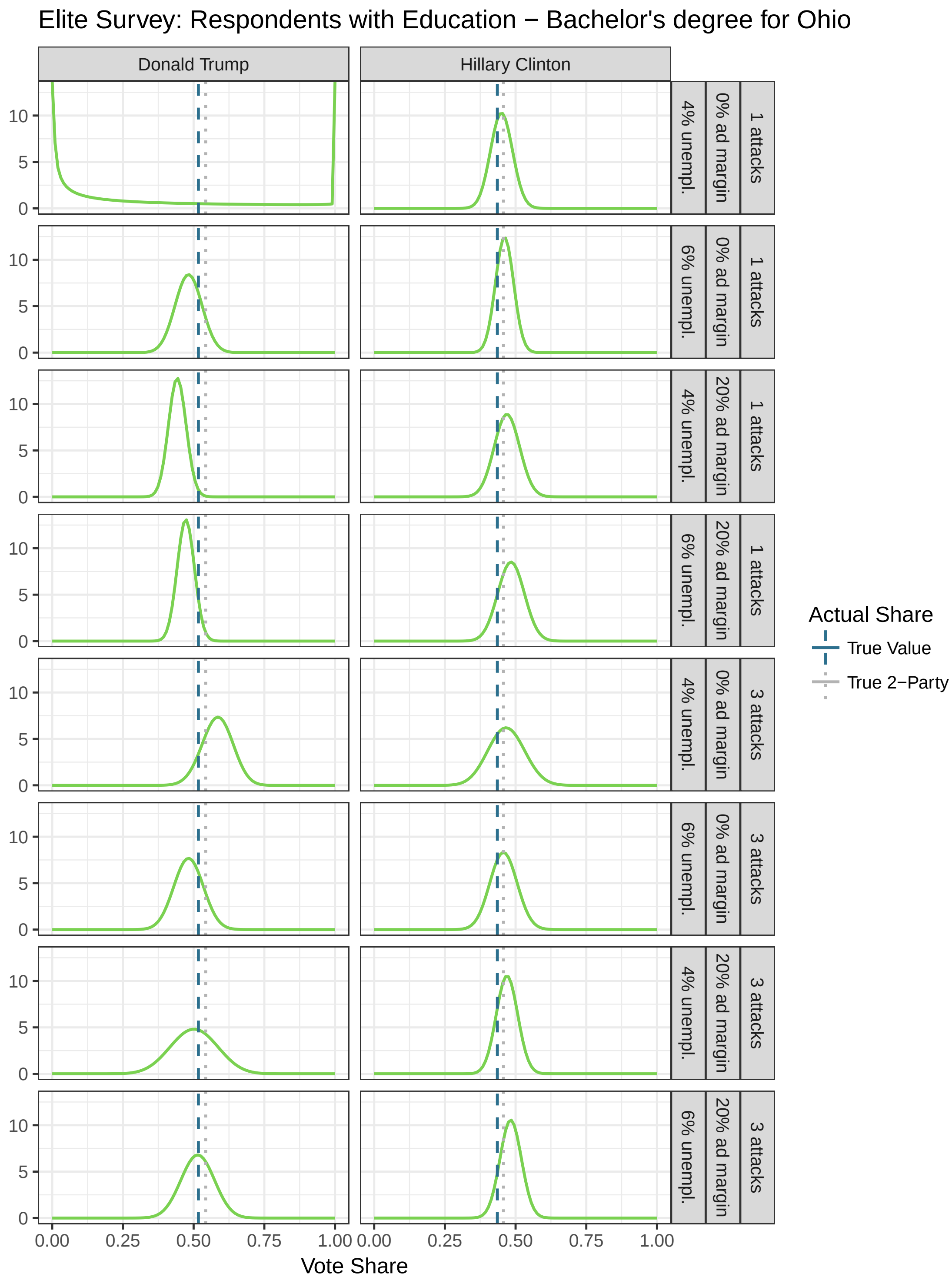


Figure 3.178: Priors with covariates: Elite Ohio Education Bachelor's degree

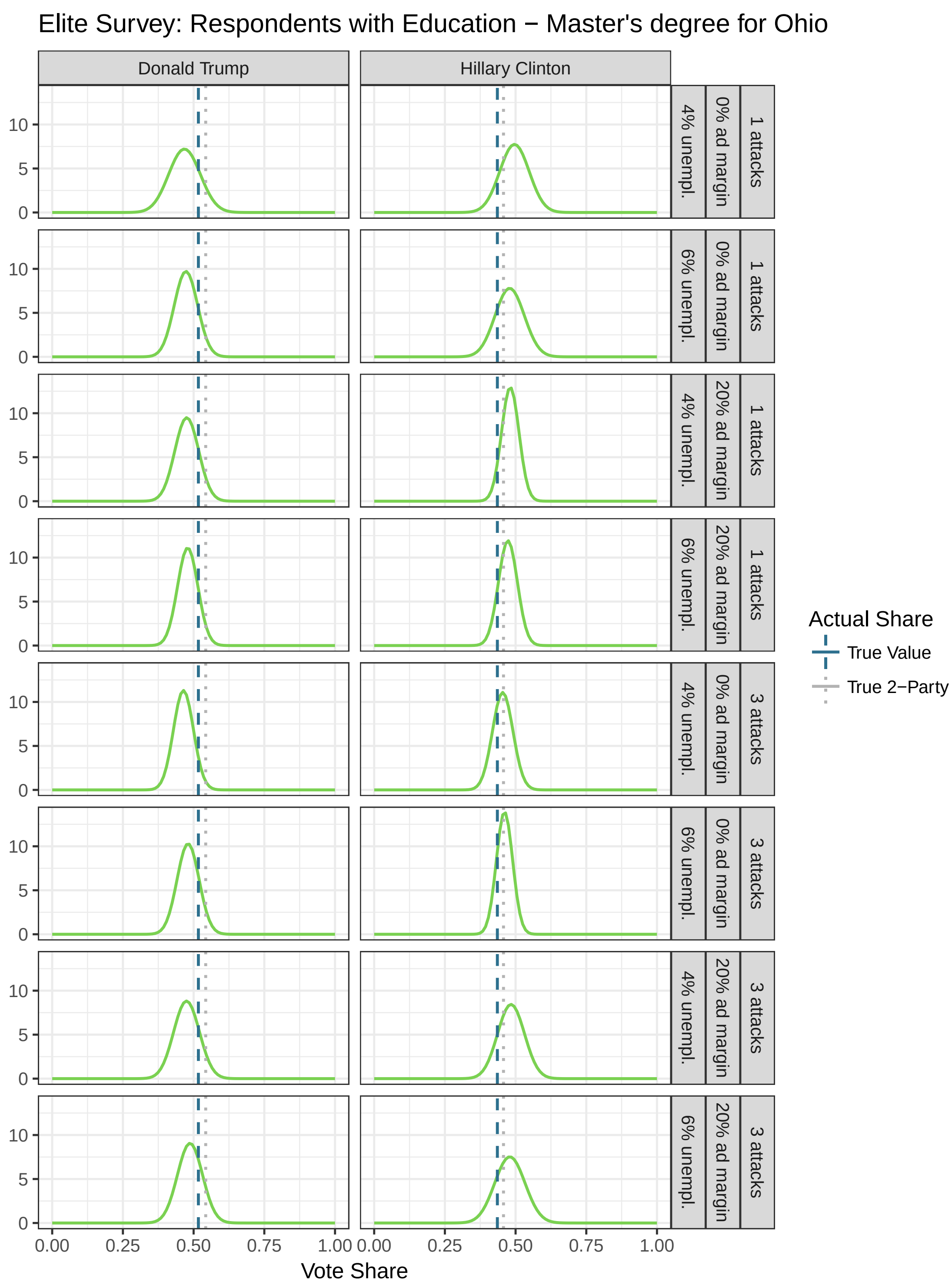


Figure 3.179: Priors with covariates: Elite Ohio Education Master's degree

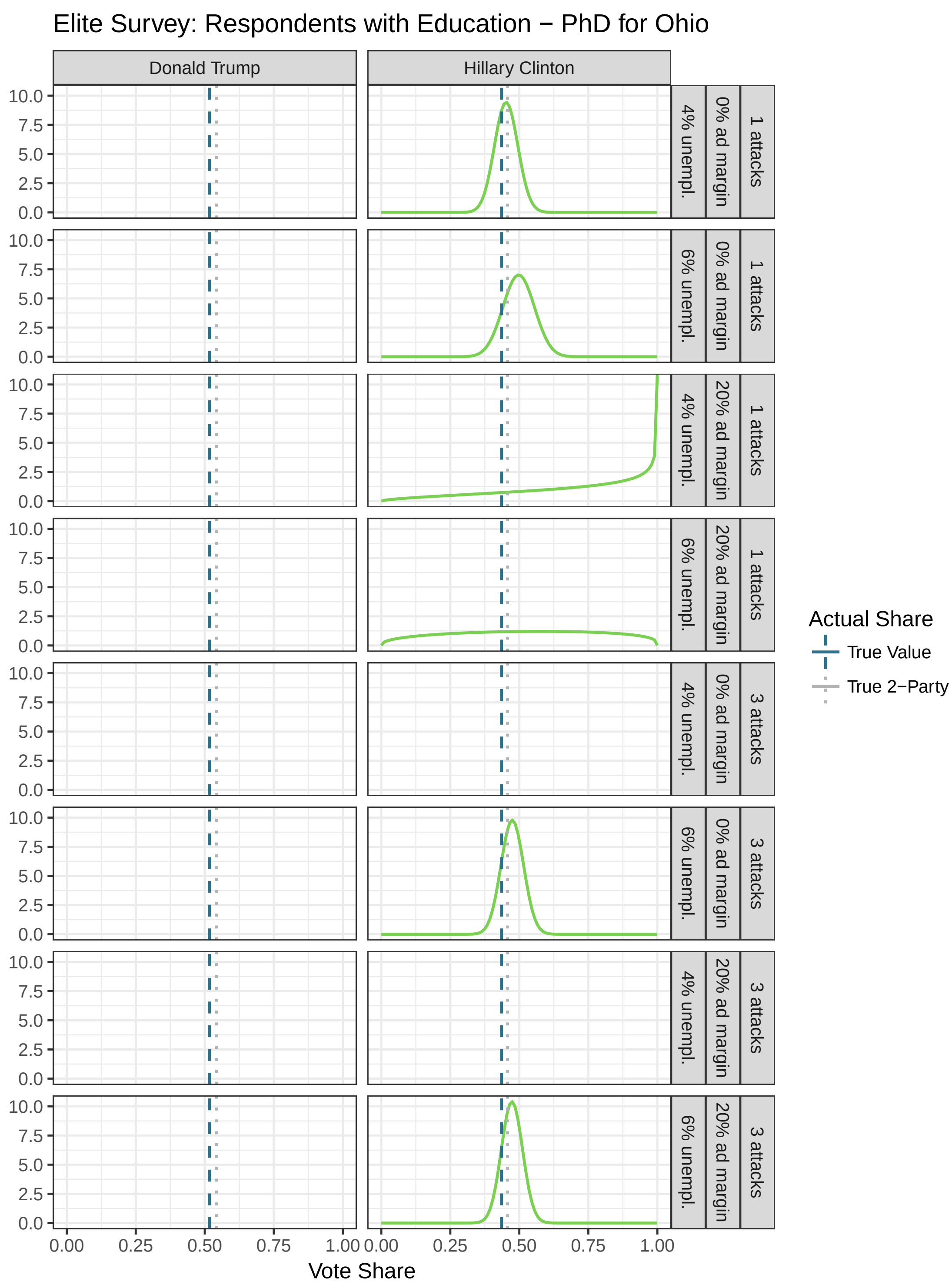


Figure 3.180: Priors with covariates: Elite Ohio Education PhD

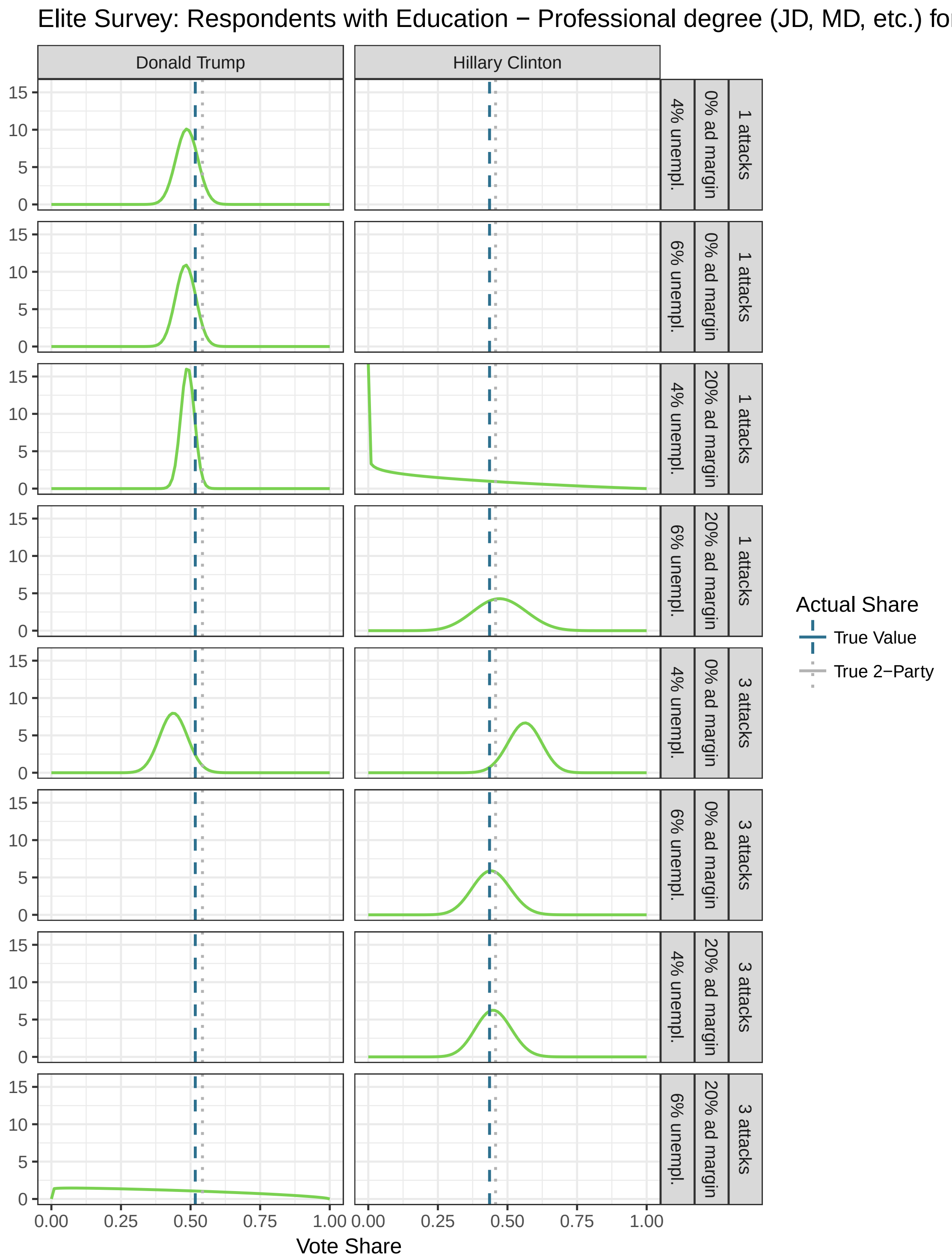


Figure 3.181: Priors with covariates: Elite Ohio Education Professional degree JD MD etc

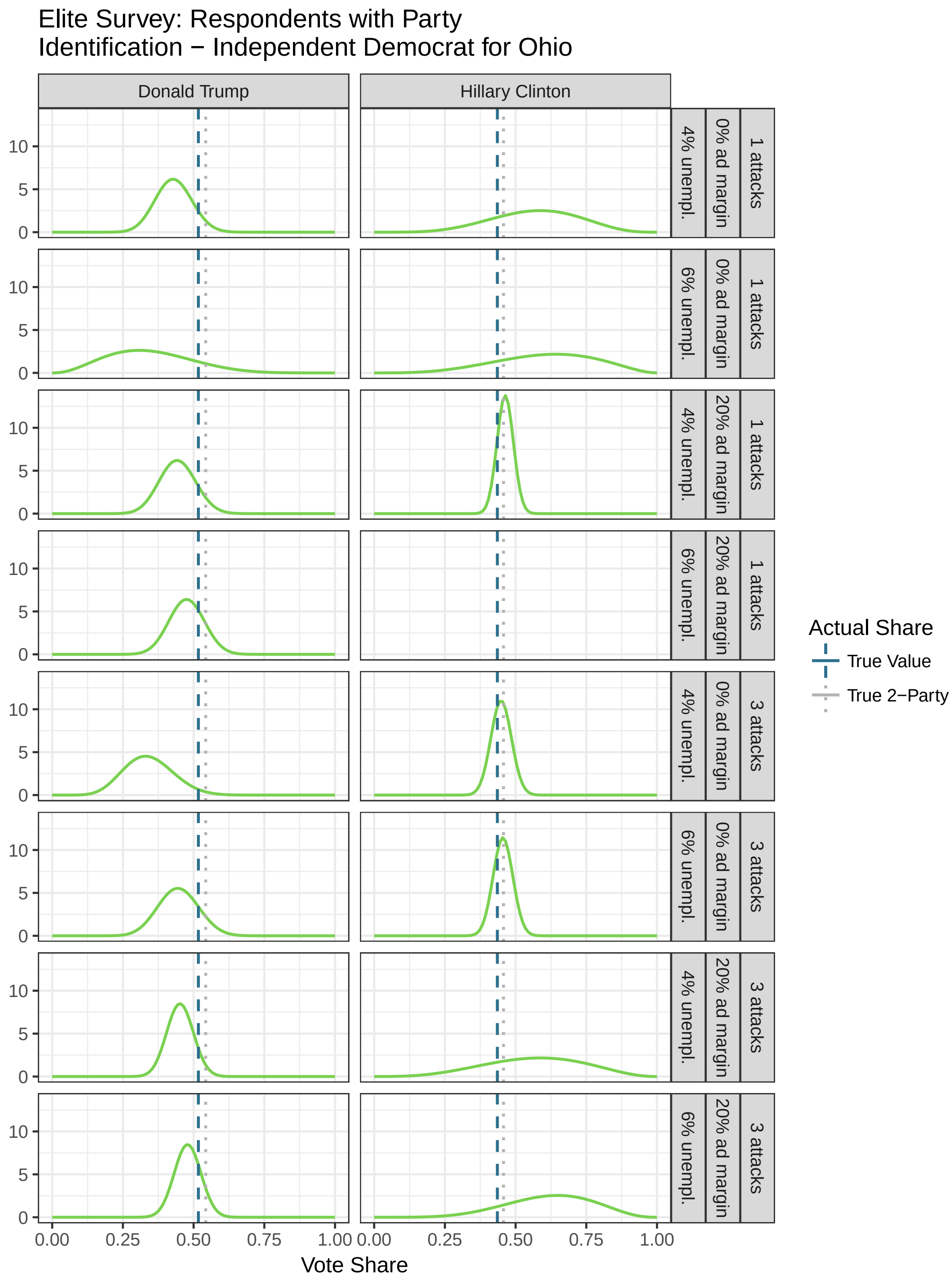


Figure 3.182: Priors with covariates: Elite Ohio Party Identification Independent Democrat

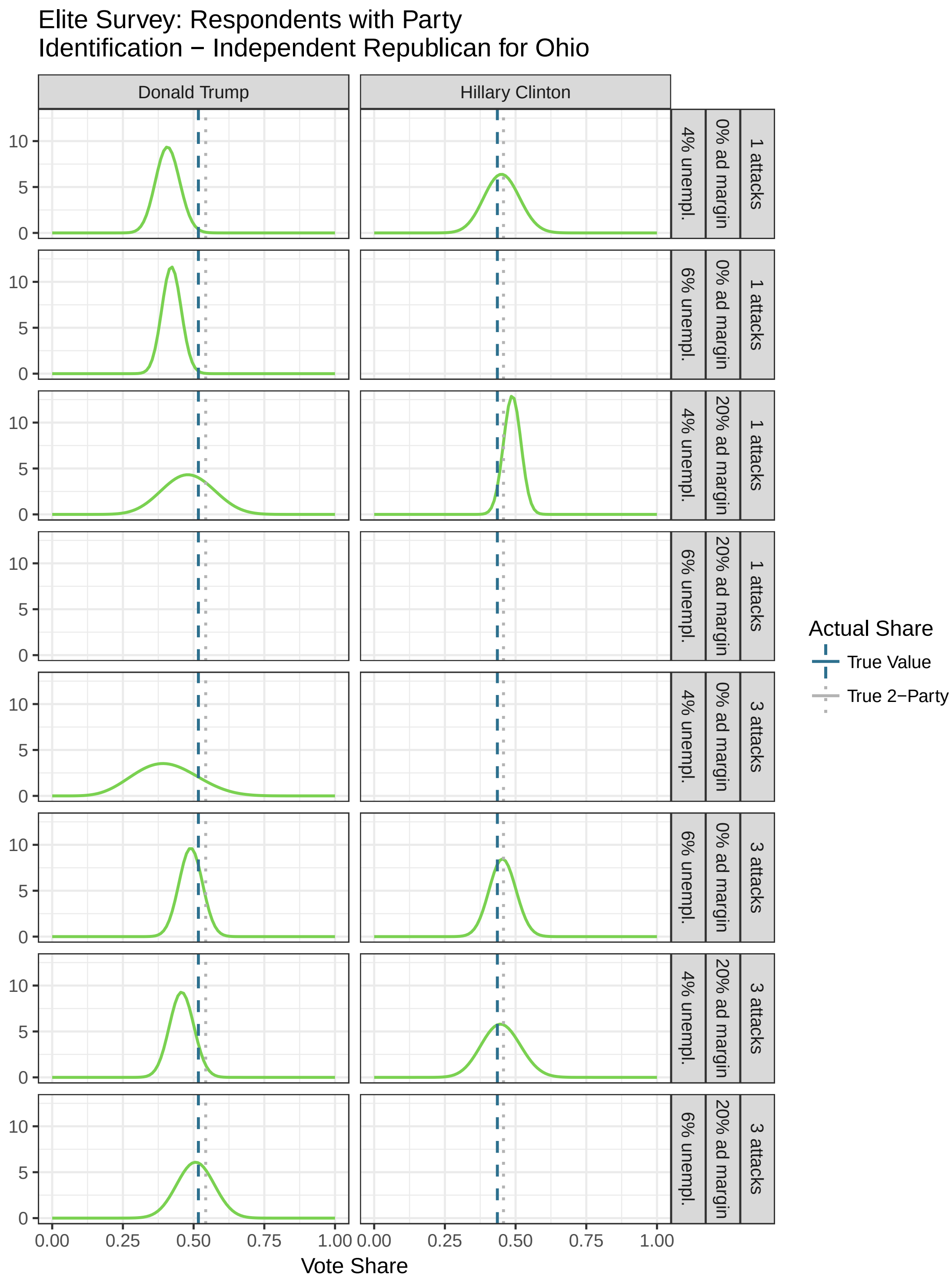


Figure 3.183: Priors with covariates: Elite Ohio Party Identification Independent Republican

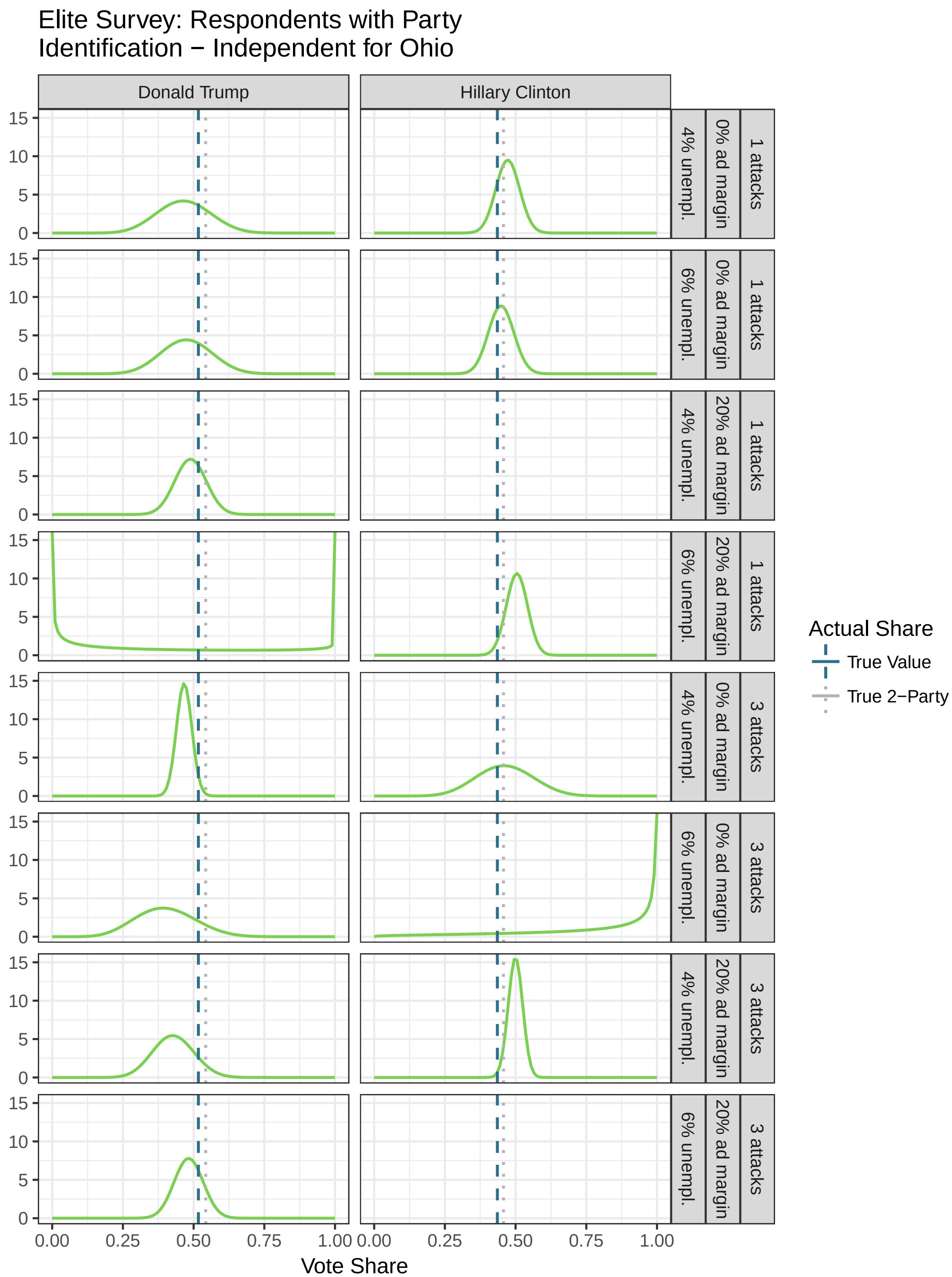


Figure 3.184: Priors with covariates: Elite Ohio Party Identification Independent

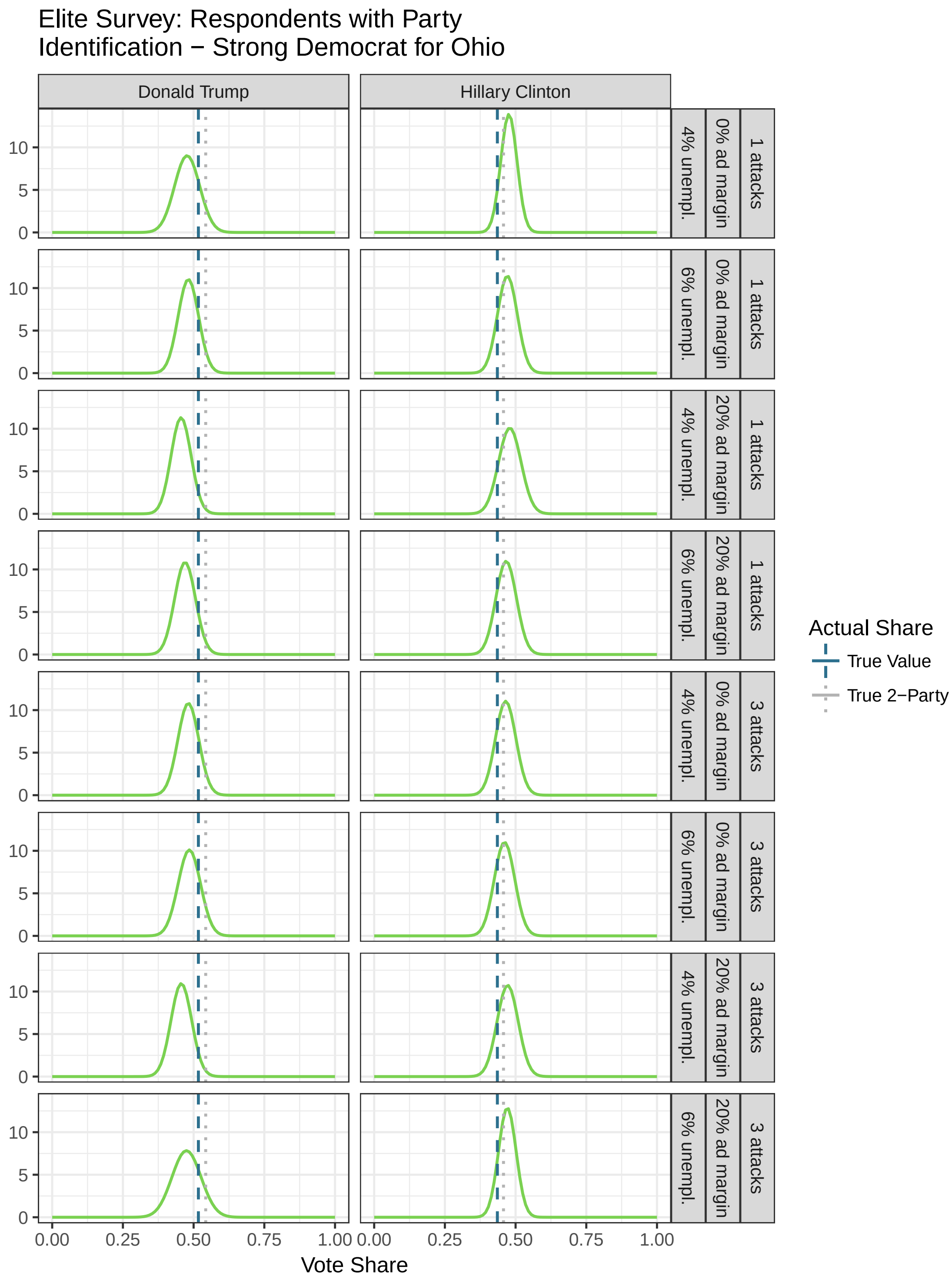


Figure 3.185: Priors with covariates: Elite Ohio Party Identification Strong Democrat

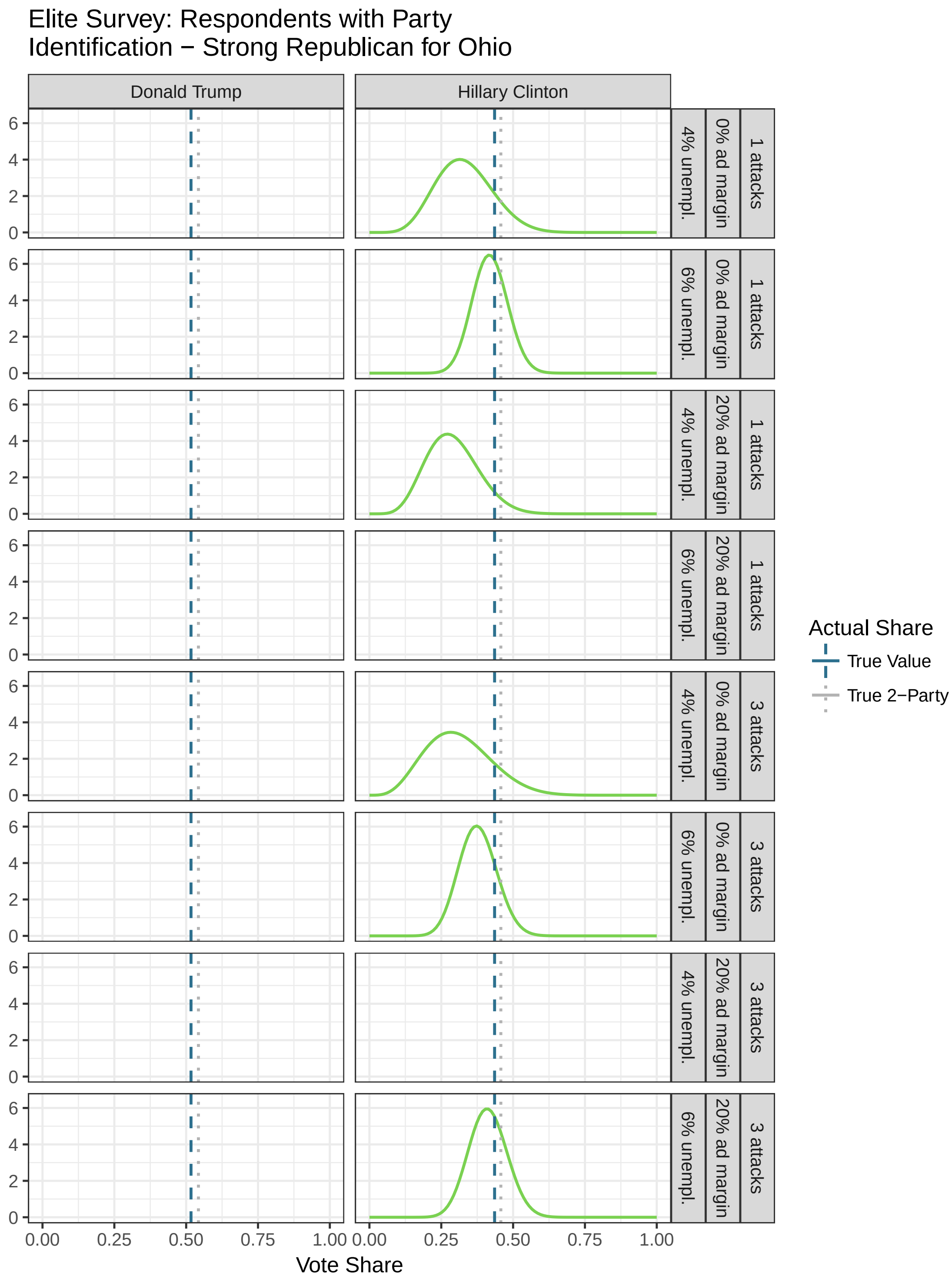


Figure 3.186: Priors with covariates: Elite Ohio Party Identification Strong Republican

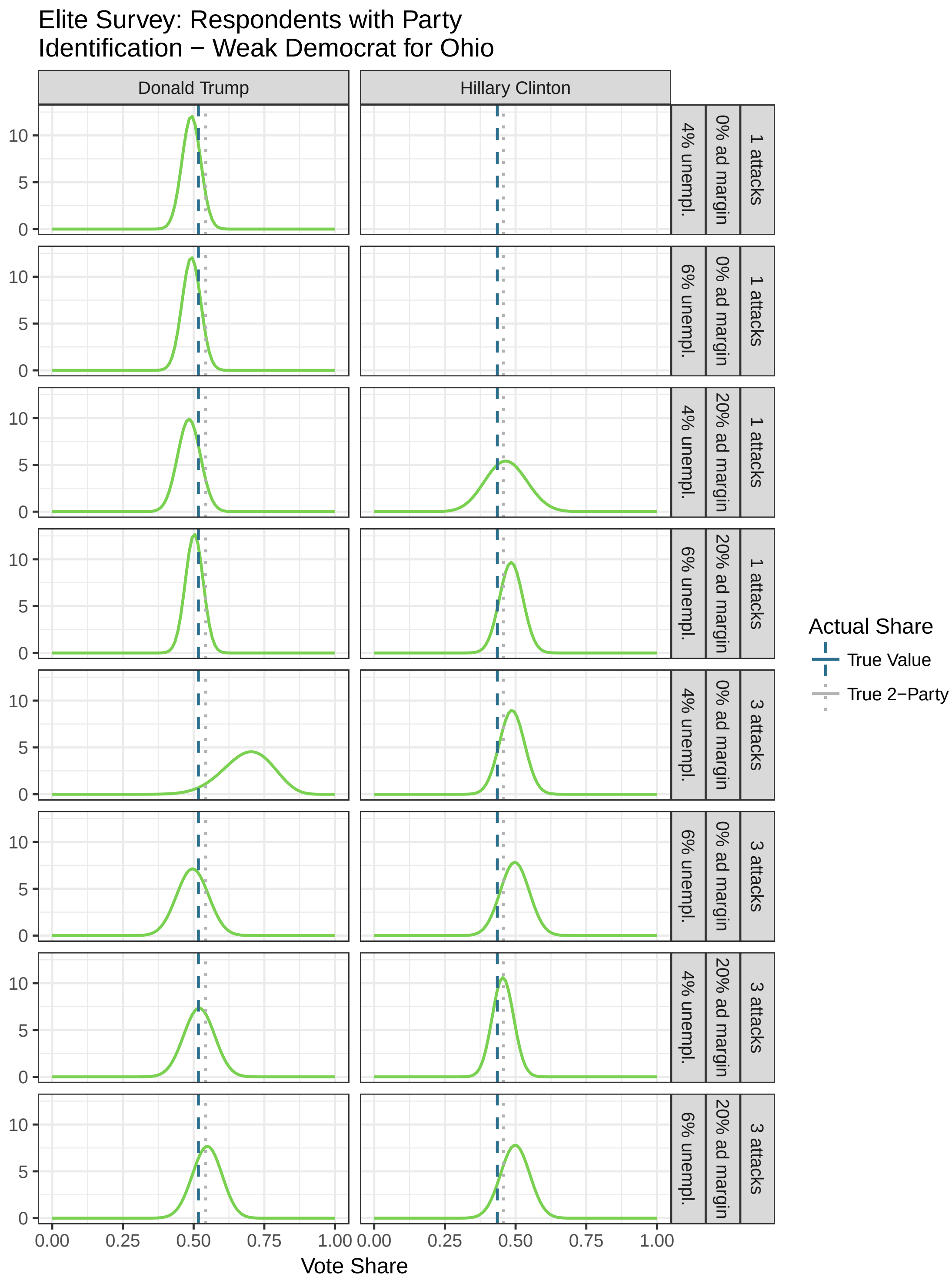


Figure 3.187: Priors with covariates: Elite Ohio Party Identification Weak Democrat

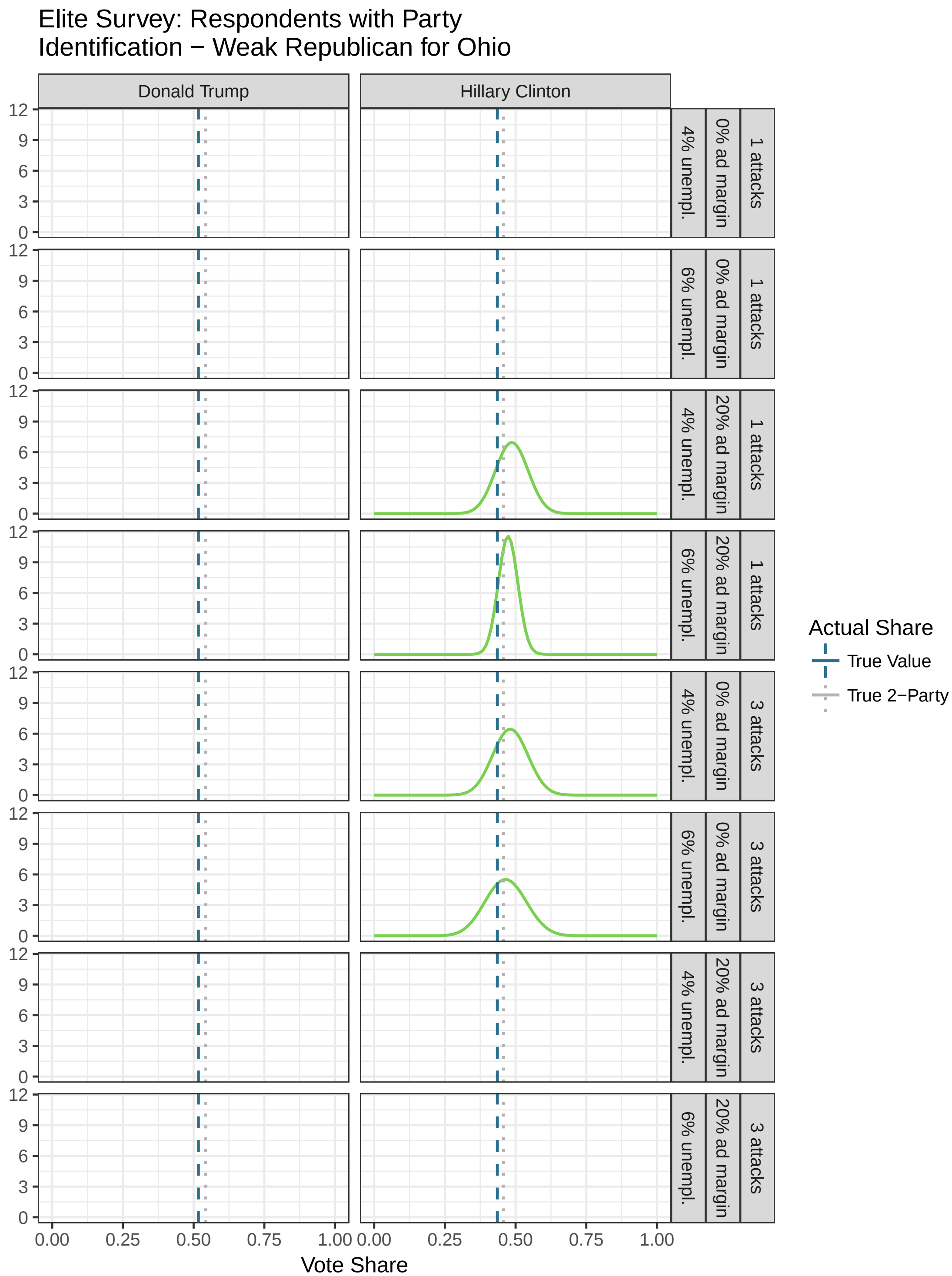


Figure 3.188: Priors with covariates: Elite Ohio Party Identification Weak Republican

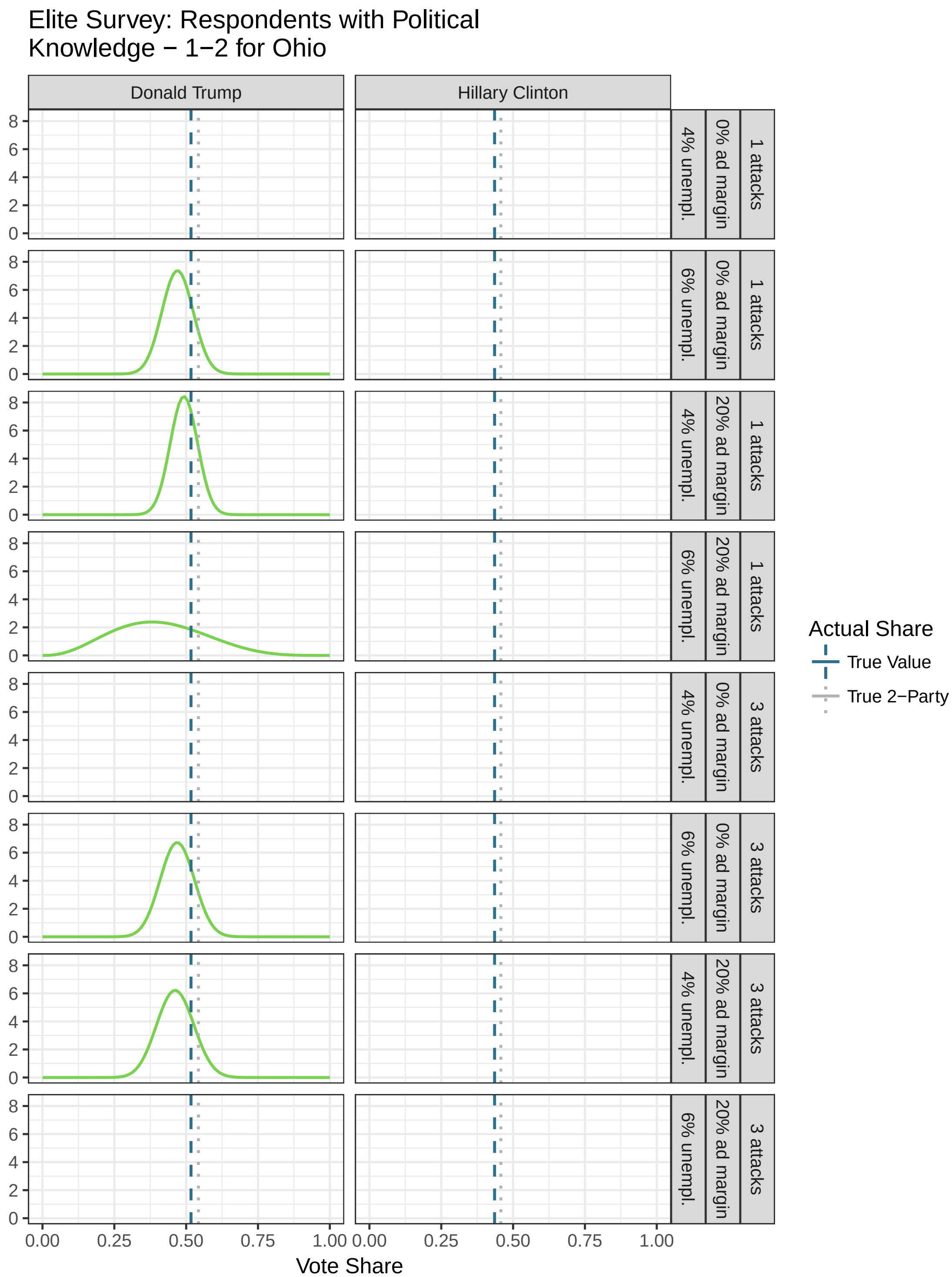


Figure 3.189: Priors with covariates: Elite Ohio Political Knowledge 1-2

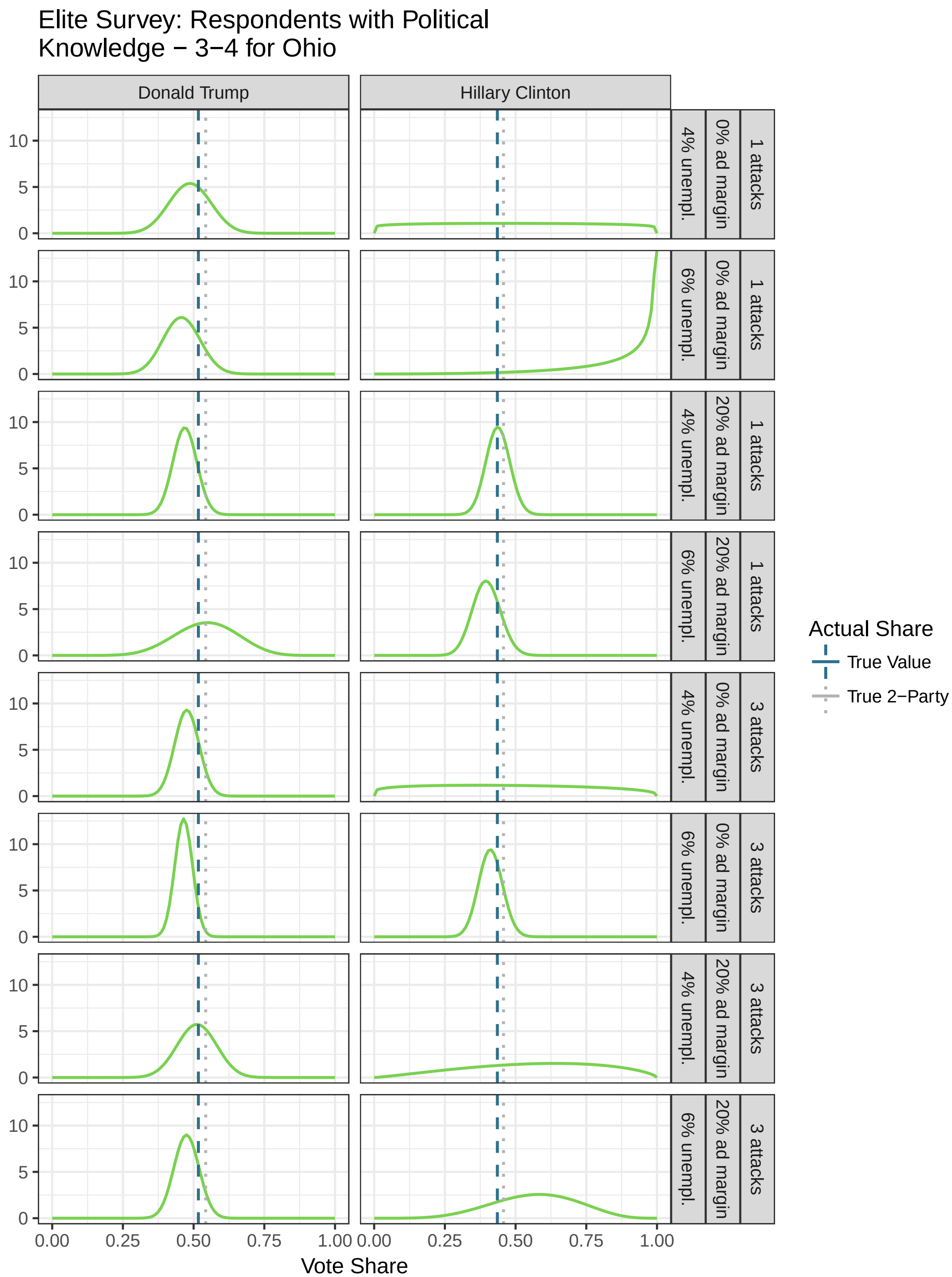


Figure 3.190: Priors with covariates: Elite Ohio Political Knowledge 3-4

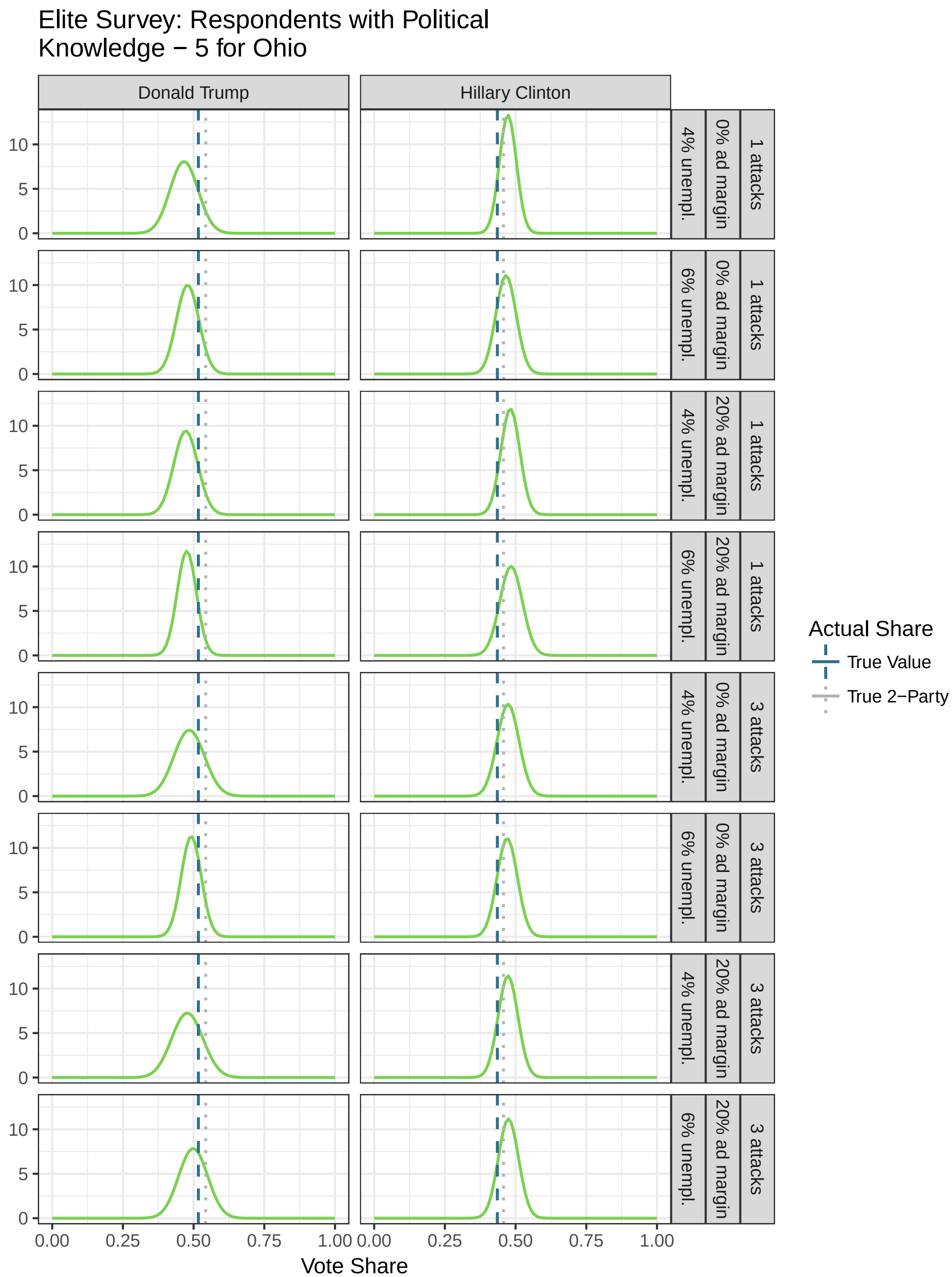


Figure 3.191: Priors with covariates: Elite Ohio Political Knowledge 5

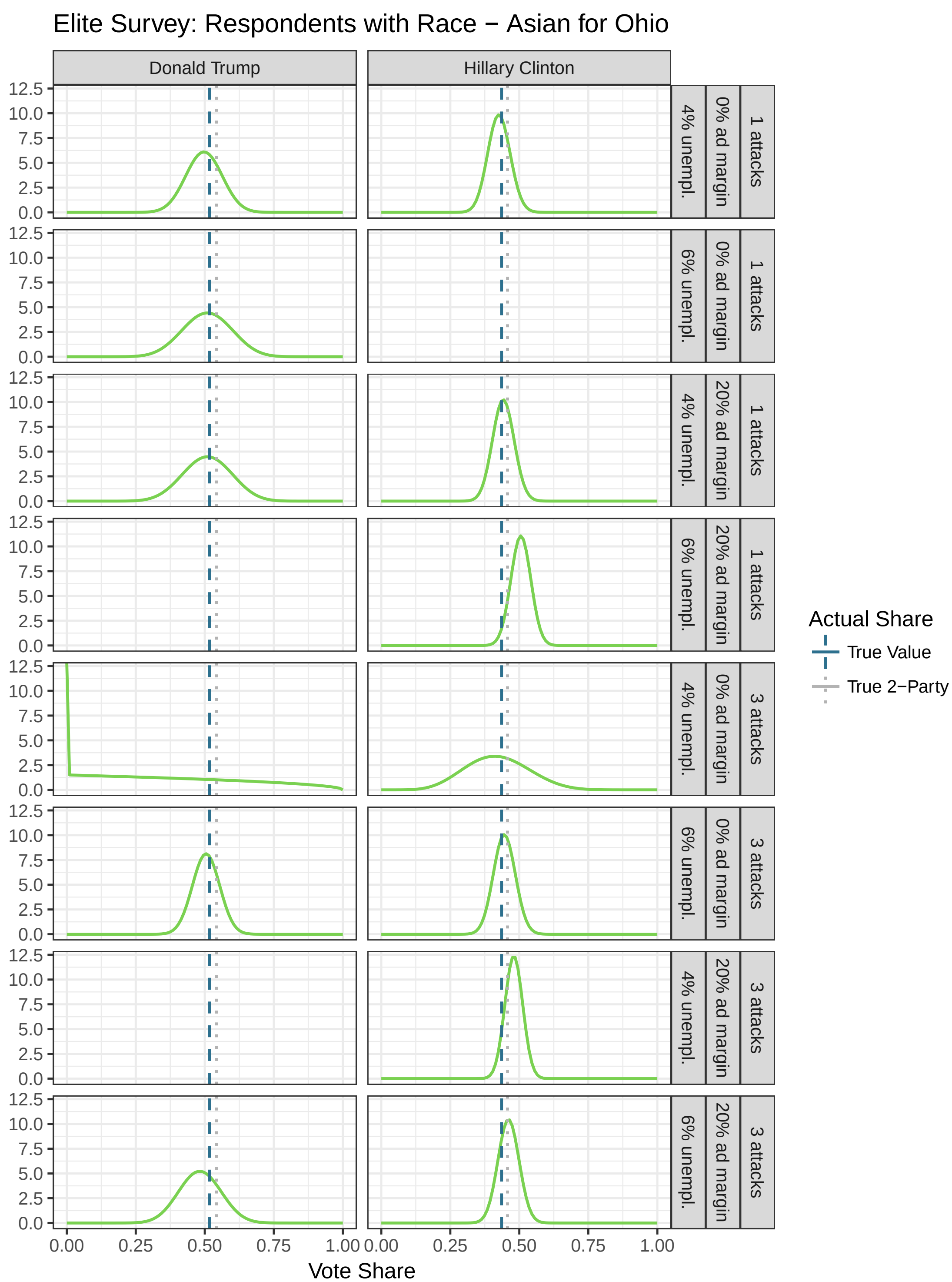


Figure 3.192: Priors with covariates: Elite Ohio Race Asian

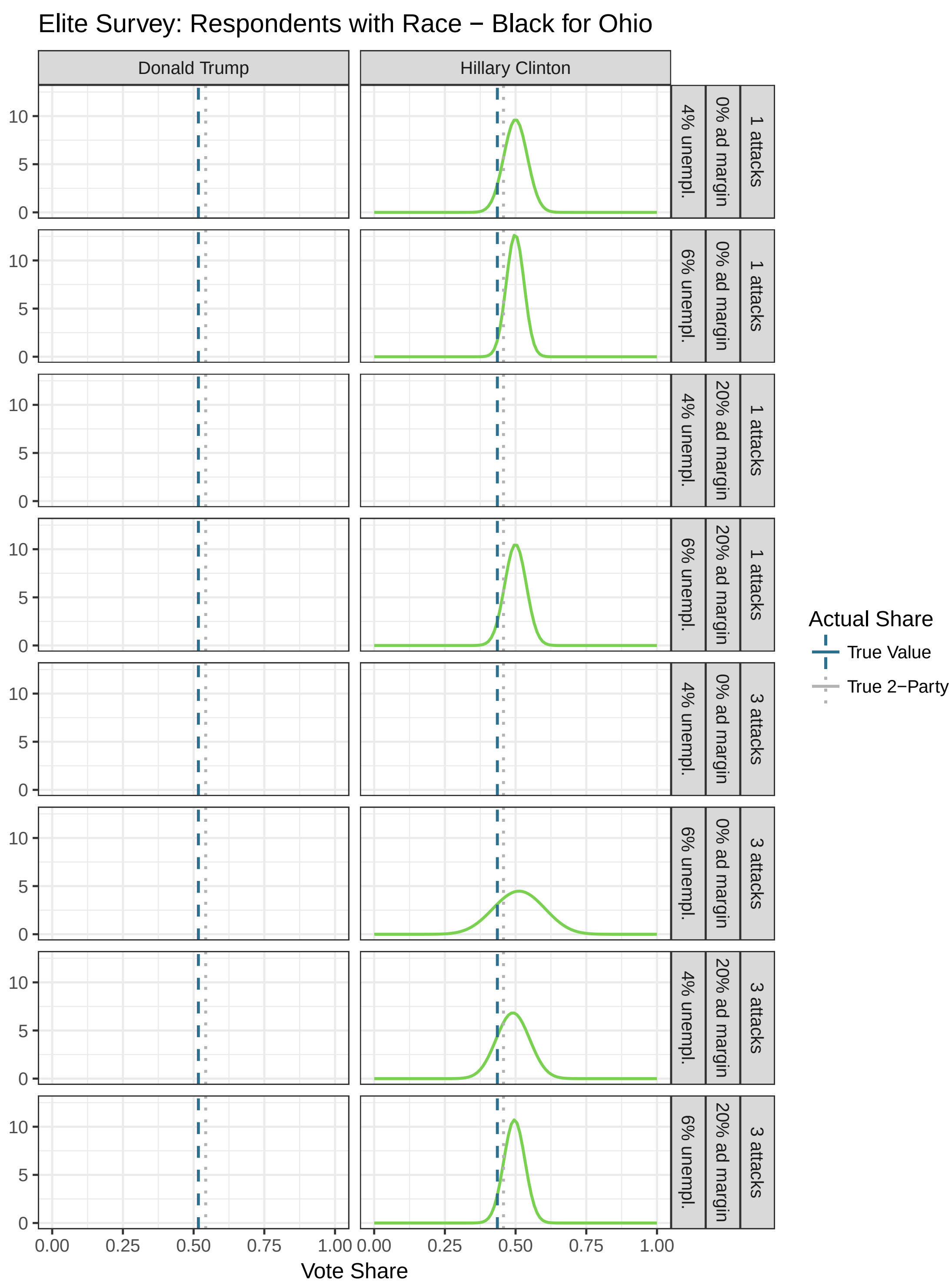


Figure 3.193: Priors with covariates: Elite Ohio Race Black

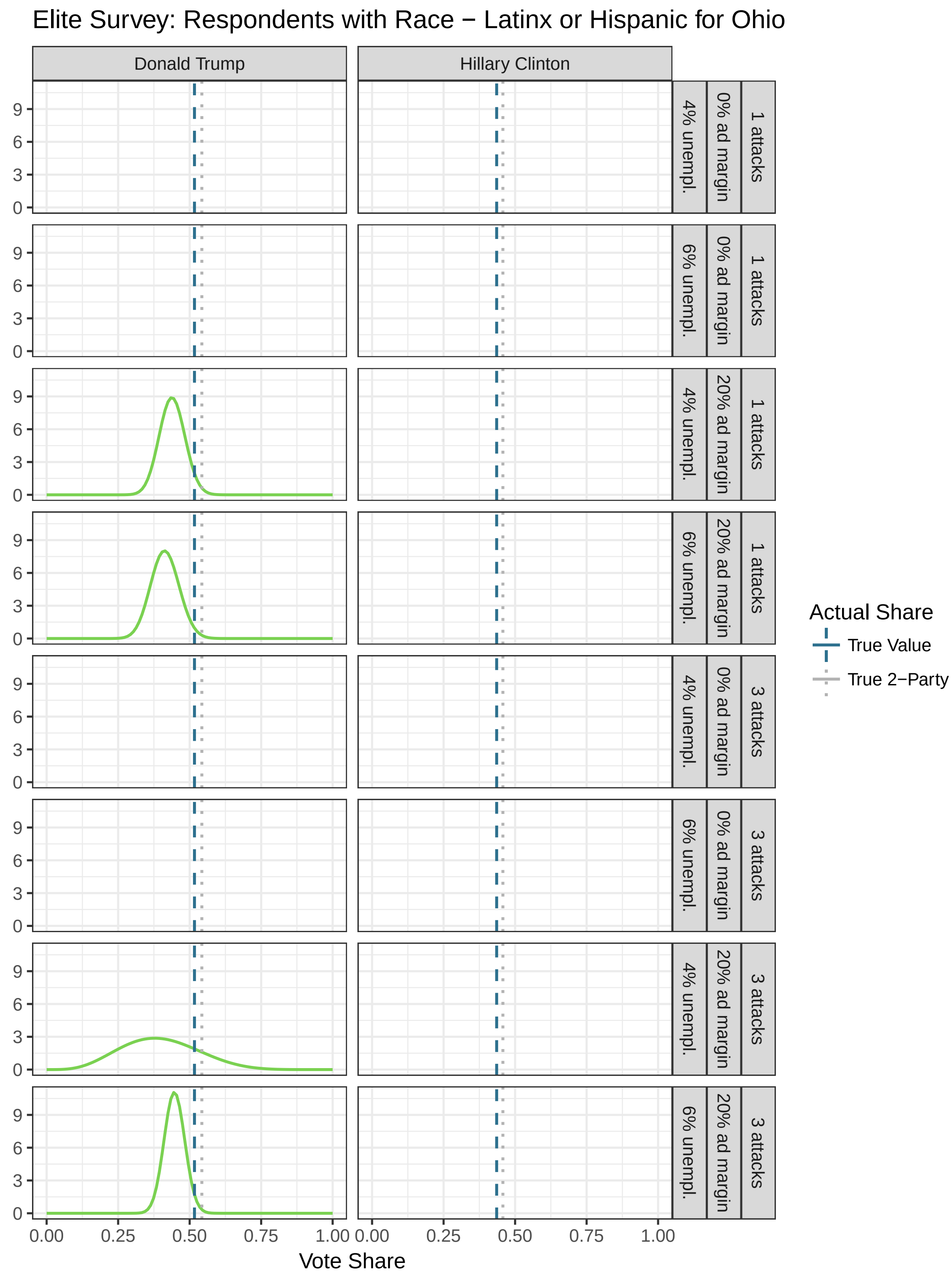


Figure 3.194: Priors with covariates: Elite Ohio Race Latinx or Hispanic

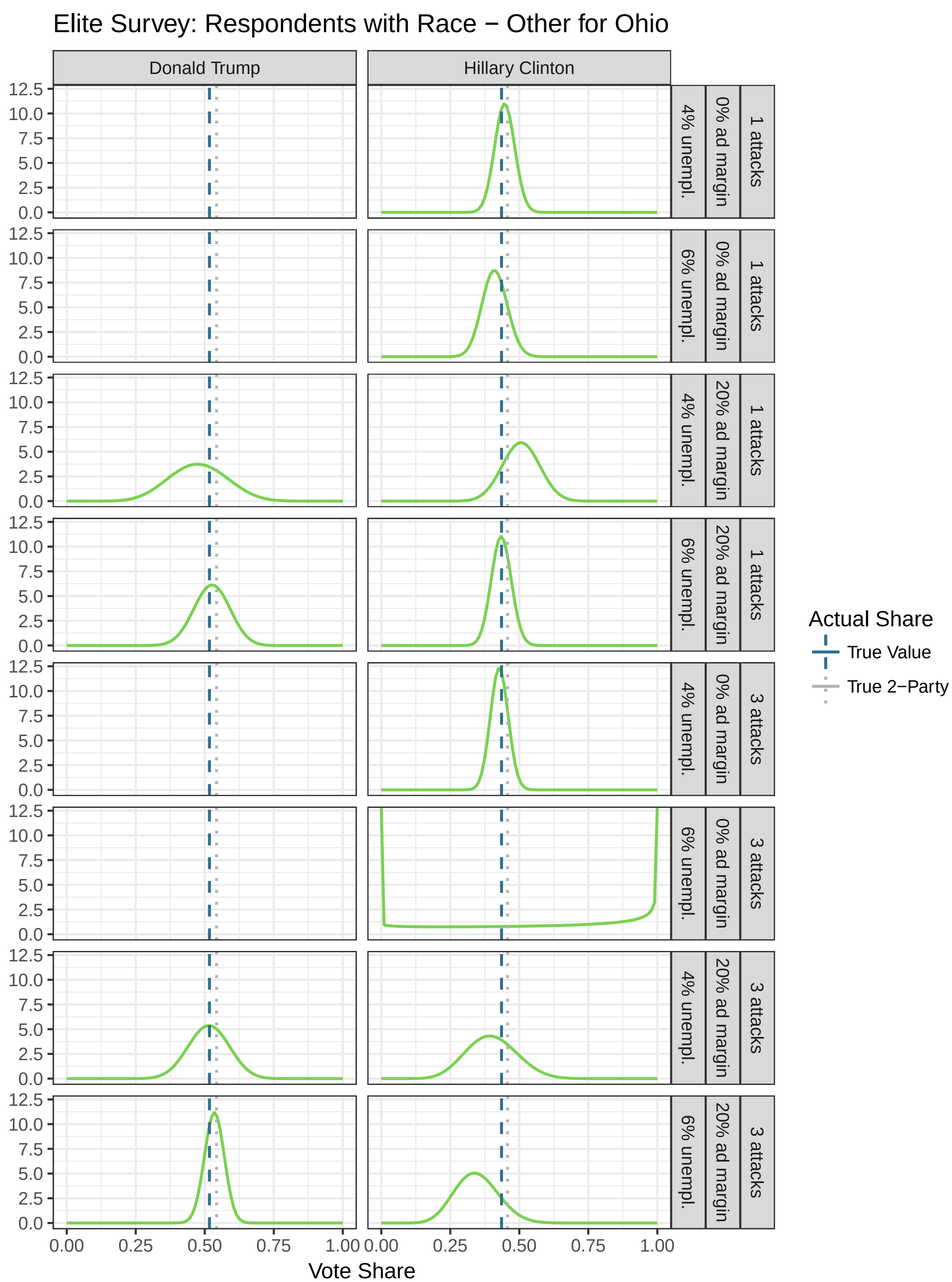


Figure 3.195: Priors with covariates: Elite Ohio Race Other

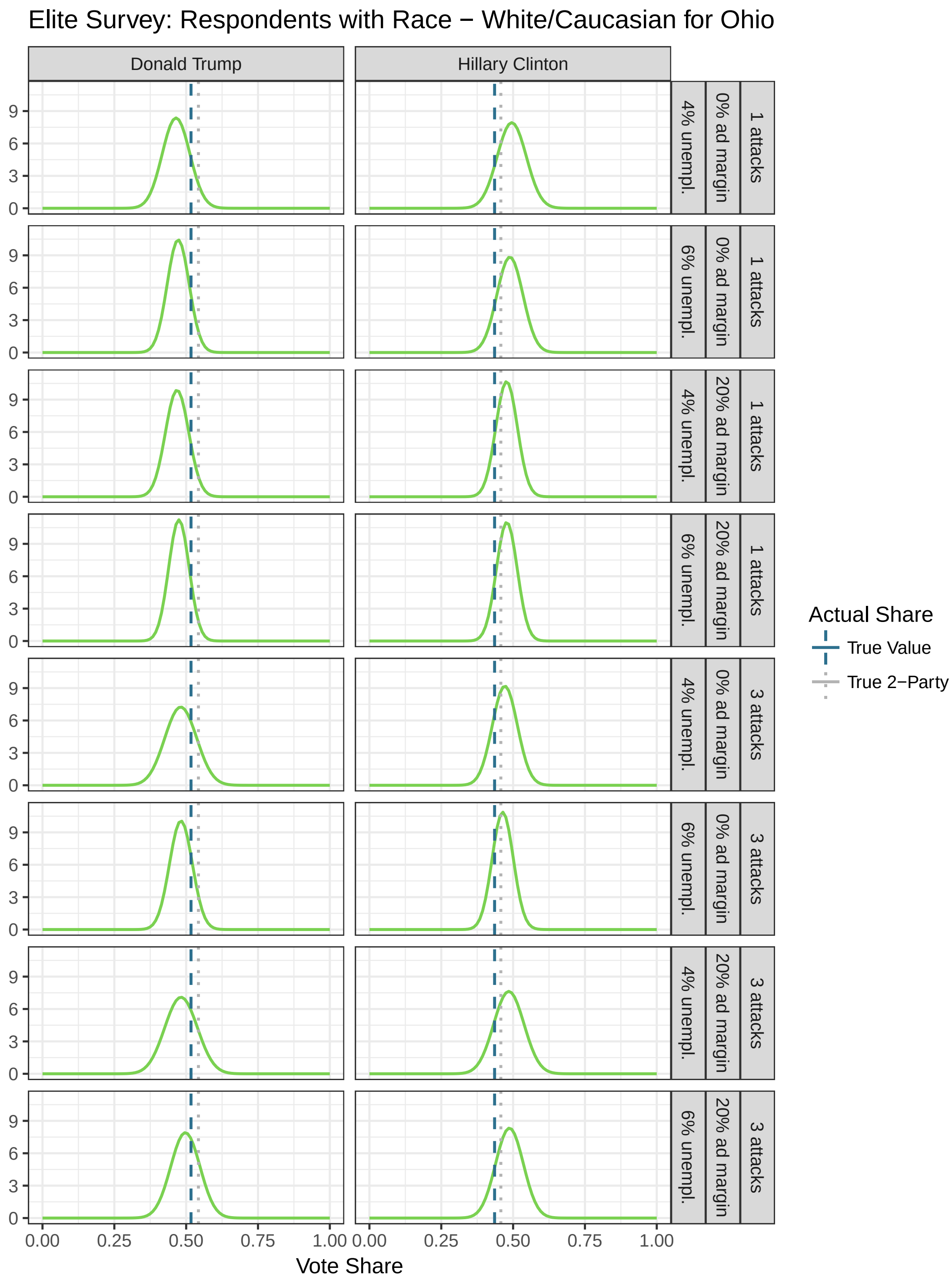


Figure 3.196: Priors with covariates: Elite Ohio Race White Caucasian

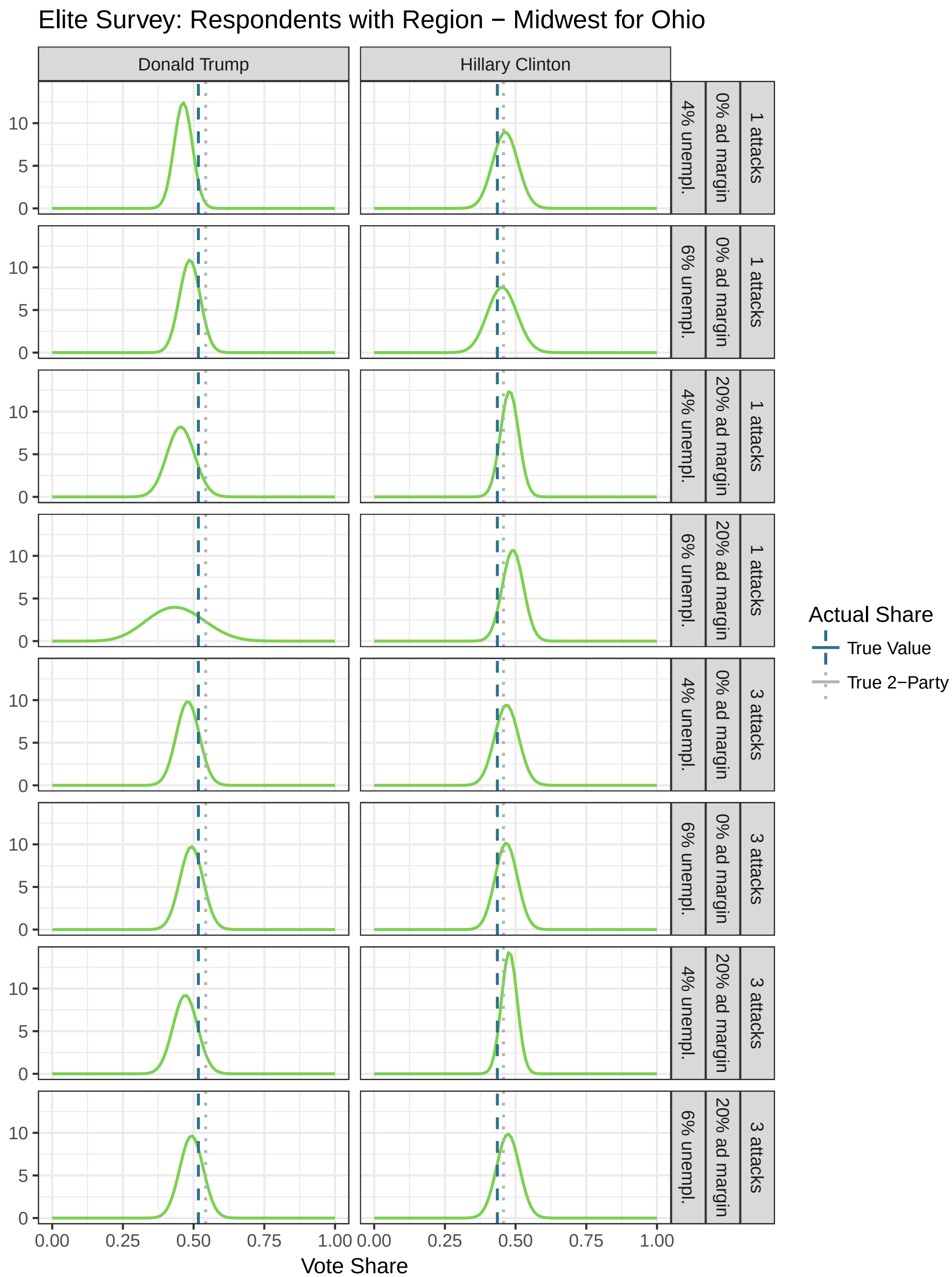


Figure 3.197: Priors with covariates: Elite Ohio Region Midwest

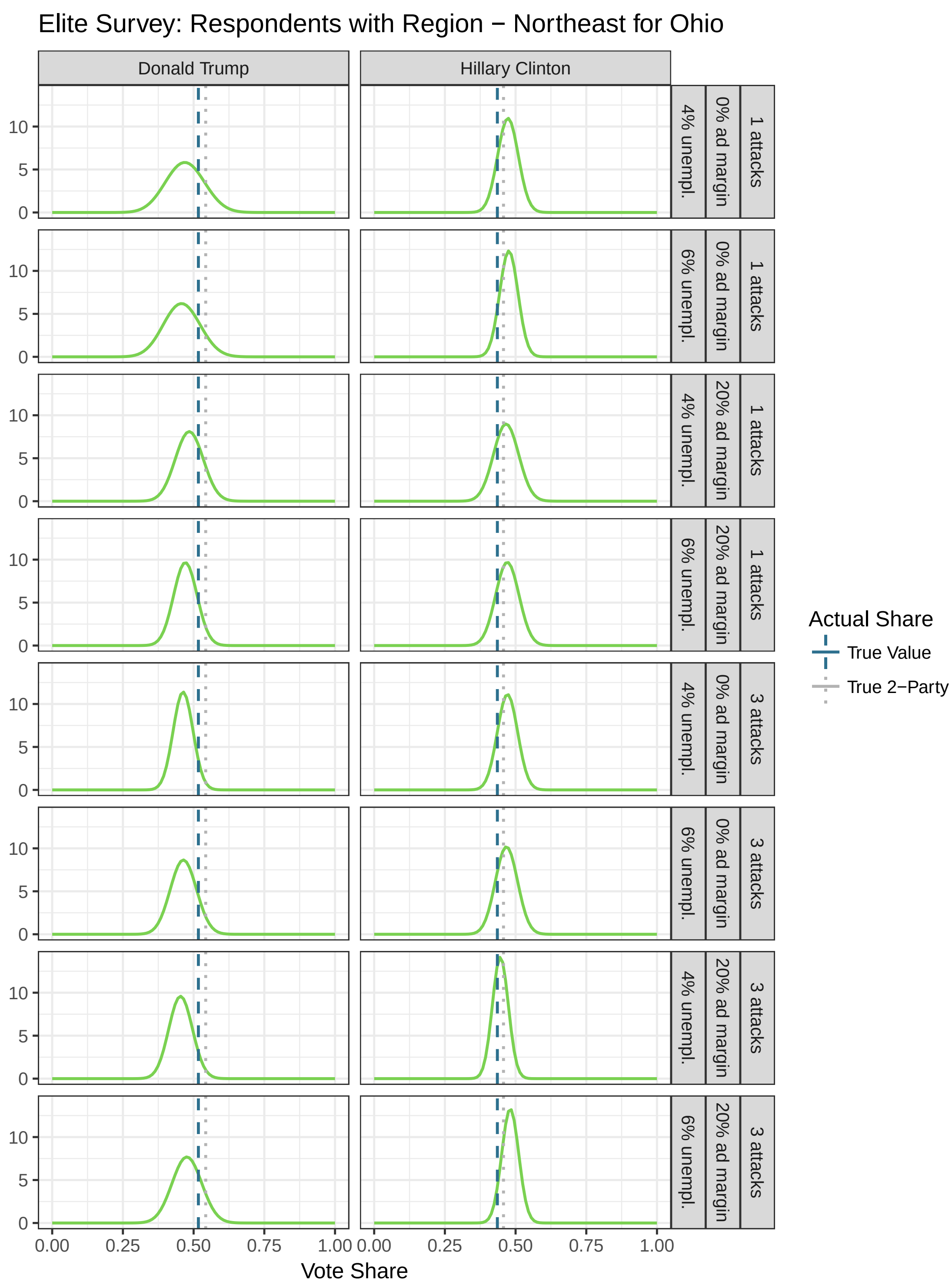


Figure 3.198: Priors with covariates: Elite Ohio Region Northeast

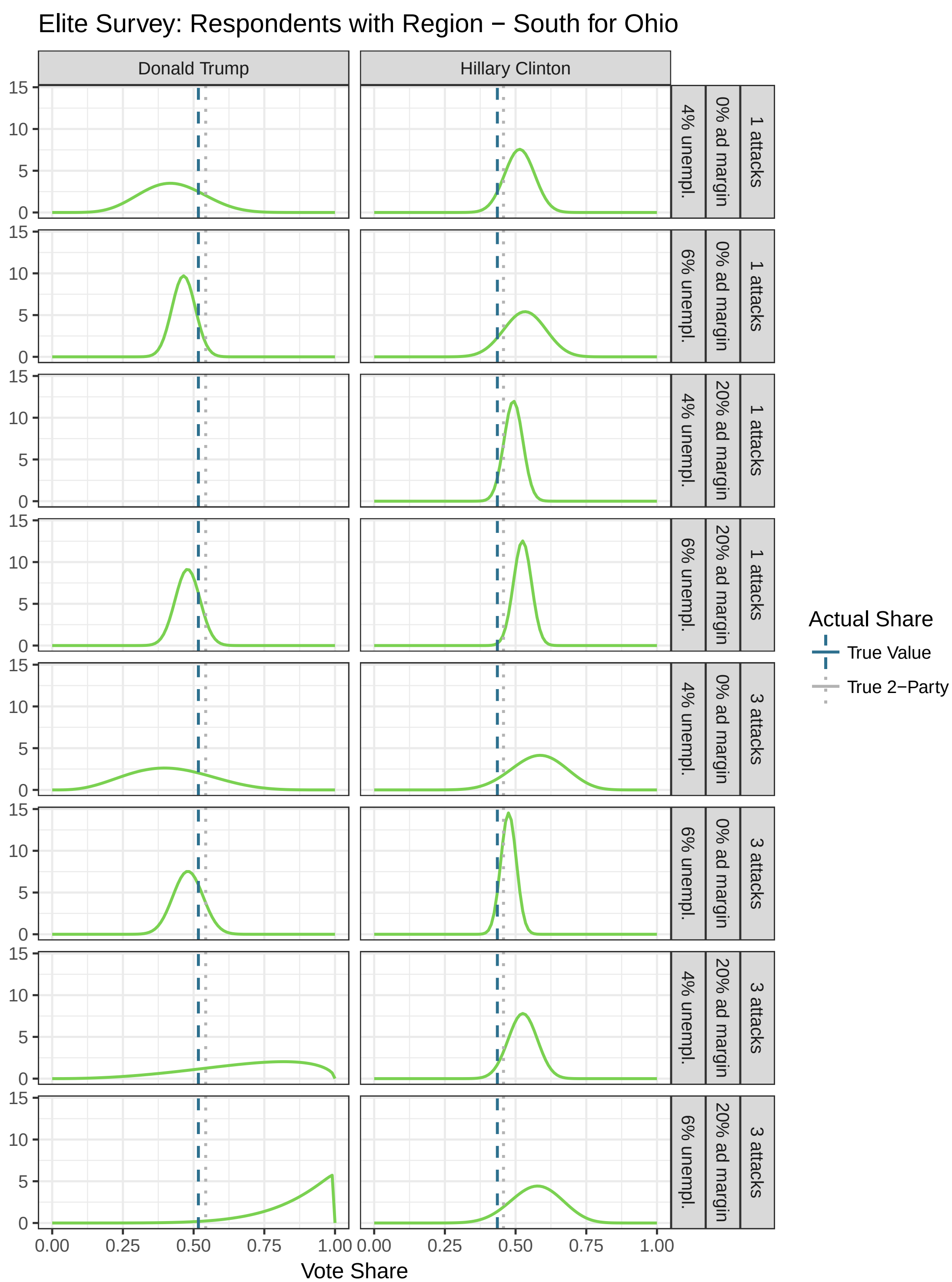


Figure 3.199: Priors with covariates: Elite Ohio Region South

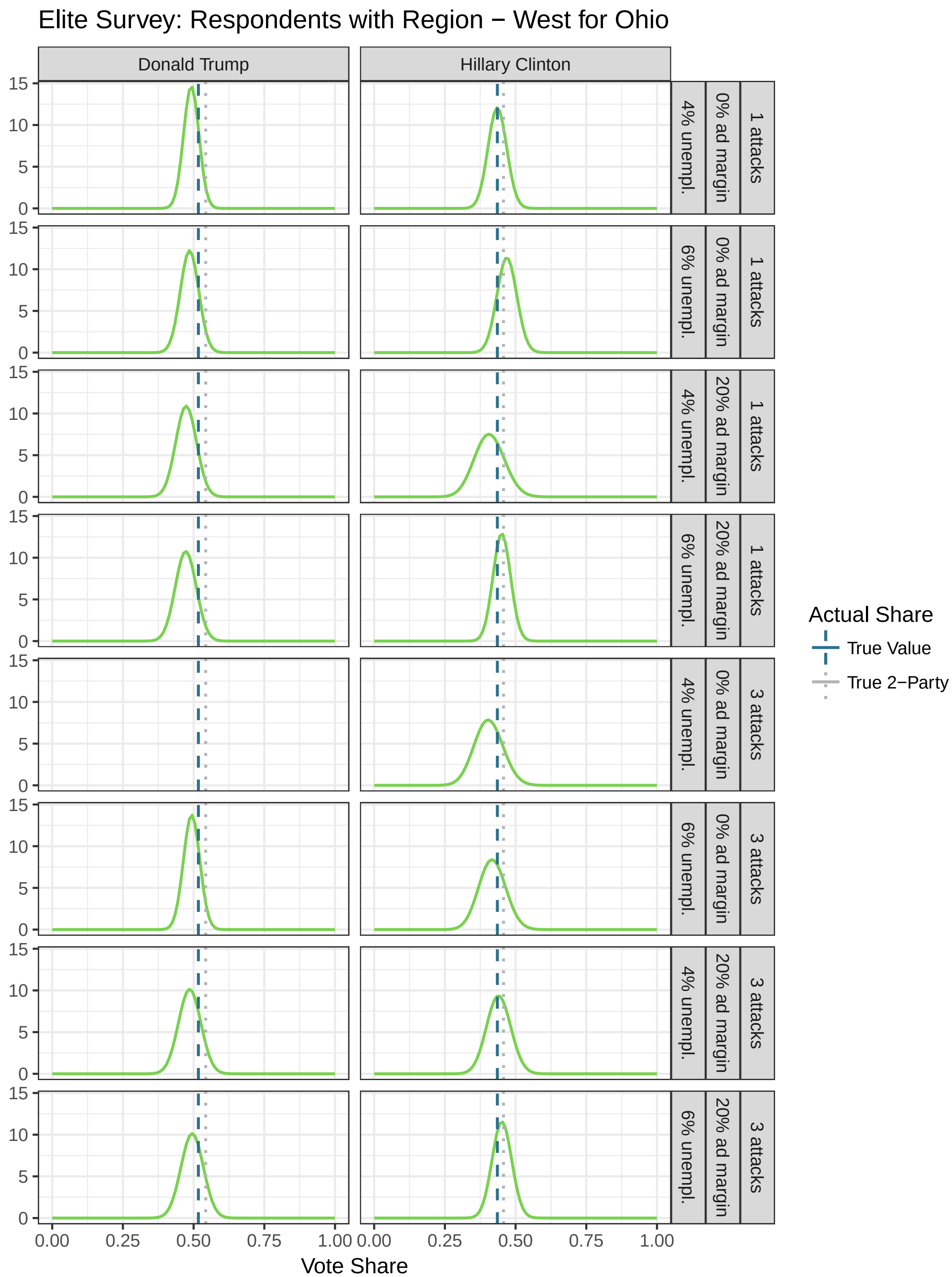


Figure 3.200: Priors with covariates: Elite Ohio Region West

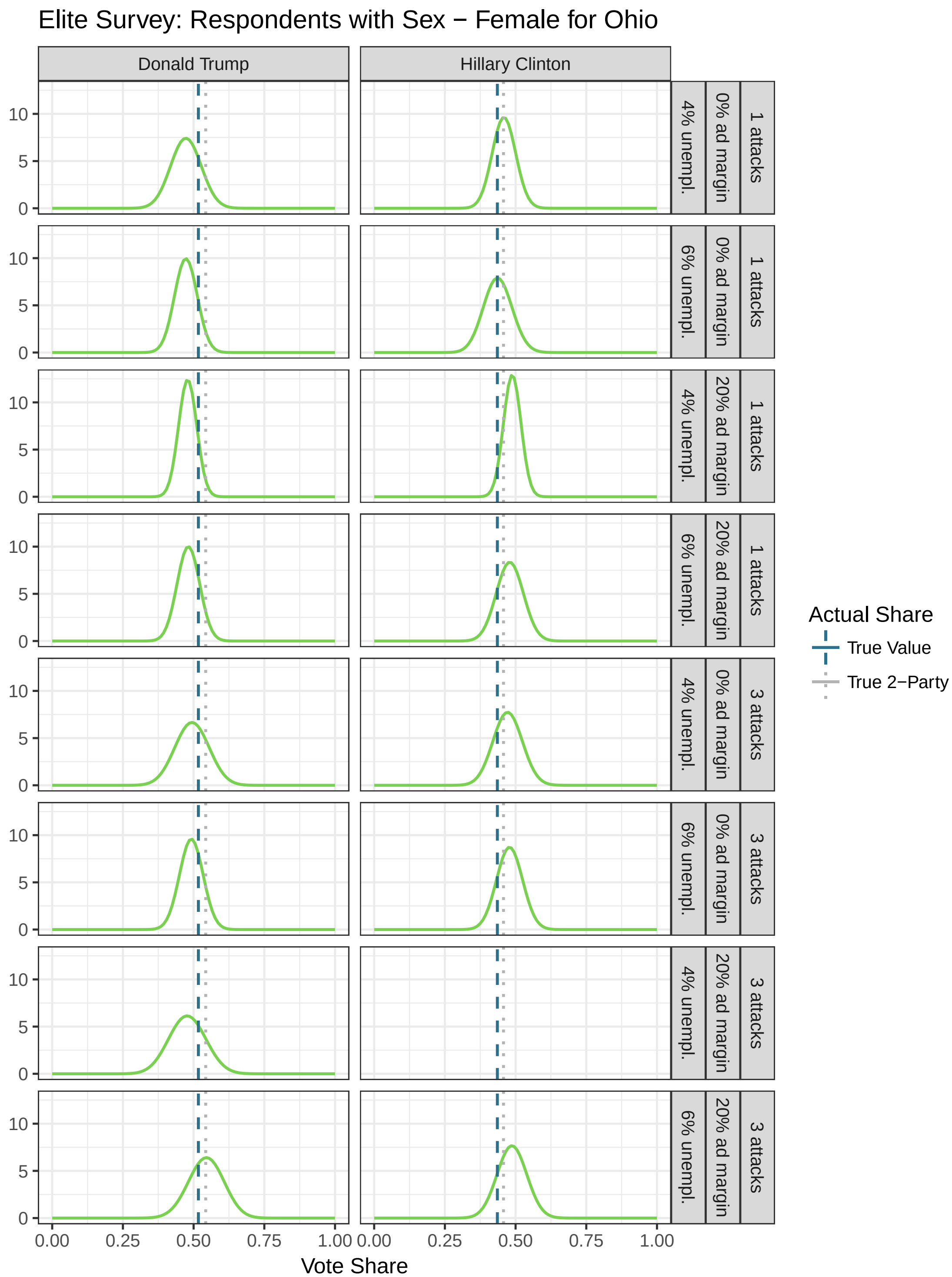


Figure 3.201: Priors with covariates: Elite Ohio Sex Female

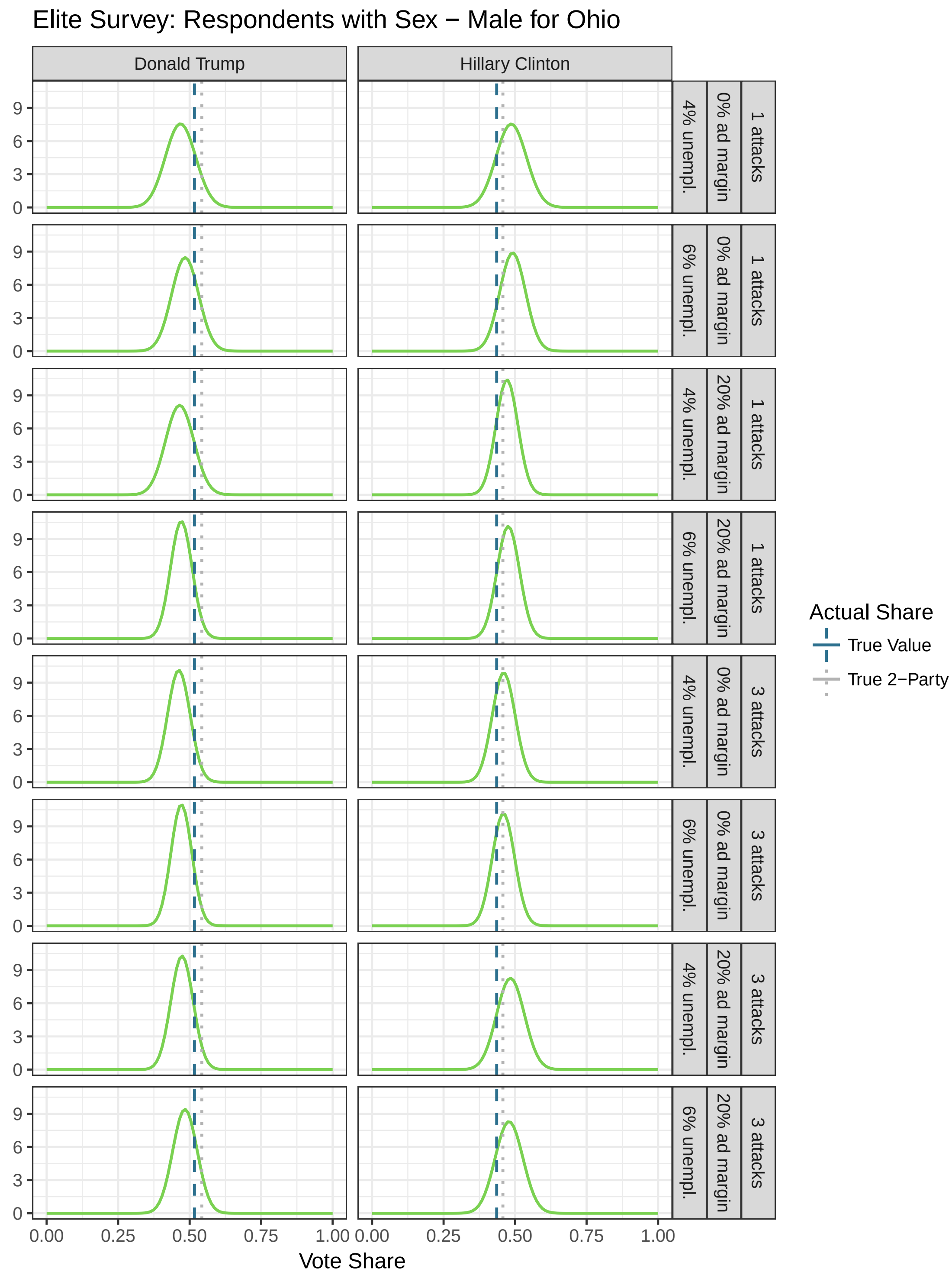


Figure 3.202: Priors with covariates: Elite Ohio Sex Male

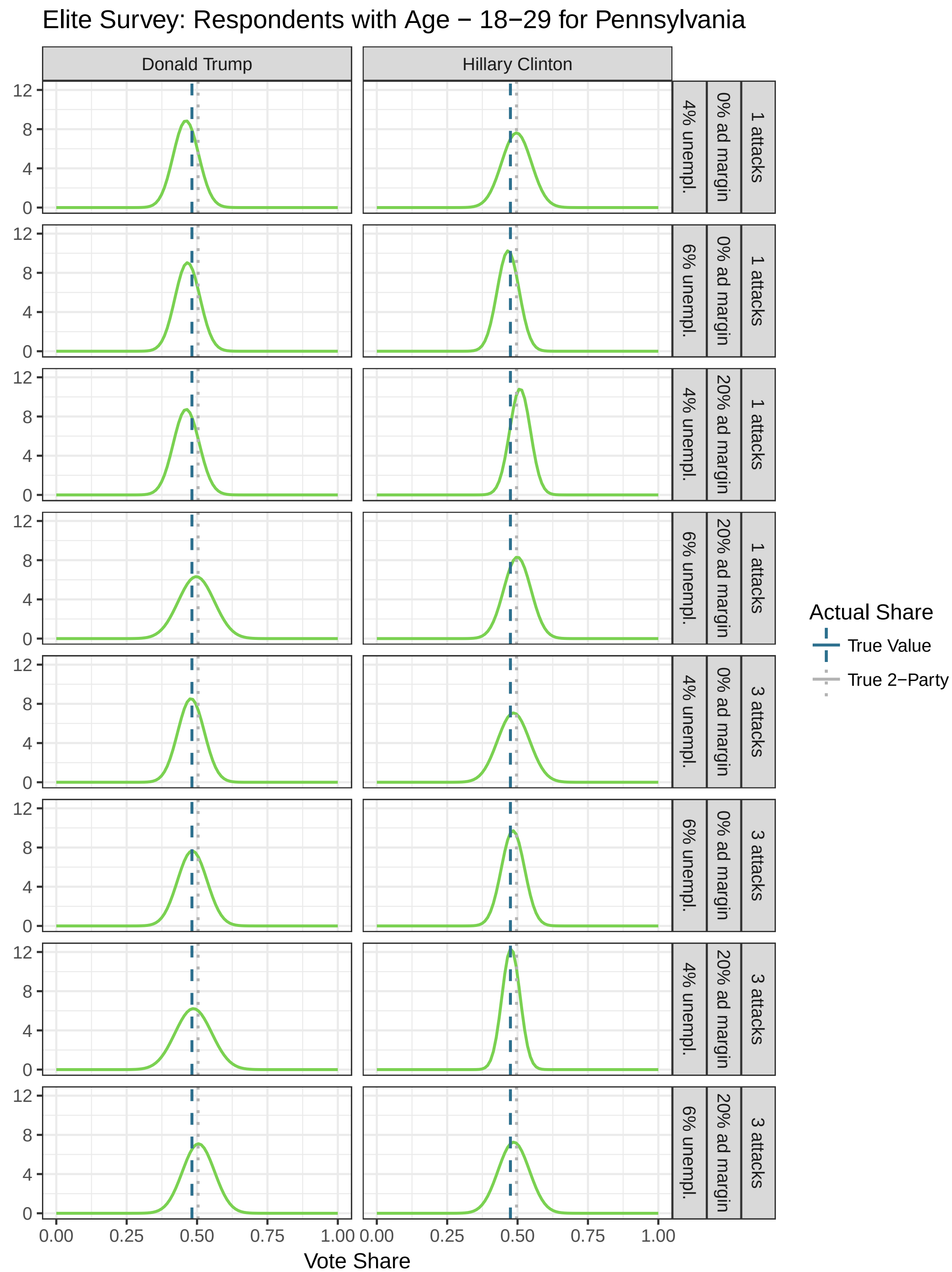


Figure 3.203: Priors with covariates: Elite Pennsylvania Age 18-29

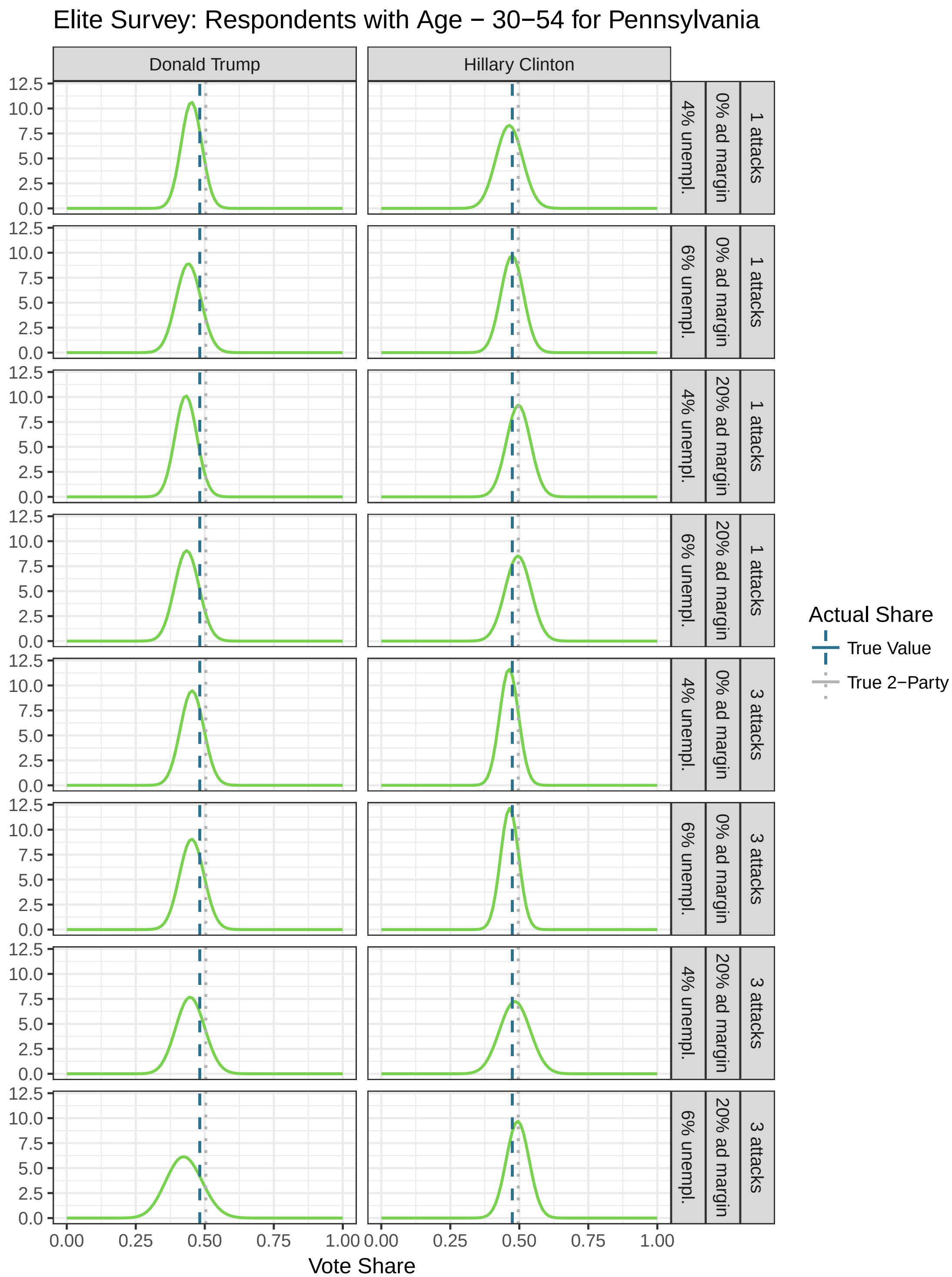


Figure 3.204: Priors with covariates: Elite Pennsylvania Age 30-54

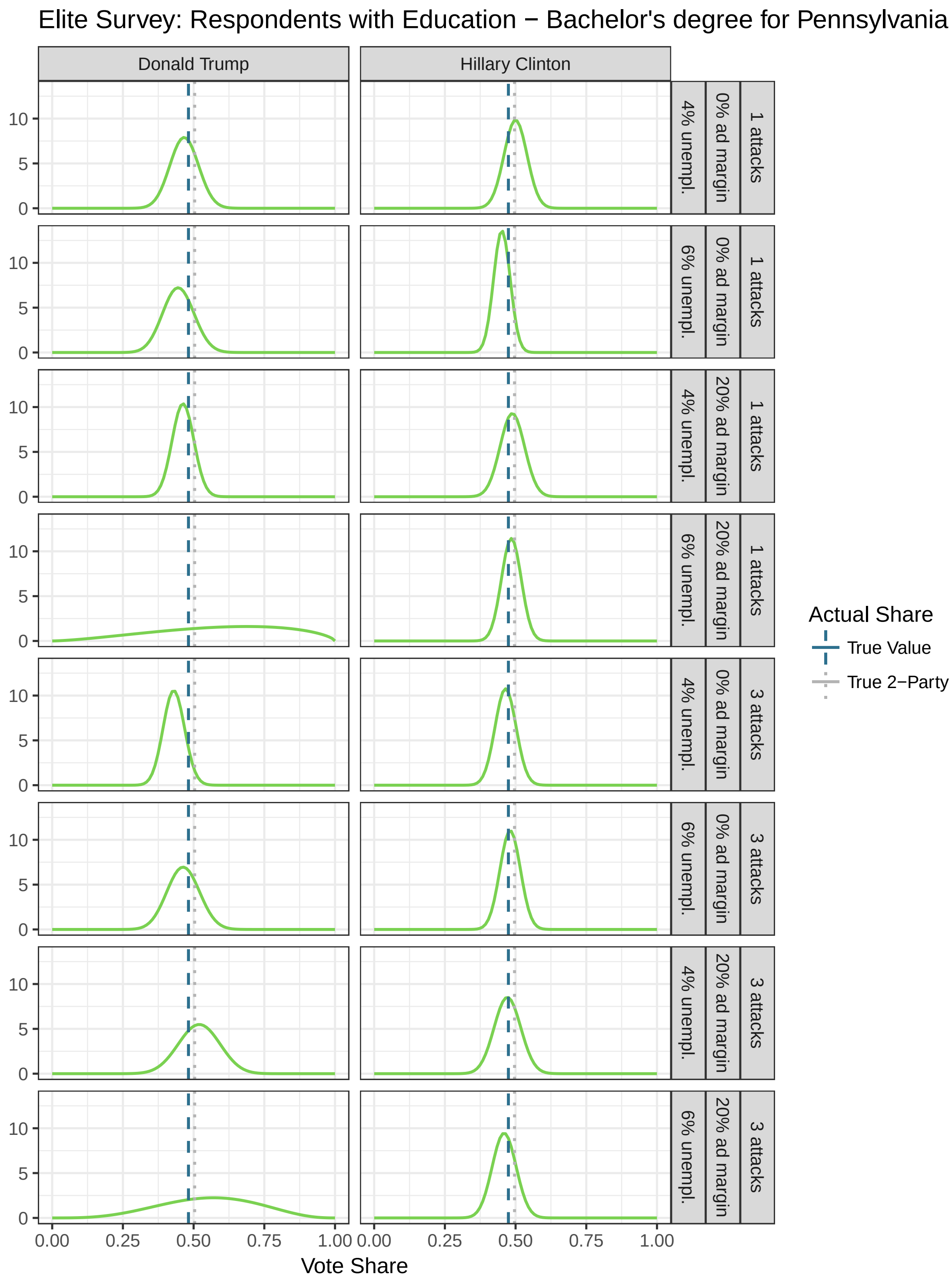


Figure 3.205: Priors with covariates: Elite Pennsylvania Education Bachelor's degree

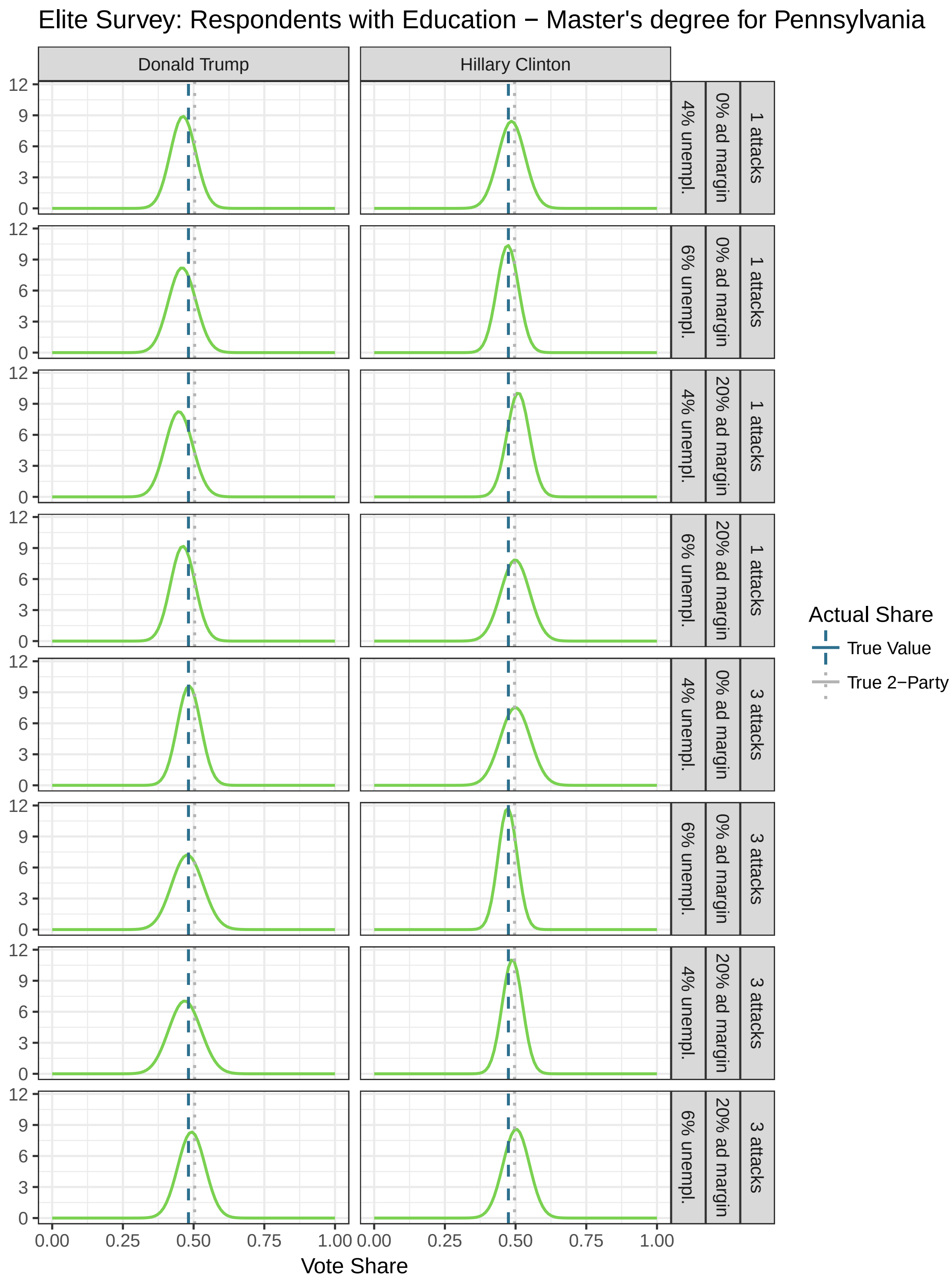


Figure 3.206: Priors with covariates: Elite Pennsylvania Education Master's degree

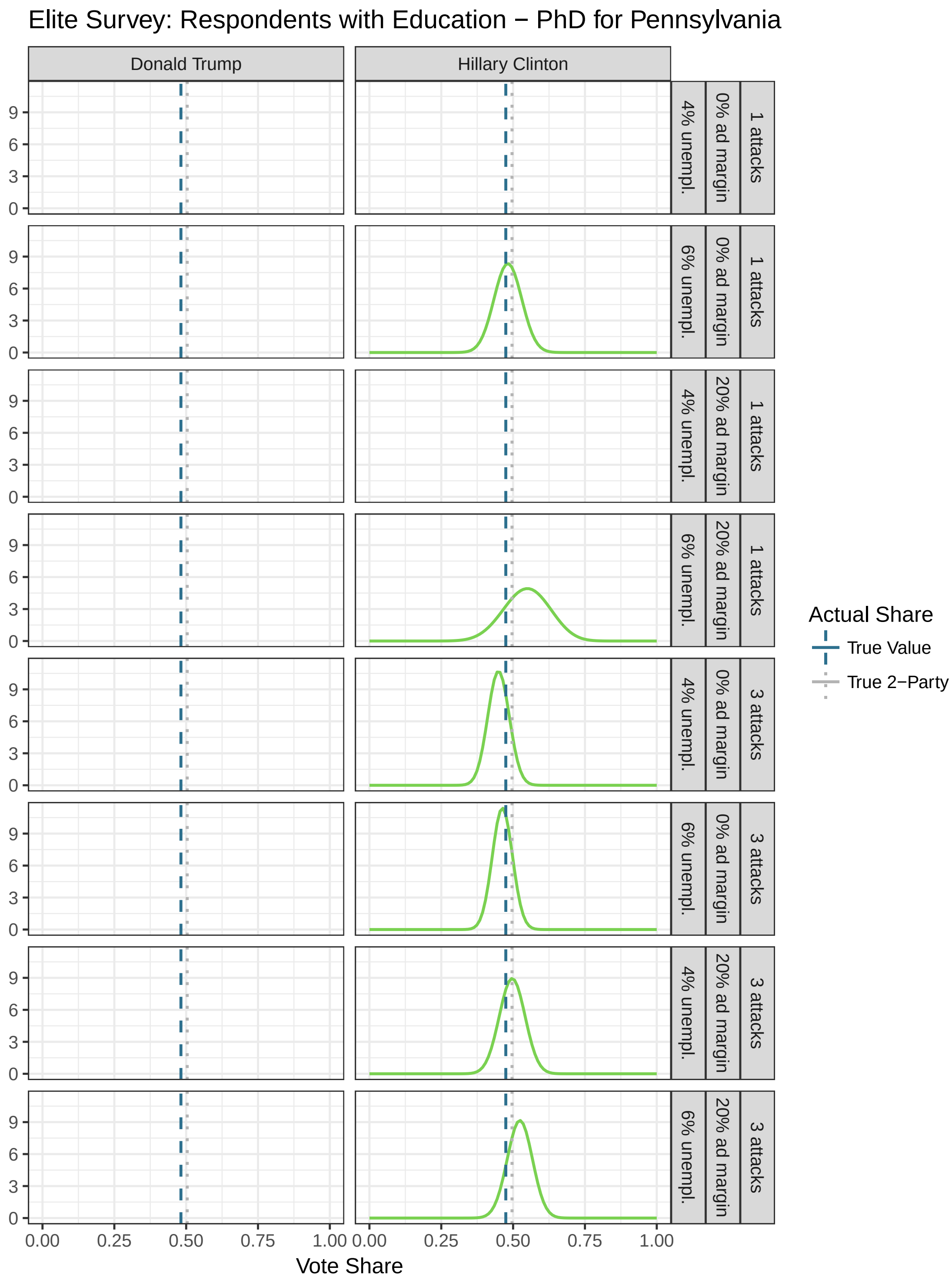


Figure 3.207: Priors with covariates: Elite Pennsylvania Education PhD

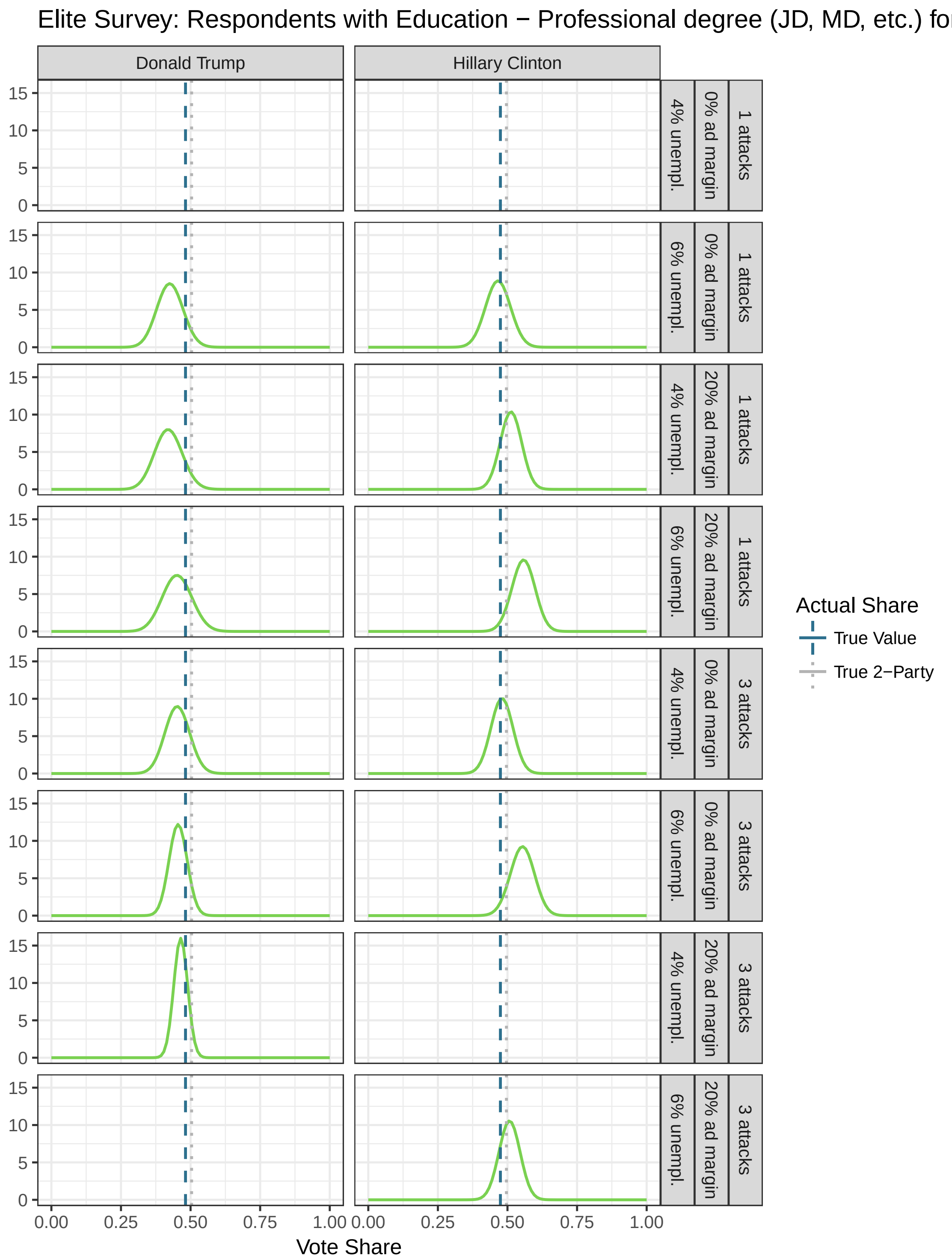


Figure 3.208: Priors with covariates: Elite Pennsylvania Education Professional degree JD MD etc

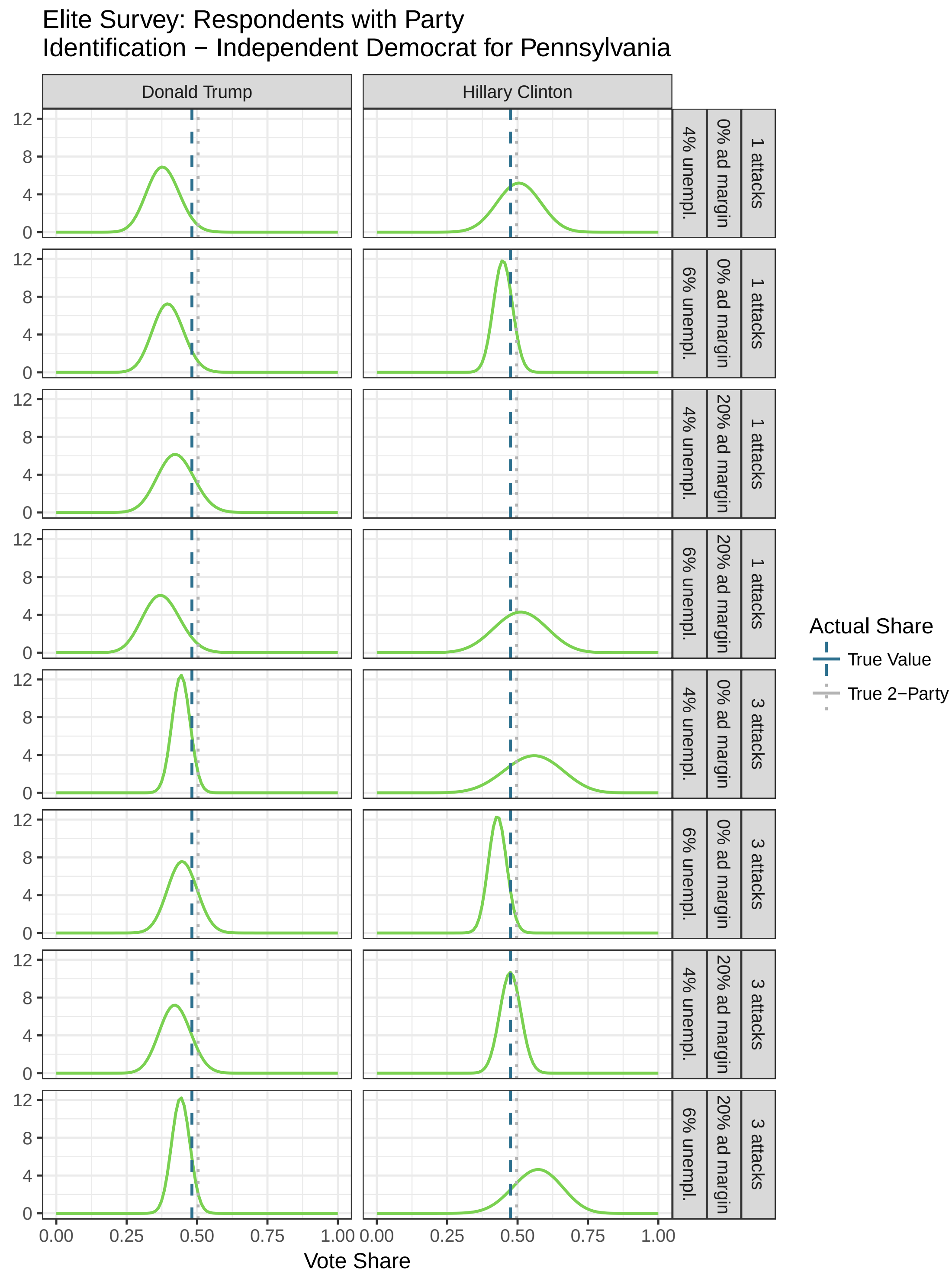


Figure 3.209: Priors with covariates: Elite Pennsylvania Party Identification Independent Democrat

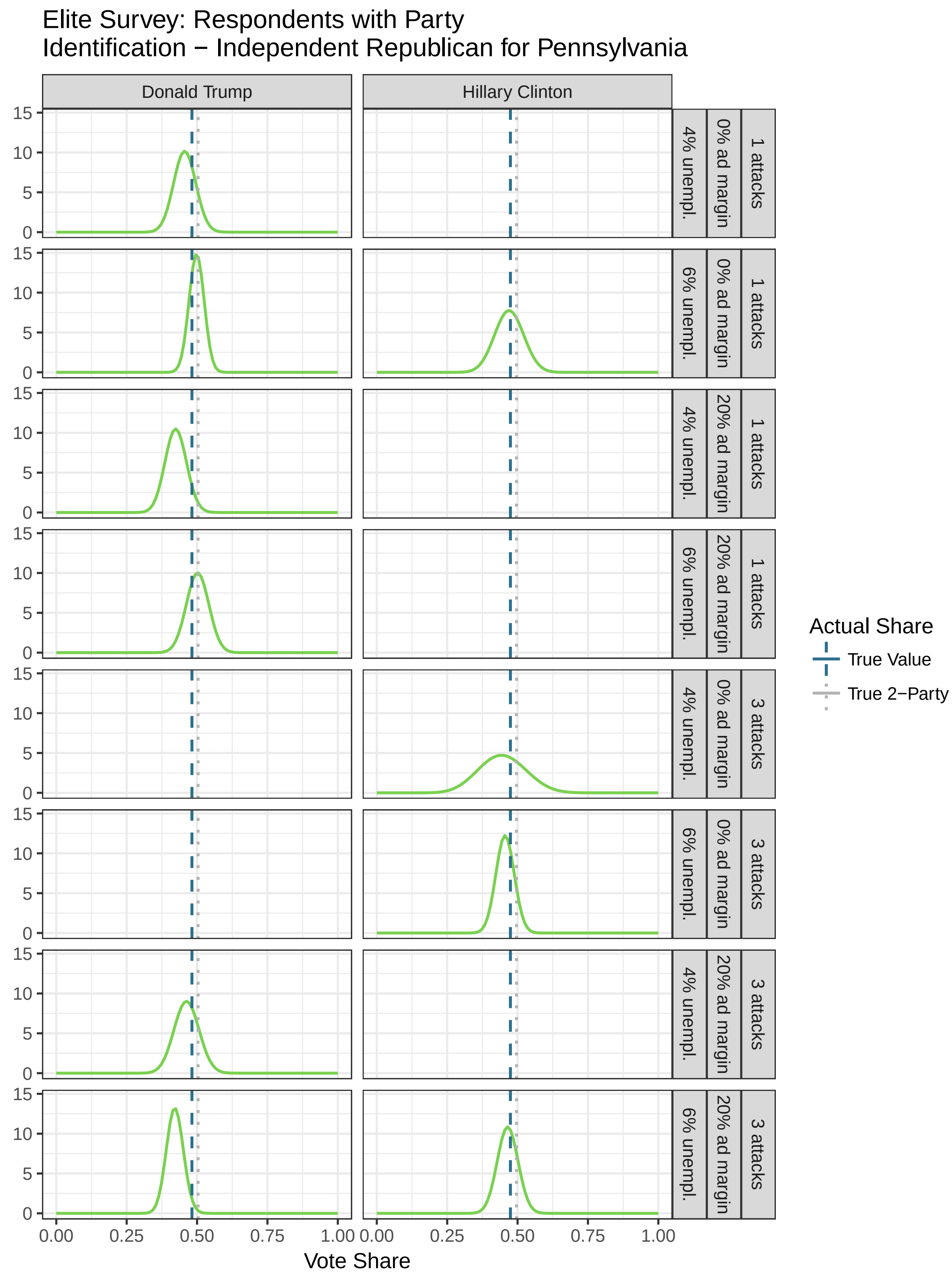


Figure 3.210: Priors with covariates: Elite Pennsylvania Party Identification Independent Republican

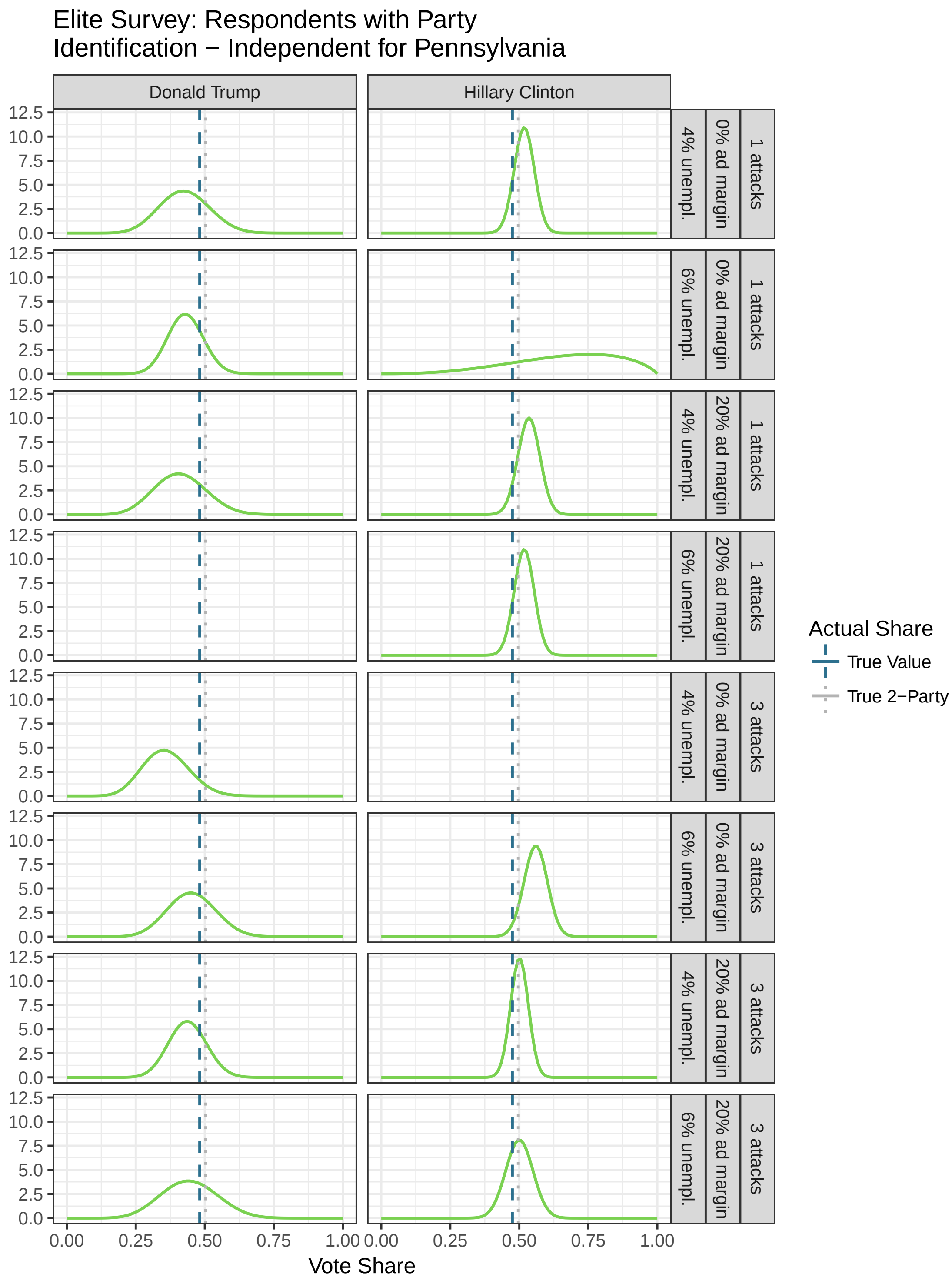


Figure 3.211: Priors with covariates: Elite Pennsylvania Party Identification Independent

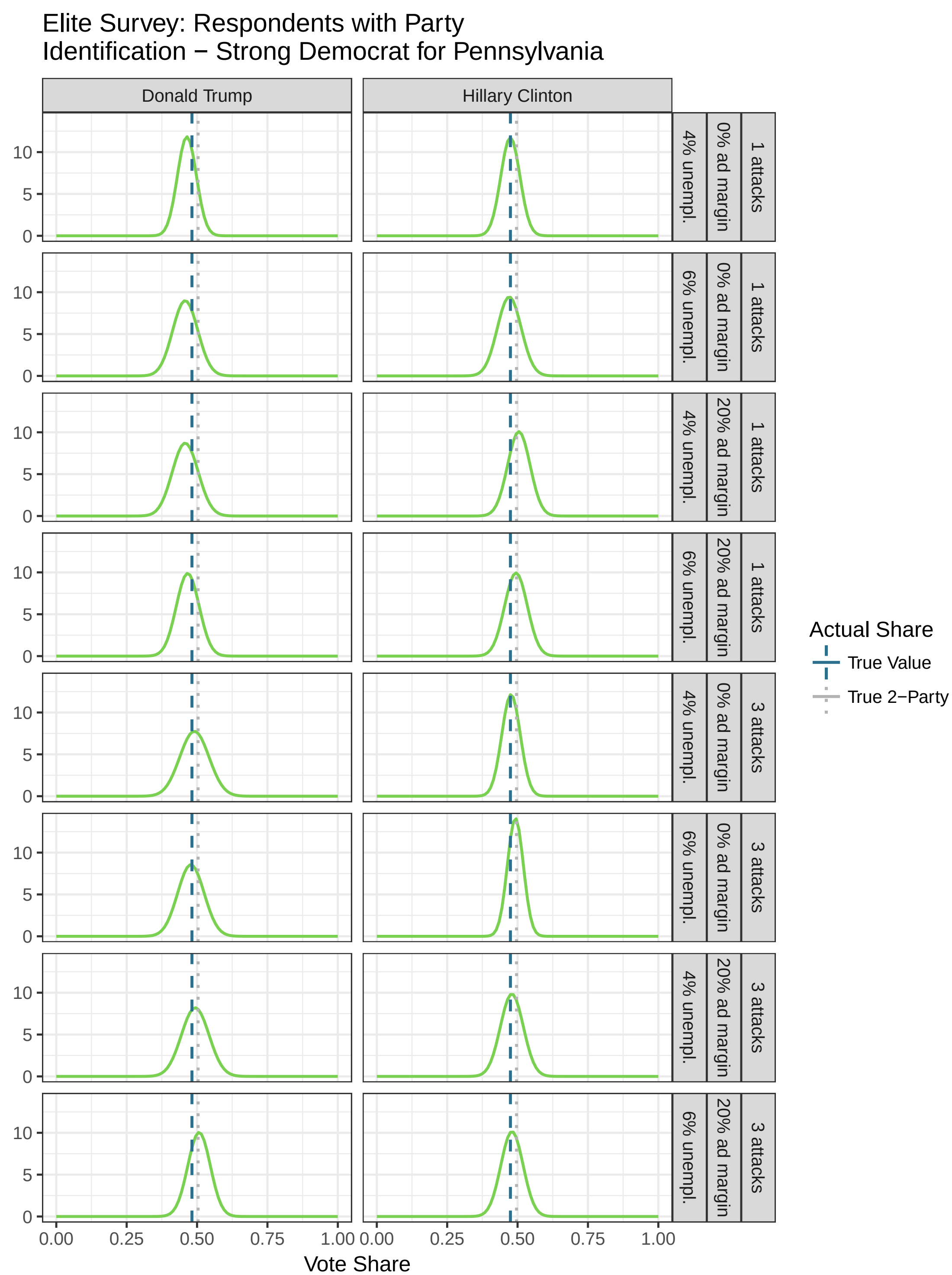


Figure 3.212: Priors with covariates: Elite Pennsylvania Party Identification Strong Democrat

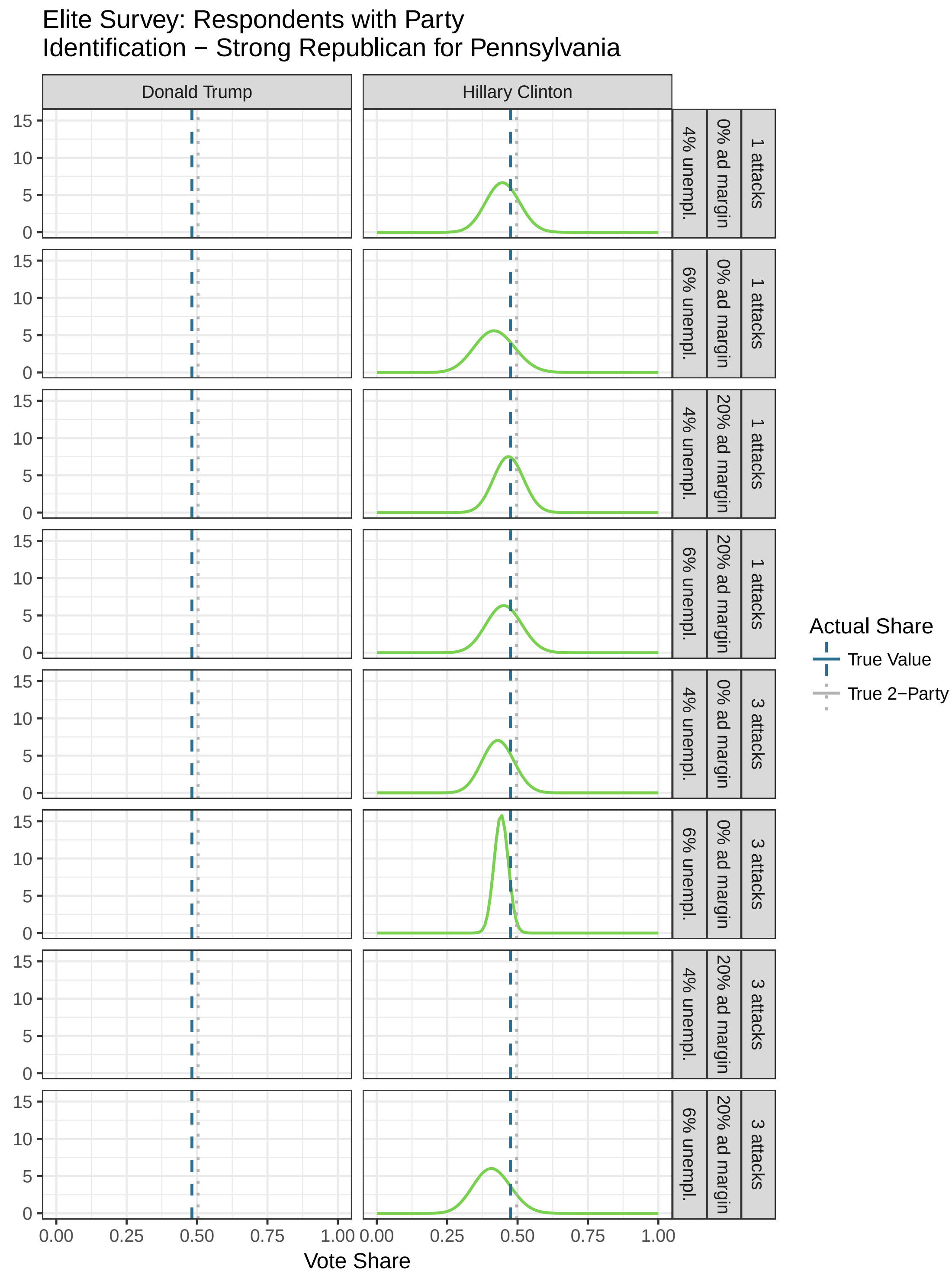


Figure 3.213: Priors with covariates: Elite Pennsylvania Party Identification Strong Republican

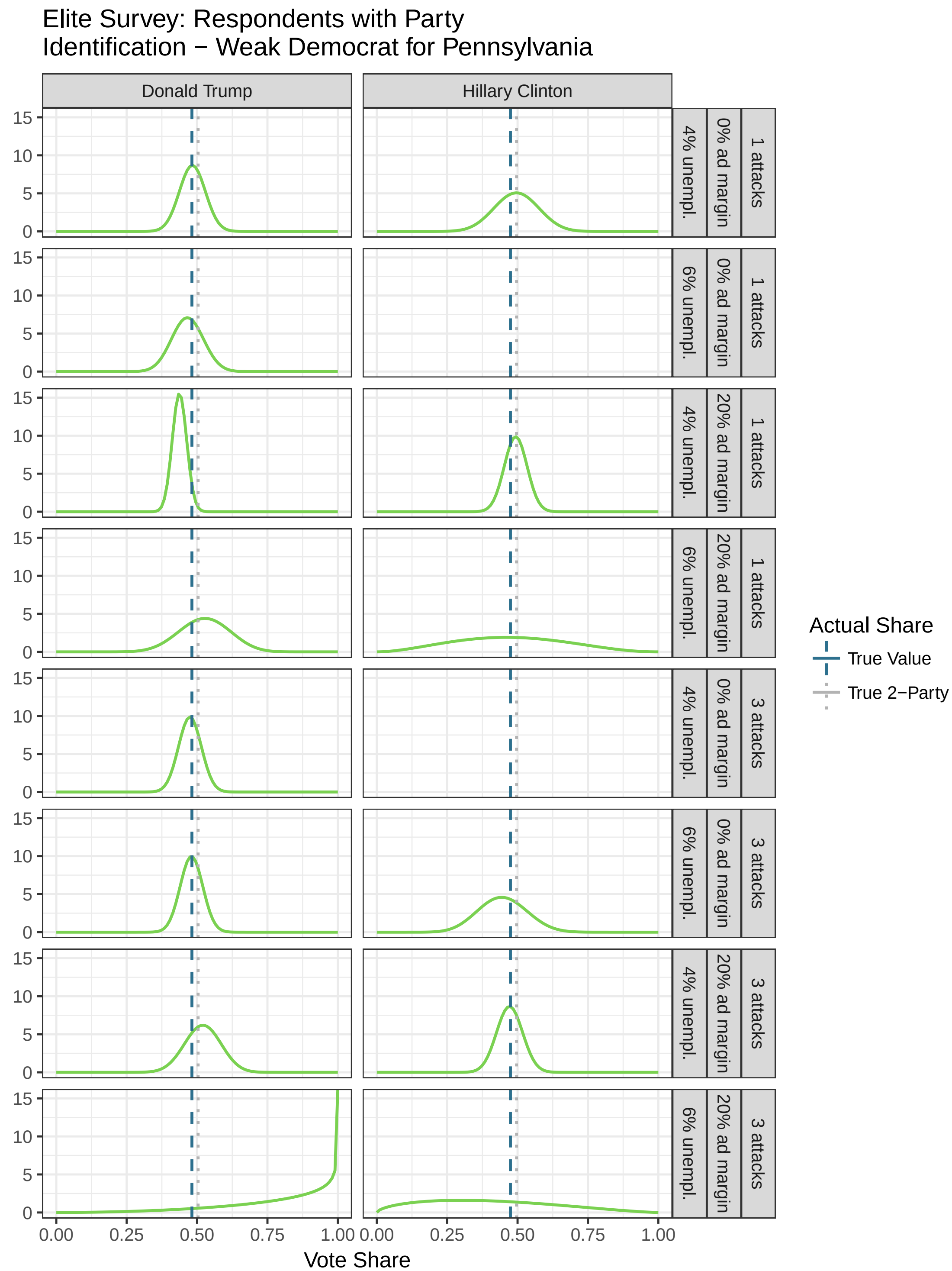


Figure 3.214: Priors with covariates: Elite Pennsylvania Party Identification Weak Democrat

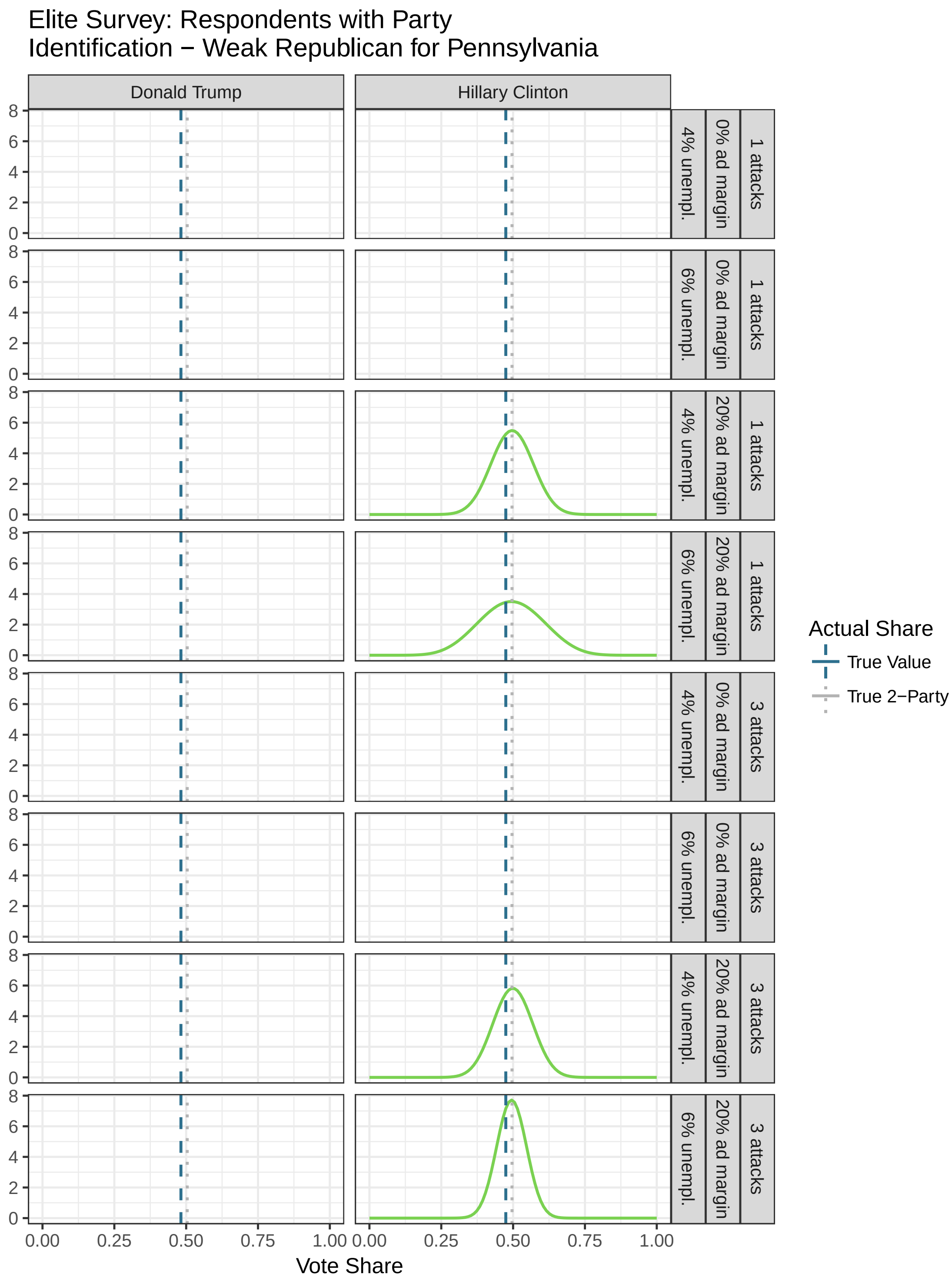


Figure 3.215: Priors with covariates: Elite Pennsylvania Party Identification Weak Republican

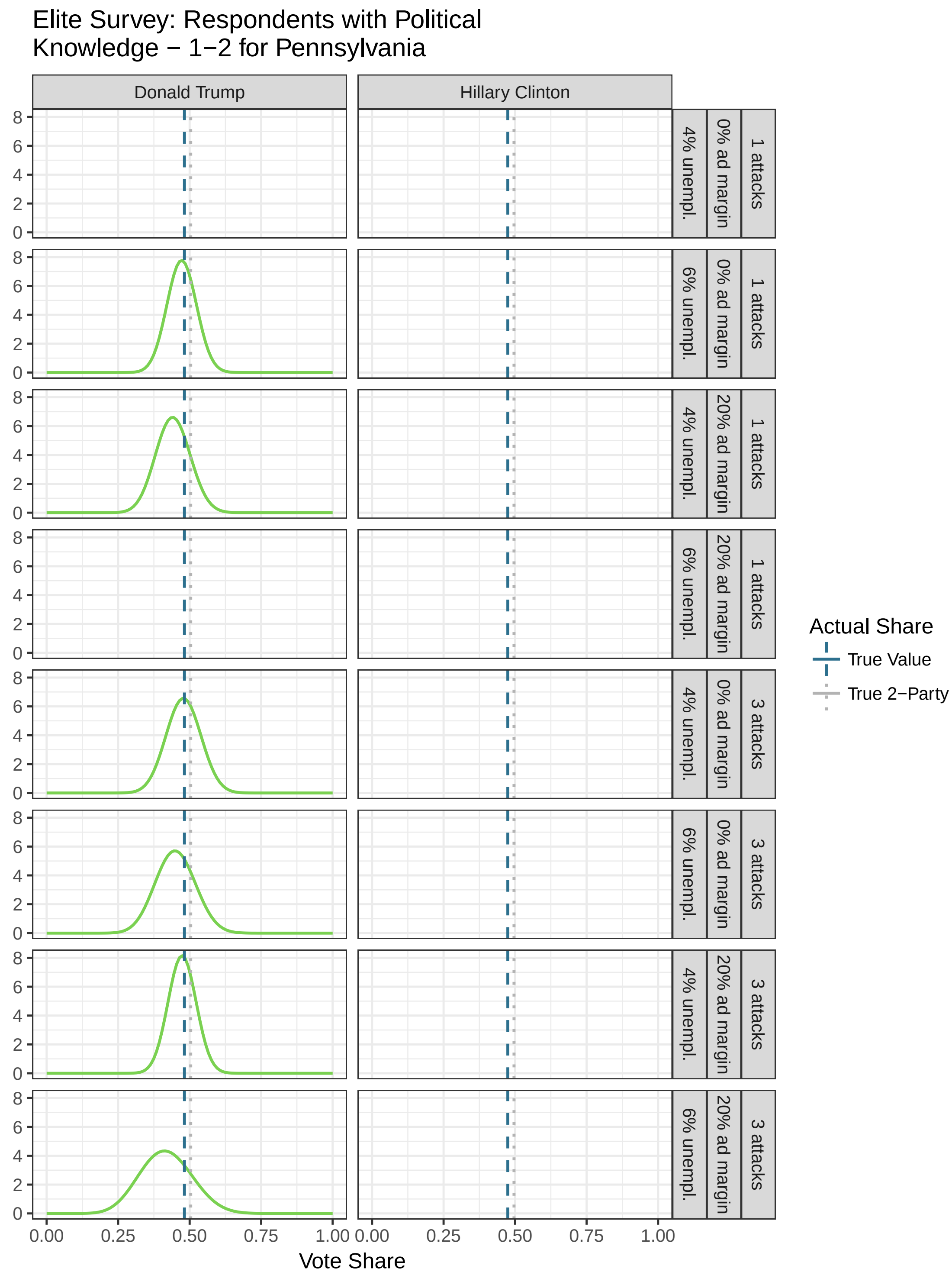


Figure 3.216: Priors with covariates: Elite Pennsylvania Political Knowledge 1-2

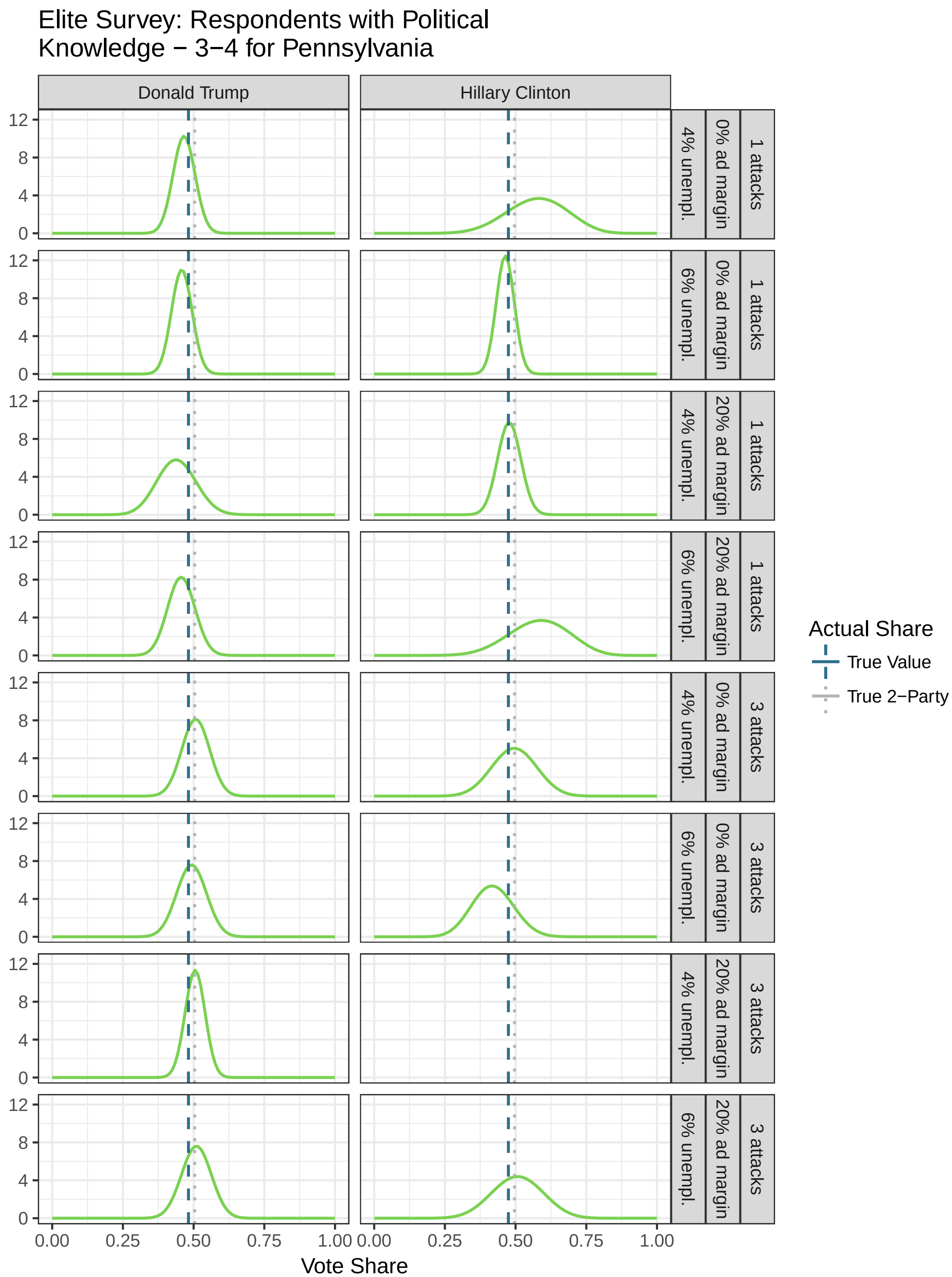


Figure 3.217: Priors with covariates: Elite Pennsylvania Political Knowledge 3-4

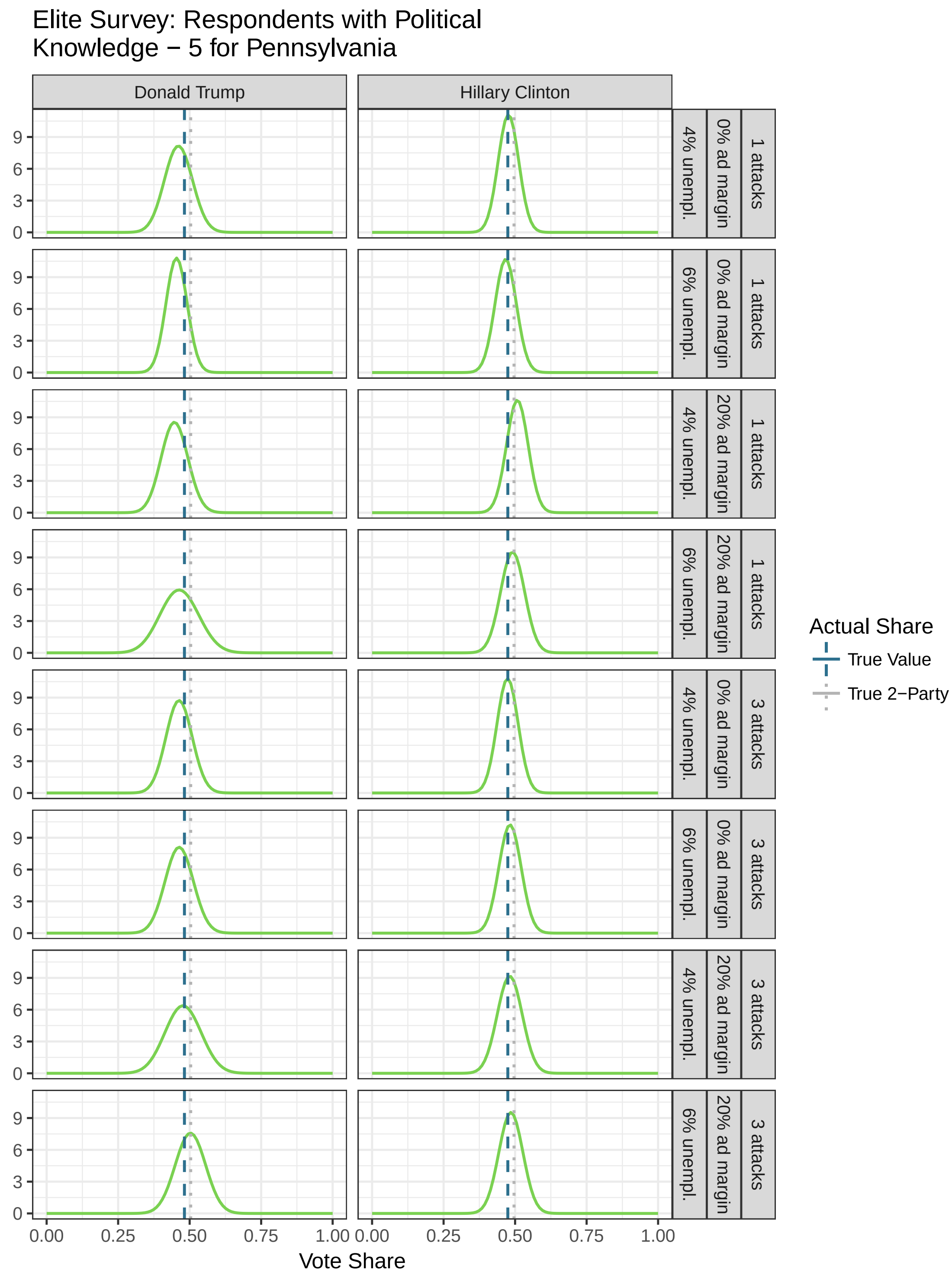


Figure 3.218: Priors with covariates: Elite Pennsylvania Political Knowledge 5

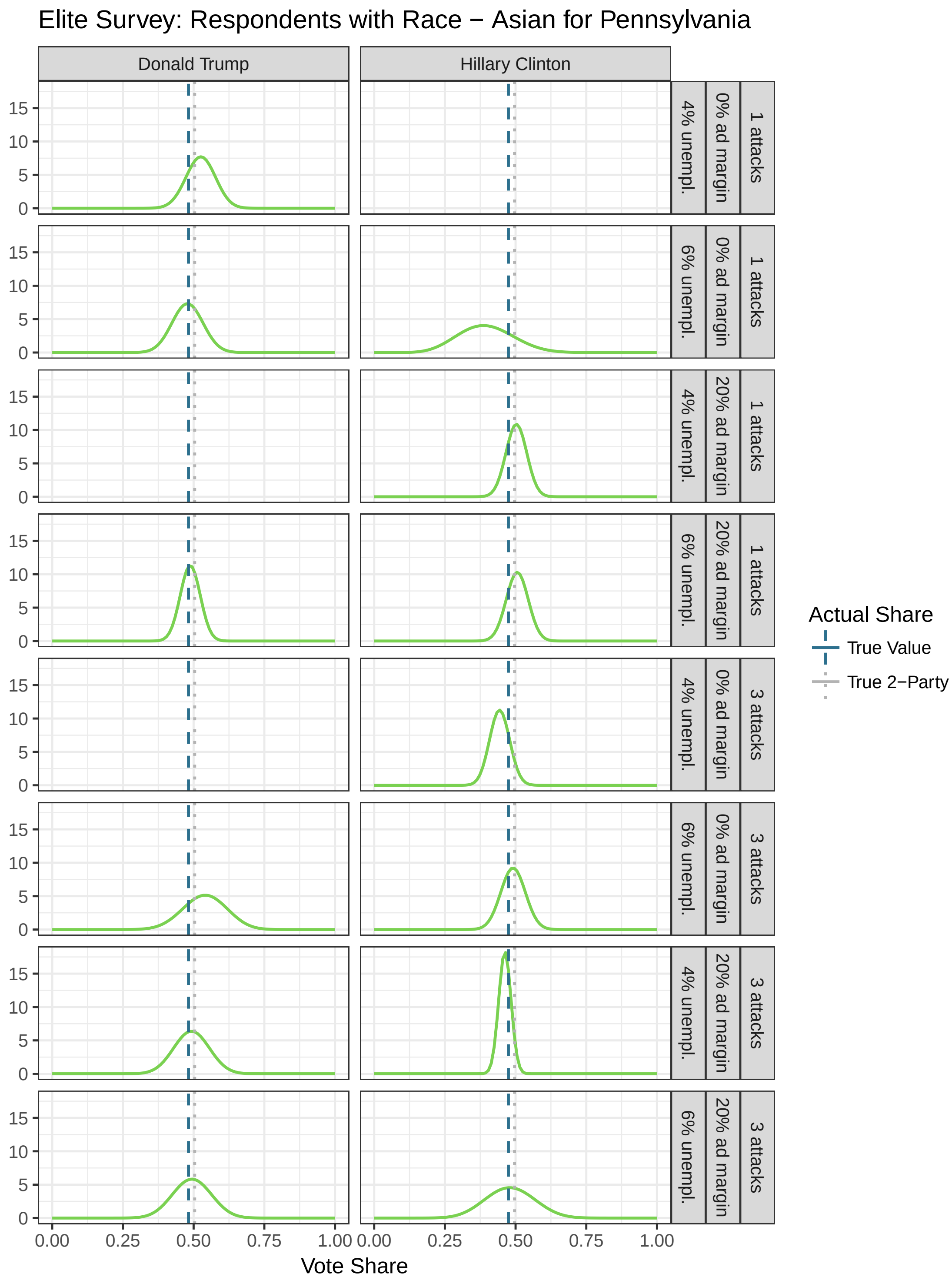


Figure 3.219: Priors with covariates: Elite Pennsylvania Race Asian

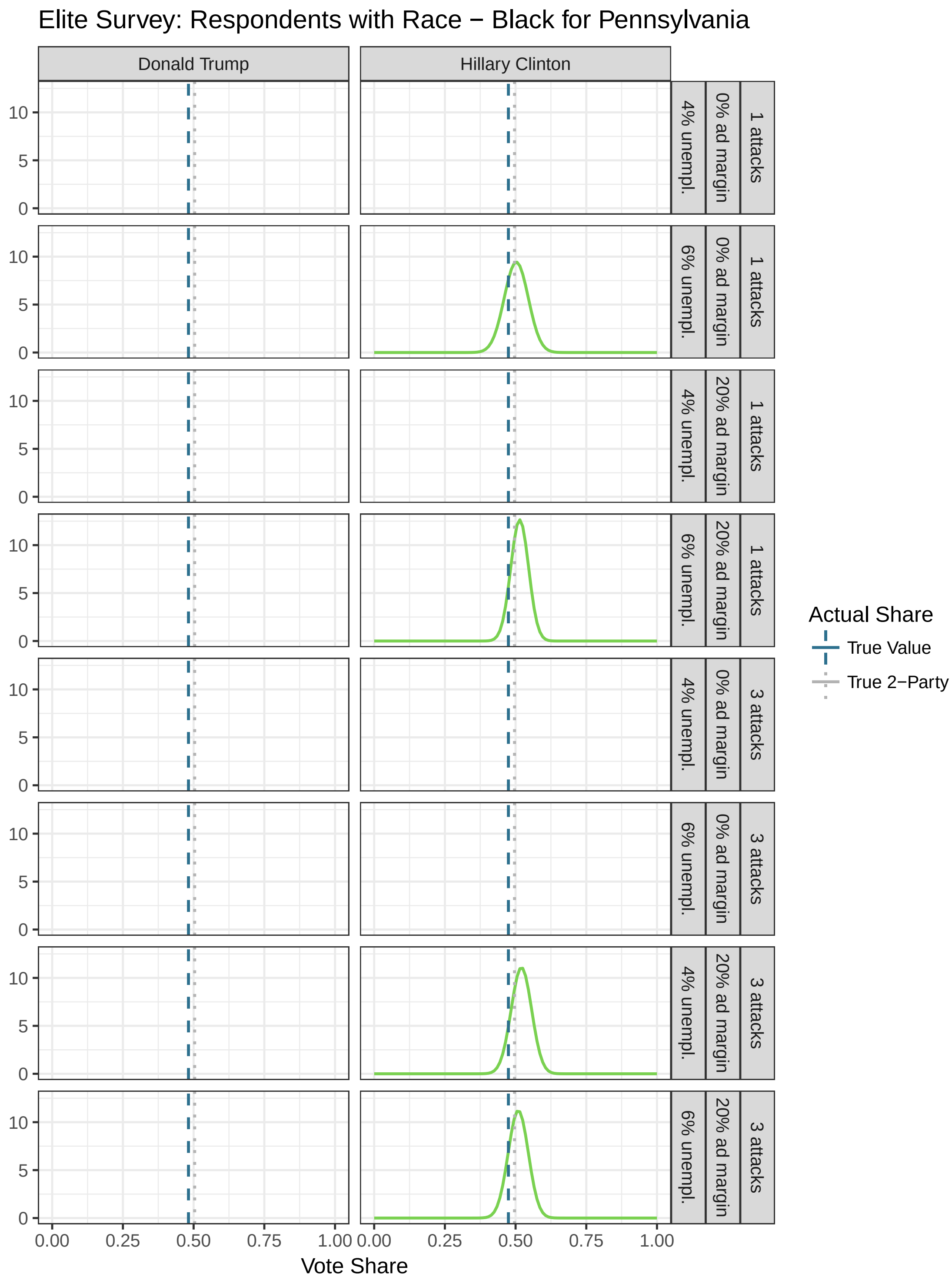


Figure 3.220: Priors with covariates: Elite Pennsylvania Race Black

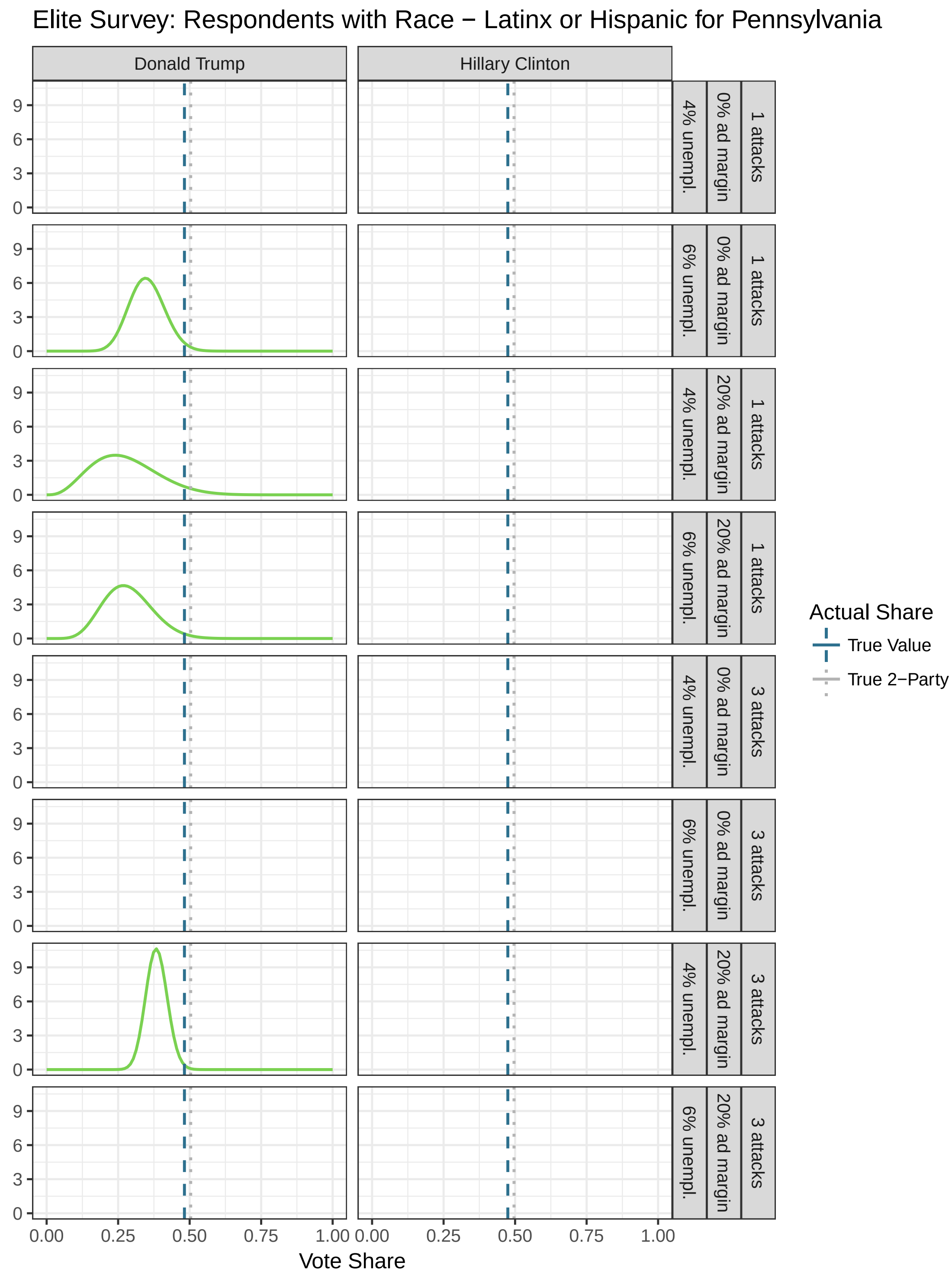


Figure 3.221: Priors with covariates: Elite Pennsylvania Race Latinx or Hispanic

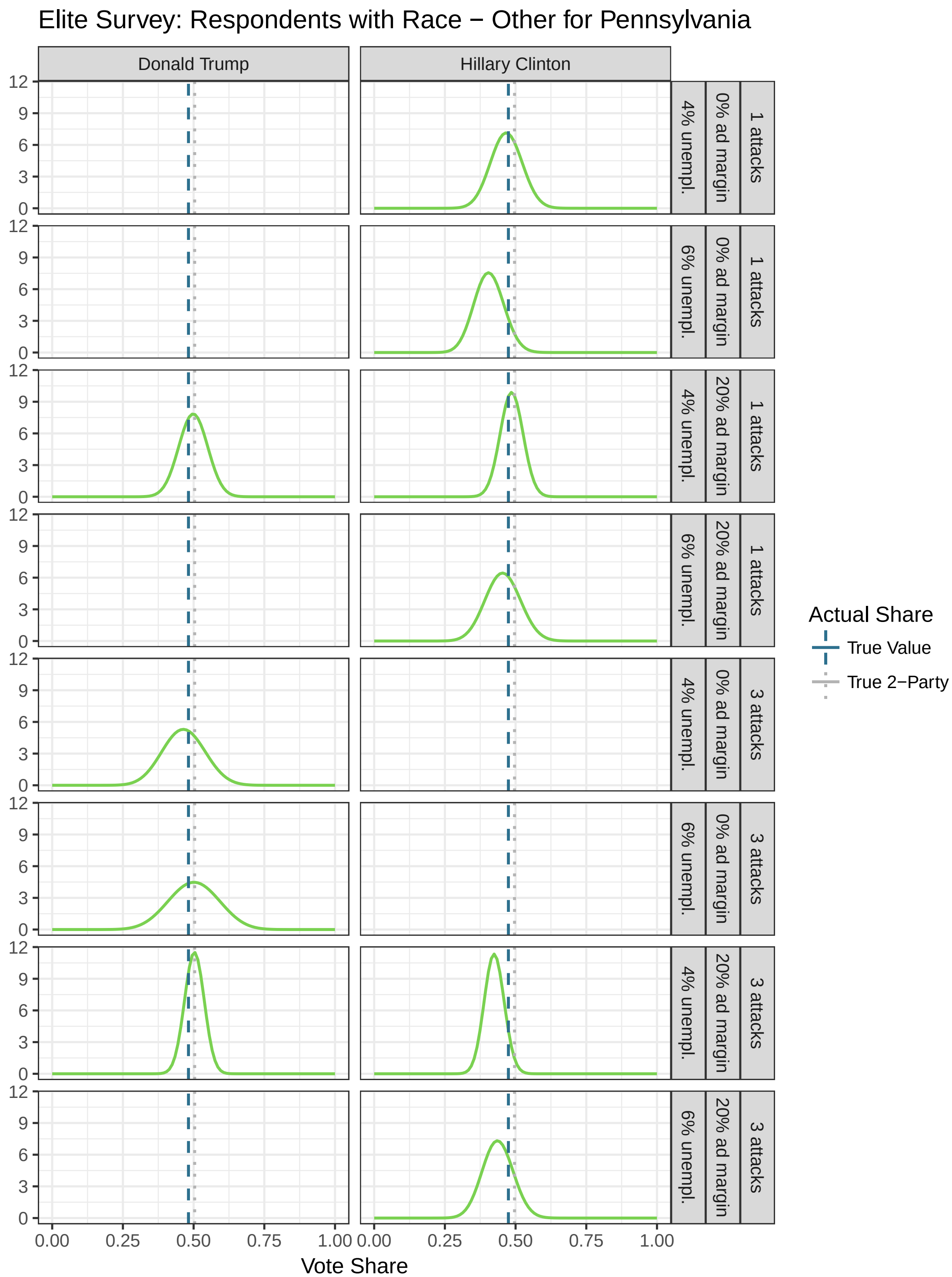


Figure 3.222: Priors with covariates: Elite Pennsylvania Race Other

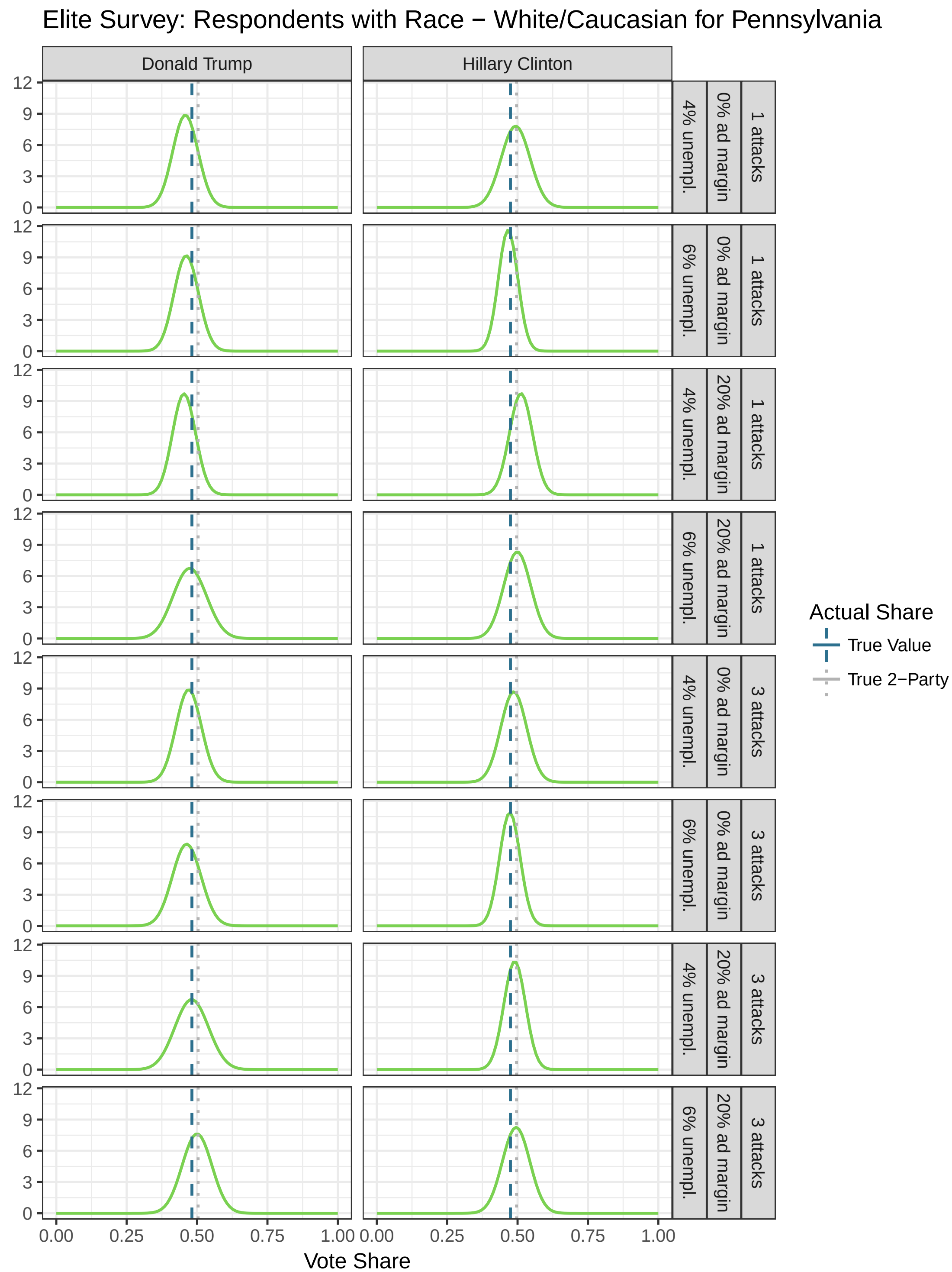


Figure 3.223: Priors with covariates: Elite Pennsylvania Race White Caucasian

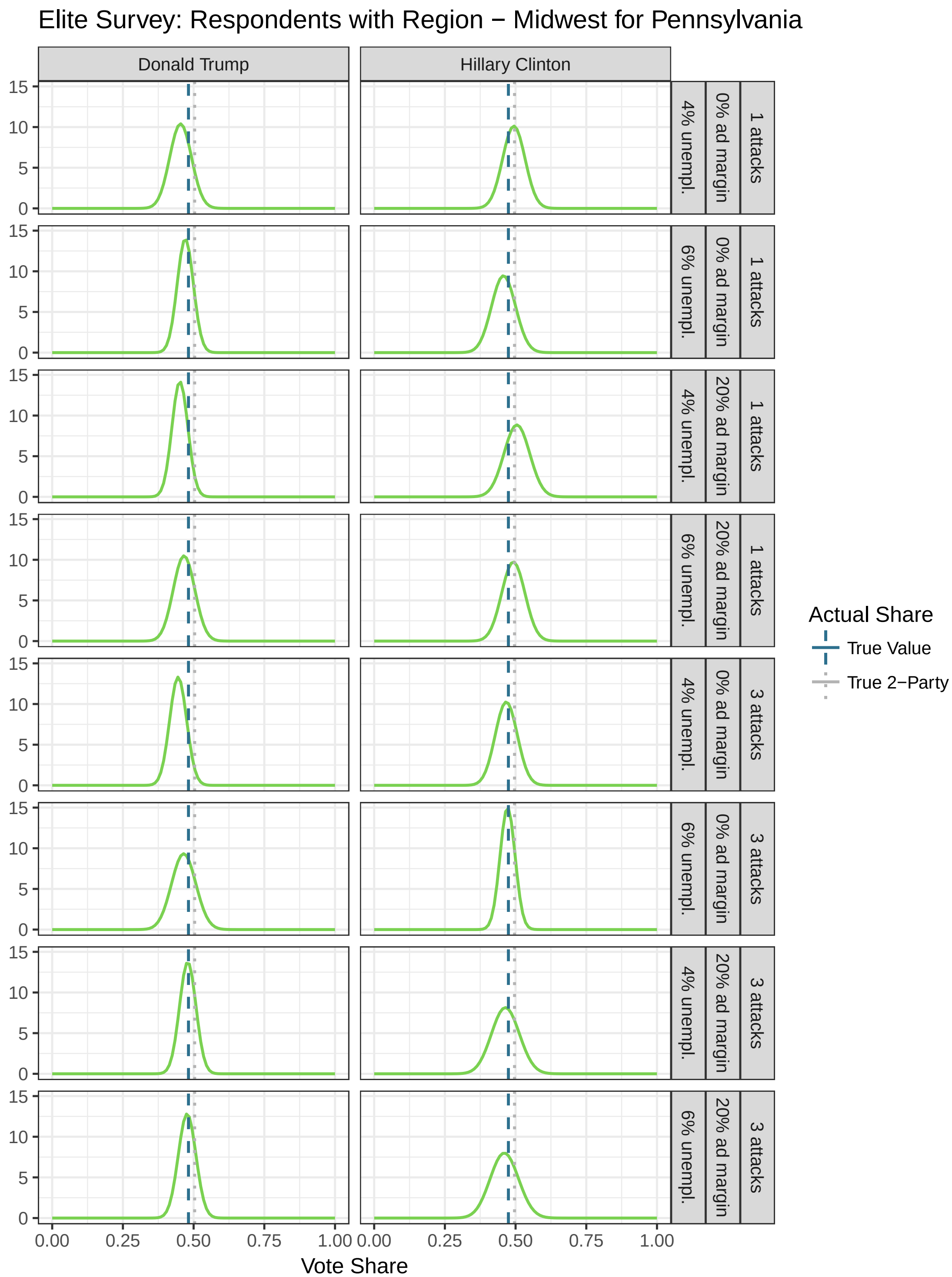


Figure 3.224: Priors with covariates: Elite Pennsylvania Region Midwest

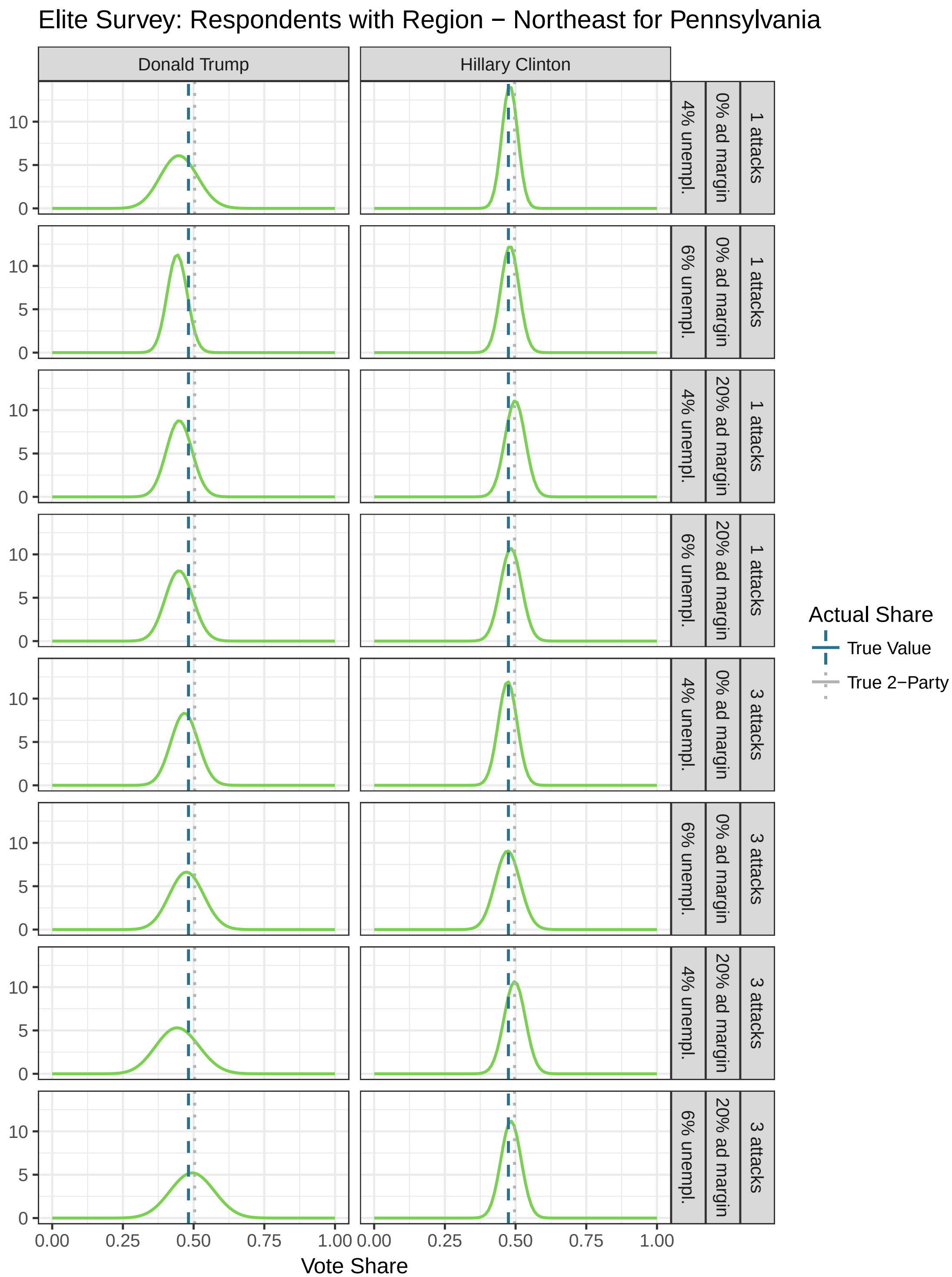


Figure 3.225: Priors with covariates: Elite Pennsylvania Region Northeast

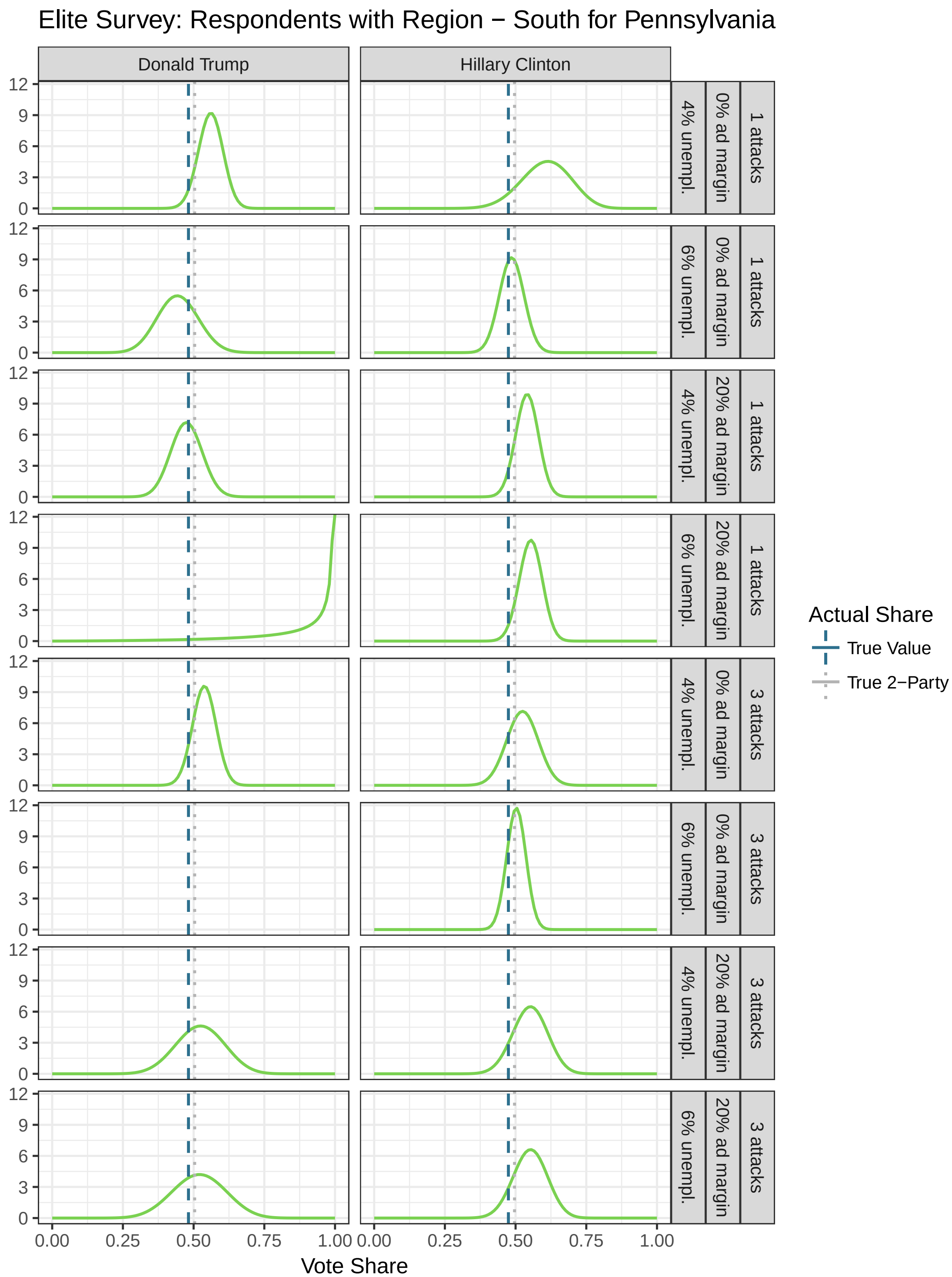


Figure 3.226: Priors with covariates: Elite Pennsylvania Region South

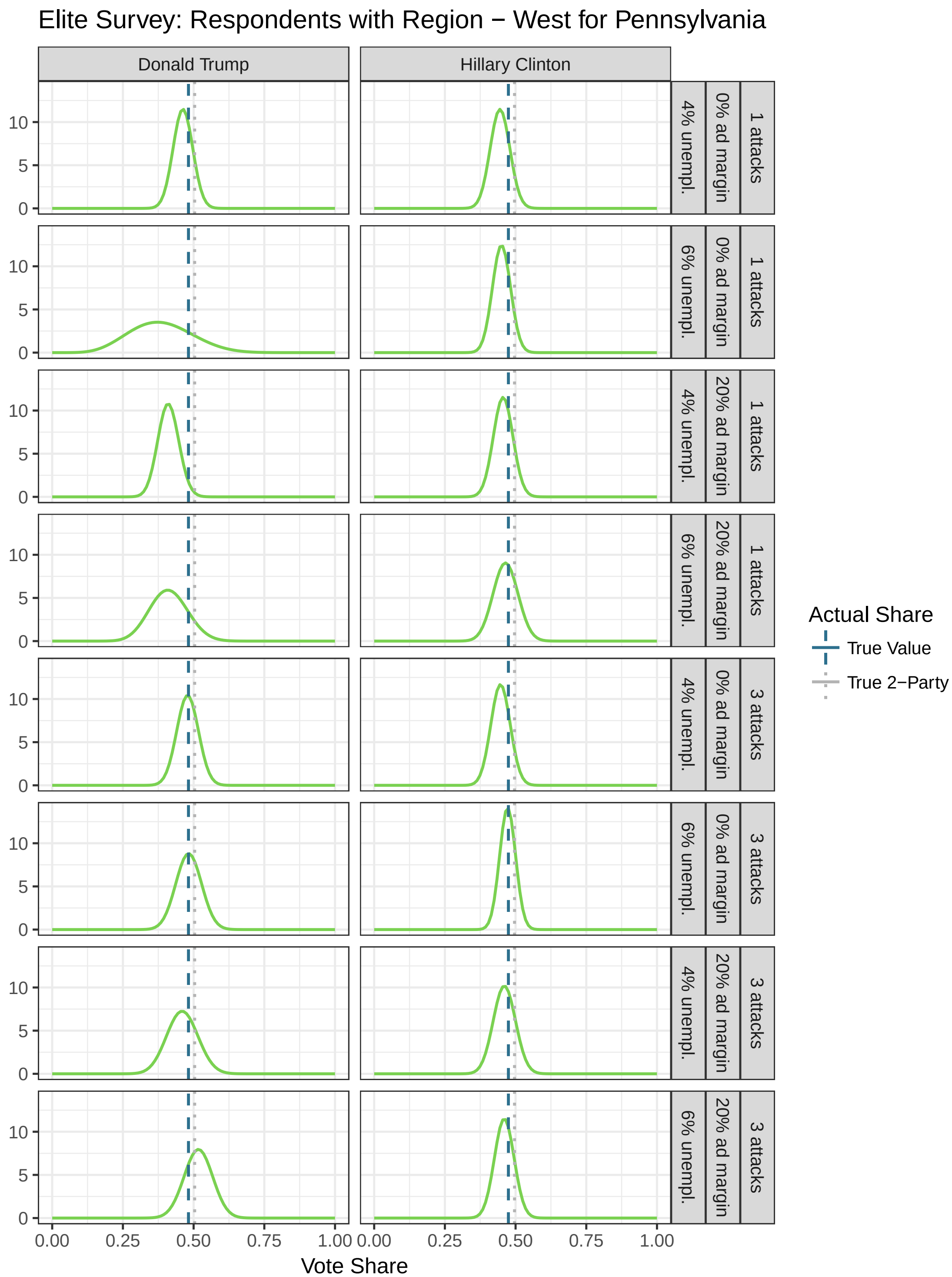


Figure 3.227: Priors with covariates: Elite Pennsylvania Region West

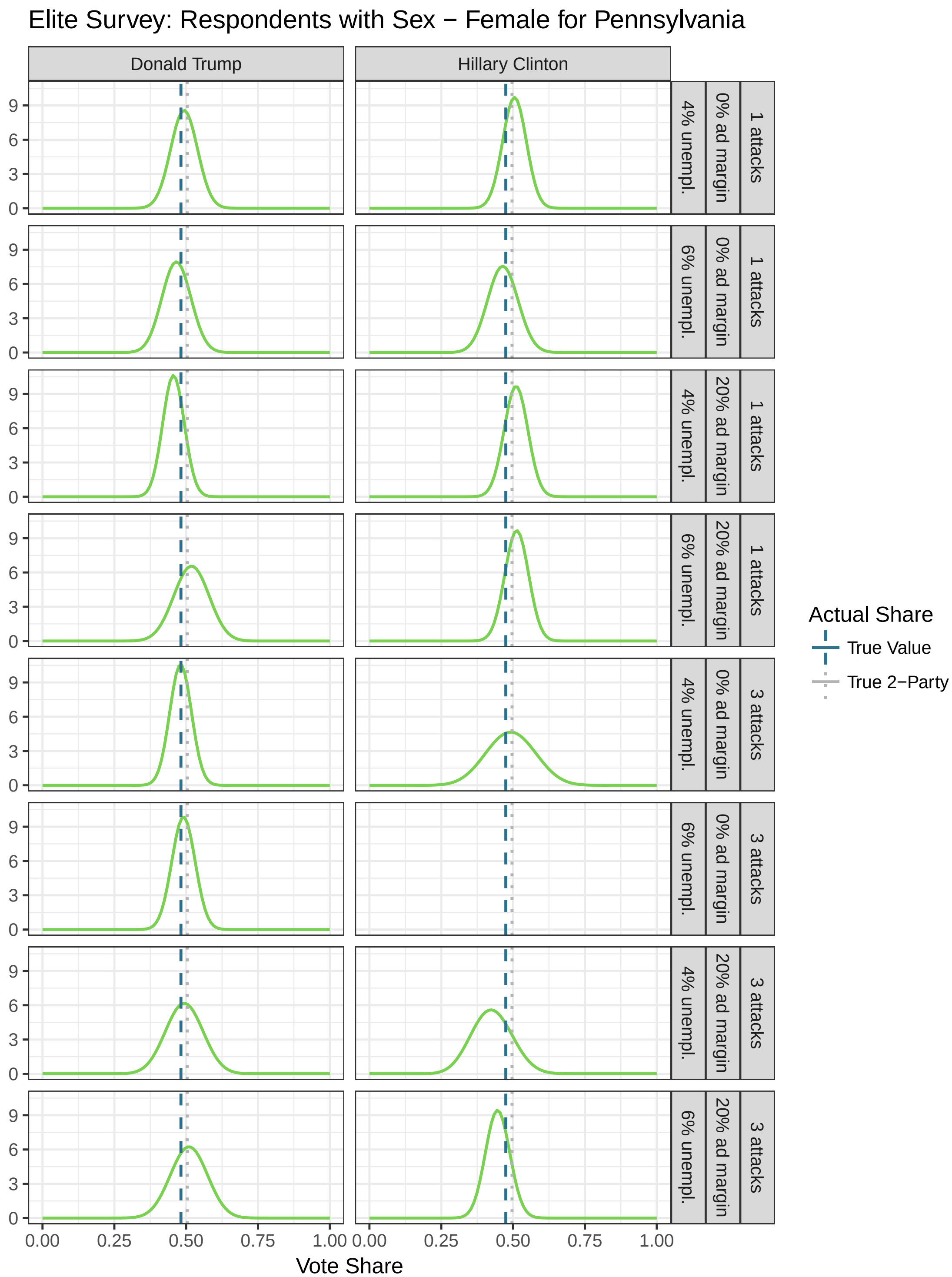


Figure 3.228: Priors with covariates: Elite Pennsylvania Sex Female

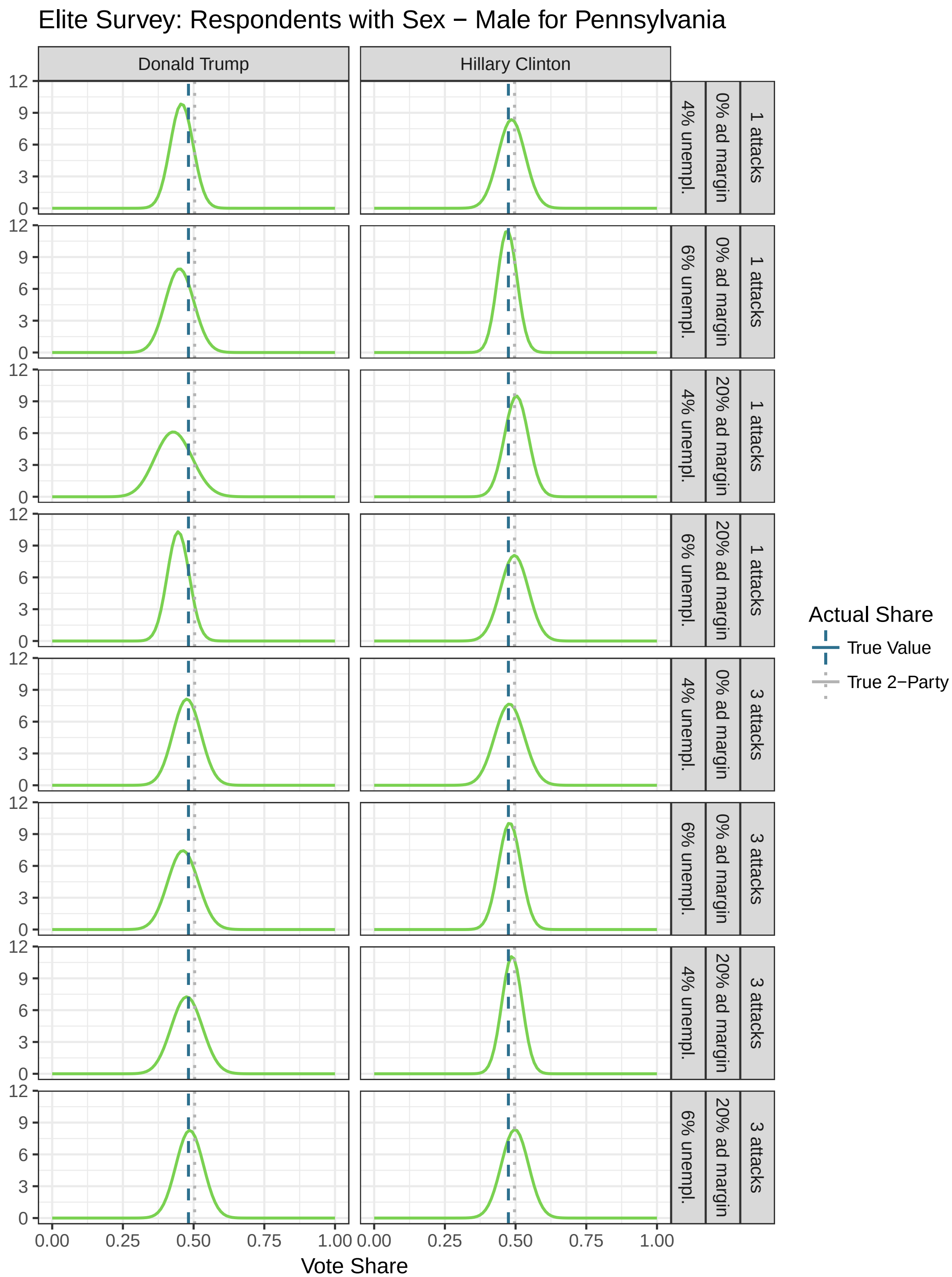


Figure 3.229: Priors with covariates: Elite Pennsylvania Sex Male

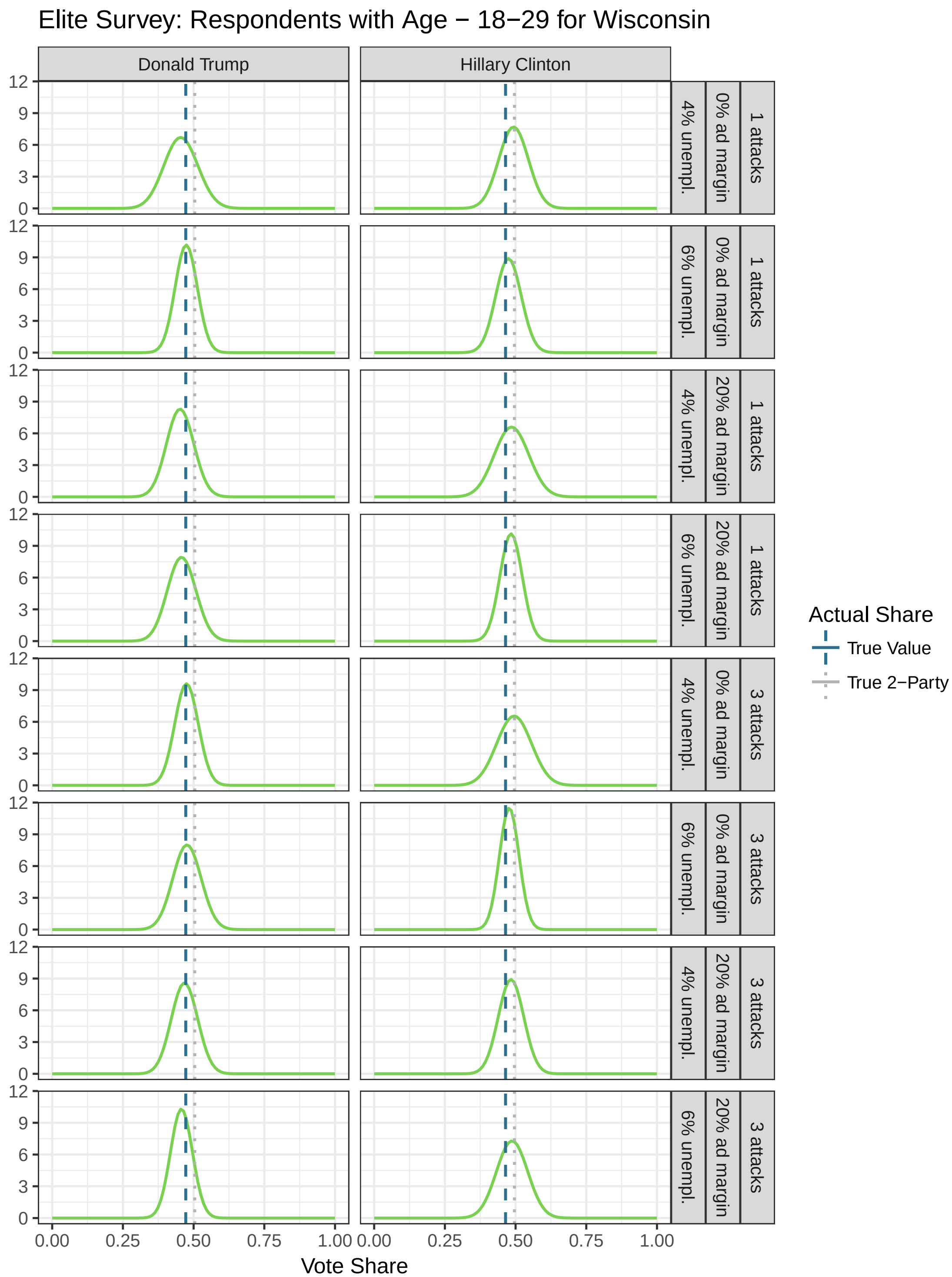


Figure 3.230: Priors with covariates: Elite Wisconsin Age 18-29

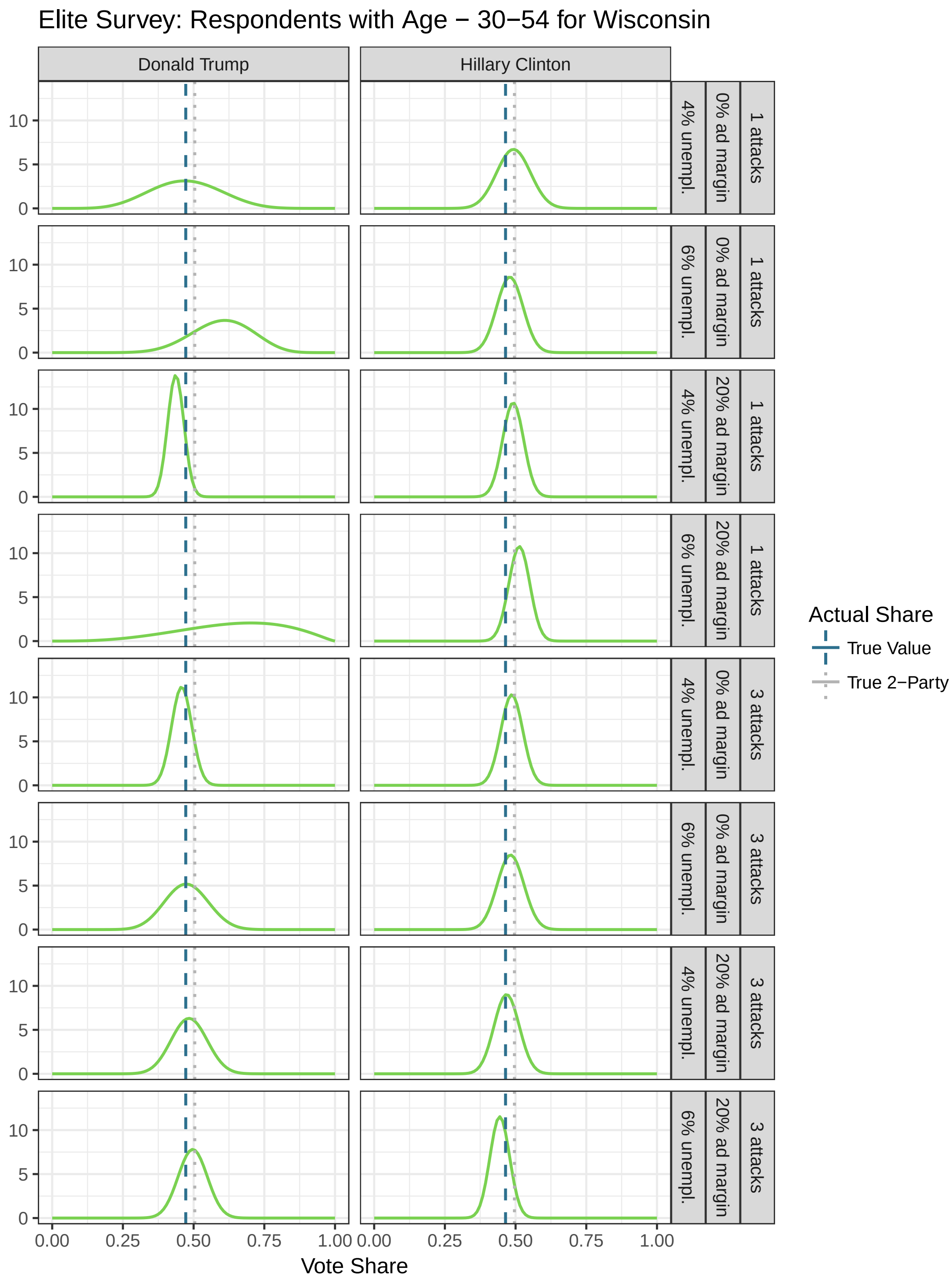


Figure 3.231: Priors with covariates: Elite Wisconsin Age 30-54

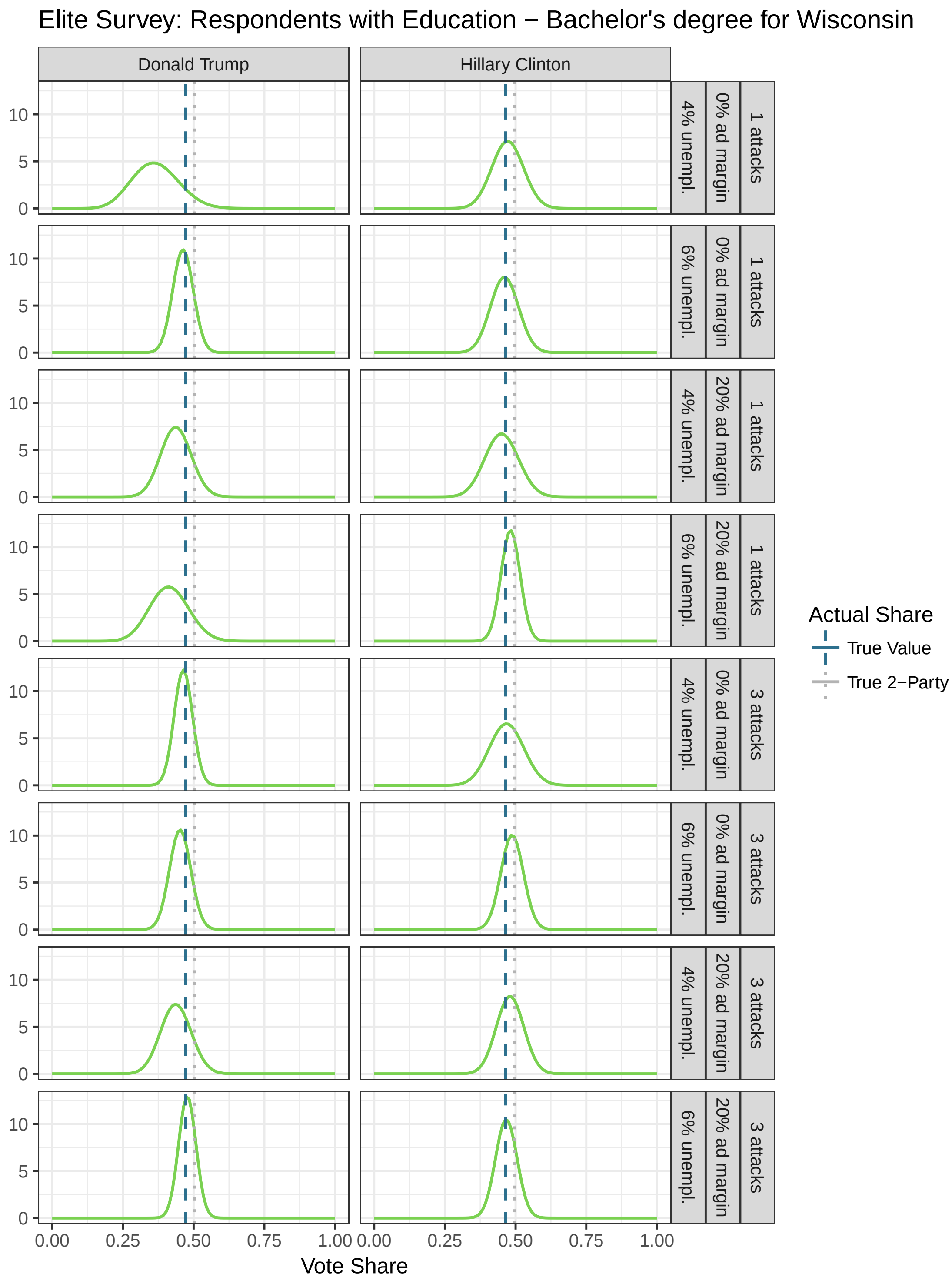


Figure 3.232: Priors with covariates: Elite Wisconsin Education Bachelor's degree

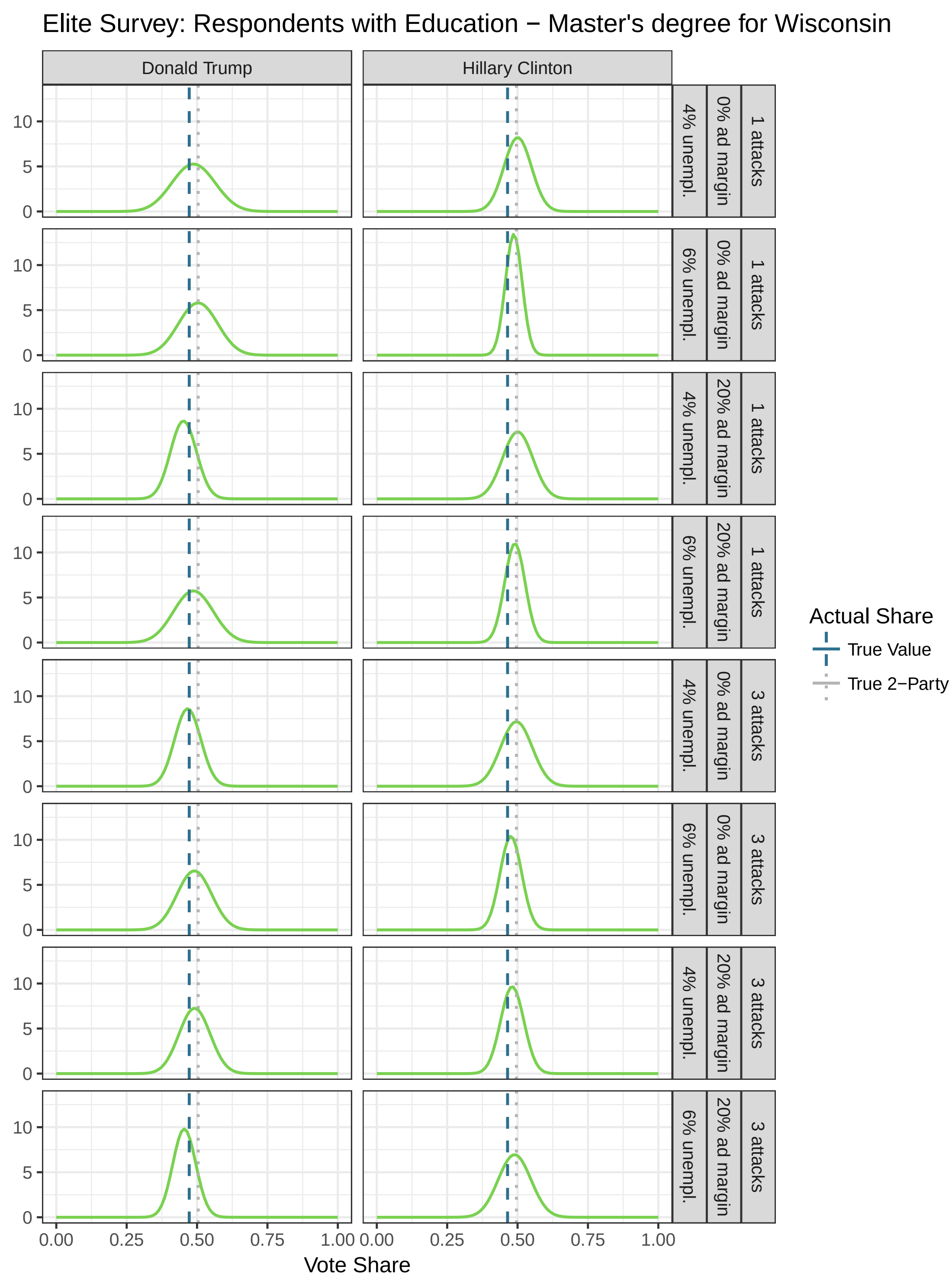


Figure 3.233: Priors with covariates: Elite Wisconsin Education Master's degree

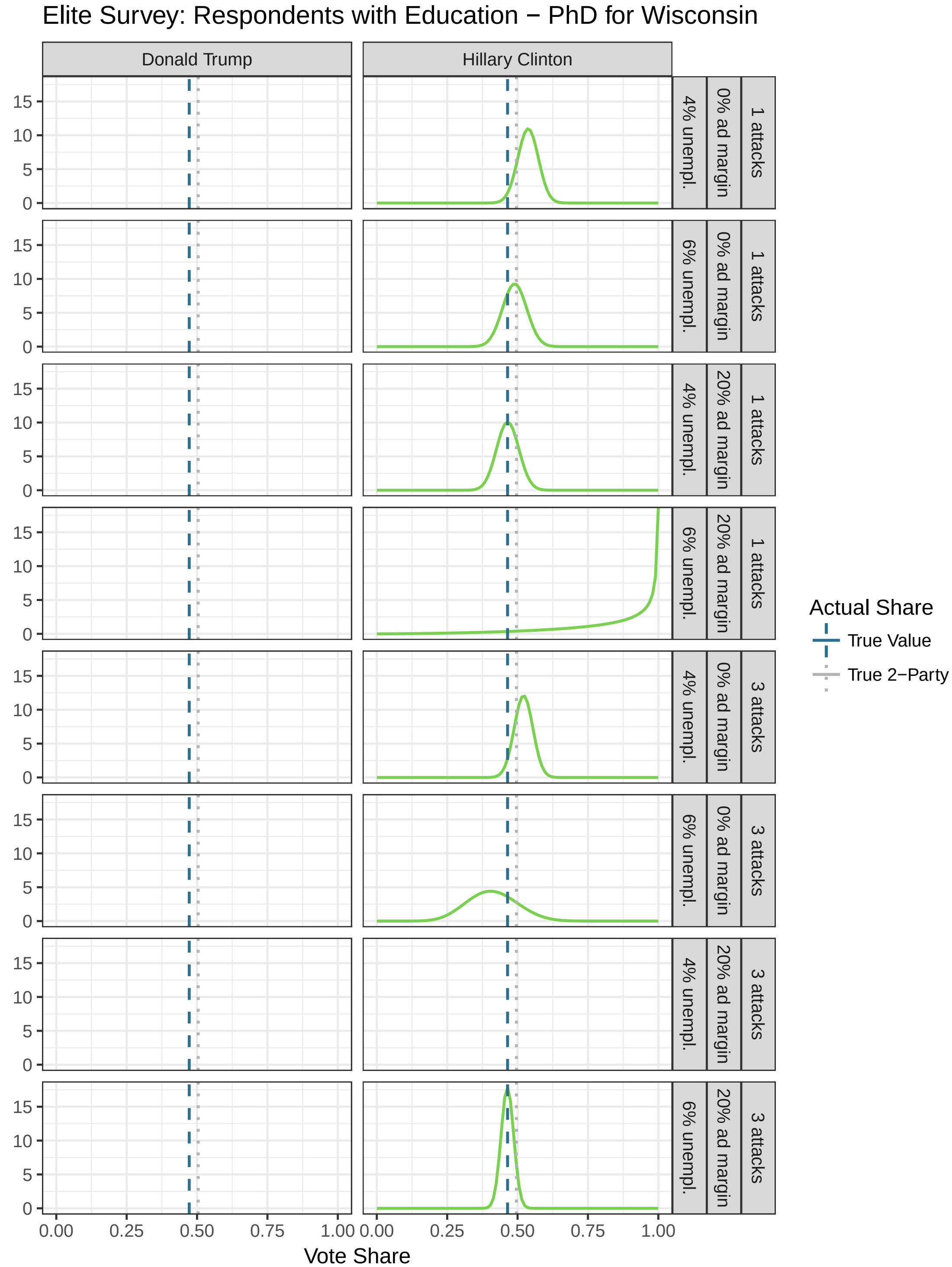


Figure 3.234: Priors with covariates: Elite Wisconsin Education PhD

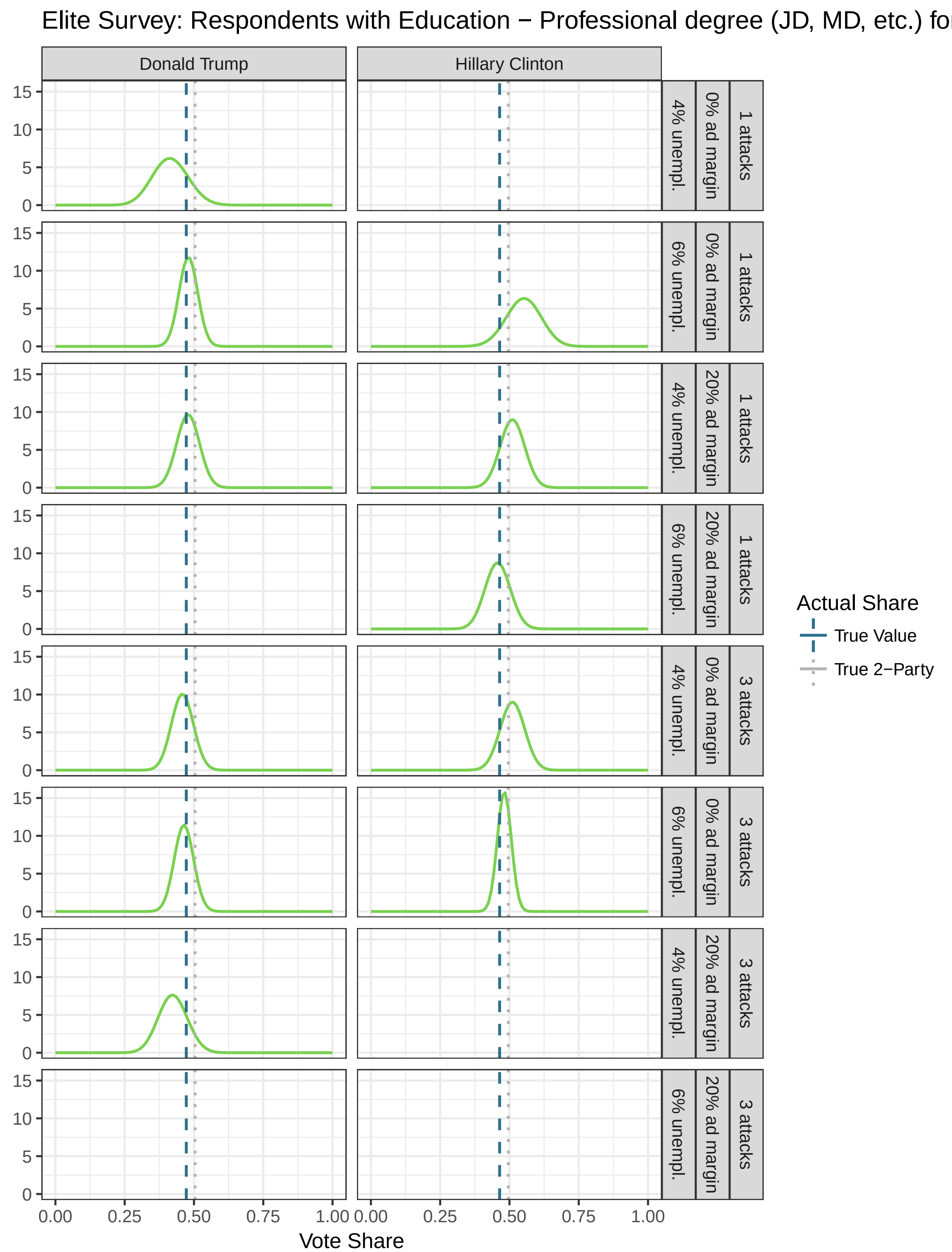


Figure 3.235: Priors with covariates: Elite Wisconsin Education Professional degree JD MD etc

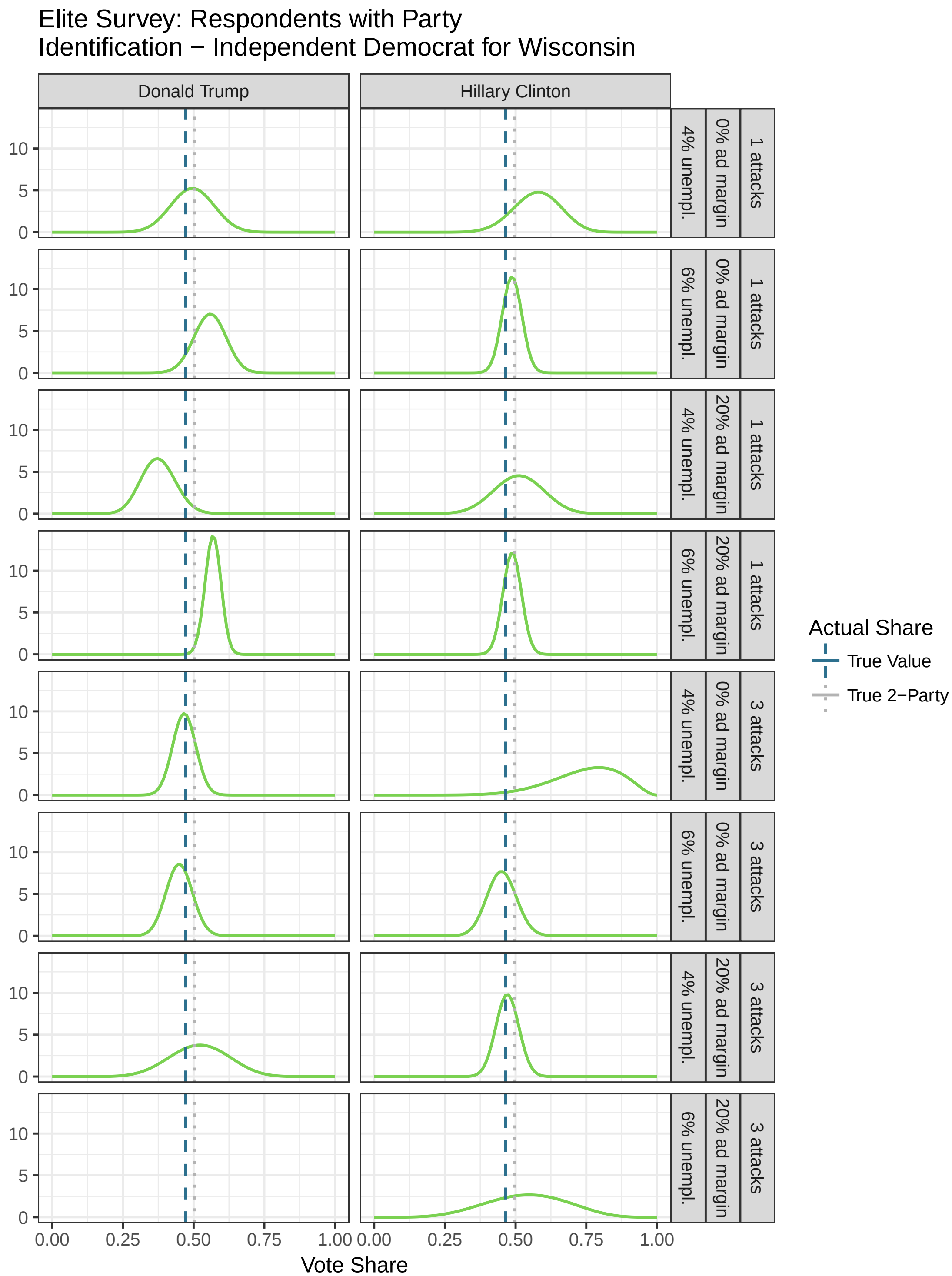


Figure 3.236: Priors with covariates: Elite Wisconsin Party Identification Independent Democrat

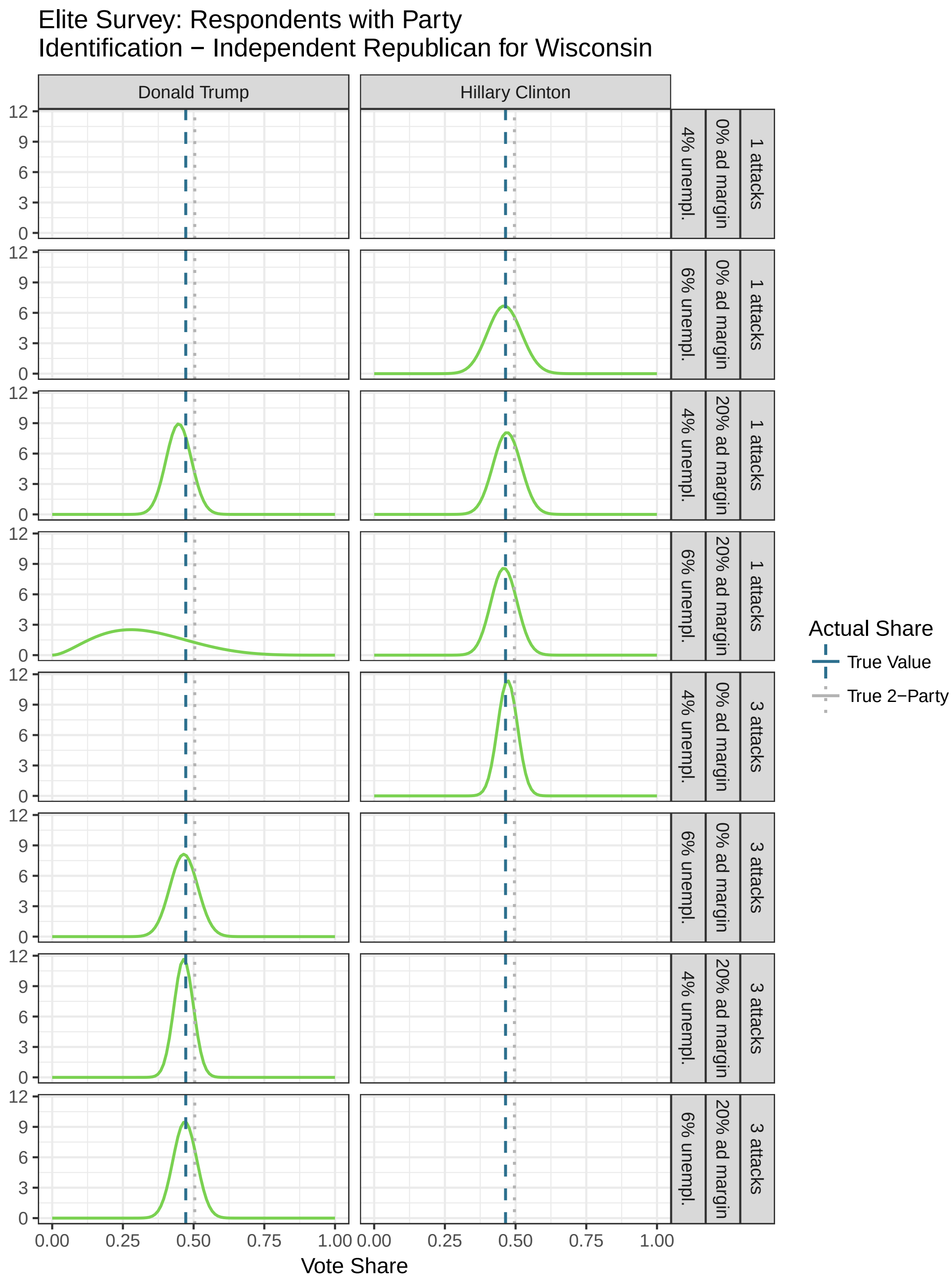


Figure 3.237: Priors with covariates: Elite Wisconsin Party Identification Independent Republican

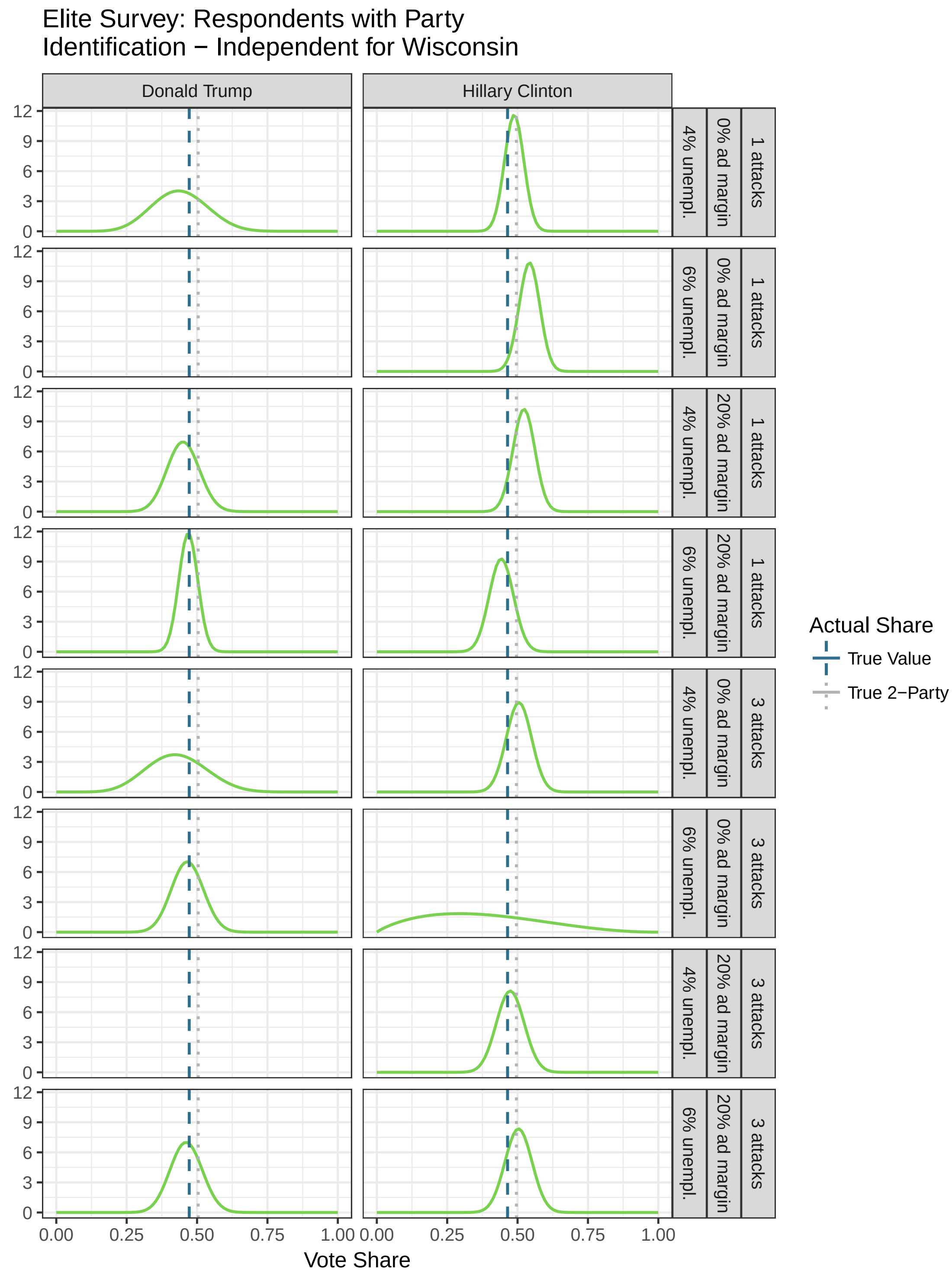


Figure 3.238: Priors with covariates: Elite Wisconsin Party Identification Independent

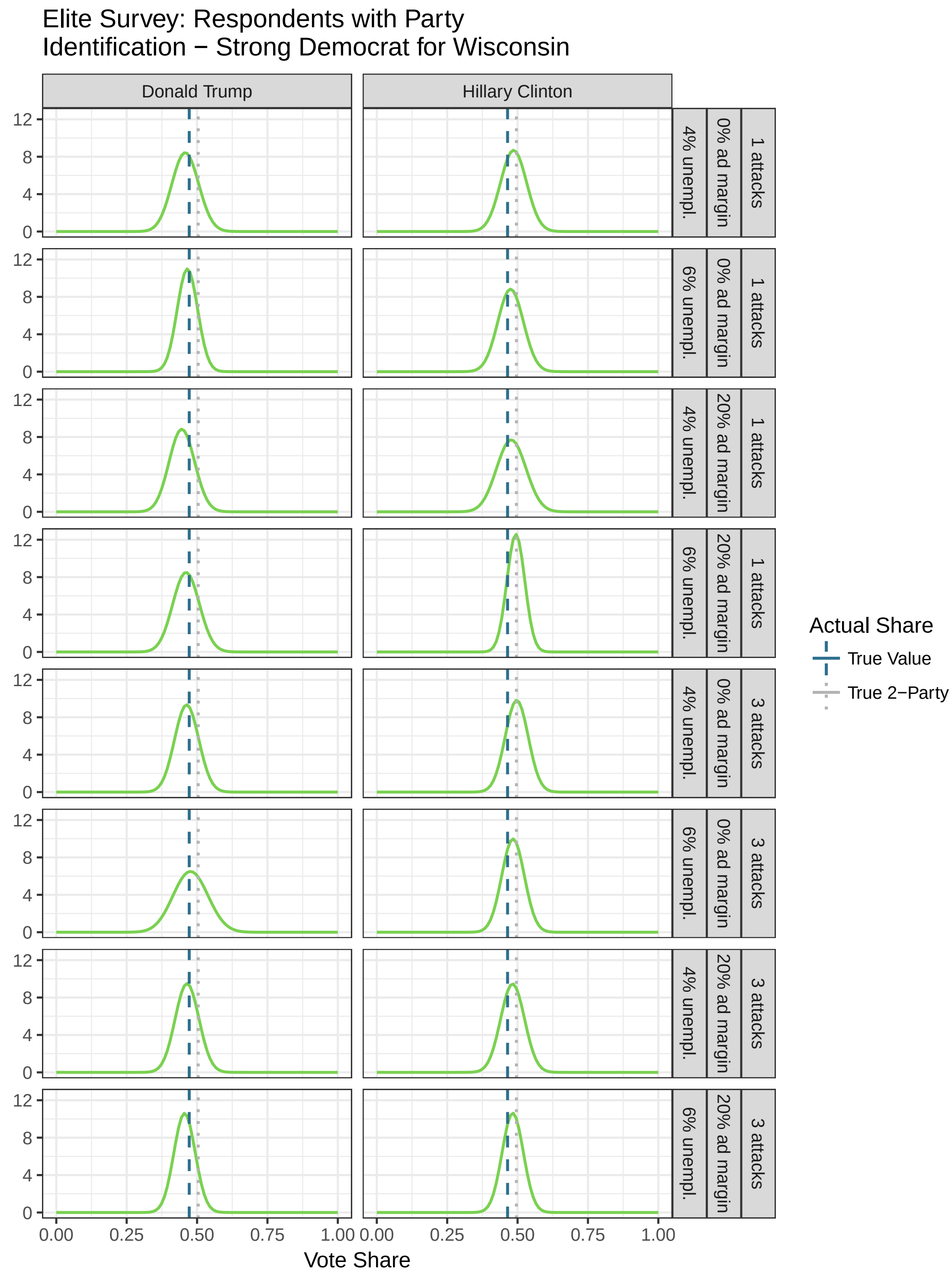


Figure 3.239: Priors with covariates: Elite Wisconsin Party Identification Strong Democrat

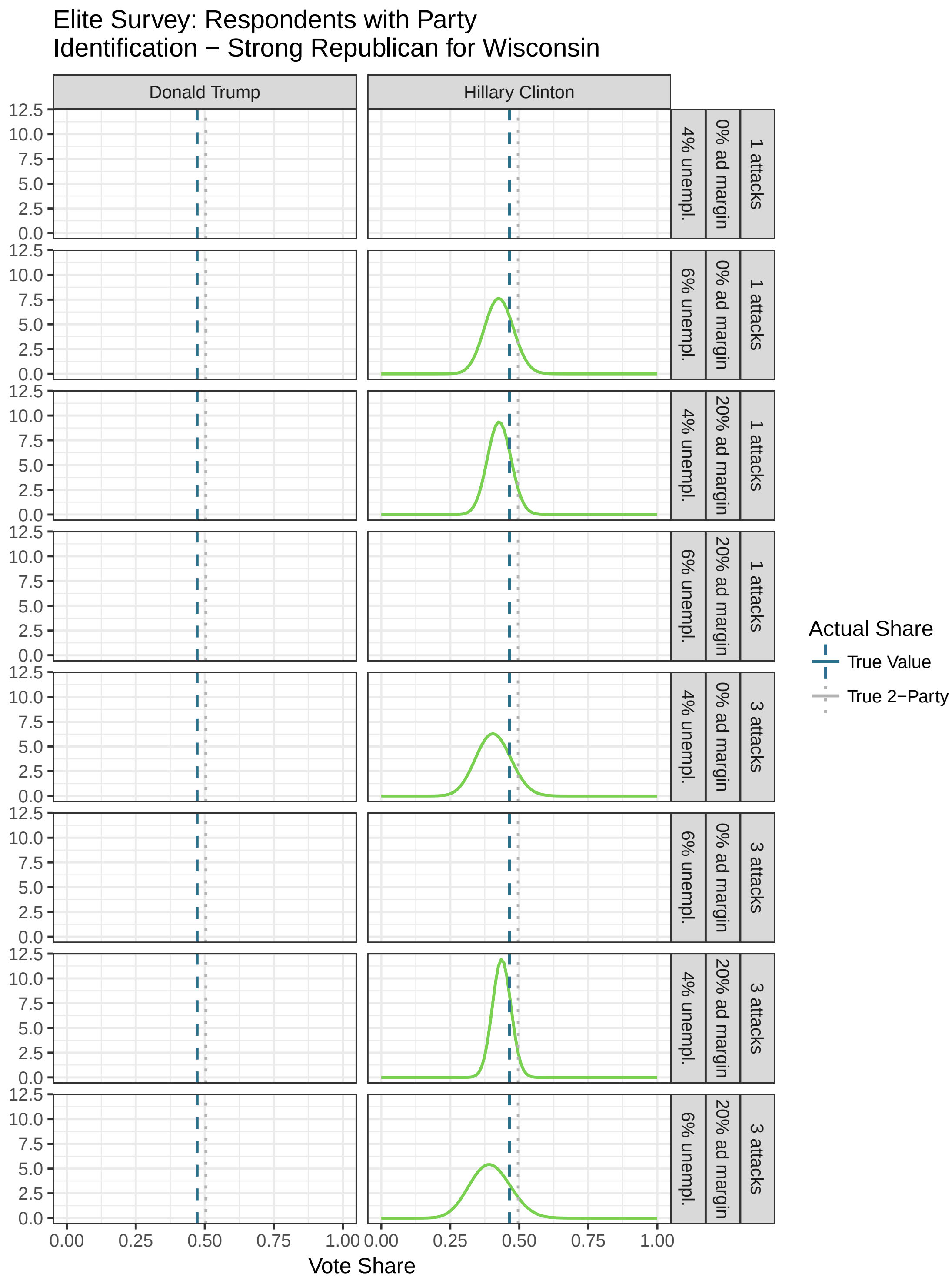


Figure 3.240: Priors with covariates: Elite Wisconsin Party Identification Strong Republican

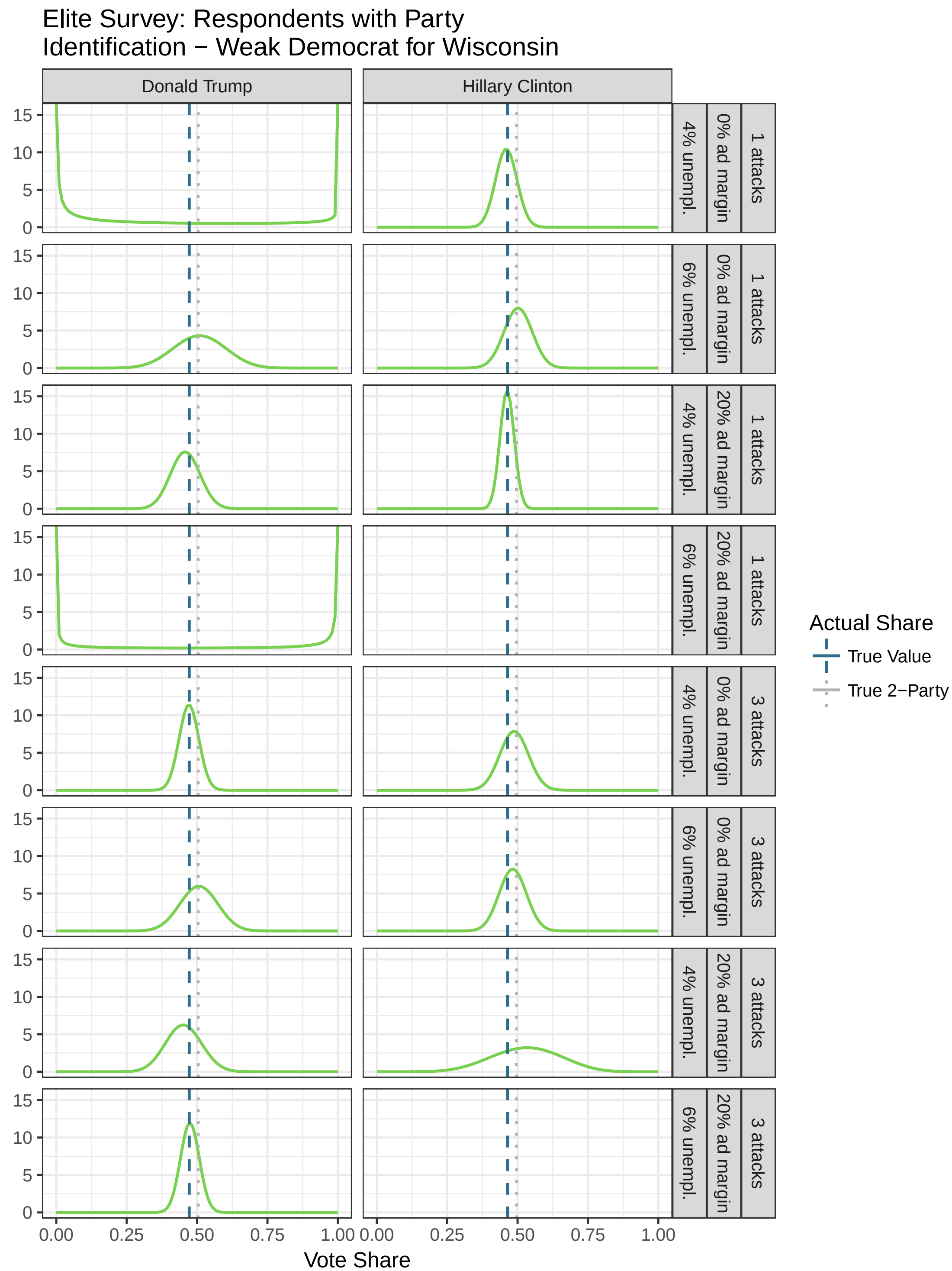


Figure 3.241: Priors with covariates: Elite Wisconsin Party Identification Weak Democrat

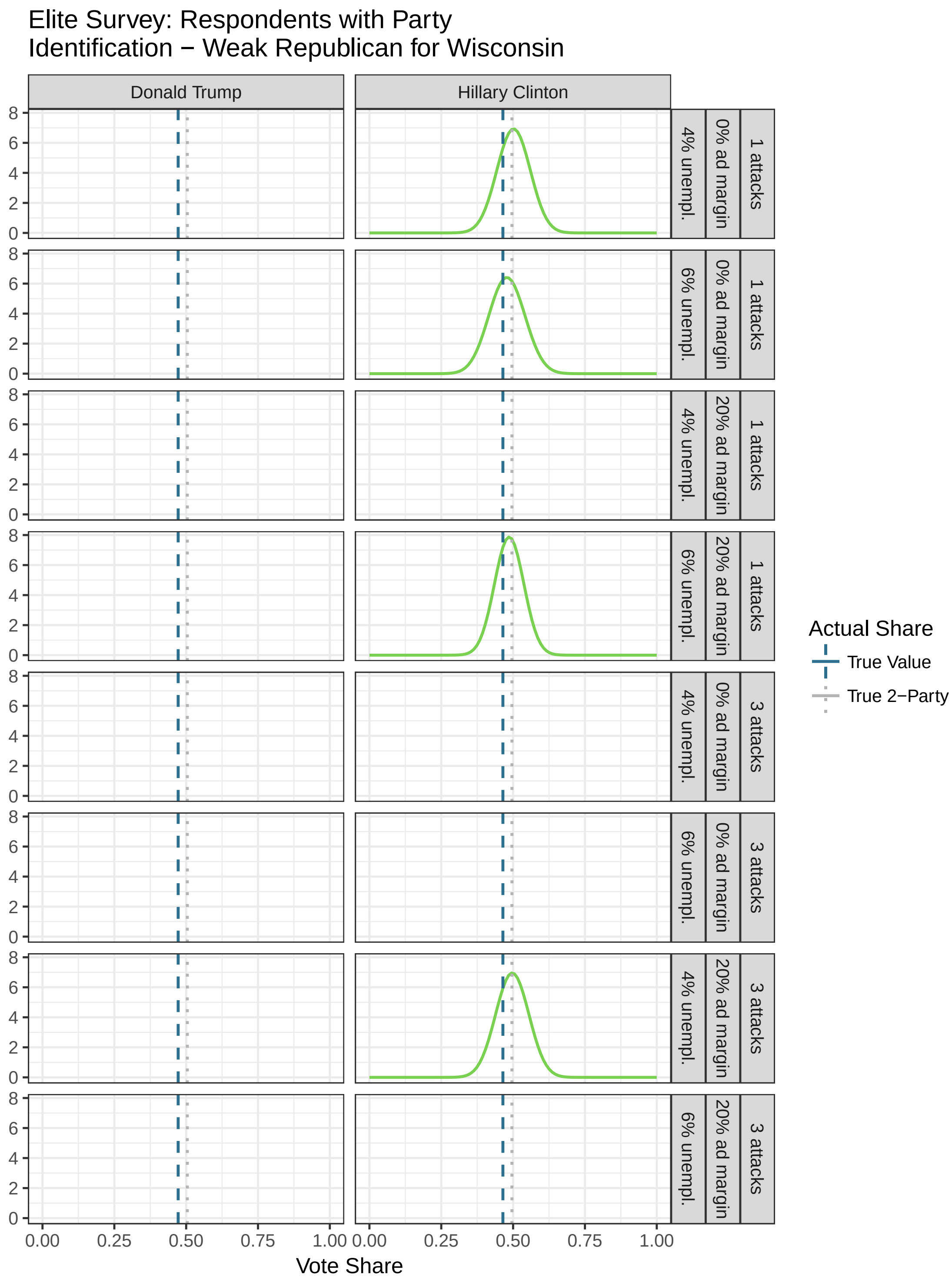


Figure 3.242: Priors with covariates: Elite Wisconsin Party Identification Weak Republican

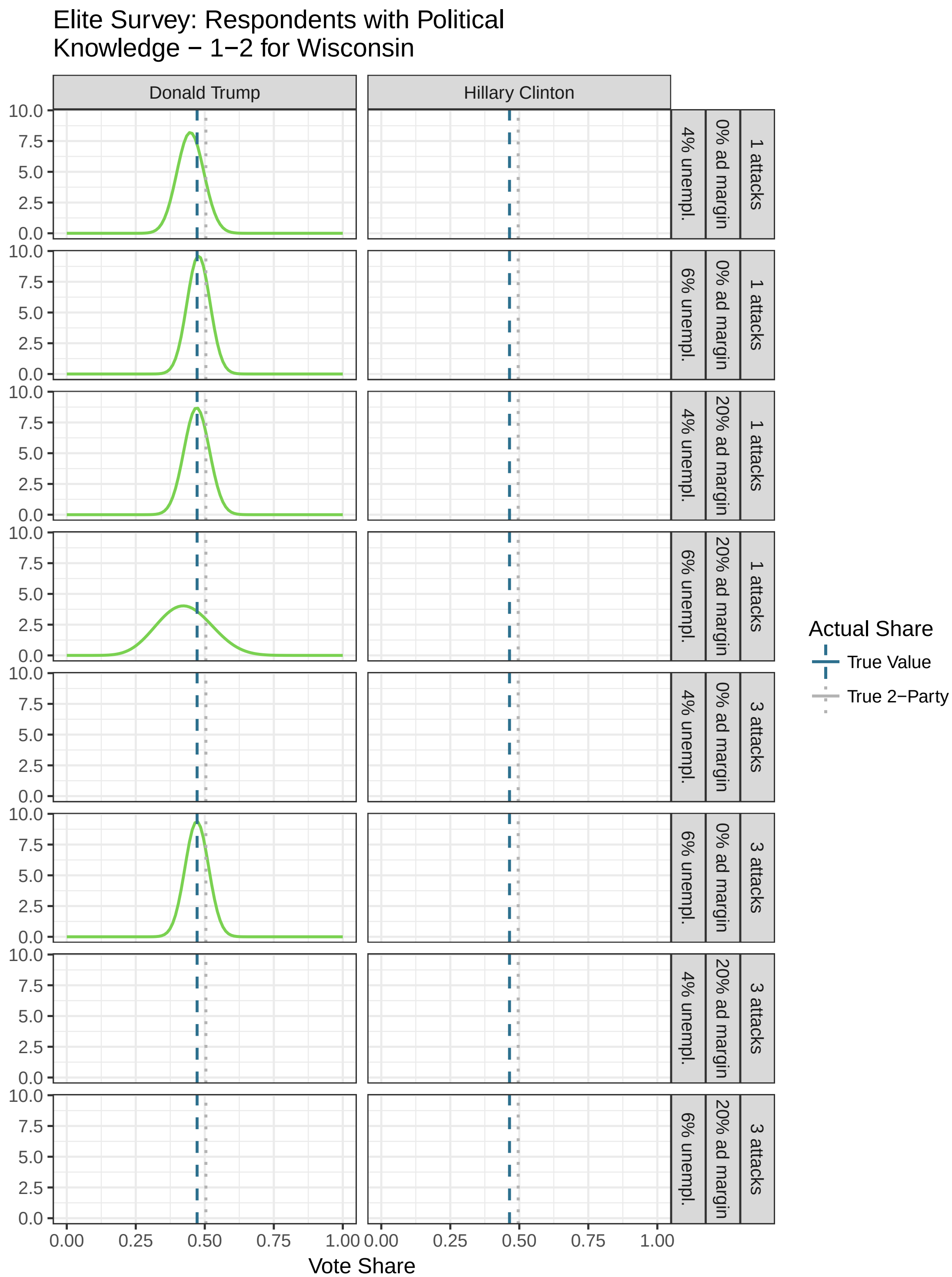


Figure 3.243: Priors with covariates: Elite Wisconsin Political Knowledge 1-2

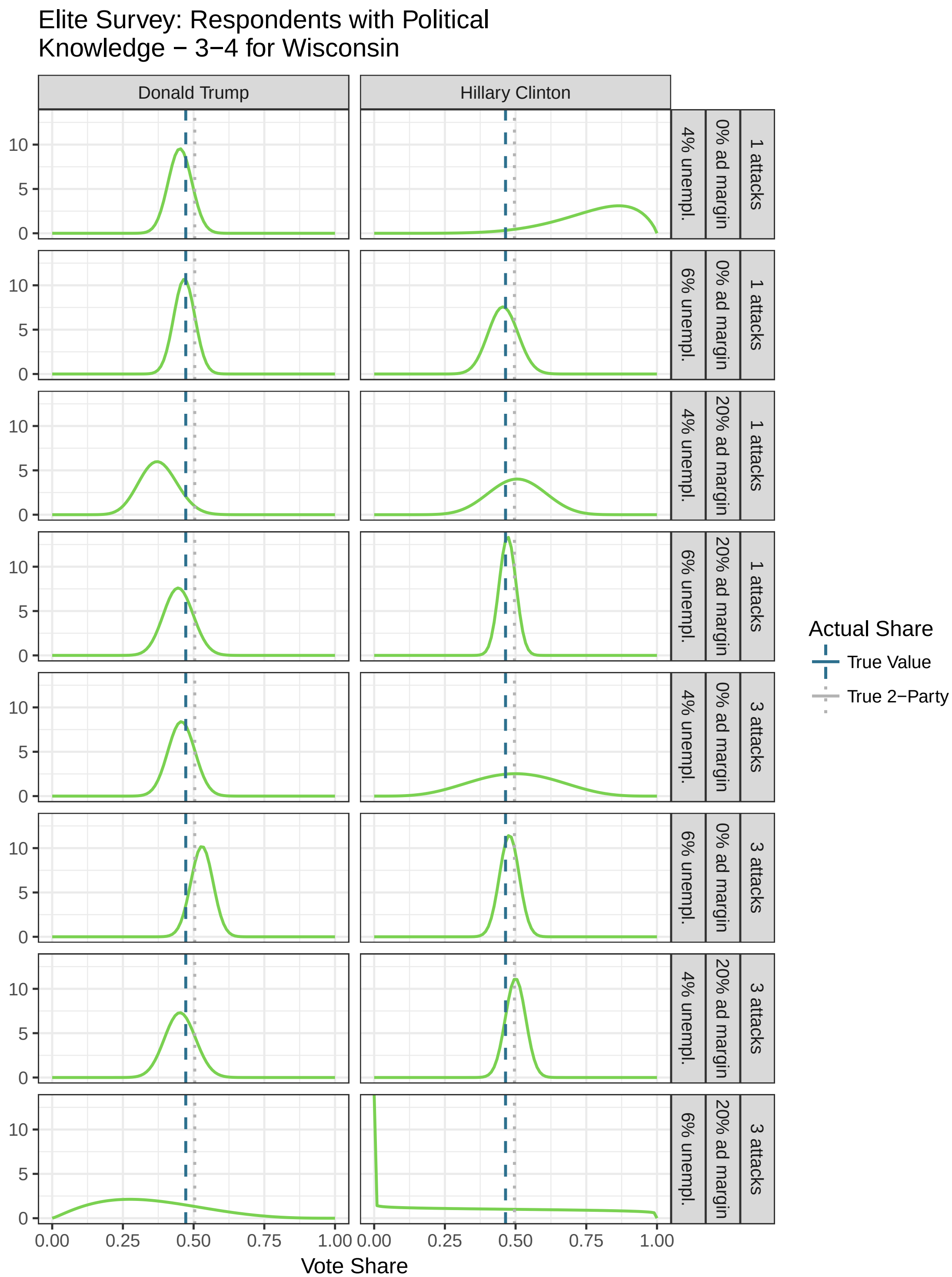


Figure 3.244: Priors with covariates: Elite Wisconsin Political Knowledge 3-4

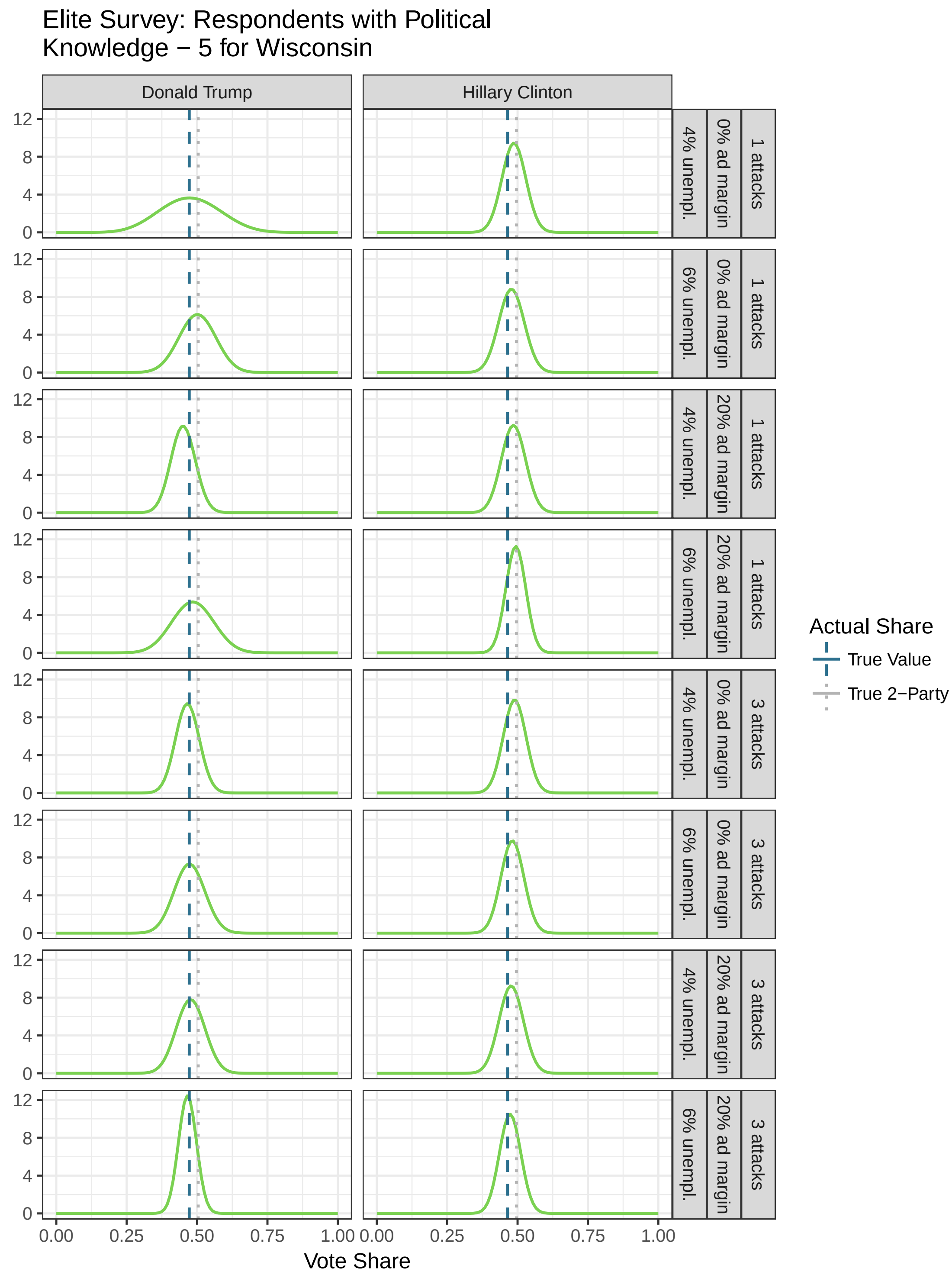


Figure 3.245: Priors with covariates: Elite Wisconsin Political Knowledge 5

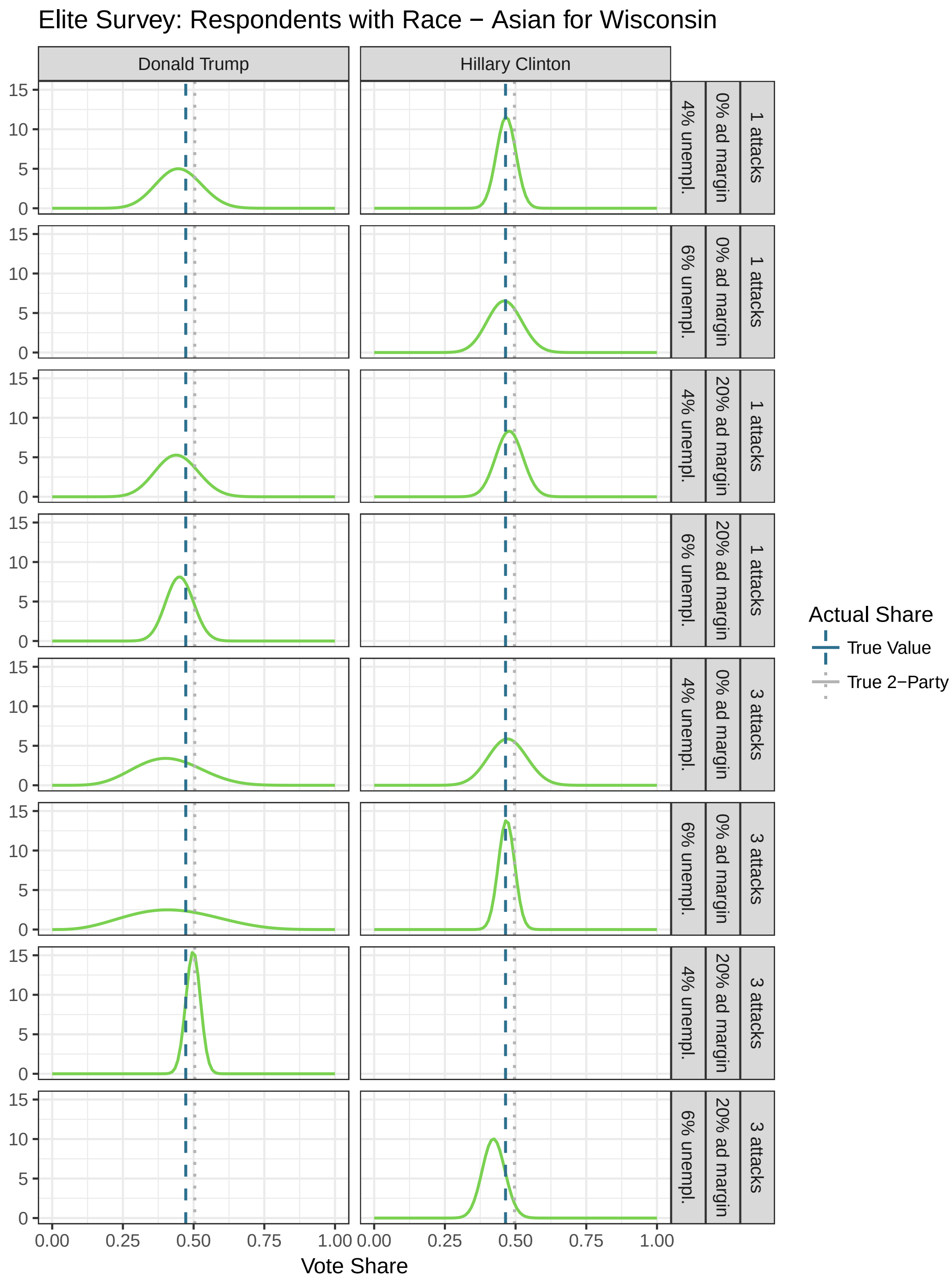


Figure 3.246: Priors with covariates: Elite Wisconsin Race Asian

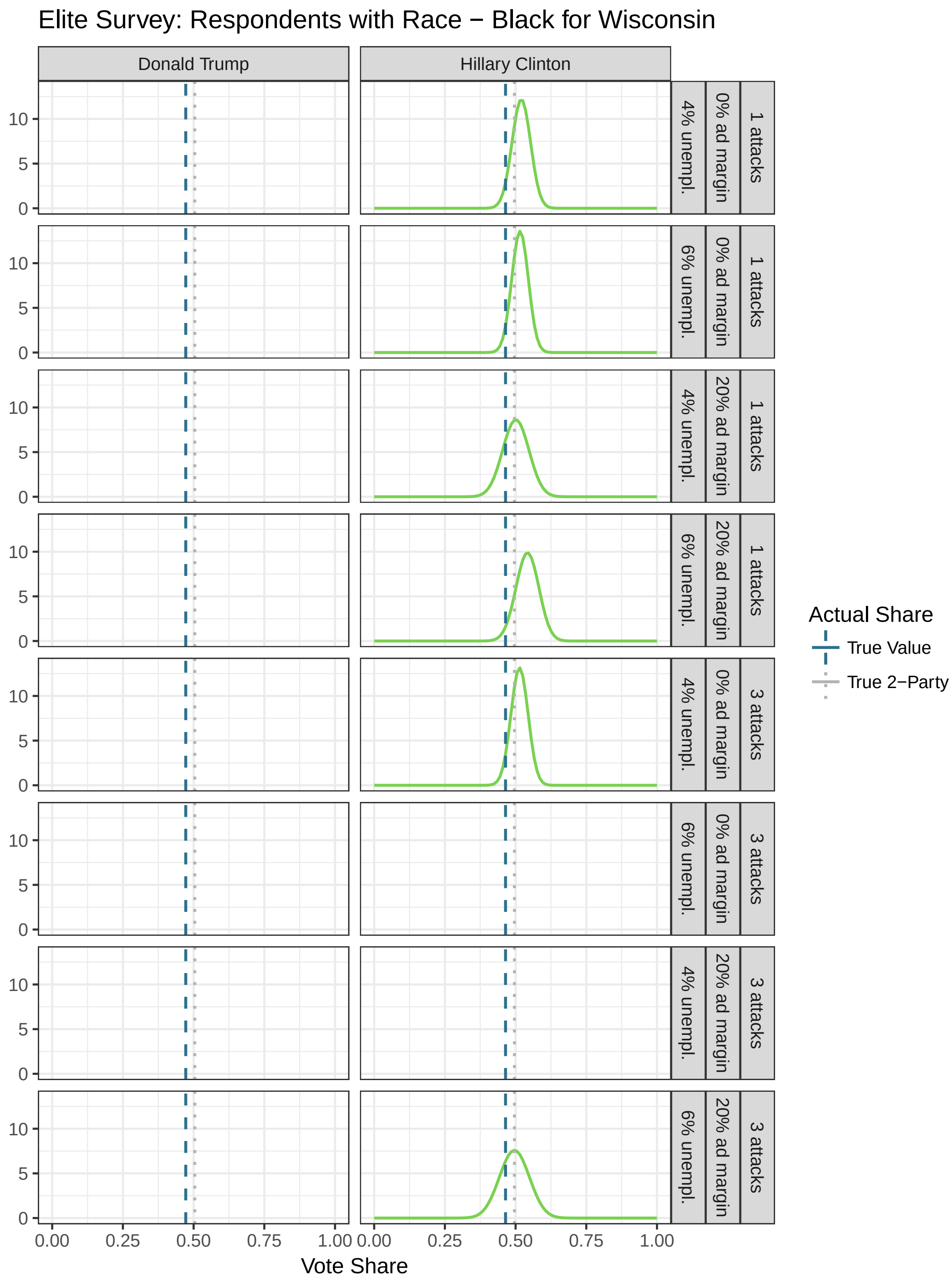


Figure 3.247: Priors with covariates: Elite Wisconsin Race Black

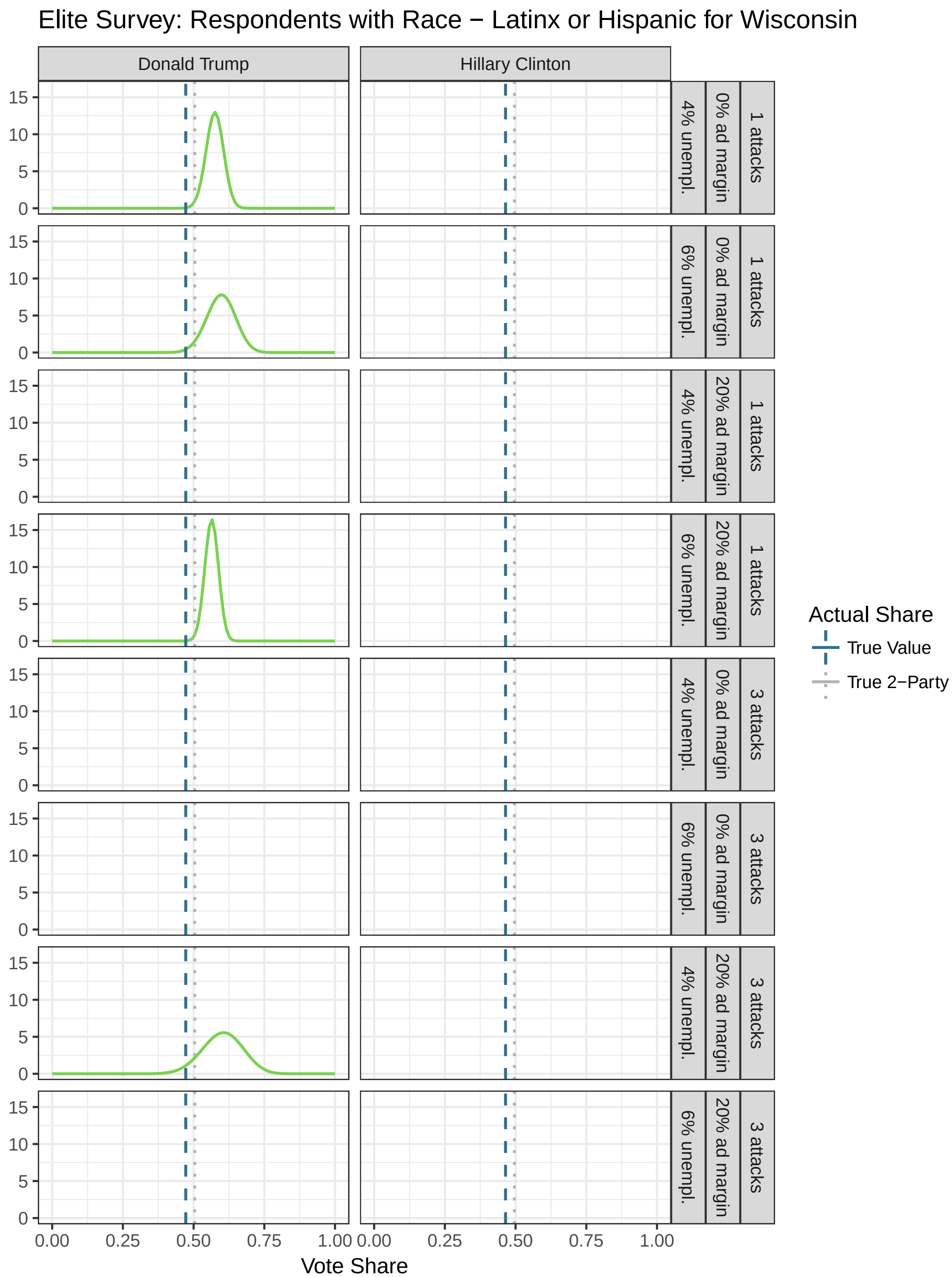


Figure 3.248: Priors with covariates: Elite Wisconsin Race Latinx or Hispanic

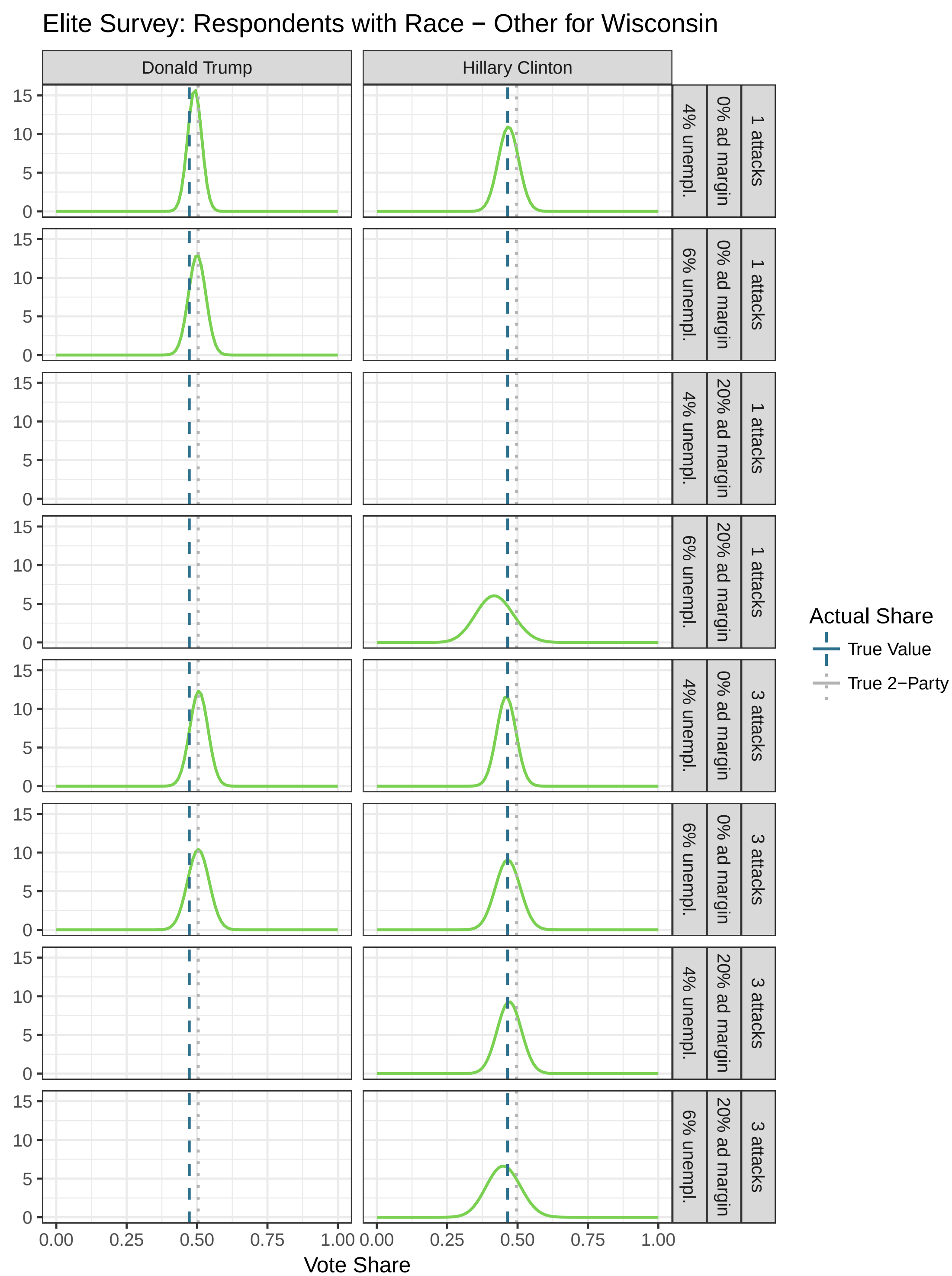


Figure 3.249: Priors with covariates: Elite Wisconsin Race Other

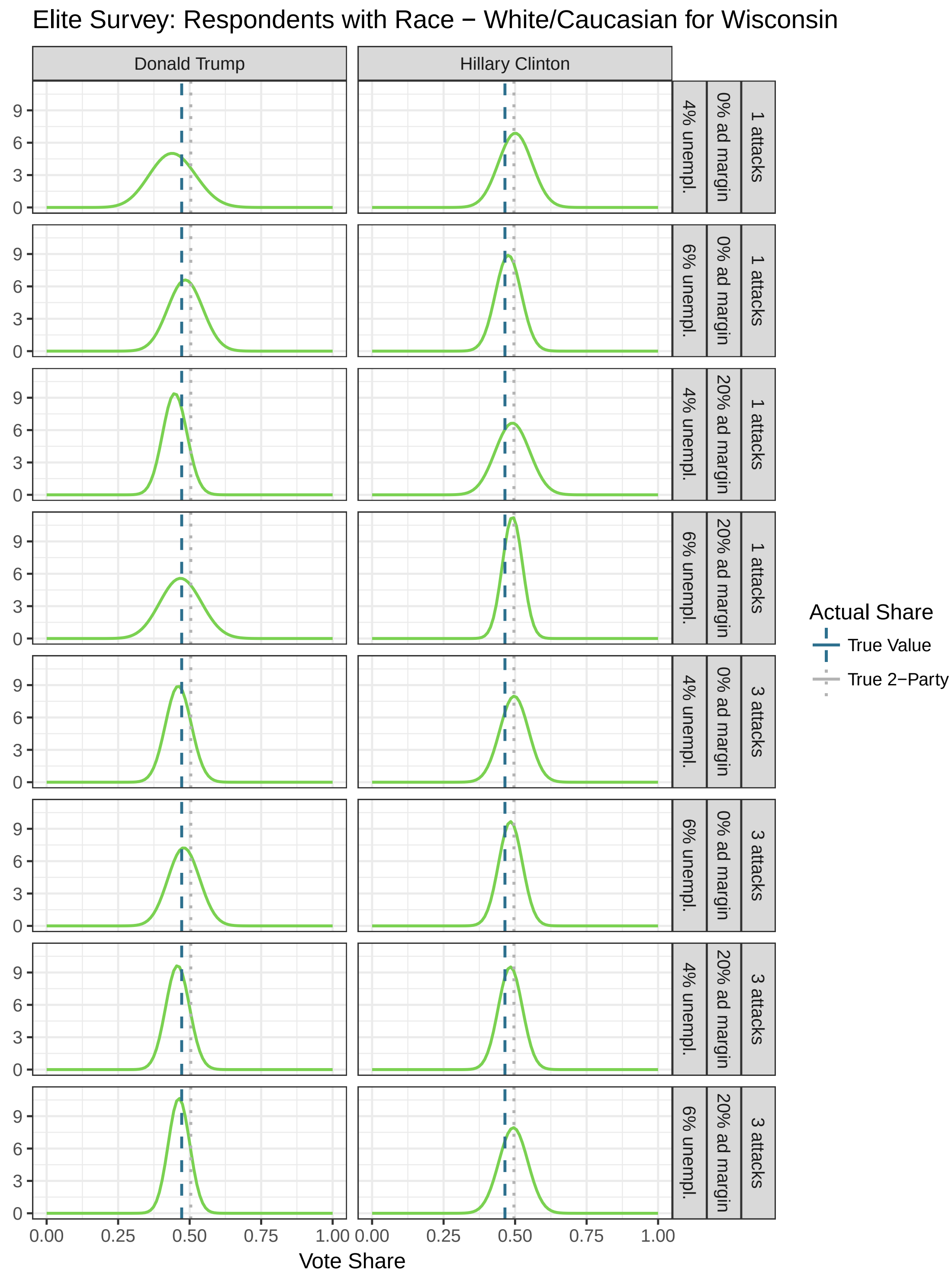


Figure 3.250: Priors with covariates: Elite Wisconsin Race White Caucasian

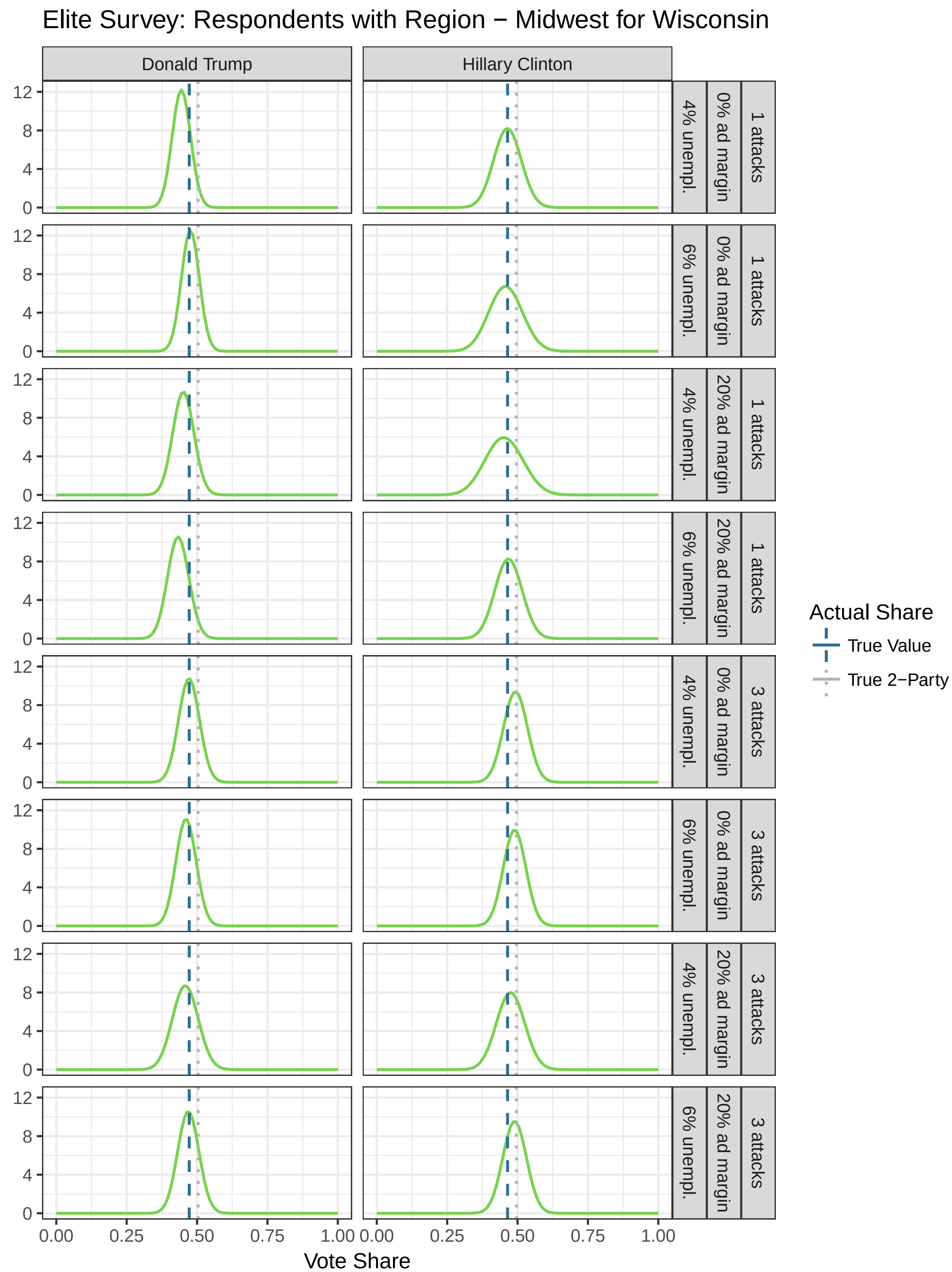


Figure 3.251: Priors with covariates: Elite Wisconsin Region Midwest

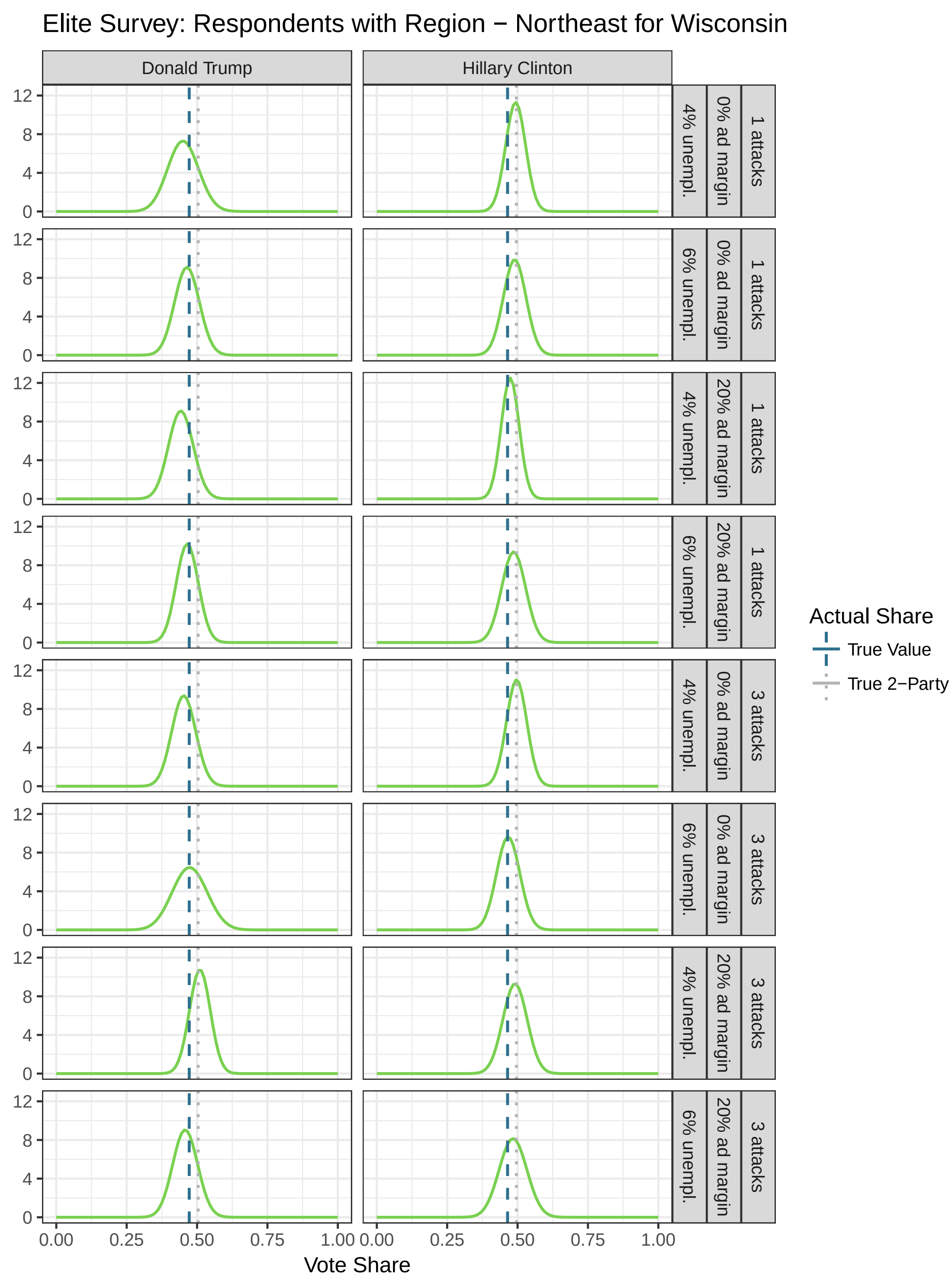


Figure 3.252: Priors with covariates: Elite Wisconsin Region Northeast

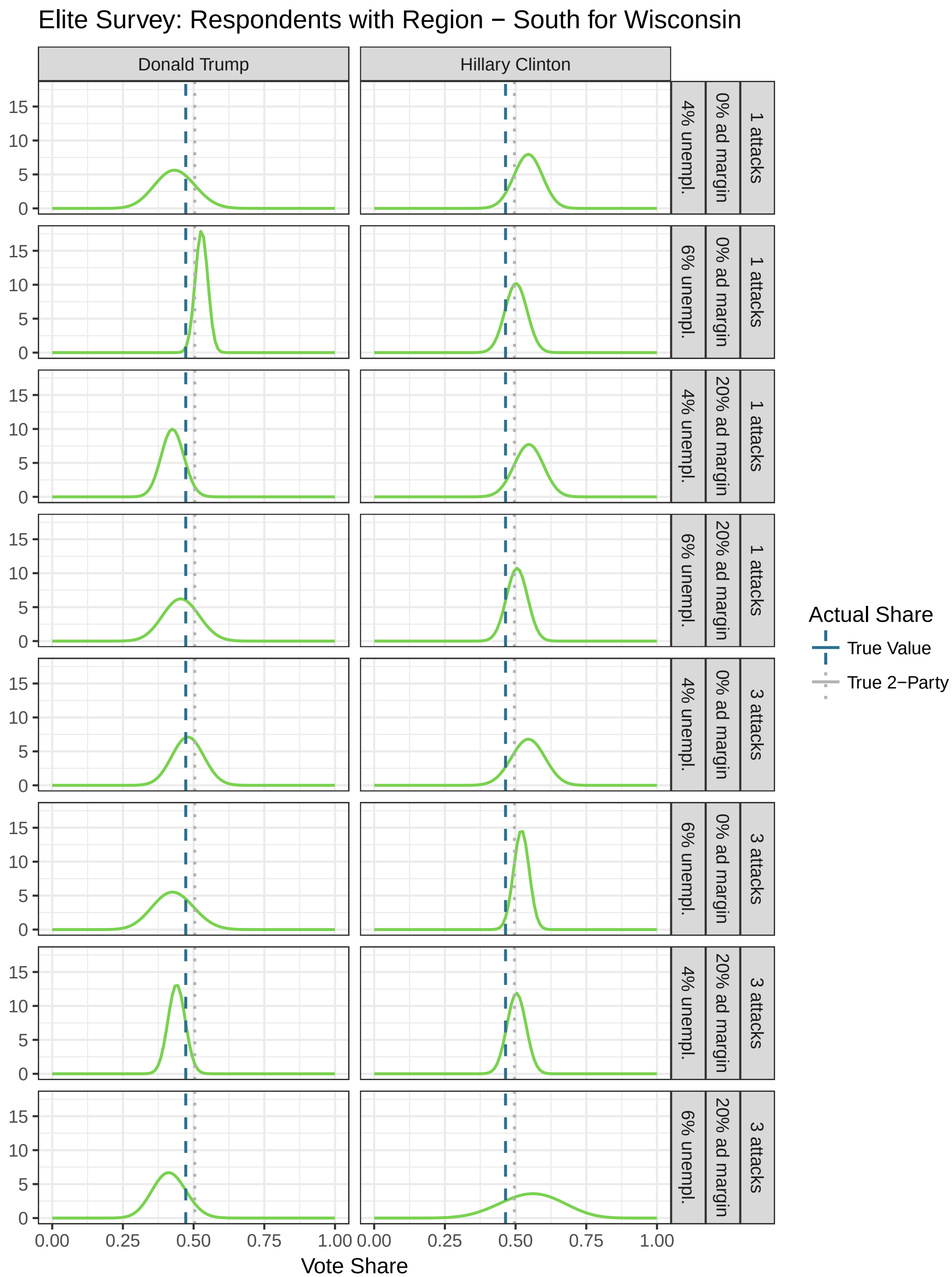


Figure 3.253: Priors with covariates: Elite Wisconsin Region South

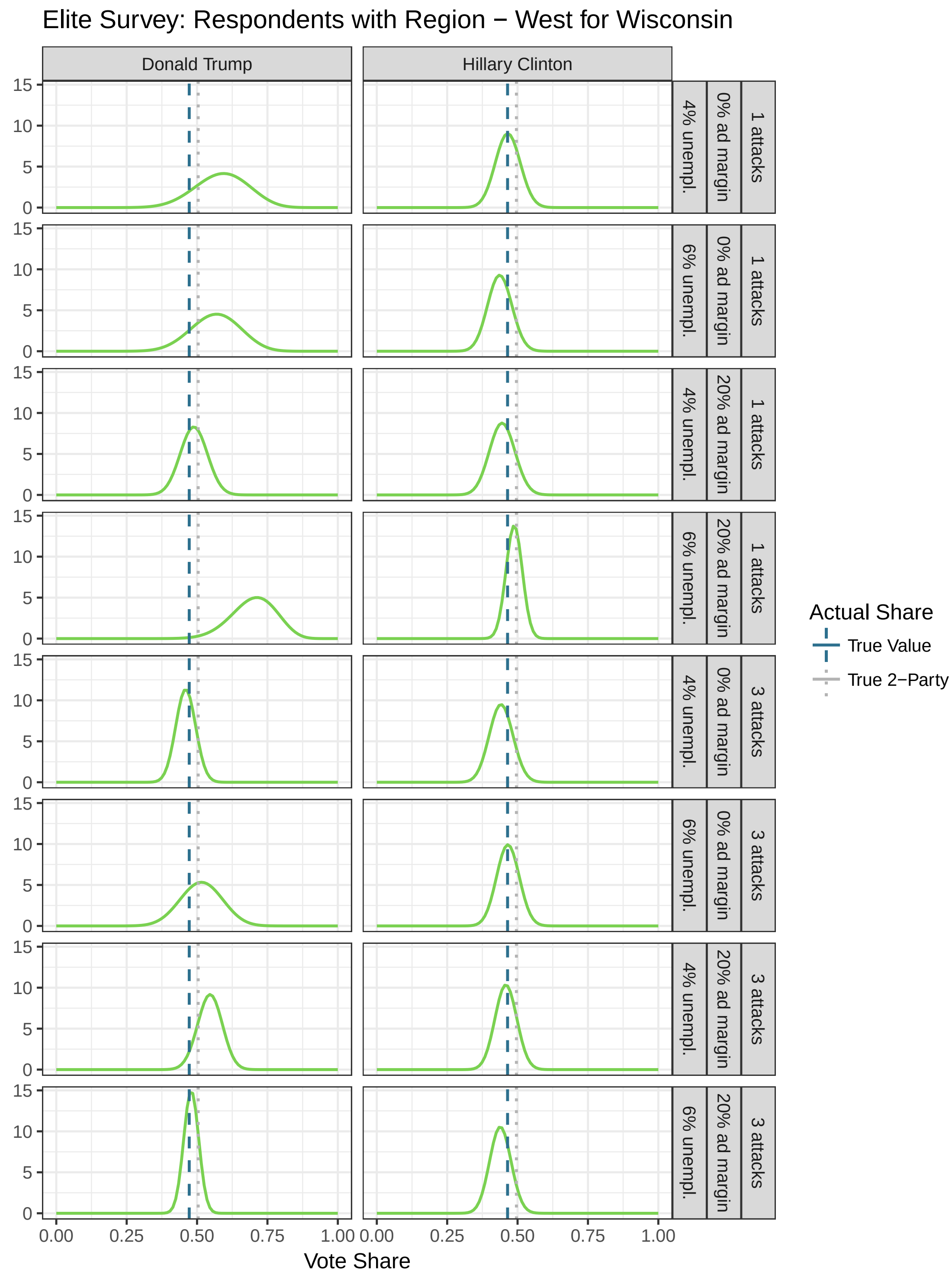


Figure 3.254: Priors with covariates: Elite Wisconsin Region West

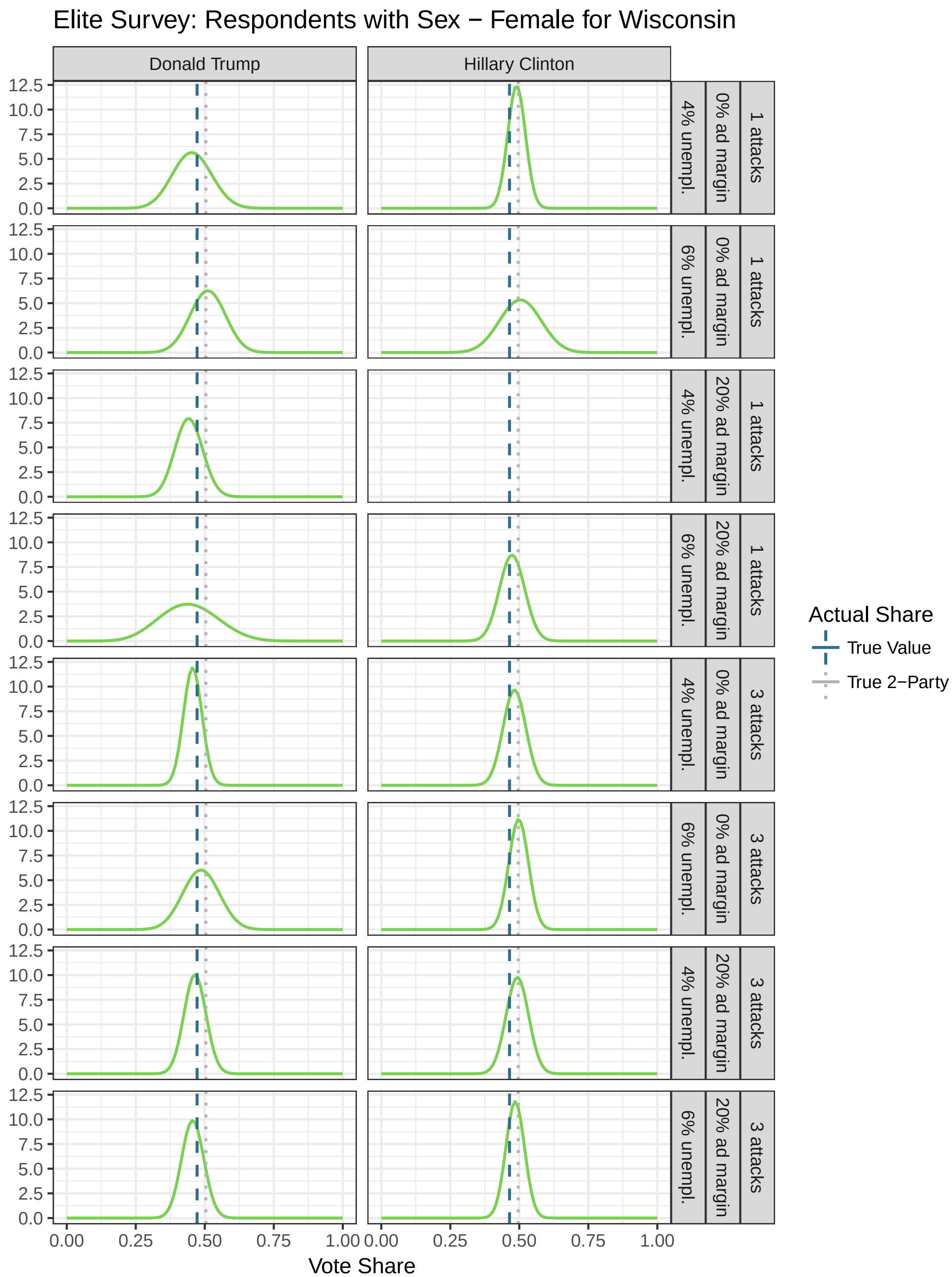


Figure 3.255: Priors with covariates: Elite Wisconsin Sex Female

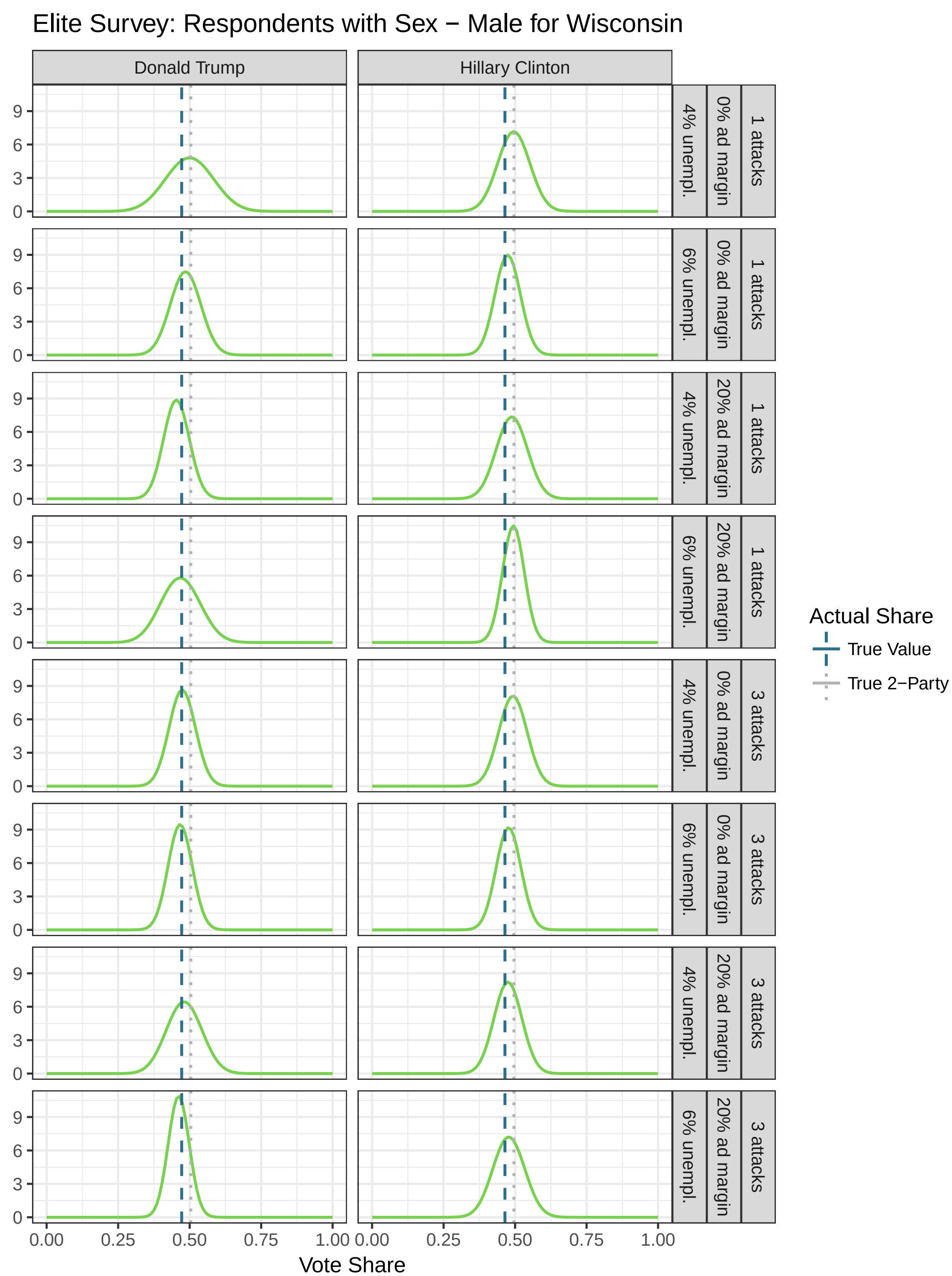


Figure 3.256: Priors with covariates: Elite Wisconsin Sex Male

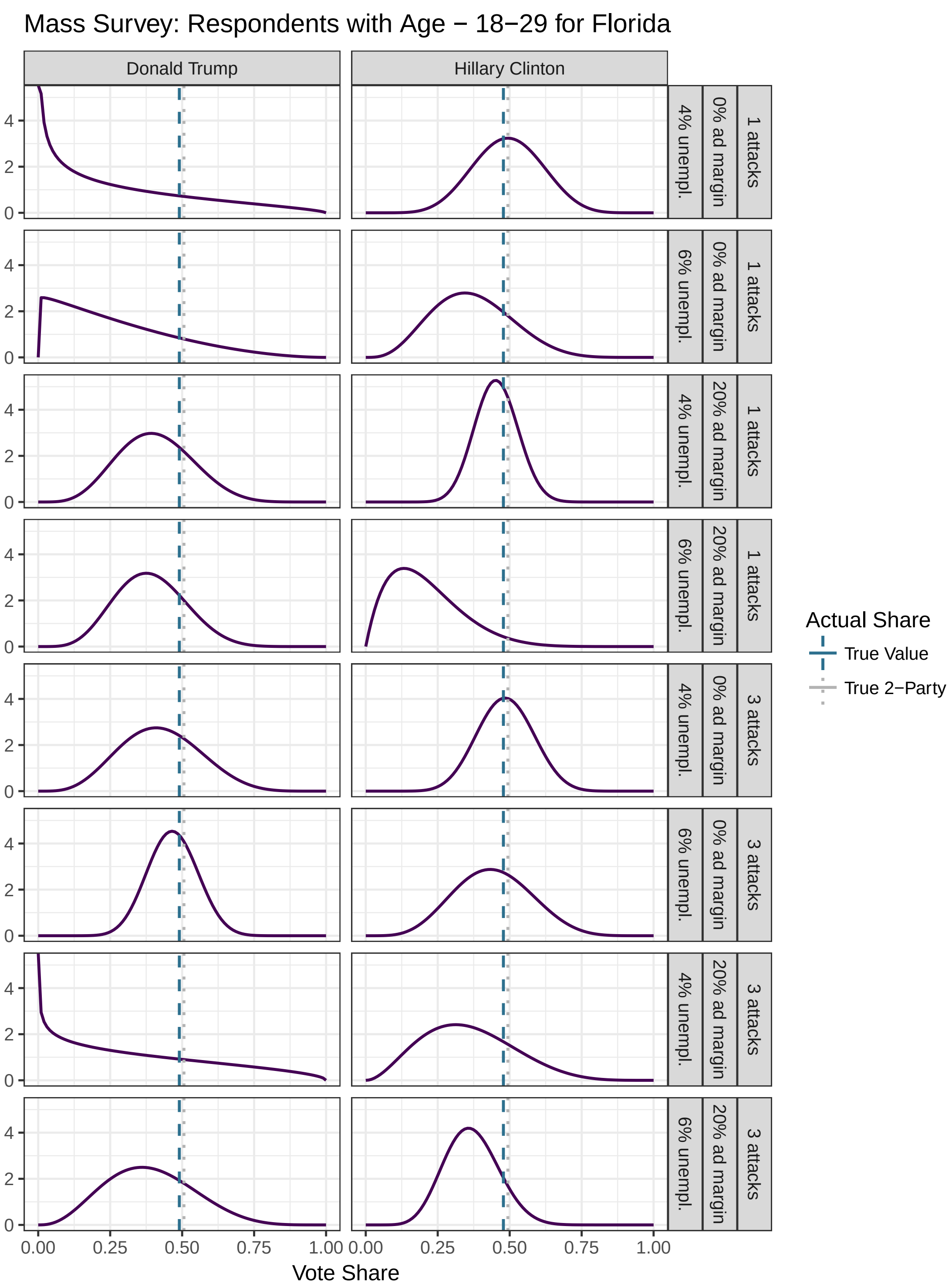


Figure 3.257: Priors with covariates: Mass Florida Age 18-29

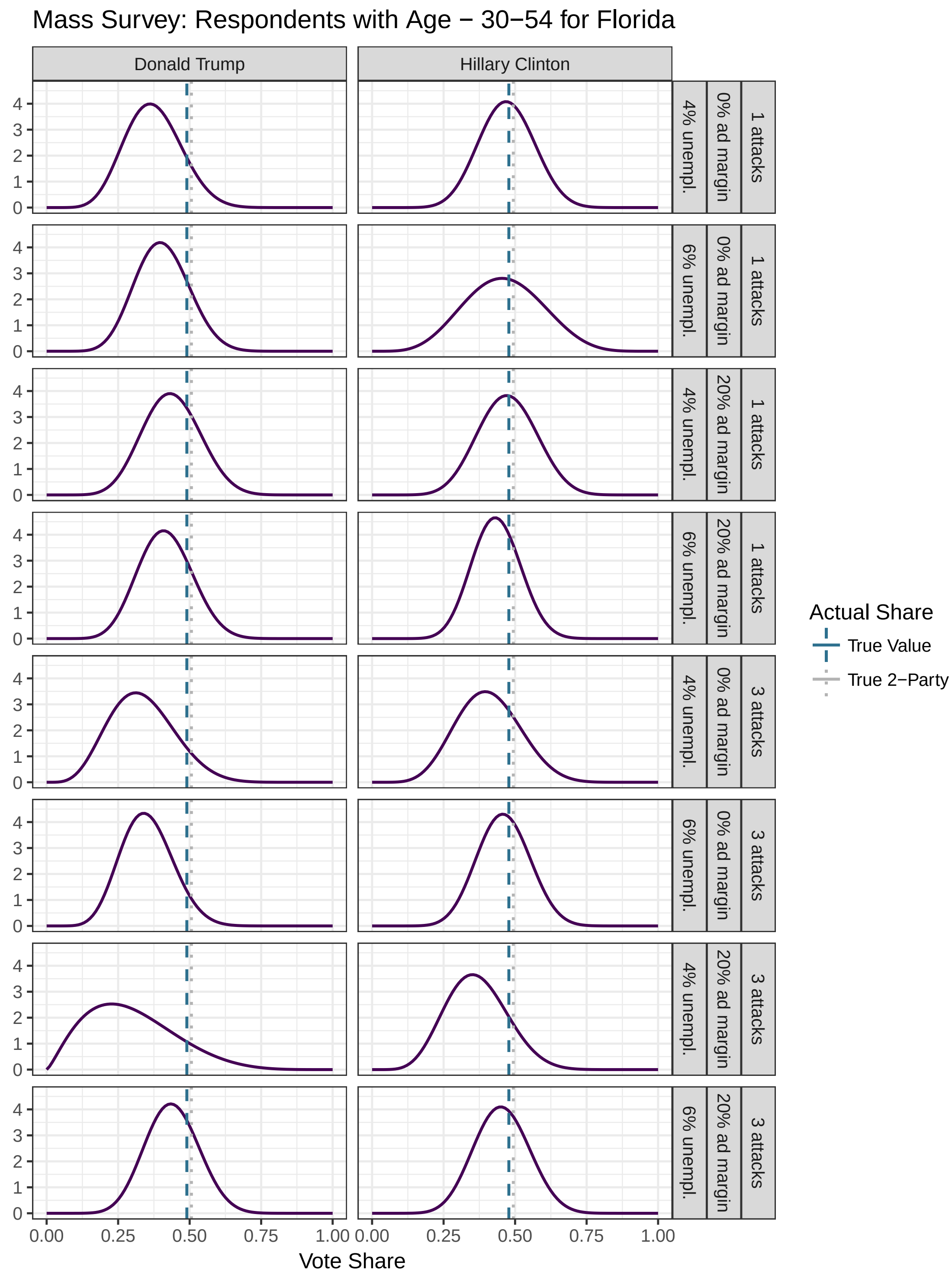


Figure 3.258: Priors with covariates: Mass Florida Age 30-54

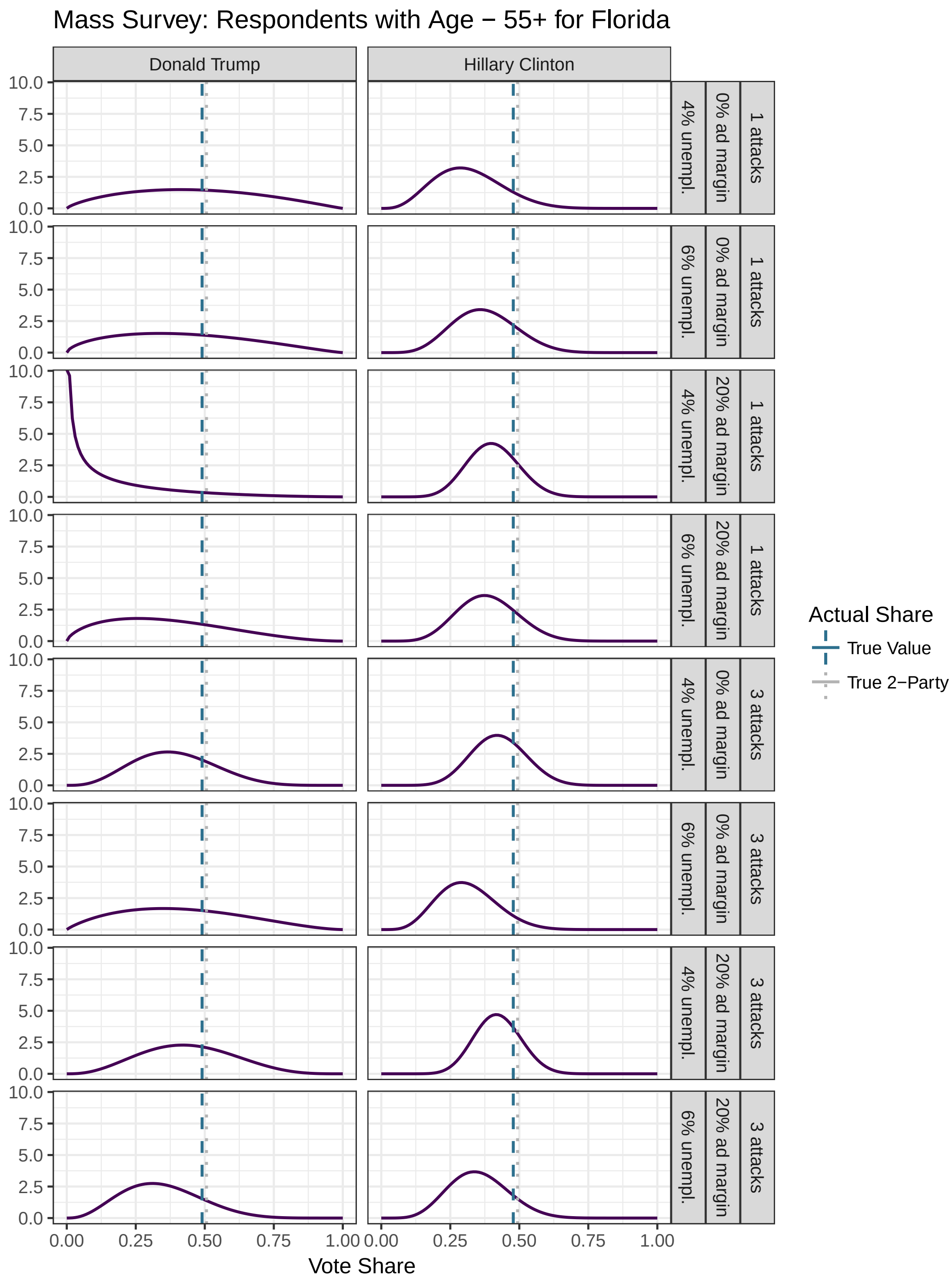


Figure 3.259: Priors with covariates: Mass Florida Age 55+

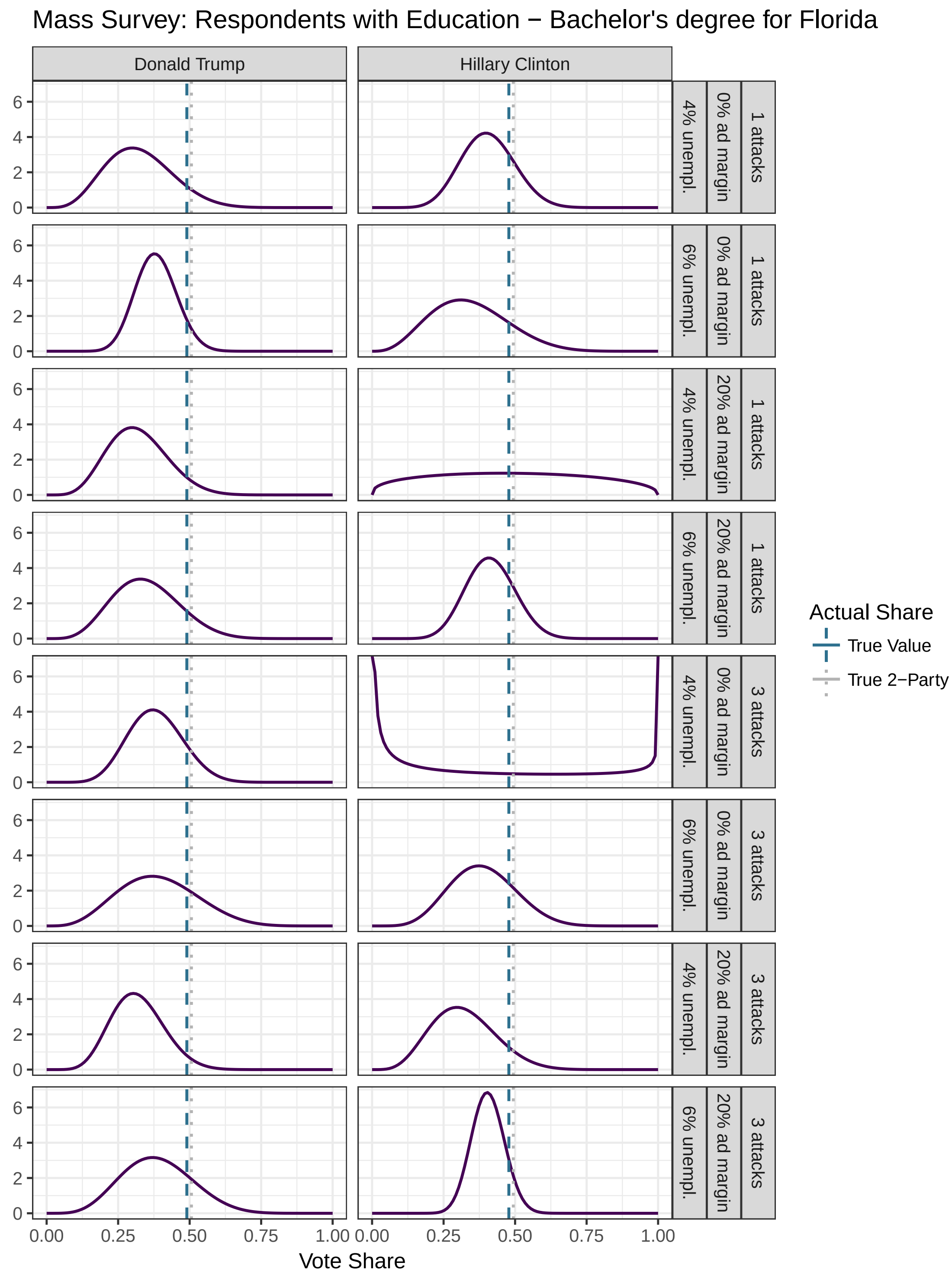


Figure 3.260: Priors with covariates: Mass Florida Education Bachelor's degree

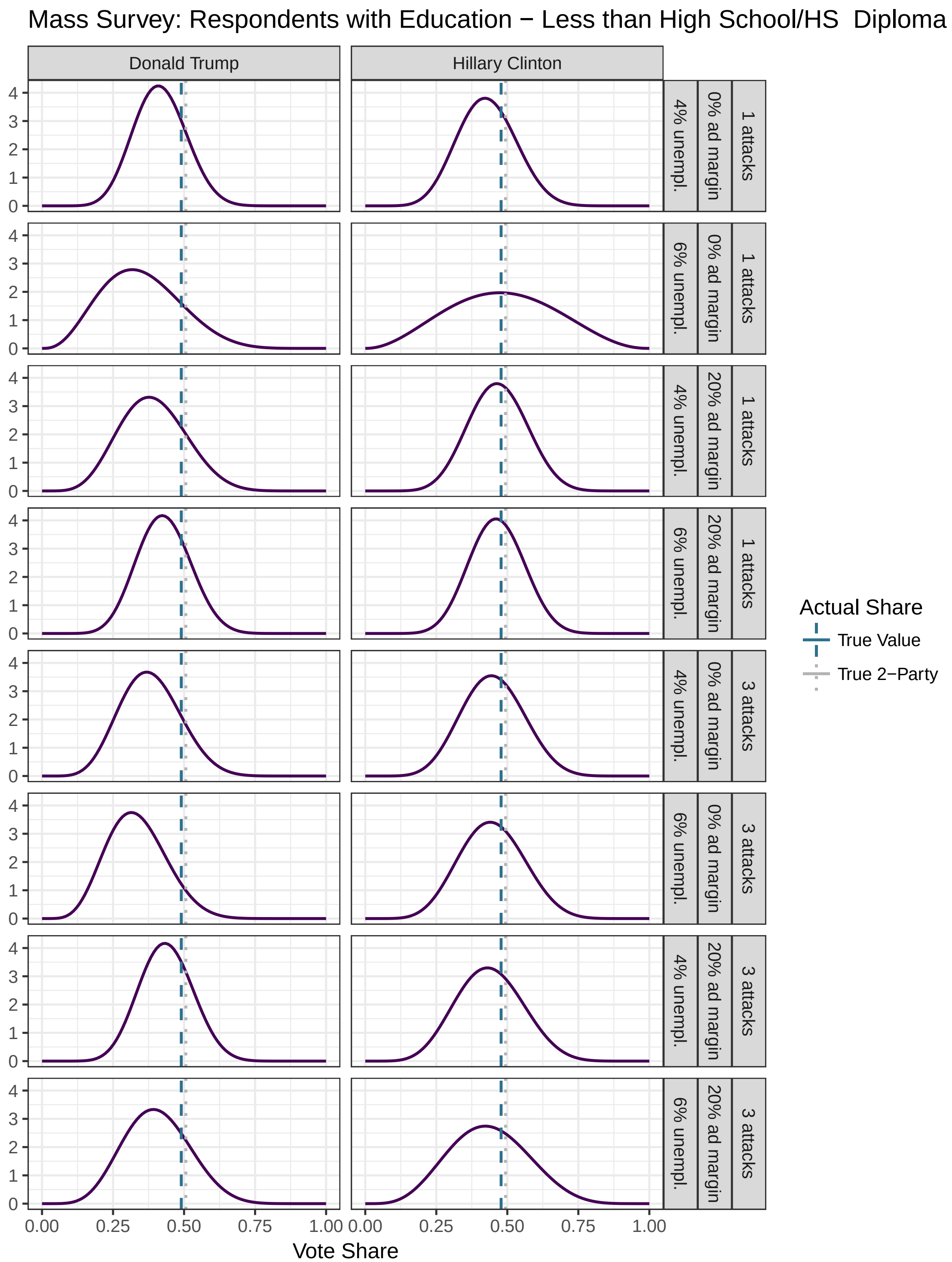


Figure 3.261: Priors with covariates: Mass Florida Education Less than High School HS Diploma

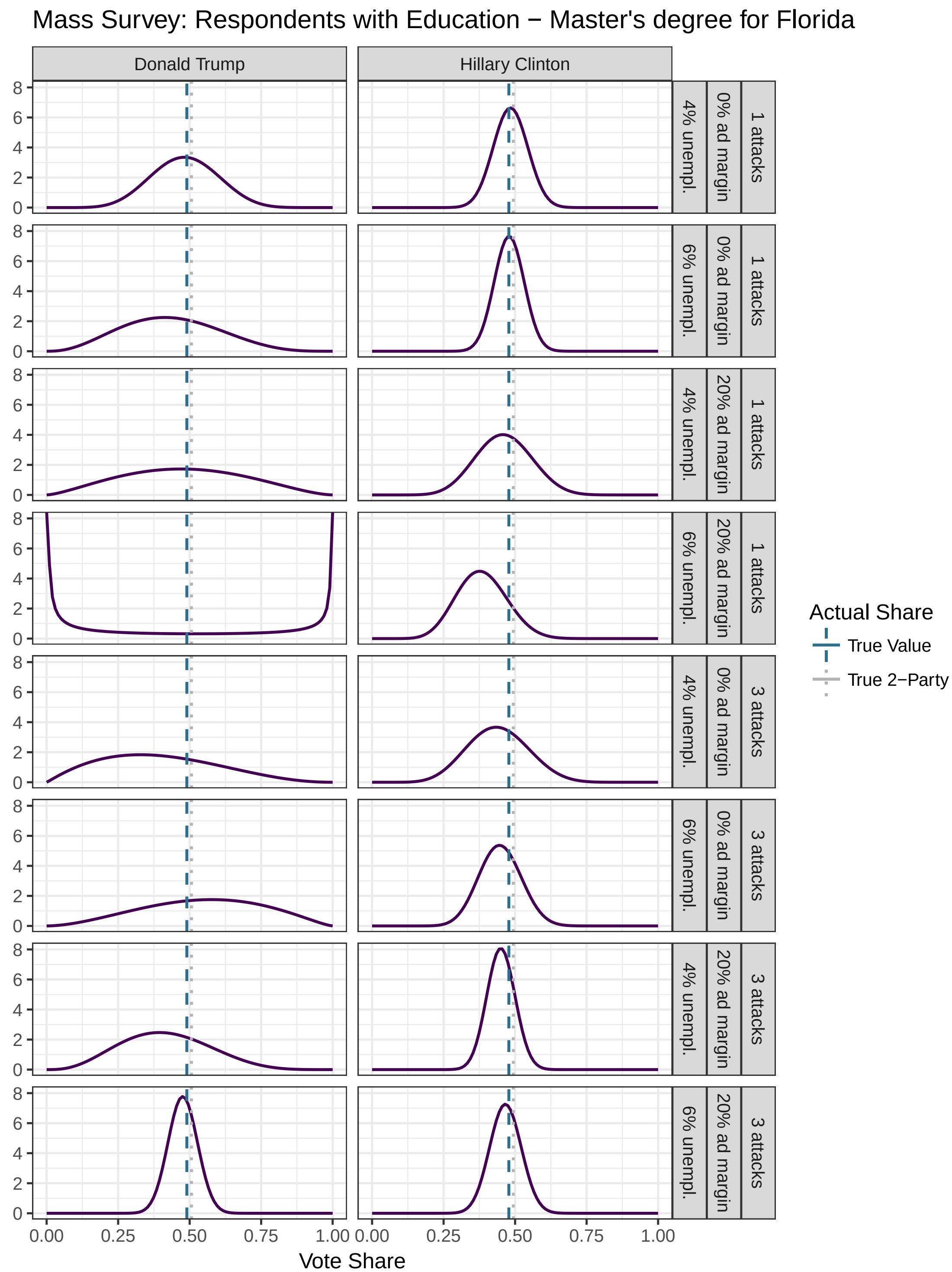


Figure 3.262: Priors with covariates: Mass Florida Education Master's degree

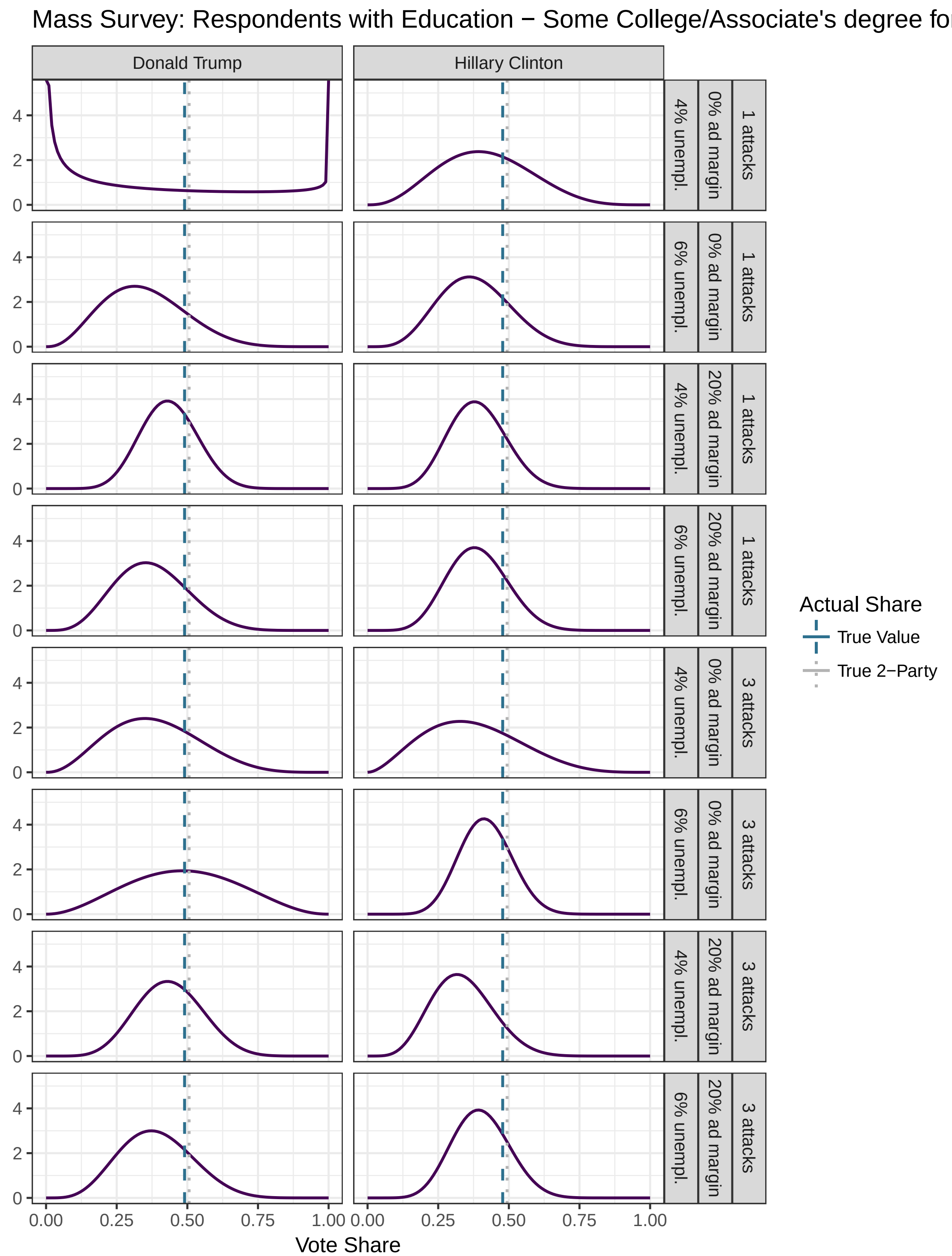


Figure 3.263: Priors with covariates: Mass Florida Education Some College Associate's degree

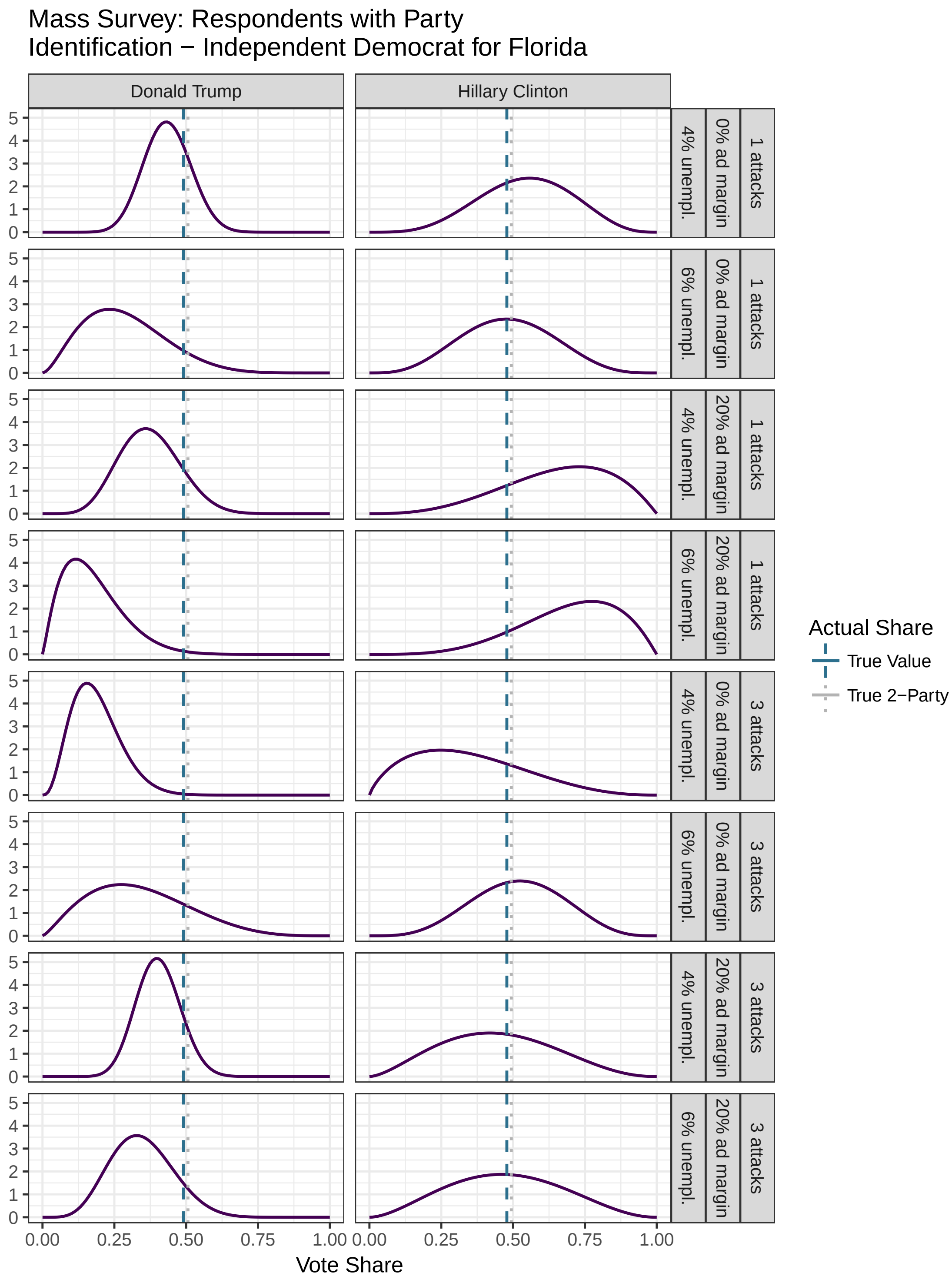


Figure 3.264: Priors with covariates: Mass Florida Party Identification Independent Democrat

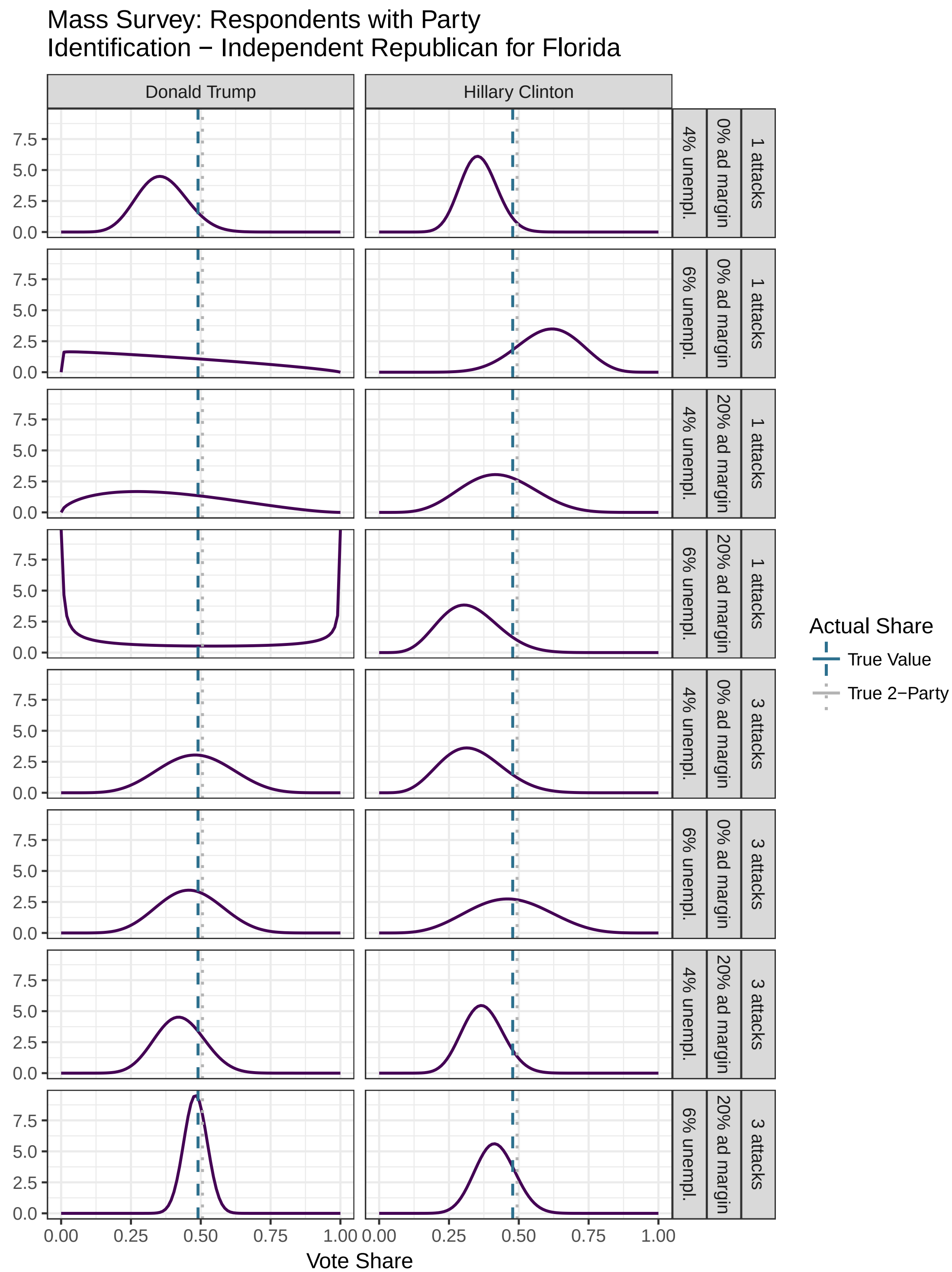


Figure 3.265: Priors with covariates: Mass Florida Party Identification Independent Republican

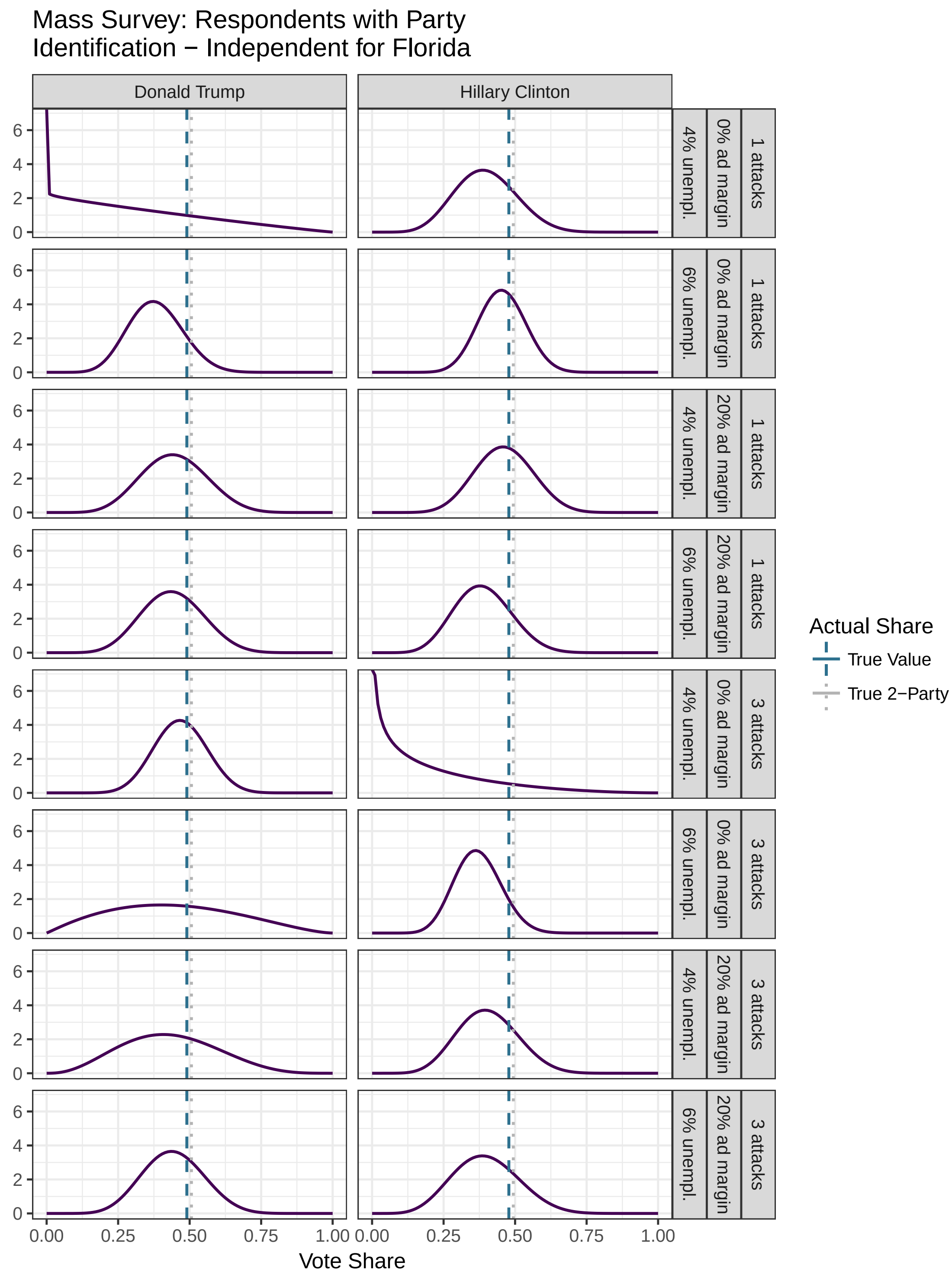


Figure 3.266: Priors with covariates: Mass Florida Party Identification Independent

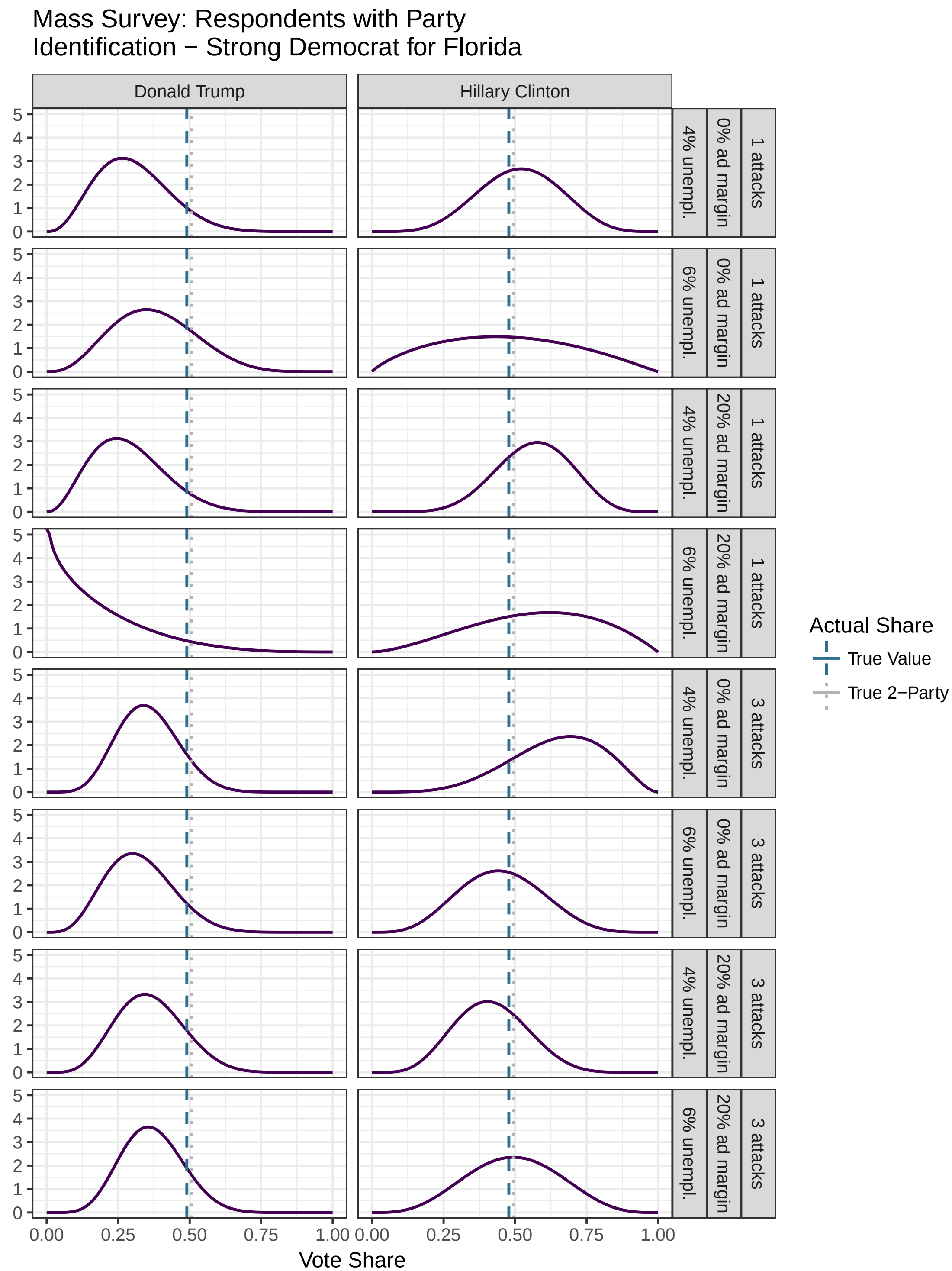


Figure 3.267: Priors with covariates: Mass Florida Party Identification Strong Democrat

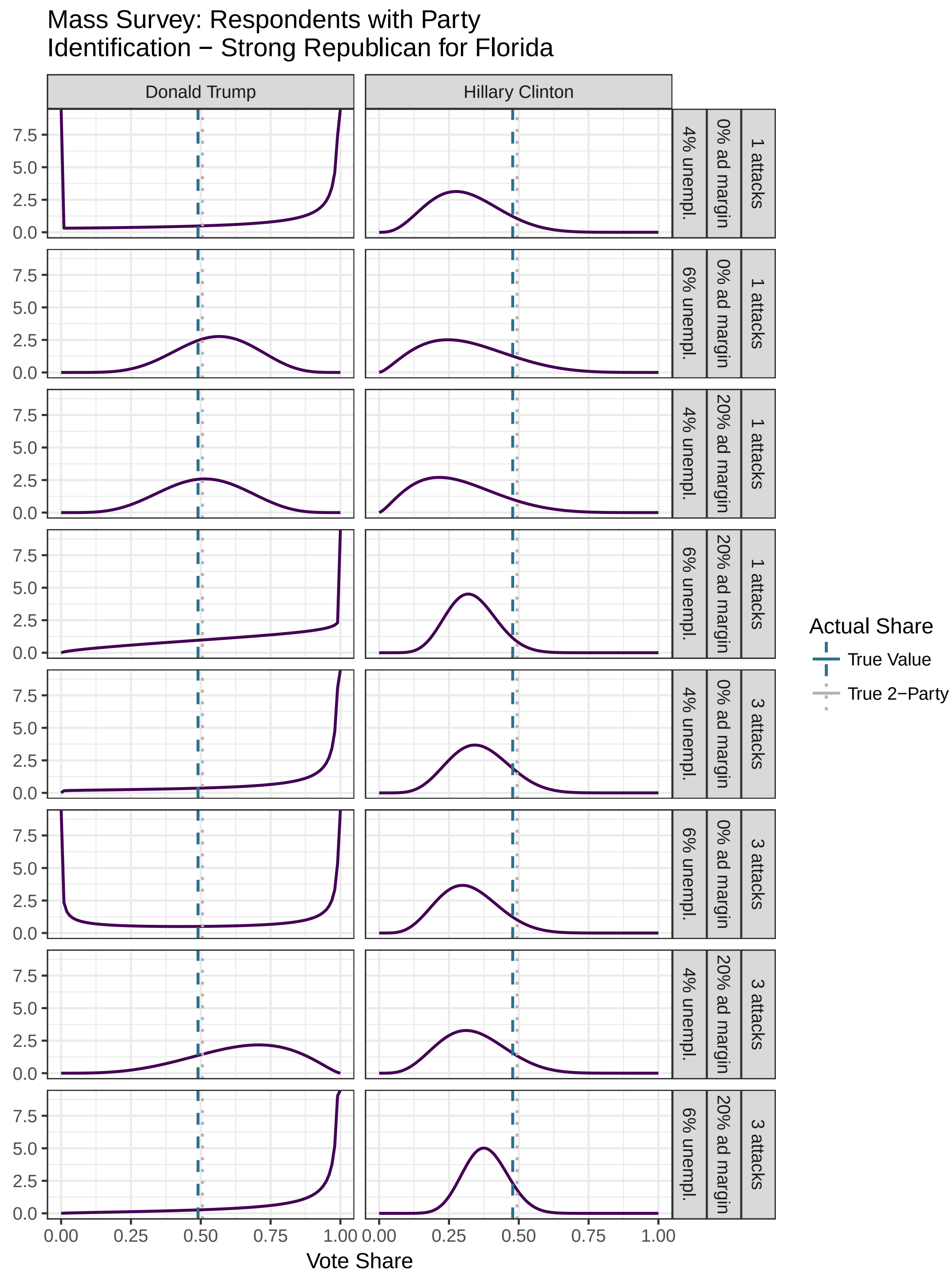


Figure 3.268: Priors with covariates: Mass Florida Party Identification Strong Republican

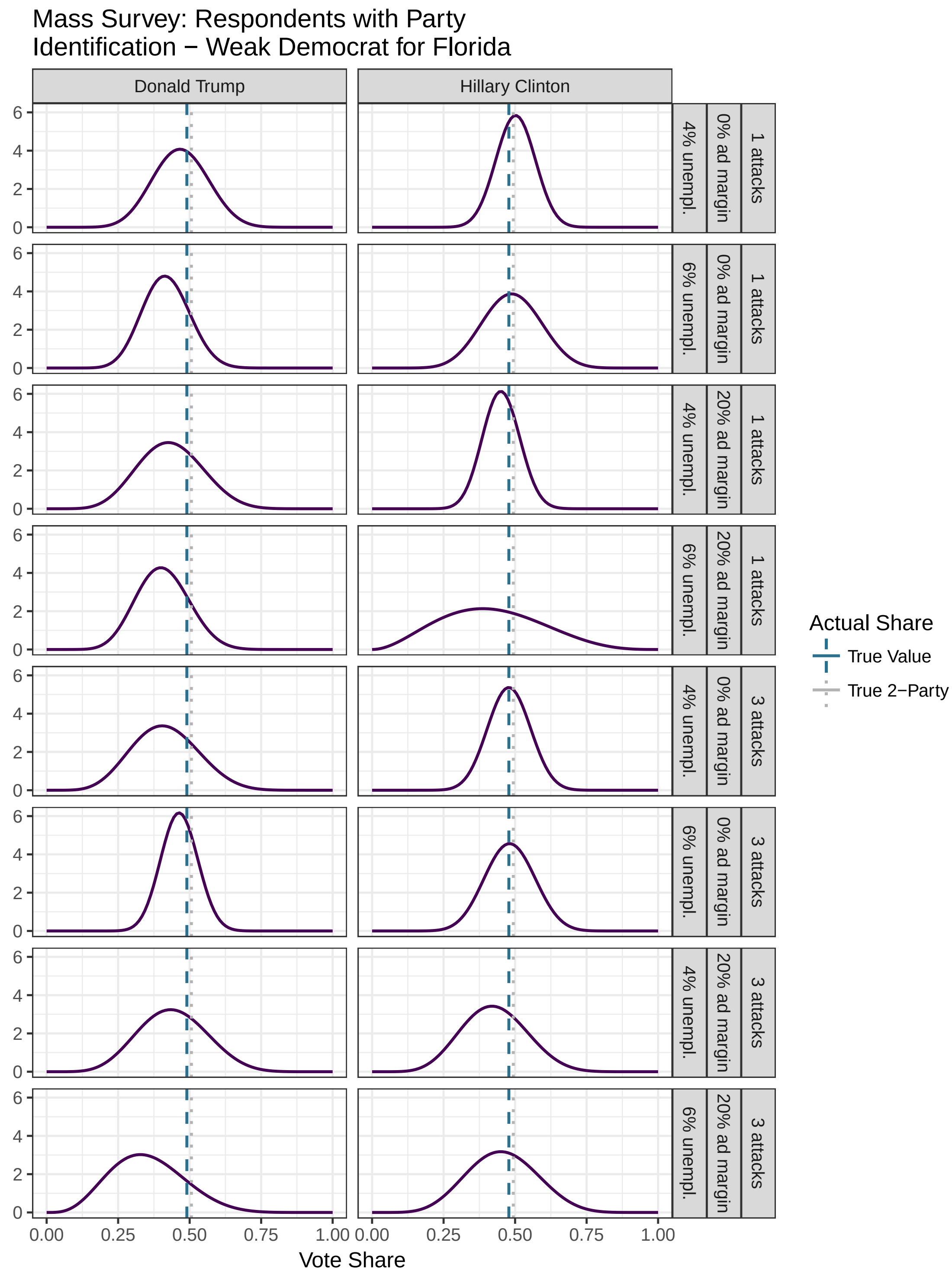


Figure 3.269: Priors with covariates: Mass Florida Party Identification Weak Democrat

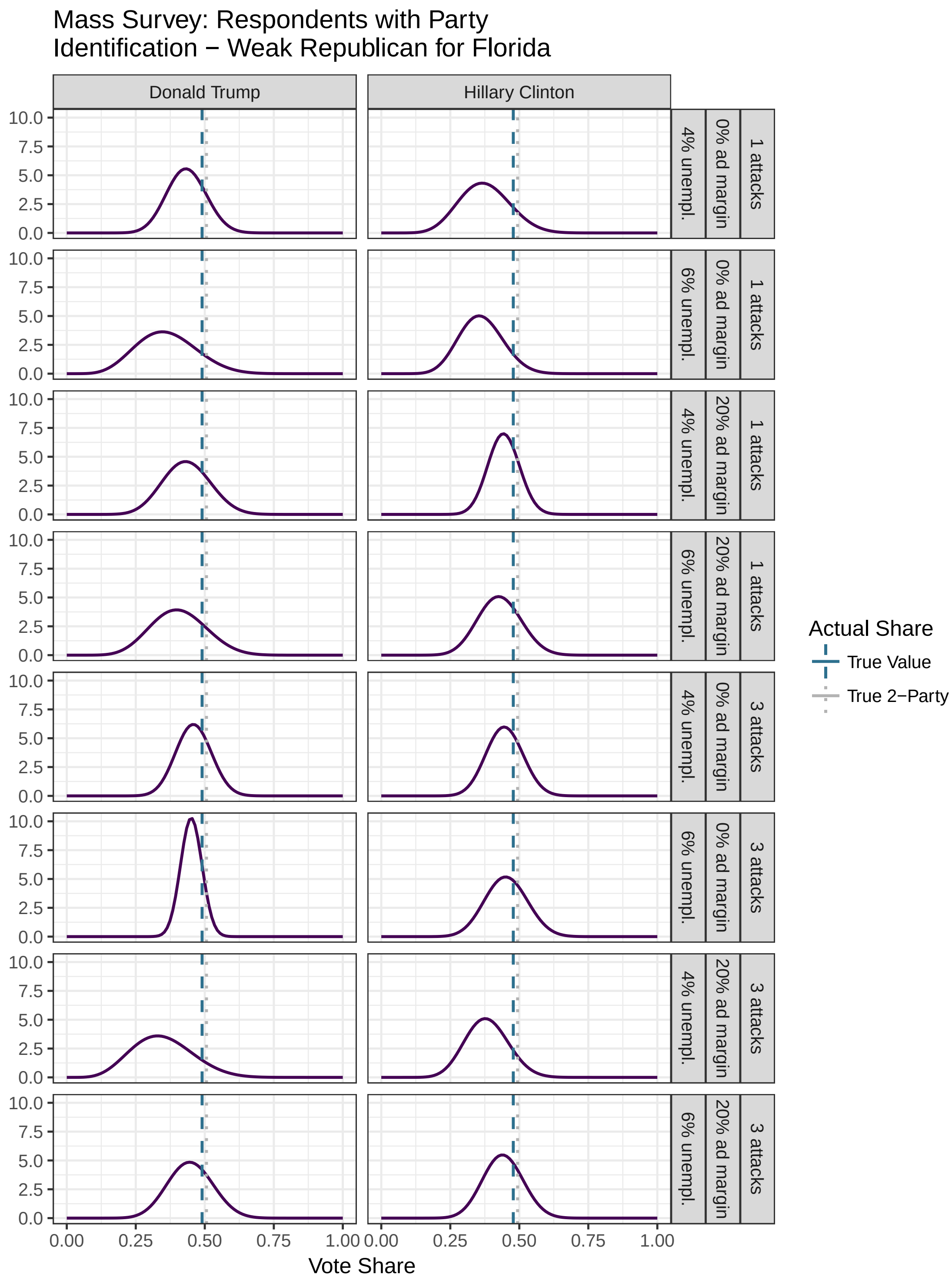


Figure 3.270: Priors with covariates: Mass Florida Party Identification Weak Republican

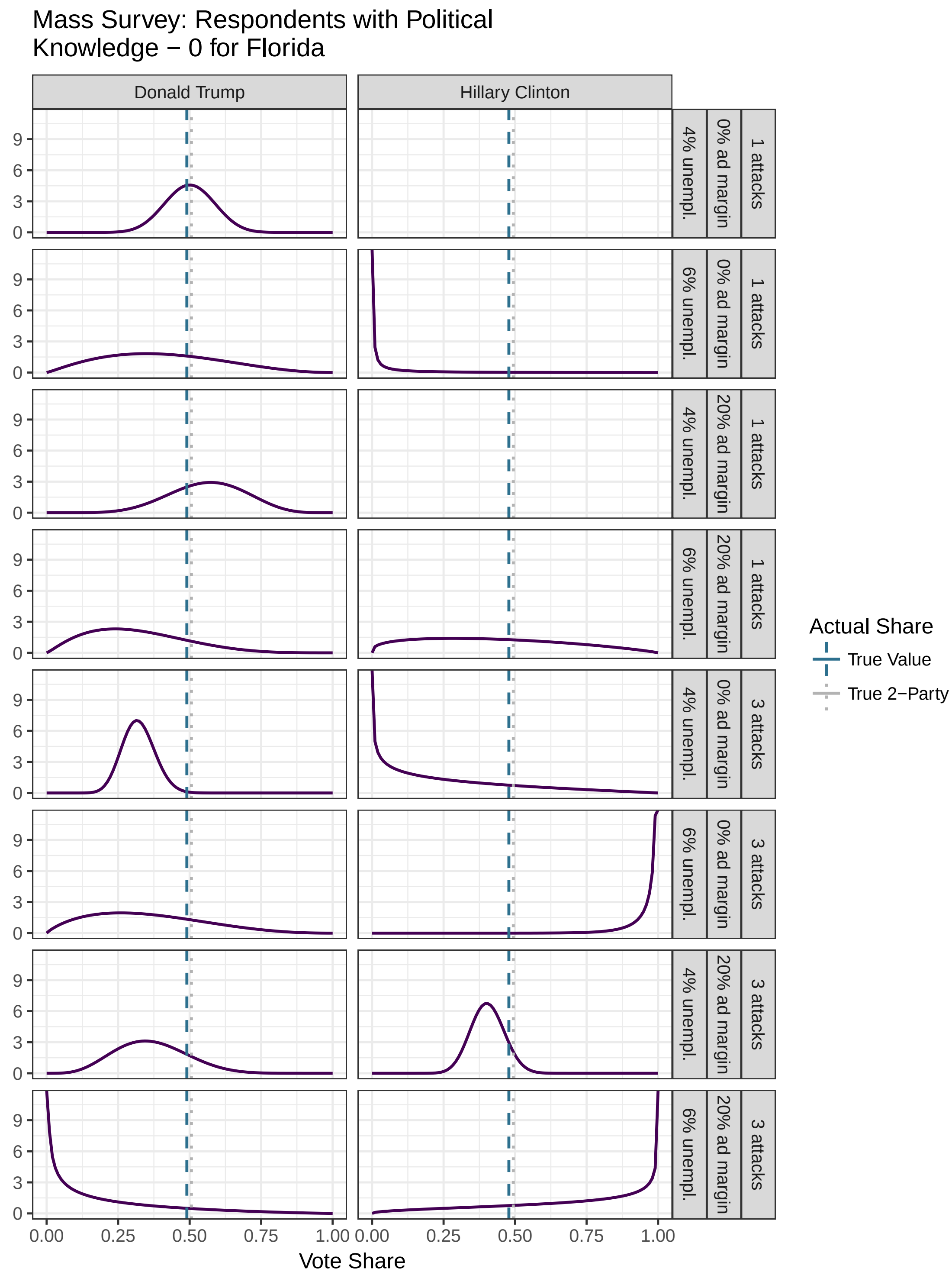


Figure 3.271: Priors with covariates: Mass Florida Political Knowledge 0

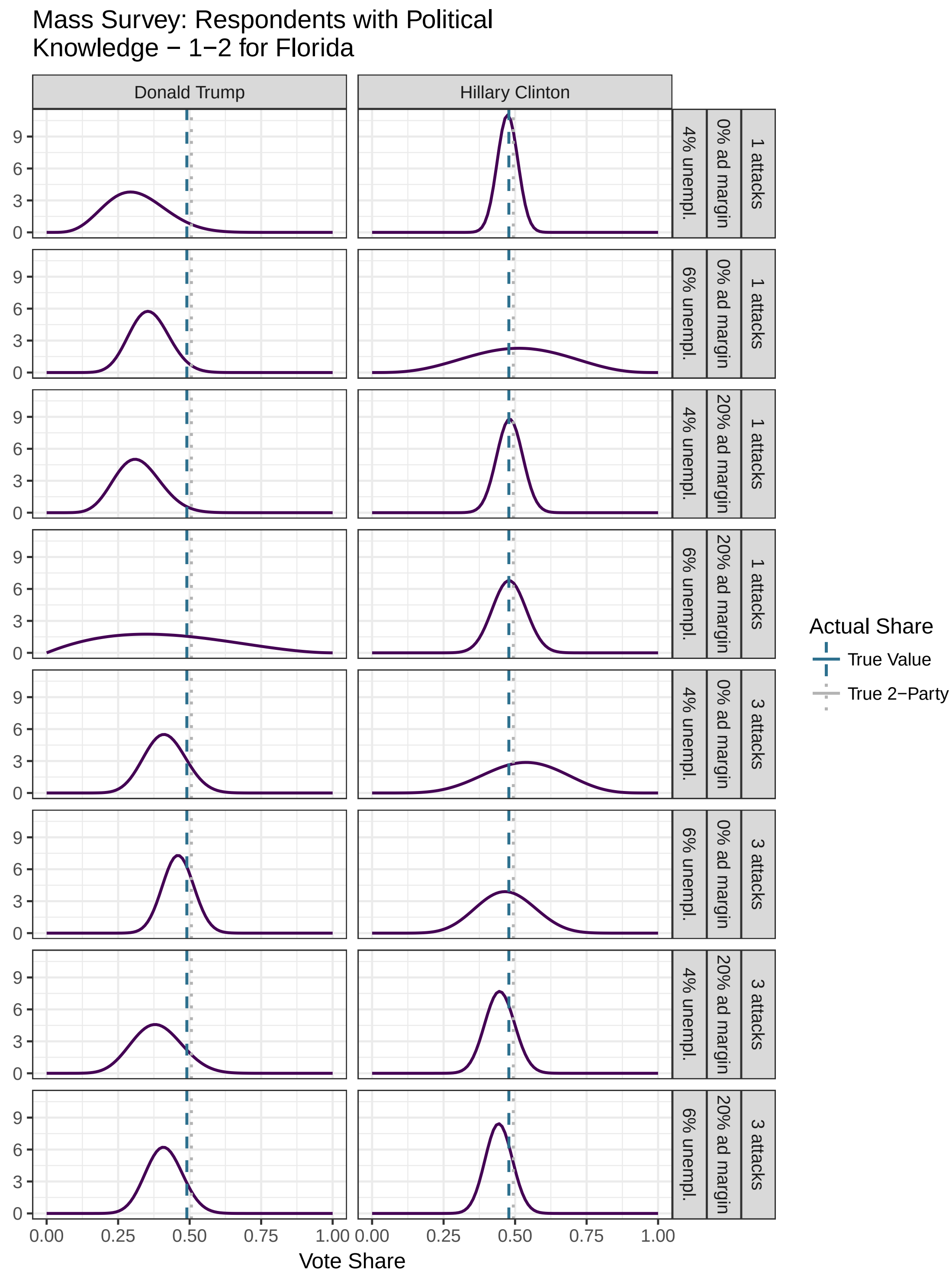


Figure 3.272: Priors with covariates: Mass Florida Political Knowledge 1-2

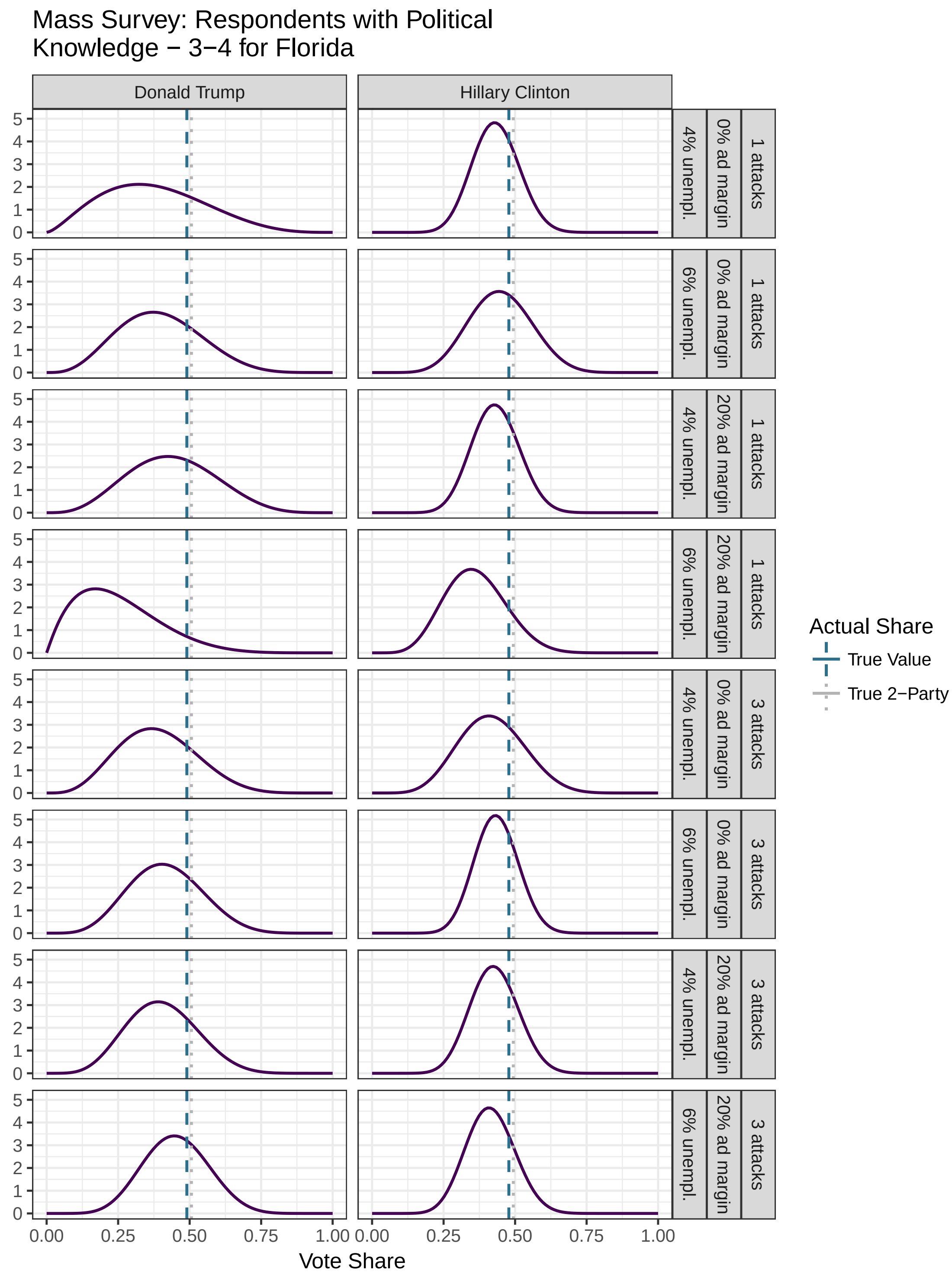


Figure 3.273: Priors with covariates: Mass Florida Political Knowledge 3-4

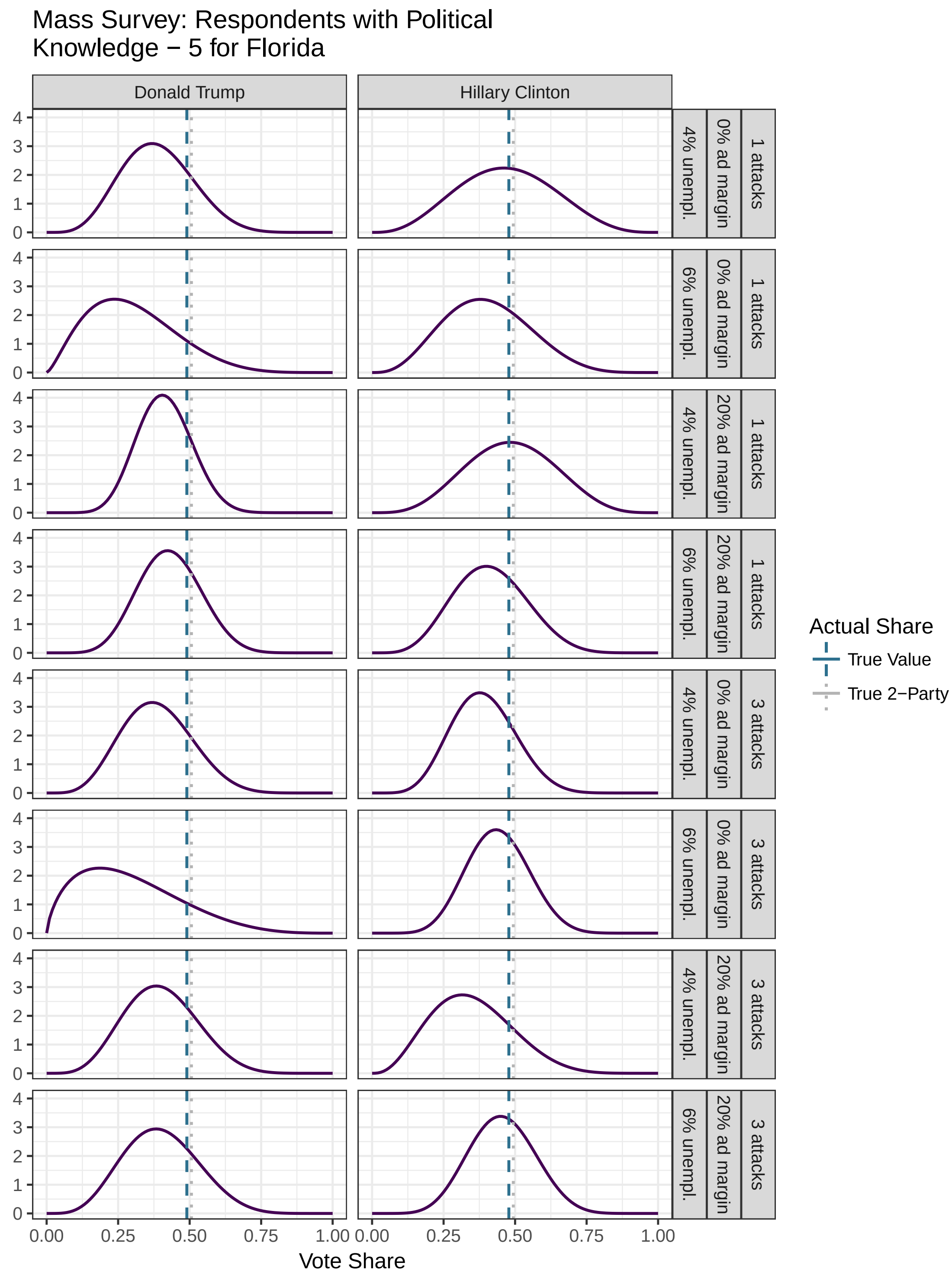


Figure 3.274: Priors with covariates: Mass Florida Political Knowledge 5

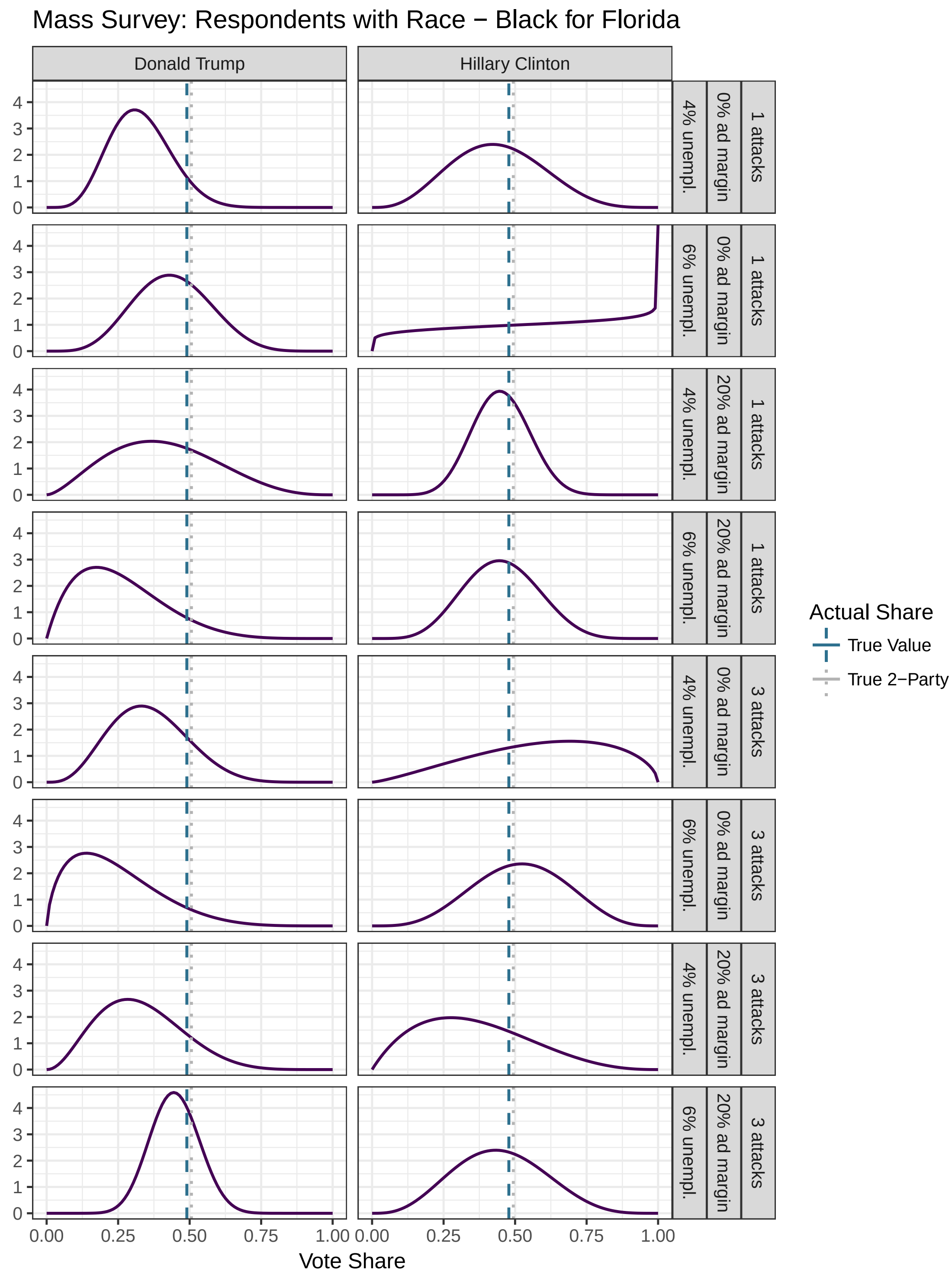


Figure 3.275: Priors with covariates: Mass Florida Race Black

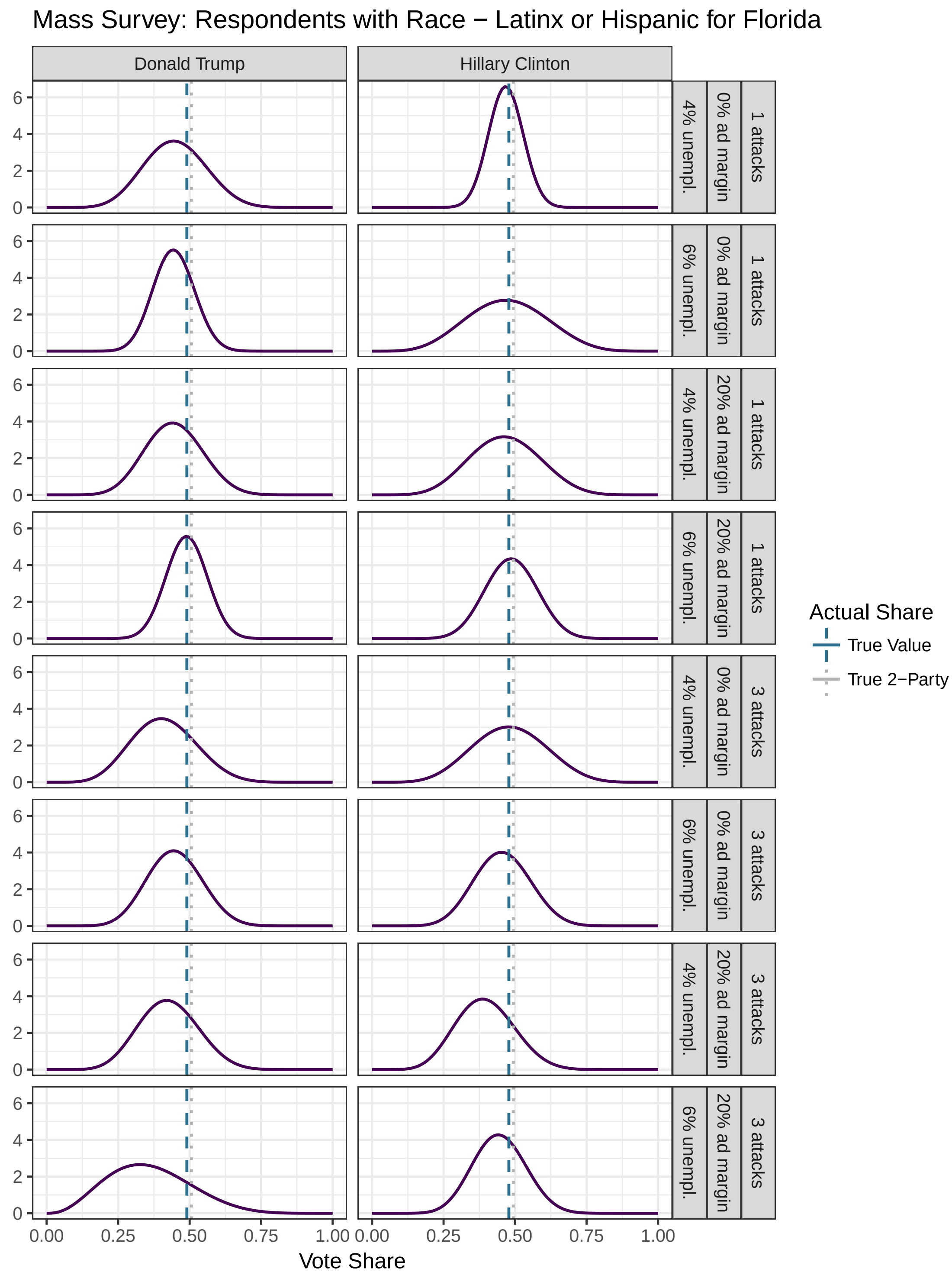


Figure 3.276: Priors with covariates: Mass Florida Race Latinx or Hispanic

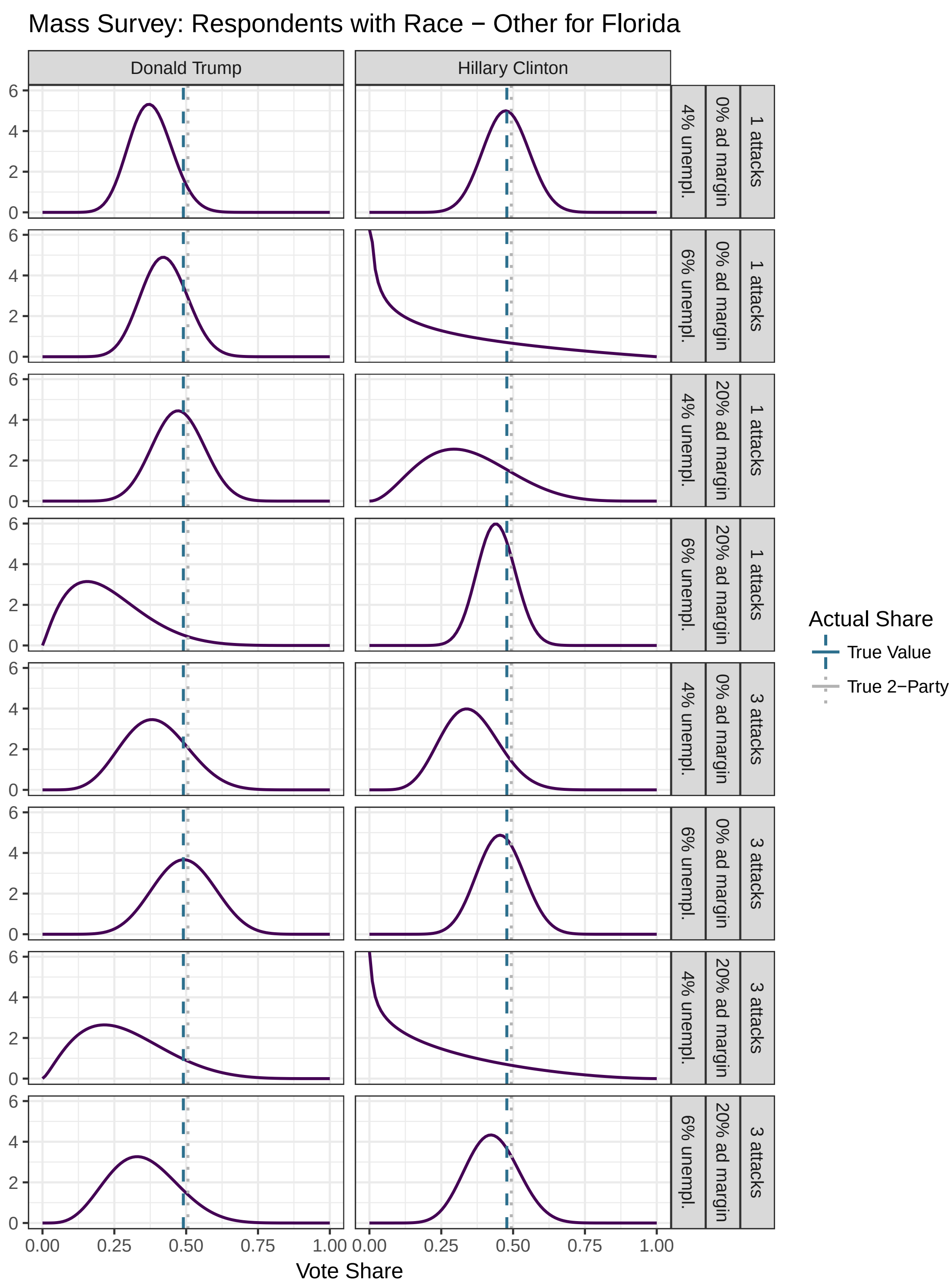


Figure 3.277: Priors with covariates: Mass Florida Race Other

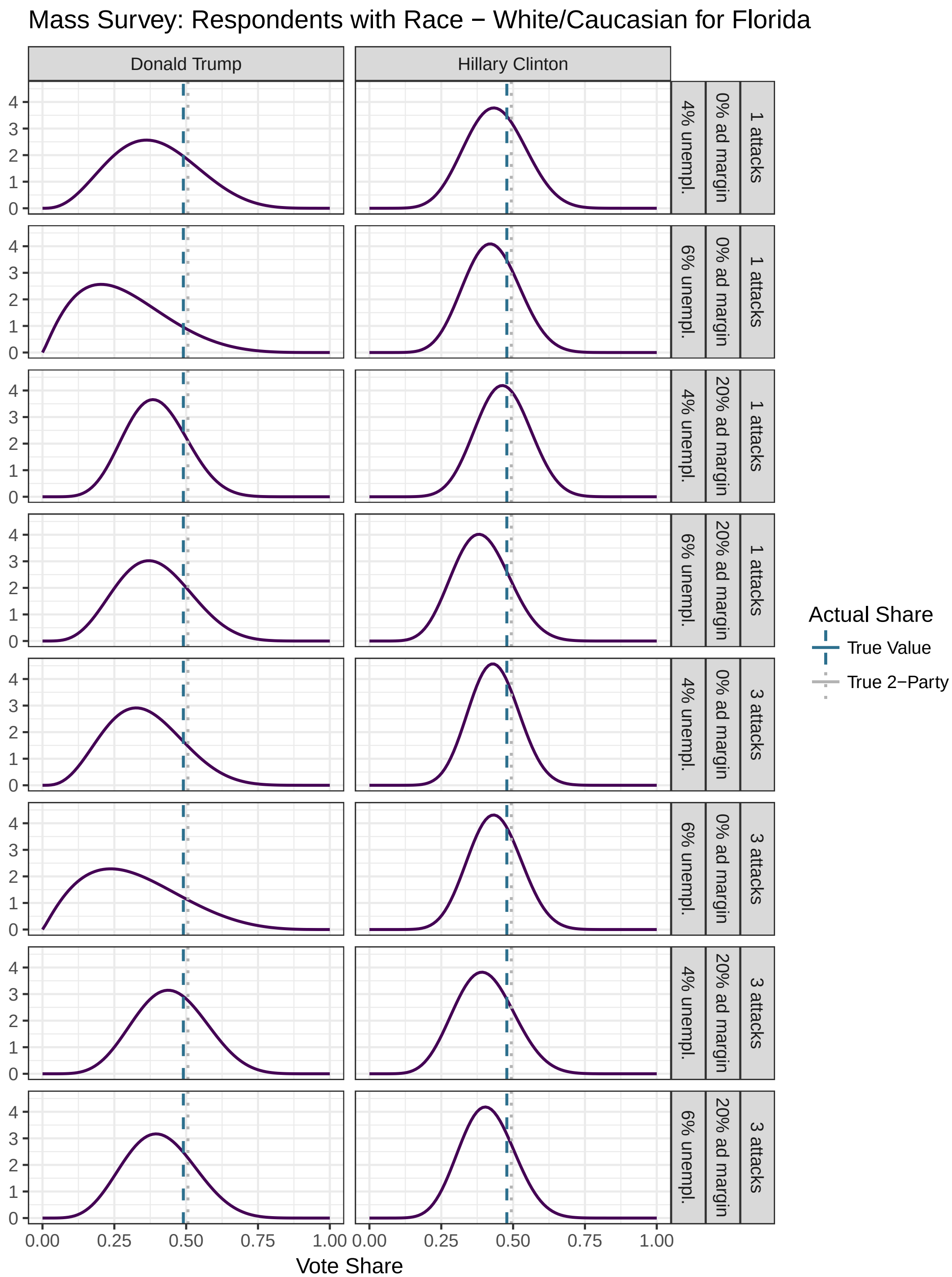


Figure 3.278: Priors with covariates: Mass Florida Race White Caucasian

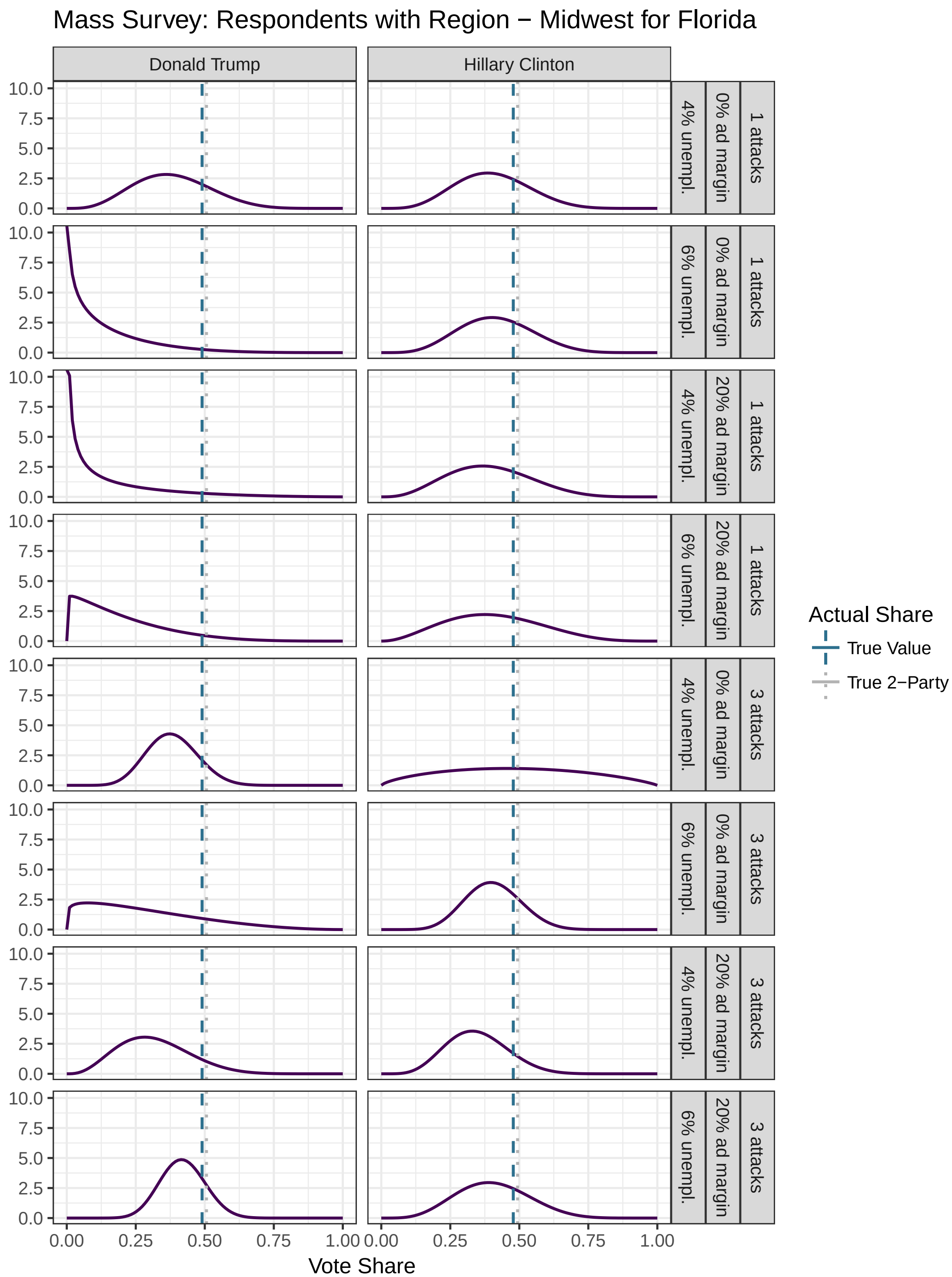


Figure 3.279: Priors with covariates: Mass Florida Region Midwest

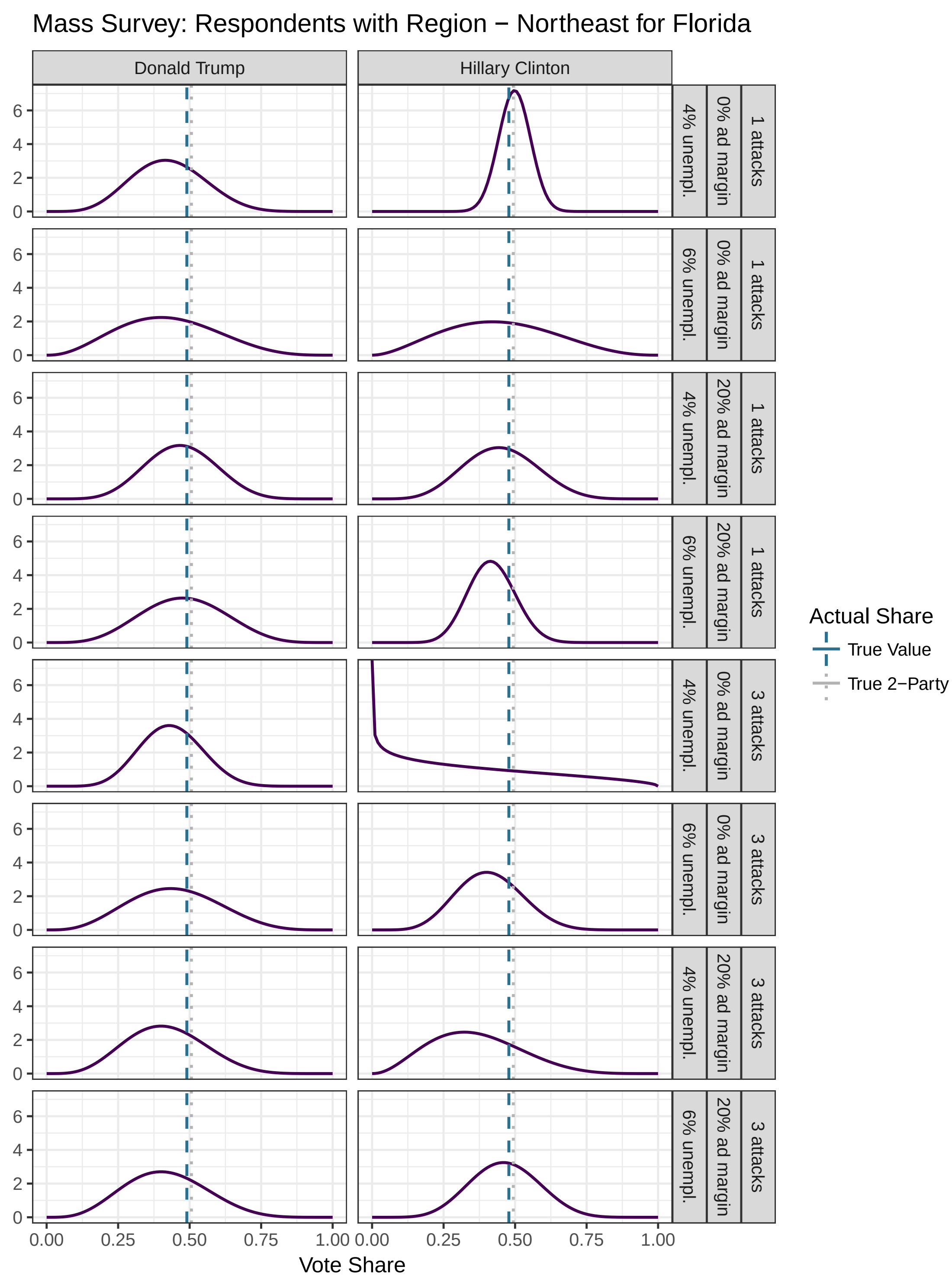


Figure 3.280: Priors with covariates: Mass Florida Region Northeast

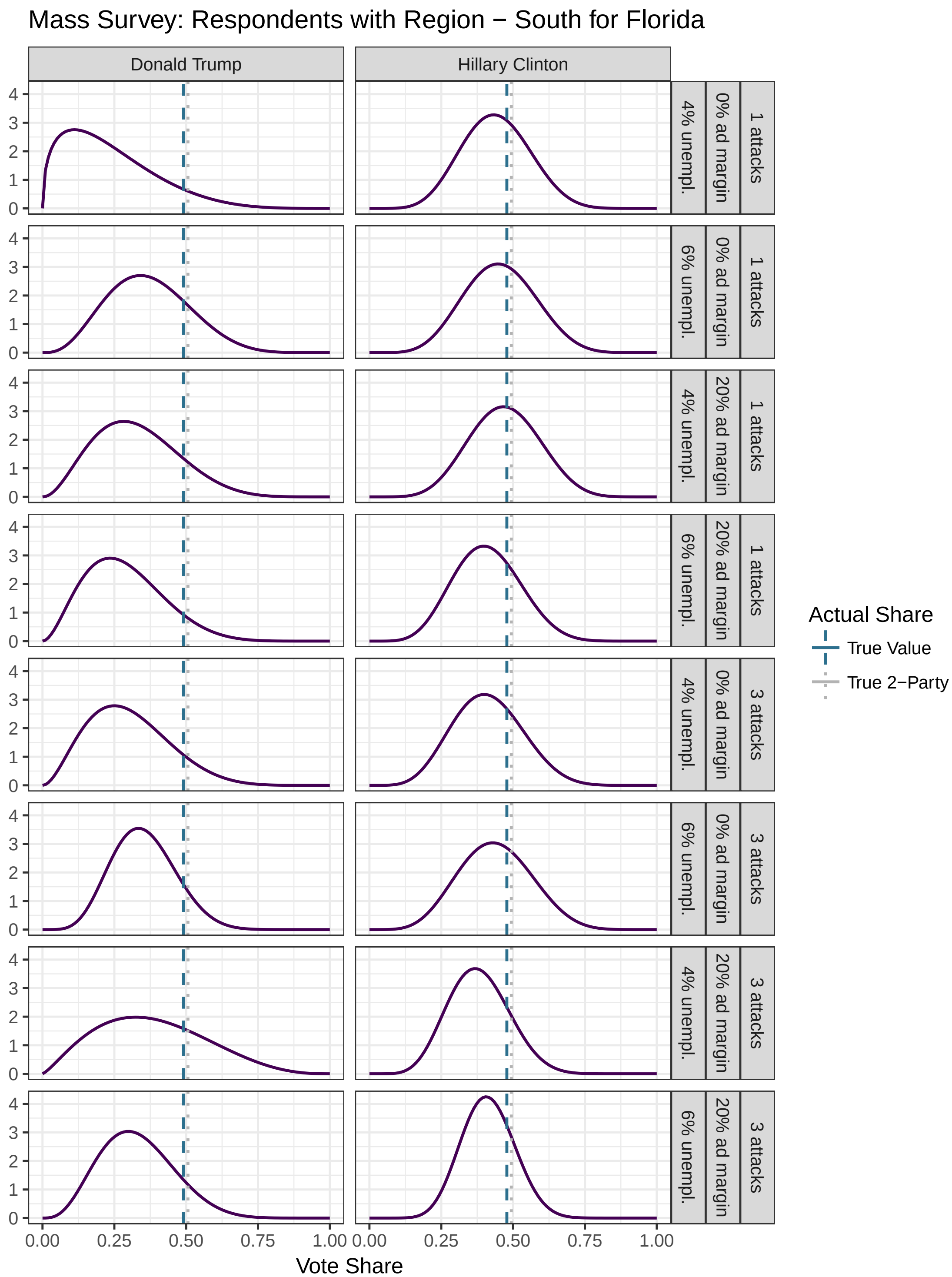


Figure 3.281: Priors with covariates: Mass Florida Region South

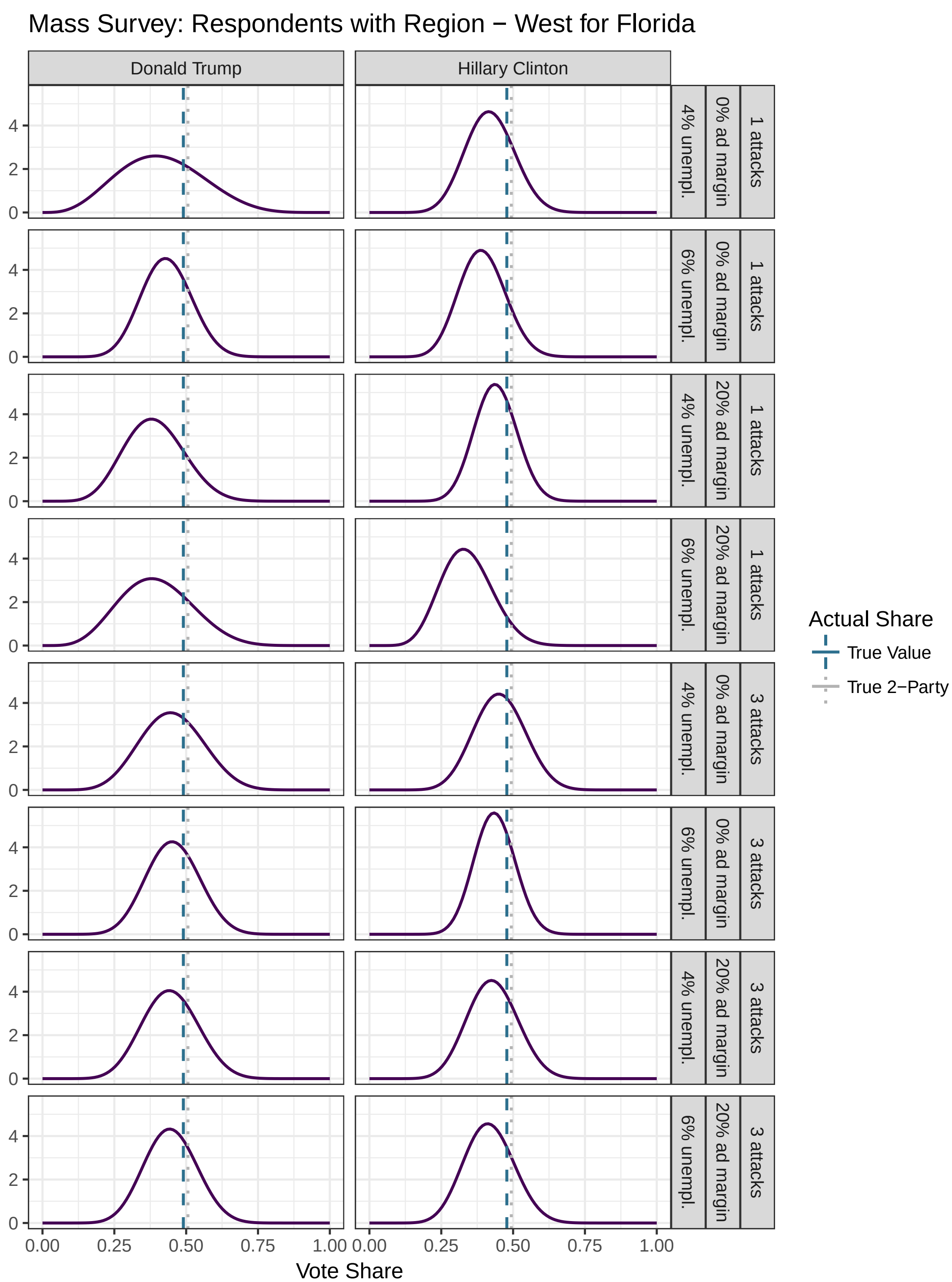


Figure 3.282: Priors with covariates: Mass Florida Region West

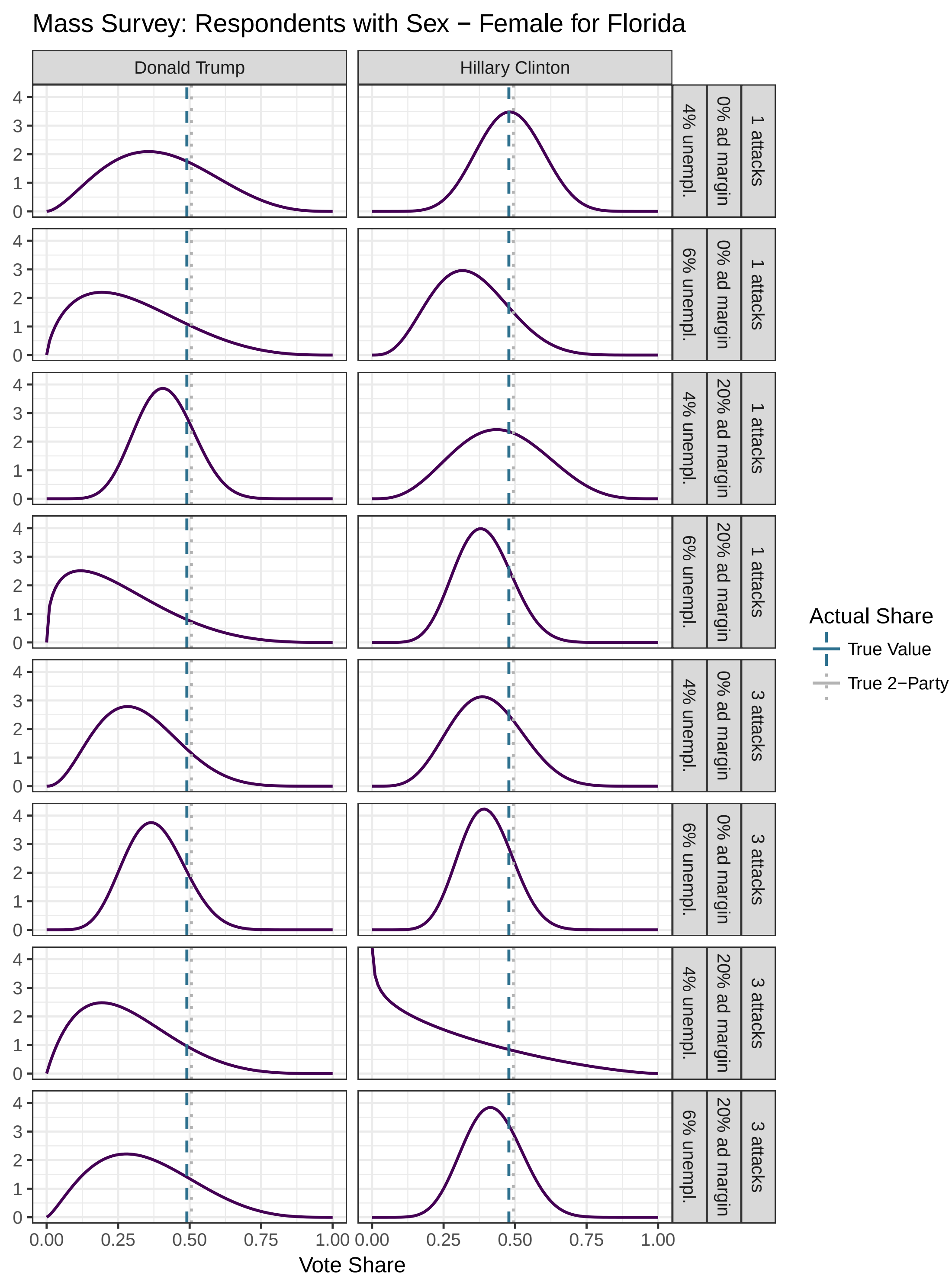


Figure 3.283: Priors with covariates: Mass Florida Sex Female

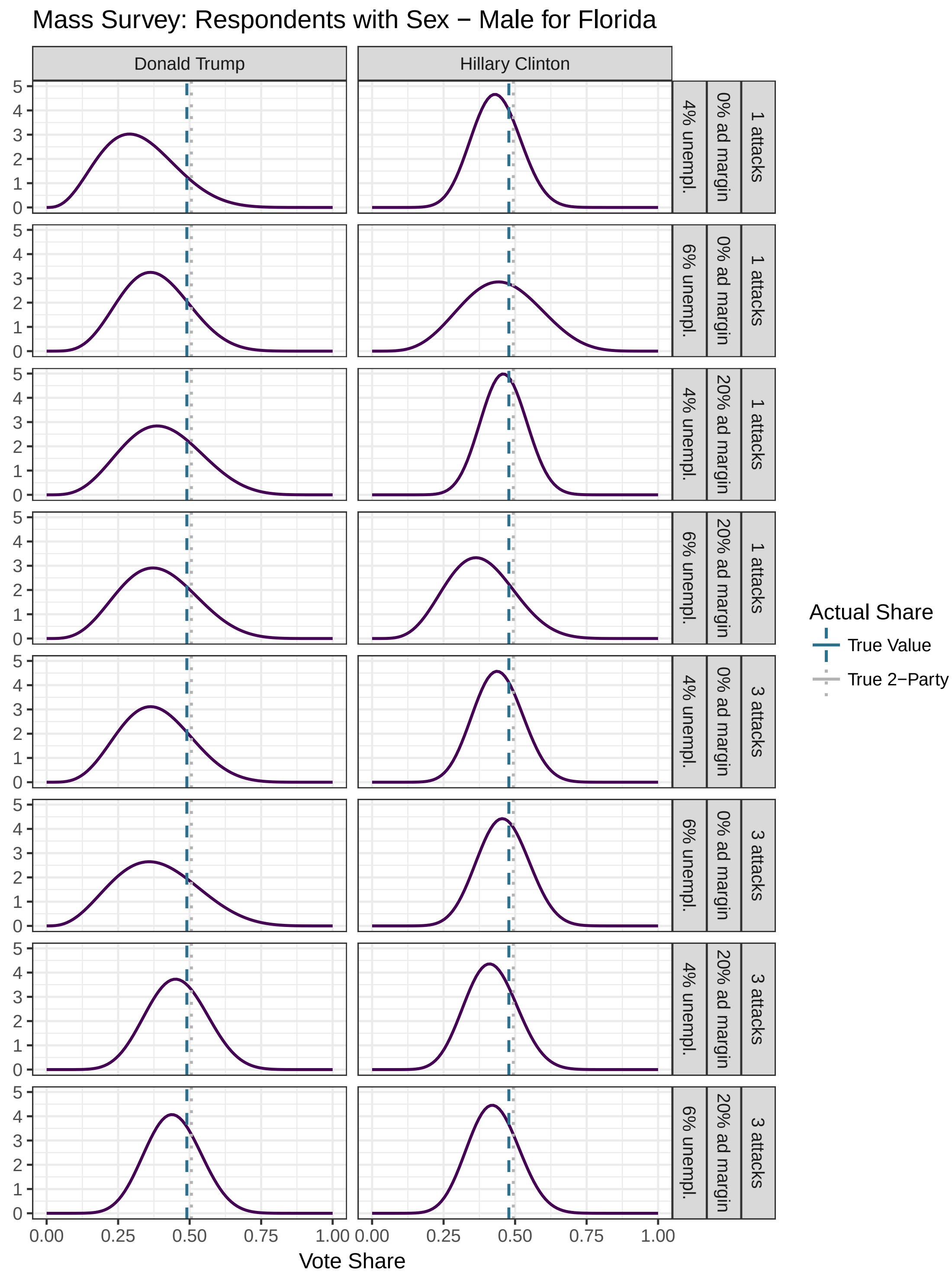


Figure 3.284: Priors with covariates: Mass Florida Sex Male

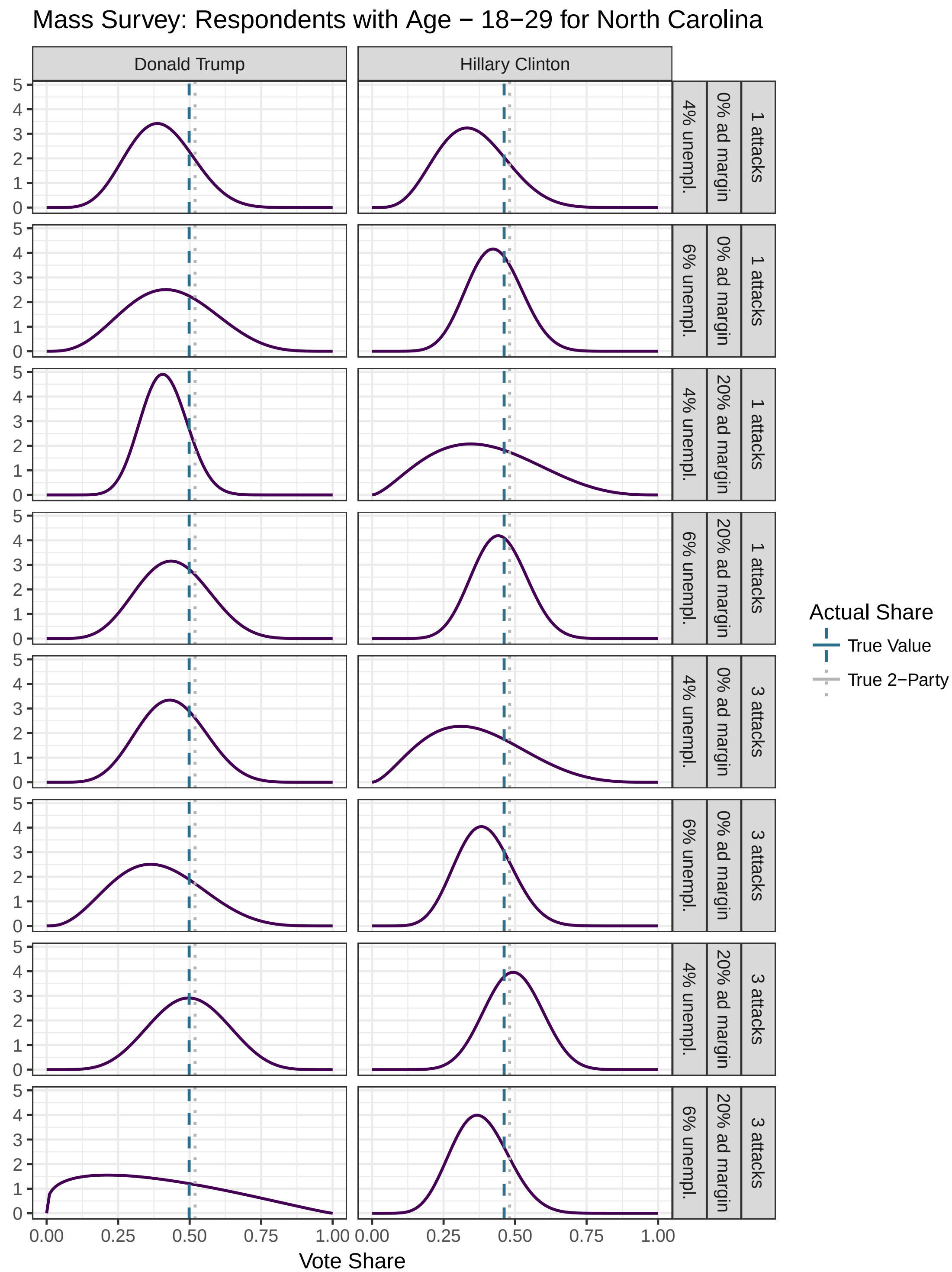


Figure 3.285: Priors with covariates: Mass North Carolina Age 18-29

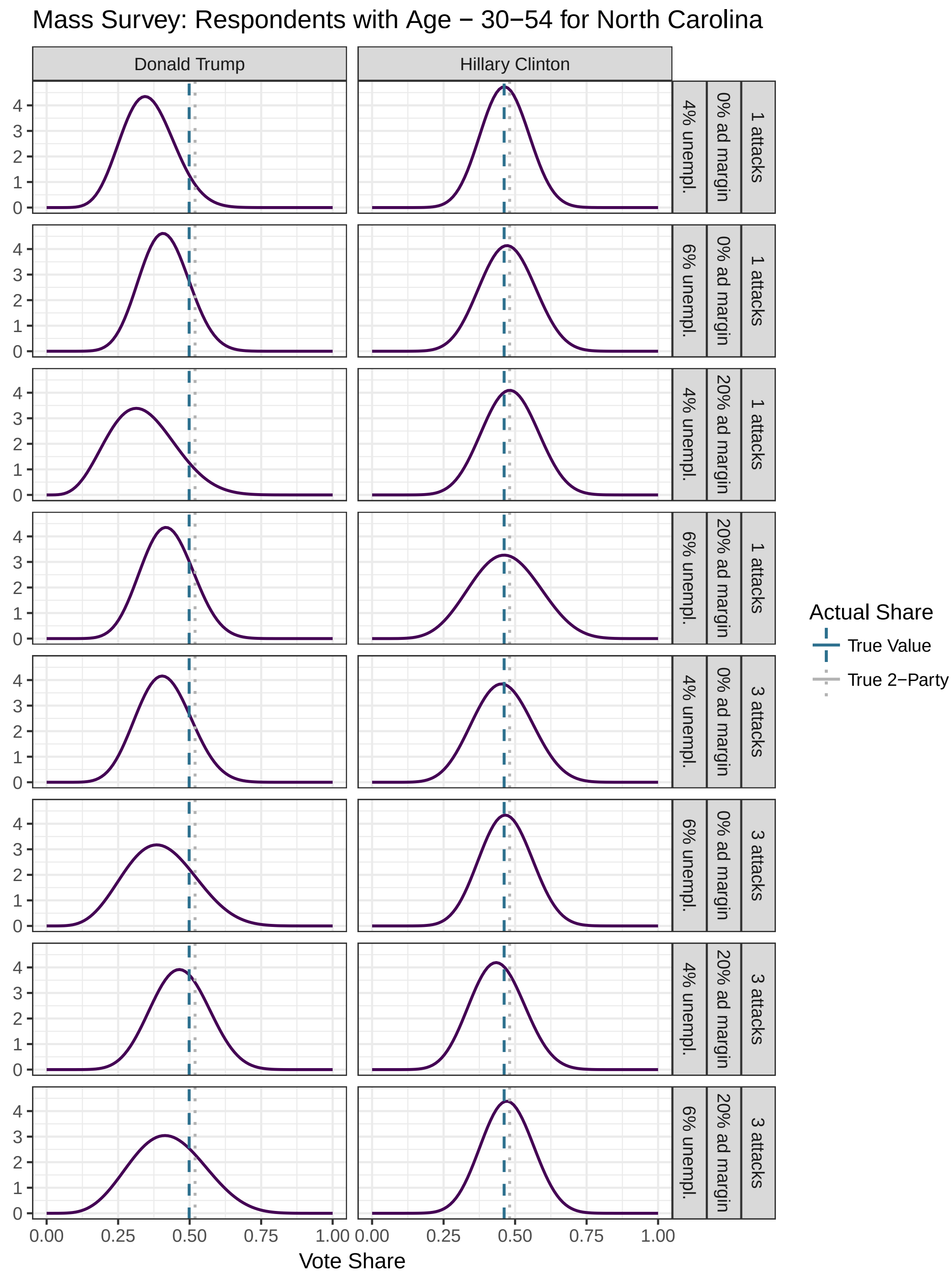


Figure 3.286: Priors with covariates: Mass North Carolina Age 30-54

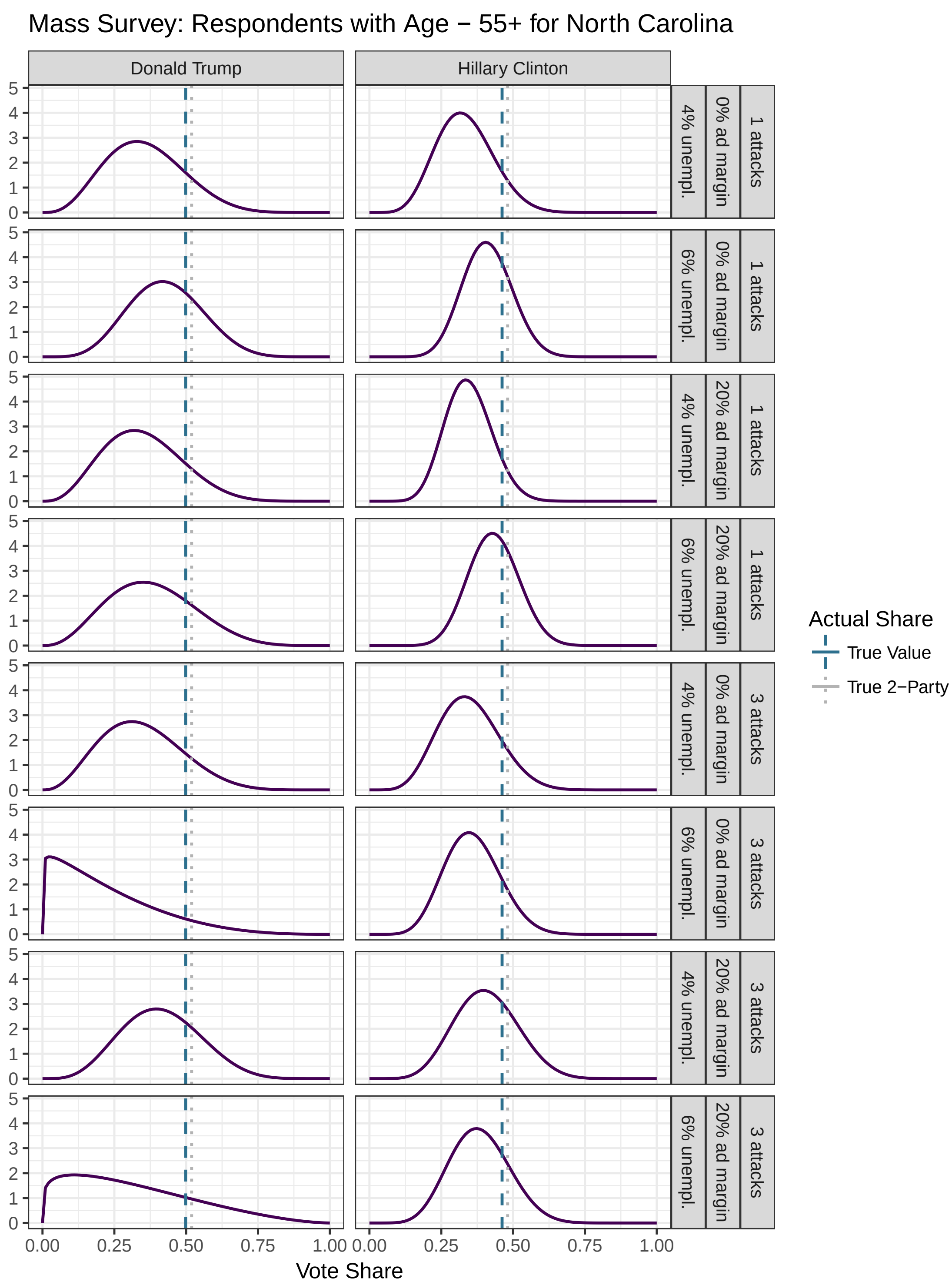


Figure 3.287: Priors with covariates: Mass North Carolina Age 55+

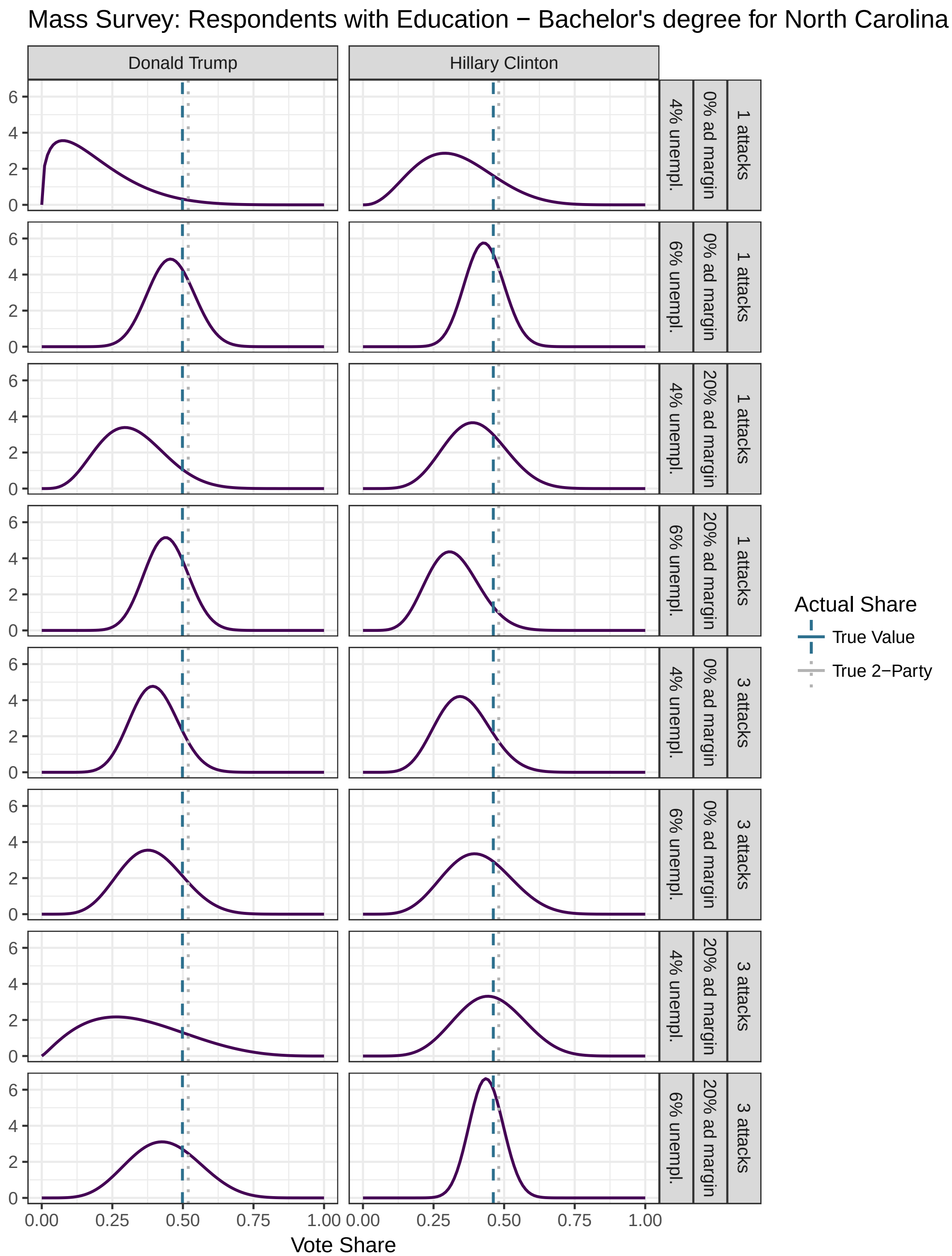


Figure 3.288: Priors with covariates: Mass North Carolina Education Bachelor's degree

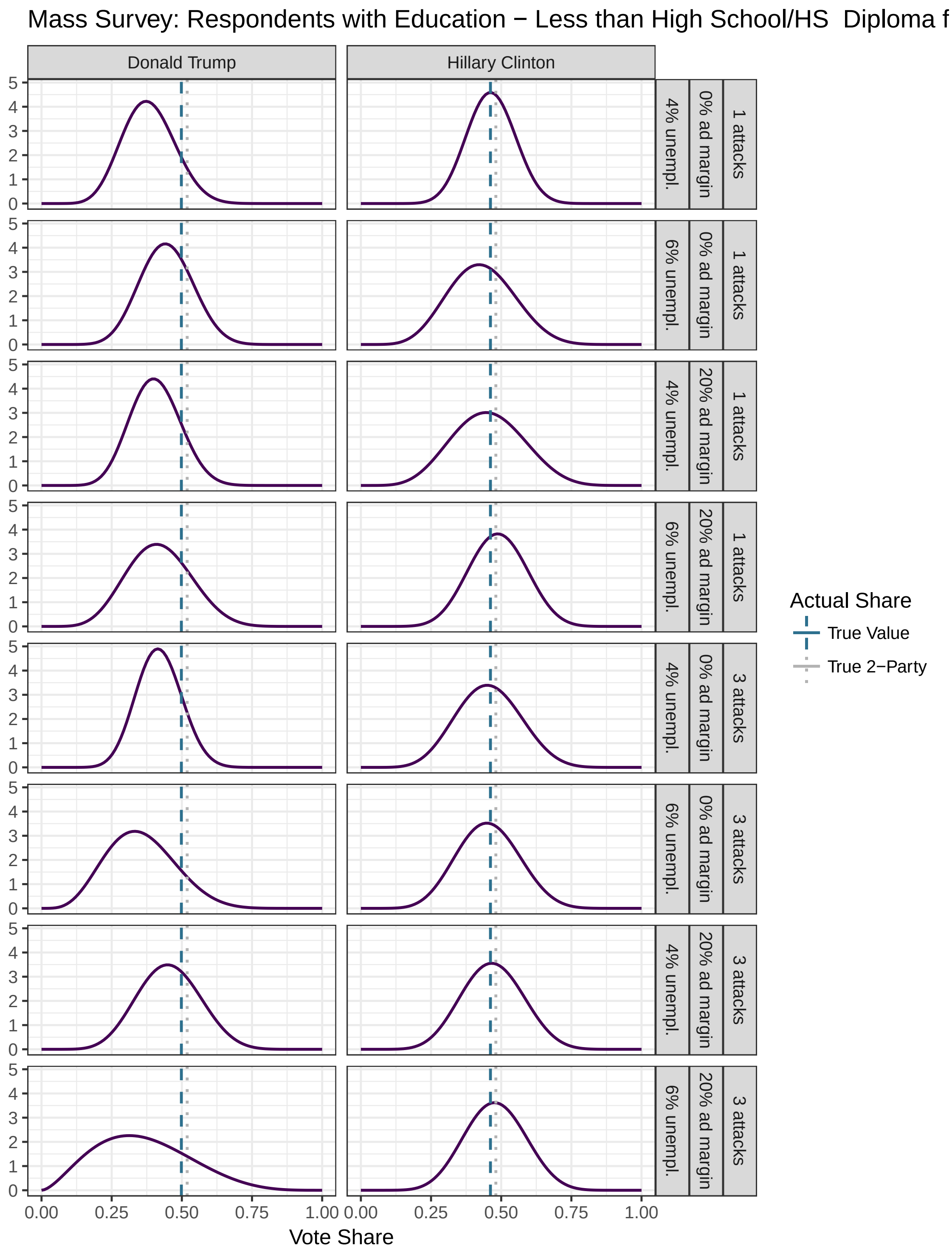


Figure 3.289: Priors with covariates: Mass North Carolina Education Less than High School HS Diploma

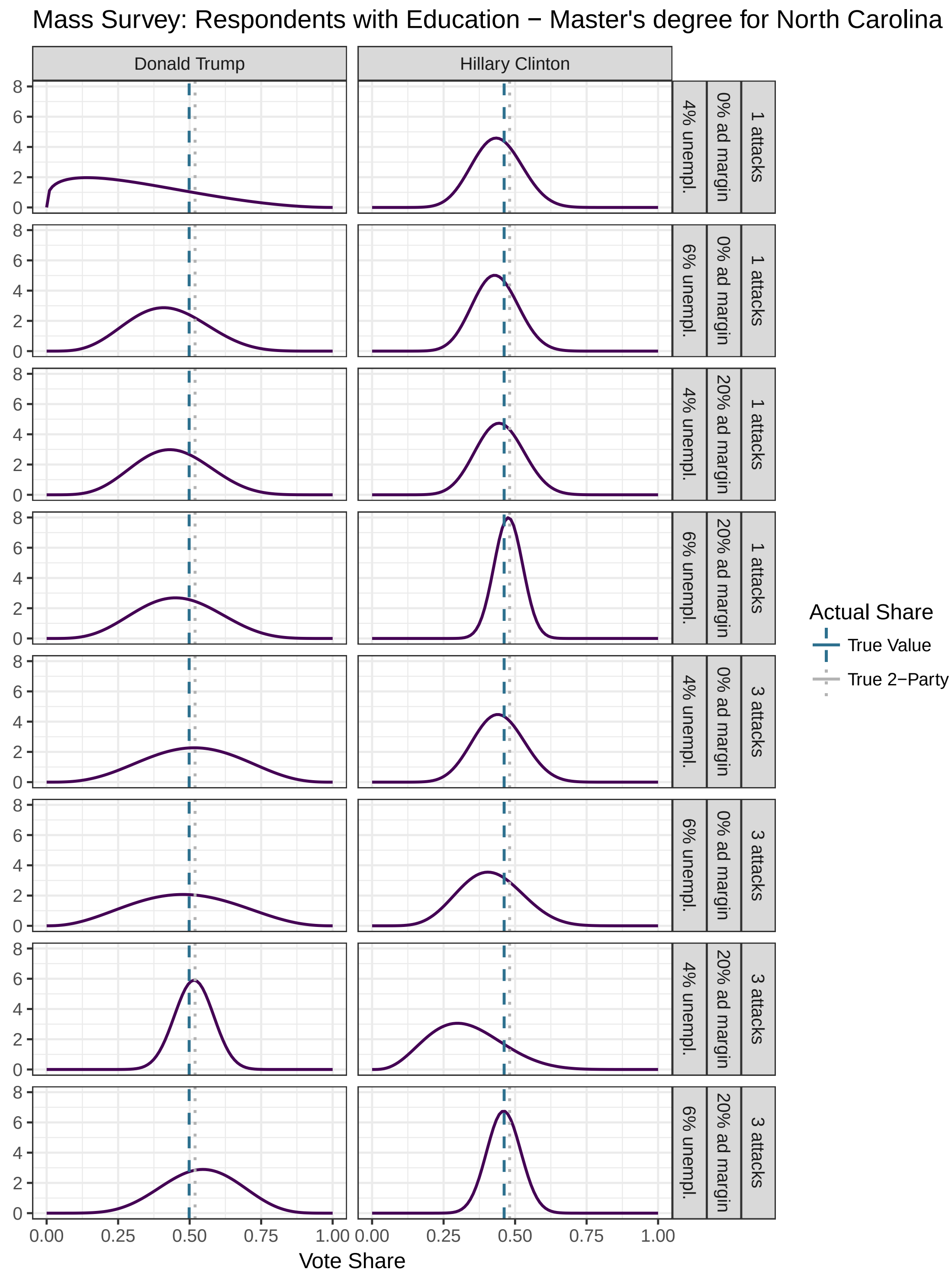


Figure 3.290: Priors with covariates: Mass North Carolina Education Master's degree

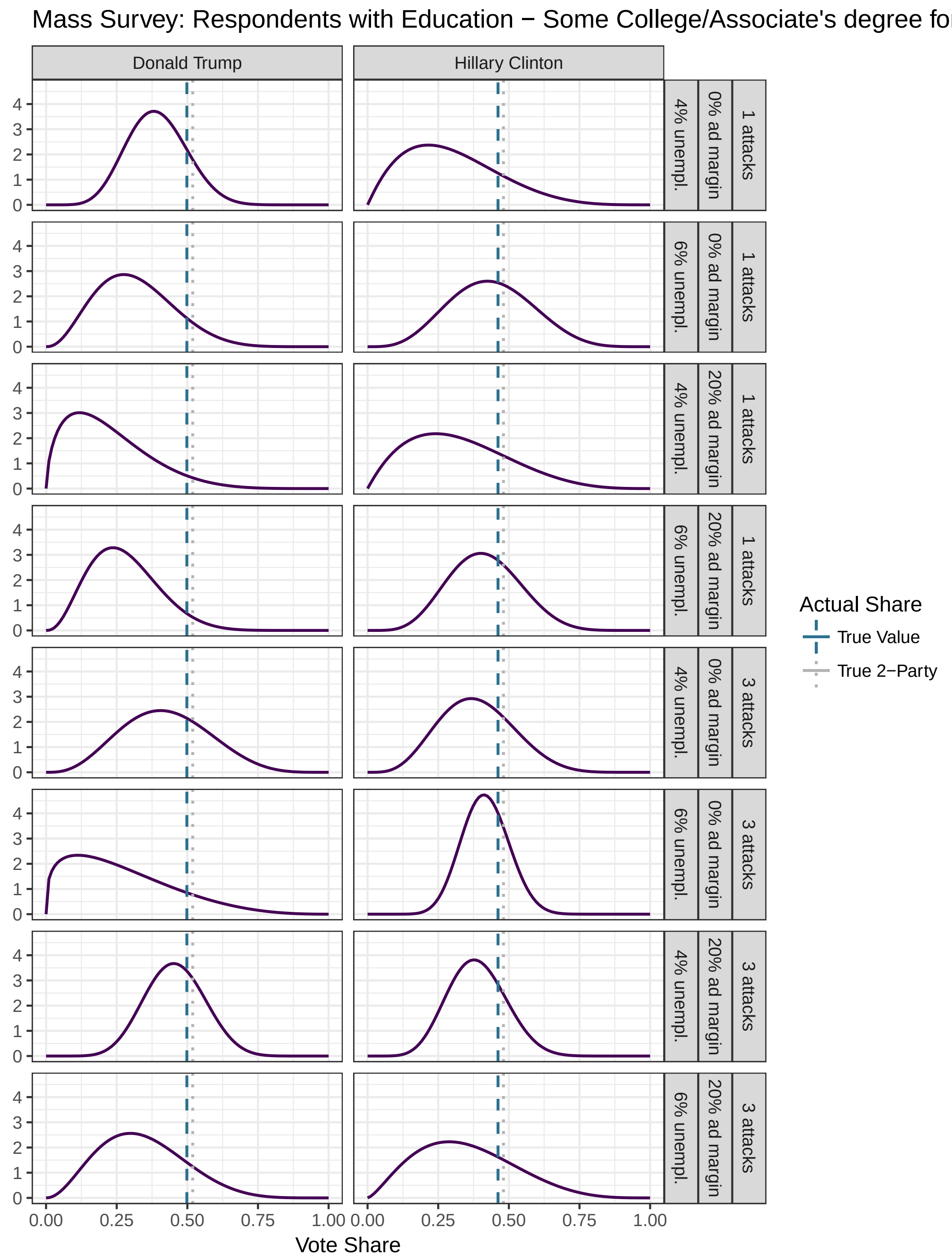


Figure 3.291: Priors with covariates: Mass North Carolina Education Some College Associate's degree

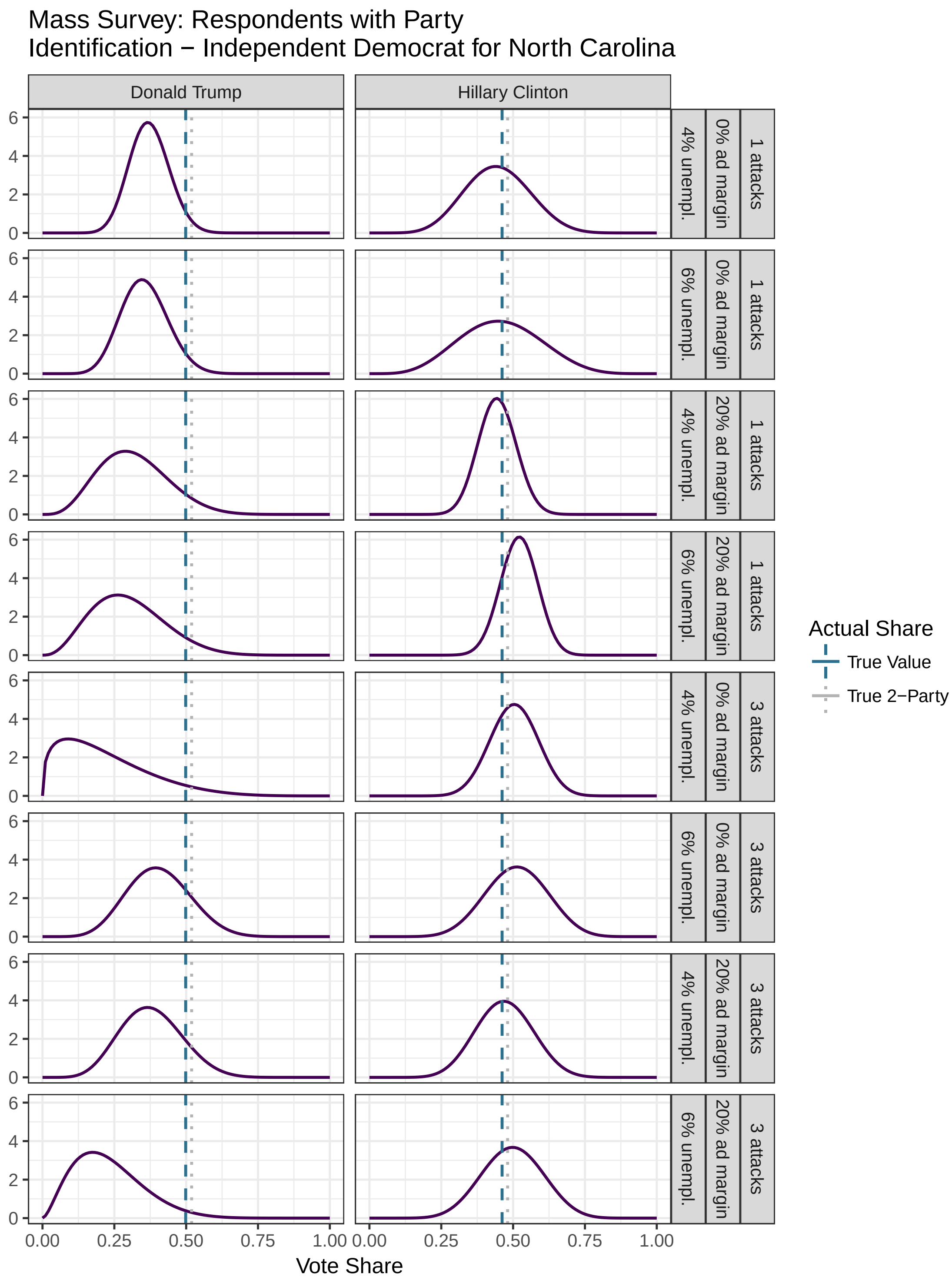


Figure 3.292: Priors with covariates: Mass North Carolina Party Identification Independent Democrat

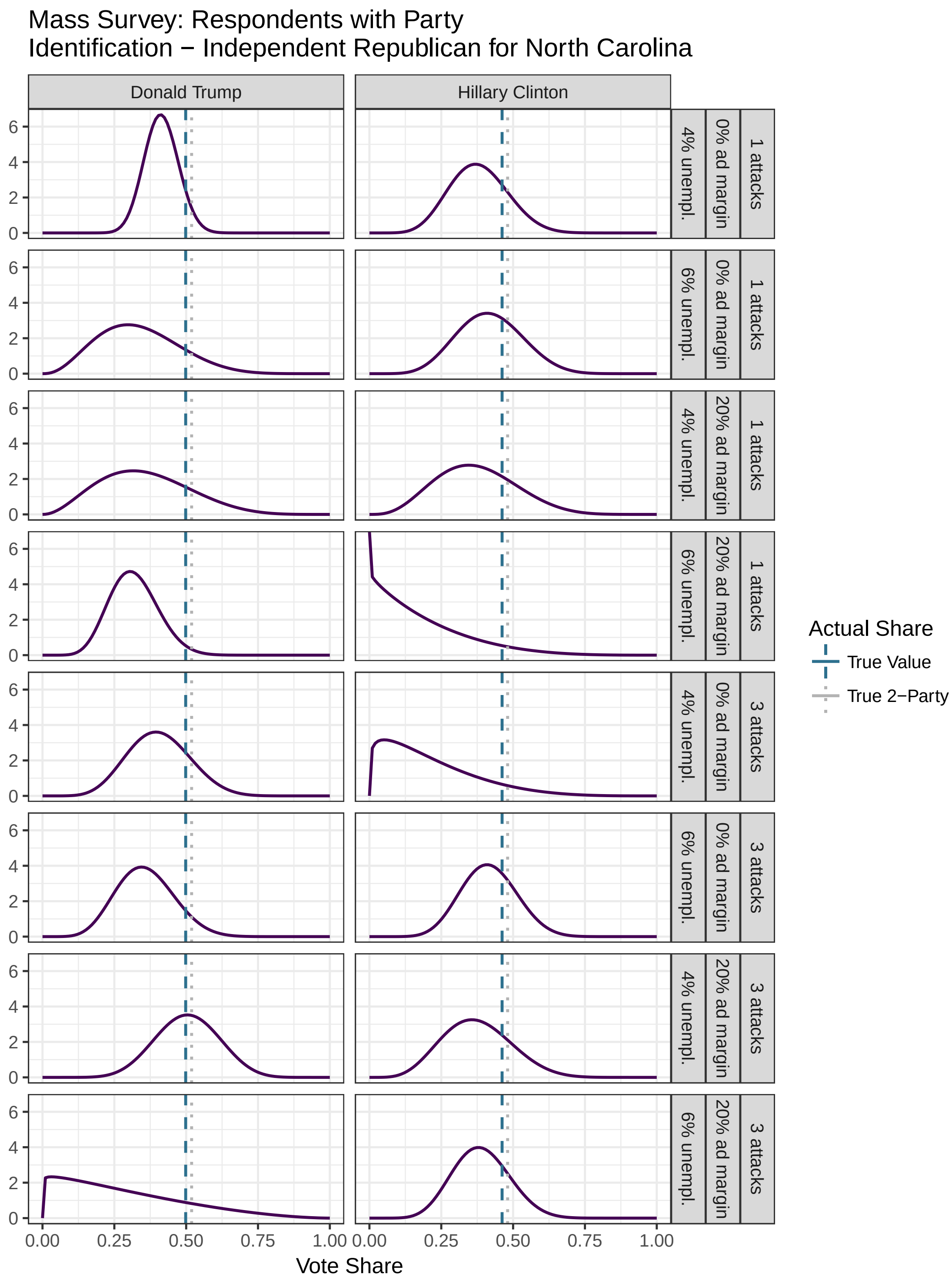


Figure 3.293: Priors with covariates: Mass North Carolina Party Identification Independent Republican

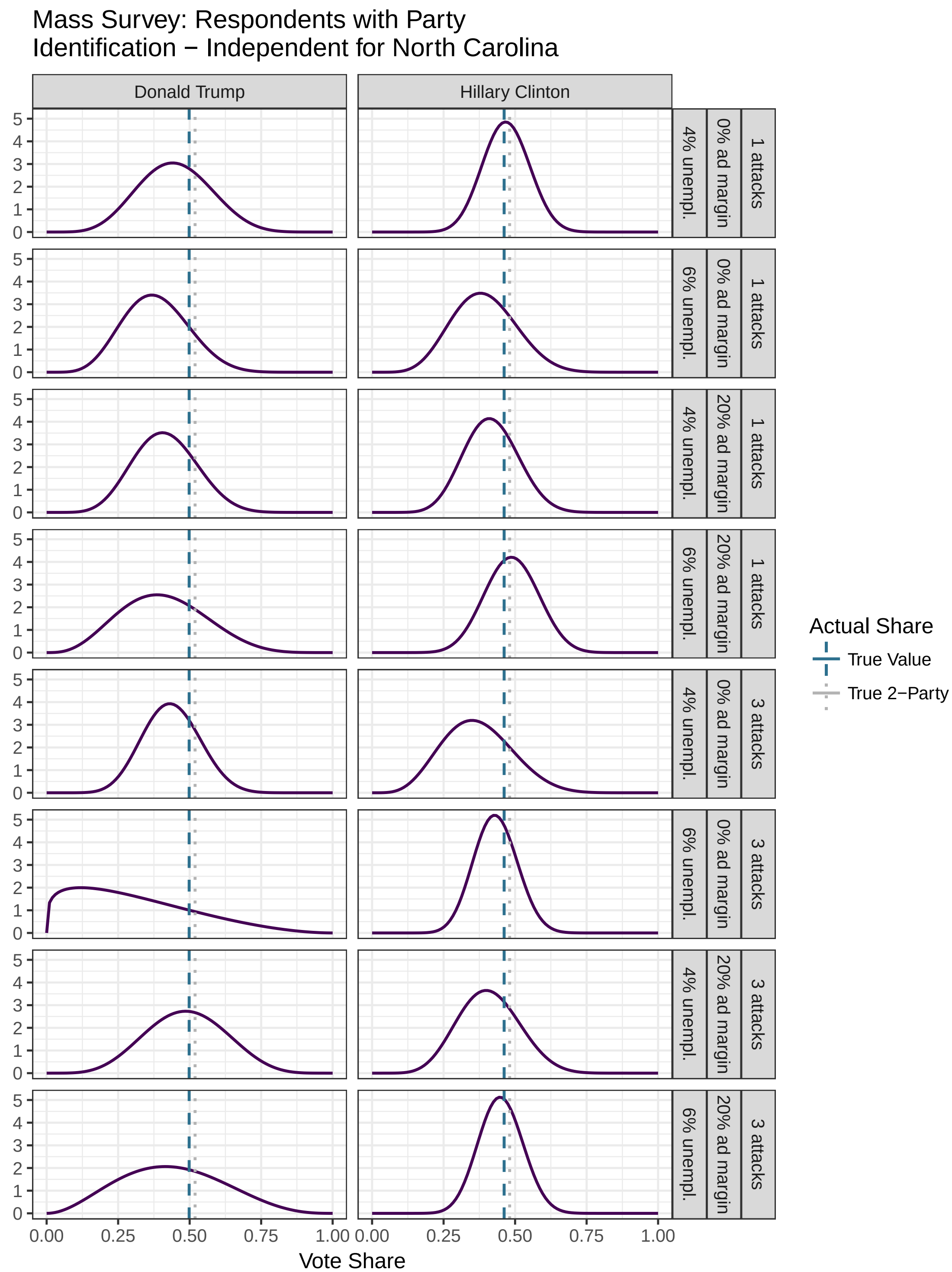


Figure 3.294: Priors with covariates: Mass North Carolina Party Identification Independent

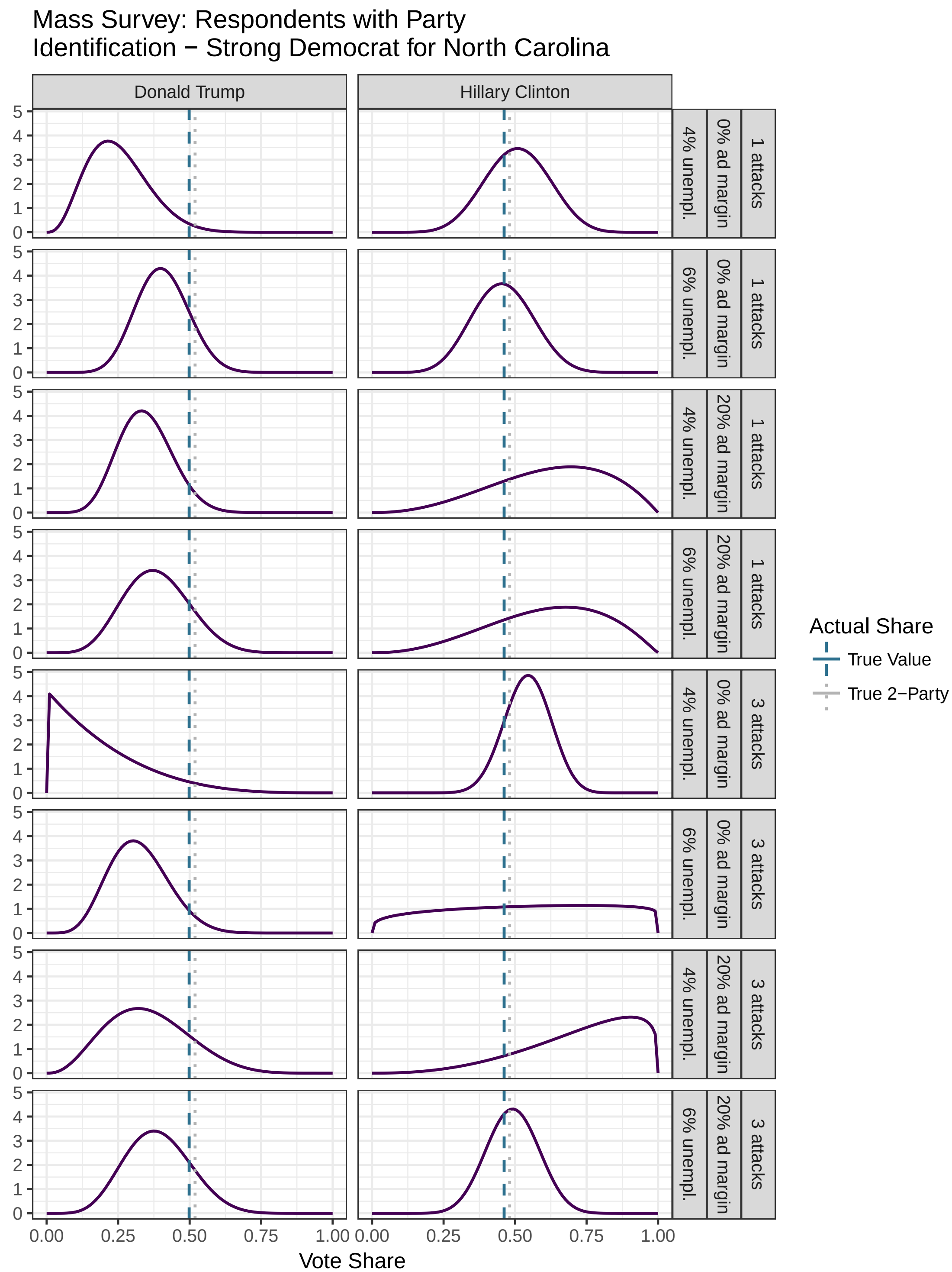


Figure 3.295: Priors with covariates: Mass North Carolina Party Identification Strong Democrat

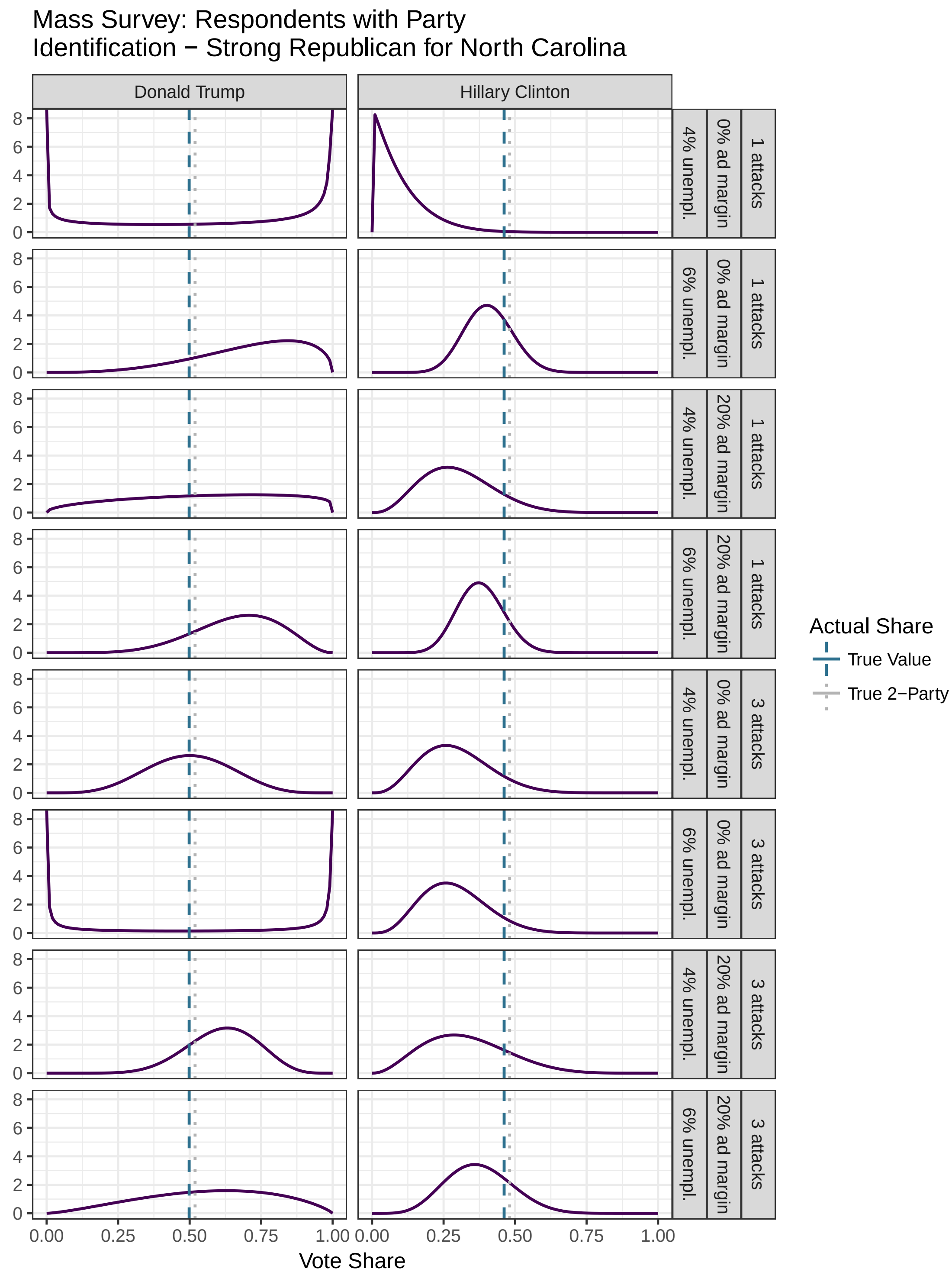


Figure 3.296: Priors with covariates: Mass North Carolina Party Identification Strong Republican

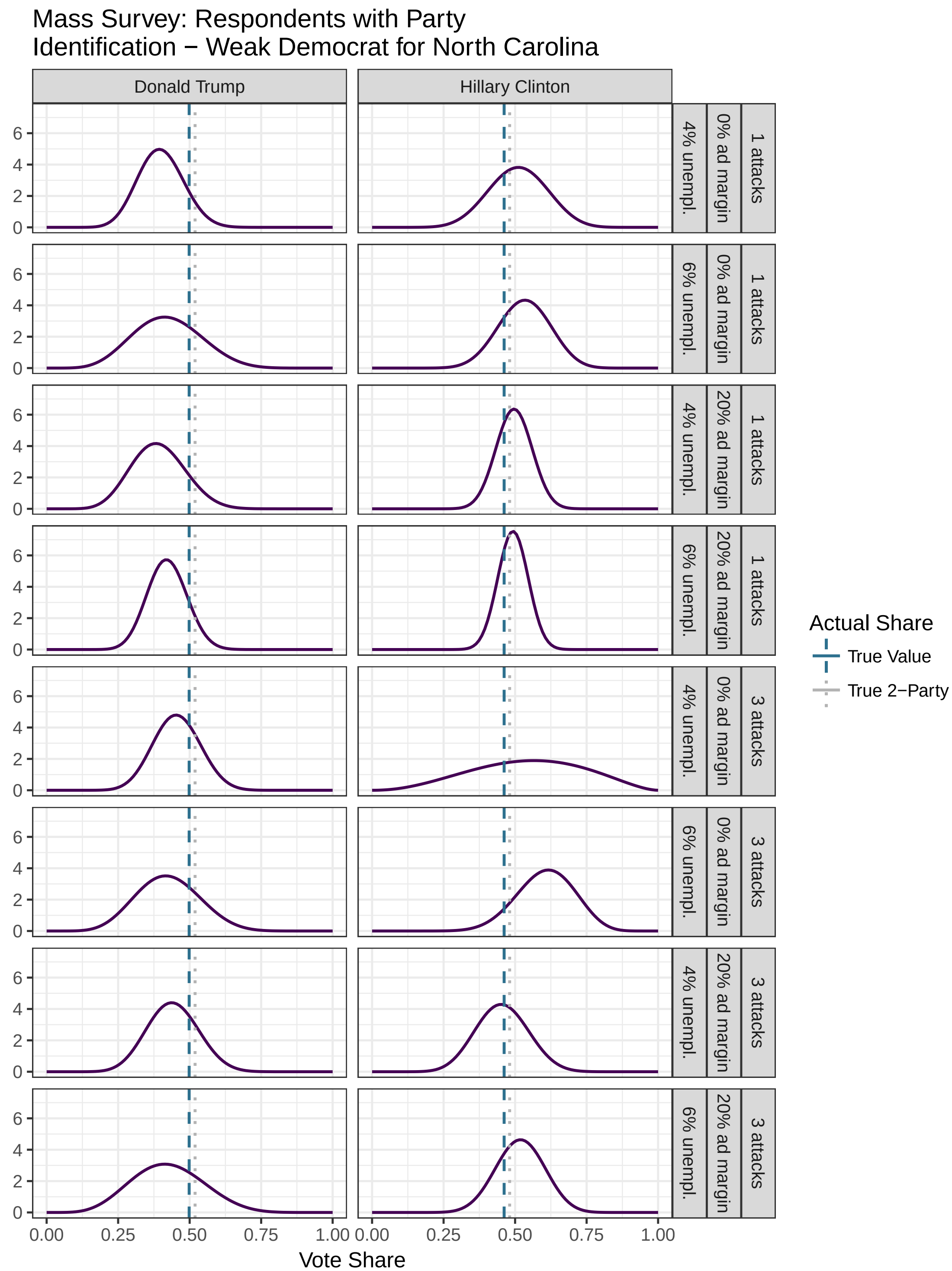


Figure 3.297: Priors with covariates: Mass North Carolina Party Identification Weak Democrat

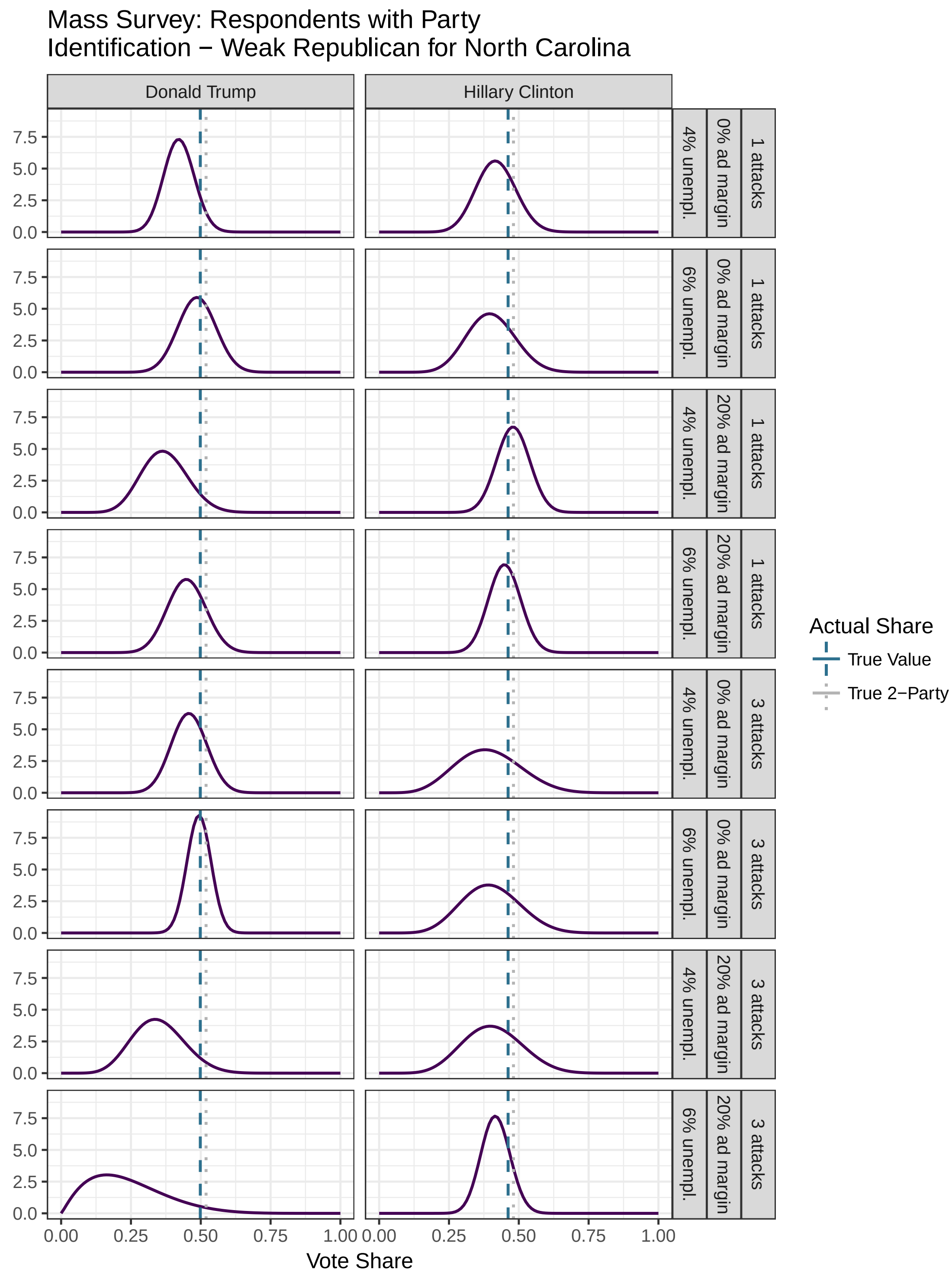


Figure 3.298: Priors with covariates: Mass North Carolina Party Identification Weak Republican

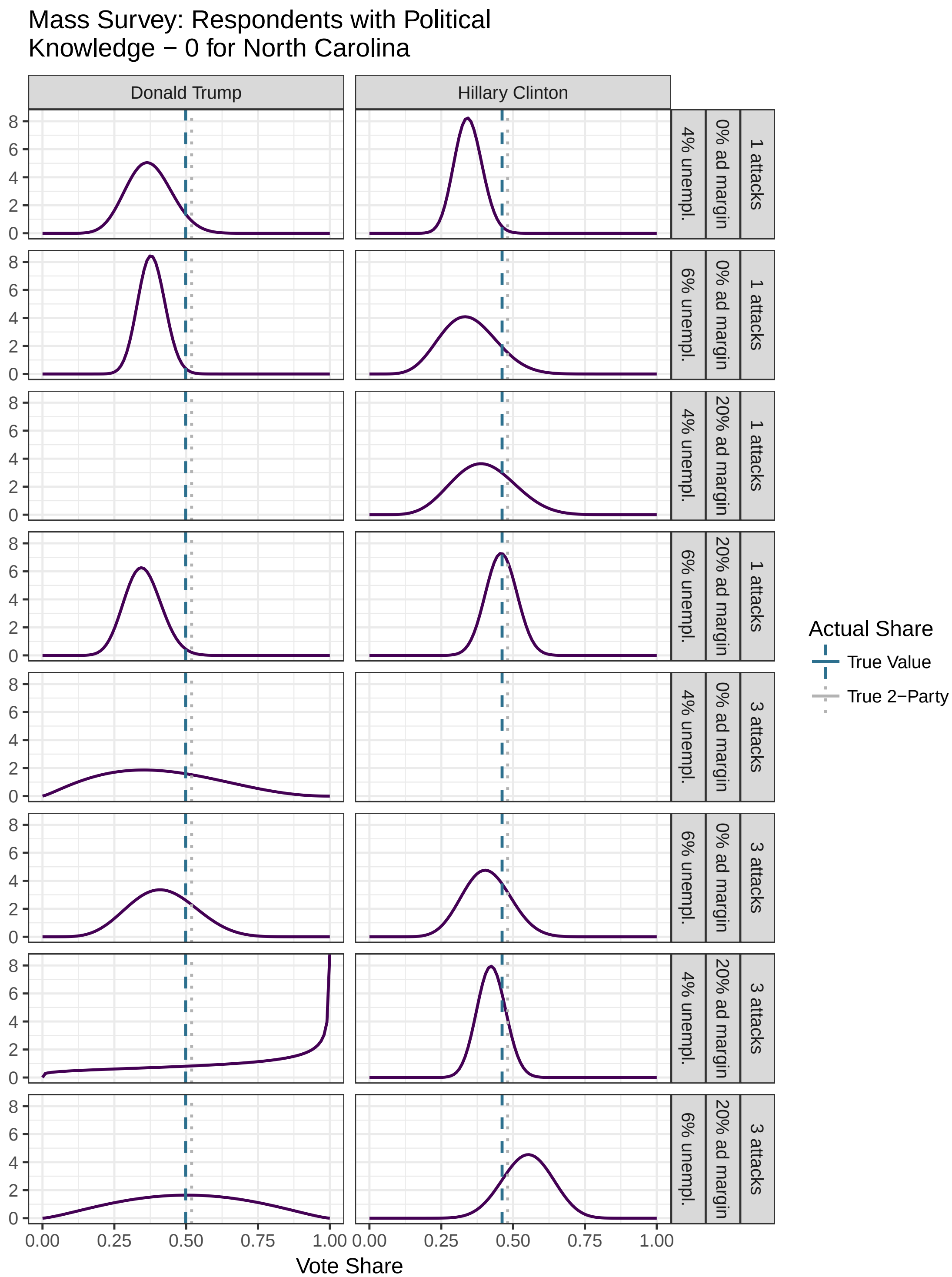


Figure 3.299: Priors with covariates: Mass North Carolina Political Knowledge 0

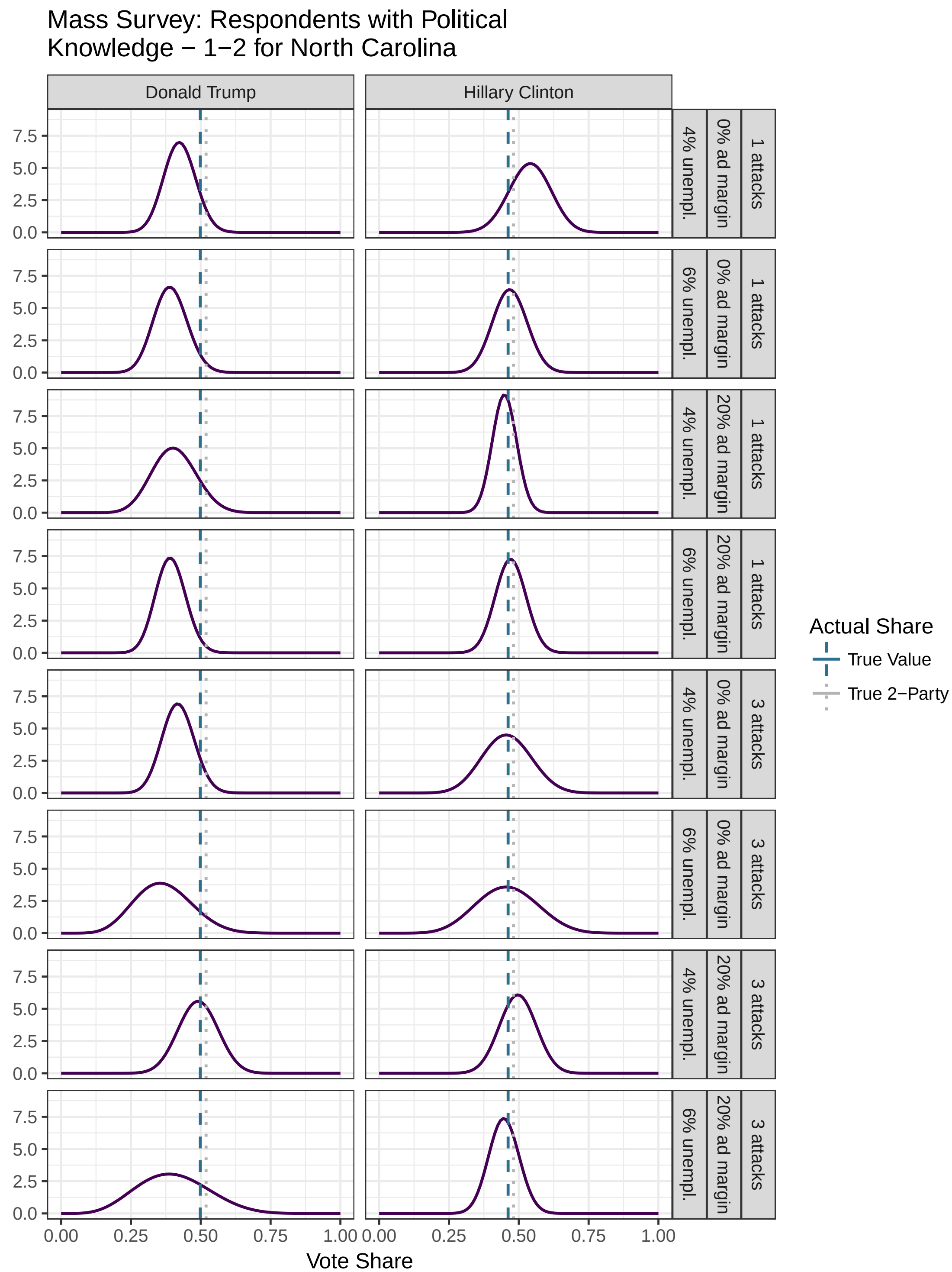


Figure 3.300: Priors with covariates: Mass North Carolina Political Knowledge 1-2

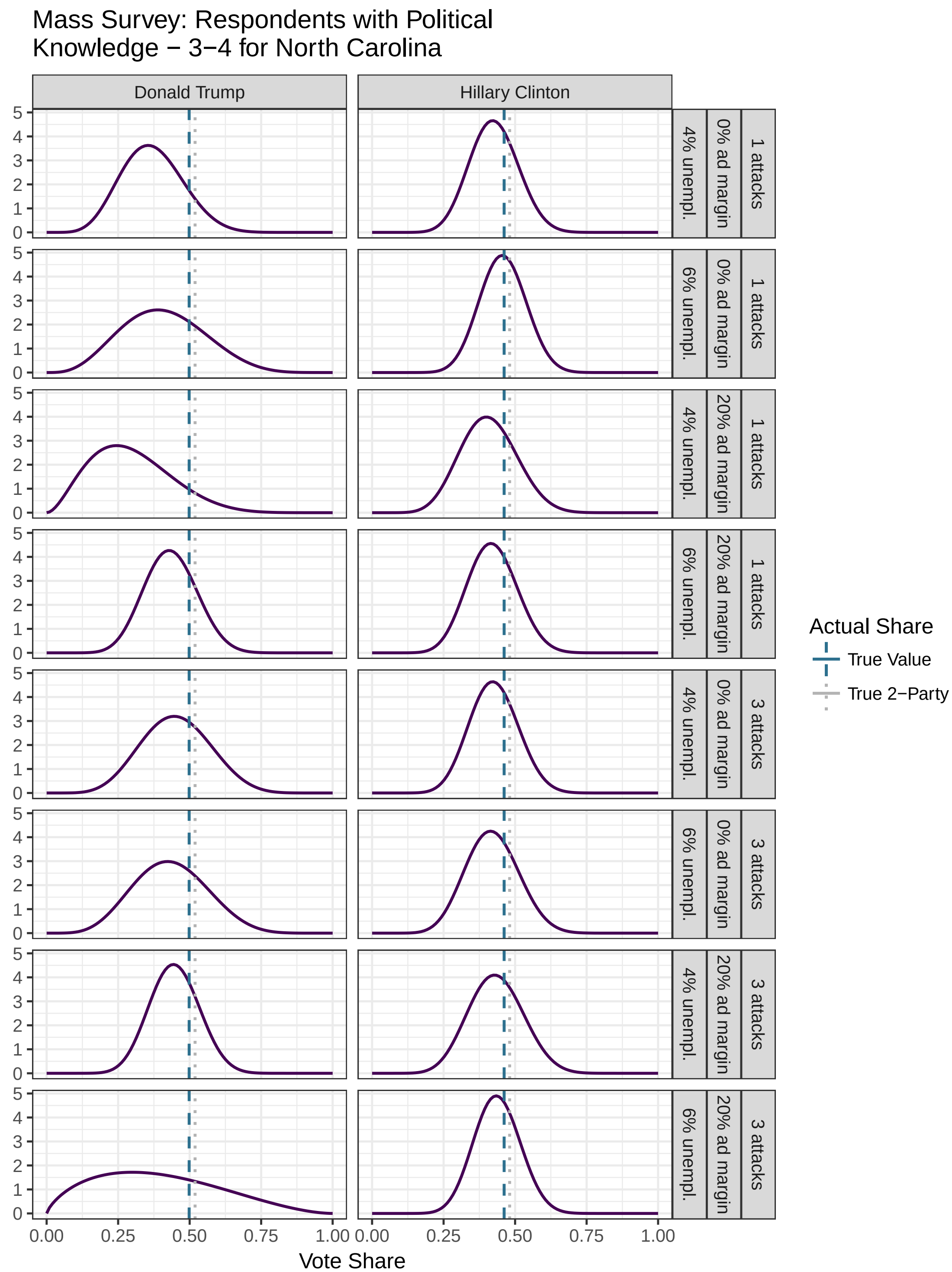


Figure 3.301: Priors with covariates: Mass North Carolina Political Knowledge 3-4

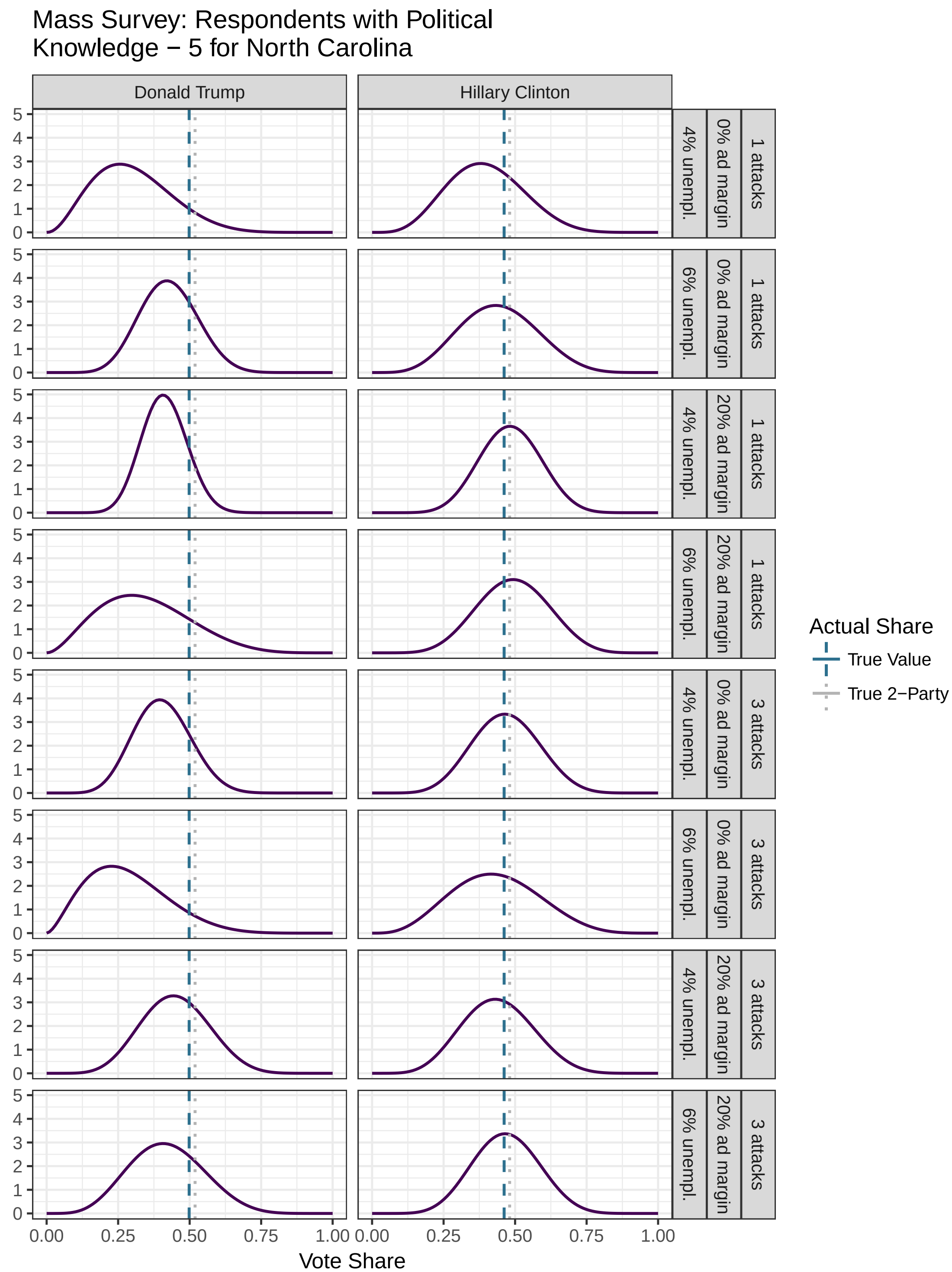


Figure 3.302: Priors with covariates: Mass North Carolina Political Knowledge 5

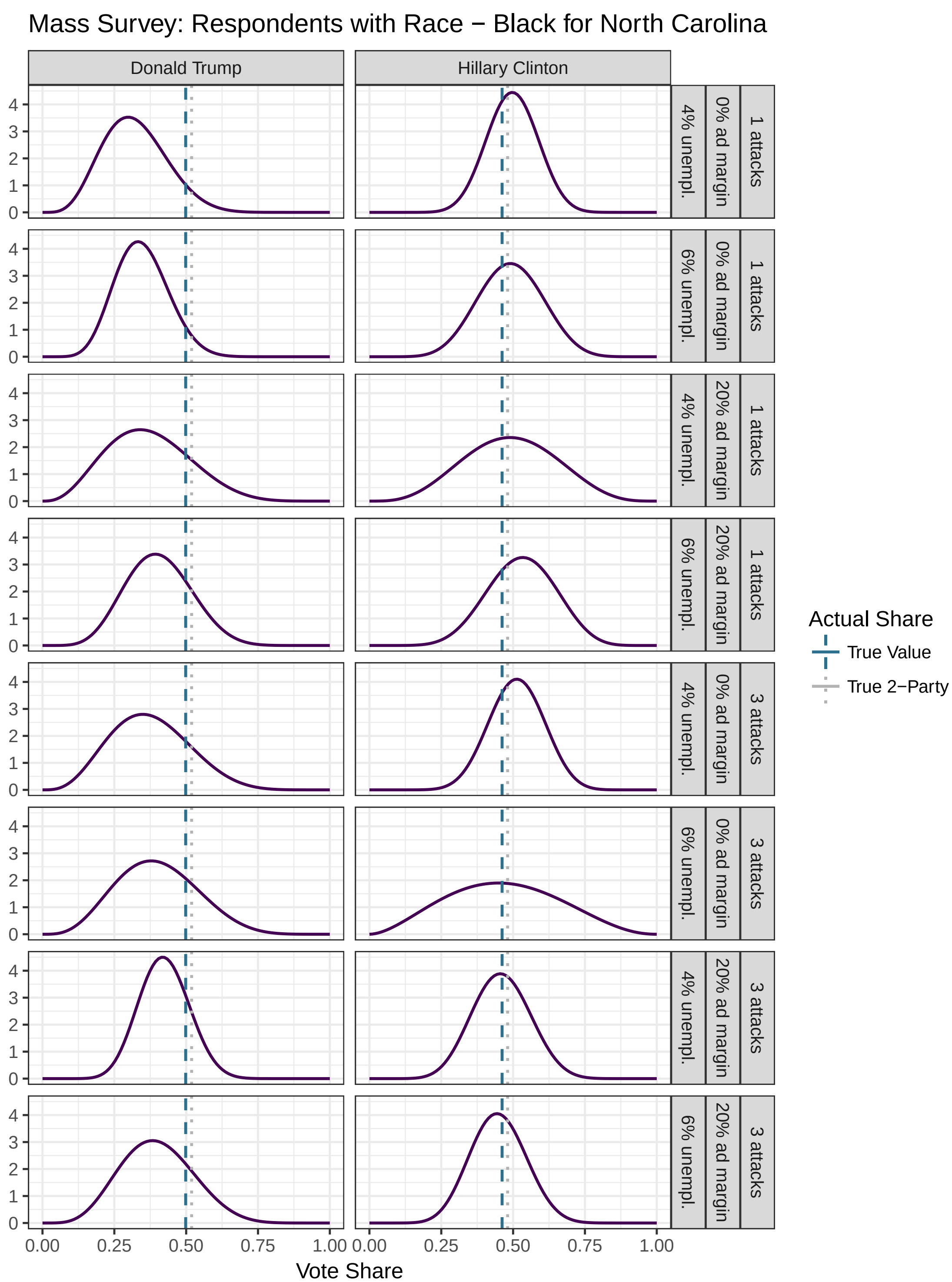


Figure 3.303: Priors with covariates: Mass North Carolina Race Black

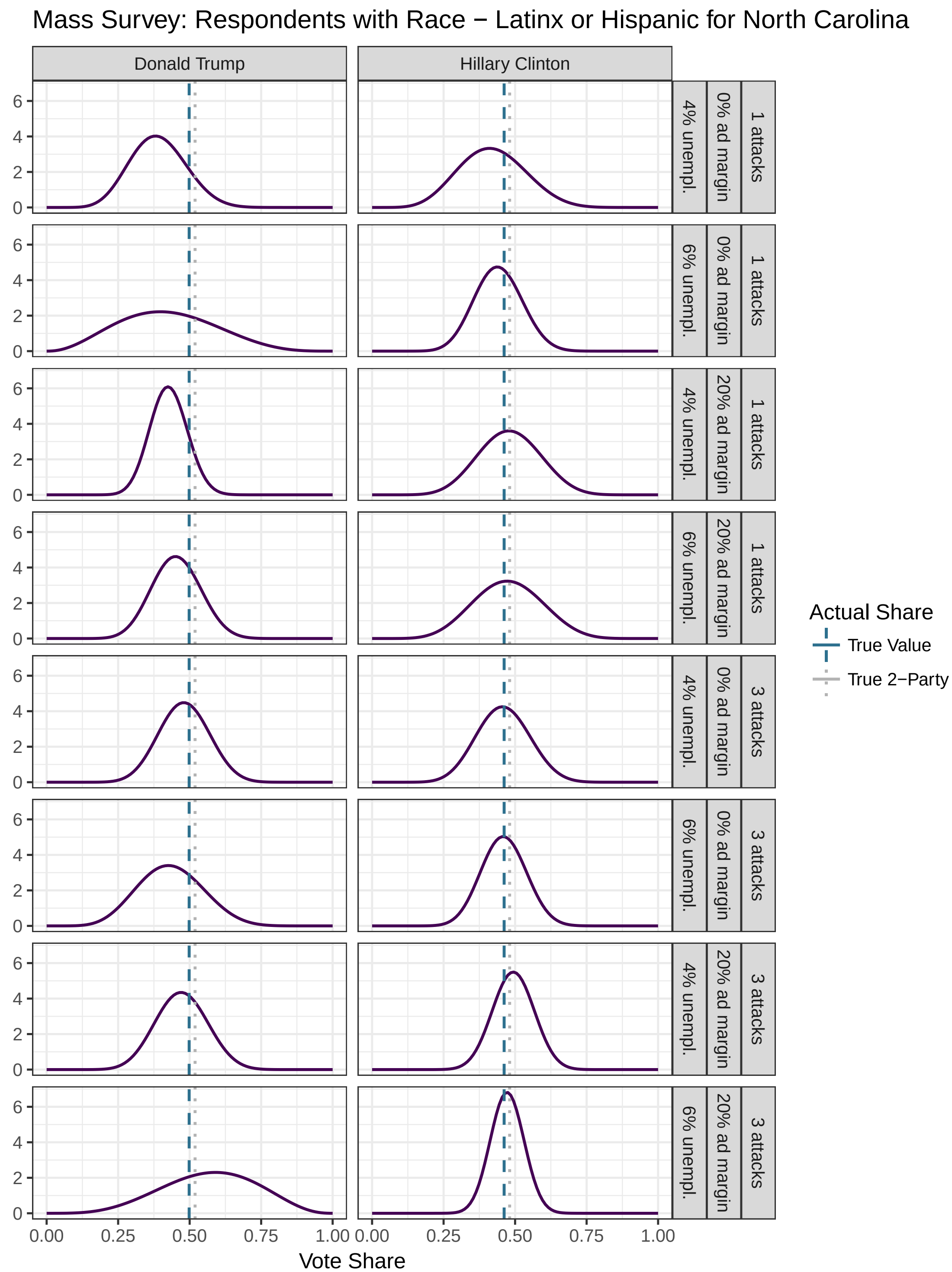


Figure 3.304: Priors with covariates: Mass North Carolina Race Latinx or Hispanic

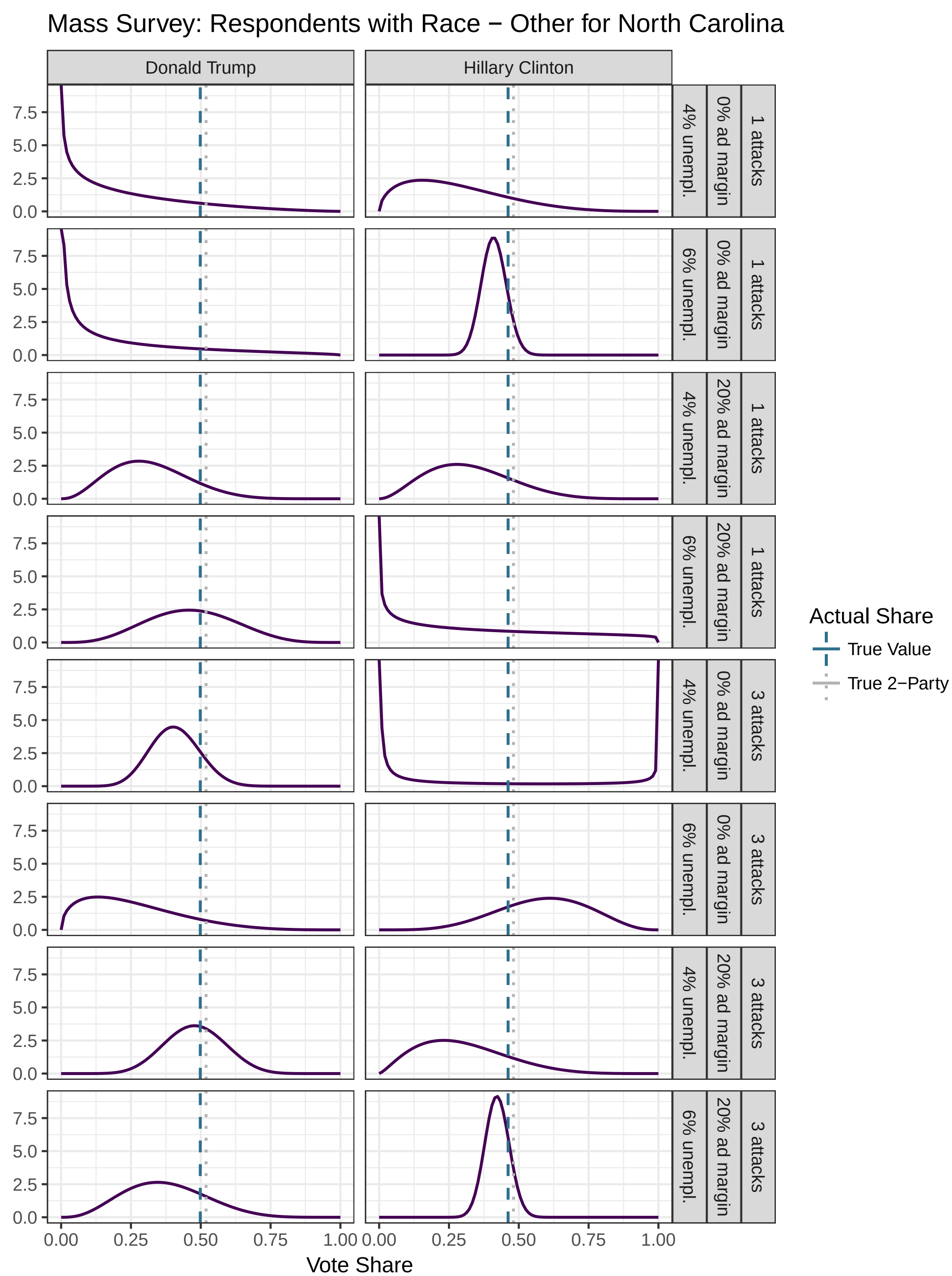


Figure 3.305: Priors with covariates: Mass North Carolina Race Other

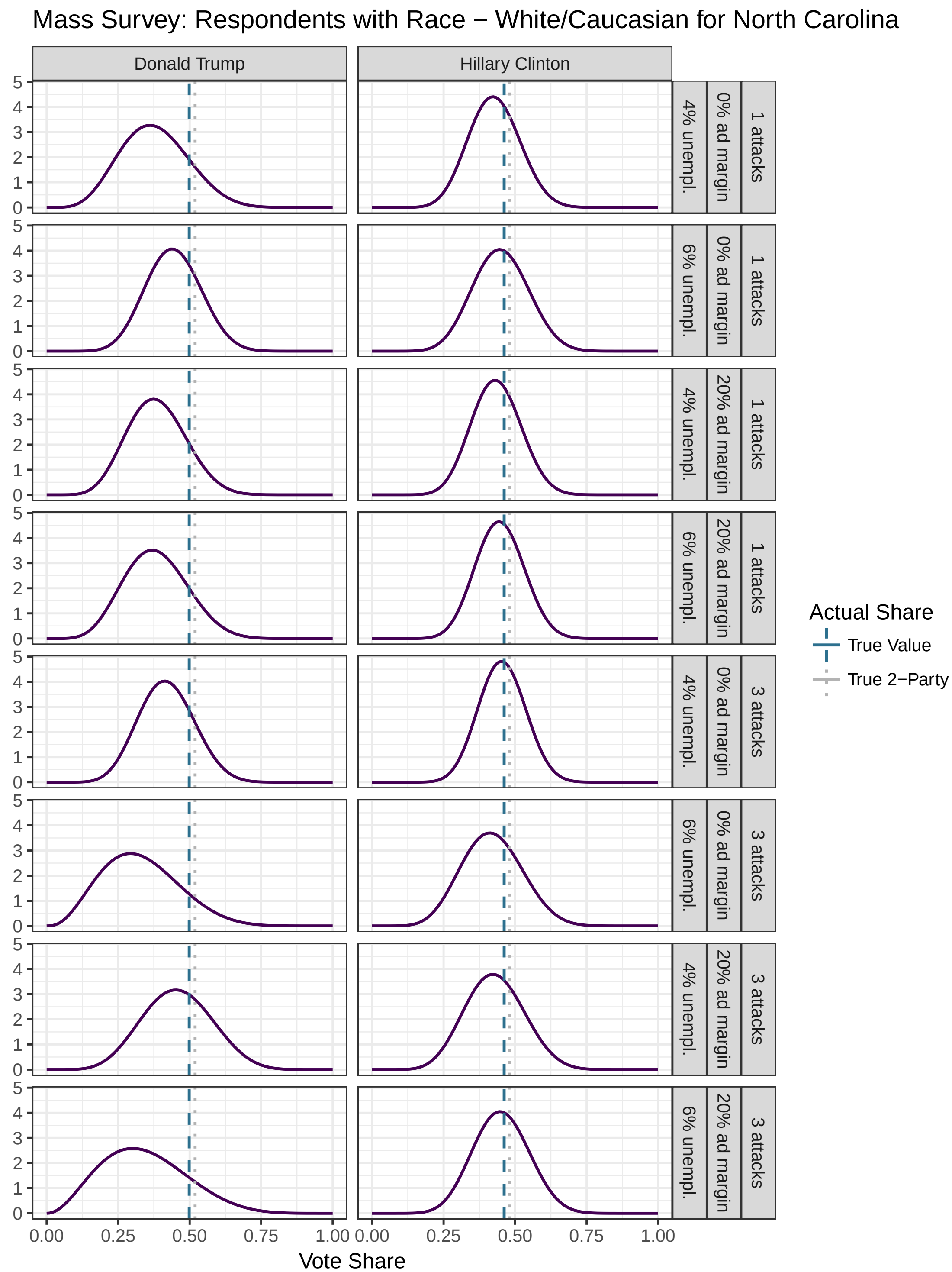


Figure 3.306: Priors with covariates: Mass North Carolina Race White Caucasian

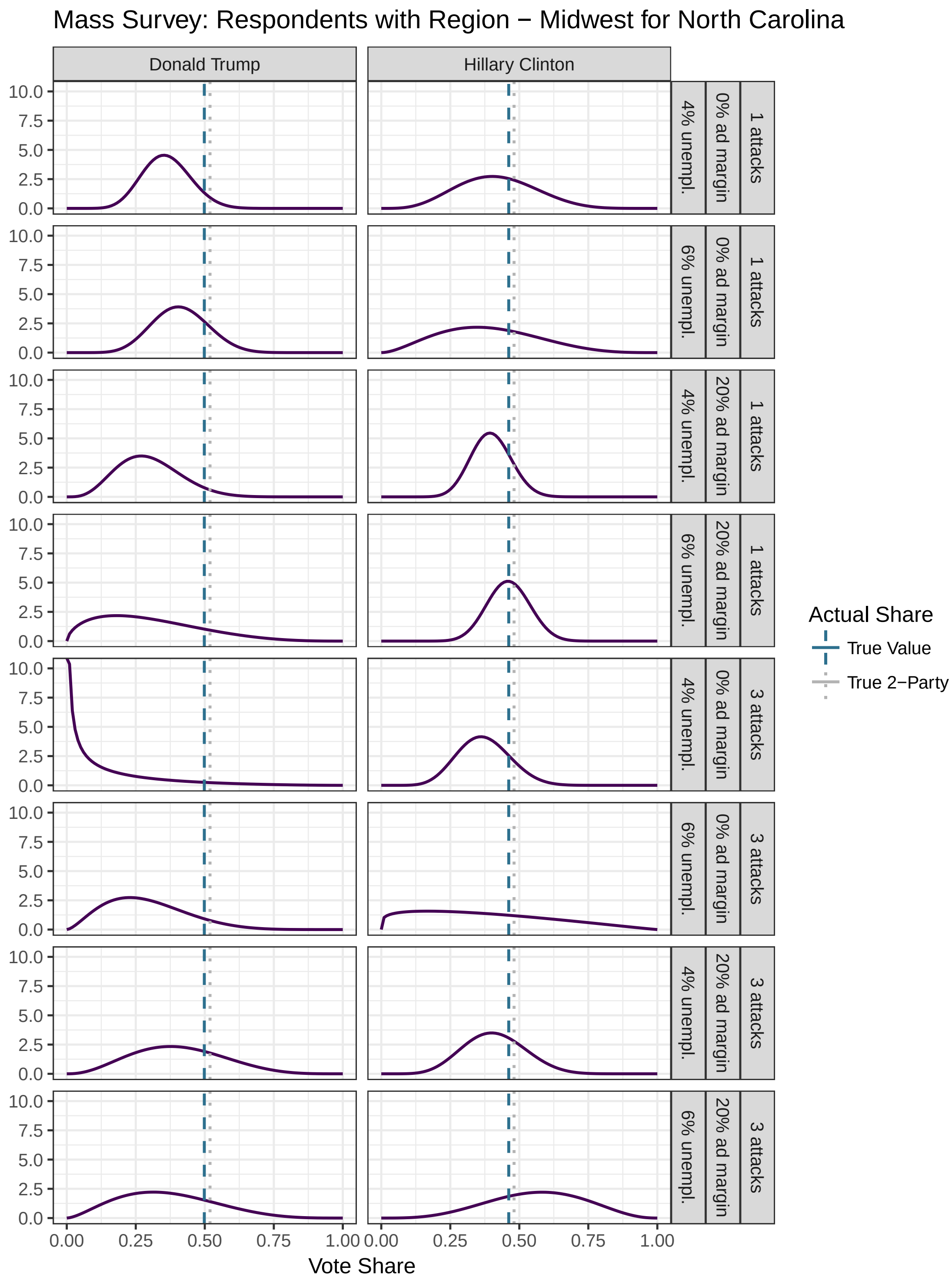


Figure 3.307: Priors with covariates: Mass North Carolina Region Midwest

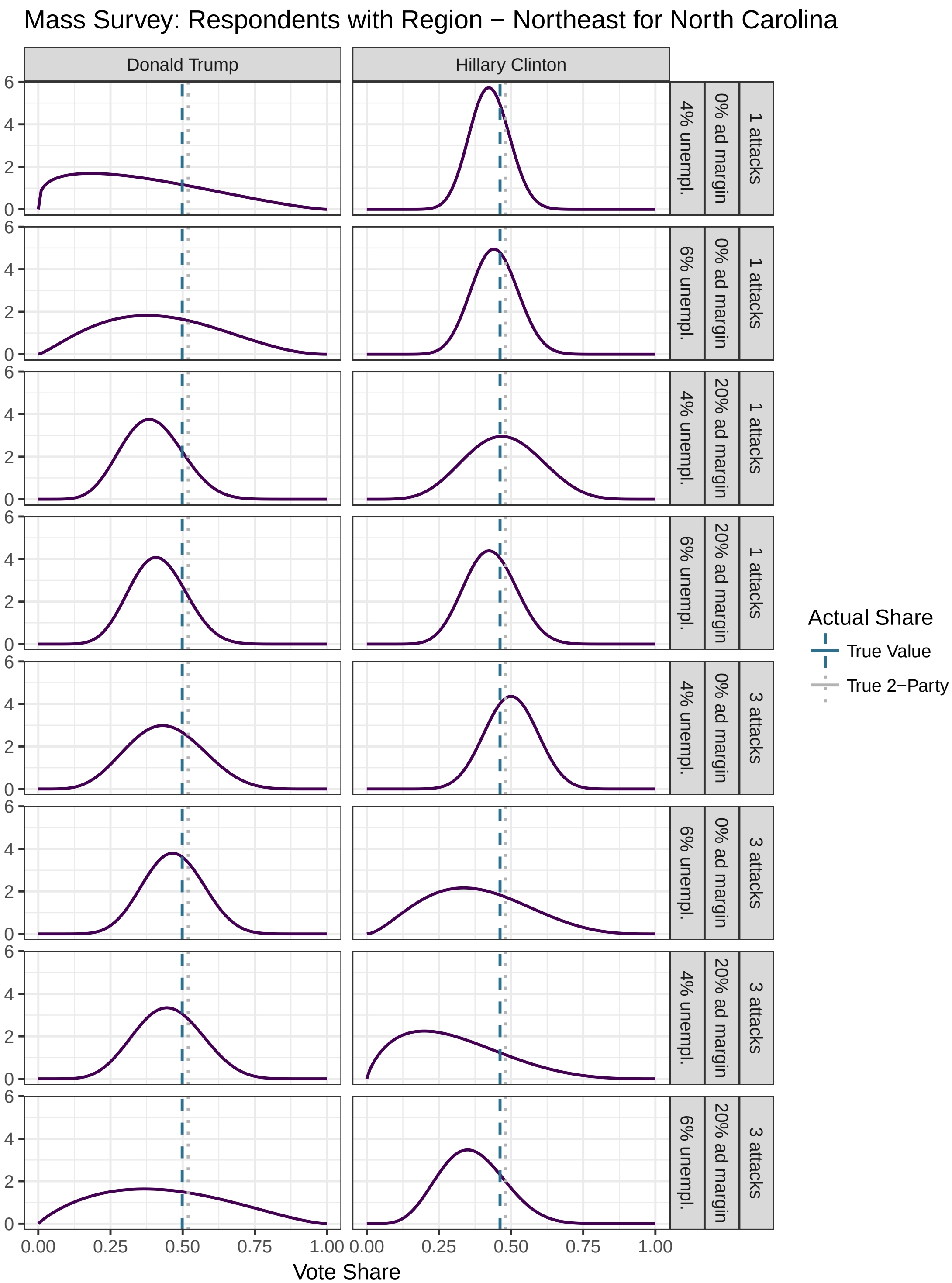


Figure 3.308: Priors with covariates: Mass North Carolina Region Northeast

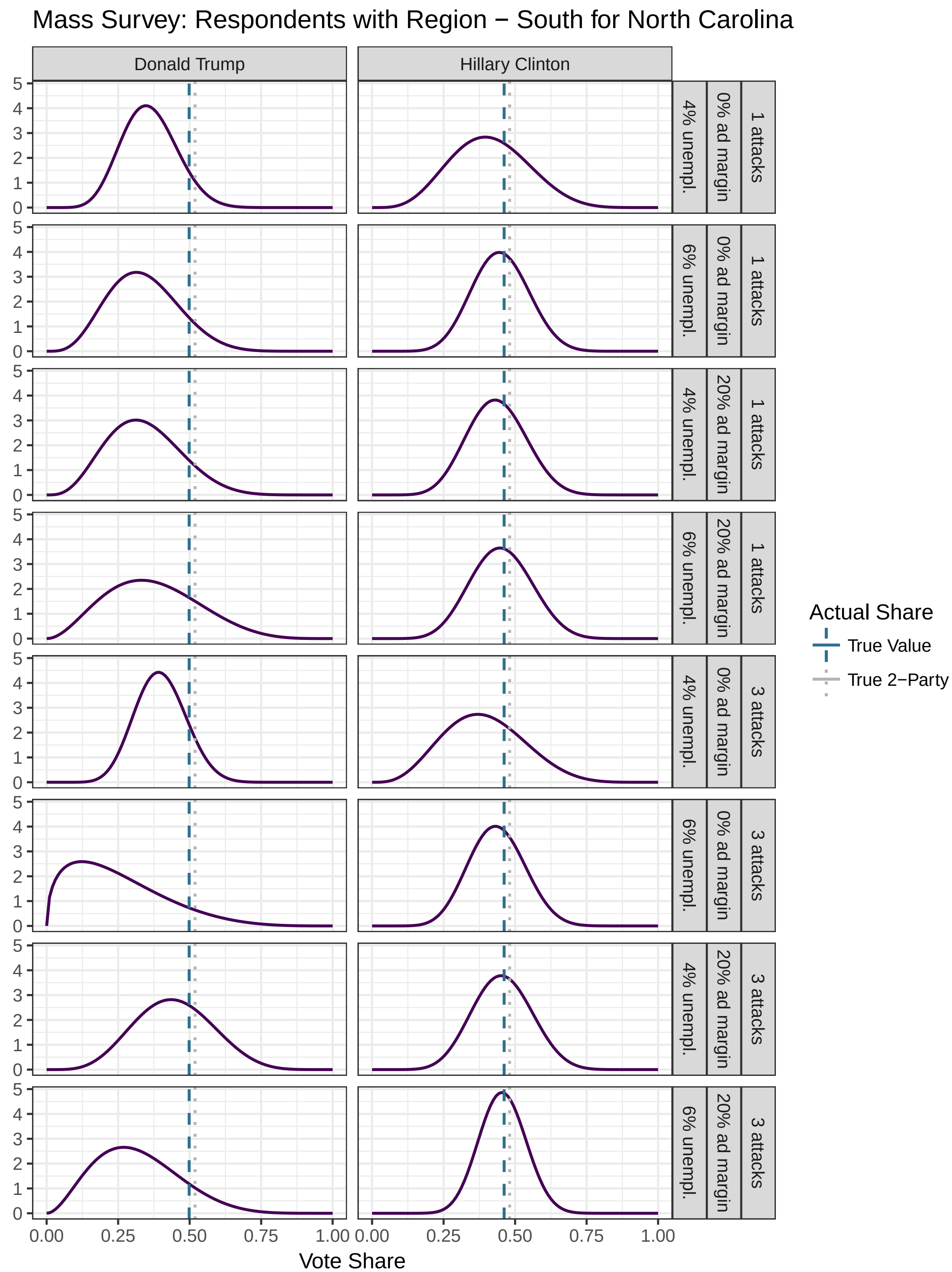


Figure 3.309: Priors with covariates: Mass North Carolina Region South

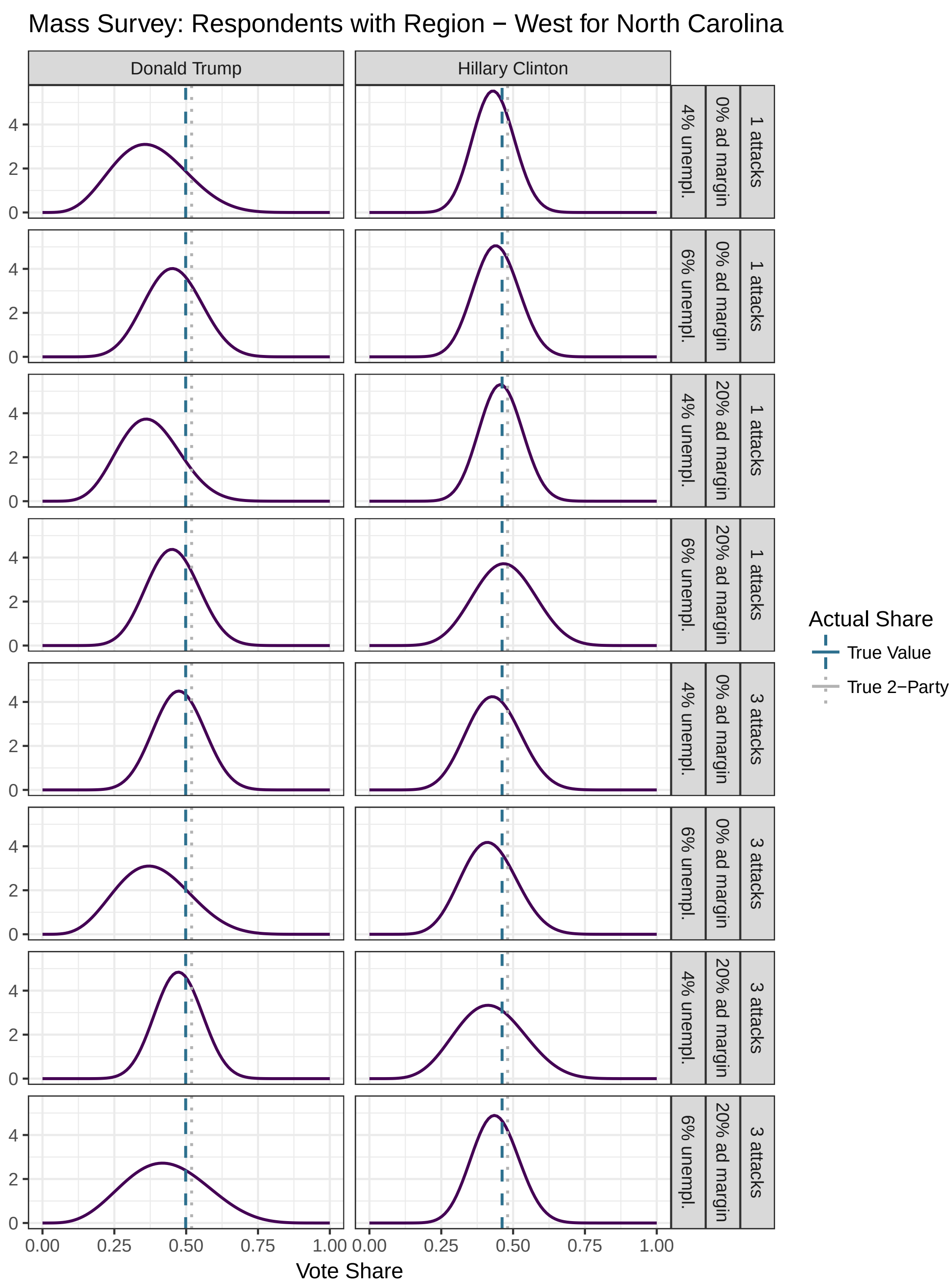


Figure 3.310: Priors with covariates: Mass North Carolina Region West

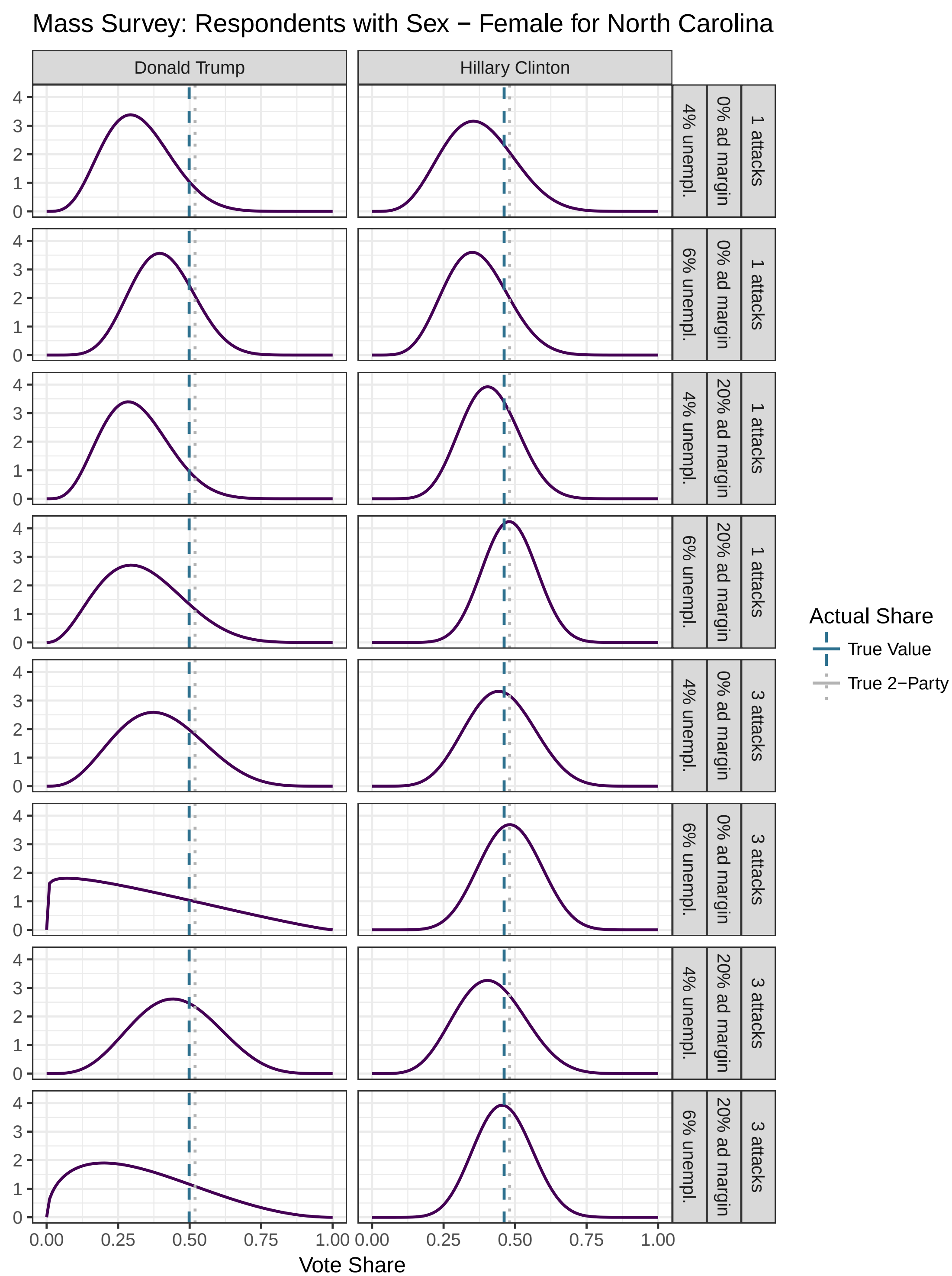


Figure 3.311: Priors with covariates: Mass North Carolina Sex Female

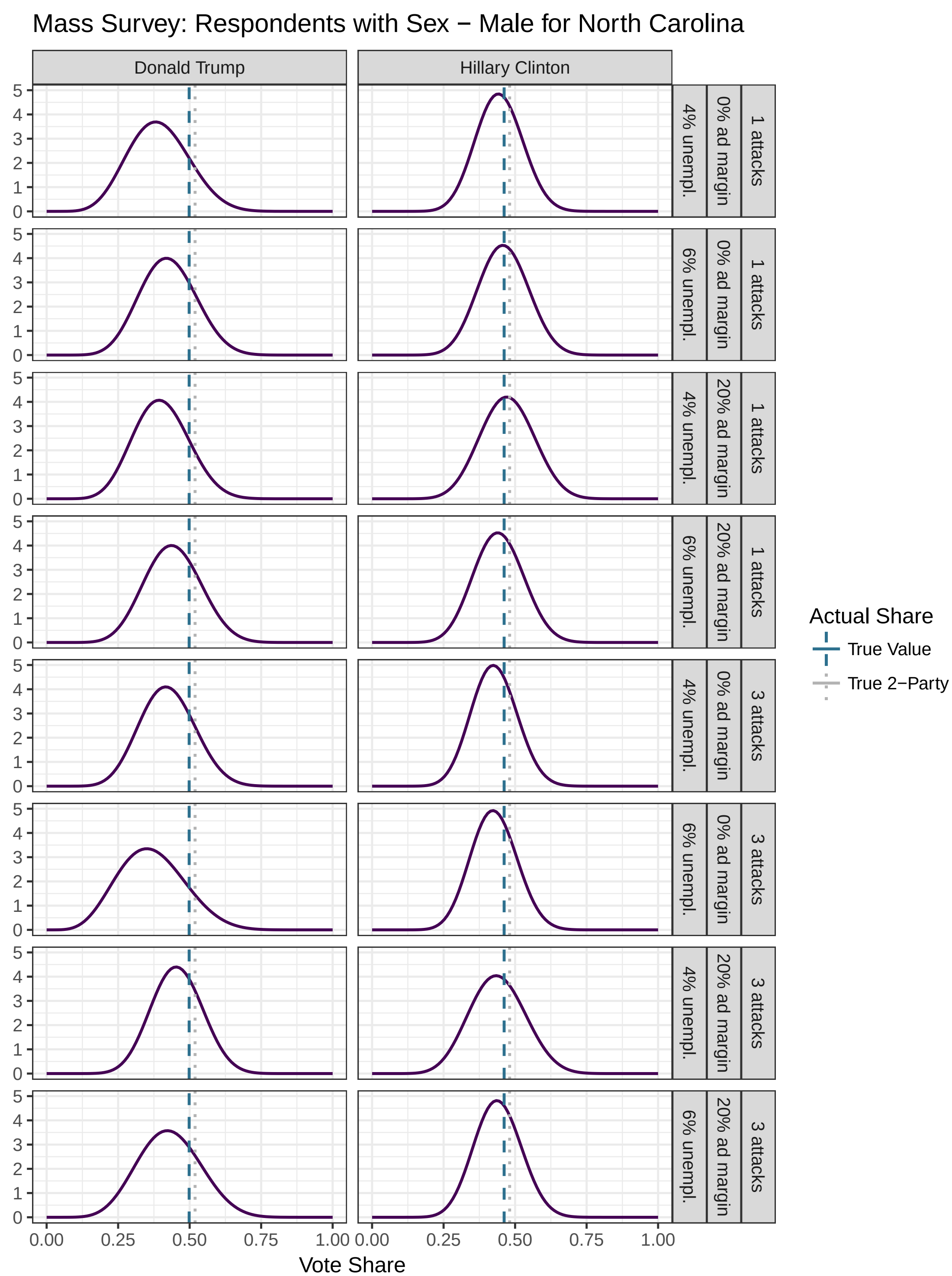


Figure 3.312: Priors with covariates: Mass North Carolina Sex Male

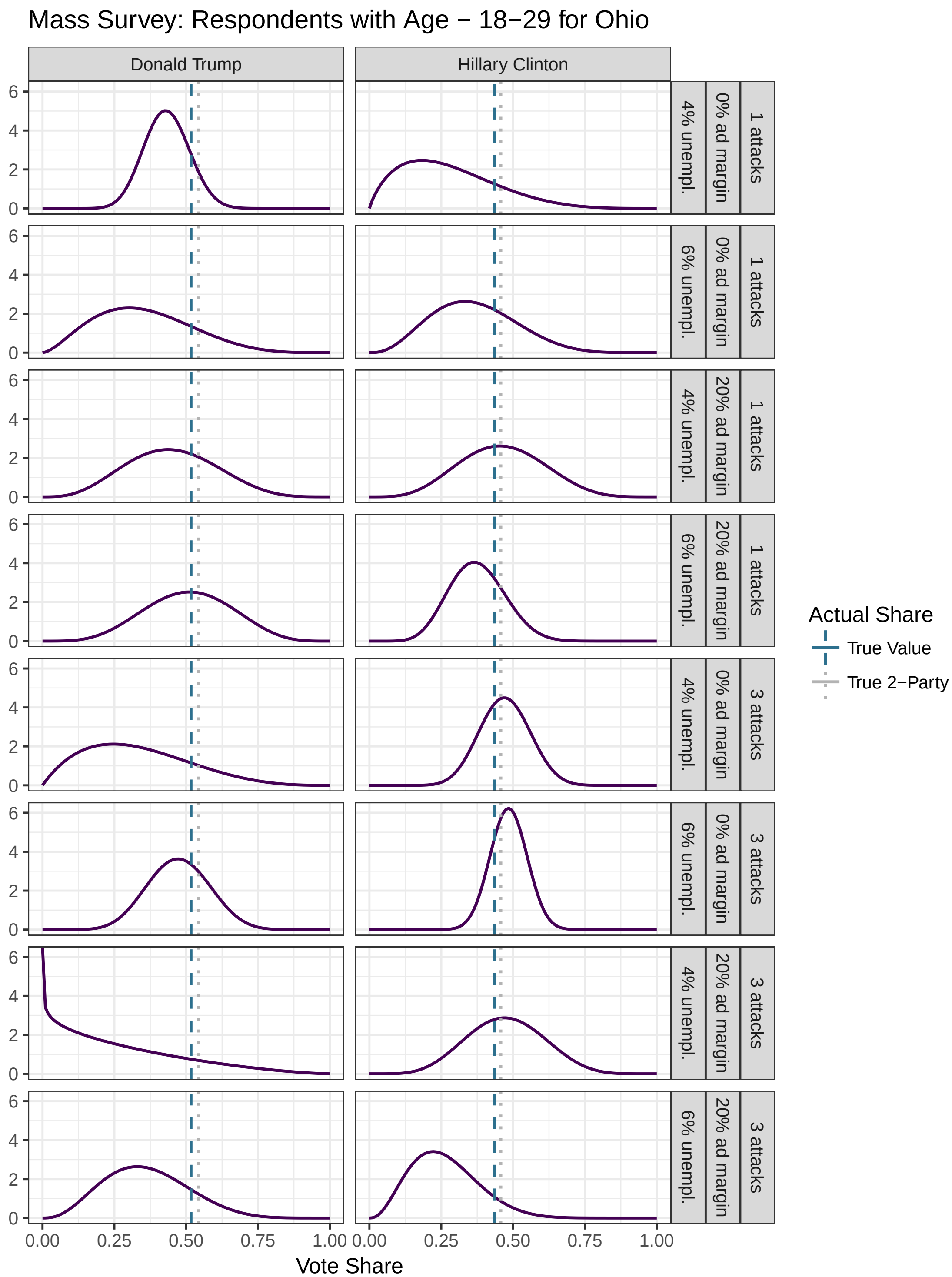


Figure 3.313: Priors with covariates: Mass Ohio Age 18-29

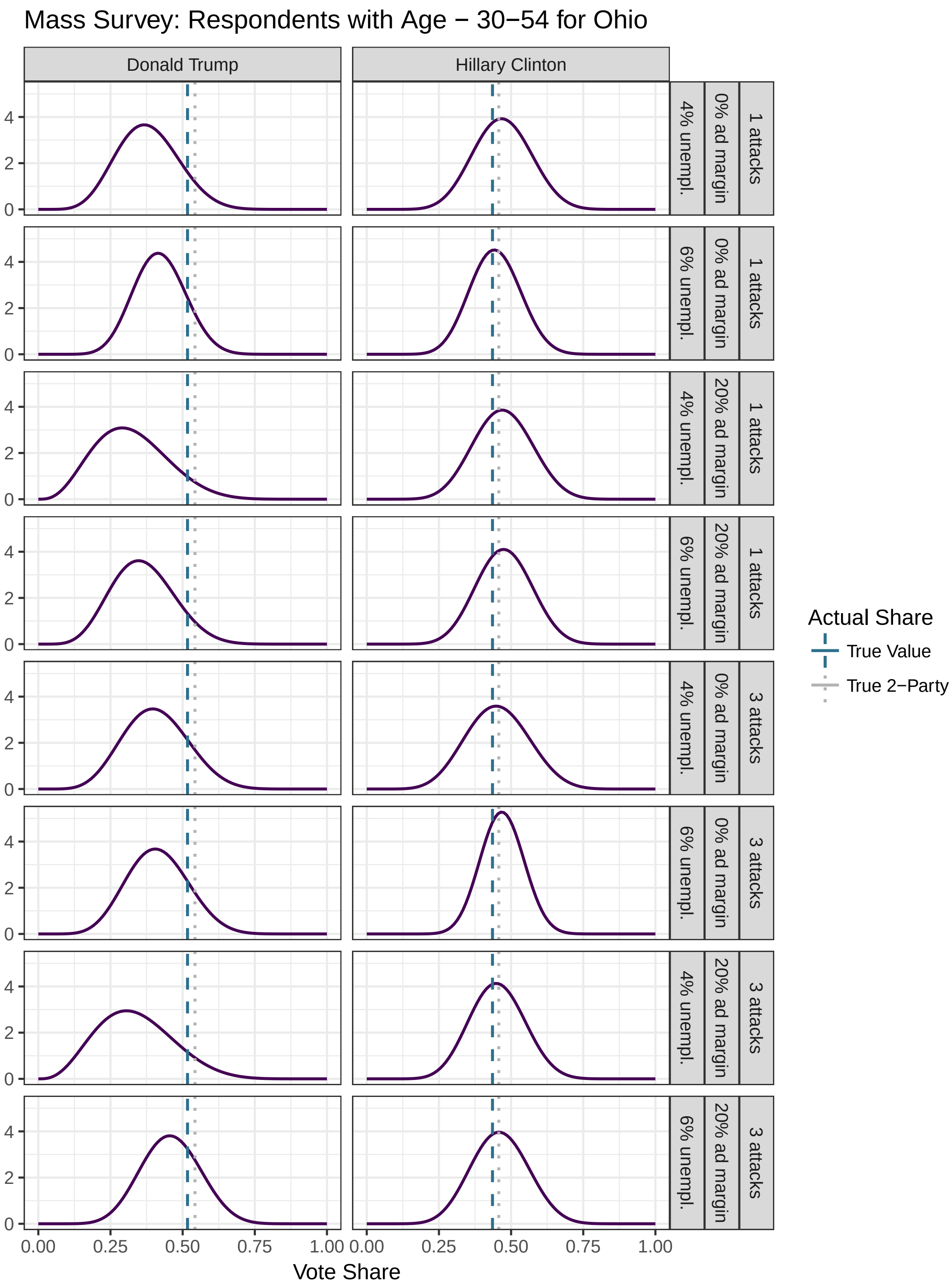


Figure 3.314: Priors with covariates: Mass Ohio Age 30-54

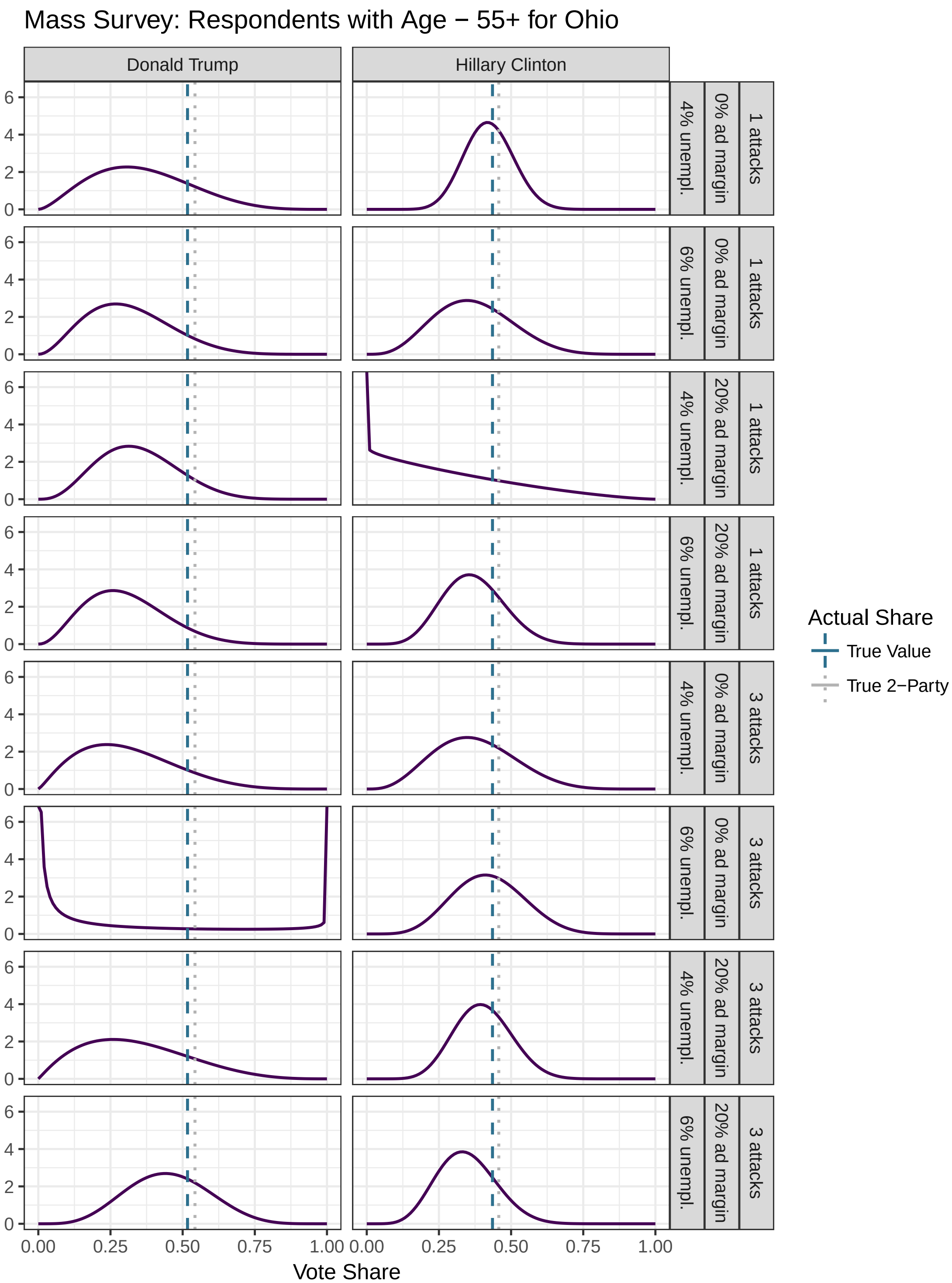


Figure 3.315: Priors with covariates: Mass Ohio Age 55+

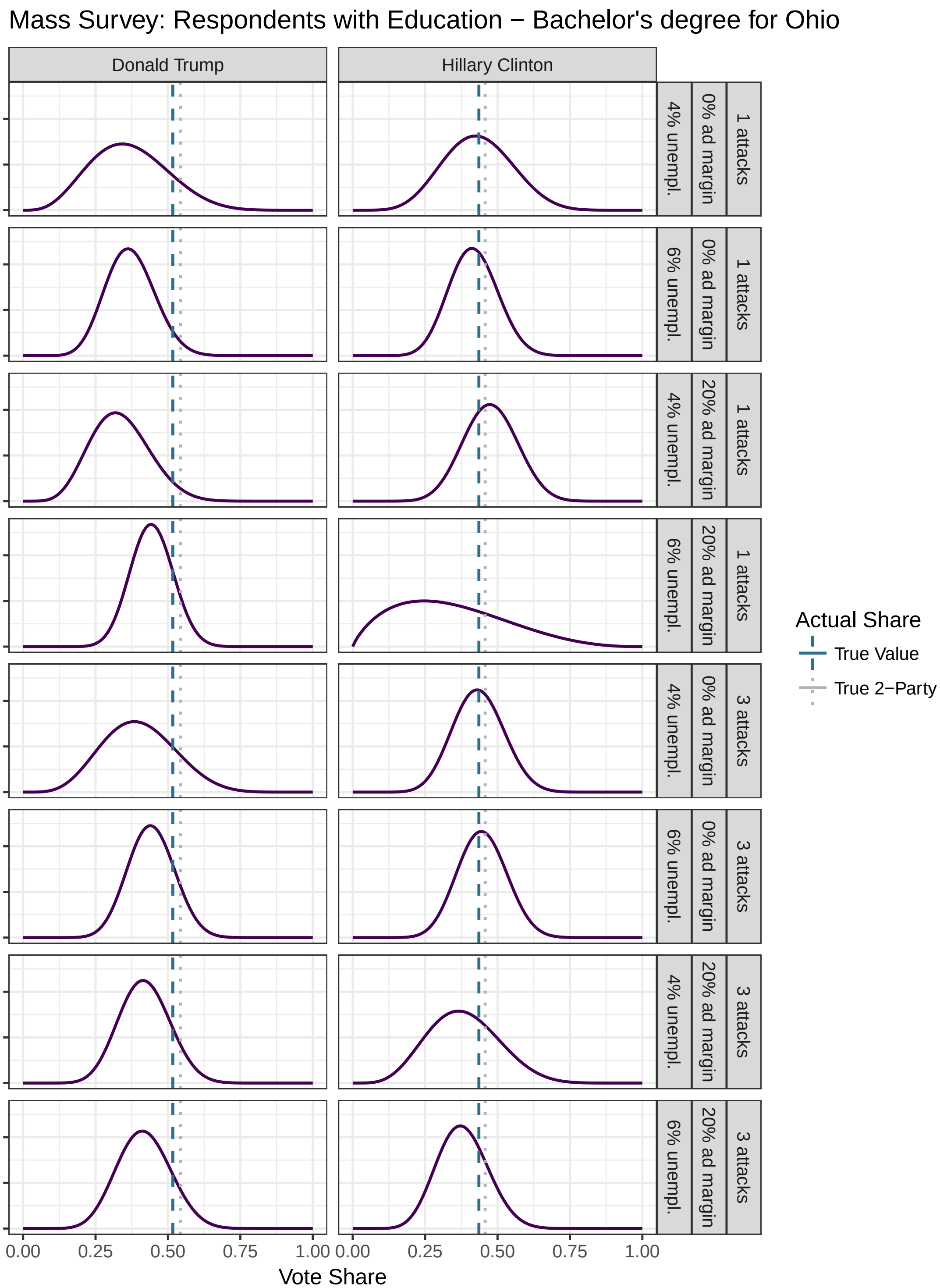


Figure 3.316: Priors with covariates: Mass Ohio Education Bachelor's degree

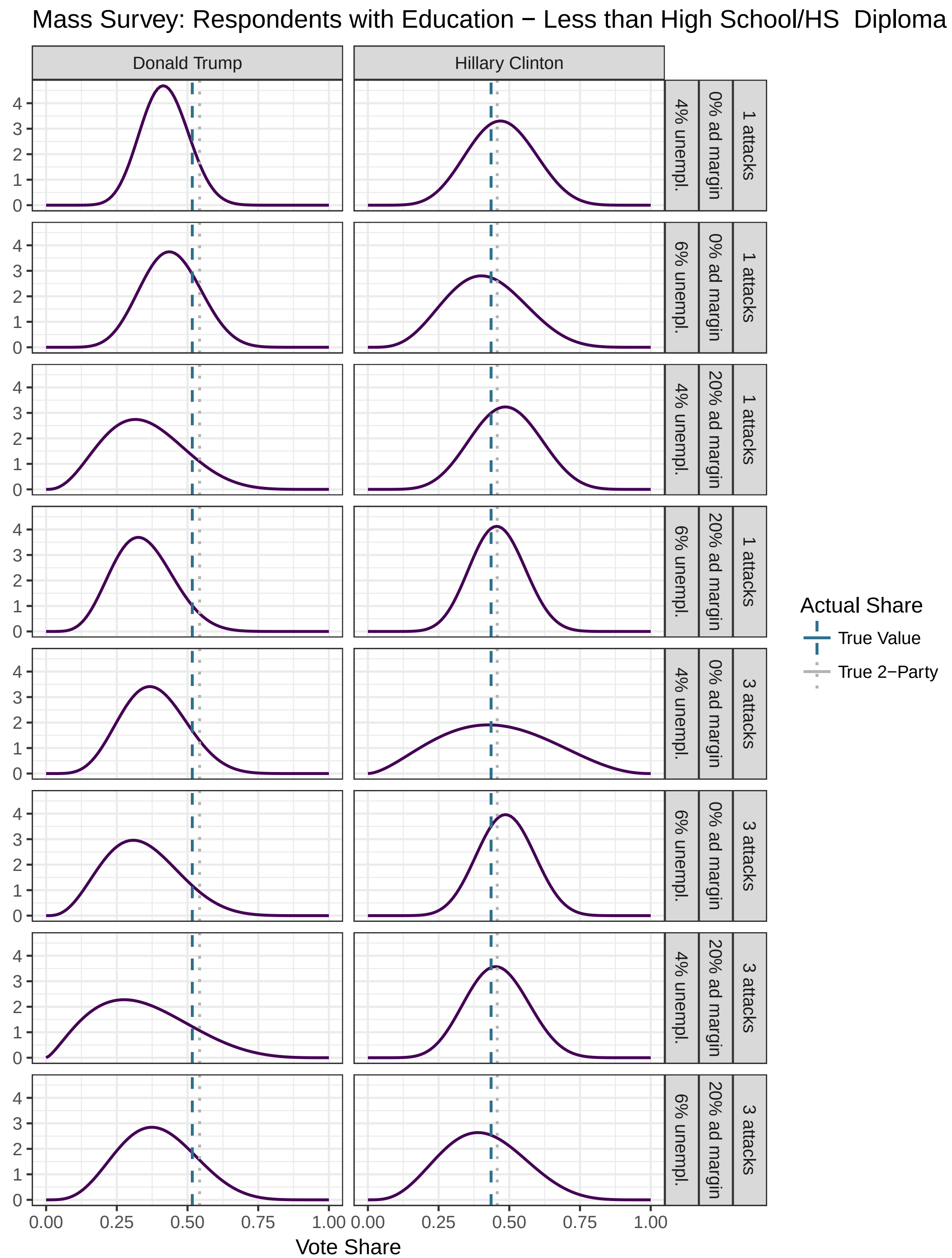


Figure 3.317: Priors with covariates: Mass Ohio Education Less than High School HS Diploma

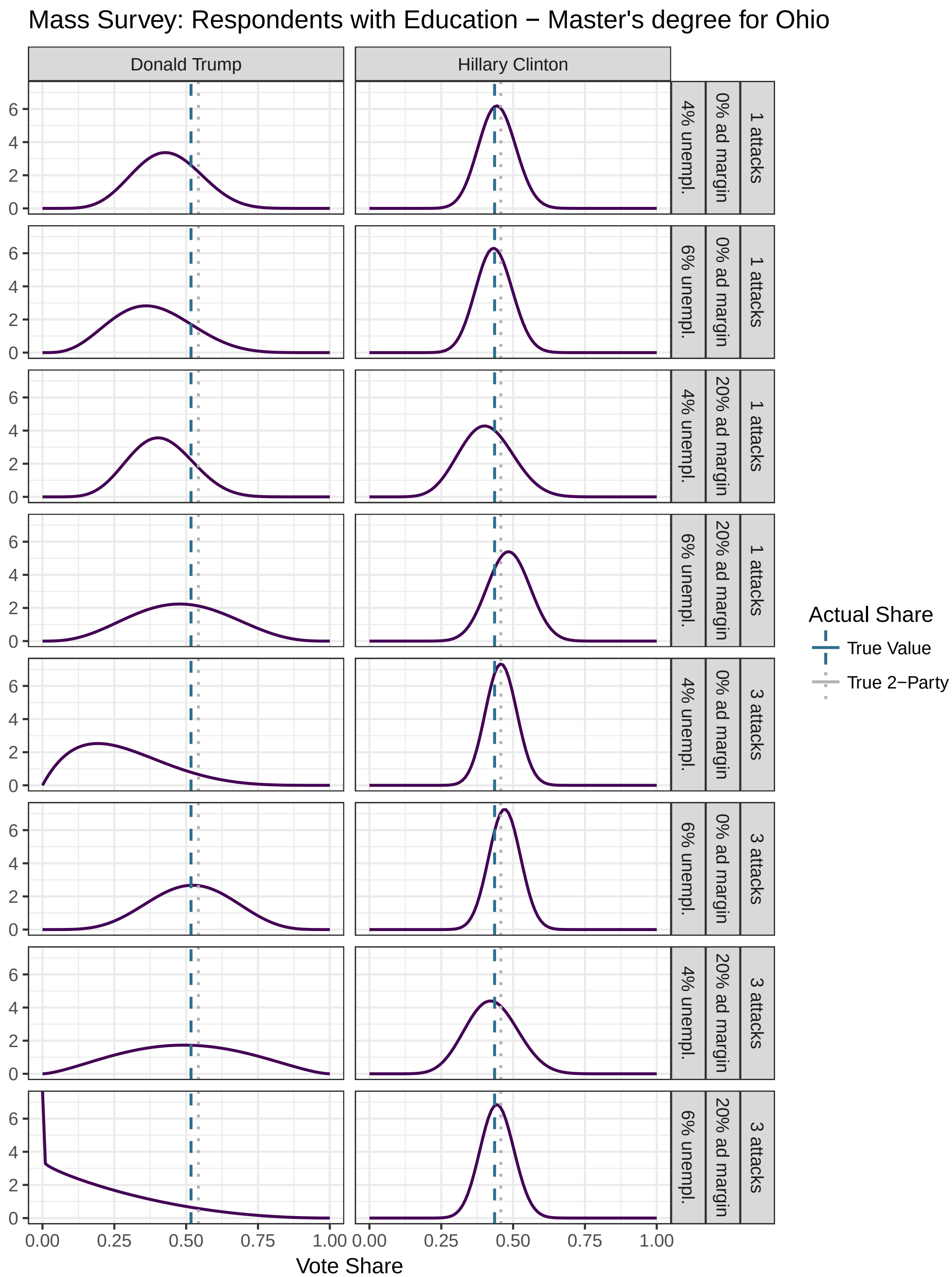


Figure 3.318: Priors with covariates: Mass Ohio Education Master's degree

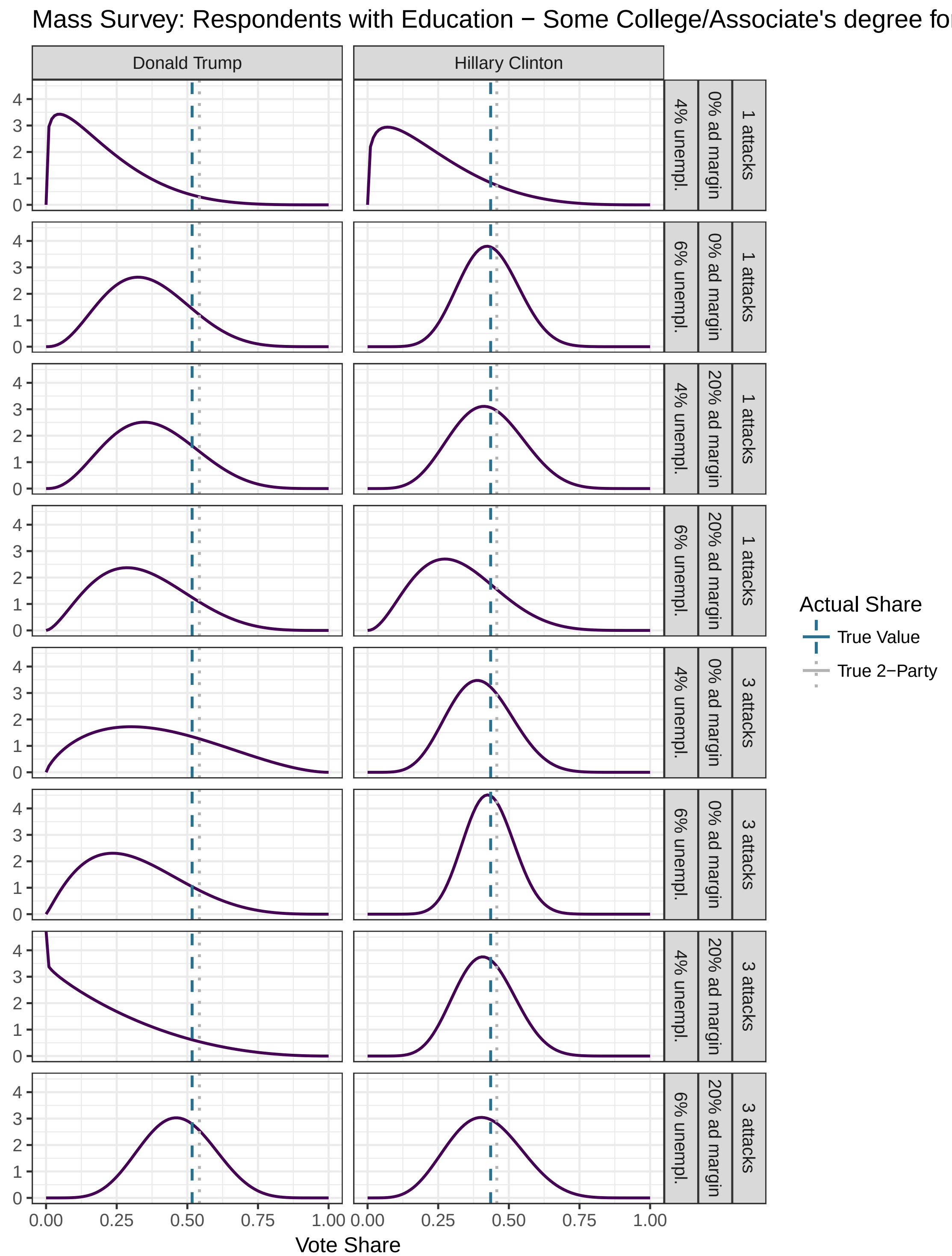


Figure 3.319: Priors with covariates: Mass Ohio Education Some College Associate's degree

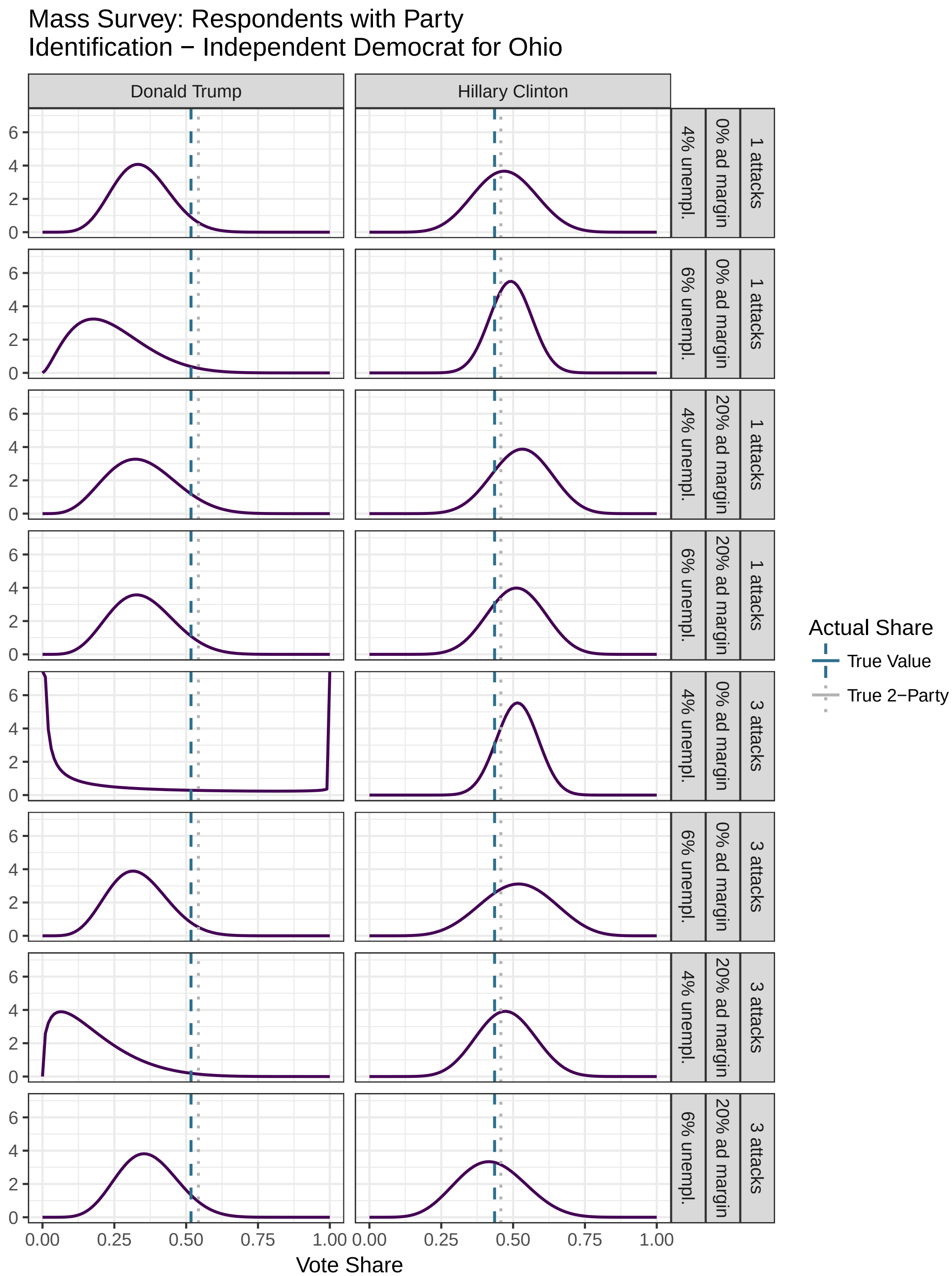


Figure 3.320: Priors with covariates: Mass Ohio Party Identification Independent Democrat

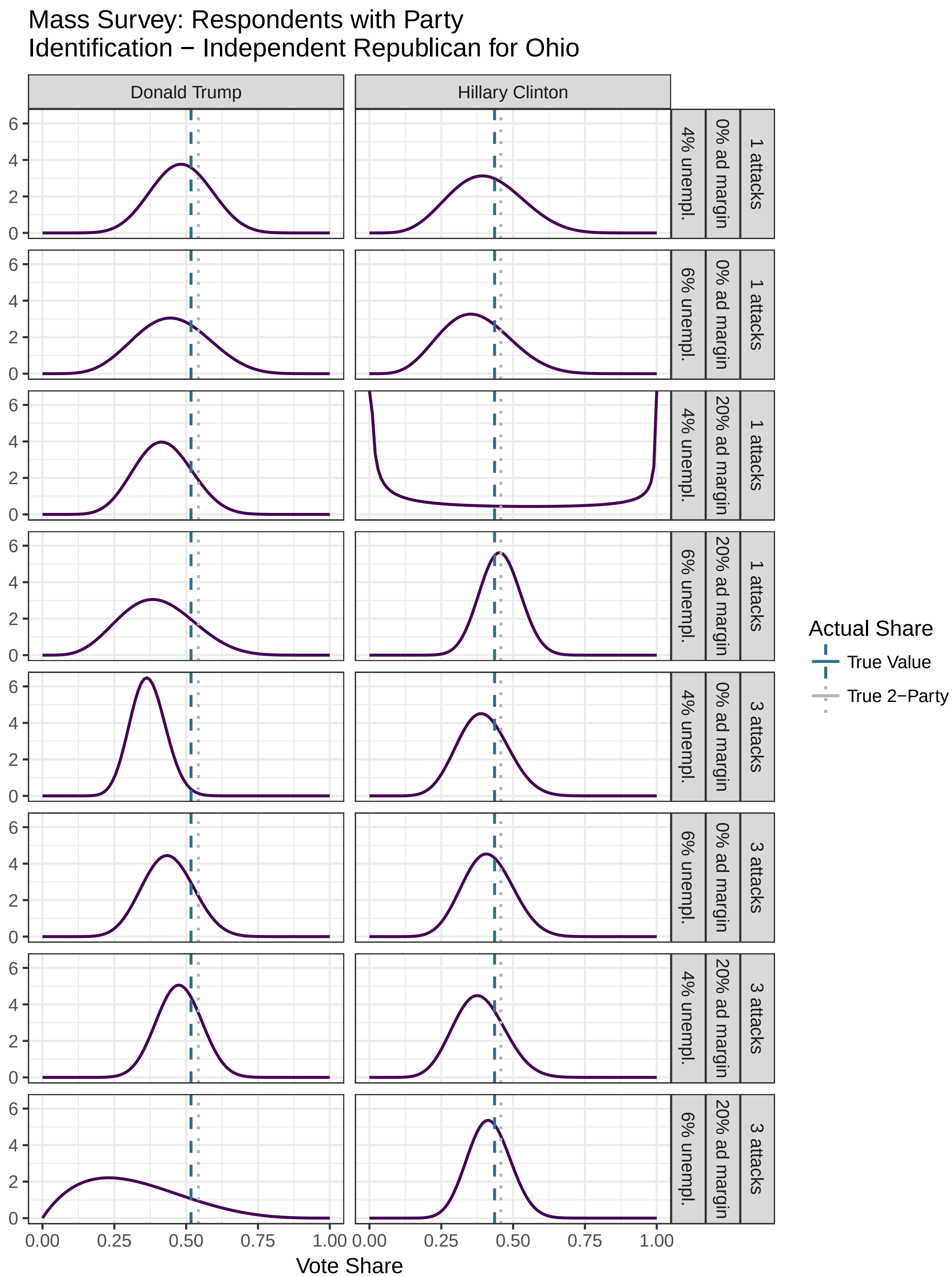


Figure 3.321: Priors with covariates: Mass Ohio Party Identification Independent Republican

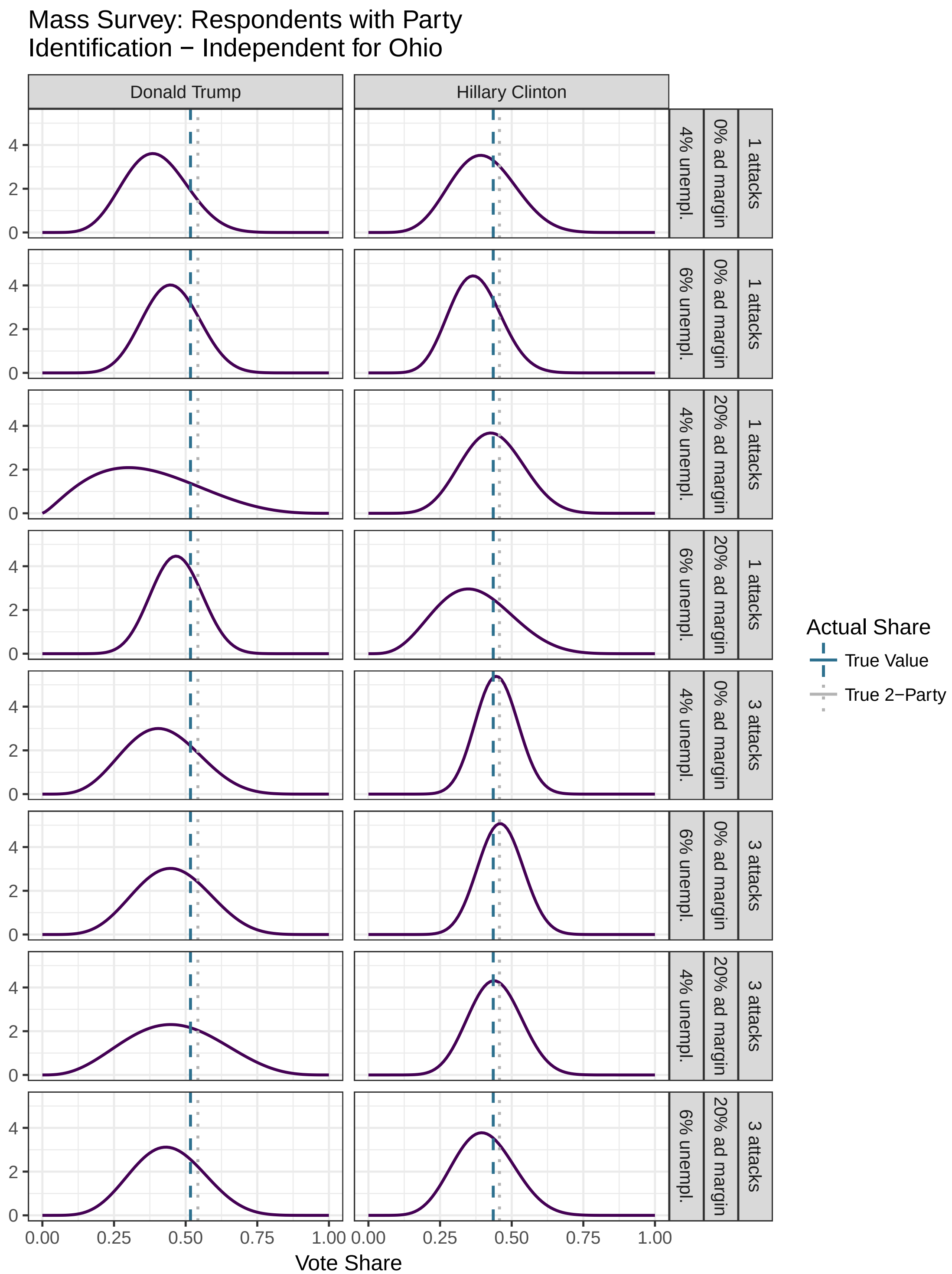


Figure 3.322: Priors with covariates: Mass Ohio Party Identification Independent

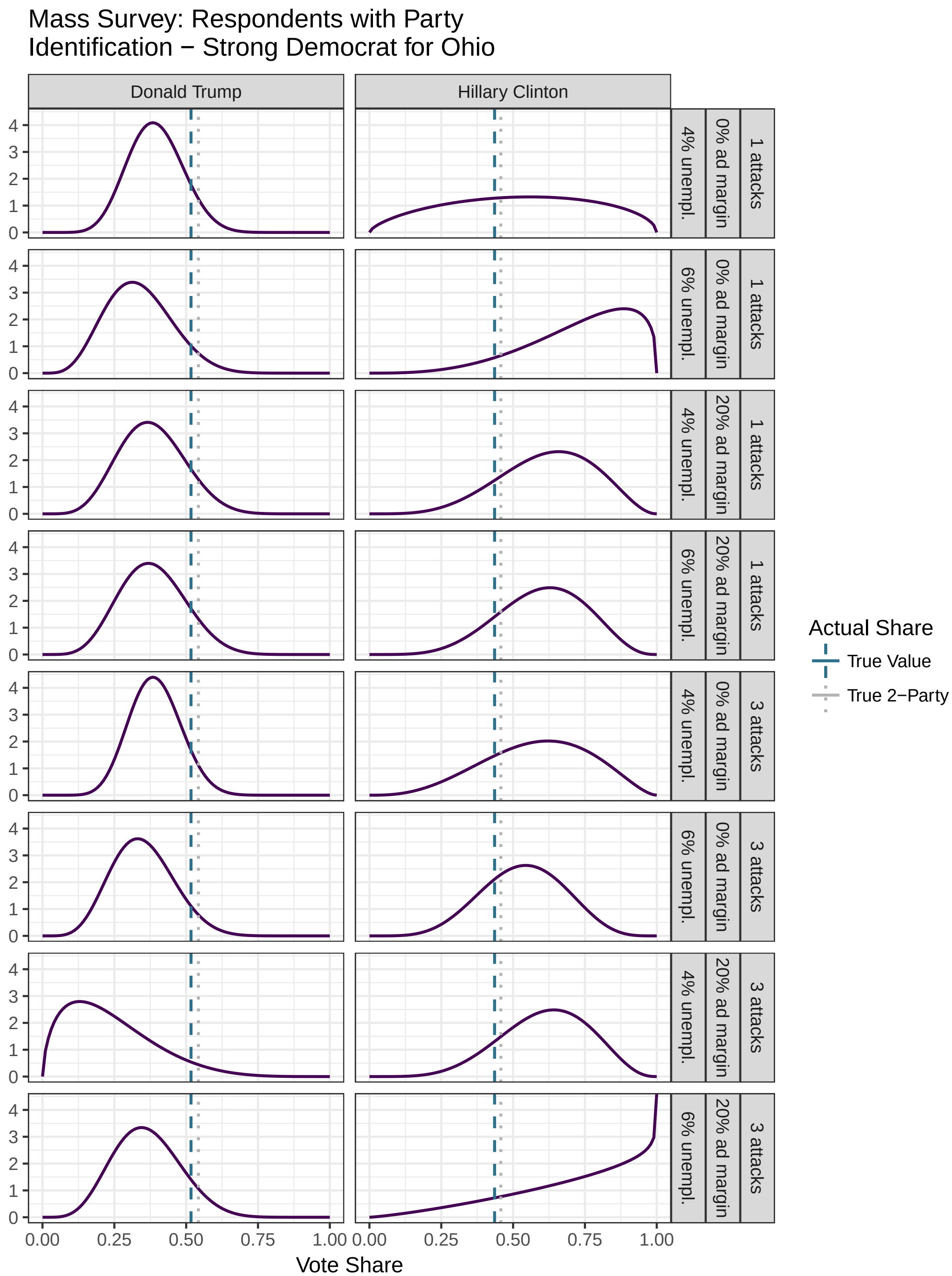


Figure 3.323: Priors with covariates: Mass Ohio Party Identification Strong Democrat

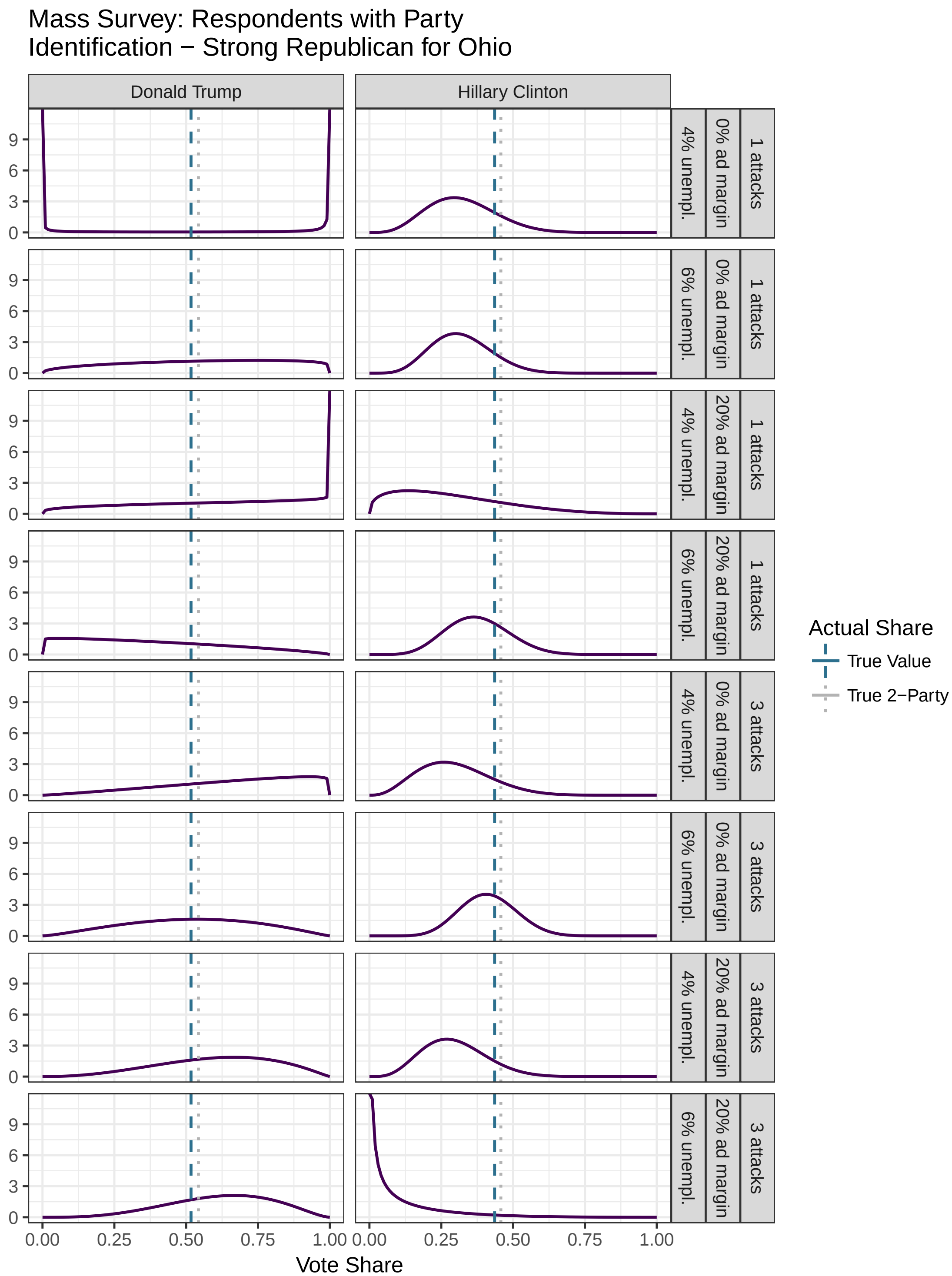


Figure 3.324: Priors with covariates: Mass Ohio Party Identification Strong Republican

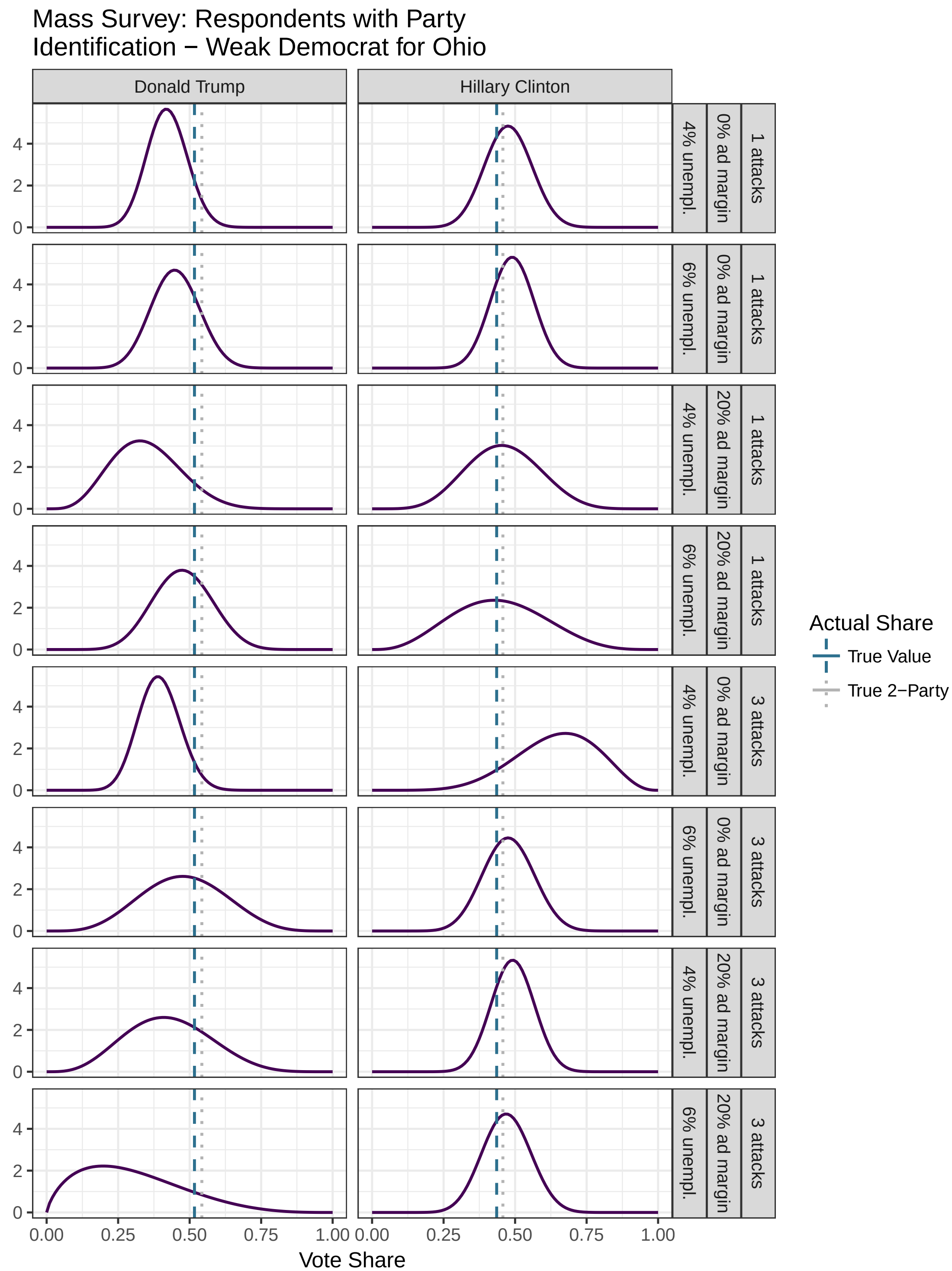


Figure 3.325: Priors with covariates: Mass Ohio Party Identification Weak Democrat

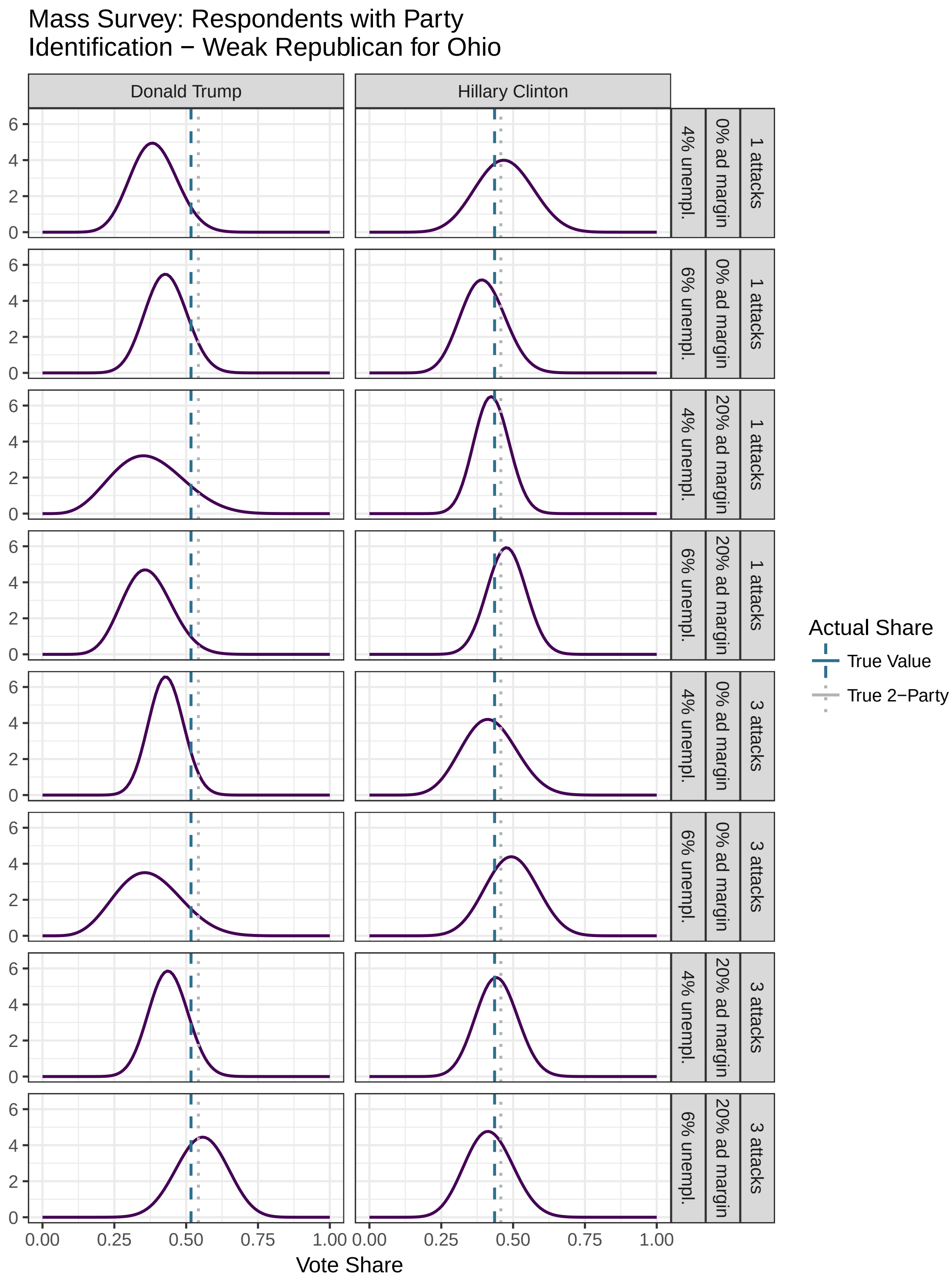


Figure 3.326: Priors with covariates: Mass Ohio Party Identification Weak Republican

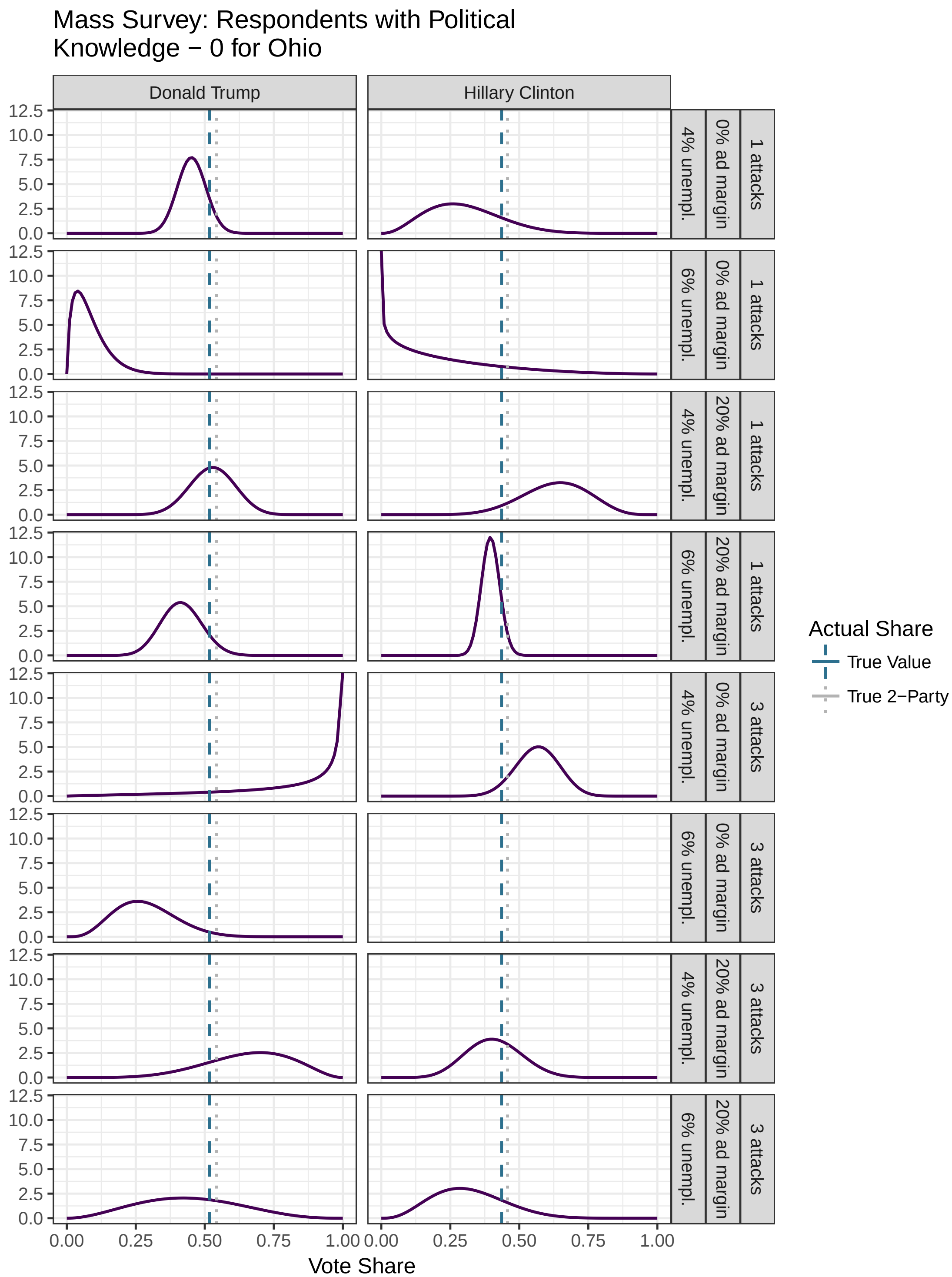


Figure 3.327: Priors with covariates: Mass Ohio Political Knowledge 0

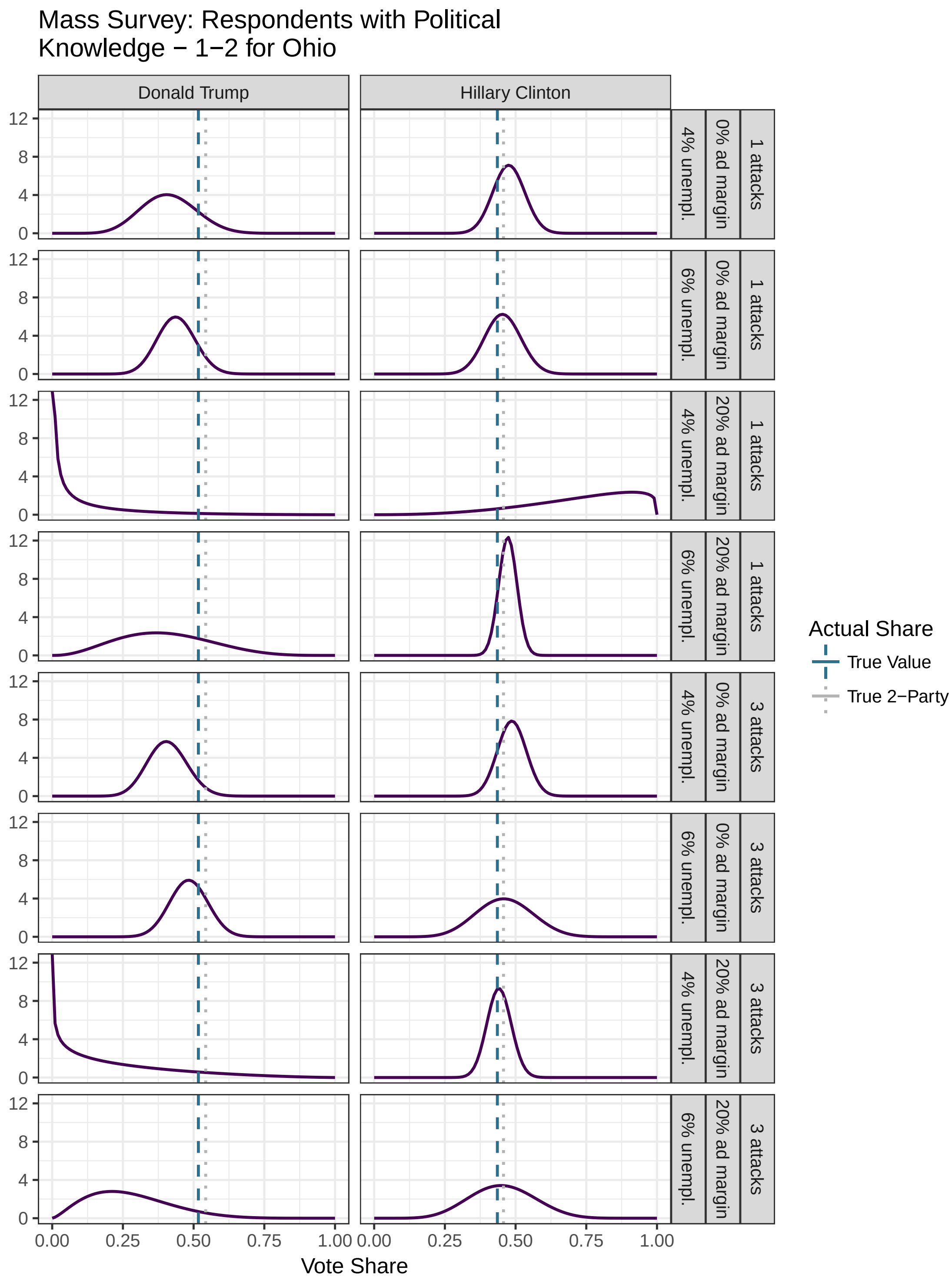


Figure 3.328: Priors with covariates: Mass Ohio Political Knowledge 1-2

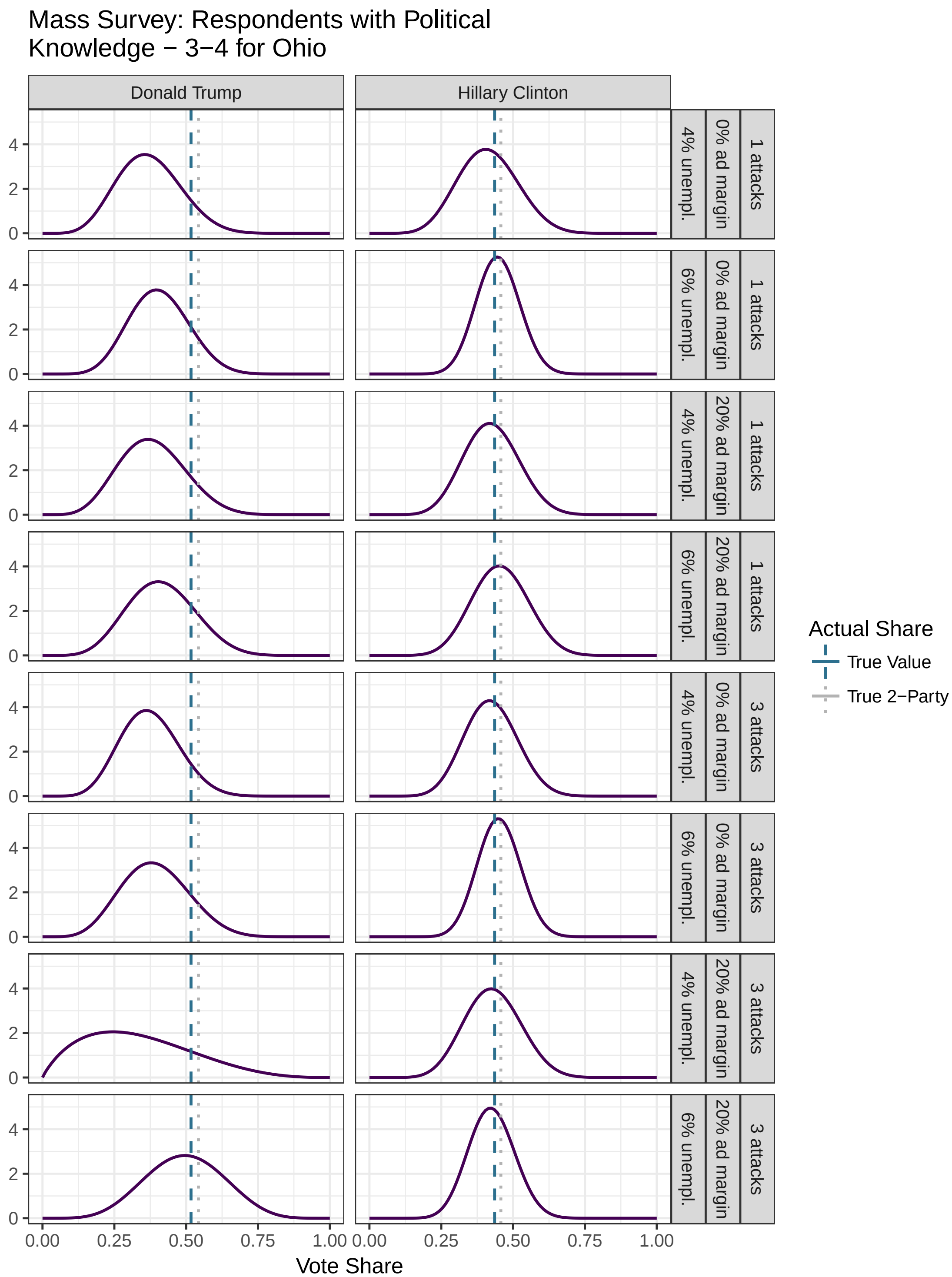


Figure 3.329: Priors with covariates: Mass Ohio Political Knowledge 3-4

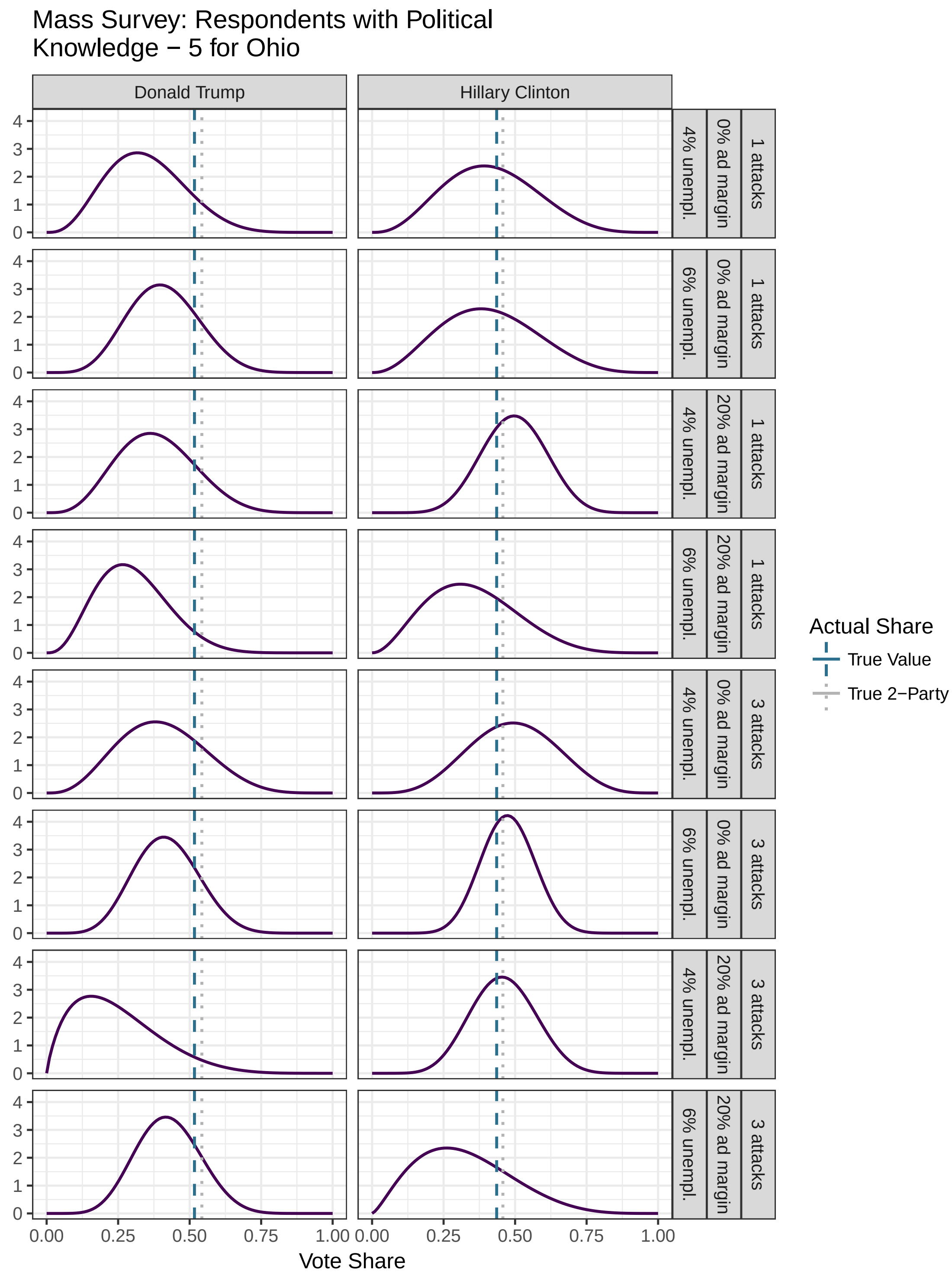


Figure 3.330: Priors with covariates: Mass Ohio Political Knowledge 5

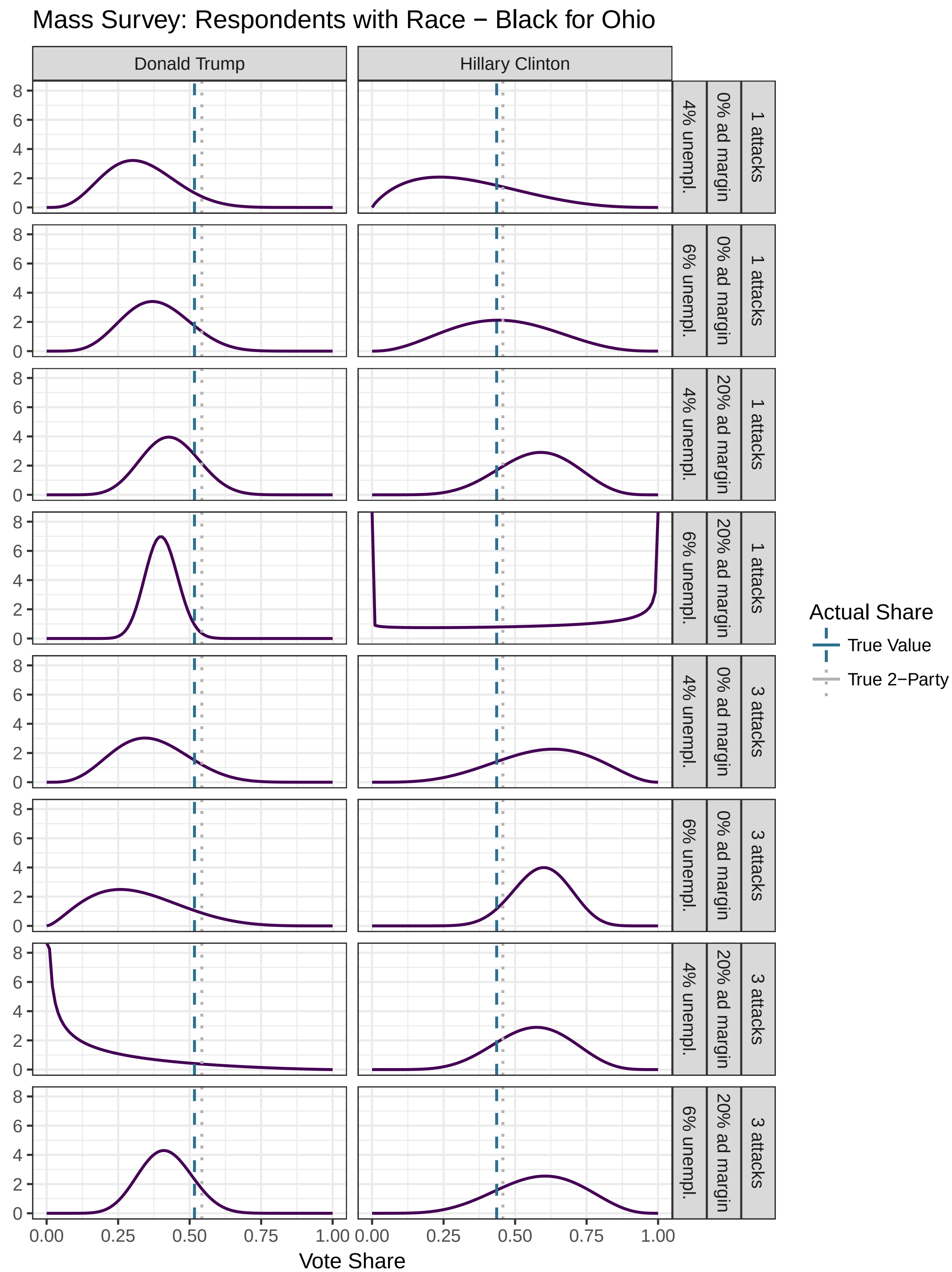


Figure 3.331: Priors with covariates: Mass Ohio Race Black

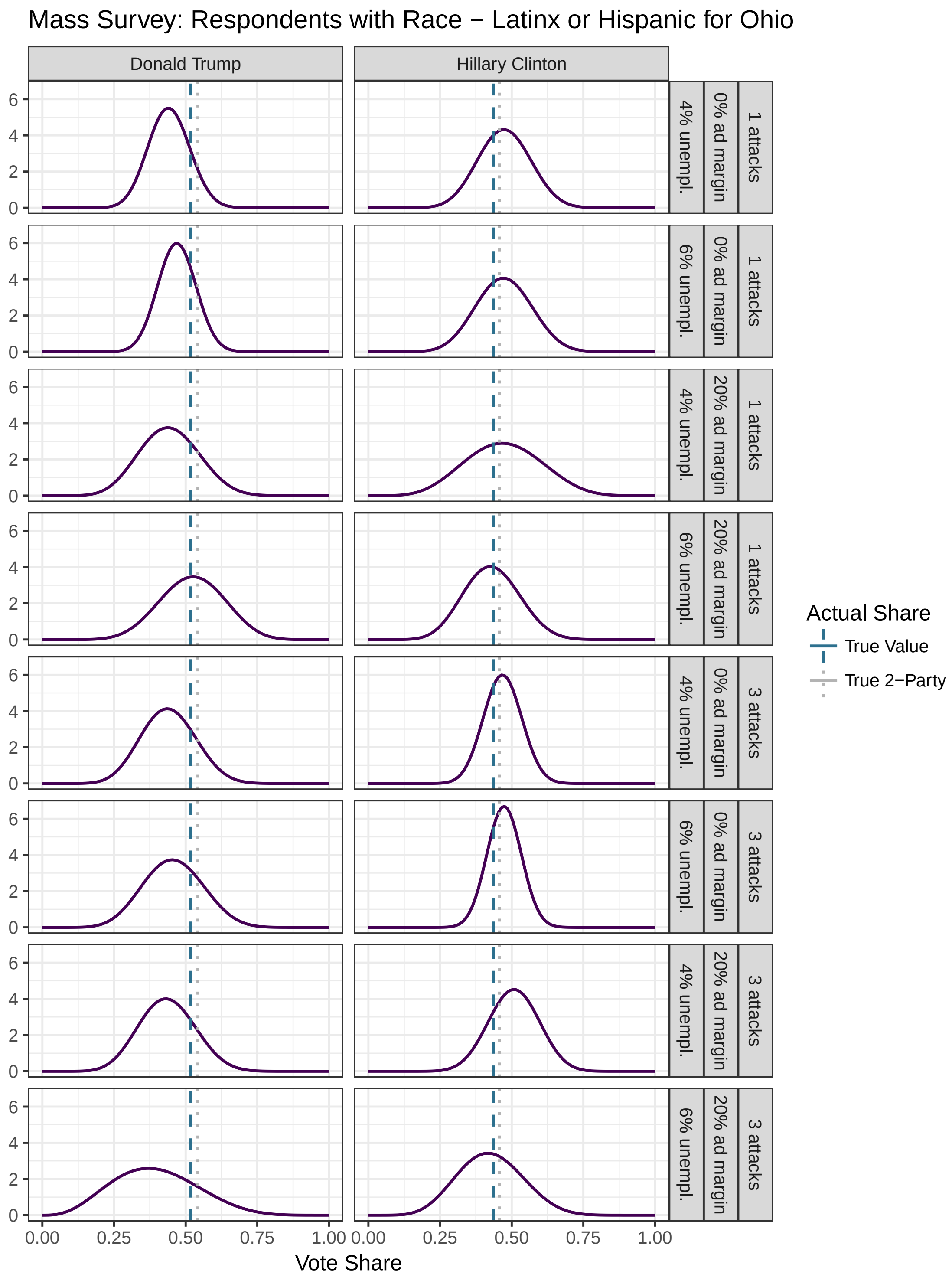


Figure 3.332: Priors with covariates: Mass Ohio Race Latinx or Hispanic

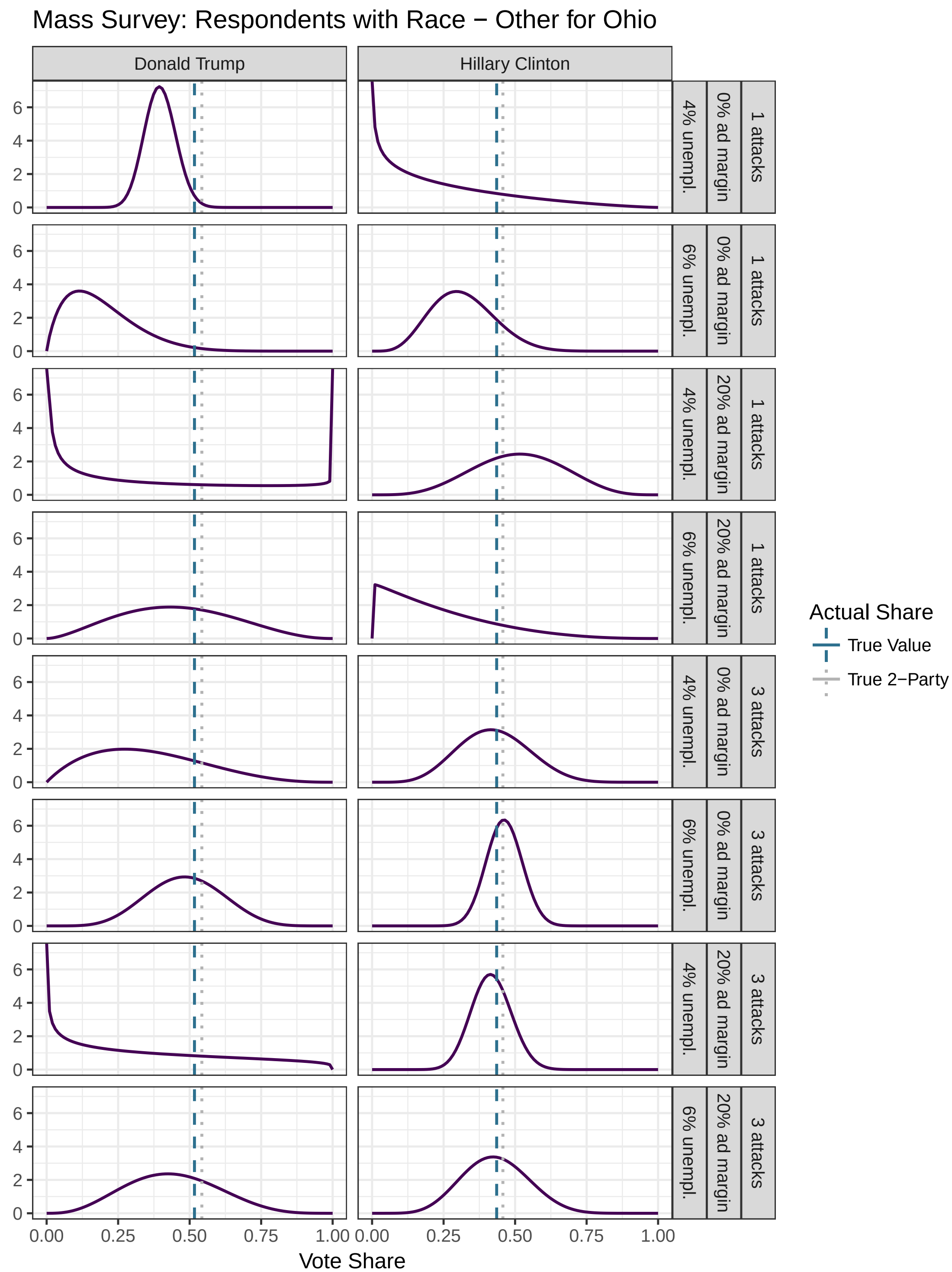


Figure 3.333: Priors with covariates: Mass Ohio Race Other

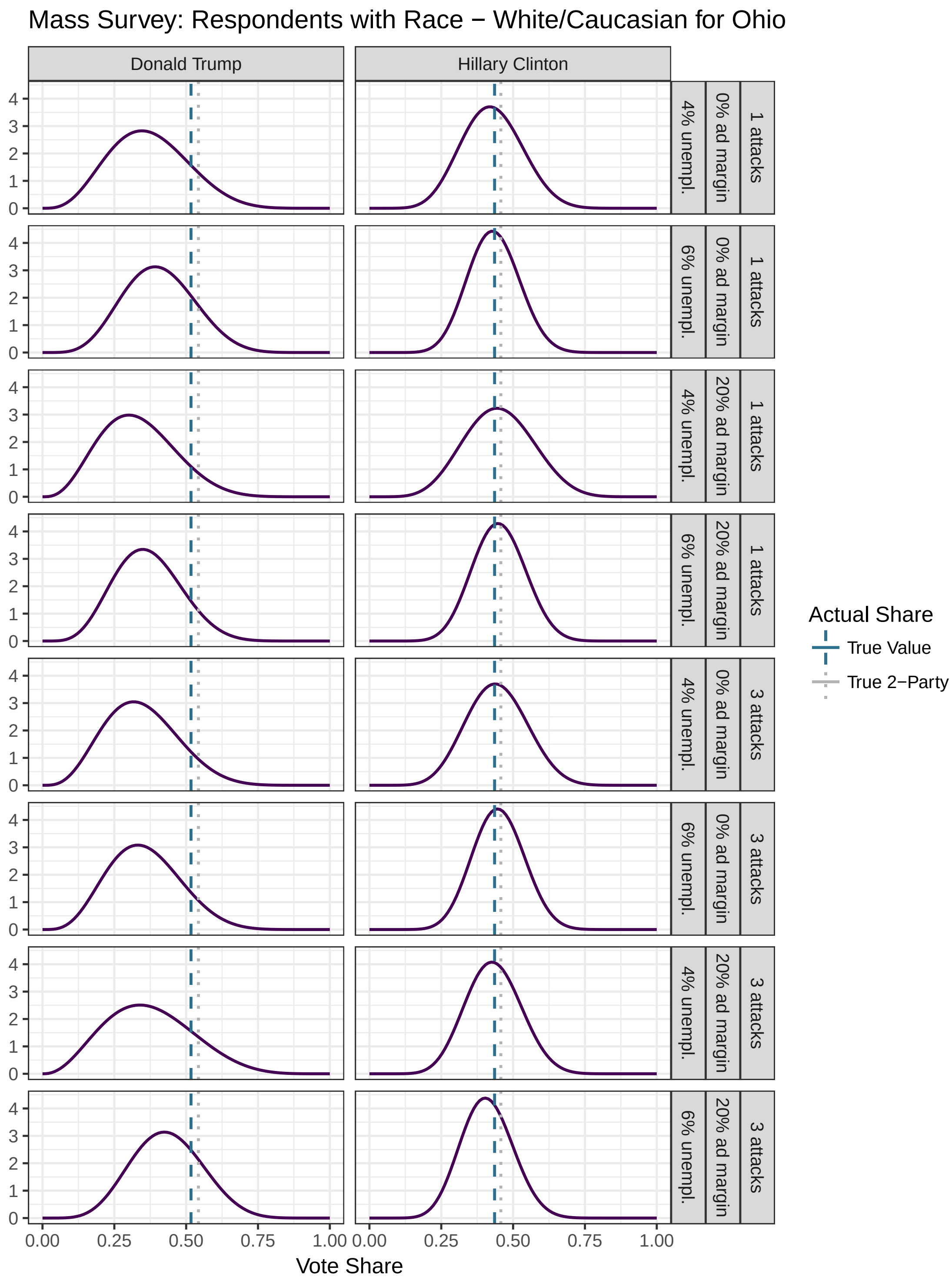


Figure 3.334: Priors with covariates: Mass Ohio Race White Caucasian

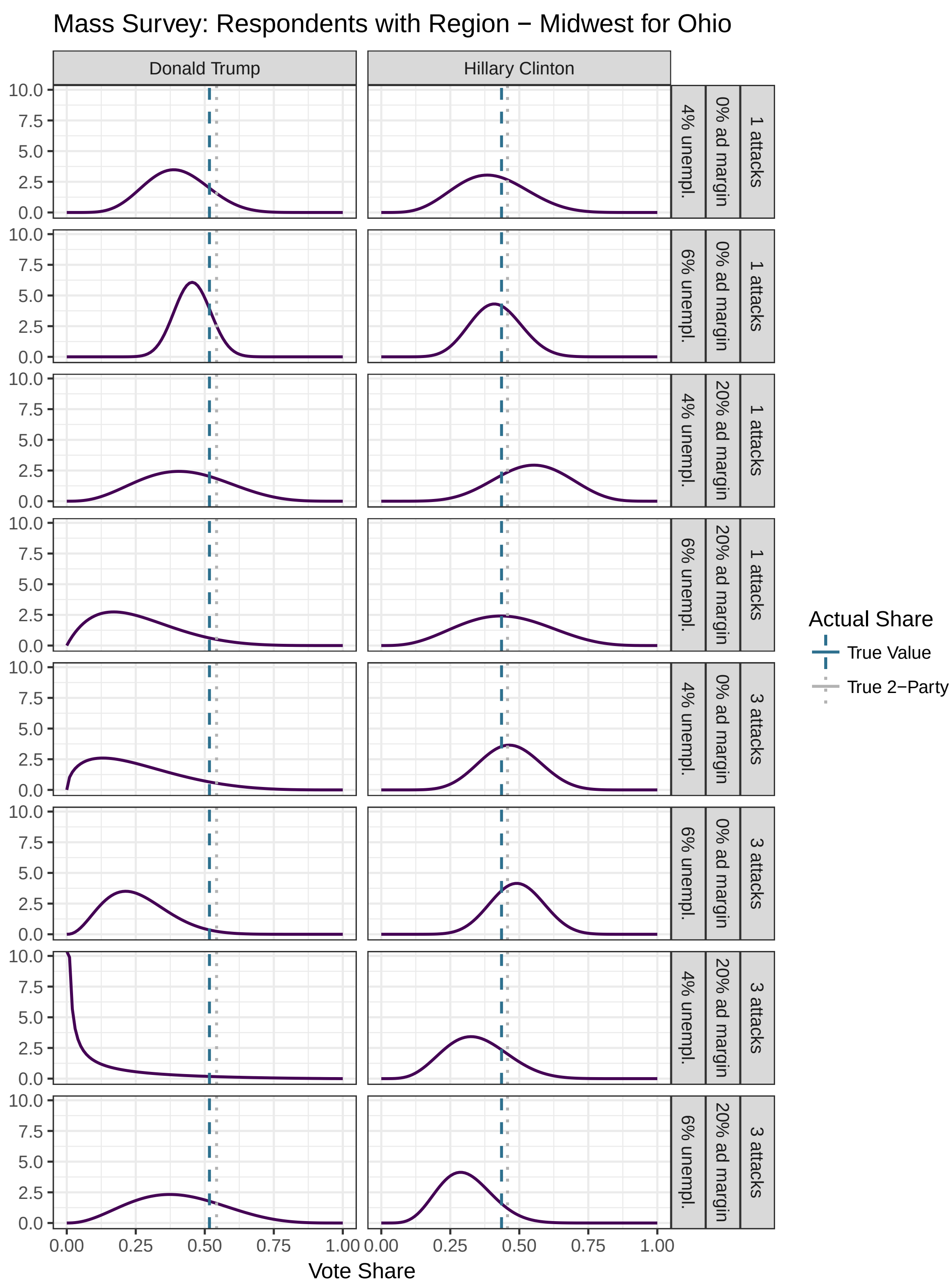


Figure 3.335: Priors with covariates: Mass Ohio Region Midwest

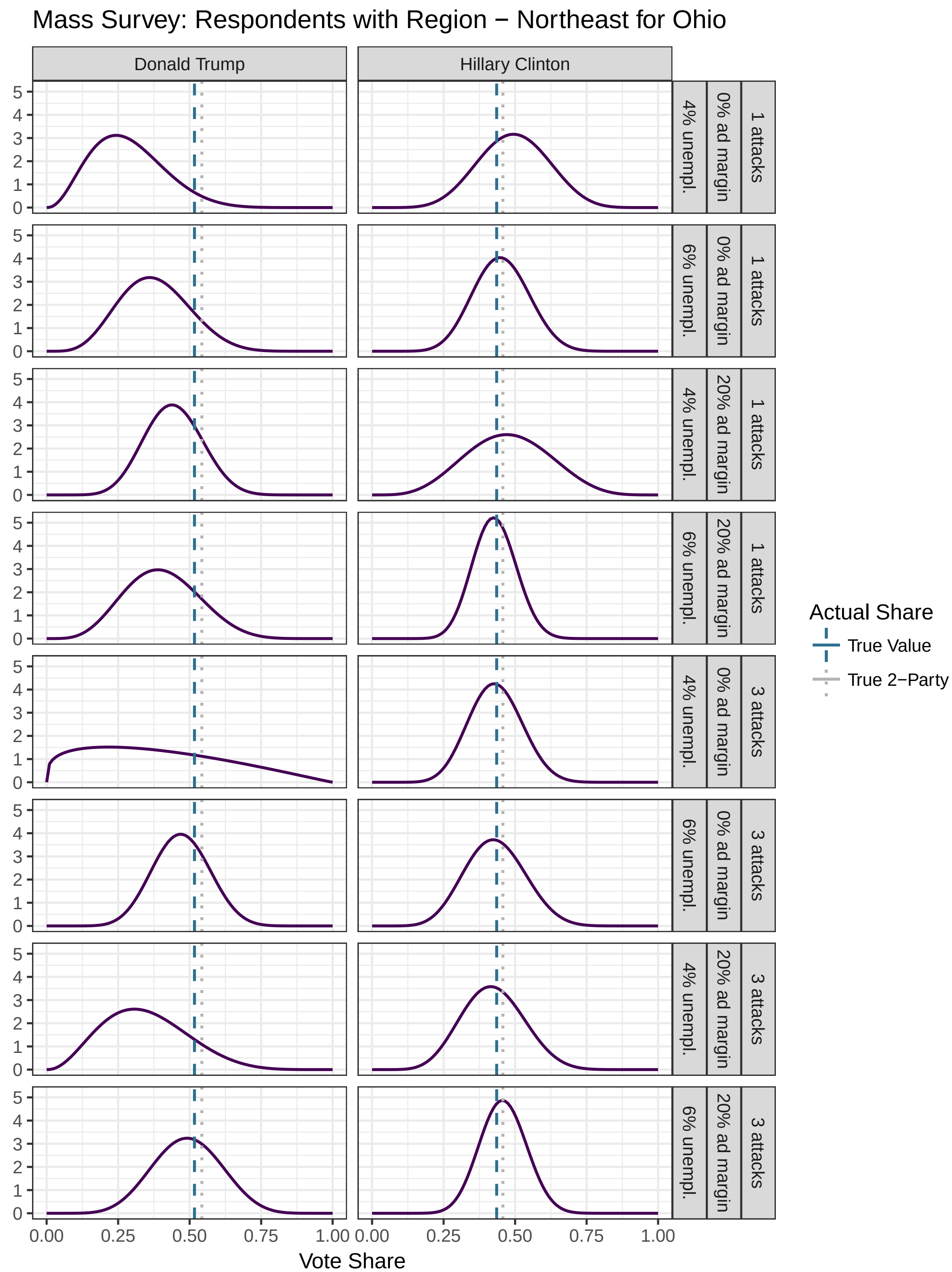


Figure 3.336: Priors with covariates: Mass Ohio Region Northeast

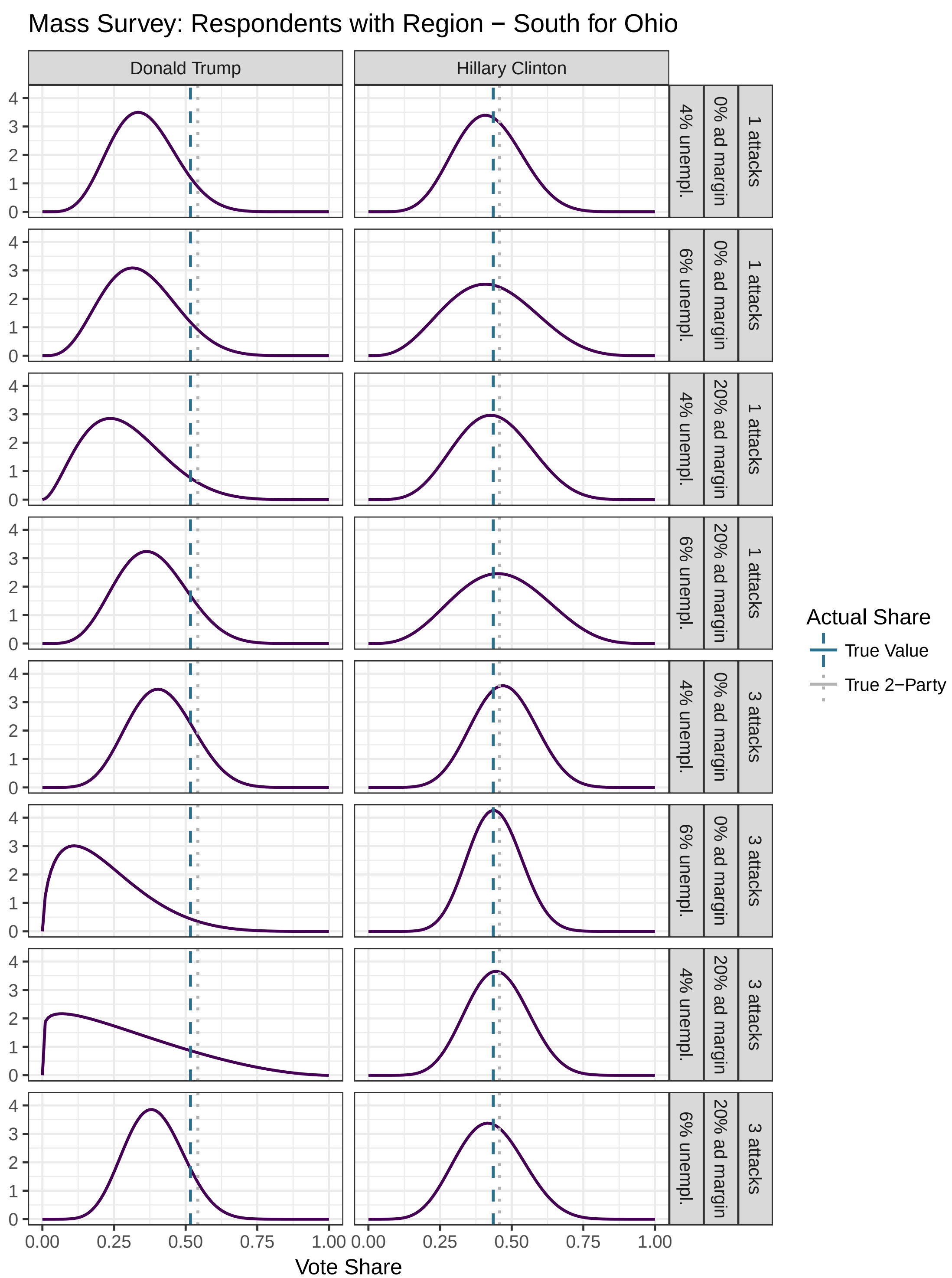


Figure 3.337: Priors with covariates: Mass Ohio Region South

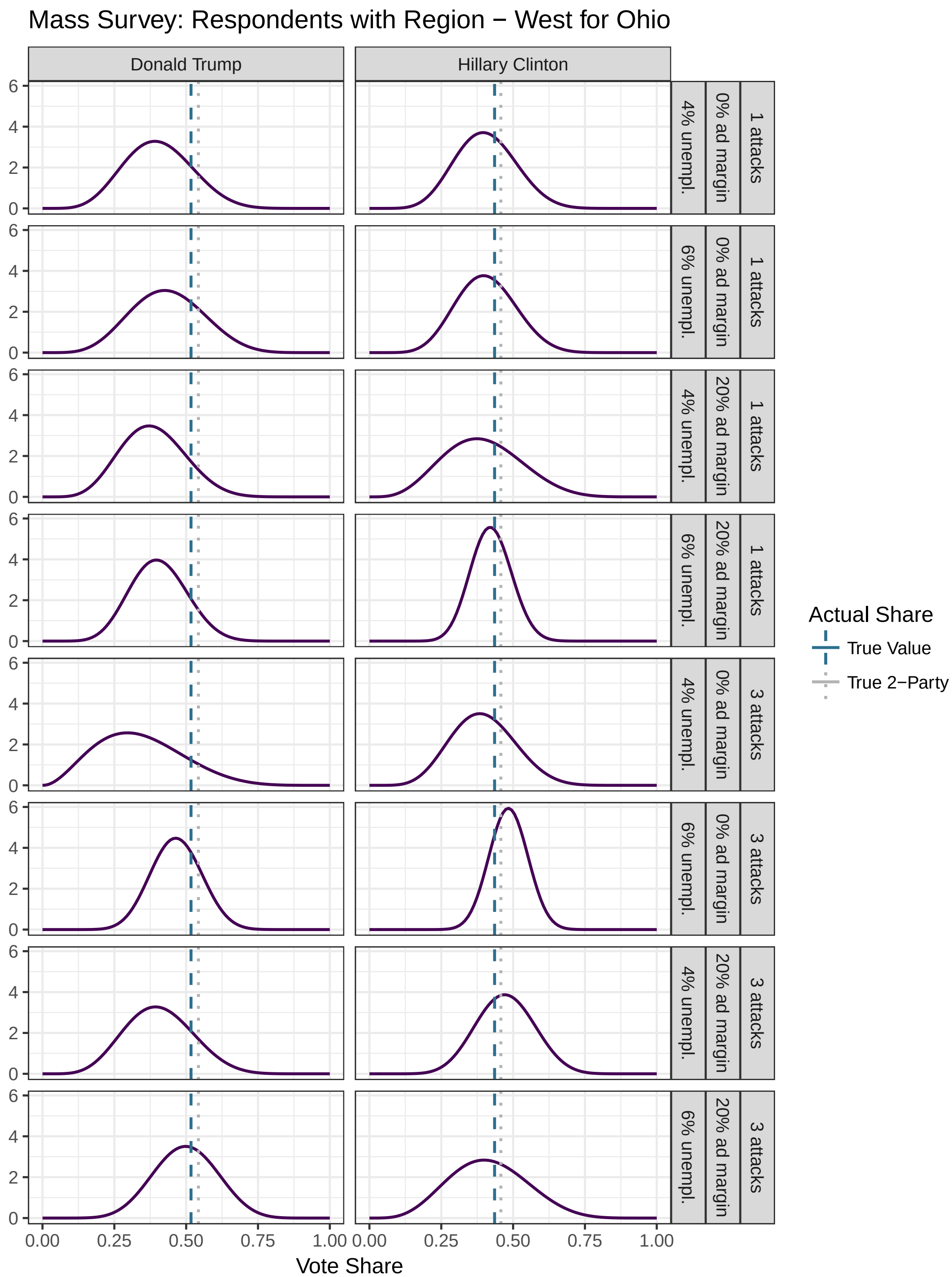


Figure 3.338: Priors with covariates: Mass Ohio Region West

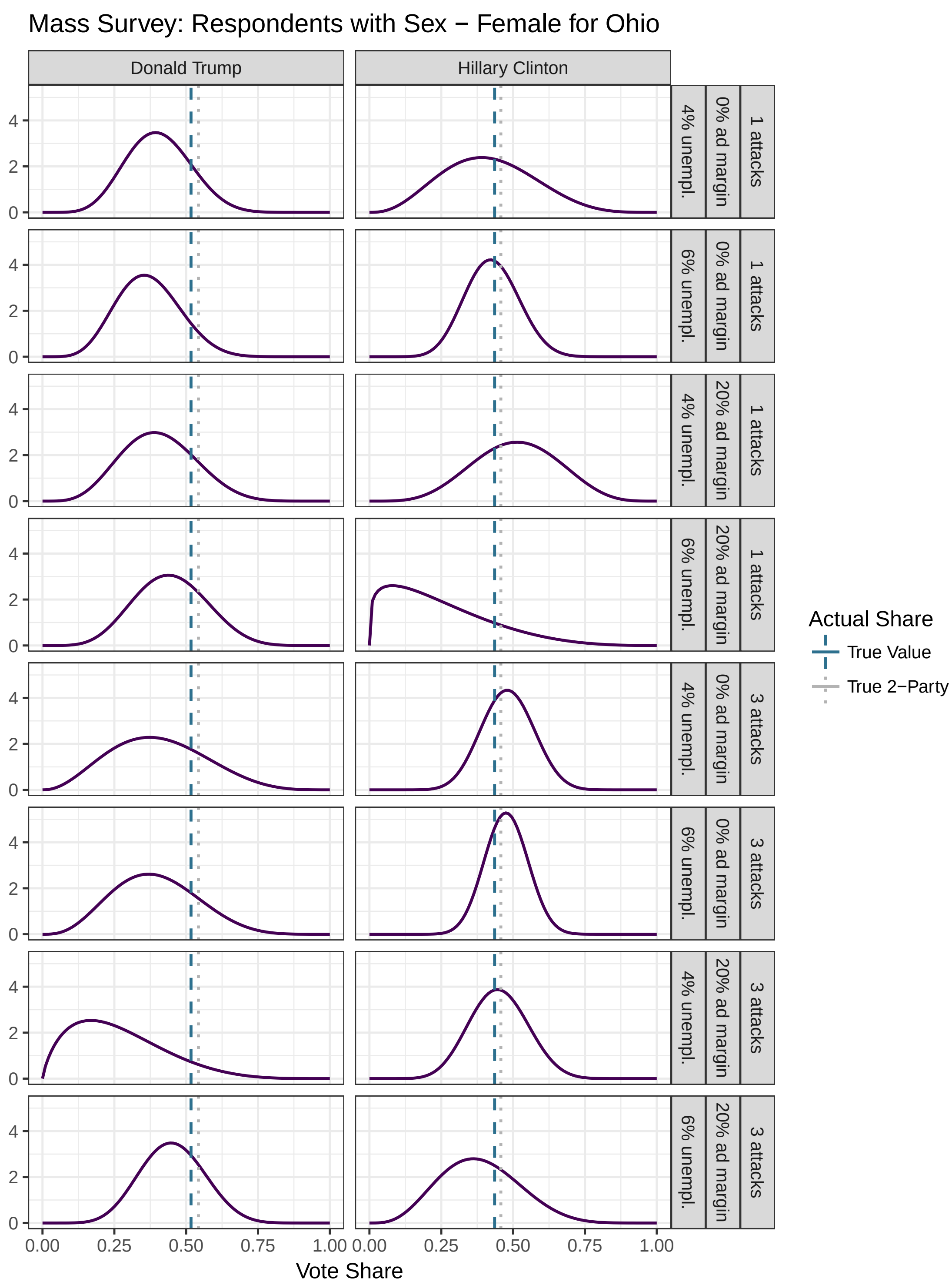


Figure 3.339: Priors with covariates: Mass Ohio Sex Female

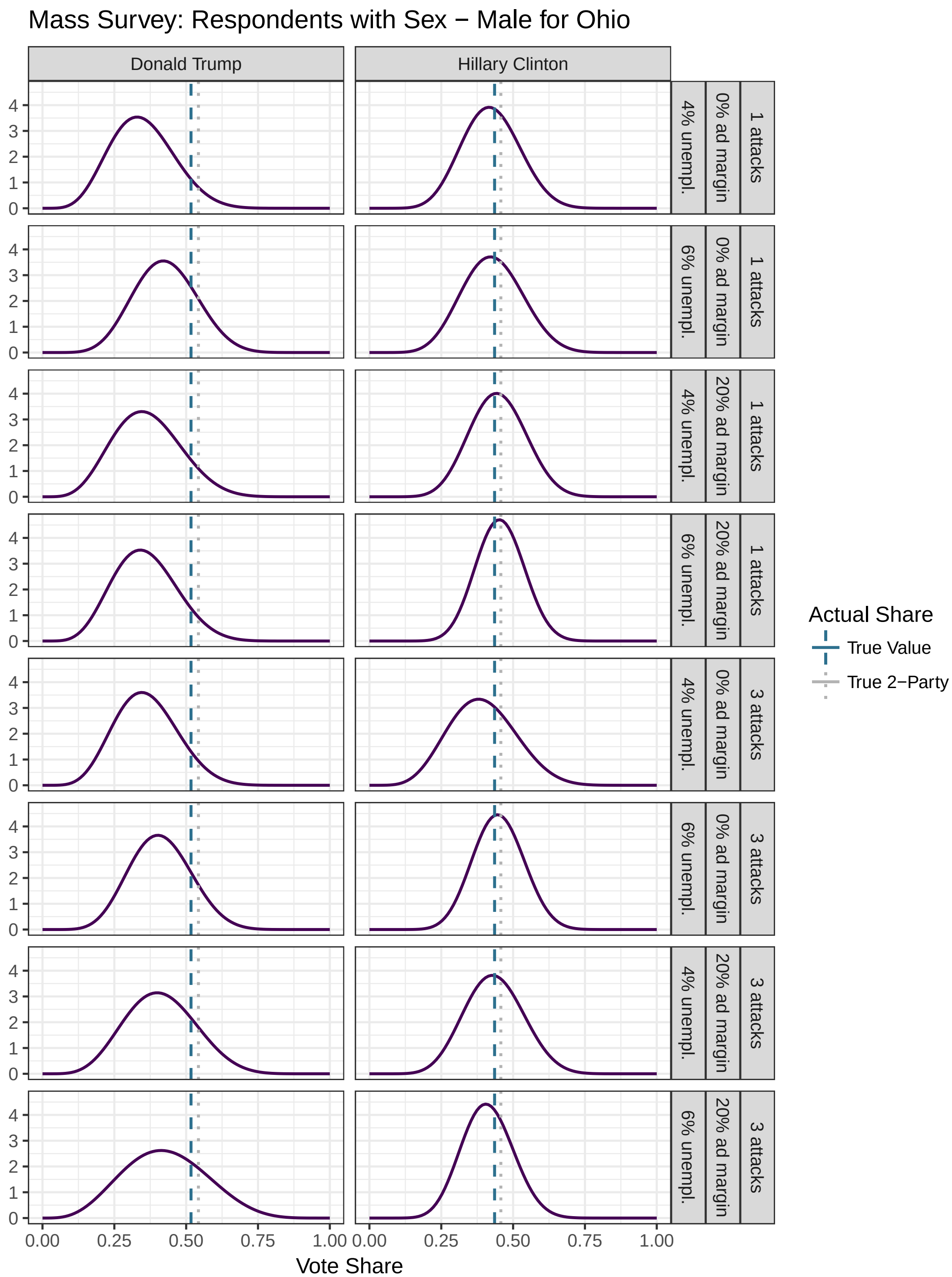


Figure 3.340: Priors with covariates: Mass Ohio Sex Male

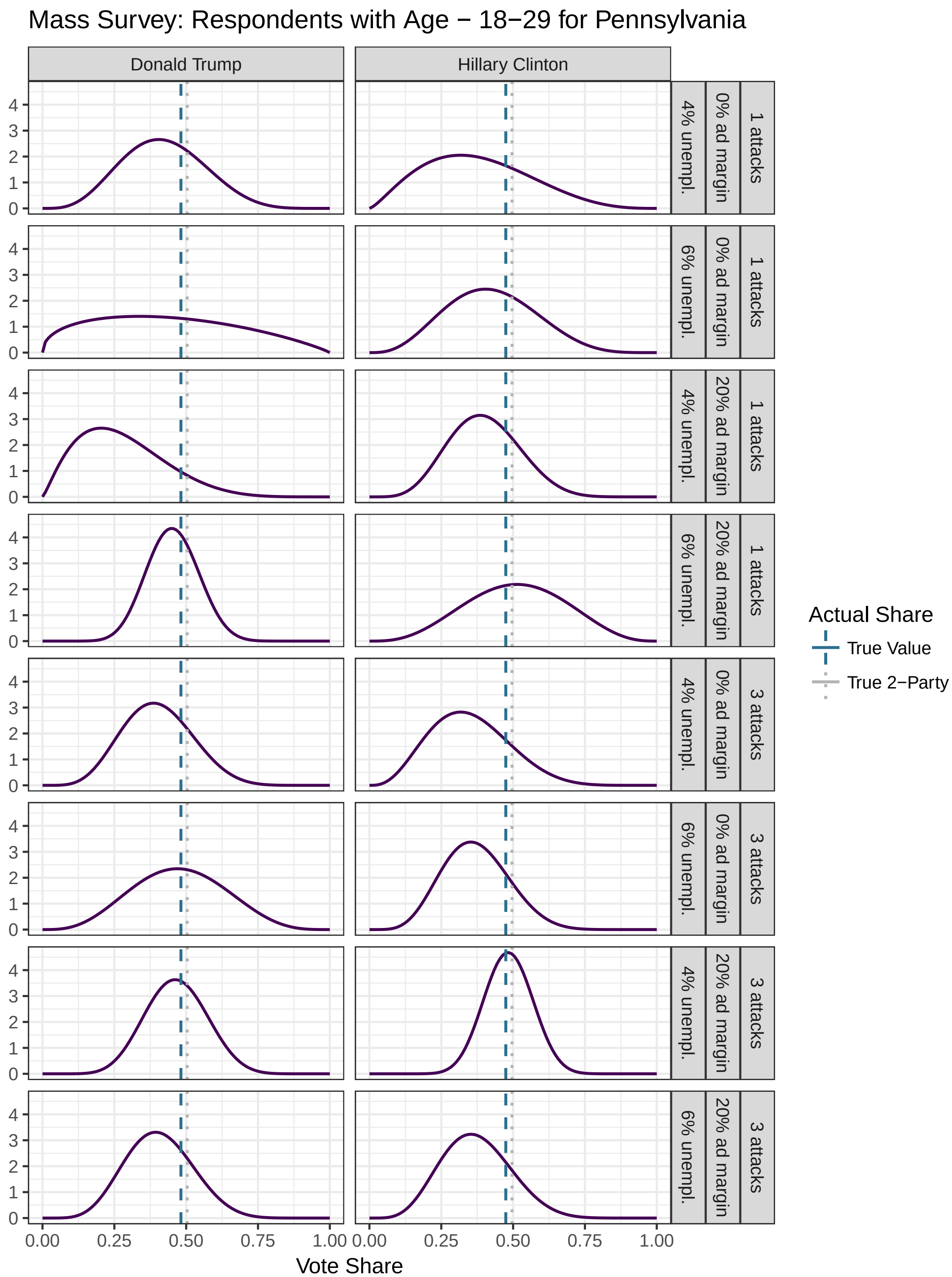


Figure 3.341: Priors with covariates: Mass Pennsylvania Age 18-29

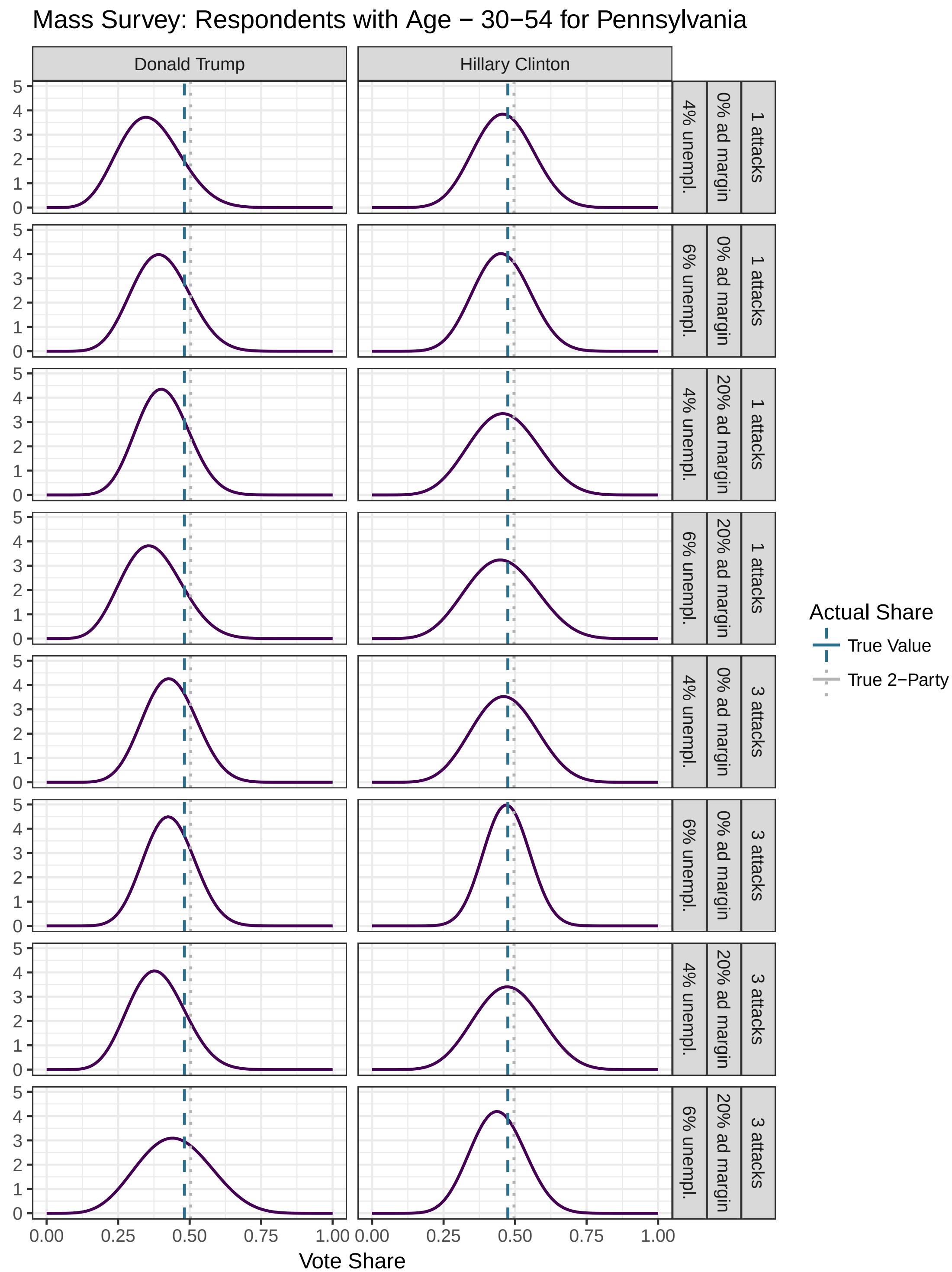


Figure 3.342: Priors with covariates: Mass Pennsylvania Age 30-54

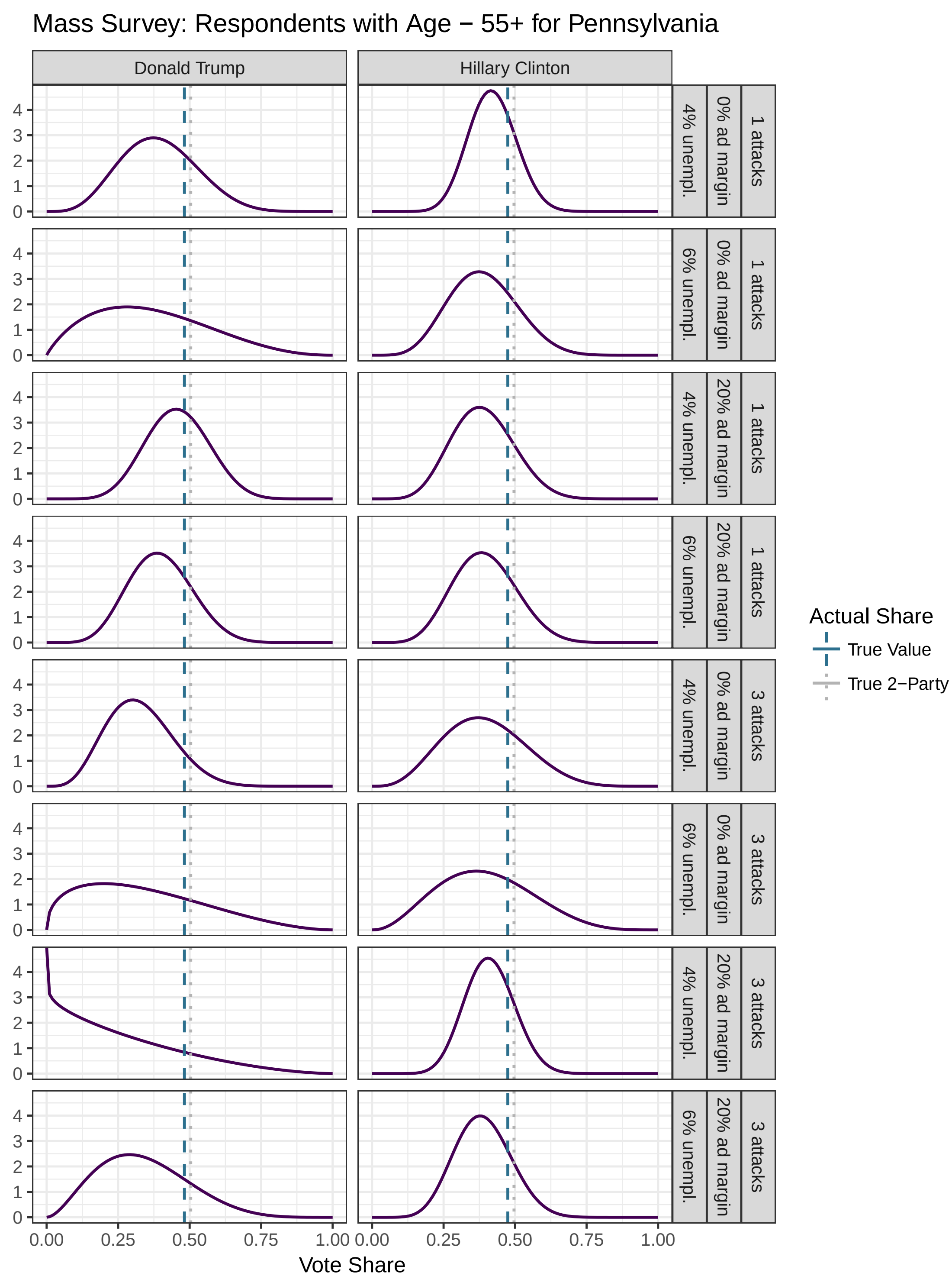


Figure 3.343: Priors with covariates: Mass Pennsylvania Age 55+

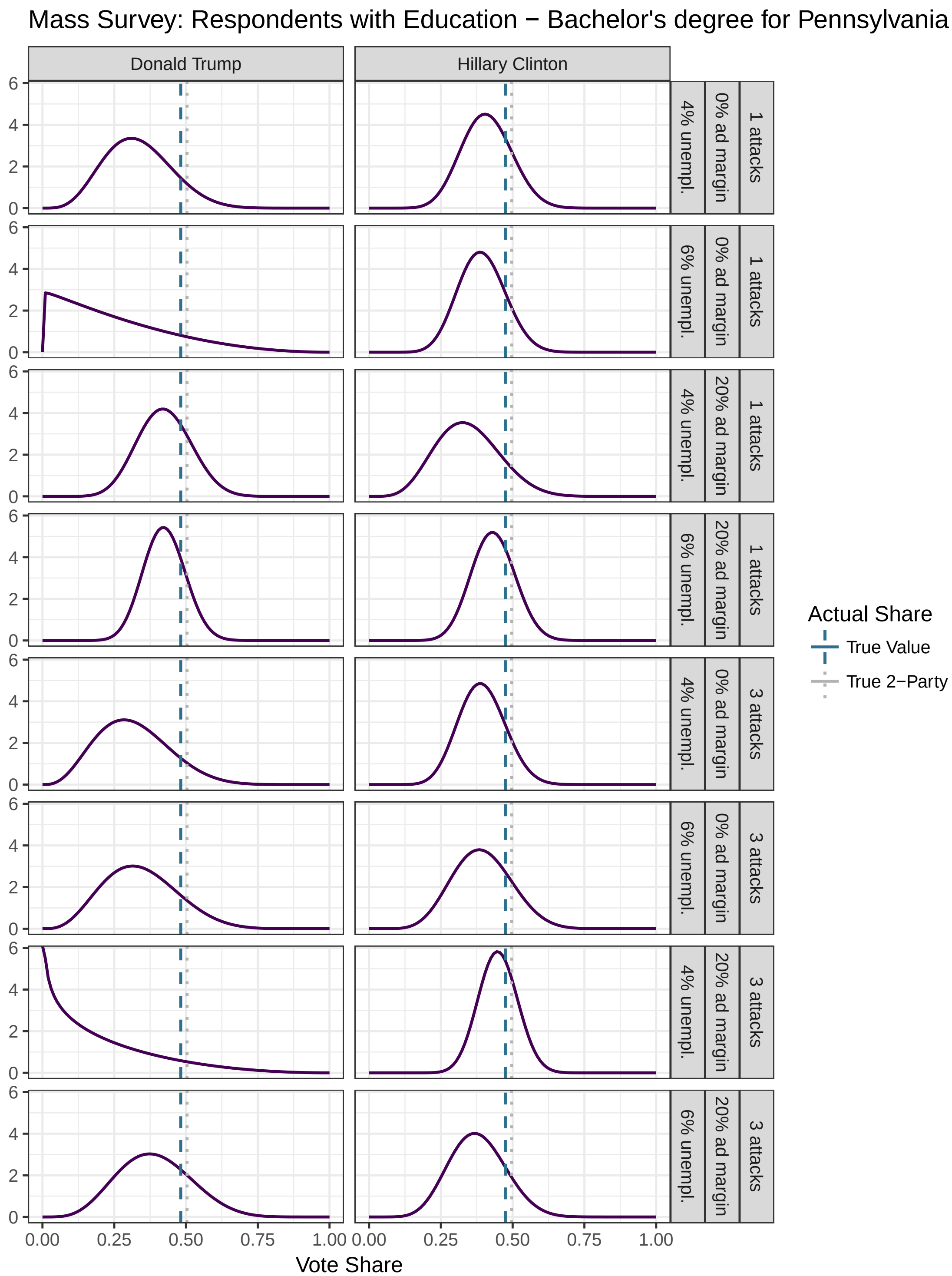


Figure 3.344: Priors with covariates: Mass Pennsylvania Education Bachelor's degree

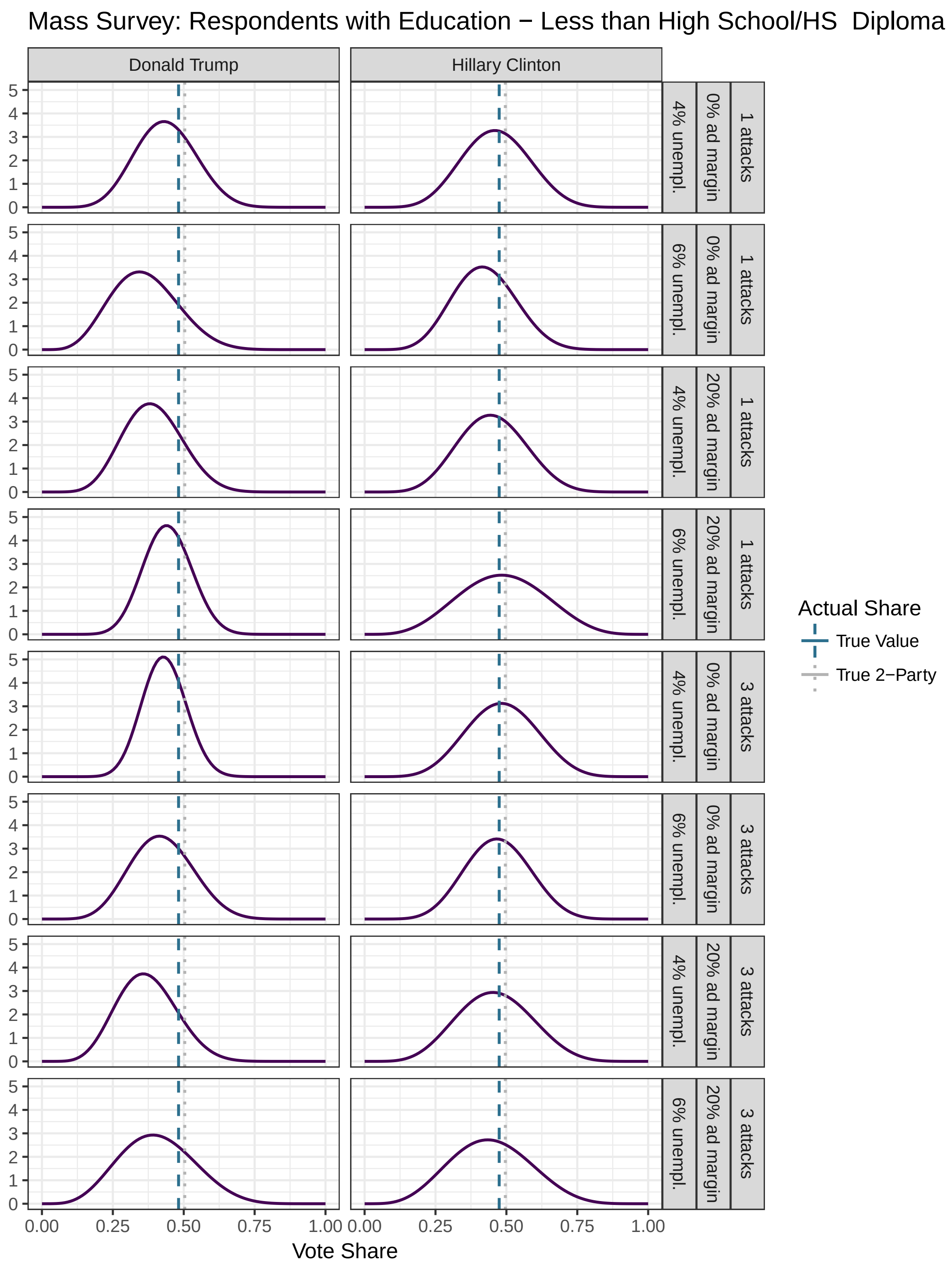


Figure 3.345: Priors with covariates: Mass Pennsylvania Education Less than High School HS Diploma

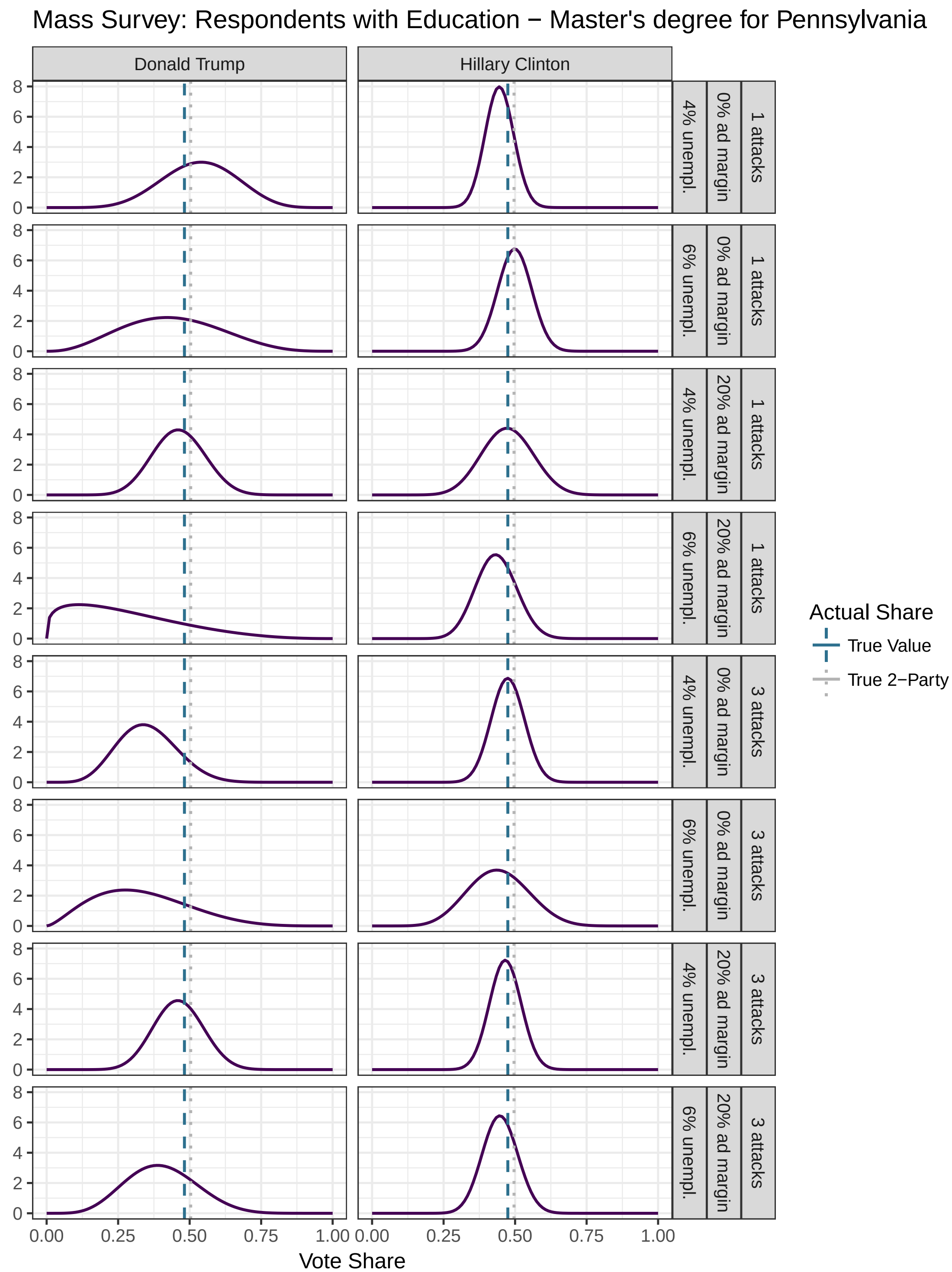


Figure 3.346: Priors with covariates: Mass Pennsylvania Education Master's degree

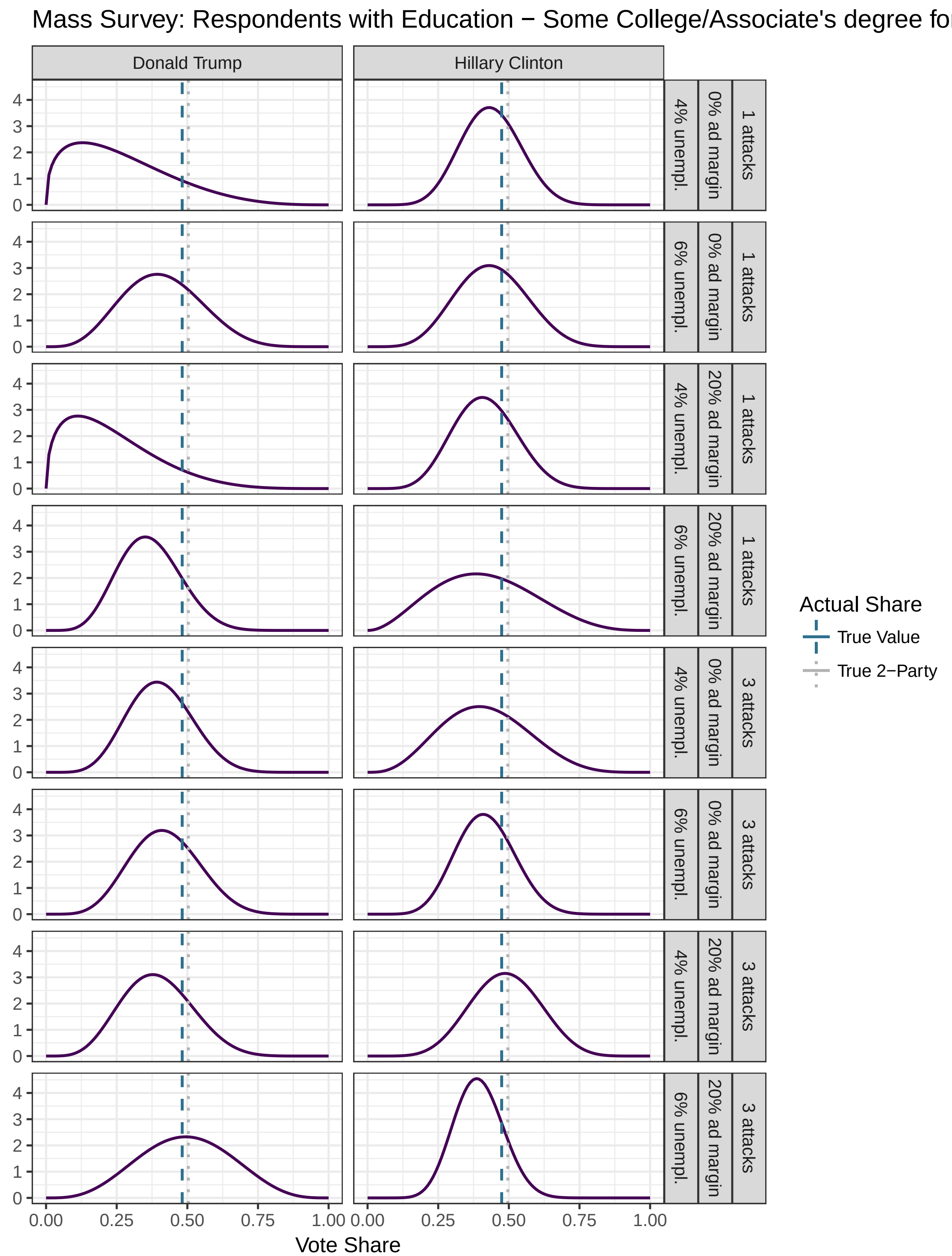


Figure 3.347: Priors with covariates: Mass Pennsylvania Education Some College Associate's degree

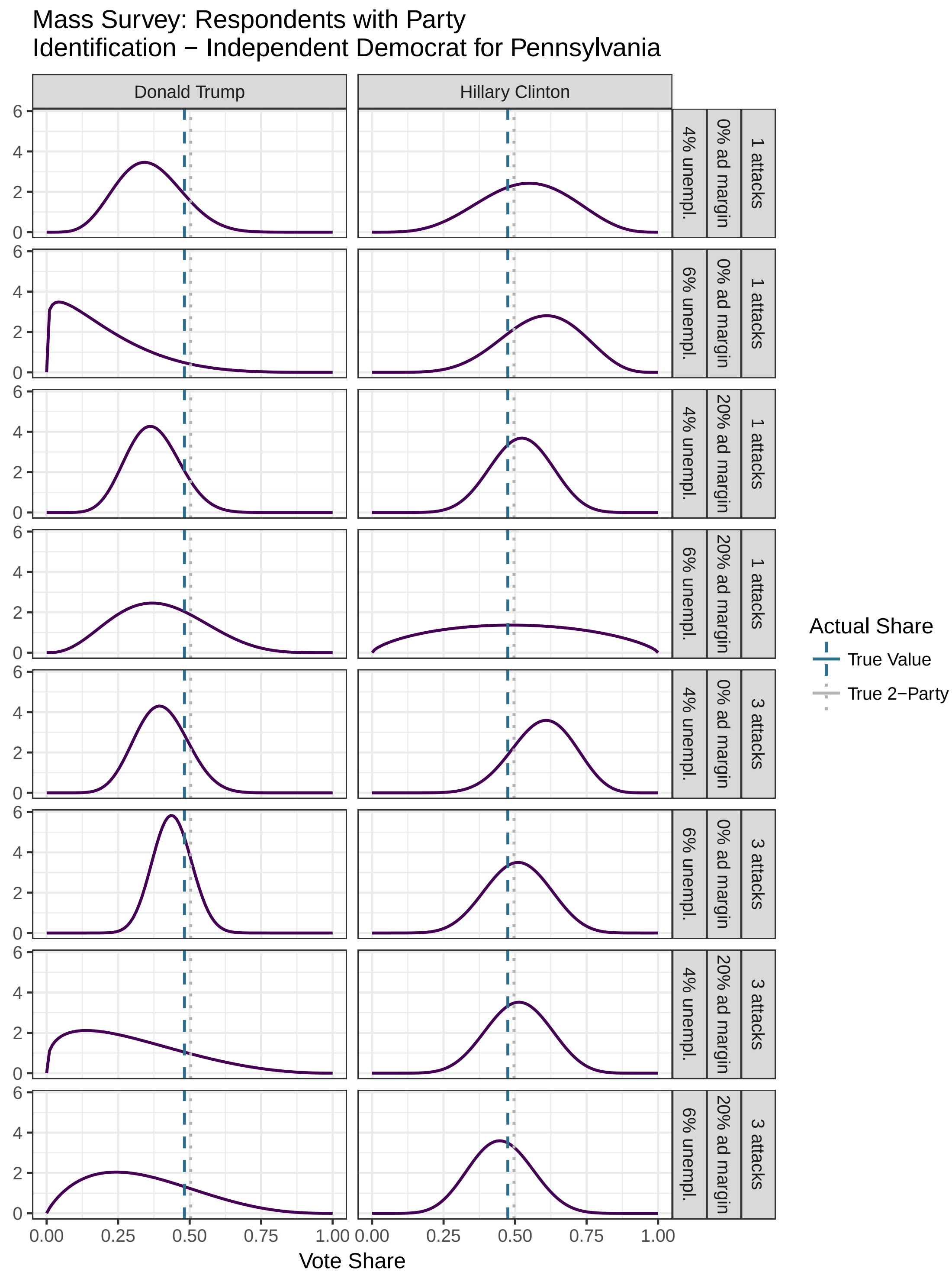


Figure 3.348: Priors with covariates: Mass Pennsylvania Party Identification Independent Democrat

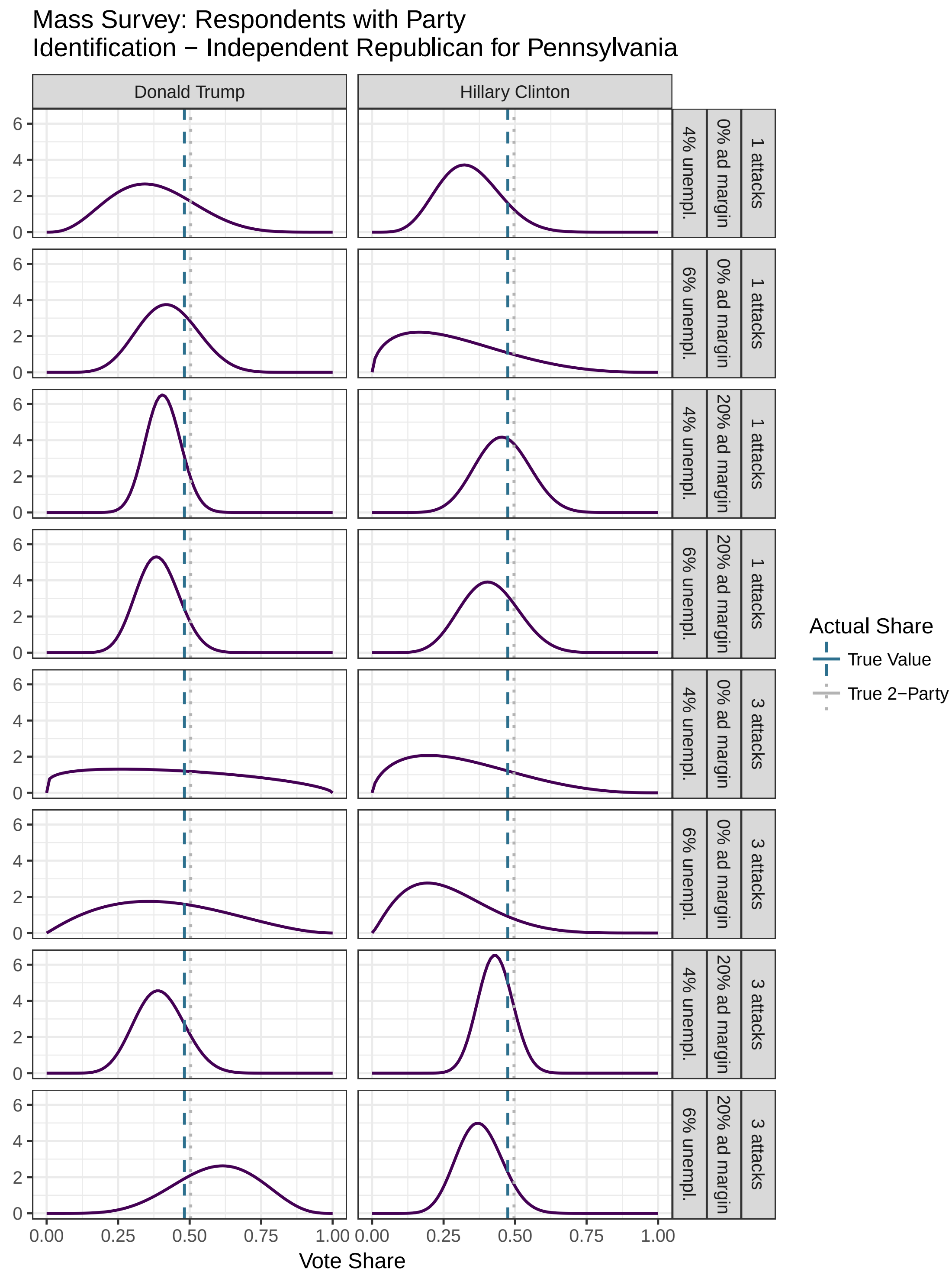


Figure 3.349: Priors with covariates: Mass Pennsylvania Party Identification Independent Republican

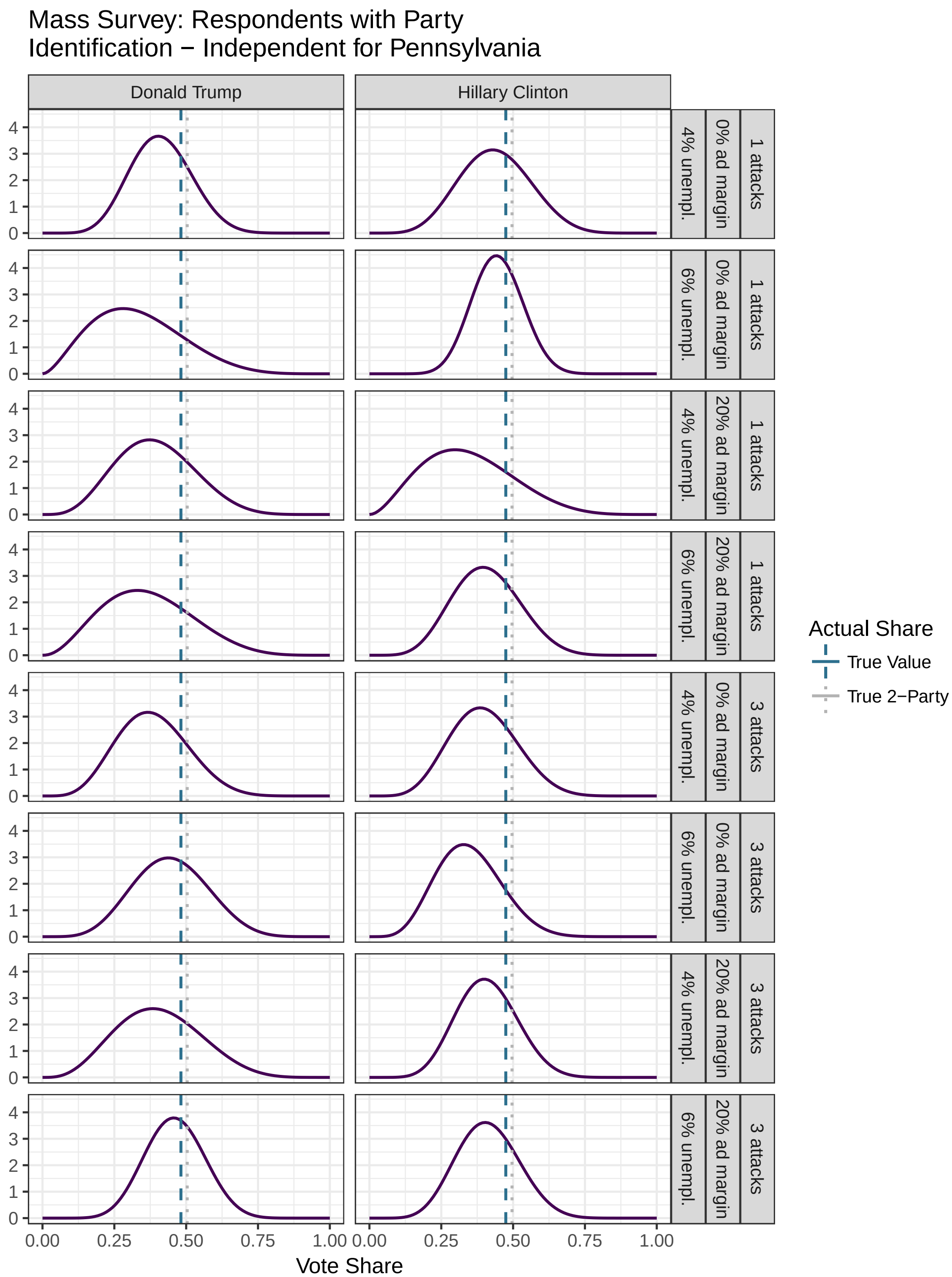


Figure 3.350: Priors with covariates: Mass Pennsylvania Party Identification Independent

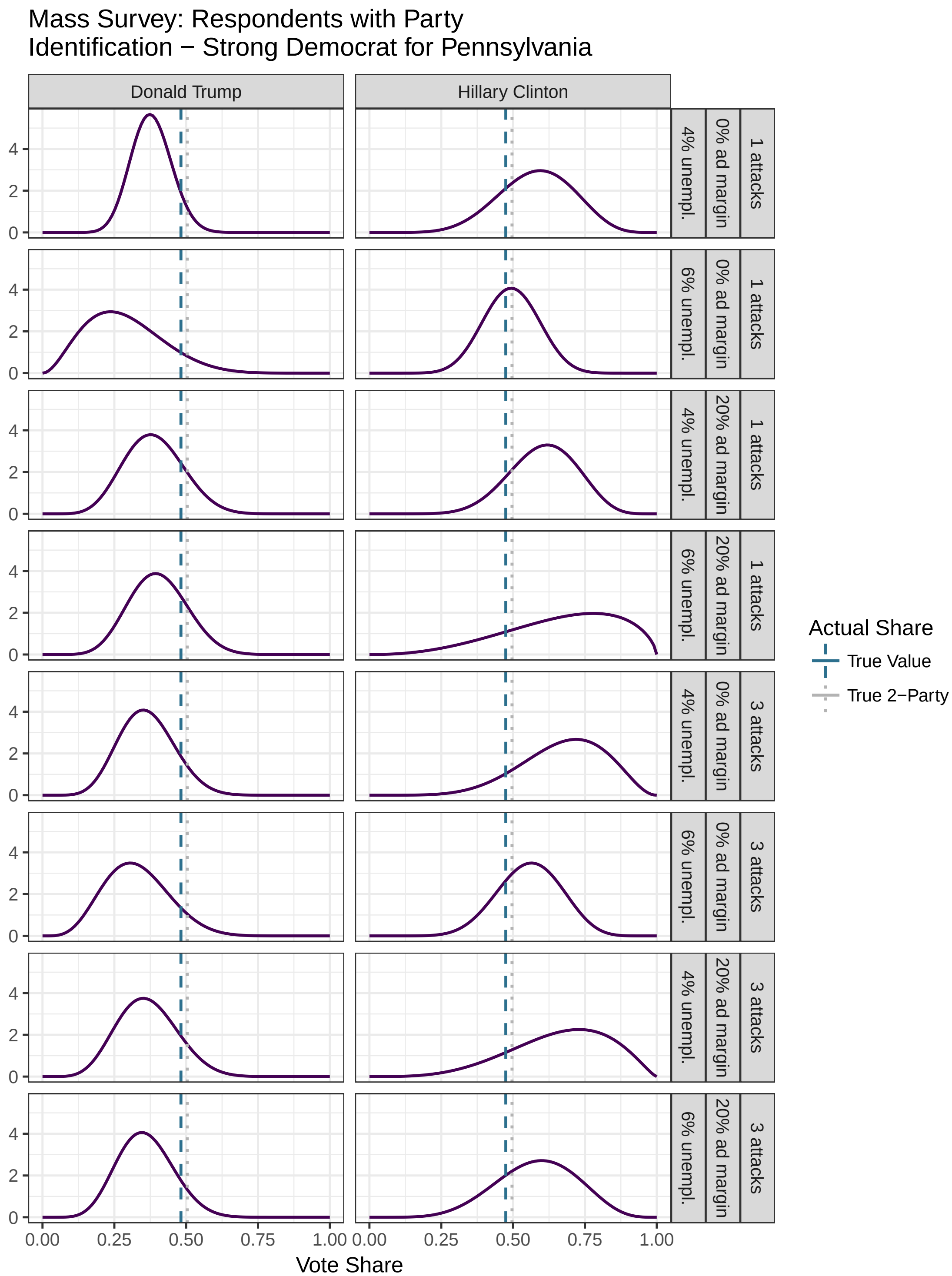


Figure 3.351: Priors with covariates: Mass Pennsylvania Party Identification Strong Democrat

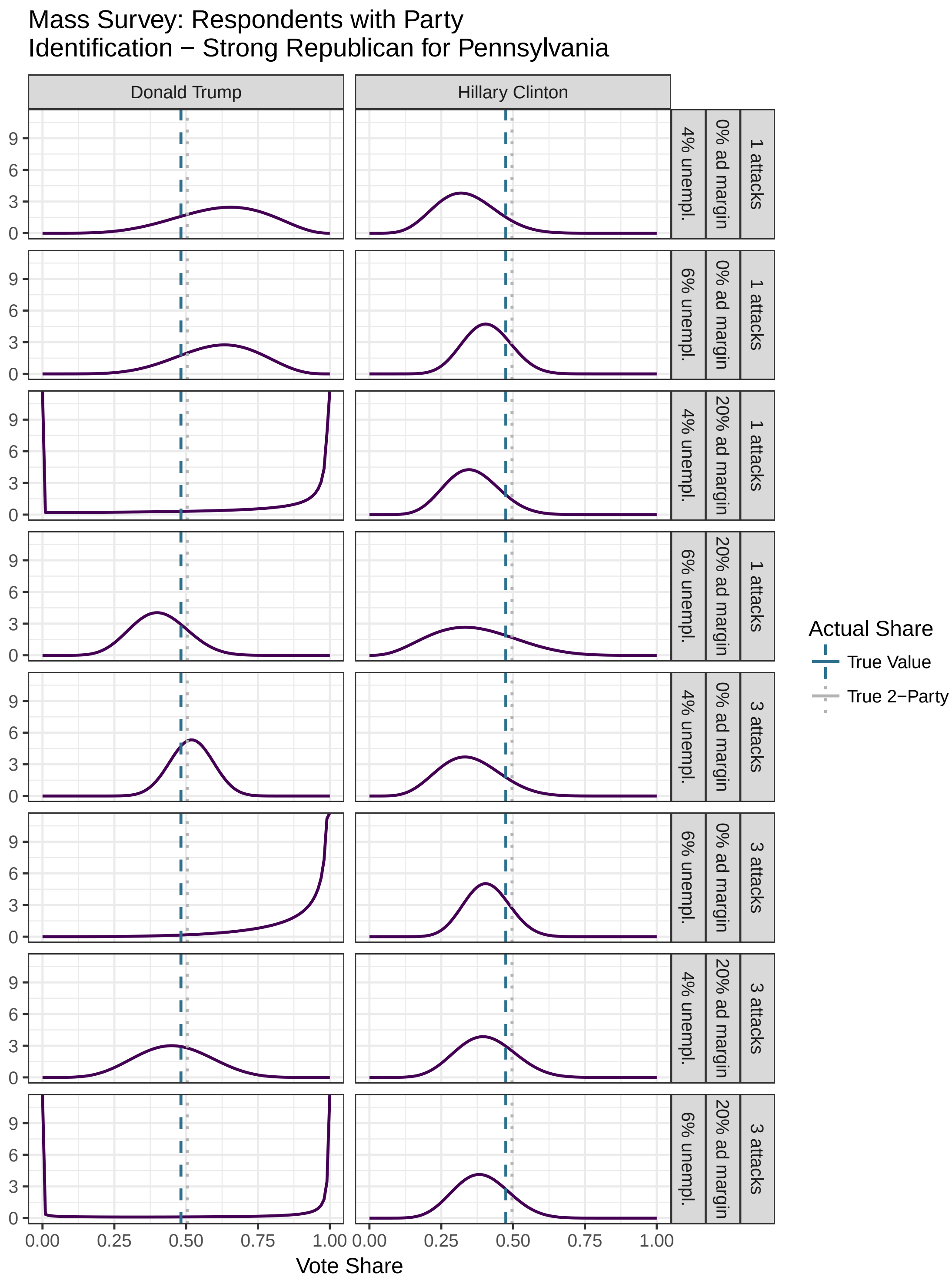


Figure 3.352: Priors with covariates: Mass Pennsylvania Party Identification Strong Republican

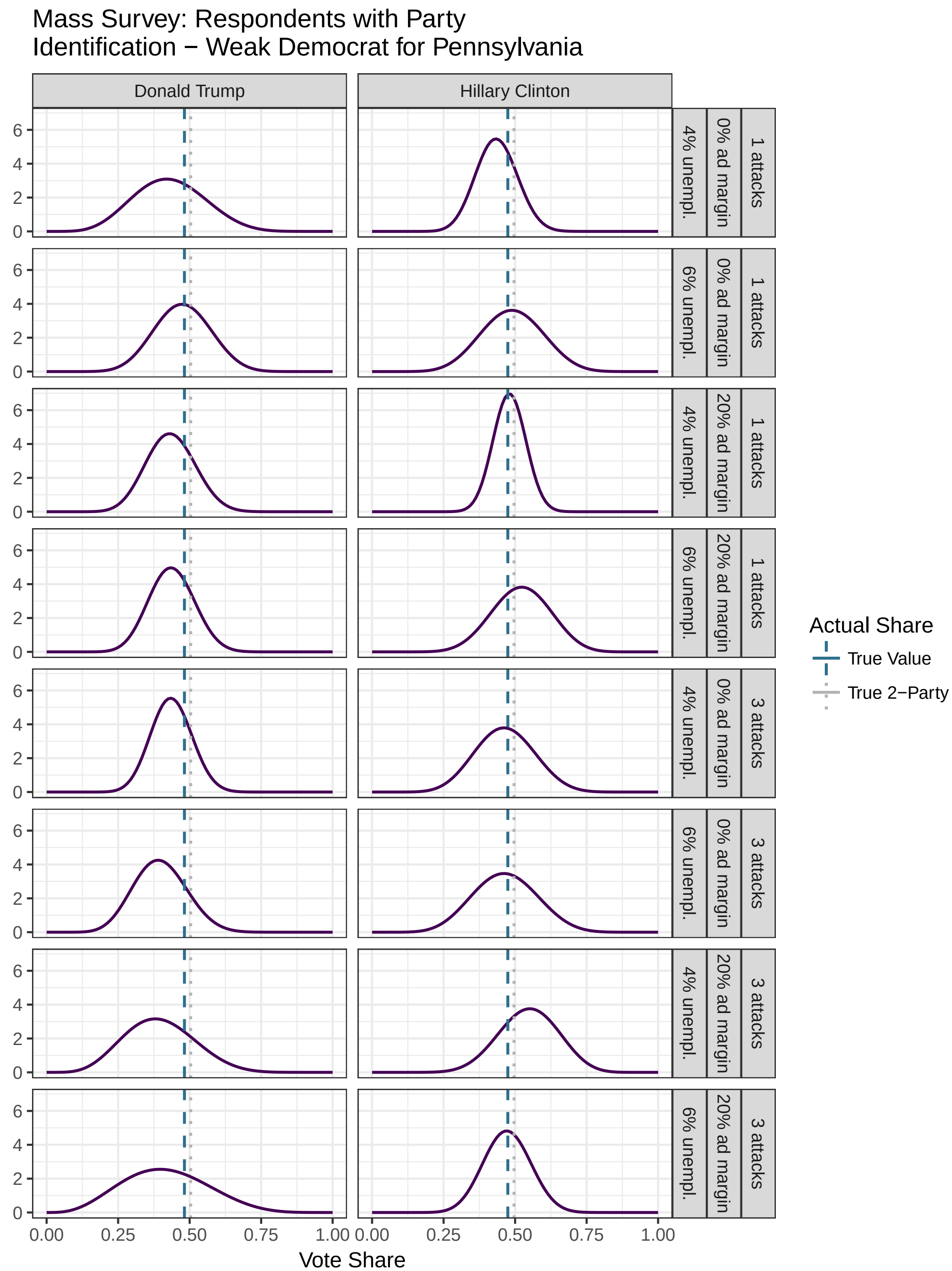


Figure 3.353: Priors with covariates: Mass Pennsylvania Party Identification Weak Democrat

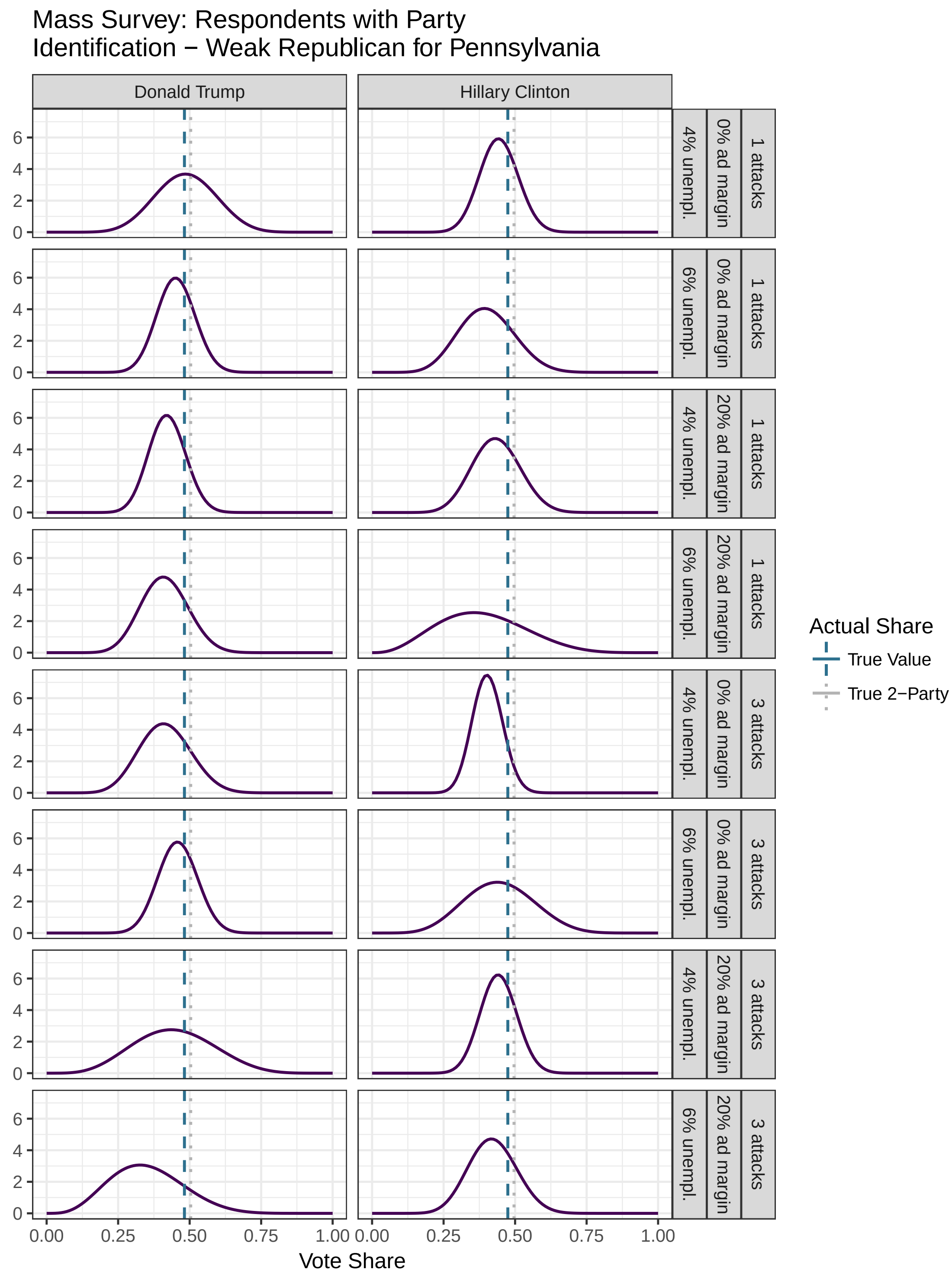


Figure 3.354: Priors with covariates: Mass Pennsylvania Party Identification Weak Republican

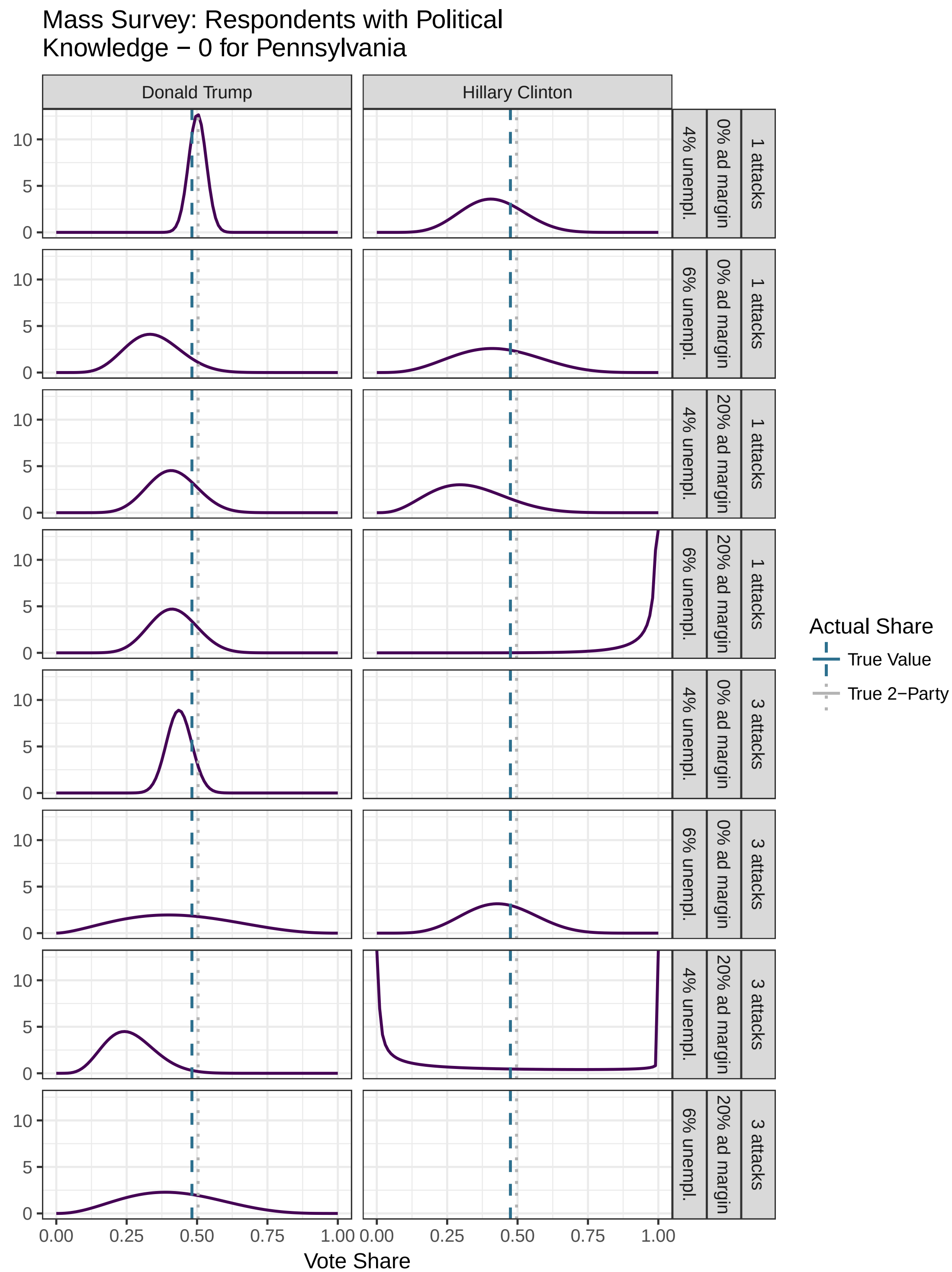


Figure 3.355: Priors with covariates: Mass Pennsylvania Political Knowledge 0

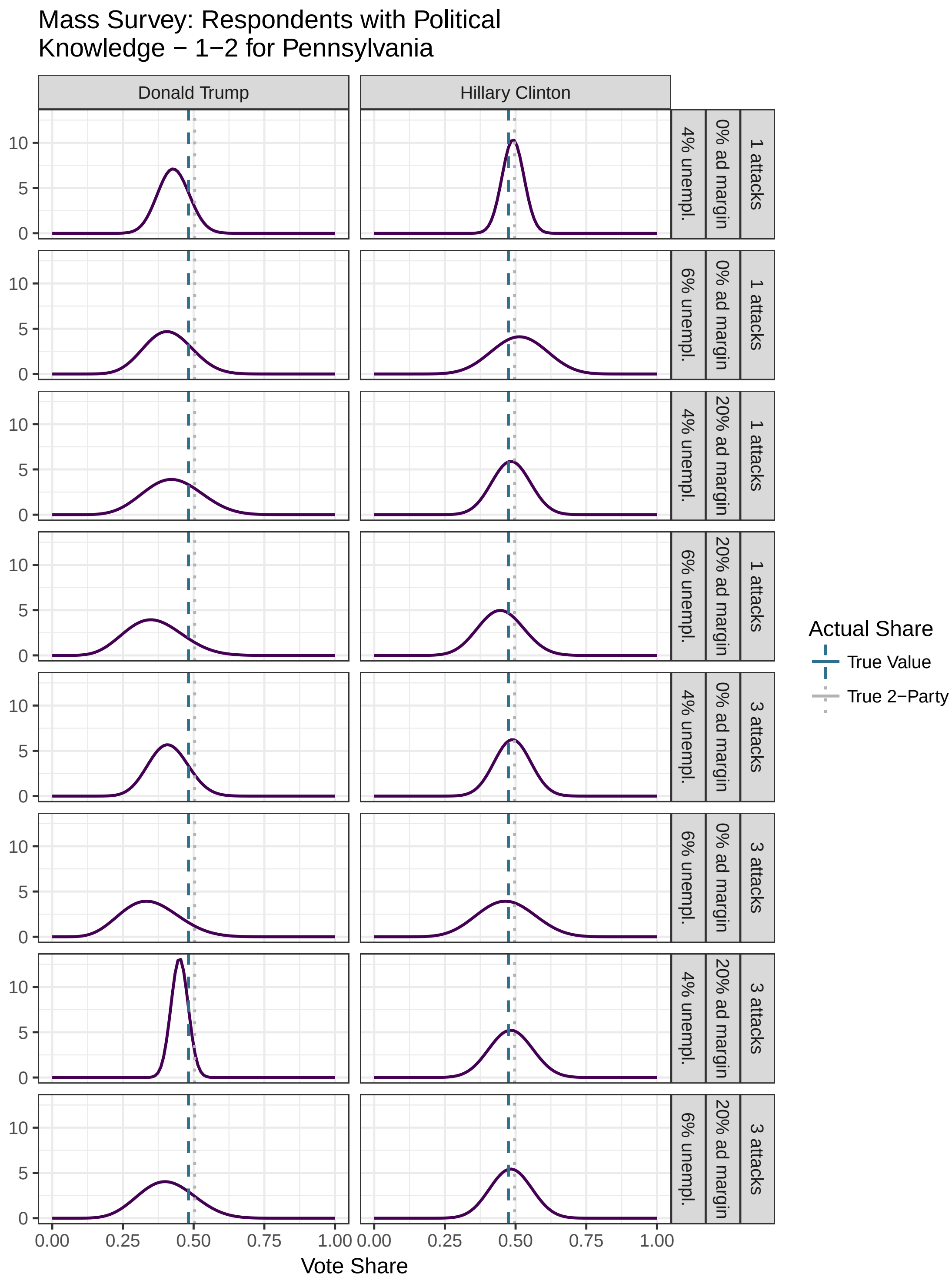


Figure 3.356: Priors with covariates: Mass Pennsylvania Political Knowledge 1-2

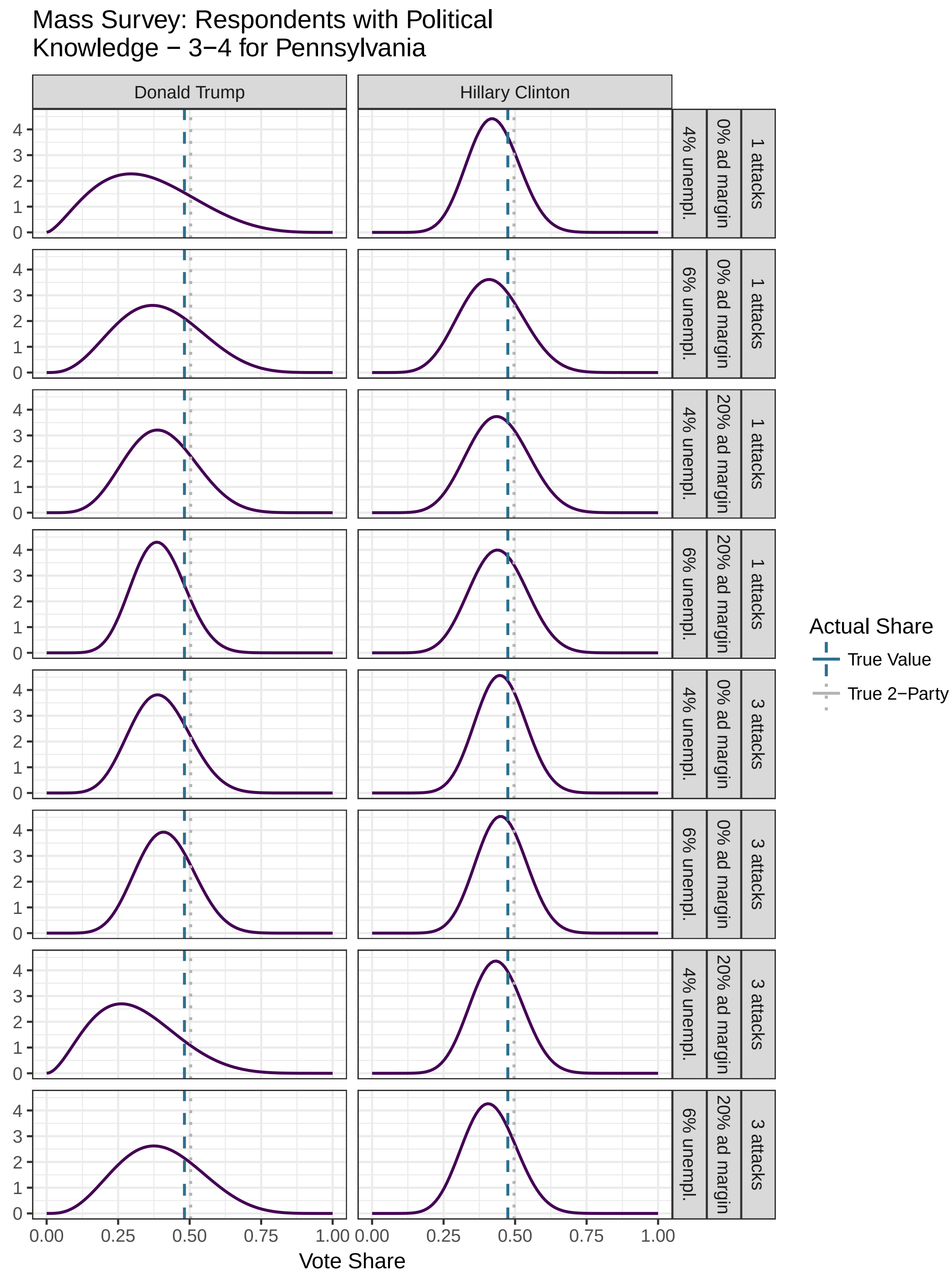


Figure 3.357: Priors with covariates: Mass Pennsylvania Political Knowledge 3-4

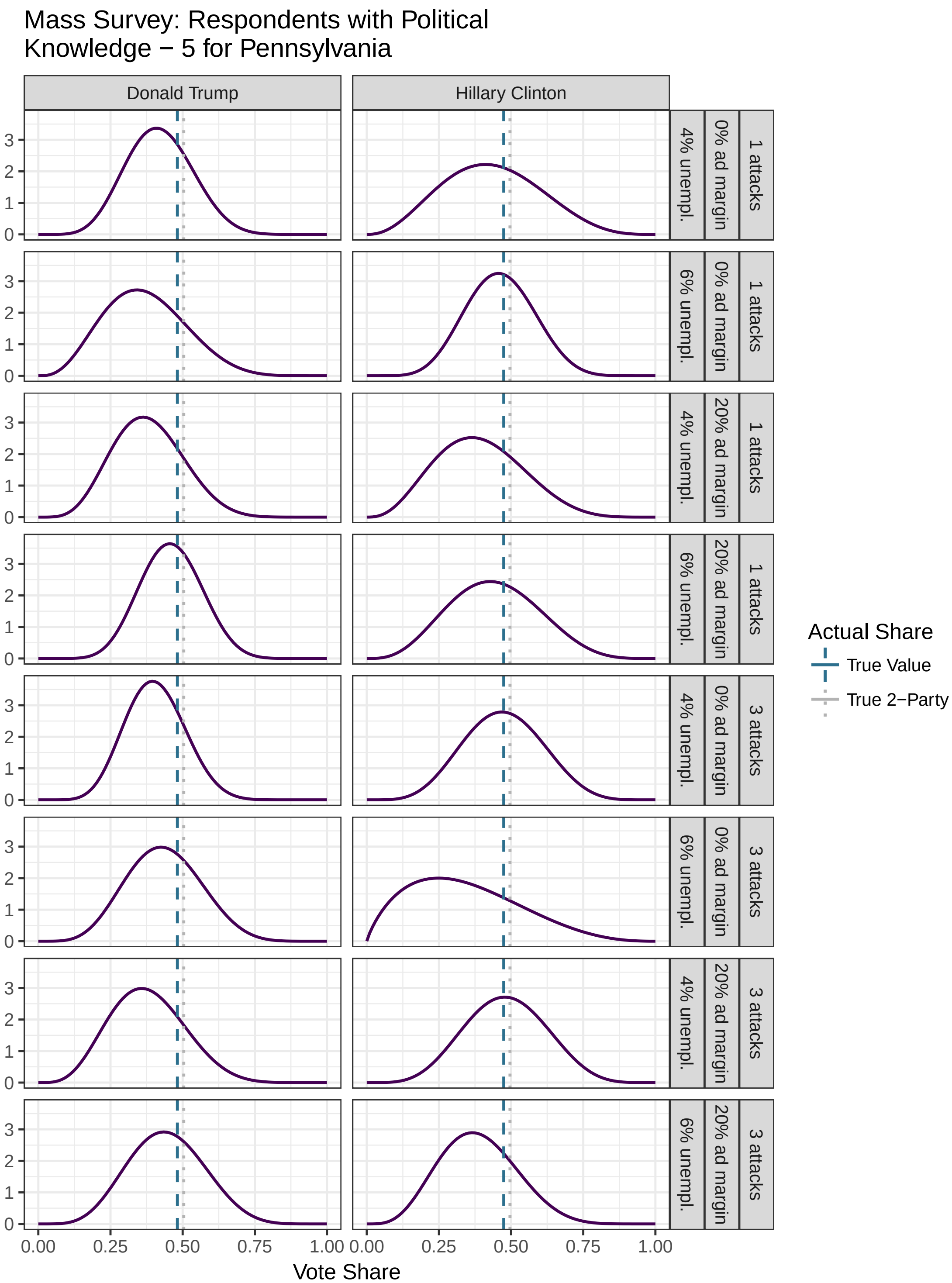


Figure 3.358: Priors with covariates: Mass Pennsylvania Political Knowledge 5

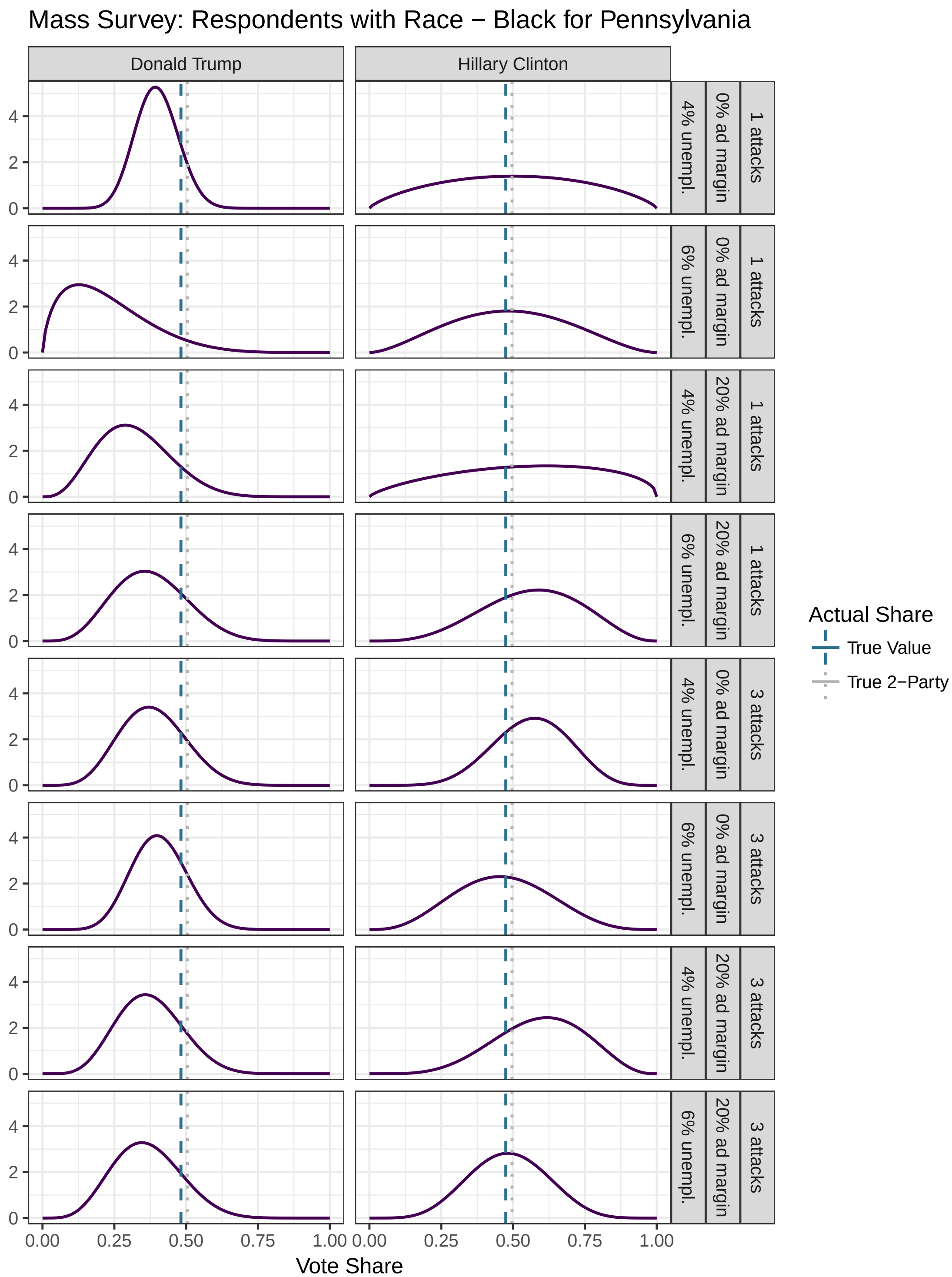


Figure 3.359: Priors with covariates: Mass Pennsylvania Race Black

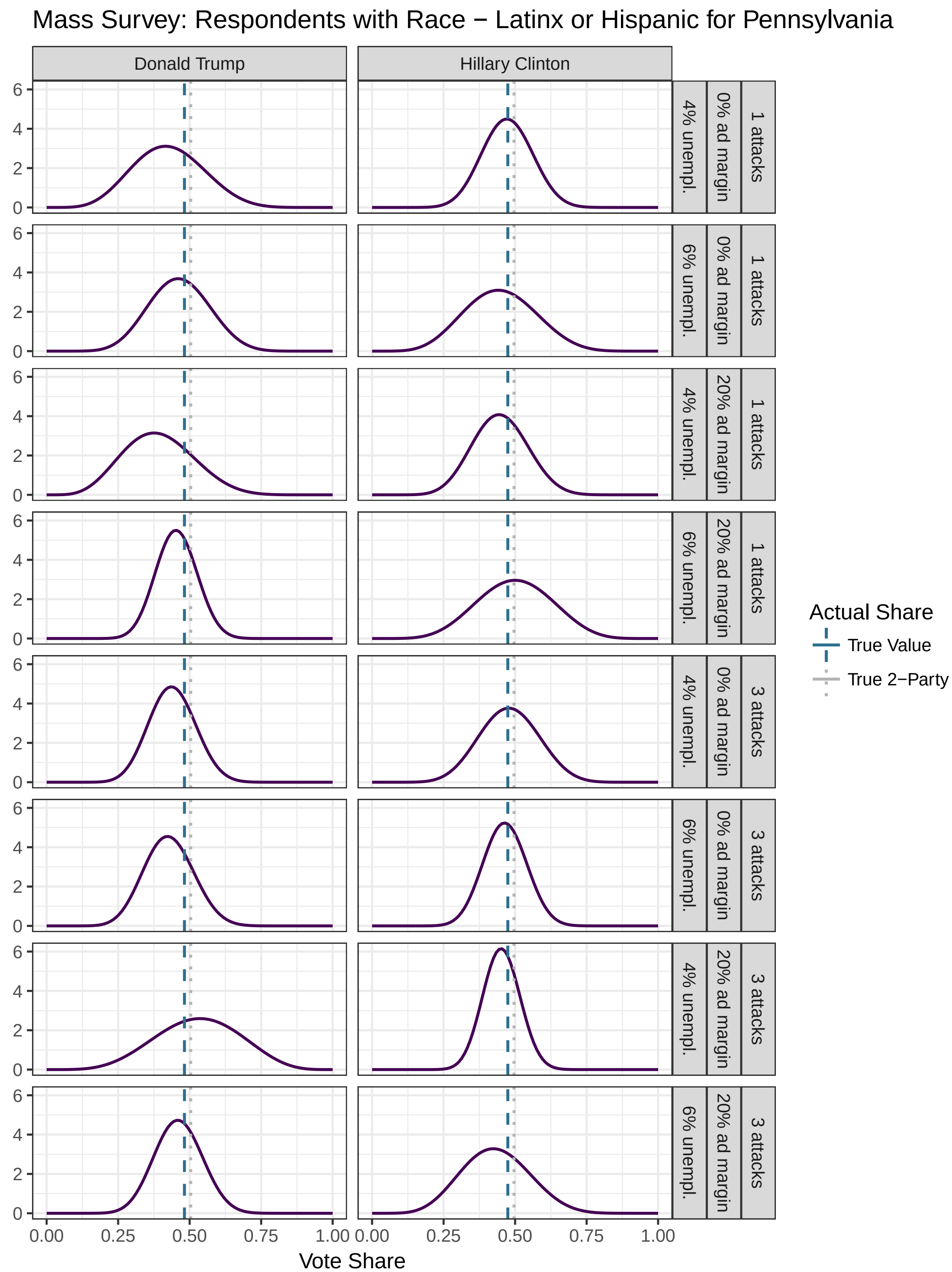


Figure 3.360: Priors with covariates: Mass Pennsylvania Race Latinx or Hispanic

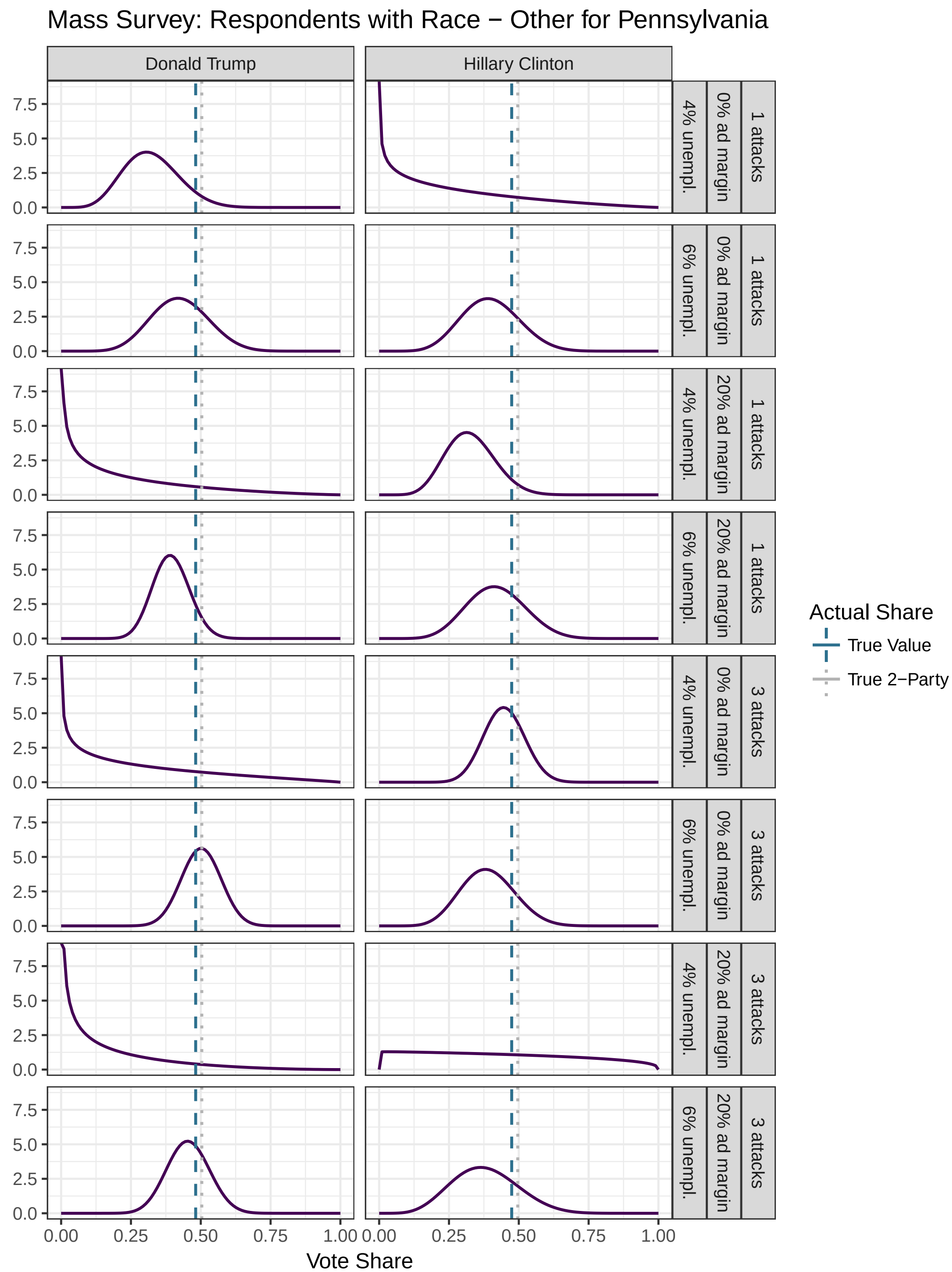


Figure 3.361: Priors with covariates: Mass Pennsylvania Race Other

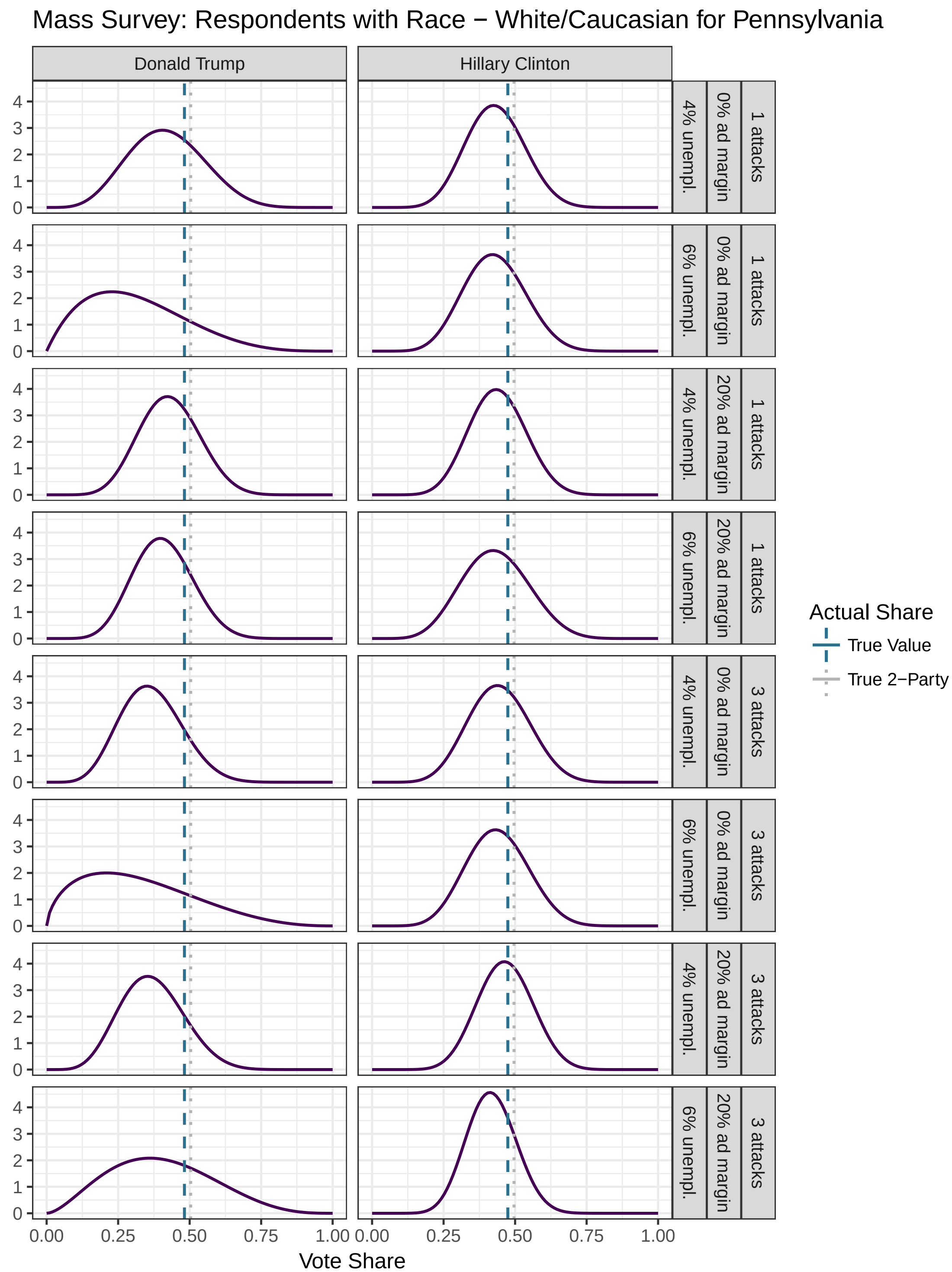


Figure 3.362: Priors with covariates: Mass Pennsylvania Race White Caucasian

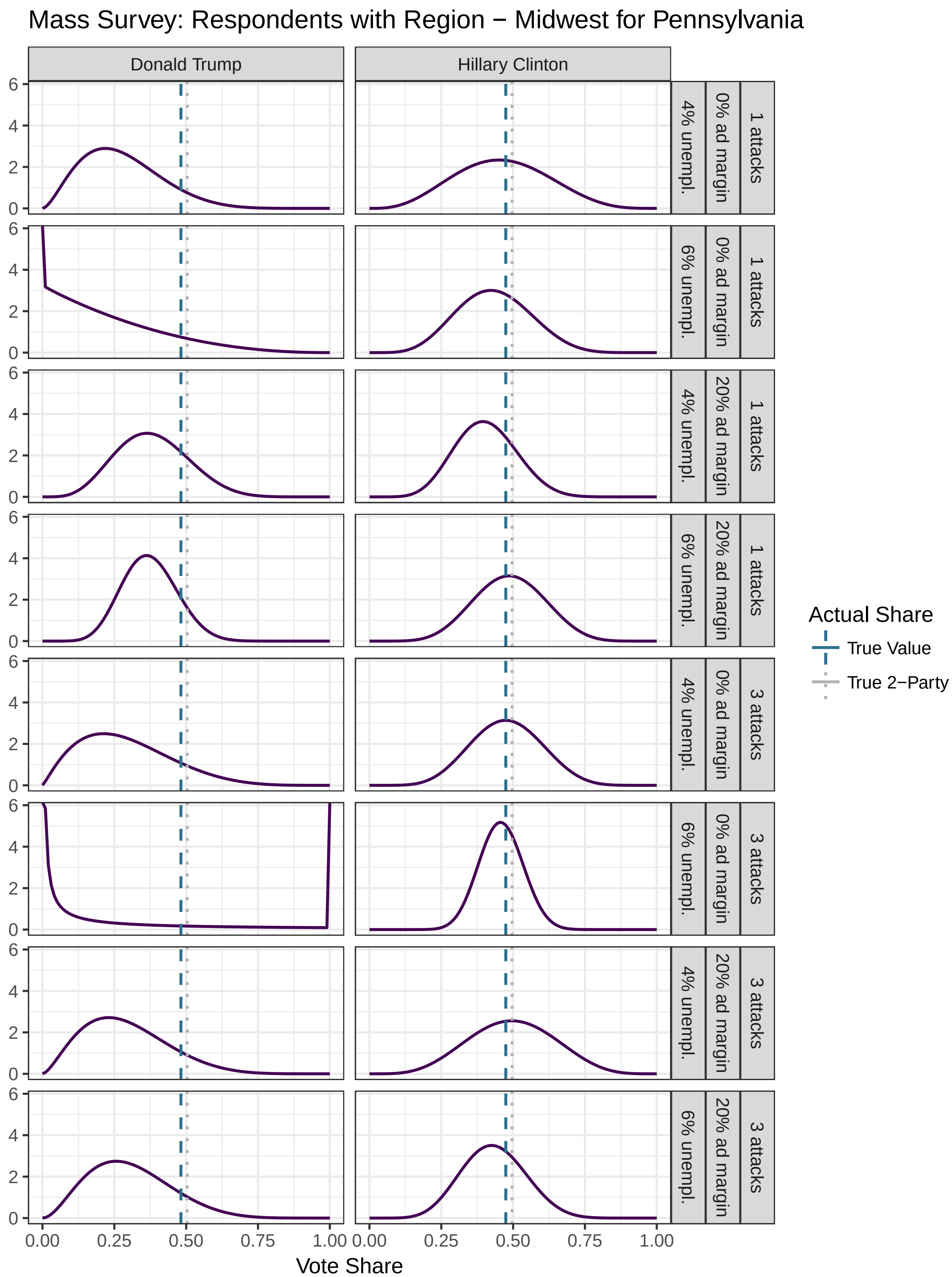


Figure 3.363: Priors with covariates: Mass Pennsylvania Region Midwest

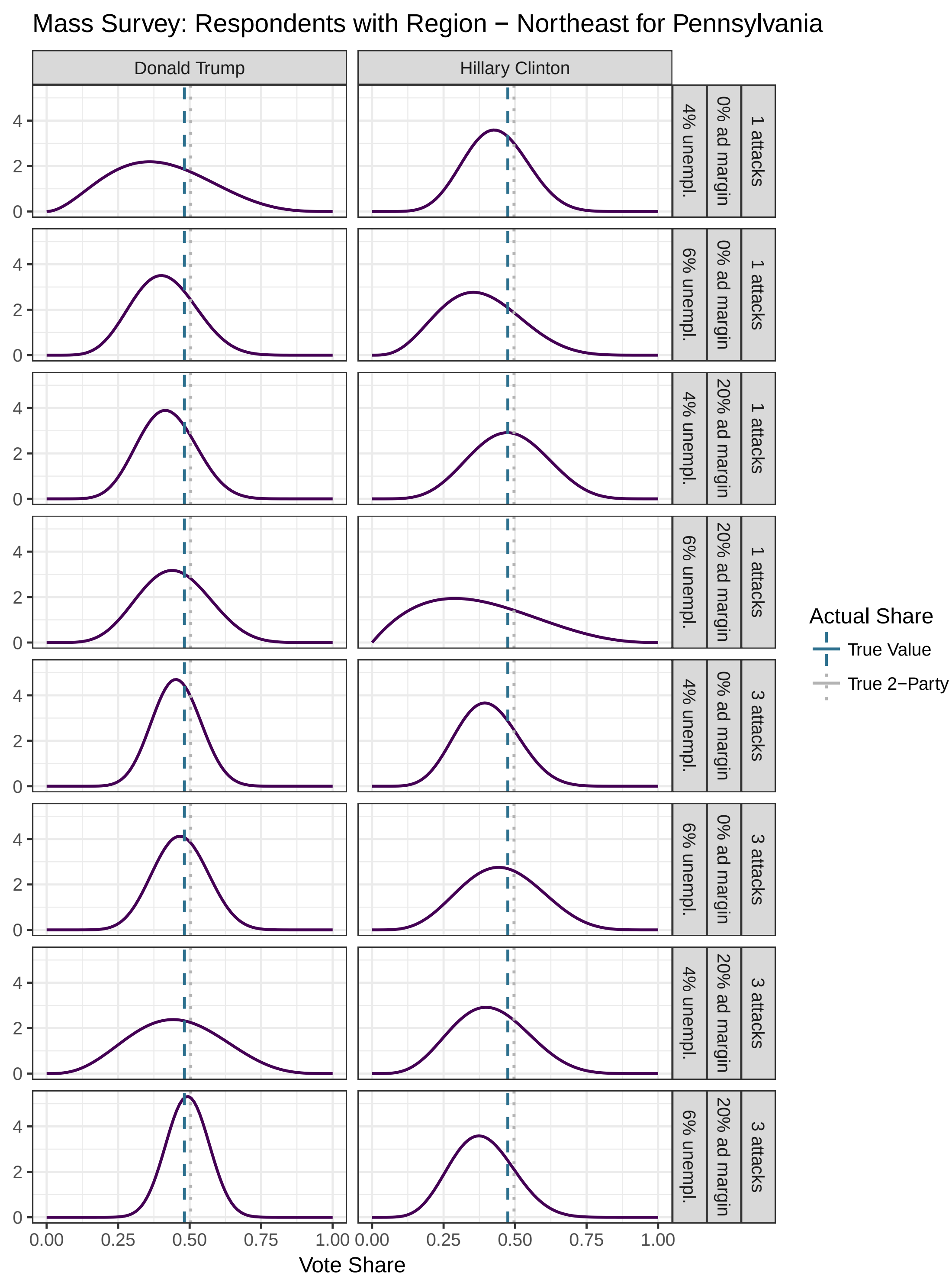


Figure 3.364: Priors with covariates: Mass Pennsylvania Region Northeast

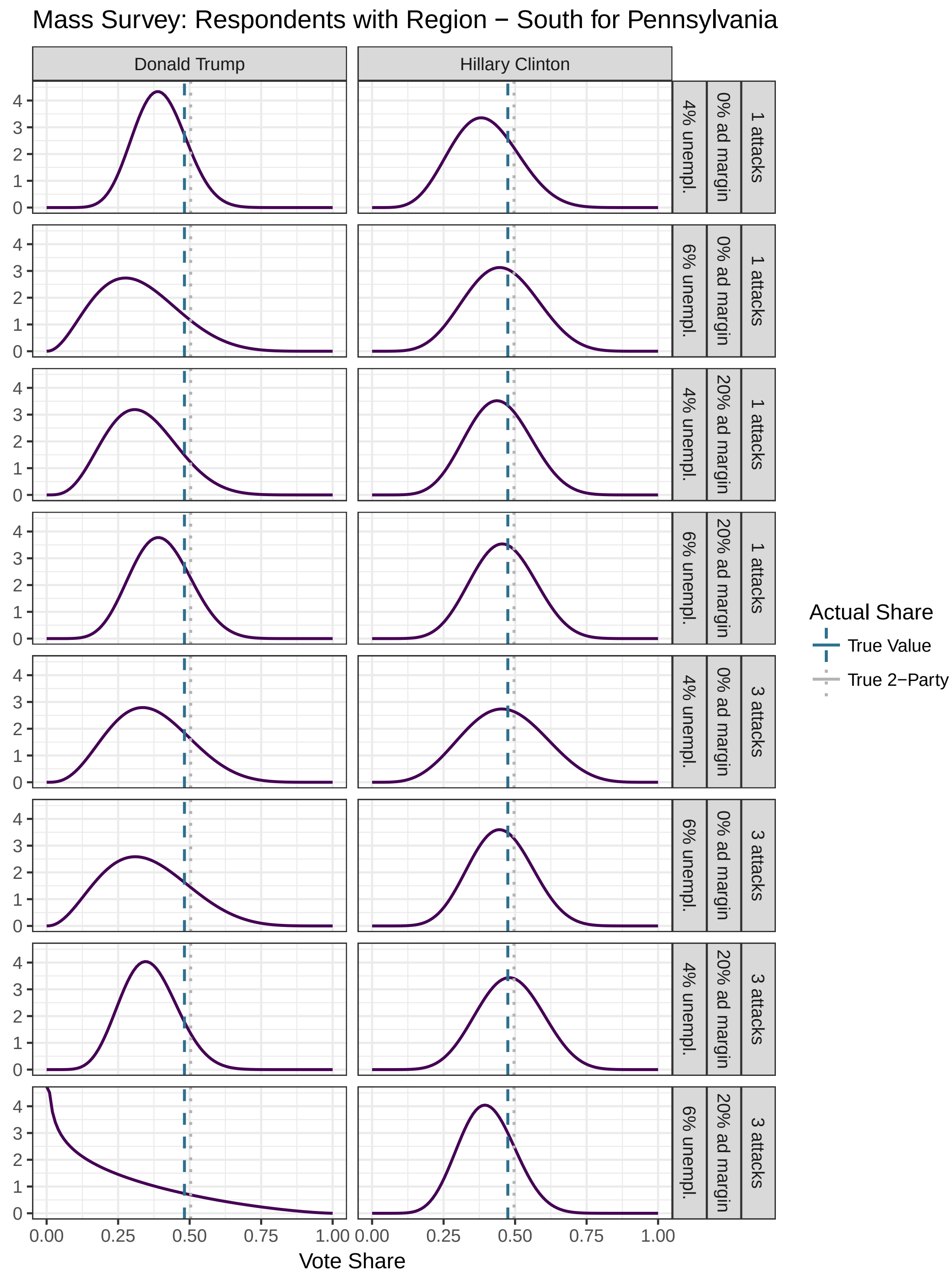


Figure 3.365: Priors with covariates: Mass Pennsylvania Region South

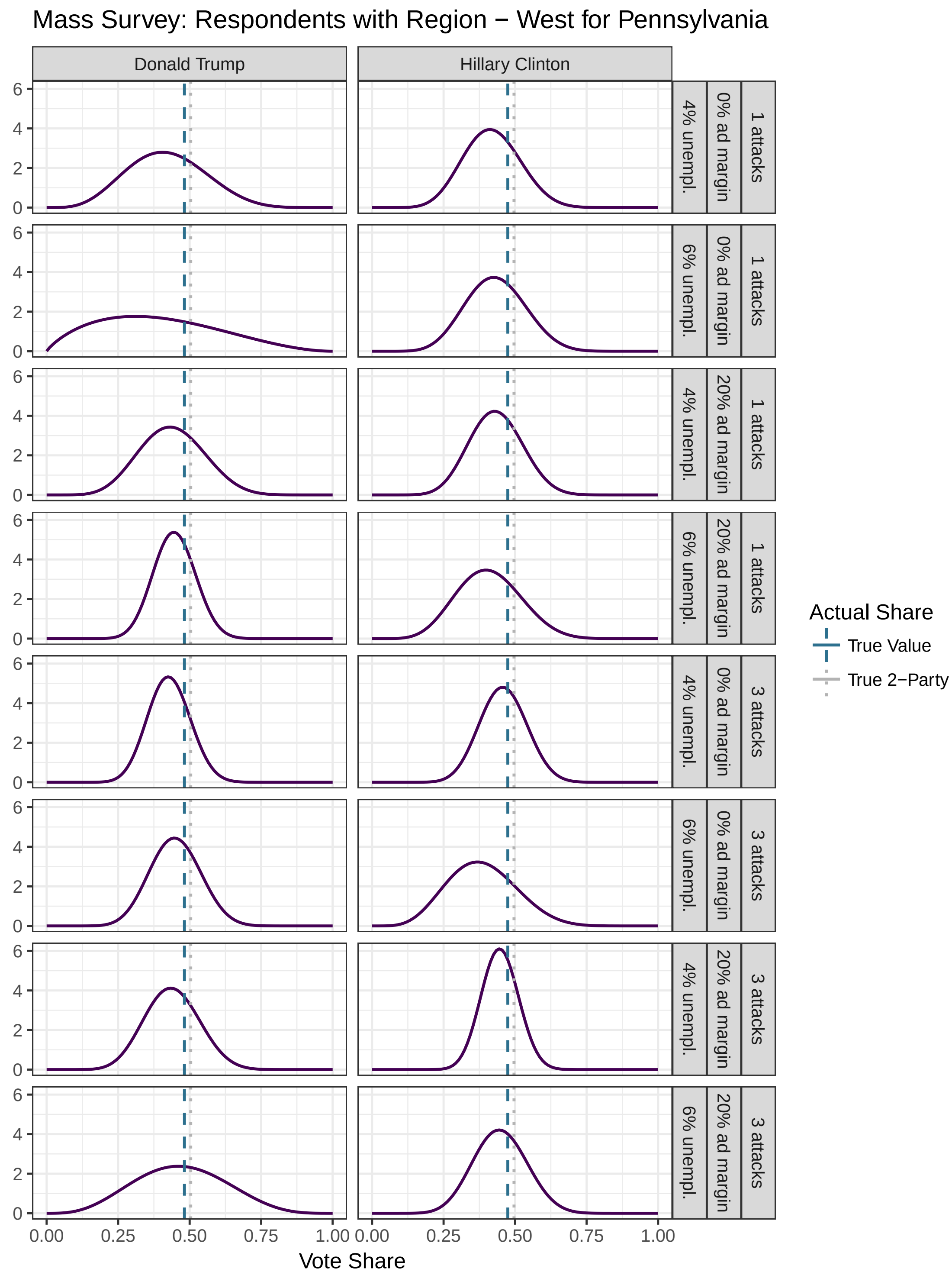


Figure 3.366: Priors with covariates: Mass Pennsylvania Region West

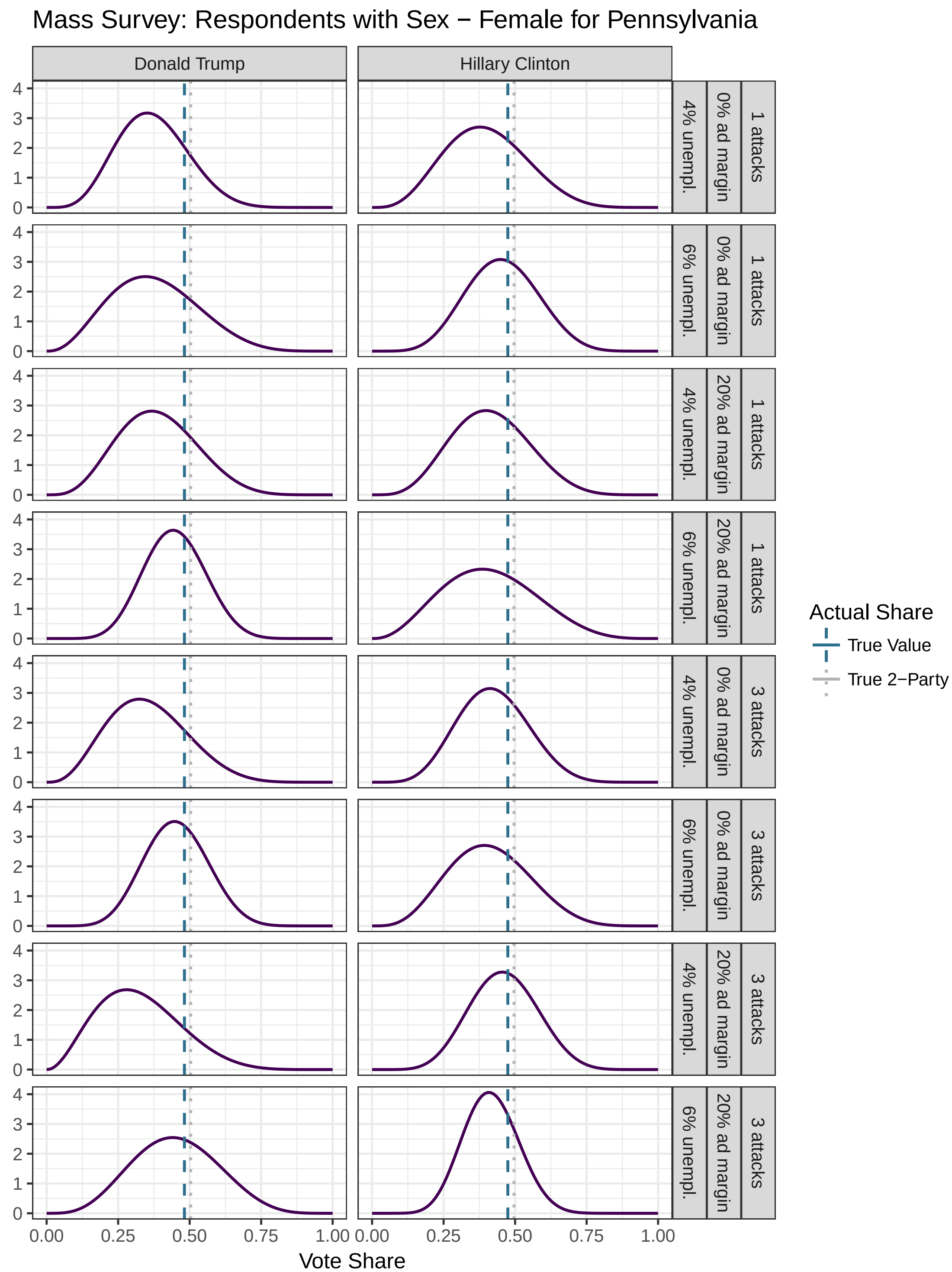


Figure 3.367: Priors with covariates: Mass Pennsylvania Sex Female

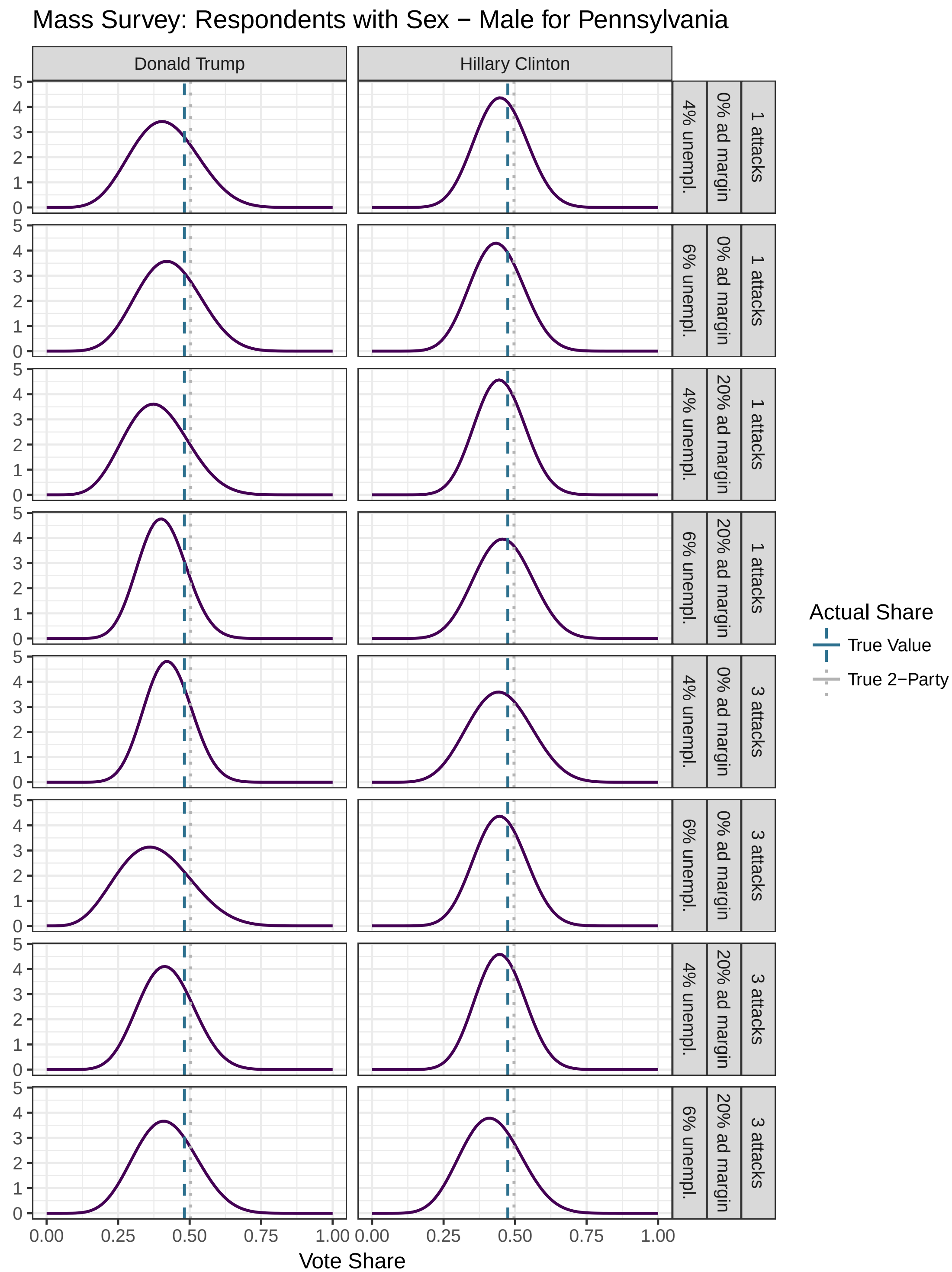


Figure 3.368: Priors with covariates: Mass Pennsylvania Sex Male

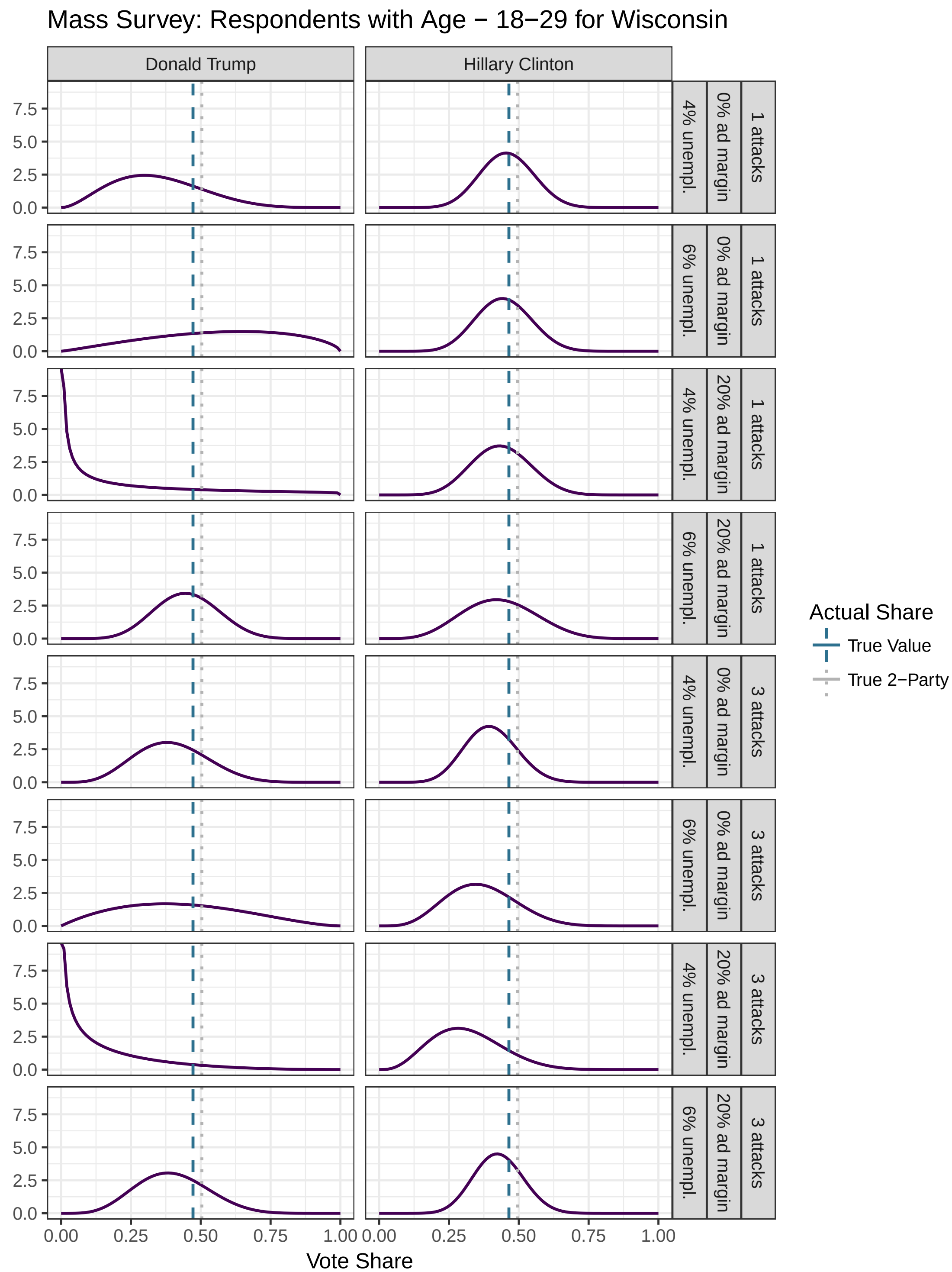


Figure 3.369: Priors with covariates: Mass Wisconsin Age 18-29

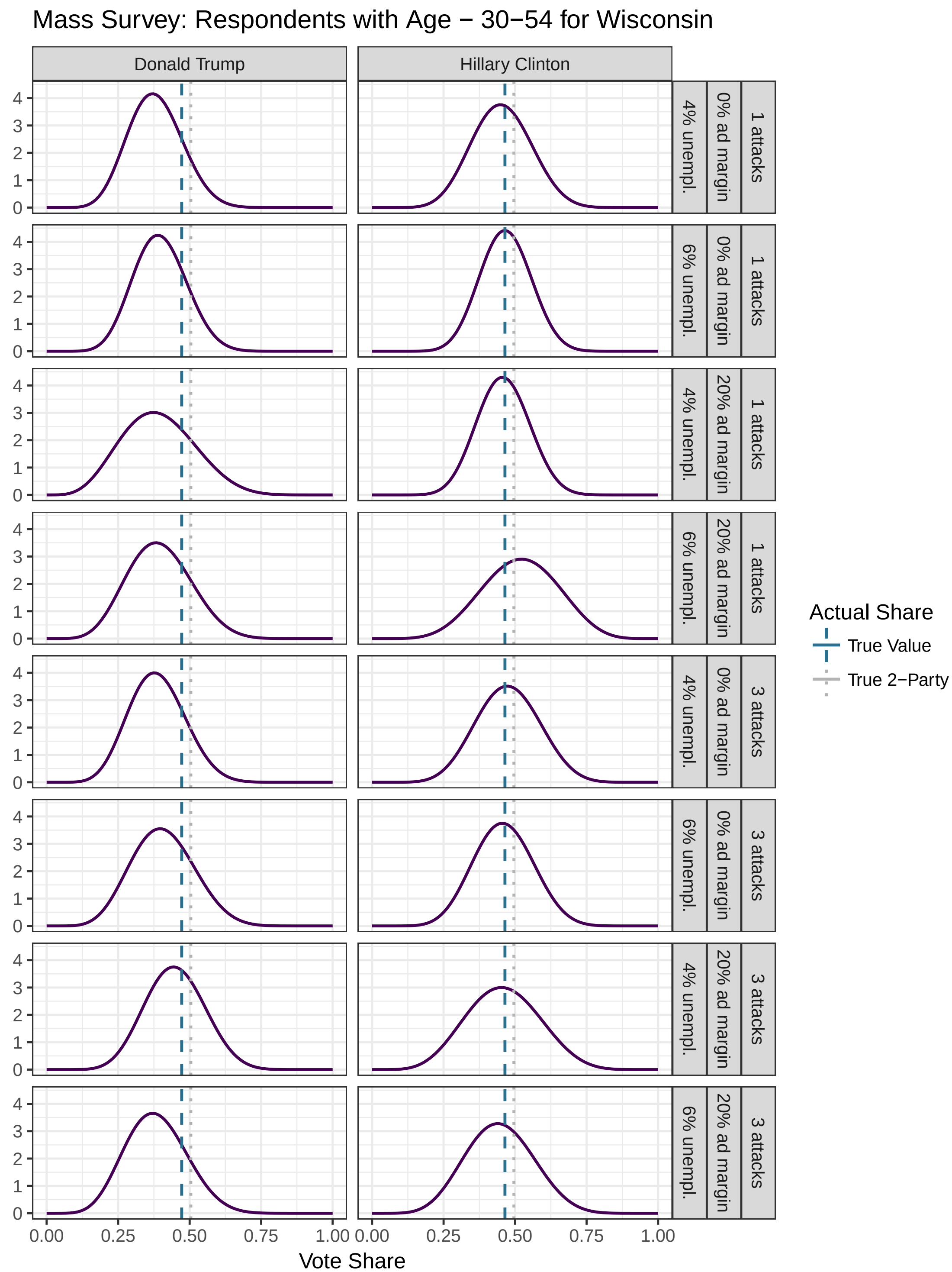


Figure 3.370: Priors with covariates: Mass Wisconsin Age 30-54

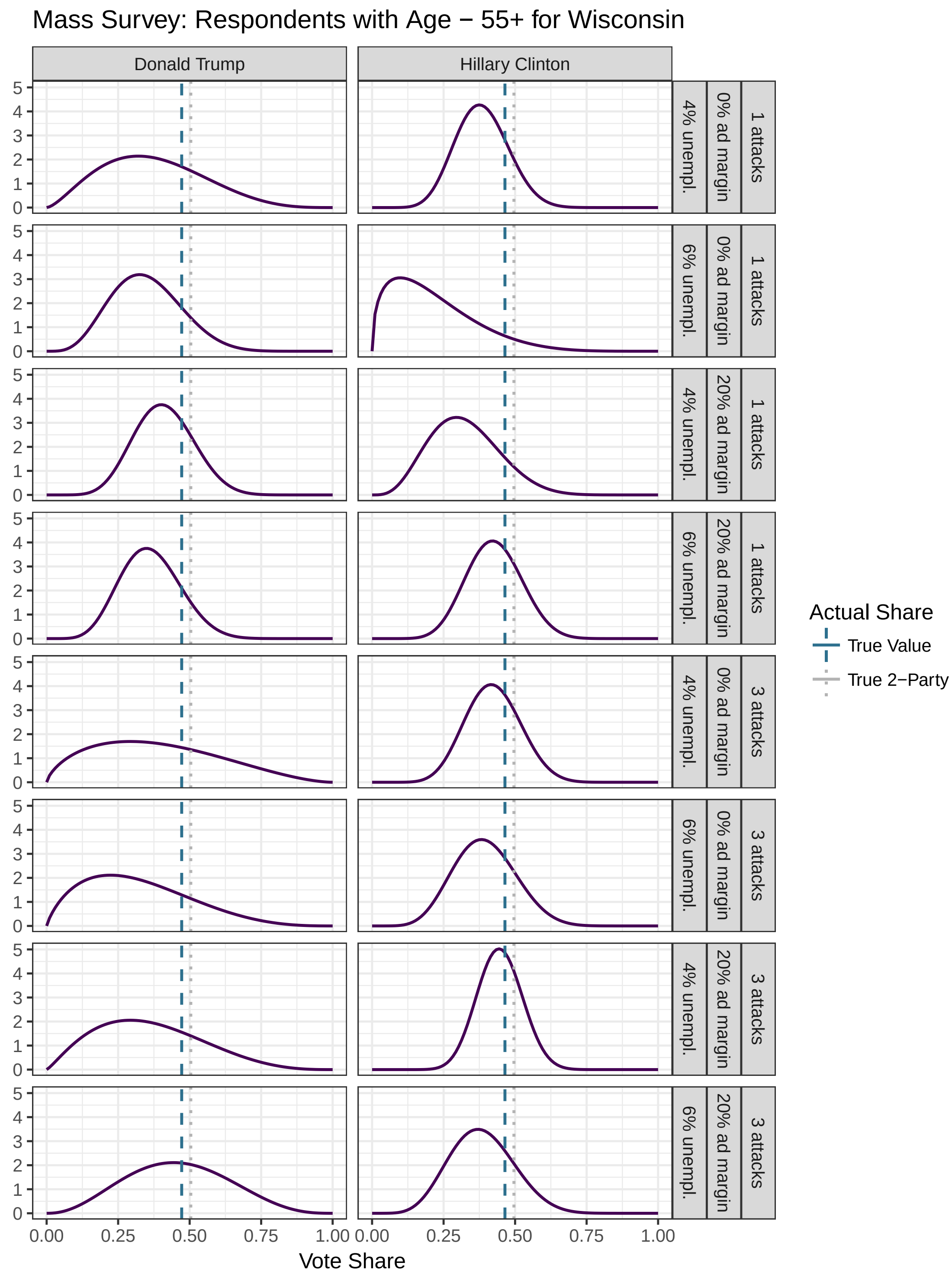


Figure 3.371: Priors with covariates: Mass Wisconsin Age 55+

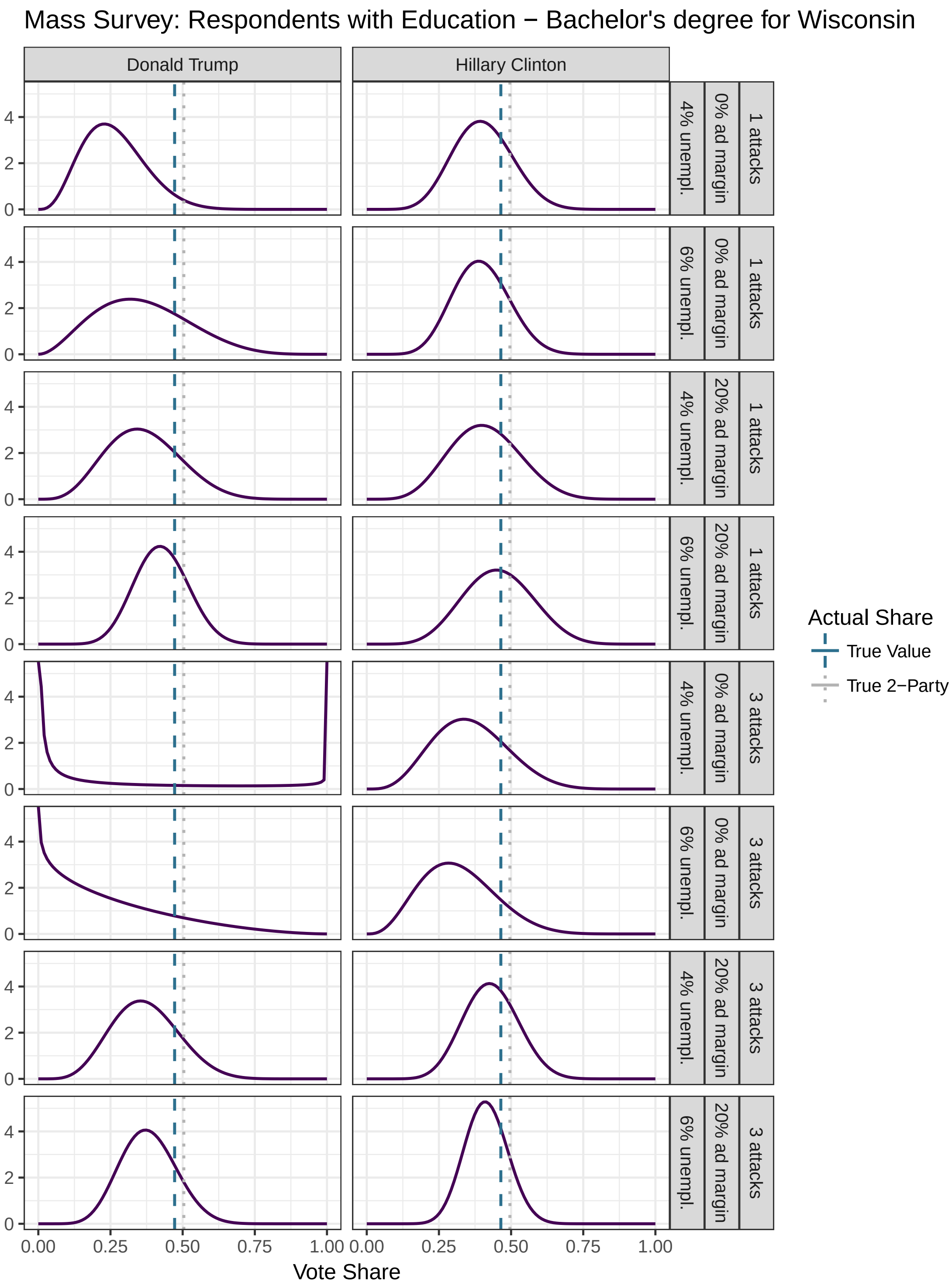


Figure 3.372: Priors with covariates: Mass Wisconsin Education Bachelor's degree

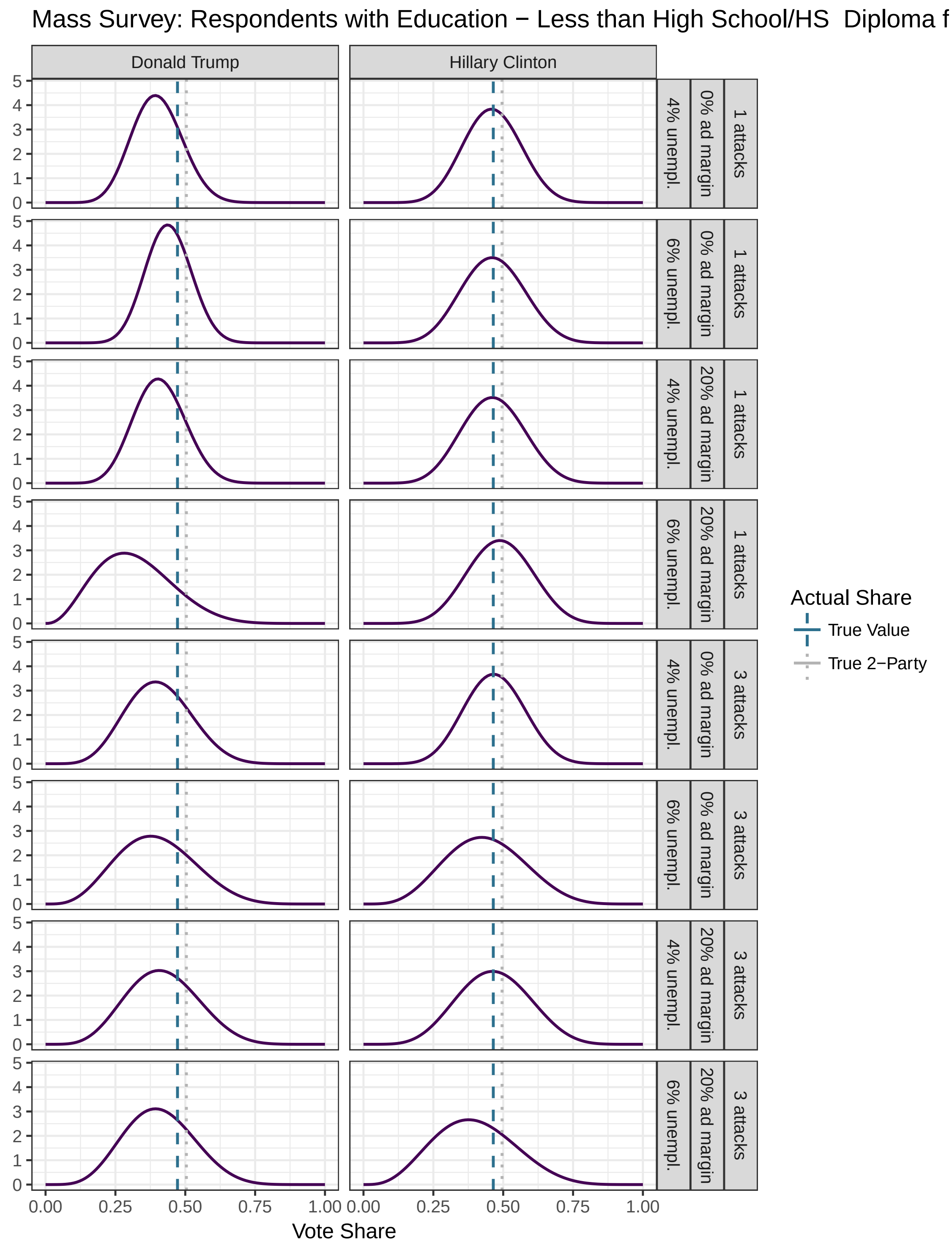


Figure 3.373: Priors with covariates: Mass Wisconsin Education Less than High School HS Diploma

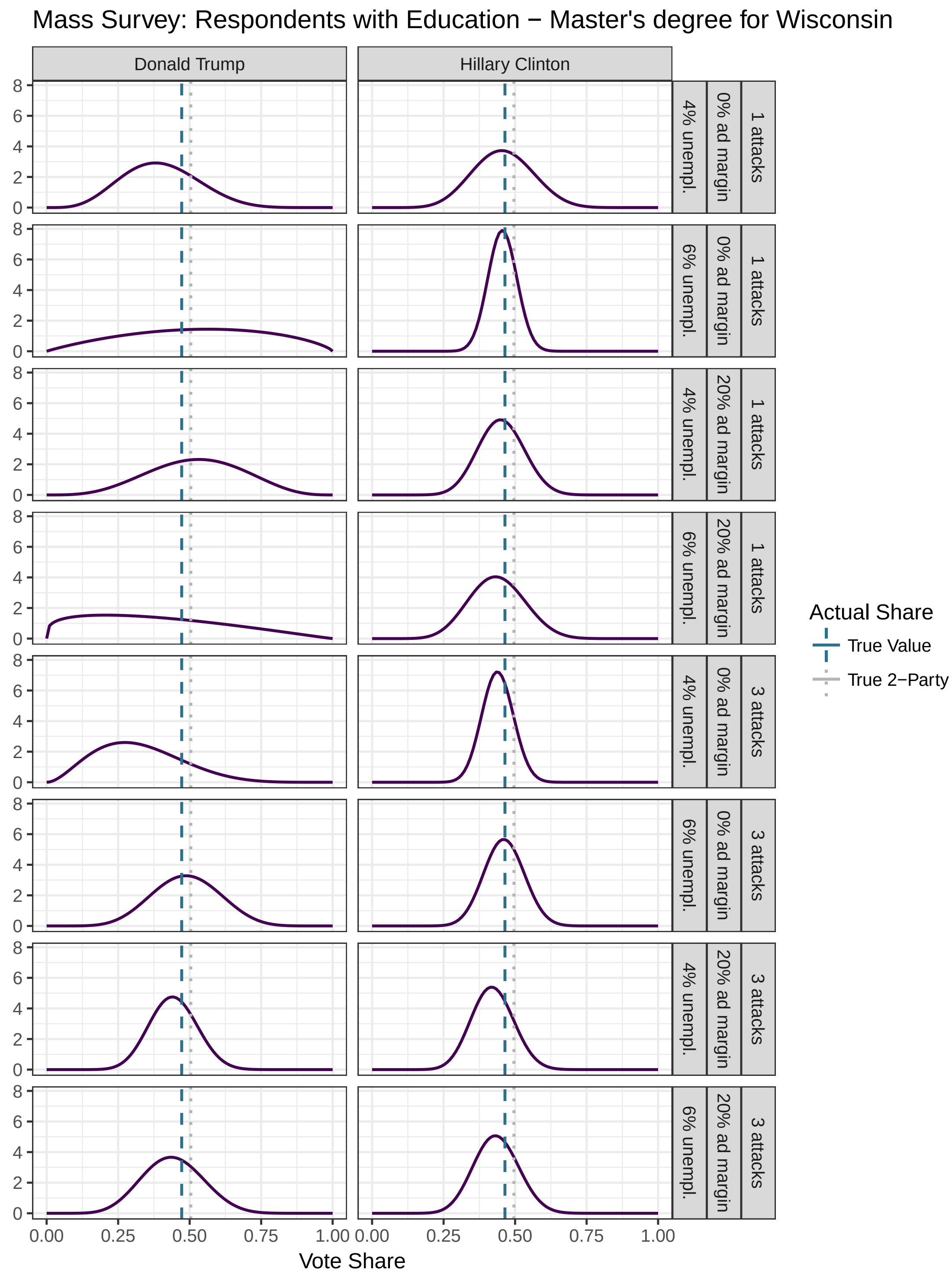


Figure 3.374: Priors with covariates: Mass Wisconsin Education Master's degree

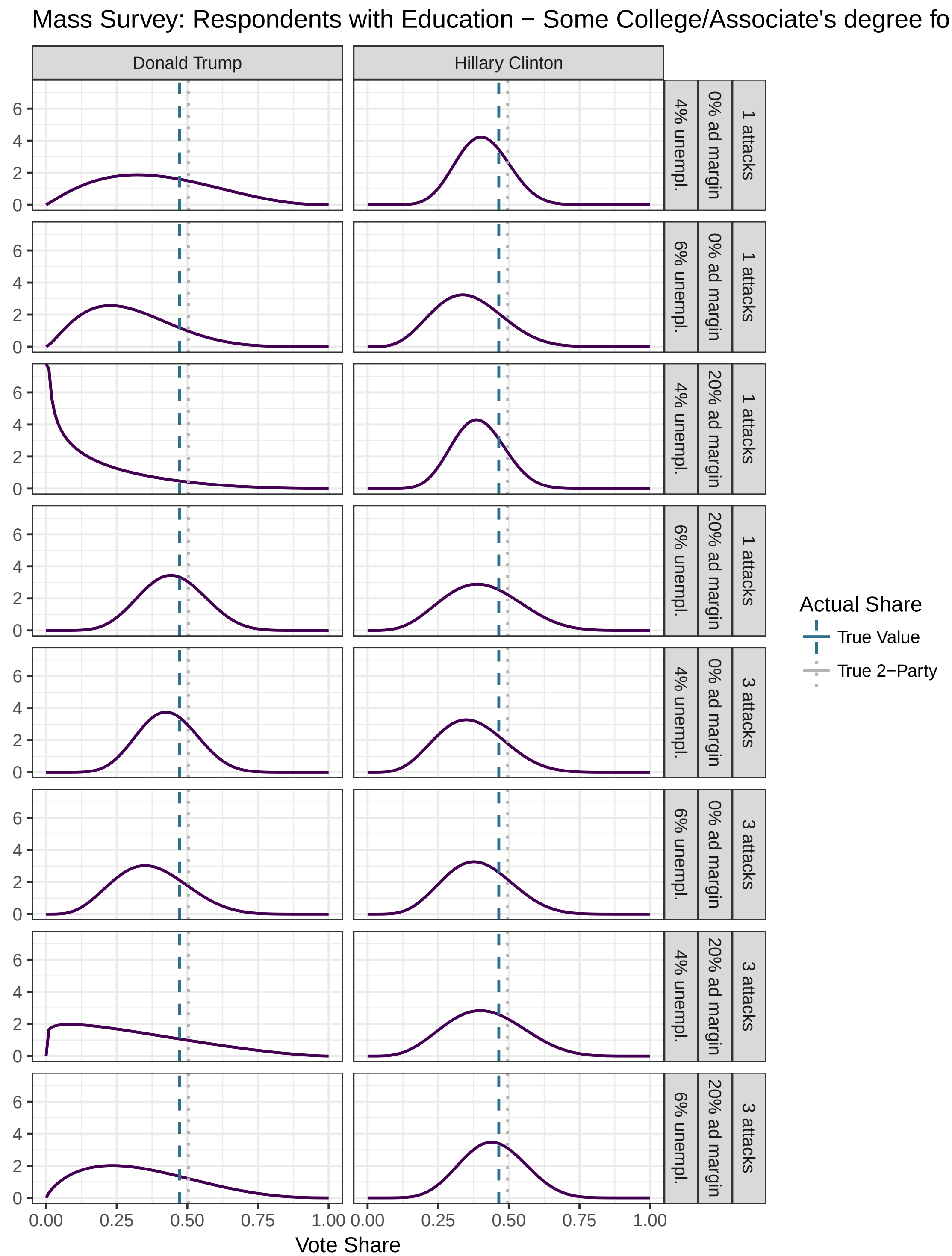


Figure 3.375: Priors with covariates: Mass Wisconsin Education Some College Associate's degree

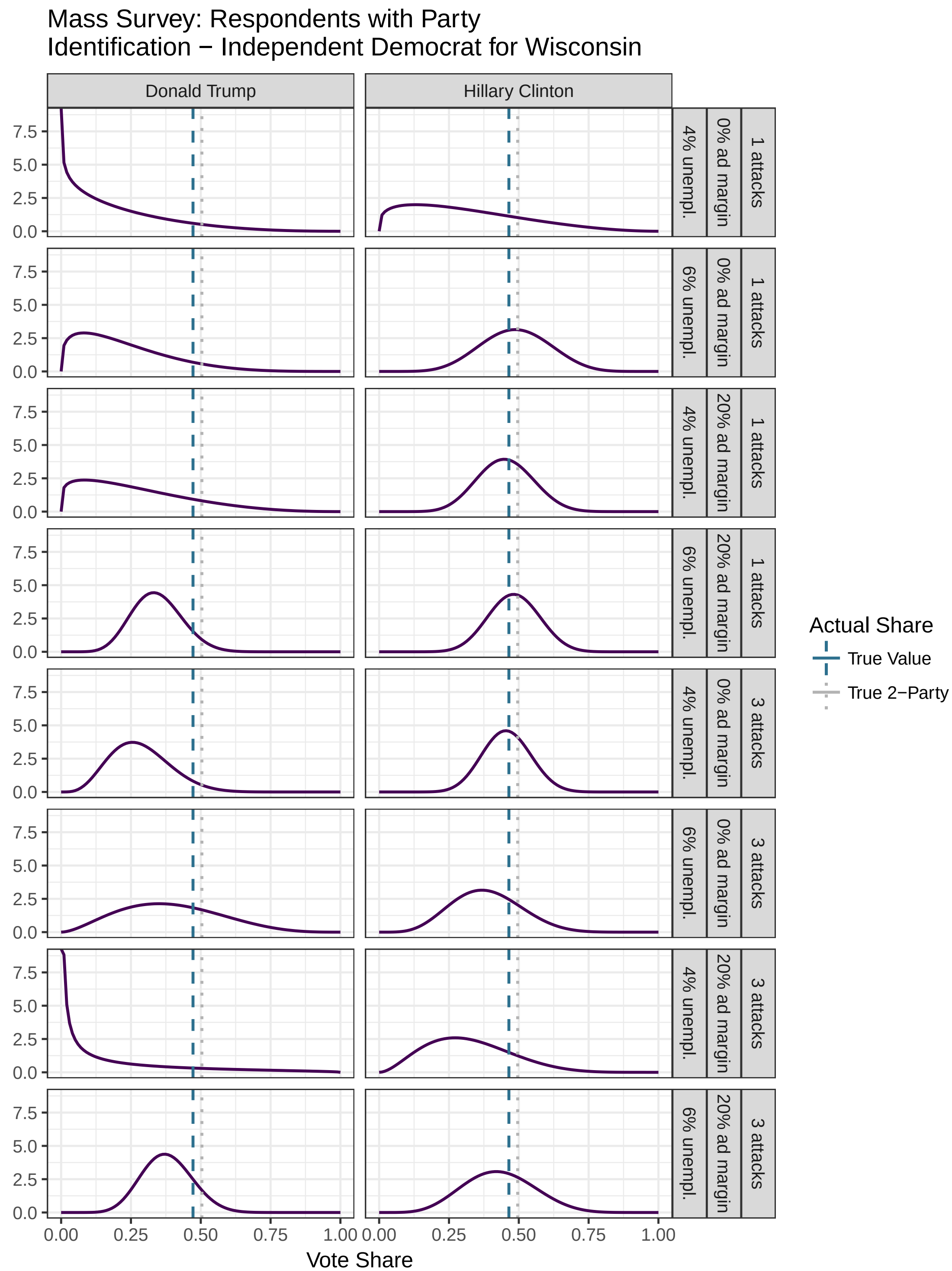


Figure 3.376: Priors with covariates: Mass Wisconsin Party Identification Independent Democrat

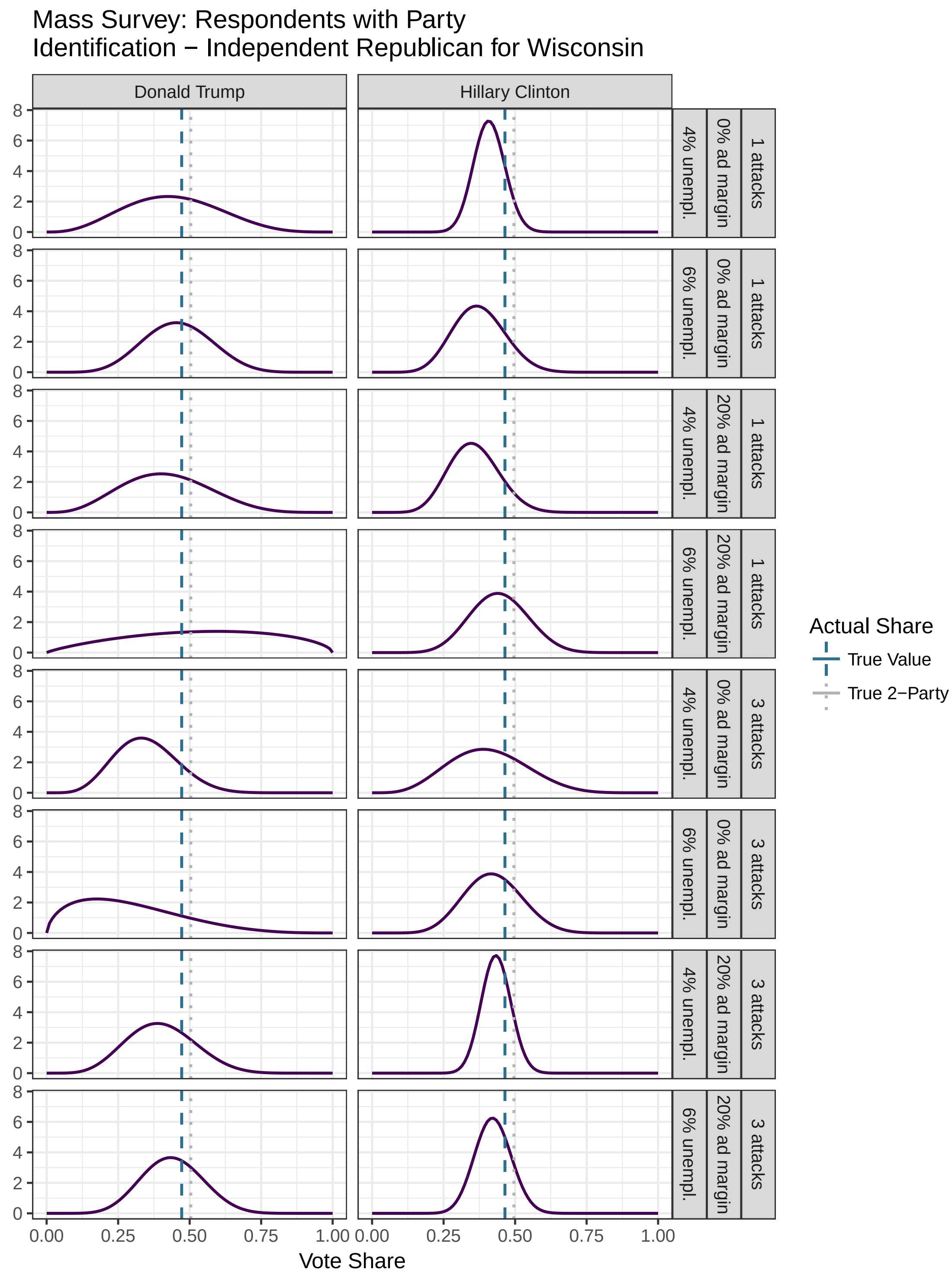


Figure 3.377: Priors with covariates: Mass Wisconsin Party Identification Independent Republican

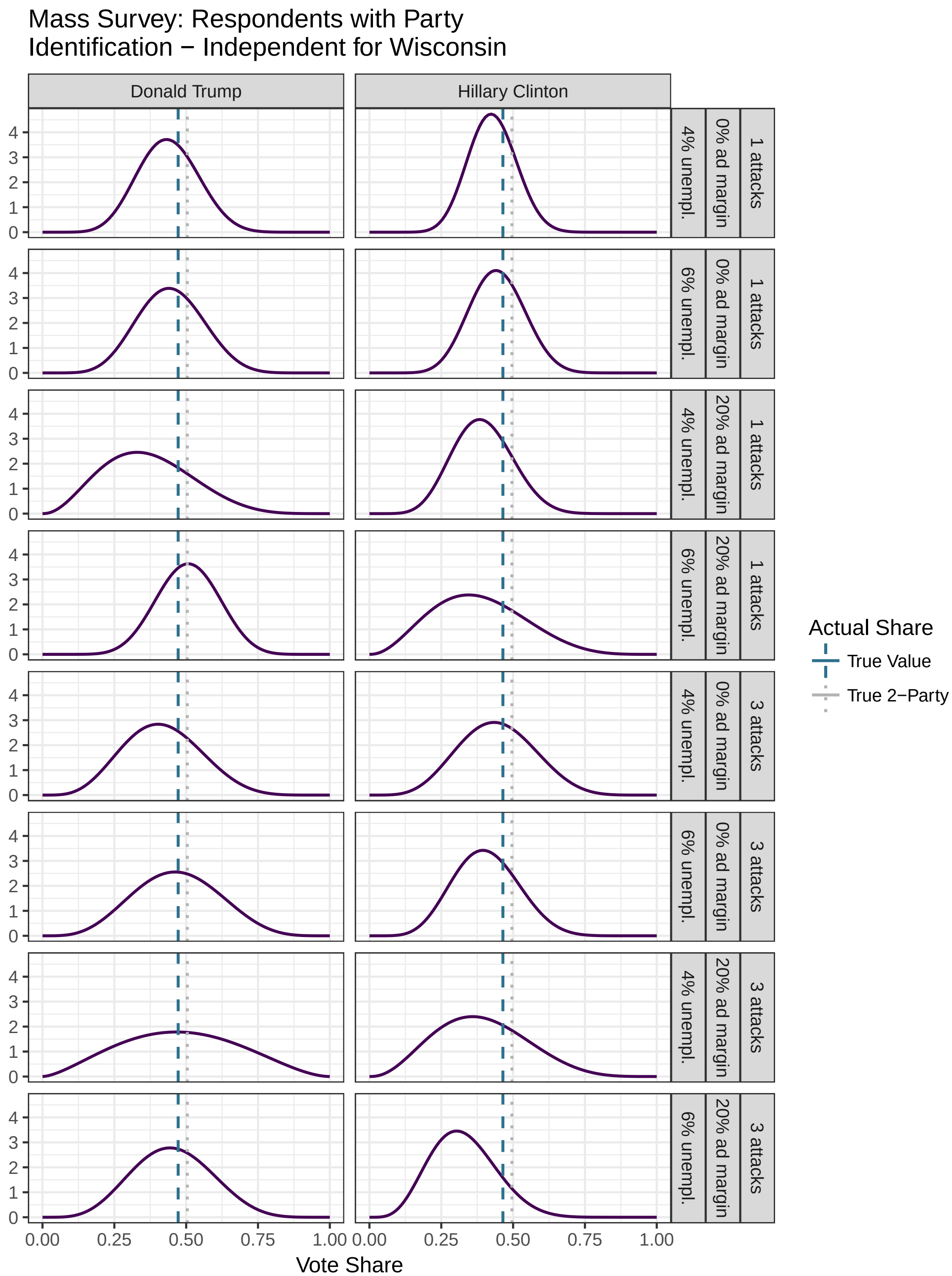


Figure 3.378: Priors with covariates: Mass Wisconsin Party Identification Independent

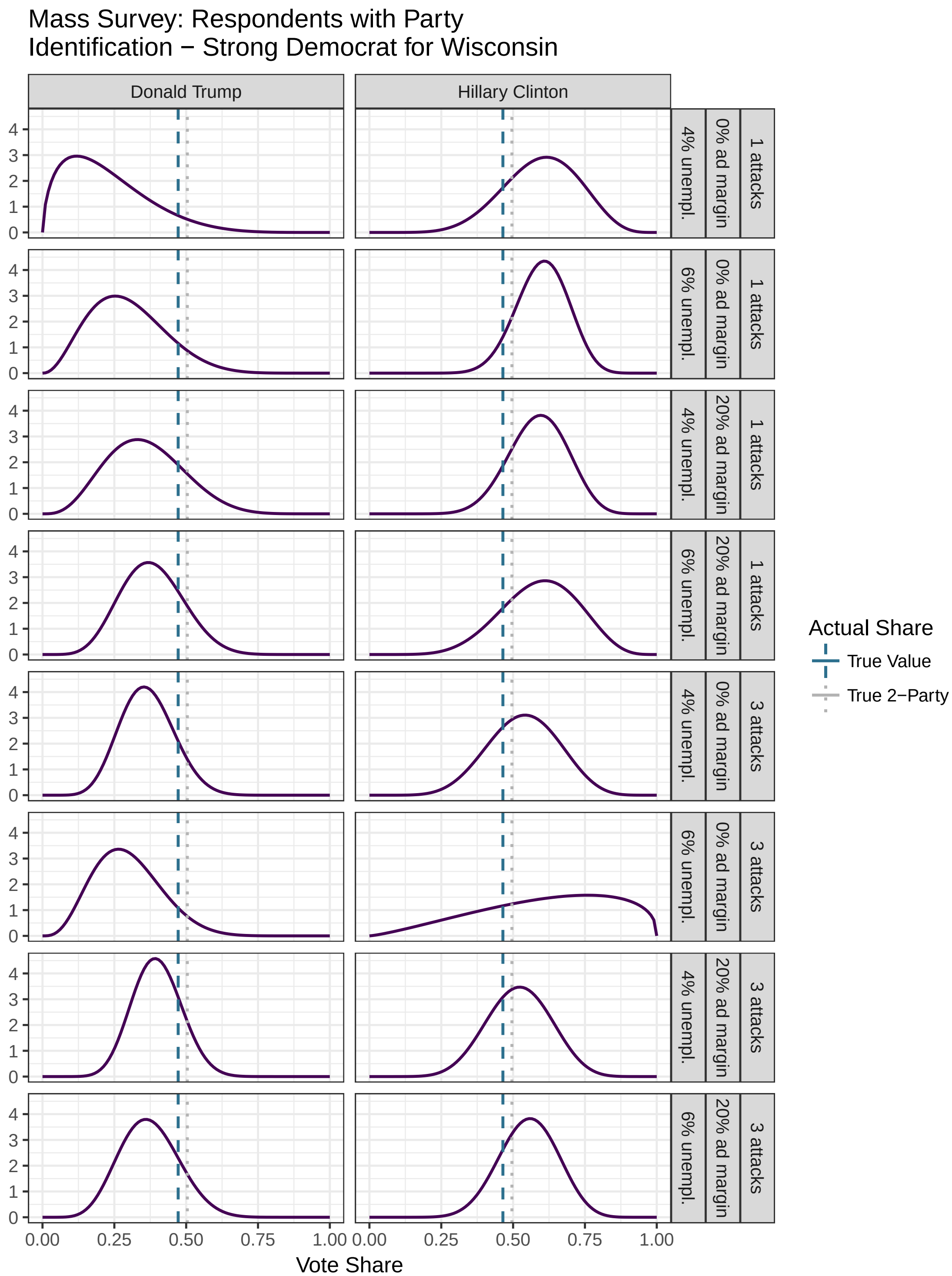


Figure 3.379: Priors with covariates: Mass Wisconsin Party Identification Strong Democrat

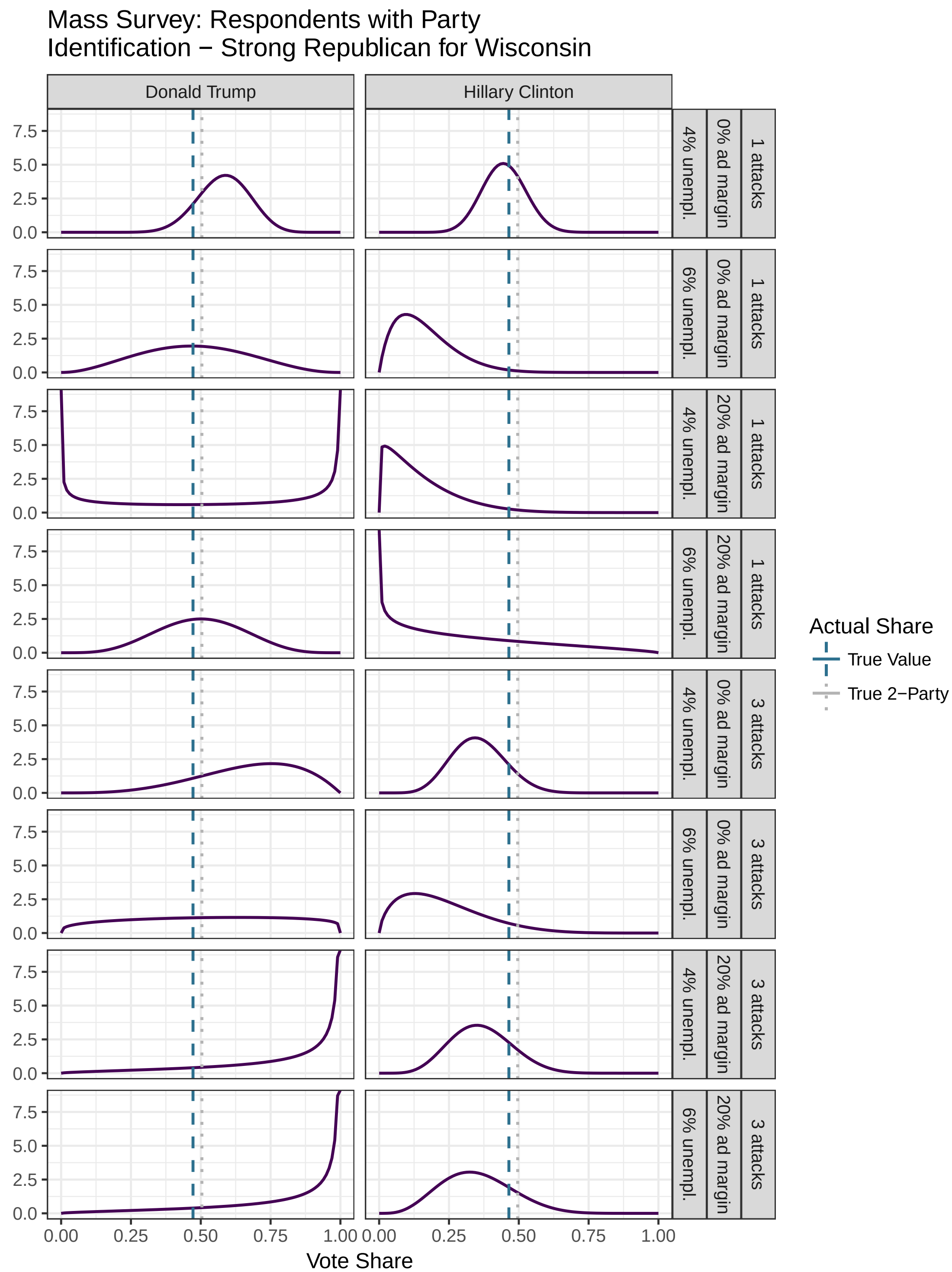


Figure 3.380: Priors with covariates: Mass Wisconsin Party Identification Strong Republican

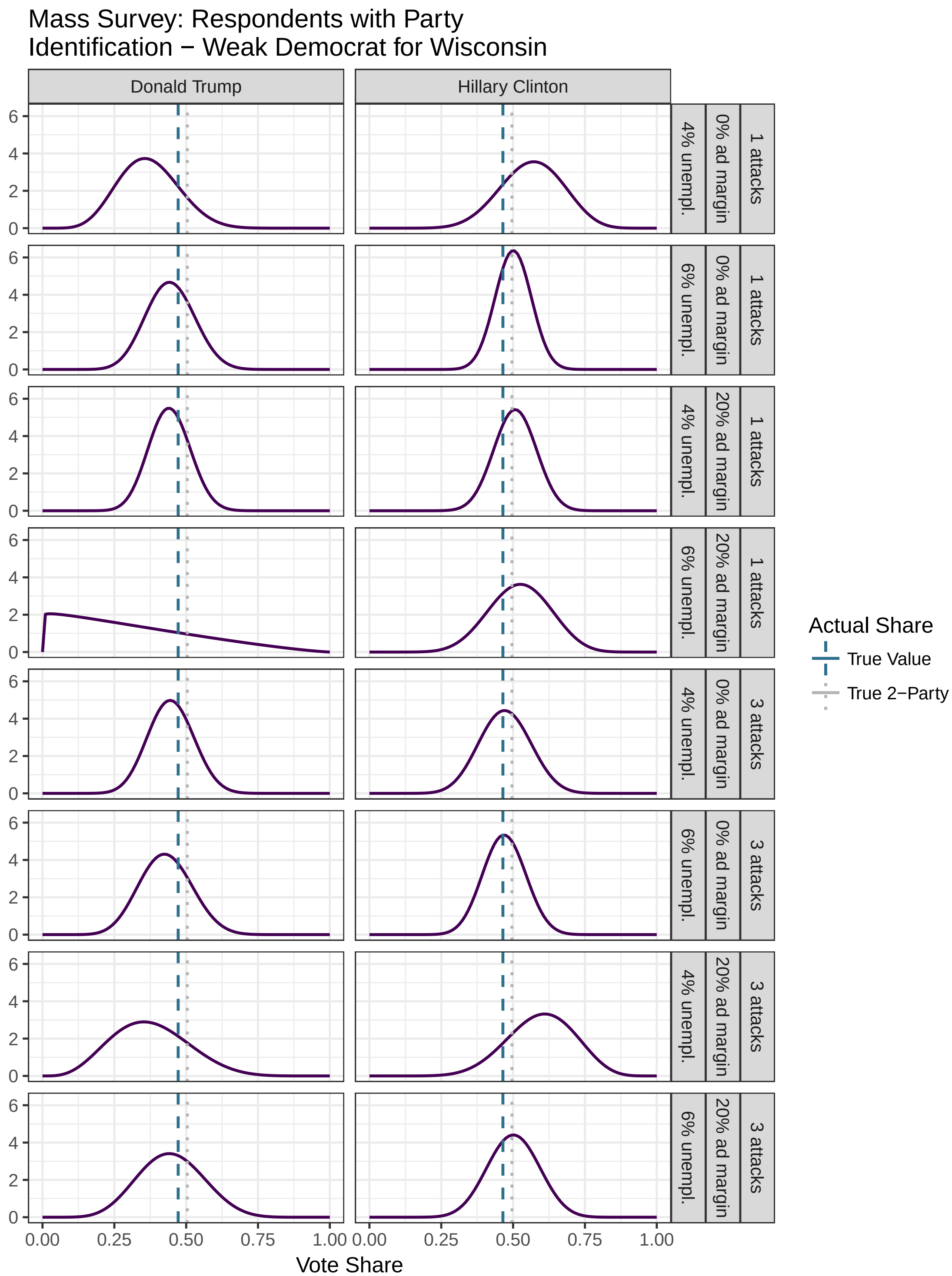


Figure 3.381: Priors with covariates: Mass Wisconsin Party Identification Weak Democrat

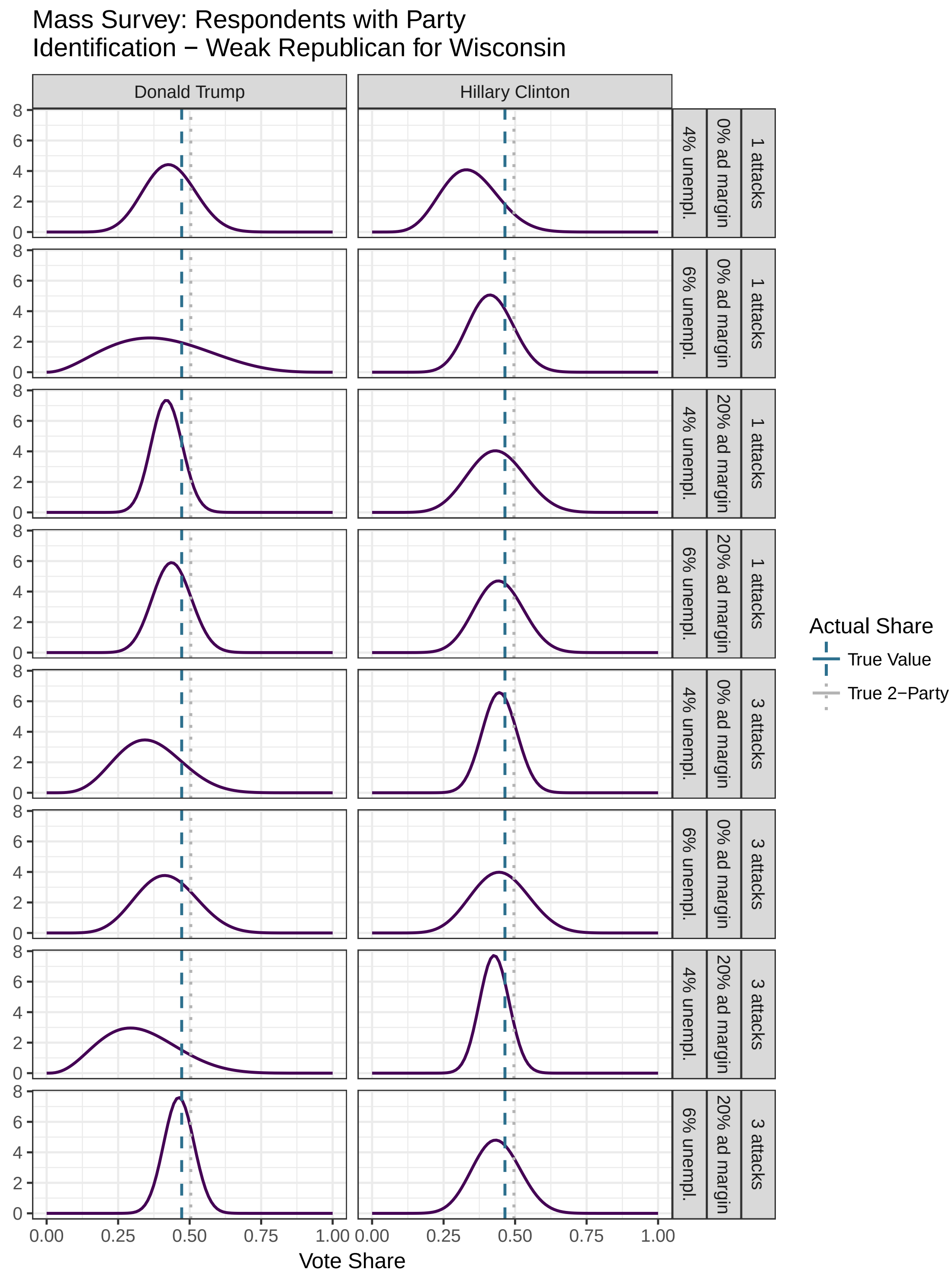


Figure 3.382: Priors with covariates: Mass Wisconsin Party Identification Weak Republican

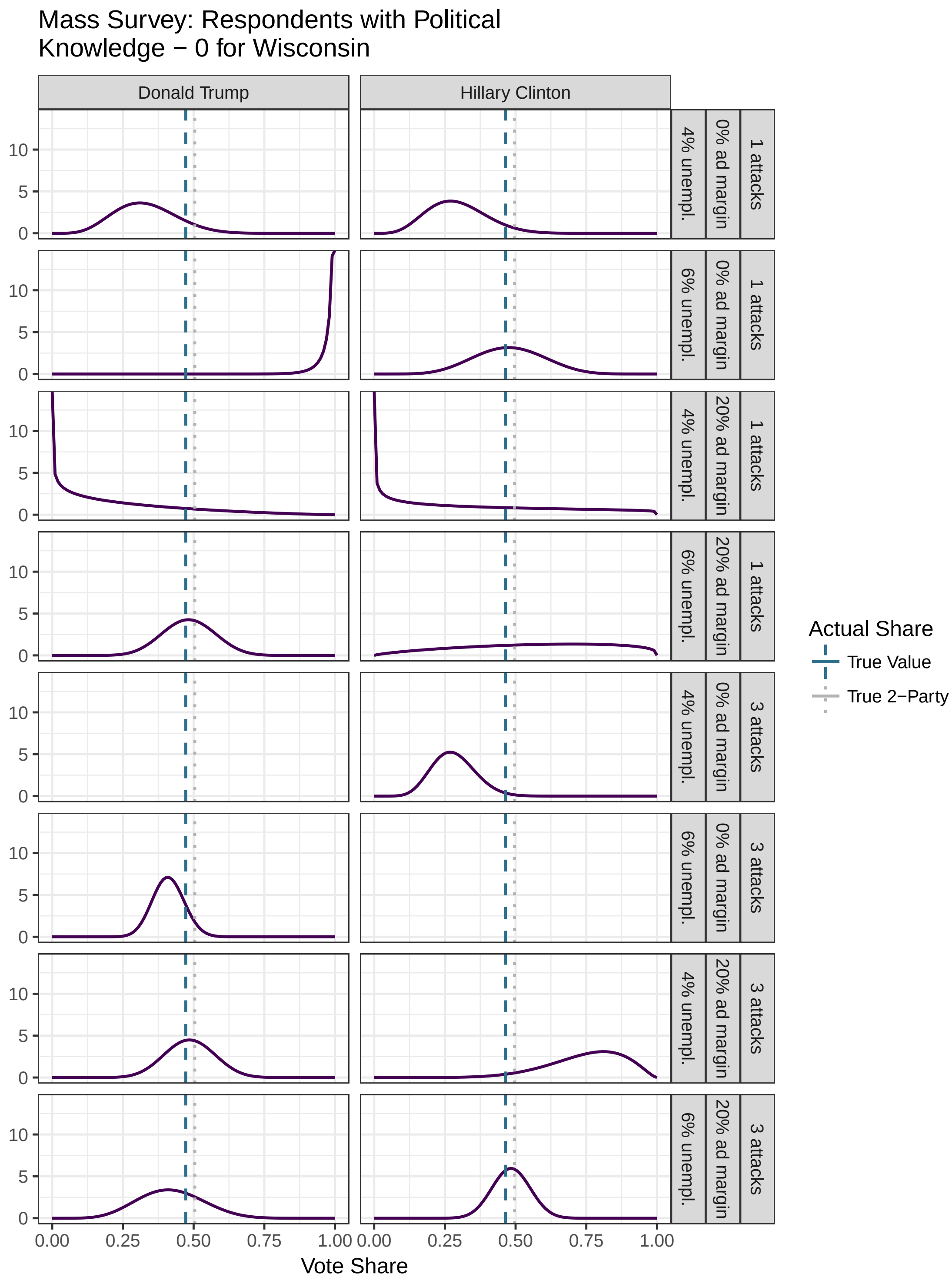


Figure 3.383: Priors with covariates: Mass Wisconsin Political Knowledge 0

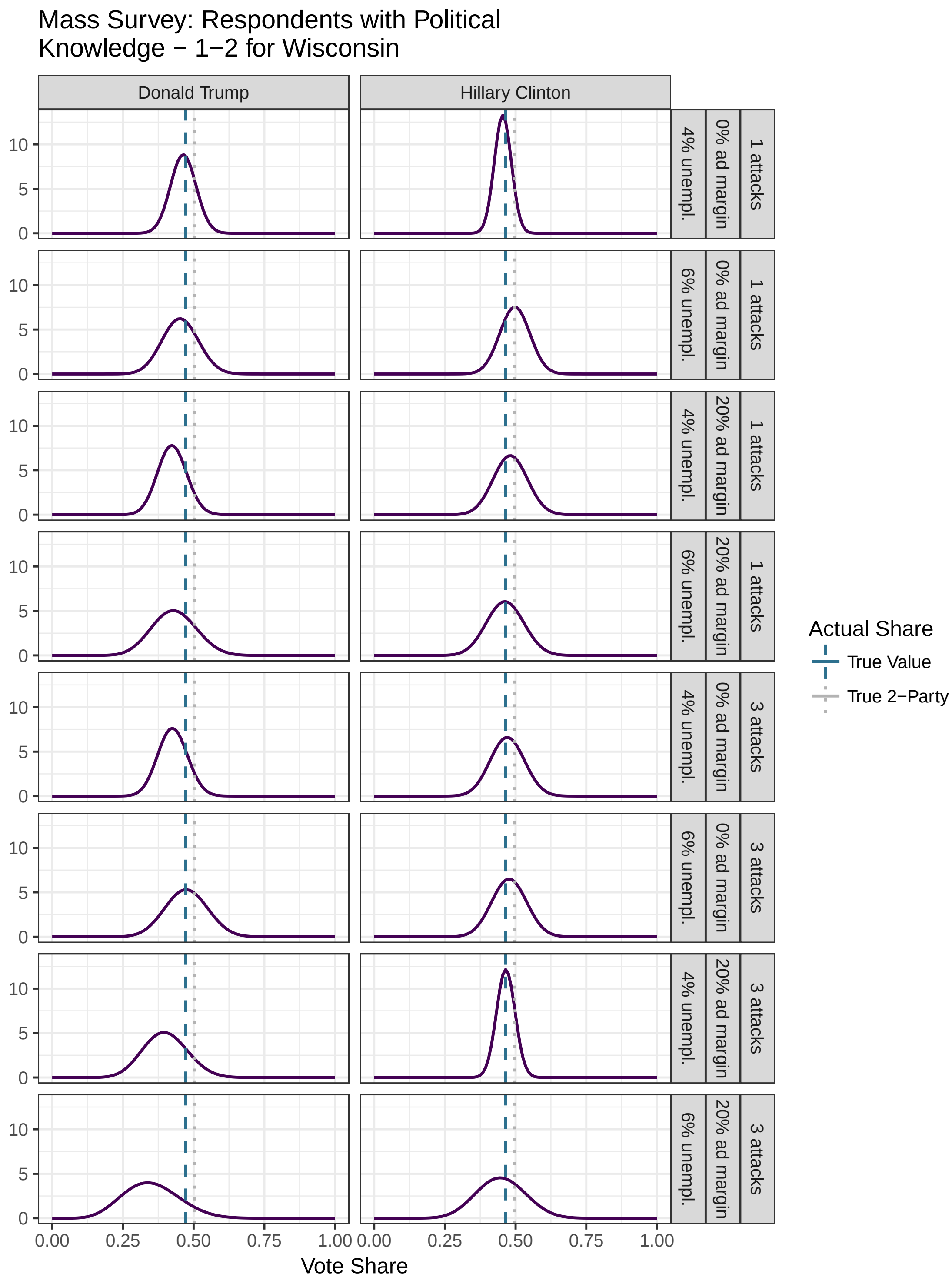


Figure 3.384: Priors with covariates: Mass Wisconsin Political Knowledge 1-2

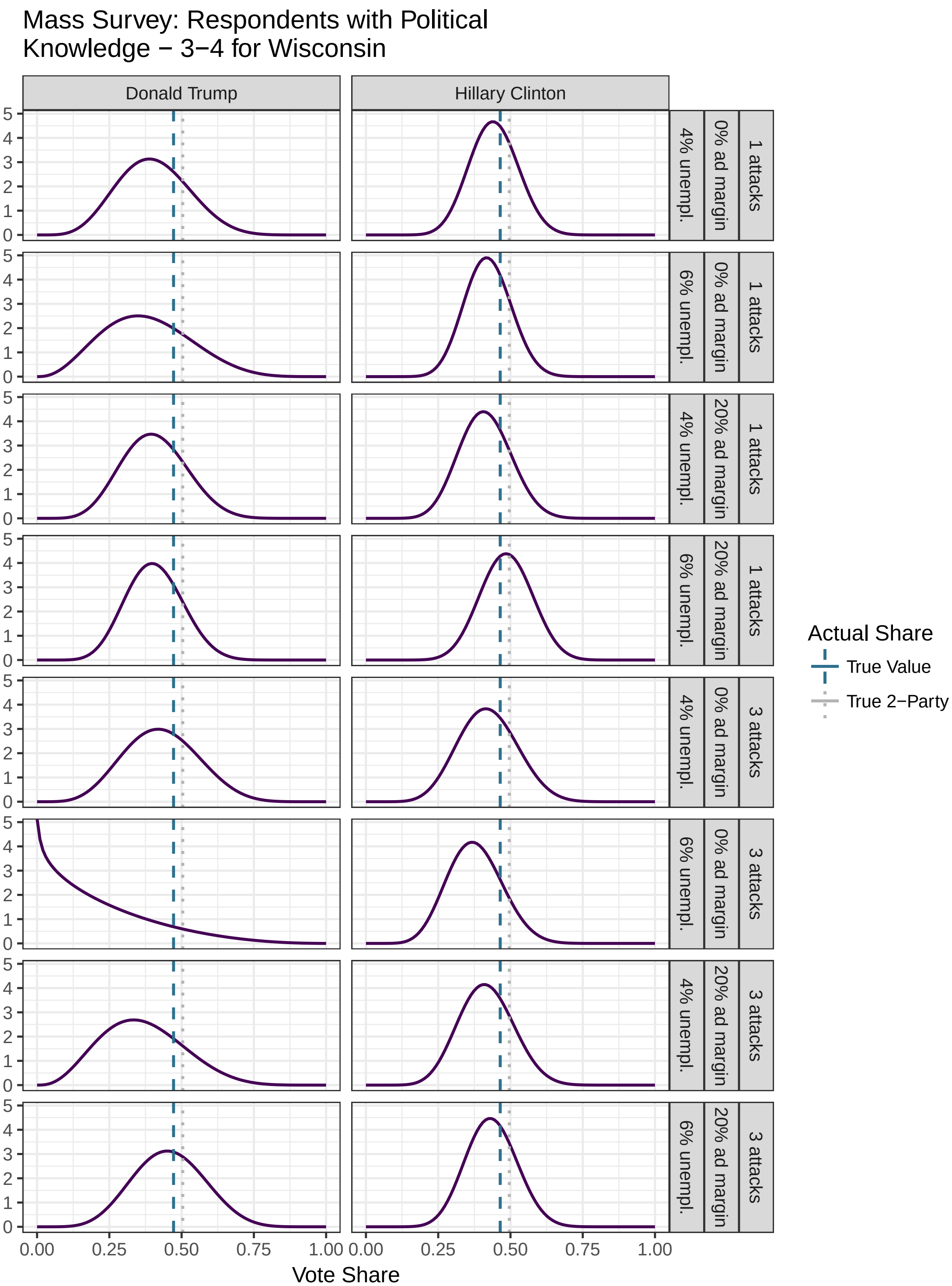


Figure 3.385: Priors with covariates: Mass Wisconsin Political Knowledge 3-4

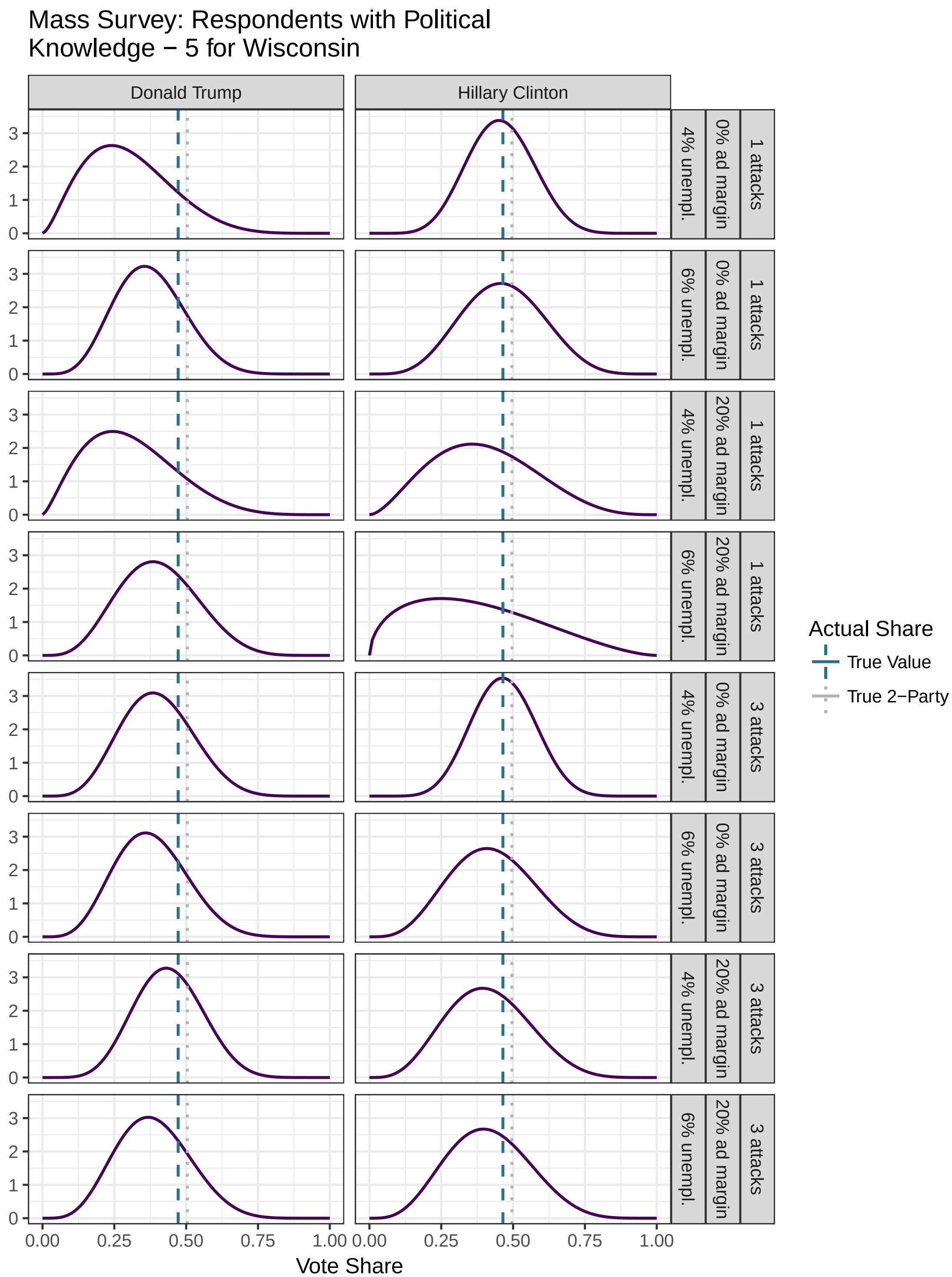


Figure 3.386: Priors with covariates: Mass Wisconsin Political Knowledge 5

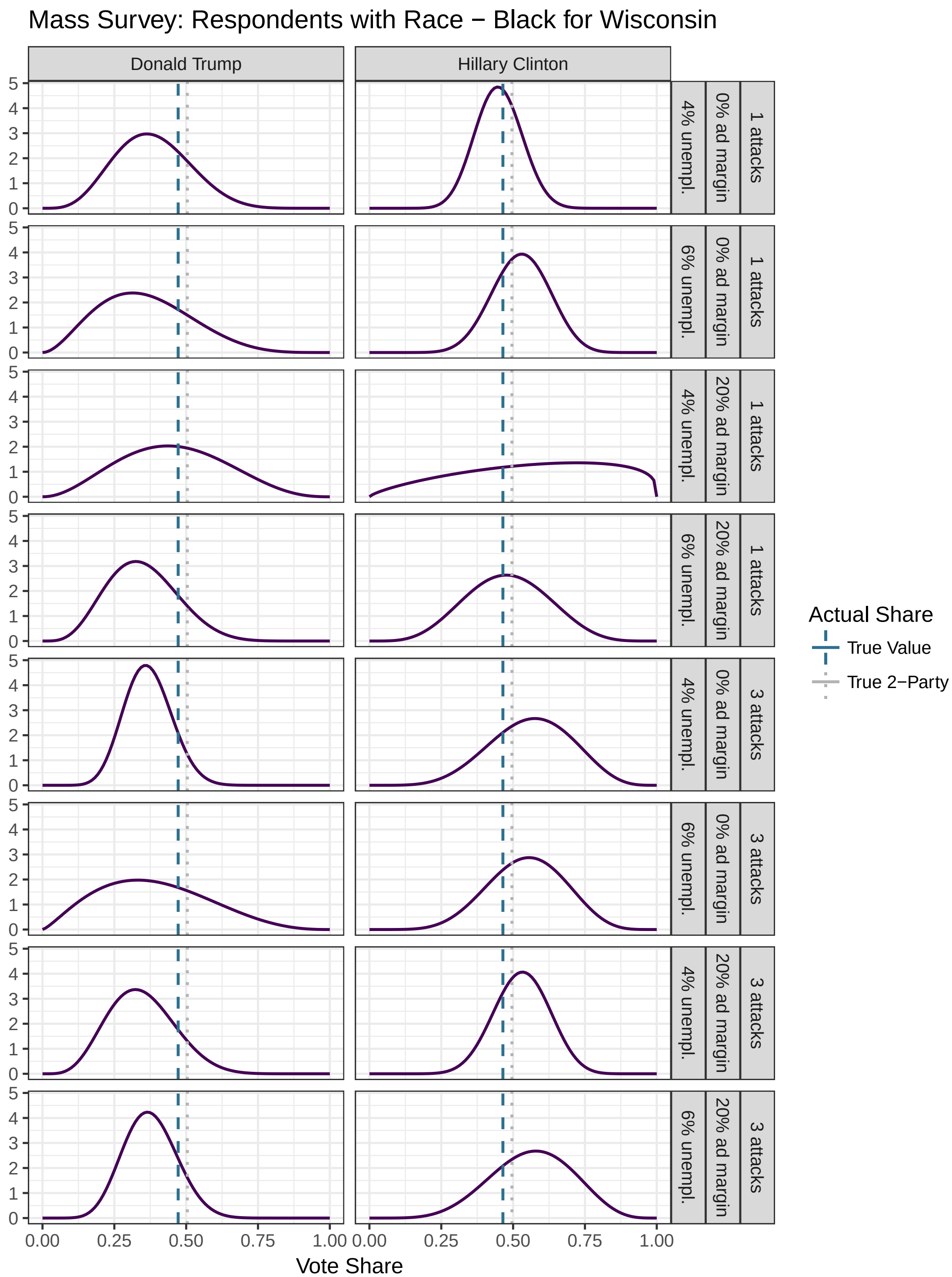


Figure 3.387: Priors with covariates: Mass Wisconsin Race Black

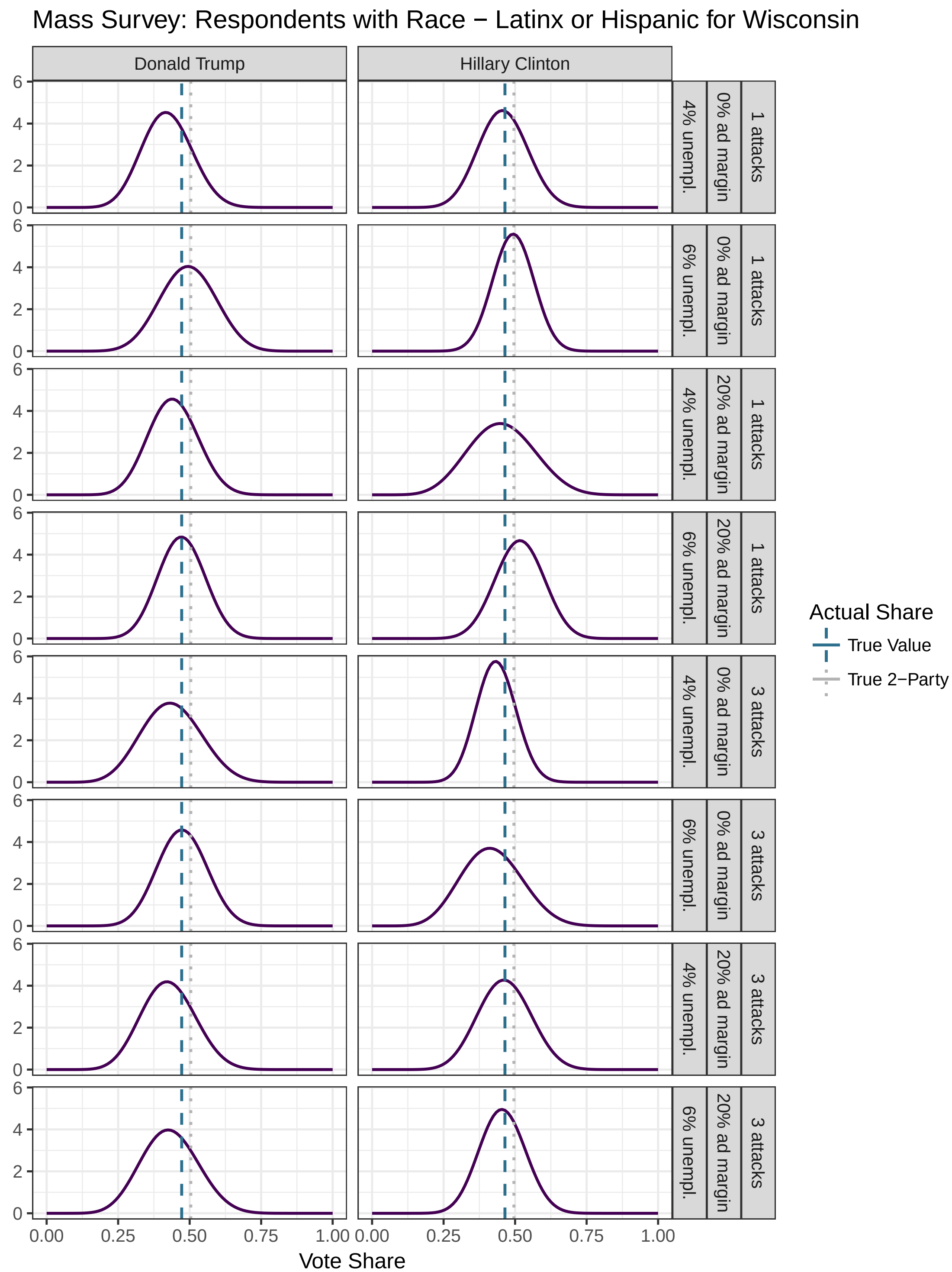


Figure 3.388: Priors with covariates: Mass Wisconsin Race Latinx or Hispanic

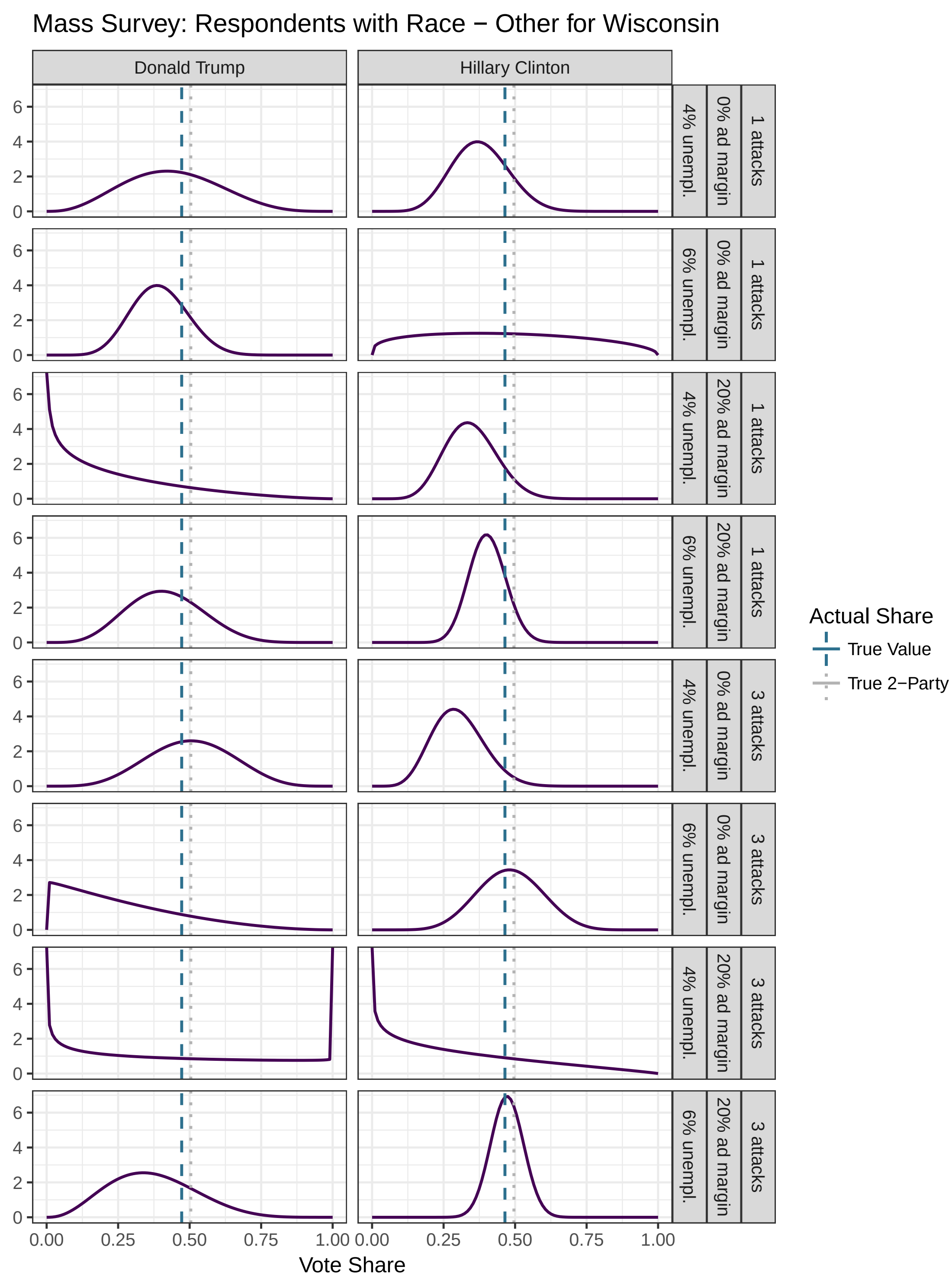


Figure 3.389: Priors with covariates: Mass Wisconsin Race Other

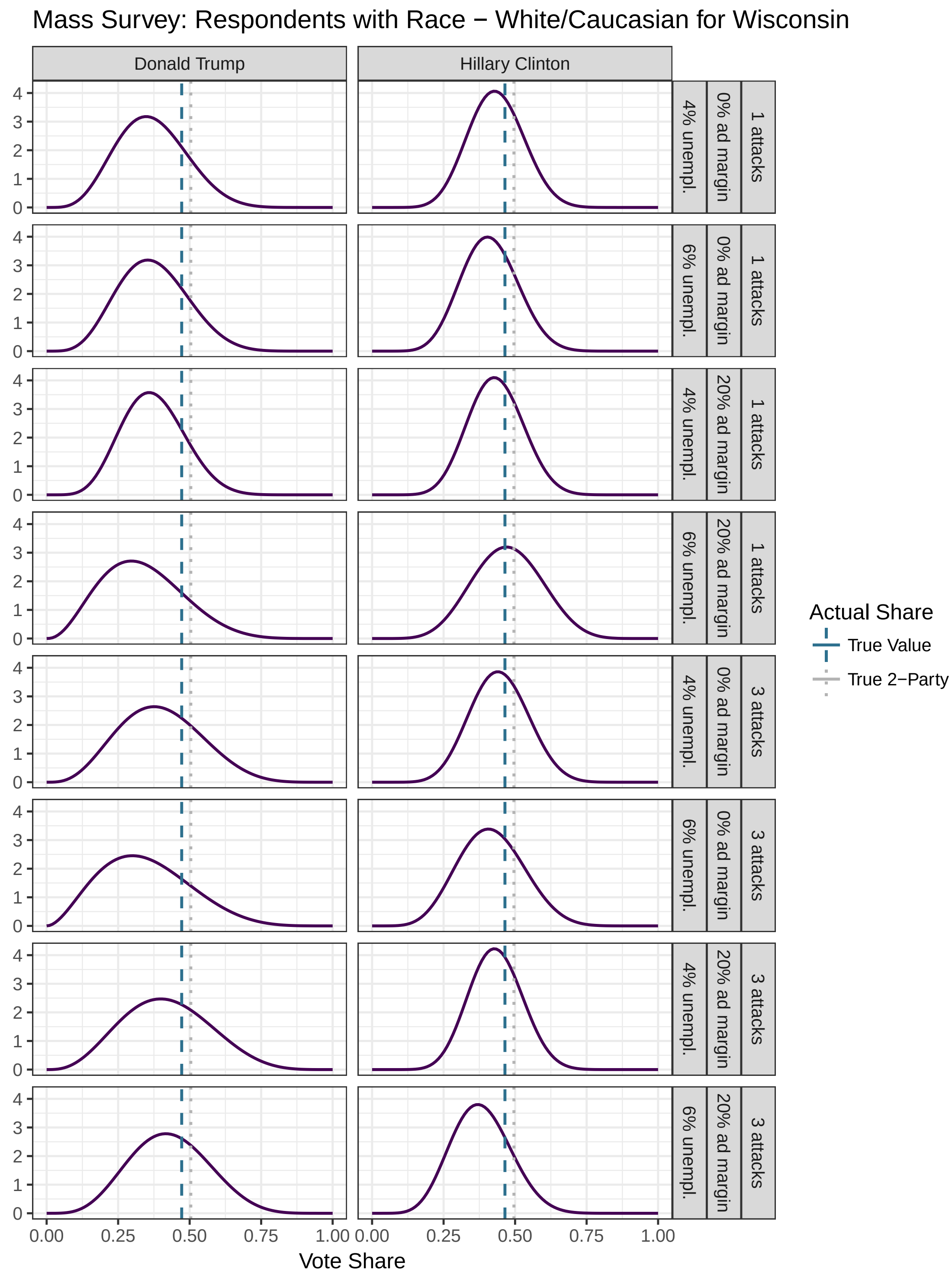


Figure 3.390: Priors with covariates: Mass Wisconsin Race White Caucasian

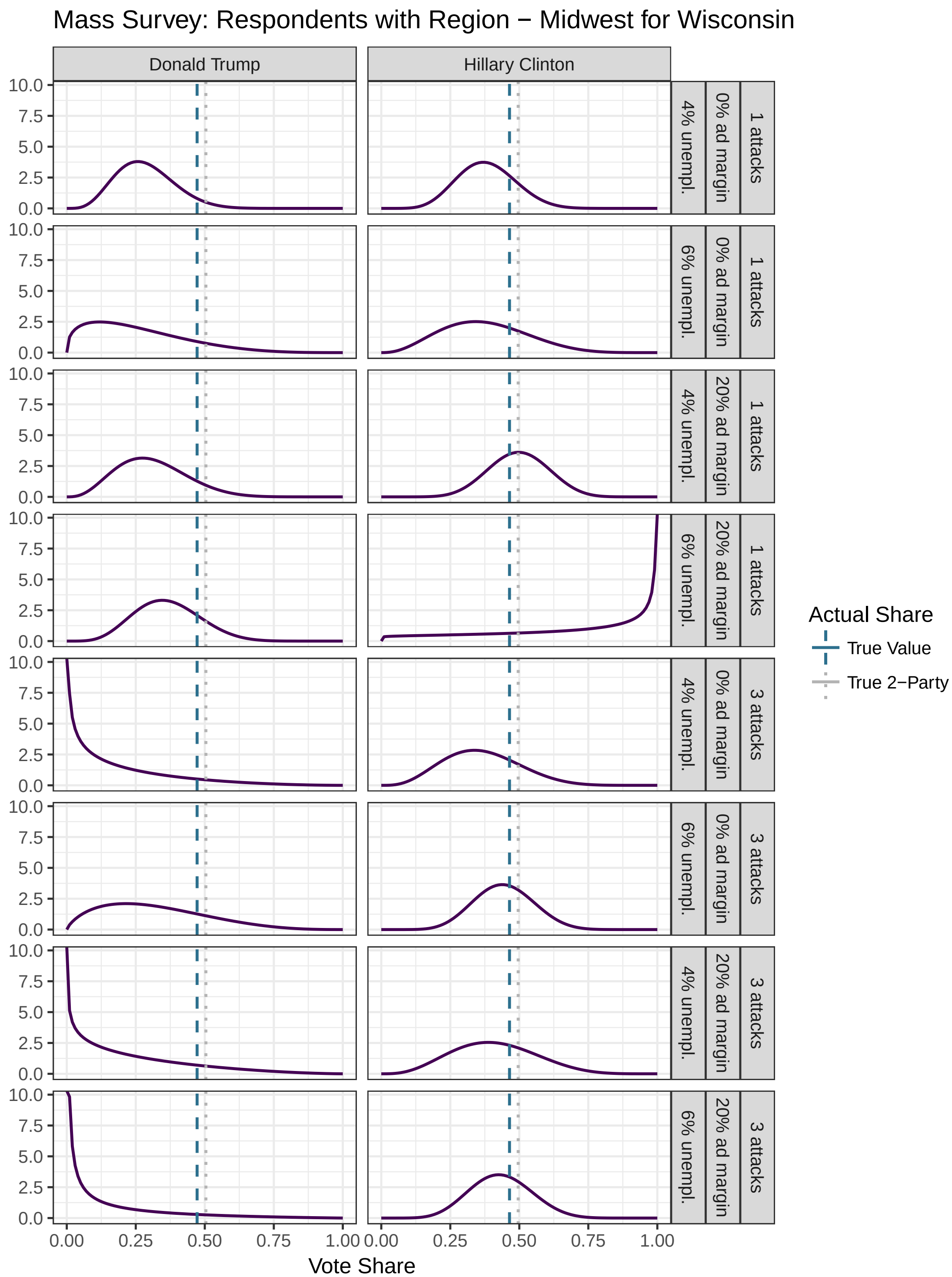


Figure 3.391: Priors with covariates: Mass Wisconsin Region Midwest

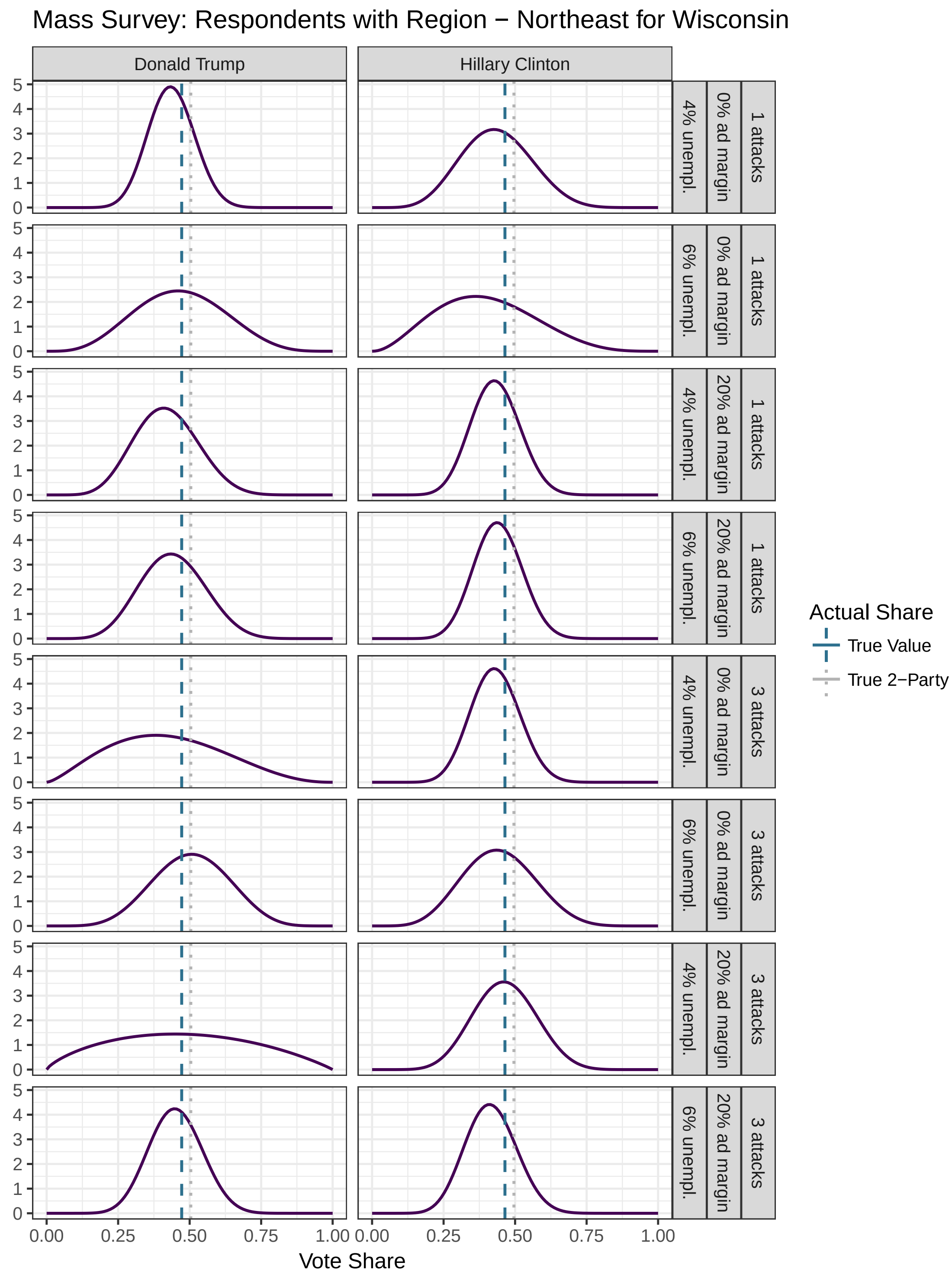


Figure 3.392: Priors with covariates: Mass Wisconsin Region Northeast

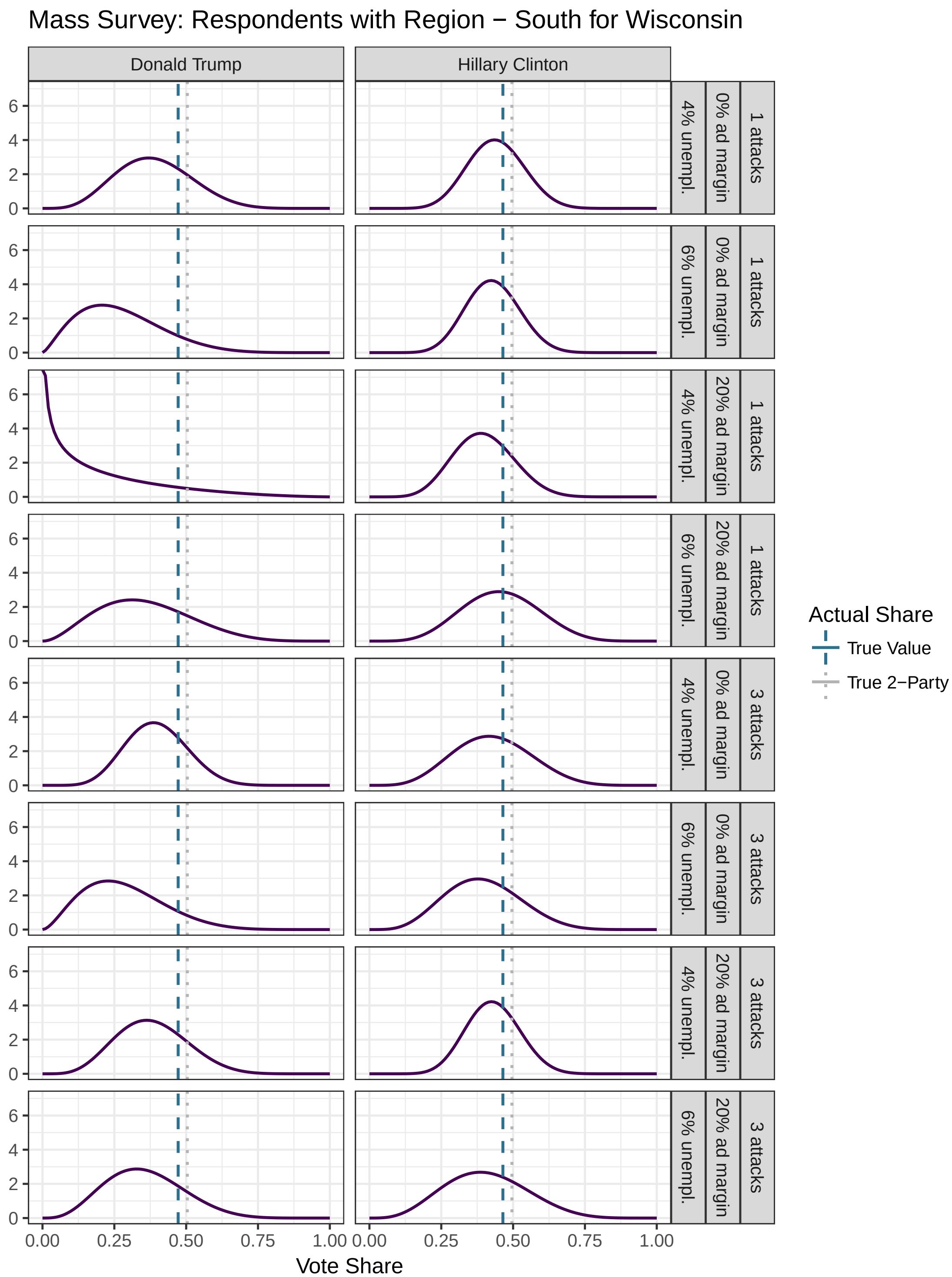


Figure 3.393: Priors with covariates: Mass Wisconsin Region South

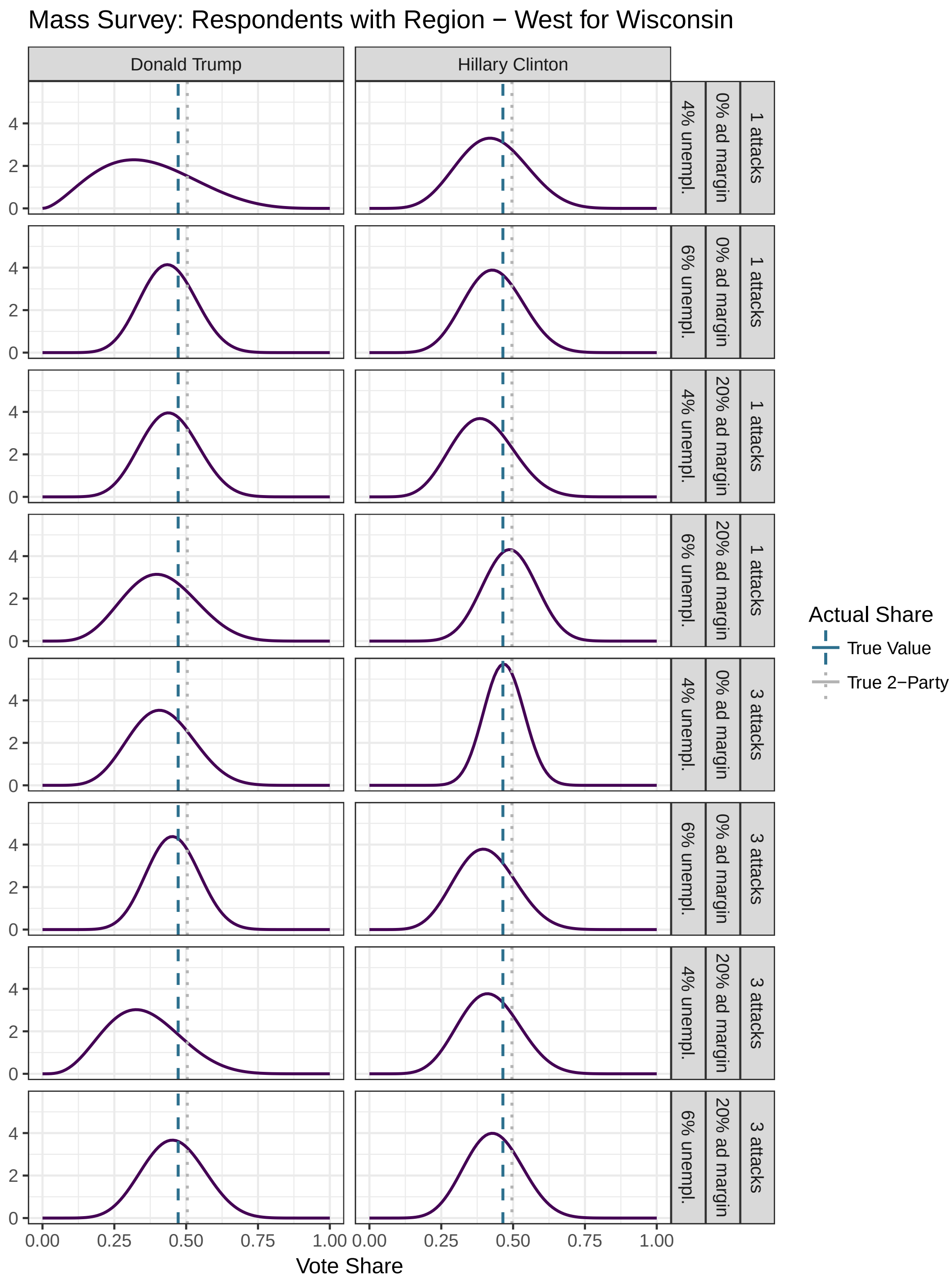


Figure 3.394: Priors with covariates: Mass Wisconsin Region West

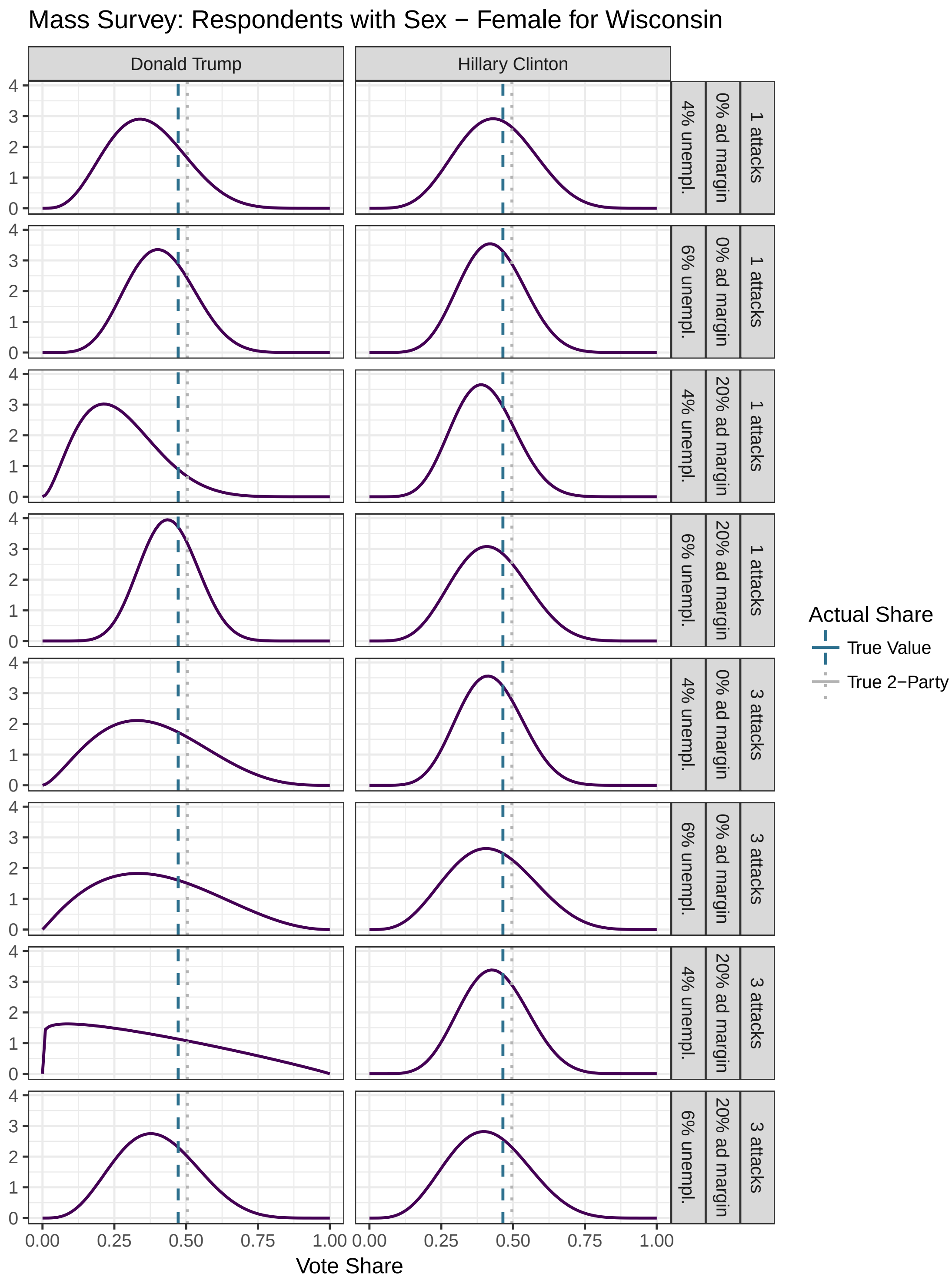


Figure 3.395: Priors with covariates: Mass Wisconsin Sex Female

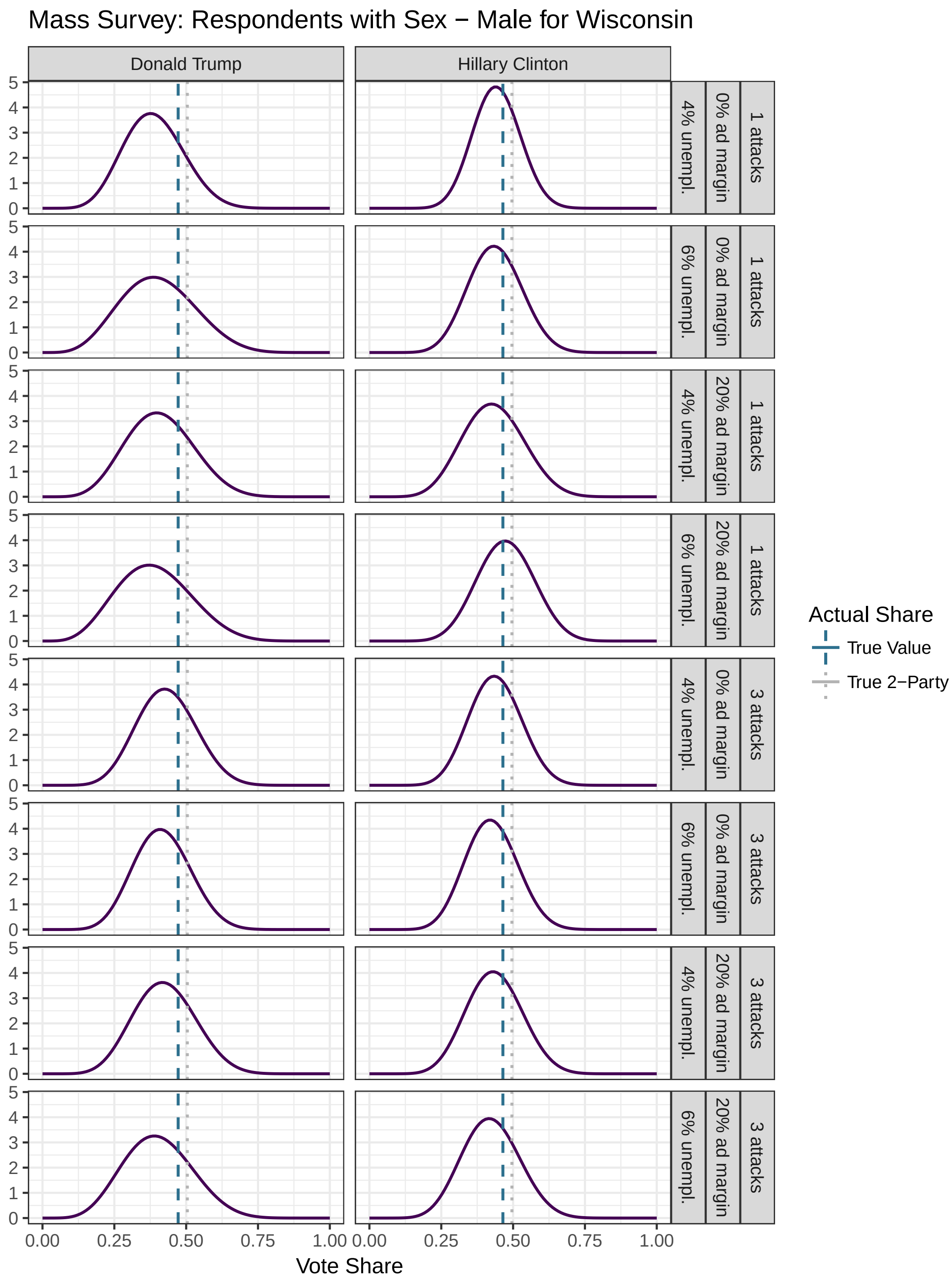


Figure 3.396: Priors with covariates: Mass Wisconsin Sex Male

Chapter 4

Careers & Causes in Authoritarian Legislatures: Clustering Text-Based Elicited Priors

4.1 Question & Motivation

What motivates participation in authoritarian legislatures? The proliferation of formal political institutions in autocratic contexts has spawned various academic attempts to explain the emergence, function, and persistence of these institutions. Yet even the most careful accounts of authoritarian legislatures, in particular, fail to provide micro-foundations for the participation of lawmakers in floor debates, roll-call votes, or query sessions. If authoritarian institutions serve a purpose other than mere "window dressing," that function is contingent on the active participation—or compliant nonparticipation—of individuals selected into the institution.

This paper seeks to understand how individual legislators participate in an authoritarian context, and to introduce a text-based prior elicitation method in order to empirically examine this phenomenon in a Bayesian framework. Methodologically, this project introduces a novel way to incorporate "expert" knowledge into the analysis of a political context with sparse data: culling the perspectives of observers, specifically through newspapers, to theoretically inform empirical analysis and to practically overcome estimation challenges.

Substantively, this examination is motivated in part by observed empirical variation in the level and type of participation in which legislators in authoritarian systems engage. Current scholarship fails to adequately explain this variation. Present theories suggest that legislatures allow dictators to provide policy concessions, or serve as a venue through which hegemonic parties co-opt. For example, Svolik suggests a broad strategy of co-optation within authoritarian parties, but does not indicate how co-optation could or should govern the behavior of party members as they function within political roles and institutions. Just as literature on democratic contexts anticipates that parties will govern legislative coalitions, one might expect that authoritarian party co-optation should govern behavior within an authoritarian legislature.

Yet this narrative provides too blunt and too remote a picture of how co-optation should influence behavior. If an authoritarian party plays a significant role in selecting participants to a legislature, what would explain variation in co-optedness? Even if variation in behavioral indications of co-optation is relatively small, should participation be taken as evidence of co-optation, or should nonparticipation? Should we expect a race to participate first and frequently in a laudatory fashion, constantly praising and supporting regime policies, or an encompassing silence punctuated by occasional objections? Current empirical scholarship cannot adequately address these types of questions because it considers the motivations of legislators only in a very limited fashion; rather, these types of questions suggest a benefit to formalizing the incentives and preferences of legislators that might explain their behavior within autocratic institutions. While intuitively, for instance, laudatory behavior requires even minimal effort compared to nonparticipation, and therefore coordination on nonparticipation is easier and more efficient, this type of coordination should increase the signaling value of any defector's laudatory participation. Exploring exactly what produces either a participation or nonparticipation equilibrium would provide insight into exactly the mechanisms through which career incentives (through a party, for example) might relate to legislative behavior in ways that the present empirical analysis cannot.

In this paper, I address these substantive considerations through two separate questions:

(1) Under what conditions do legislators participate at all in parliament?

(2) Having participated, what leads some legislators to participate more than others?

I adopt an empirical approach that allows me to disaggregate motivations for each of these questions while maintaining the inclusion of key explanatory variables from the literature. In particular, this project evaluates the extent to which career concerns weigh against policy preferences in MPs' choice to participate by examining the role that their past experience and commitments, their party affiliation, and their demographic characteristics play in their participation. Moreover, departing from previous work in this vein that has investigated entrenched, dominant-party autocratic systems, this project examines legislative participation in the authoritarian context of Myanmar. Myanmar's parliament, newly established in 2011, incorporates a diverse set of competing interests vying for power —members of the former military junta, the new ruling party, the traditional and widely popular democratic opposition, ethnic minority parties—and includes MPs with a wide range of experience apart from government bureaucracy. This diversity lends itself well to interrogating both career- and policy-related dimensions of legislative participation in an authoritarian context.

More broadly, articulating the microfoundations of behavior in authoritarian legislatures is necessary to understand the ways in which these authoritarian institutions sustain themselves or evolve endogenously. To the extent that authoritarian rulers delegate some control over policy to legislative bodies, or allow for substantive changes to policy regime in consultation with legislators, these institutions are a potential site for retrenchment or democratization. Specifically, how legislators balance their career objectives within an authoritarian system relative to their policy preferences has implications for what types of policy obtain in authoritarian systems in the short term, as well as how citizen constituencies can achieve substantive representation in the long term. This paper aims to address this question initially through legislative participation, since without any participation or contention over policy positions, the sole policy-making apparatus of the state is the ruling coalition.

The paper proceeds by addressing some of the key literatures that offer theoretical and empirical insights into legislatures in autocracies. I draw these literatures into dialogue with an empirical case, that of Myanmar's Pyithu Hluttaw (lower house of parliament), in order to identify independent variables that plausibly explain variation in legislative participation. Following the empirical analysis, I discuss additional extensions, both theoretical and empirical, and offer initial conclusions.

4.2 Literature & Case

The question this paper engages several subliteratures that investigate the role of institutions, and particularly legislatures, under autocracy.

4.2.1 Institutions & Legislatures under Autocracy

Jennifer Gandhi argues that institutions in autocratic contexts are established to solve particular political problems (Gandhi 2008), ideally with the aim of increasing the odds of regime survival (Gandhi and Przeworski 2007). Institutions in Gandhi's narrative allow dictators to co-opt and offer concessions to potential opposition. Bruce Bueno de Mesquita et al. echo this argument in their articulation of the "selectorate theory," in which the population determining a leader's prospects for survival (the winning coalition) is targeted for particular concessions because they enable the dictator to remain in power (Bueno de Mesquita et al. 2003). Autocrats can thus use institutions to gather information (Manion 2013) and distribute rents or provide policy concessions as a means of achieving continued tenure in office, particularly when repression is costly (Magaloni and Kricheli 2010, 126, 129-130). Given this, the question remains how legislatures in particular function to dispense concessions, as well as how individual legislators within that venue might go about maximizing the concessions they receive or the payoffs they obtain.

Wright and Escribà-Folch (2011) emphasize that although legislatures themselves may contribute to authoritarian stability, parties within those legislatures can serve a destabilizing function. This assertion aligns with the notion that legislatures can serve as a locus of negotiation with elites, or at a minimum provide for monitoring of whether authoritarians fulfill obligations (i.e., payments or concessions) (Boix and Svolik 2013; Jensen, Malesky, and Weymouth 2013; Gehlbach, Sonin, and Svolik 2016). Wright and Escribà-Folch (2011) focus on the threat of democratization in particular as a potential outcome of destabilization, but their disaggregation of the function of legislatures relative to parties provides important insights for evaluating the significance of parliamentary participation under autocracy. Specifically, while legislatures can provide a credible commitment by authoritarians to power-sharing or resource distribution (Magaloni 2008), parties are an overlapping institution that can destabilize authoritarian rule by influencing this process of allocation (Wright and Escribà-Folch 2011, 308). That is, how individuals aggregate their interests and act within the legislative context in the form of parties has significant implications for regime survival and stability in authoritarian systems. Going beyond the observation of much early work on authoritarian systems—that institutions matter—requires investigating not just *how* these institutions matter for regime outcomes, but also the ways in which electoral institutions, parties, and legislatures serve overlapping and competing functions in particular contexts (Art 2012; Morse 2012; Reuter and Robertson 2014).

Core to a structural-functional evaluation of authoritarian institutions, furthermore, is an assessment of the incentives that individuals have and how they act accordingly. As Reuter and Robertson (2014, 237) note, the narrative of cooptation in authoritarian legislatures put forth by the literature does not adequately address not just the need to assuage the authoritarian elite but also the imperative to counter broader unrest through pork or concessions (minimal "representation"). Under what conditions opposition or other members of parliament comply with expectations merely to provide spoils to constitutents, rather than more comprehensively impacting the policy-making process, is left unexamined (245-246). The distinction in these types of behaviors, rather, would be most evident in legislative participation, whether through floor speeches, bill authorship or sponsorship, or questioning of the authoritarian elite.

The primary empirical paper that provides both theoretical discussion of the issue of participation in authoritarian legislatures and evidence of its variation is Malesky and Schuler's 2010 paper, "Nodding or Needling: Analyzing Delegate Responsiveness in an Authoritarian Parliament." Their paper utilizes data on legislator behavior from biannual query sessions of the Vietnamese National Assembly from 2007–2012 to evaluate the assertion of the authoritarian institutions literature that autocratic regimes establish parliaments to "co-opt" opposition and negotiate policy concessions that allow for regime stability and longevity. The paper notes a tension in the key assumption of this authoritarian institutions literature: the parliament must facilitate discussion, but not so much as to jeopardize regime stability and the authority of the ruling party (Malesky and Schuler 2010, 482). Malesky and Schuler appeal to empirical evidence to assess how one authoritarian regime might strike this balance.

The paper draws on transcripts from all query sessions in the 12th VNA, beginning in 2007 (four sessions). These query sessions include 776 questions put to 13 ministers by 162 out of a total 493 delegates (483). The authors conducted a content analysis of these transcripts, coding questions according to whether they were "critical" of the minister, addressed local issues, or referenced the provincial constituency (483). The authors then paired these data with their 2009 dataset of delegate biographies to assess delegates' behavior as it relates to their individual characteristics. They also integrated province-level information to evaluate delegate behavior in the context of the constituency the delegate "represents" (492).

The authors estimate the effects of delegate and province-level characteristics on two dependent variables—number of times spoken and number of questions asked—using a negative binomial specification, and on three further log-transformed dependent variables—percentage of critical questions, percentage mentions of local issues, and percentage uses of the word "constituency"—using OLS. In particular, they are interested in investigating co-optation and delegate behavior through three mechanisms. First, delegates nominated by *central authorities* rather than provincial commissions should exhibit greater upward accountability in their behavior. There were 876 total candidates nominated for 493 seats in the VNA. The data includes 165 delegates who were nominated by the central VCP in Hanoi, while 711 were nominated by provincial electoral commissions (Malesky and Schuler 2010, 488). Despite this apparent variation, however, the authors also note that "the Central Election Board makes it clear that it expects the centrally nominated candidates to prevail" and mention evidence that "they resort to some level of electoral engineering to achieve this result..." (488). In general, this involves sending central nominees to less competitive districts. Relatedly, delegates from *competitive districts* are more critical than those elected in safe seats. Finally, *full-time* delegates, who also sit on legislative committees or manage provincial legislative offices, exhibit more participation, particularly concerning local issues, and more critical behavior (483). Full-time delegates—a relatively new status adopted as 30% of all delegates after the 2002 term—conduct research for and draft legislation, which gives them greater influence on policy (489). The authors argue that the influence of these factors on delegate behavior provides evidence in favor of the co-optation thesis: delegates respond to incentives and limitations imposed by Vietnamese leaders.

While providing a very thorough treatment of the actions of VNA members, the investigation of mechanisms of cooptation in Malesky and Schuler (2010) is limited by context. In particular, Vietnam's dominant-party system makes claims about the incentives actors face and the path for advancement they would likely pursue very clear. While this is advantageous for offering interpretations of VNA members' behavior, it lacks external validity to authoritarian contexts where multiple parties contend for power and resources in a less stable environment. In such a context, co-optation may appear more multi-faceted, and interpreting nonparticipation as acquiescence is at best more complicated.

4.2.2 Parties within Legislatures

Magaloni and Kricheli reiterate some of the arguments made by authors evaluating legislatures as they more deeply examine the continued prevalence of single-party authoritarian regimes worldwide (Magaloni and Kricheli 2010). Party institutionalization under autocracy, they argue, can allow the dictator to bargain over policy and mobilize support (124–125). While other institutions enable more effective monitoring of this distribution of spoils, parties can serve both to mobilize mass support and to engage and co-opt elites. As Milan Svolik notes, citing Geddes, "parties are the vehicles through which the regime rewards its supporters" (Svolik 2012, 163). Much like parties in democratic systems, which leverage organizational characteristics to provide benefits to political entrepreneurs seeking office (Aldrich 1995), parties in authoritarian systems, Svolik argues, work by hierarchically assigning tasks and benefits, controlling appointments, and selectively recruiting and repressing (Svolik 2012, 163). These "incentive structures" encourage participation and "sunk political investment" by members that engenders co-opted behavior and participation (163). Applying this theoretical argument to the conclusions drawn from the empirical evidence in the previous section, then, becoming a delegate is just one further step in a series that the VCP has designed in order to encourage sunk investment.

Svolik particularly discusses career appointments as an inducement for individuals to join parties (165). The hierarchical distribution of benefits then obtains upon joining the party, which generates an endogenously reinforcing set of incentives to continue to cooperate with and invest in the party (172). Co-optation emerging from the party, therefore, also involves party influence over other areas of the state and economy that influence the set of options available in contrast to membership, while also influencing perceptions of the party's longevity that are necessary for investing in the first place (Svolik 2012, 179, 182).

These arguments need not mean, however, that party institutions precede or supercede legislative institutions. In principle, for example, the party should select and co-opt those who are ideologically similar, and repress those whose preferences are more distant, since "party-based co-optation exploits natural creed aspirations within the general population to marginalize actual, ideological opposition" (183–184). Given this, co-optation could actually occur *after* election or selection to a legislative body like the VNA, where instead the VCP has preliminarily chosen potential political operatives but uses the legislative venue to further assess their policy preferences or ideological proclivities in order to further select candidates for advancement. In this context, as mentioned previously, active participation rather than nonparticipation would be a critical signal of compliance and co-optation.

4.2.3 Selection: Individual Characteristics & Types

In contrast to these institution-level narratives about behavior under authoritarianism, Besley's emphasis on the role of "good types" in positions of political decision-making power further highlights potential for individual MPs to vary in their legislative participation (Besley 2006). Besley demonstrates that inefficiencies and "government failure" can result from the ignorance of politicians, undue influence, or variations in the quality of leadership (48–53, 59–70). In contrast to the preceding discussions of institutional influences, Besley's key contribution is his inclusion of individual capacity and quality factors alongside an analysis of structural constraints. Besley argues that "some individuals can implement policies more cheaply or may even have more insight into what works" and furthermore, "some policy makers may be better at carrying out the citizens' wishes" (69). The question, then, revolves not only around how to establish institutions that effectively constrain policy actors, but how to develop incentives and parallel institutions (such as electoral institutions) that enable the selection of better types of policy makers, and can thereby also influence the quality of policy.

While Besley's work might appear to have only limited applications to autocratic legislatures because it assumes strong electoral institutions, his emphasis on the types of individuals, their competencies, and preferences within institutions bears further investigation even in autocratic contexts. Meritocratic selection institutions, selection for human capital, and the availability of educational opportunities might explain at least part of the variation in legislative participation even under autocracy, as these kinds of characteristics indicate more firmly held policy preferences and beliefs.

4.3 Empirical Analysis

4.3.1 Data Description

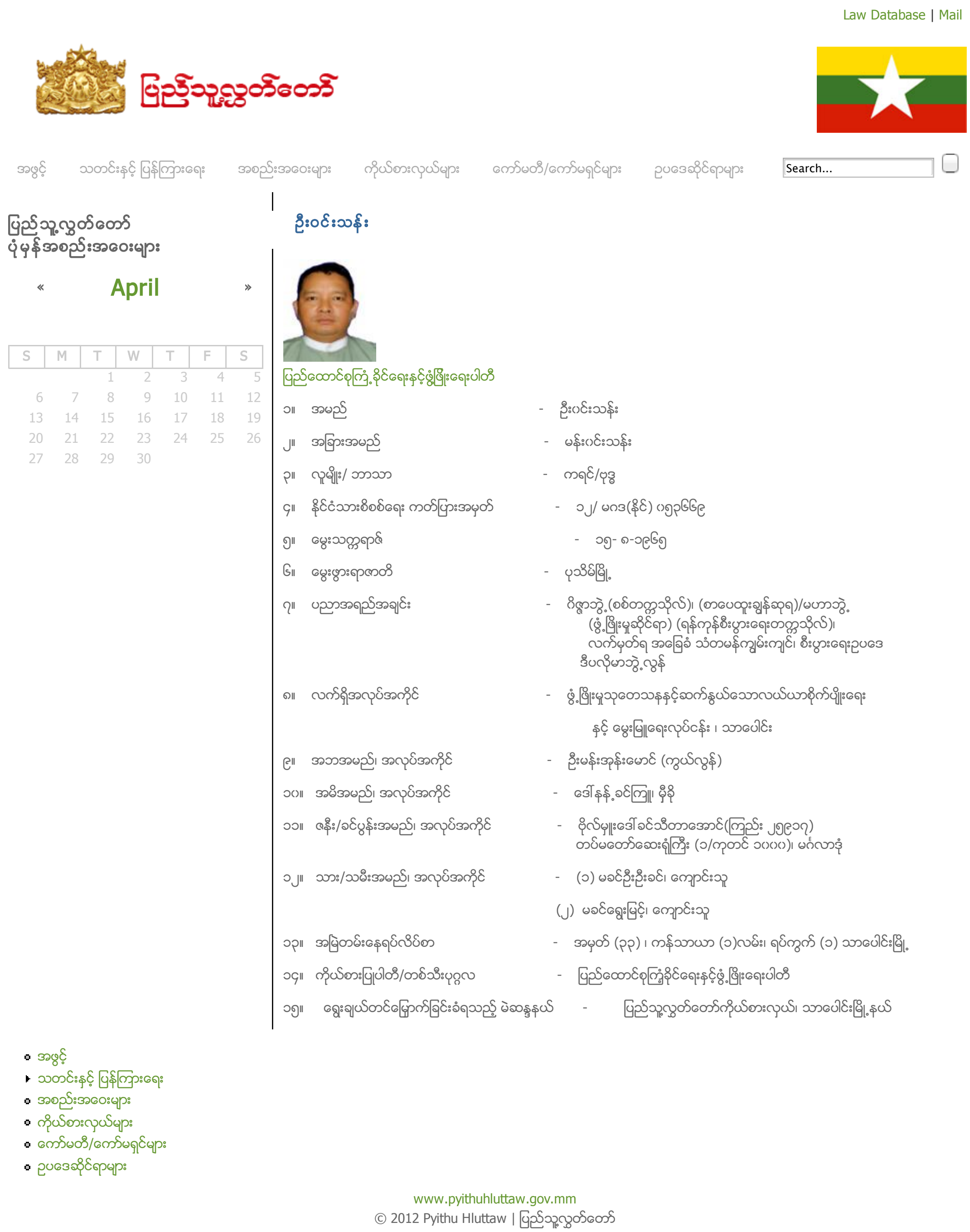
In this paper, I use data concerning the participation of the 508 delegates in Myanmar's Pyithu Hluttaw (People's Parliament, or lower house). Applying the theoretical insights from each of these literatures to the Myanmar case offers several benefits. In particular, it allows for greater variation in ethnic identification, party, and career experience than the Vietnam data afford. Because legislators are more diverse in their characteristics, they also plausibly differ to a greater extent in their co-optedness and policy preferences, providing a stronger test of the theories explored above. Likewise, Myanmar's parliament includes a greater diversity of parties than the VNA, which will facilitate conclusions that would apply to other multiparty authoritarian contexts more easily.

Following Malesky and Schuler's approach, I evaluate participation using the number of questions asked of ministers on particular policy areas. Question counts for the first 5 sessions of parliament (2011–2013) were gathered from AltSEAN (the Alternative ASEAN Network), while biographical data were synthesized from several sources, including the Open Myanmar Institute's database of MPs and the database available through the Pyithu Hluttaw itself.

Figure 4.1 below represents the type of data available through the Pyithu Hluttaw. While the analysis that follows leverages many of the variables provided in MP profiles such as the one presented below, several others will require inclusion in later analysis. Specifically, each MP profile provides information about the MP's name (and any aliases), ethnicity, religion, date of birth, place of birth, education (level of education by degrees earned, as well as subject of focus), current occupation, parents' names and occupation(s), spouse's name and occupation, children (number, sex, current occupation), permanent address, party, and constituency.

While a focus on question counts not only aligns the analysis in this paper with the investigation undertaken by Malesky and Schuler and allows for comparison between results, the theoretical justification for focusing on questions remains strong on its own. As mentioned previously, participation in an authoritarian legislature has the potential to either facilitate endogenous change by increasing policy concessions, or shore up authoritarian power by playing into co-optation. Questions asked of ministers allow for more specificity in identifying relevant policy areas, and are more aligned with the project of evaluating micro-foundations for participation since they represent the actions of single MPs with respect to ministers (unlike bill proposals) and can later be evaluated for content, whether critical or laudatory (unlike roll-call voting).

For example, during the initial parliamentary session in 2011, representative



Law Database | Mail

ပြည်သူ့လွှတ်တော်

အဖွင့် သတင်းနှင့် ပြန်ကြားရေး အစည်းအဝေးများ ကိုယ်စားလှယ်များ ကော်မတီ/ကော်မရှင်များ ဥပဒေဆိုင်ရာများ

Search...

ပြည်သူ့လွှတ်တော်
ပုံမှန်အစည်းအဝေးများ

« April »

S	M	T	W	T	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

ဦးဝင်းသန်း

ပြည်ထောင်စုကြံ့ခိုင်ရေးနှင့်ဖွံ့ဖြိုးရေးပါတီ

၁။	အမည်	- ဦးဝင်းသန်း
၂။	အခြားအမည်	- မန်းဝင်းသန်း
၃။	လူမျိုး/ ဘာသာ	- ကရင်/ဗုဒ္ဓ
၄။	နိုင်ငံသားစိစစ်ရေး ကတ်ပြားအမှတ်	- ၁၂/ မဂဒ(နိုင်) ဂ၅၃၆၆၉
၅။	မွေးသက္ကရာဇ်	- ၁၅- ၈-၁၉၆၅
၆။	မွေးဖွားရာဇာတိ	- ပုသိမ်မြို့
၇။	ပညာအရည်အချင်း	- ဂိုဏ္ဍာဘွဲ့(စစ်တက္ကသိုလ်)၊ (စာပေထူးချွန်ဆုရ)/မဟာဘွဲ့ (ဖွံ့ဖြိုးမှုဆိုင်ရာ) (ရန်ကုန်စီးပွားရေးတက္ကသိုလ်)၊ လက်မှတ်ရ အခြေခံ သံတမန်ကျွမ်းကျင်၊ စီးပွားရေးဥပဒေ ဒီပလိုမာဘွဲ့လွန်
၈။	လက်ရှိအလုပ်အကိုင်	- ဖွံ့ဖြိုးမှုသုတေသနနှင့်ဆက်နွယ်သောလယ်ယာစိုက်ပျိုးရေး နှင့် မွေးမြူရေးလုပ်ငန်း ၊ သာပေါင်း
၉။	အဘအမည်၊ အလုပ်အကိုင်	- ဦးမန်းအုန်းမောင် (ကွယ်လွန်)
၁၀။	အမိအမည်၊ အလုပ်အကိုင်	- ဒေါ်နန့်ခင်ကြူ၊ မှီခို
၁၁။	ဇနီး/ခင်ပွန်းအမည်၊ အလုပ်အကိုင်	- ဗိုလ်မှူးဒေါ်ခင်သီတာအောင်(ကြည်း ၂၅၉၁၇) တပ်မတော်ဆေးရုံကြီး (၁/ကုတင် ၁၀၀၀)၊ မင်္ဂလာဒုံ
၁၂။	သား/သမီးအမည်၊ အလုပ်အကိုင်	- (၁) မခင်ဦးဦးခင်၊ ကျောင်းသူ (၂) မခင်ရွှေးမြင့်၊ ကျောင်းသူ
၁၃။	အမြဲတမ်းနေရပ်လိပ်စာ	- အမှတ် (၃၃) ၊ ကန်သာယာ (၁)လမ်း၊ ရပ်ကွက် (၁) သာပေါင်းမြို့
၁၄။	ကိုယ်စားပြုပါတီ/တစ်သီးပုဂ္ဂလ	- ပြည်ထောင်စုကြံ့ခိုင်ရေးနှင့်ဖွံ့ဖြိုးရေးပါတီ
၁၅။	ရွေးချယ်တင်မြှောက်ခြင်းခံရသည့် မဲဆန္ဒနယ်	- ပြည်သူ့လွှတ်တော်ကိုယ်စားလှယ်၊ သာပေါင်းမြို့နယ်

- အဖွင့်
- သတင်းနှင့် ပြန်ကြားရေး
- အစည်းအဝေးများ
- ကိုယ်စားလှယ်များ
- ကော်မတီ/ကော်မရှင်များ
- ဥပဒေဆိုင်ရာများ

www.pyithuhluttaw.gov.mm
© 2012 Pyithu Hluttaw | ပြည်သူ့လွှတ်တော်

Figure 4.1: MP Profile of U Win Than, USDP representative to the Pyithu Hluttaw for Thabaung constituency

U Soe Thein from Kalewa in Sagaing asked the following question:

> ကျောက်မီးသွေးသုံးစက်ရုံများနိုင်ငံတစ်ဝှမ်းတွင်တည်ဆောက်၍နိုင်ငံသားများအားအဖိုးနှုန်း
> ချိုသာသောလျှပ်စစ်မာတ်အားဖြန့်ဖြူးရောင်းချပေးမည့်အစီအစဉ်ရှိမရှိ။
>
> "Do you have plans to make available/extend relatively cheap electrical power to citizens located in areas throughout the country where coal processing plants are built?"

This type of question illustrates the potential to raise critical concerns about regime policies: the USDP and its related military entities have forged relationships with foreign firms to establish industrial plants, like coal-processing plants, that pose significant health and environmental risks to citizens. At the same time, however, representative U Soe Thein is not from a traditional opposition party like the National League for Democracy (NLD), which might be interested in constituent service for reasons of principle or commitment to democratic ideals. Rather, U Soe Thein is a representative from the National Unity Party—a party created to support General U Ne Win that competed against the NLD in the infamous 1990 elections, after which NLD party members were rounded up for arrest and never allowed to take power. While further analysis of the content of questions like these, asked in parliament, is critical to understanding the true nature of representation in Myanmar's authoritarian legislature, this example indicates that the questions themselves have the potential to signal meaningful participation.

4.3.2 Negative Binomial vs. Zero-Inflated Negative Binomial: Theoretical Justifications

In addition to using parliamentary data from Myanmar, I diverge from the modeling choices adopted in Malesky and Schuler's initial analysis of the VNA data. Malesky and Schuler note that "the NBREG [negative binomial regression] is preferable to a *Poisson* distribution for capturing the count nature of the data because the high number of delegates with zero speeches leads to over dispersion in the data. In both cases, the unconditional variance is higher than the mean, which violates the *Poisson* assumption that they are equal..." (Malesky and Schuler 2010, 494). Beyond this statistical point, however, they also seek to make inferences from the number of "zeros" in their data, saying that "the decision not to speak implies an individual choice by delegates" where "the high number of nonspeakers provides important insights into how a regime such as Vietnam might use its Assembly for co-optation but still maintain control over the proceedings" (493). This paper endeavors to more rigorously investigate the role of "zeros" in legislative participation and the story they tell about co-optation in authoritarian parliaments. I find that there are strong theoretical reasons and preliminary empirical evidence to consider a zero-inflated model.

While the negative binomial specification addresses apparent overdispersion in models like Malesky and Schuler's, the overdispersion evident in participation in authoritarian parliaments is caused by "excess" zeros in the dependent variable, which instead support a zero-inflated negative binomial. The zero-inflated estimation strategy posits a dual data-generating process for zeros arising in the data, where some observations are "always zeros" and some are merely "sometimes zeros." In terms of Malesky and Schuler's data, co-optation may cause some delegates to be "always zeros"—individuals who never speak because their presence in the legislature is contingent on the support of the Vietnamese Communist Party. Others, meanwhile, may simply not speak due to lack of interest, expertise, or information in a policy area during a particular query session. This is consistent with some trends in the data, such as the fact that the Minister of Agriculture and Rural Development and the Minister of Industry and Commerce received many questions

—119 and 105 respectively—whereas other ministers (Transportation, Home Affairs) received very few (Malesky and Schuler 2010, 493). This disparity supports the notion that there are some areas of governance that are relevant to delegates and/or some in which they feel they have expertise and can participate in querying. This should be no less true for my data from the Myanmar legislature.

A zero-inflated negative binomial allows for this potential diversity among the "zeros" in the data. Likewise, the zero-inflated negative binomial more reasonably models a "hurdle" that delegates confront when choosing to speak or abstain. Unlike the negative binomial, which posits a single process governing the number of times a delegate speaks, the zero-inflated negative binomial allows for a logistic process governing the probability that a delegate speaks at all, and a separate non-truncated negative binomial governing the number of times that a delegate speaks, given that s/he has spoken.

Specifically, the zero-inflated negative binomial (ZINB) distribution is defined as:

$$P(Y = y_i) = \begin{cases} p_i + (1 - p_i)(1 + \frac{\lambda}{\tau})^{-\tau}, & \text{if} \quad y_i = 0 \\ (1 - p_i)\frac{\Gamma(y_i + \tau)}{y_i!\Gamma(\tau)}(1 + \frac{\lambda}{\tau})^{-\tau}(1 + \frac{\tau}{\lambda})^{-y_i}, & \text{if} \quad y_i > 0 \end{cases}$$

where the negative binomial segment approximates a Poisson as $\tau \to \infty$.

The regression model therefore takes the form:

$$\log(\lambda_i) = x_i^\top \beta \qquad \text{and} \qquad \text{logit}(p_i) = z_i^\top \gamma, \qquad (i = 1, 2, ..., n)$$

Theoretically, these elements of the zero-inflated negative binomial align well with the expected behavior of MPs in authoritarian legislatures. Their behavior should be the result of a utility maximization calculation where minimal participation is bounded at zero, but this need not imply that only one "type" of individual does not speak. Malesky and Schuler rely on an assumption that the speakers and non-speakers in their data are different delegates, whereas there is no reason to assume that even "responsive" delegates who are not co-opted by the ruling always speak in every session. In particular, a zero-inflated model that distinguishes the "sometimes zero" delegates is better suited to indicate the factors that explain why certain delegates speak more and are more "responsive," given that they have spoken. Using a negative binomial alone constrains the analysis of responsiveness versus co-optation to a single measure—the number of times a delegates speak—rather than allowing for co-optation to function both as a depressant on the probability that a delegate speaks at all and a constraint on how many times they speak when they choose to do so, as the zero-inflated model does. Furthermore, the variables that Malesky and Schuler's work identifies as indicating co-optation (being nominated by the VCP Central Committee, for example) do not suggest "elasticity" in the concept of co-optation that a negative binomial specification implies. In that formulation, either co-optation is less "binding" on some delegates who do choose to speak or co-optation merely indicates whether a delegate speaks at all, in which case a logistic model would suffice. A zero-inflated negative binomial provides a more flexible and robust way to investigate how cooptation functions relative to policy preferences in authoritarian legislatures.

These types of results may not be consistent with a story in which co-optation does not operate through or exclusively within the legislature per se, but rather within the party structure itself. For example, if Svolik's theory of party co-optation under autocracy is correct, party members who are nominated by the ruling party or dictator should be more co-opted than those who are actively engaged in politics but not beholden to the party structure for nomination to run for a legislative seat. The Myanmar legislative data are ideally suited to address this theoretical concern, however, since both the ruling Union Solidarity and Development Party (USDP) and the former military have credible claims on power. Further extensions on the work in this paper will allow for a more thorough examination of this dynamic.

4.3.3 Negative Binomial vs. Zero-Inflated Negative Binomial: Empirical Justifications

Empirically, furthermore, finding overdispersion when using count data does not suffice to justify a negative binomial specification, as multiple processes may result in overdispersion. As Zorn (1996) indicates, positive contagion (in which some values values of y_i are not independent), particularly where a larger than expected number of zeros occur, may resemble overdispersion. In particular, "excess zeros caused by a two-part data generating process 'trick' [tests for overdispersion] into indicating the presence of overdispersion when, controlling on the transition stage, little or none is present" (13). More pointedly, as Zorn describes, "because they rely on the initial estimation of a (misspecified) Poisson model and fail to take account of the conditional nature of the count variable, these tests are of no value when the data generating process takes a dual regime form" (13). While the zero-inflated negative binomial allows for overdispersion in the count portion of the model as well (because it does not assume zero-truncation), it fundamentally accounts for this possible dual-data-generating process. A zero-inflated model distinguishes between zeros arising because of the dual-data-generating process relative to contagion (12). In Malesky and Schuler's analysis, therefore, estimating a zero-inflated negative binomial should better account for distinctions between co-opted delegates and those who simply do not have much to say.

Visualizing the data on participation in the Myanmar Pyithu Hluttaw provides further support for a model accounting for a "hurdle" or "transition." As Figure 4.2 below indicates, a logged version of the dependent variable, questions (including a small constant), shows bimodality. This suggests that the rate of decline in counts of the dependent variables are not as steep or steady as a Poisson process would indicate.

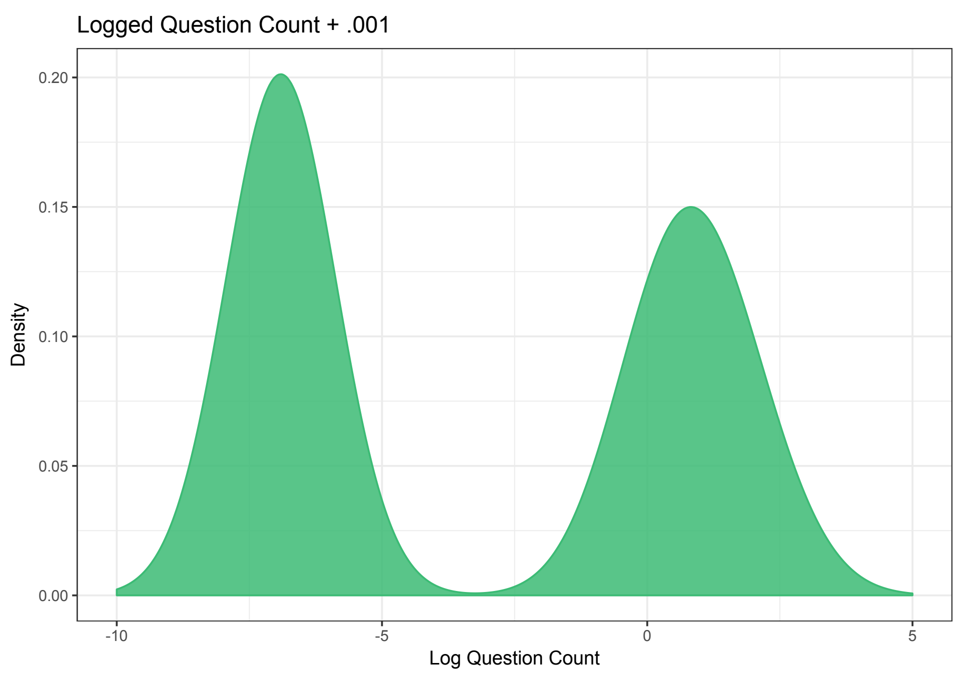


Figure 4.2: Logged Dependent Variable Shows Bimodality in Participation Data

4.3.4 Estimation Strategy

In accordance with these results, I estimate a zero-inflated negative binomial model of participation in Myanmar's Pyithu Hluttaw (lower house), using the number of questions asked of ministers in the 2011–2013 parliamentary sessions as a dependent variable. I maintain many of the same demographic explanatory variables that existed in the Malesky and Schuler analysis in order to assess how explanatory factors might differ across model specifications. In particular, I utilize indicators of age, sex, party affiliation, ethnicity, and prior career sector. Because the ZINB specification is much more computationally intensive, however, I combine several categories within the explanatory factor variables. Specifically, for the party variable, I combine ethnic and opposition parties, aside from the main opposition party (the NLD), and allow the military to be a separate category. I reduce ethnicity to merely Burman vs. non-Burman, combining nearly 30 ethnic groups each with relatively few corresponding representatives. Finally, careers are reduced to 4 sectors—civil service, educated professionals, agriculture, and others (arts, media, tourism, etc.) —where military is the omitted category. I omit region control variables because they generate issues with separation but do not contribute much to the analysis otherwise. Furthermore, in accordance with the recommendation in Gelman et al. (2008), variables are demeaned and rescaled to ease computation. In addition, pursuing this estimation in a Bayesian framework allows for the specification of relatively more informative prior distributions for these covariates, which can work to control issues of quasi-perfect separation. That is, with a series of factor variables on the right-hand side in the analysis, and with so few MPs who do participate in the parliamentary sessions, the model often "assumes" it can perfectly predict participation for some categories of individuals, leading to unrealistic estimates. For example, in this analysis, MPs with a prior career in agriculture are few, and many of them participate, leading to an overestimation in the model of how often such a hypothetical MP would speak. Utilizing more informative priors can effectively "bound" the estimates to avoid these unrealistic outcomes.

Theoretically, the ZINB specification allows for a distinction between the predictors of MPs speaking at all, relative to the predictors of actors speaking a certain number of times given that they speak. In specifying priors for each of these parts, I use the weakly informative priors suggested in Gelman et al. (2008) for the "zero" or logistic part of the model ($\mathrm{Cauchy}(0, 2.5)$, with $\mathrm{Cauchy}(0, 10)$ for the intercept), and weakly informative ($N(0, 10)$) priors for the rescaled variables on the count side of the model. Distinguishing these two parts or sides of the model should allow for a more nuanced investigation of co-optation. For example, in Malesky and Schuler (2010), the assumption is that non-speaking representatives are co-opted into silence. Distinguishing between "always zeroes" who never speak and "sometimes zeroes" who speak only on particular issues or at particular times can help identify those who, perhaps, are driven by policy interests. Again, in Malesky and Schuler's analysis, desires to move up in or be part of the regime party drive co-optation behavior. Yet if co-optation is best aligned with silence, we should observe regime party as a major driver of whether to speak or not, in the negative direction. One could also imagine a logic of co-optation, however, whereby a desire to please party elites leads to more speaking in order to "pander" or demonstrate alignment with party ideals. In this case, party could positively influence both the probability of speaking, and the probability of speaking more frequently. Yet if instead MPs do attempt to effect policy change from within an authoritarian parliament, prior career background in particular policy areas or sectors could positively predict speaking. Because these policy areas correspond to a limited number of ministers, it would also likely not positively predict speaking a large number of times.

4.3.5 Weakly Informative Priors Results

The following graphs provide the posterior distributions for each factor level of the covariates. To compensate for the complexity of the model and the number of parameters relative to the number of observations in covariate categories, I estimate this initial model with weakly informative priors using Stan, with 200,000 iterations and a burn-in period of 50,000. The densities for many coefficients center on or near 0. Particularly on the zero side of the model, there are many "long tails" in the posterior distributions evidencing the effects of partial separation. For example, because there are so few women in the dataset and the excluded category is male, the sex variable reflects uncertainty in a long tail from the separation process. Likewise, the covariate for having a prior career in agriculture has the most pronounced separation as a result of few individuals in that category and most of them having spoken multiple times. This leads to the model concluding, effectively, that someone with a past career in agriculture will always speak, even though this does not align with reasonable expectations.

From the zero side of the model, we can primarily see that all included career categories positively predict speaking. This could indicate, as previously discussed, that MPs are motivated to speak out of policy concern. Because former civil service members also are more likely to speak, however, these motivations could be disaggregated by prior career; that is, former civil service members may indeed be speaking out of career concern, wishing to show party loyalty or alignment in order to move up in political ranks. Furthermore, the category encompassing educated professionals in business, law, industry, education, etc., have a very positive association with speaking. This could indicate that those with greater expertise in policy areas, or greater education overall, are more likely to speak.

On the count side of the model, the effects are reversed. Career variables are weakly negatively associated with speaking a greater number of times. That is, while those in the career categories are more likely to speak, they are less likely to speak a large number of times. This is consistent with the notion that these individuals are policy motivated, and will likely eventually speak, but only to the few ministers responsible for the policy areas of interest. Furthermore, surprisingly, party is not a strong predictor of increased speaking behavior, or of complete silence. Members of the military are not likely to speak at all, or speak regularly if they do, but party only weakly supports the probability to speak, with the mean near 0, especially with respect to the regime party, the USDP. On the count side of the model, however, USDP and especially the main opposition party, the NLD, are associated with speaking significantly less, even having spoken. While these preliminary results suggest interesting outcomes at odds with prior theoretical expectations, drawing conclusions is challenging because of the quasi-perfect separation issue.

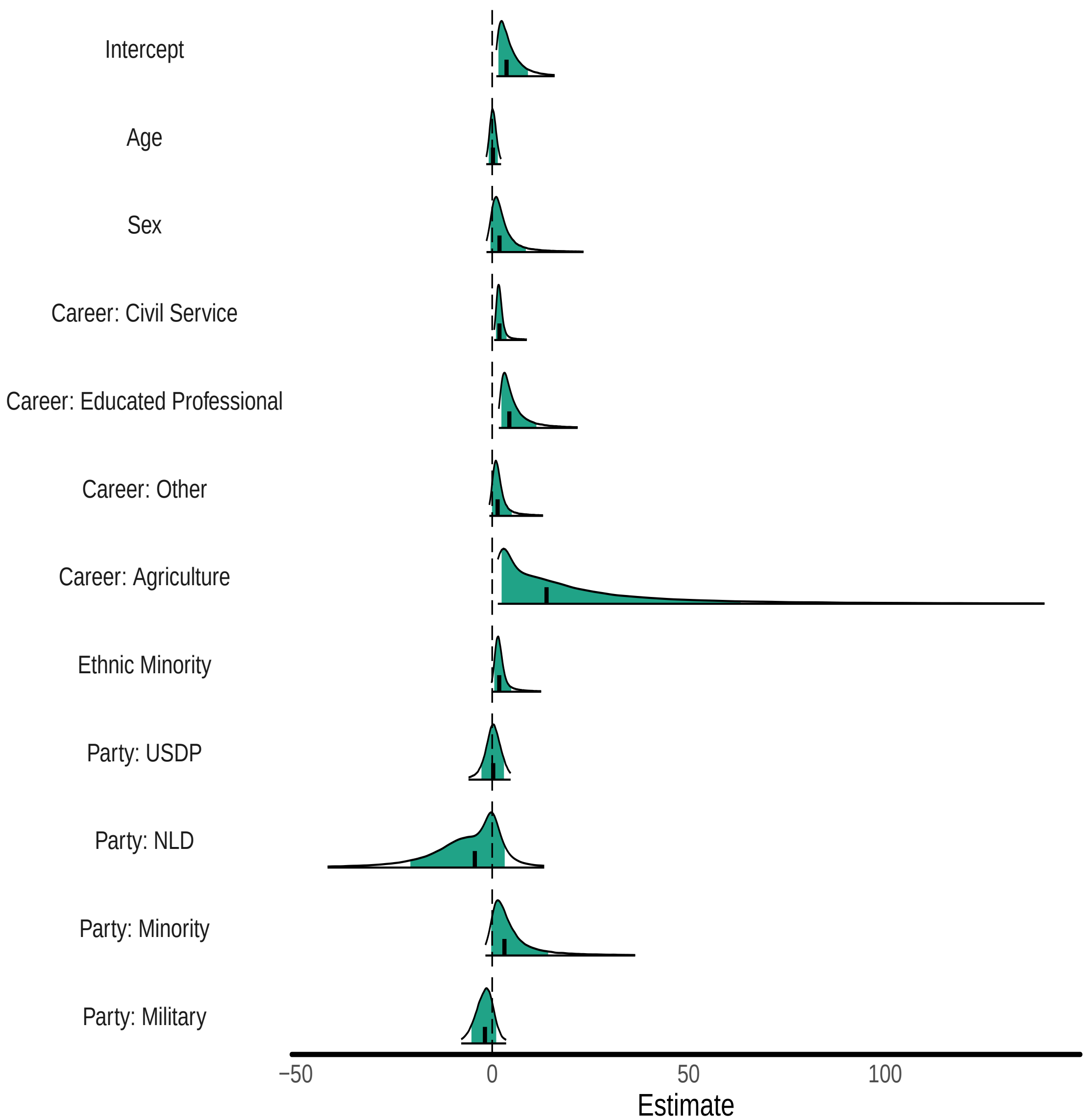


Figure 4.3: Zero Coefficient Posteriors with 80% Credible Intervals

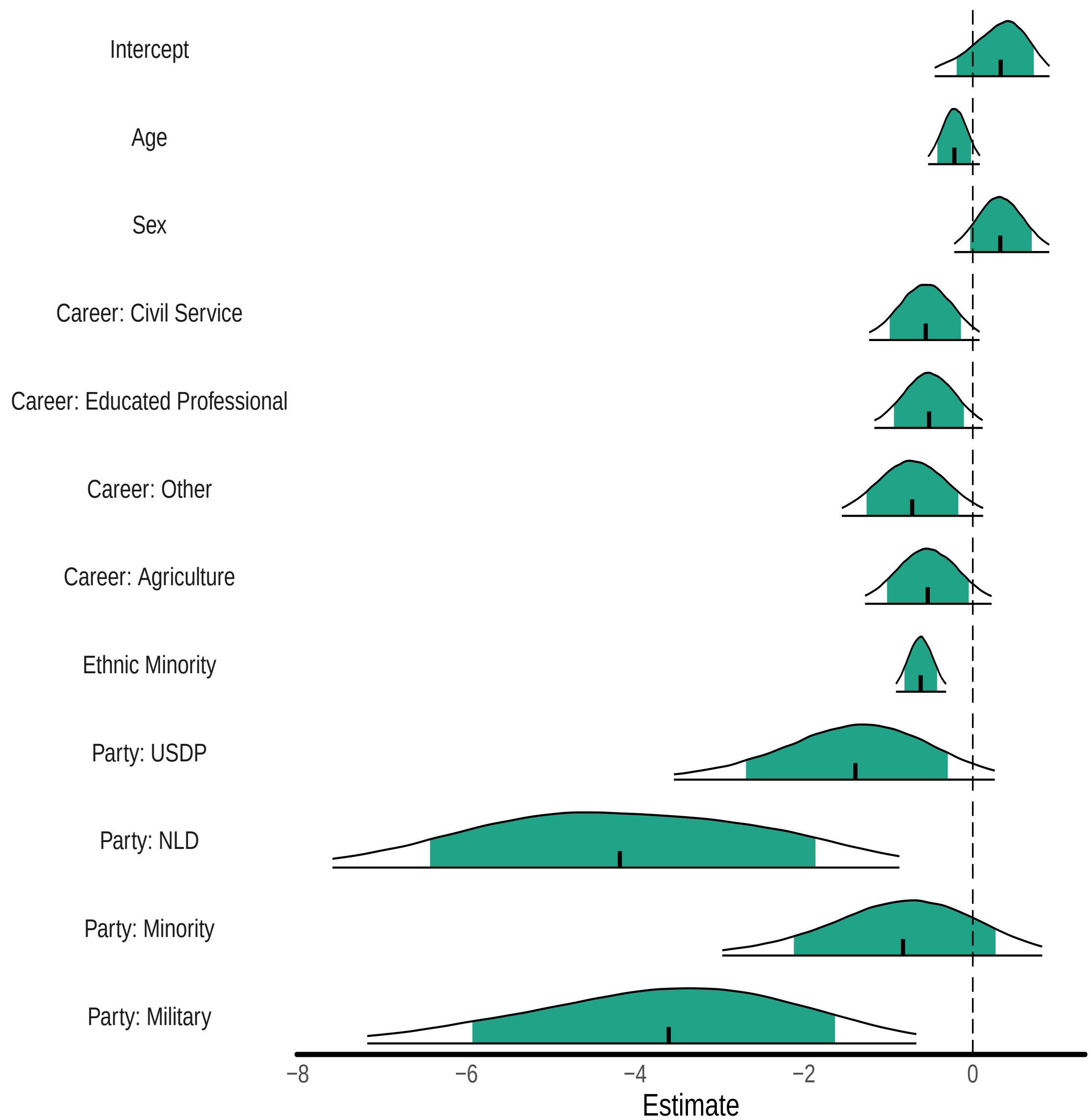


Figure 4.4: Count Coefficient Posteriors with 80% Credible Intervals

4.4 Refining Priors: Text-Based Elicitation

The approach in the previous preliminary analysis utilizes only minimally informative priors that are taken from standard statistical approaches, rather than tailored priors that reflect the state of knowledge about how these MP characteristics should influence participation in the legislature. In order to generate more principled and refined priors for this analysis, I pursue a novel elicited-priors approach leveraging data from text sources, specifically news outlets with coverage of Myanmar politics. These divergent priors from several different sources will then be combined using a Dirichlet Process method in order to identify the various "schools of thought" that exist among the news sources, and apply all of these perspectives to evaluate MP participation (see Appendix for further discussion of this method). Rather than using an interview or focus group method for eliciting priors in this case, this news-based approach represents an opportunity to incorporate a wider array of perspectives without concerns for obfuscation or self-censorship dynamics in interpersonal elicitation. That is, because these priors are collected post hoc, without interacting with newspaper reporters or editorial boards, the "experts" in question do not have the opportunity to directly evade questioning in an elicitation setting. As will be discussed in later sections, this text-based approach can later be evaluated against and supplemented with the results of a survey-based elicitation technique. Preliminary validation with elicited priors from three experts via online survey are included.

Elicited priors utilize the knowledge of "experts" to generate refine and improve posterior estimates in Bayesian analyses. In authoritarian regimes, in particular, however, relying on individuals who have preferential access to politics or whose credentials would typically qualify them as "experts" might generate significantly biased expectations. Likewise, because the information environments in authoritarian regimes are so restrictive, equally plausible but differing perspectives might arise that would suggest expanding the pool of possible "experts" in order to generate the most informed perspective. This is analogous to a situation in which individual in authoritarian regimes are receiving correlated signals about the true state of the world—while none are necessarily complete perspectives, each individual has some component of the "truth." In the service of expanding this pool of "experts," this paper incorporated elicited priors from news sources. In mapping these sources onto the more traditional methods of elicitation, one can imagine a news source's editorial board as the subject of elicitation, and the published information in the paper as the set of information that would be used to generate a prior probability distribution. In this case, because the published articles represent more diffuse information than what might be elicited in an interview, elicitation will focus on identifying prior means and variances for each covariate factor level, rather than fully specifying a prior probability distribution.

4.4.1 Text Sources for Elicitation

For this analysis, I scraped text from 4,126 articles related to Myanmar's parliament. To do this, I first identified important Myanmar sources, such as *The Irrawaddy* and *Democratic Voice of Burma*, and scraped articles directly from their website with the search "parliament AND (Myanmar or Burma)" for the time period 2010–2013. This time period encompasses the same time period as the question count data, and also includes the period from 2010–2011 before Myanmar's parliament was newly established, which could encompass articles that express expectations for behavior in the parliament. Second, I scraped archived articles from aggregators (such as East View Information Services) using the same search conditions, and selected sources that had the greatest amount of coverage of Myanmar's parliament.

In addition, to ensure that a diversity of sources were included, I added coverage from *New Light of Myanmar*, the English-language version of the main Myanmar government news source *Myanmar Alinn*, even though only 8 articles had coverage of the parliament. Once all of these articles were scraped, I did further validation to ensure that all articles did pertain to Myanmar's parliament. For non-Myanmar sources, I included only articles that contained "Myanmar," "Burma," and relevant terms for parliament in their titles. This brings the total number of articles to 4,042. The number of articles by source and source type are reflected in Table 4.4.1 below.

Articles captures through this search range in topic from announcements about NLD party reorganization to the candid thoughts of an Indian diplomat, Mitra Vashishta, about the prospects for democratization and the shortcomings of the NLD that were caught in a Wikileaks release. Two articles discussing these topics, both from 2010 and featured in the *Democratic Voice of Burma* illustrate the ways in which observers may have formed prospective opinions about behavior in parliament: the NLD had boycotted the initial elections into the 2011 session of parliament as a result of legal restrictions, and if the general sentiment were that they were in a weakened position relative to newly institutionalizing autocrats, one might expect that NLD members later elected to the Pyithu Hluttaw via by-elections would participate less, for example. Likewise, publications like Thailand's *The Nation* published articles about injuries induced by police at an anti-mining protest in northwestern Myanmar in 2012. This type of article might support a number of conjectures concerning participation: perhaps MPs with past careers in the police force are more vocal in parliament in line with their enforcement of anti-protest regulations,[1] or perhaps some educated professionals who are part

[1]The police have traditionally operated separately from Myanmar's powerful military, so their enforcement of regulations against group congregation could be interpreted as alignment with military objectives, or co-optation.

News Source	Country	Type	Num. Articles
Agence France Presse	France	International wire service	219
Associated Press	USA	International wire service	31
Bangkok Post	Thailand	Independent	32
Democratic Voice of Burma	Myanmar	Opposition, Exiled	1,860
The Hindu	India	Independent	16
Irrawaddy	Myanmar	Opposition, Exiled	1,235
Mizzima	Myanmar	Opposition, Exiled	354
Narinjara	Myanmar	Ethnic (Rakhine)	127
The Nation	Thailand	Independent/Anti-Thaksin	52
New Light of Myanmar	Myanmar	Government	8
Shan Herald	Myanmar	Ethnic (Shan)	50
Xinhua	China	Government	57

Table 4.1: Newspaper Sources and Article Counts

of the business community participate more in an effort to protect their economic interests. The elicitation process described in the following section presents one approach to evaluating these types of relationships between socio-political dynamics and legislative behavior.

In addition, all included articles are in English, whether natively or by translation (as in the case of Xinhua, which includes articles both from Xinhua's China coverage and their English-language Hong Kong coverage). Including only articles in English both increases the number and diversity of news sources and ensures that the target audience for the sources is more comparable. A further extension of this elicitation approach could conduct text analysis in Burmese, although significant difficulties with natural language processing (NLP) in the Burmese language present challenges. These challenges are further discussed in the Appendix.

Covariate	Search Terms
Party: None	independent, independent member, independent candidate
Party: USDP	USDP, Union Solidarity and Development Party, Union Solidarity
Party: NLD	NLD, opposition party, democratic opposition, national league, Suu Kyi
Party: Ethnic/Minority	ethnic party, opposition party, minority party
Party: Military	military member, military MP
Ethnicity: Burman	Burman, Barmar, Bamar
Ethnicity: Minority	ethnic, minority, race, ethnicity
Career: Military	military, armed forces, army, generals
Career: Civil Service	civil service, civil servant, government worker, government employee
Career: Edu. Professional	business, industry, economics, lawyer, doctor, engineer, education, teacher, educated, medicine, law, legal, educated professional
Career: Other	art, media, tourism, artist, travel, tourist, police
Career: Agriculture	agriculture, agricultural, crop, farm, farmer, farming
Sex	female, women, sex, woman, gender, daw, woman MP, female member

Table 4.2: Article Search Terms

4.4.2 Eliciting Prior Moments: Sentiment and Word Count

In order to elicit prior information from these news sources, I focus on eliciting prior means and variances related to each covariate level. To conduct this elicitation, for each covariate level, I restrict the text corpus to include only articles that contain words or phrases related to that covariate level. These search terms are reflected in Table 4.4.2 below.

To identify means and variances that reflect each source's prior opinion on the effect of each covariate on the propensity of MPs to ask questions in parliament, I use both sentiment analysis and direct word counts. Sentiment analysis forms the basis for prior means and variances on the logit part of the ZINB model. For the sentiment analysis, I evaluate the positive or negative sentiment of each word in an article by matching the words to Bing Liu's sentiment dictionary (Liu 2004), then calculating net sentiment for each article using (number of positive words minus the number of negative words). The prior mean for the logit side of the model is then the sample mean of sentiment across all articles for a given source that relate to a particular covariate level. Because sentiment can range from negative to positive, it encompasses both situations that encourage participation and those that would depress it. For example, if a source indicates overwhelmingly positive sentiment for articles relating to the NLD party, we might imagine that the environment is more open for NLD participation and those representatives are more likely to speak. The prior variances for each "expert" on the logit side of the model simply take sample variance of sentiment across all articles for each "expert" source as well.

As can be seen in the plot of overall sentiment by news source below, which illustrates sentiment (y-axis) by article (x-axis) for each source, sentiment varies significantly across articles and across sources in directions that align with expectations. Opposition/independent sources like *Democratic Voice of Burma (DVB)* and *The Irrawaddy* have mixed positive and negative coverage of the parliament, whereas, for example, *New Light of Myanmar* coverage is overwhelmingly positive.

Likewise, viewing sentiment regarding the NLD, for example, demonstrates variation in coverage across sources. While coverage of the main democratic opposition party is overwhelmingly positive in sources like *The Bangkok Post* and *The Hindu*, it is much more mixed in opposition sources and international sources. *New Light of Myanmar* does not cover the NLD at all in their articles about parliament. Aligning these measures of sentiment with expectations of each source, this would indicate that neighboring country sources might anticipate NLD members to participate more in parliament, whereas other international and opposition sources would have more mixed expectations.

On the count side of the model, word count is used to measure mean and variance. Inclusion in the limited corpus related to a particular covariate level means that at least one word from the search must be included in each article. Once an article is included, the number of words or 2-grams relevant to the covariate level

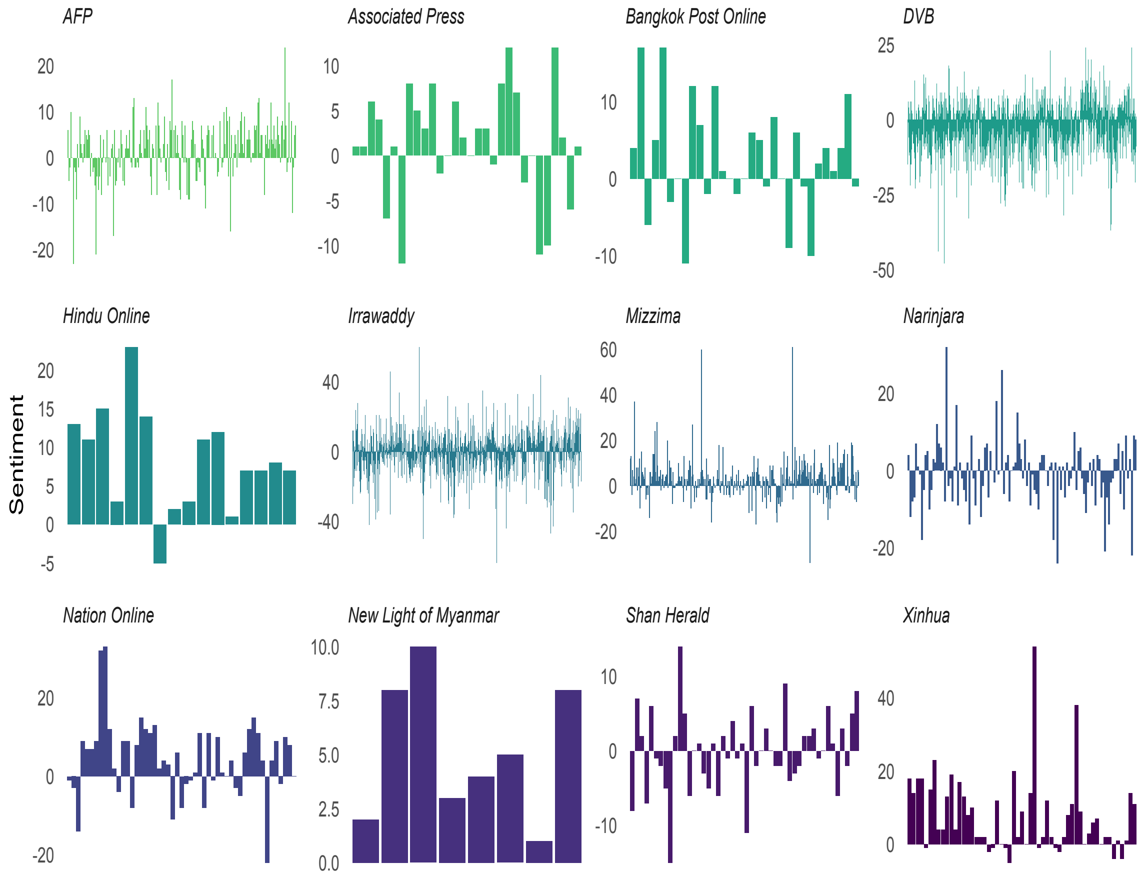


Figure 4.5: Overall Sentiment by News Source

are counted in that article. Prior mean for a covariate level for a given source is given by the sample mean for the number of words or 2-grams across all articles by that source. Prior variance is the variance of word count across all articles for that source. As on the logit side of the model, sources vary in the word counts that they allocate to each covariate, as can be seen in the figure below which represents word count related to NLD for each source.

For both the logit and count sides of the model, sentiment and word count values are rescaled to be in the appropriate range (e.g., a set of articles may contain 1,529 mentions of "NLD" but the maximum number of questions asked is 14, so

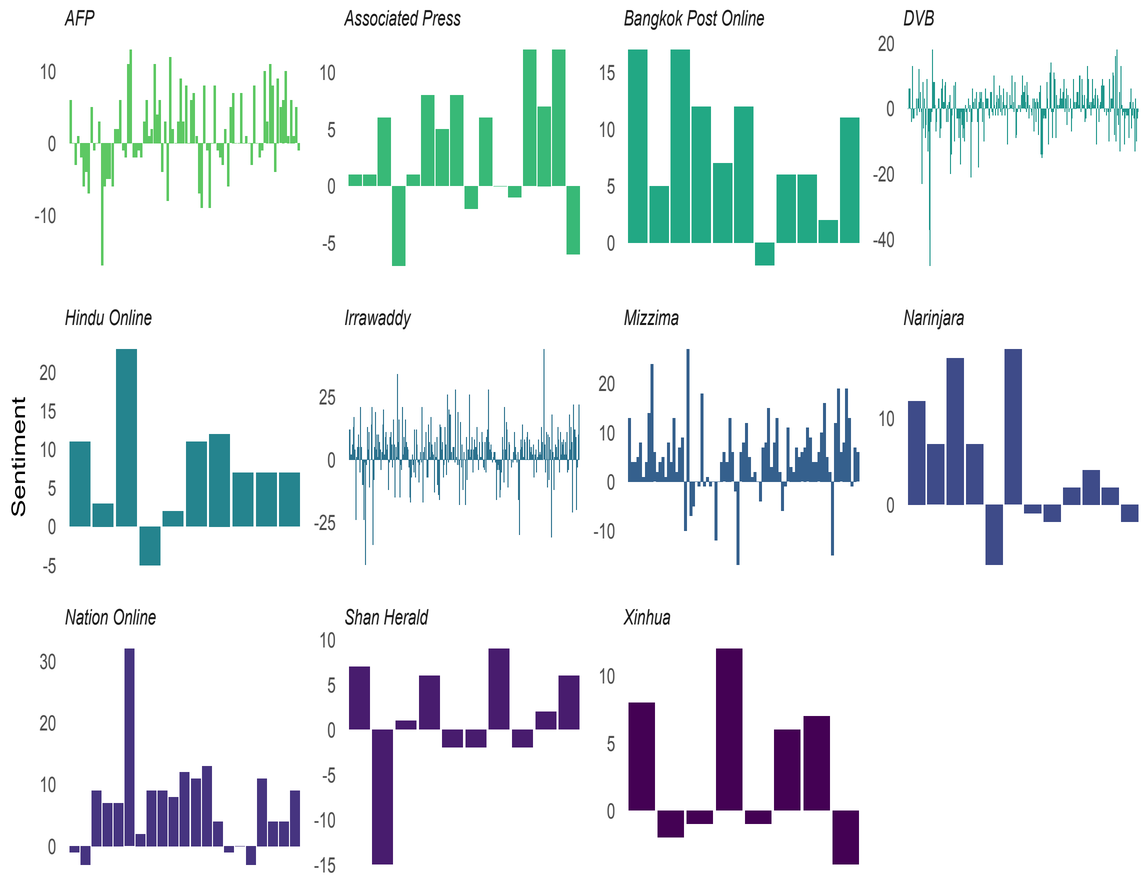


Figure 4.6: Sentiment by News Source, party = "NLD"

this word count value would be rescaled to match the range of question counts). Once these values are rescaled, a prior mean and prior variance is generated for each covariate level on both the logit and count sides of the model. The figure below illustrates these differing prior means and variances for each source, on both sides of the ZINB model, related to NLD. This example illustrates the significant variation that occurs in particular on the logit (zero) side of the model where prior means differ by source and variances do not overlap for the most part. These differences will be leveraged through their aggregation in the Dirichlet Process and will be useful in disentangling the "always zeros" from the "sometimes zeros."

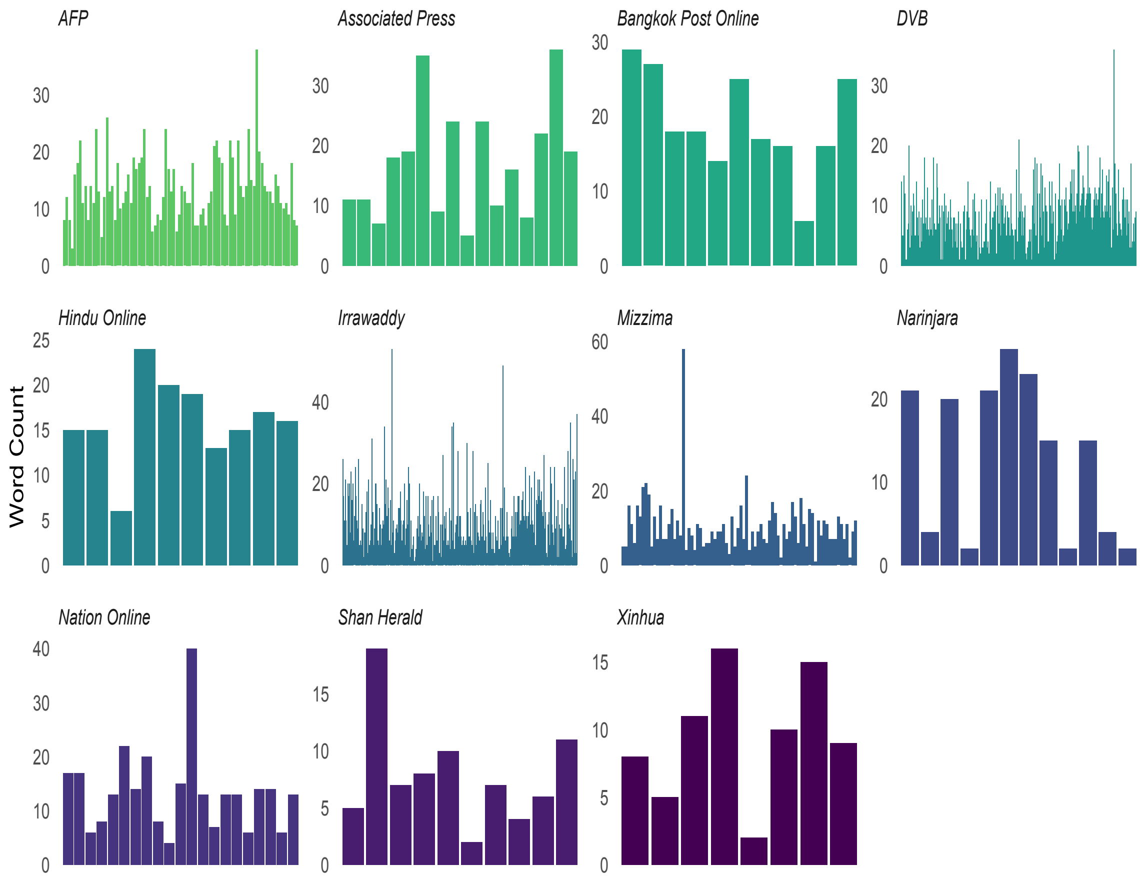


Figure 4.7: Word Count by News Source, party = "NLD"

4.4.3 Dirichlet Process Approach with ZINB Model

Having elicited these prior means and variances from each news source related to the data, I then aggregate priors using a Dirichlet Process approach (as discussed in the Appendix), and as represented in the formulation below. From the perspective of the Dirichlet Process, the elicited means and variances from each expert j, μ_j^* and Σ_j^*, are part of an underlying normal-inverse-Wishart data-generating process,[2] and the concentration parameter α for the Dirichlet Process is initially set to 1 to reflect a prior that all "expert" sources are at first their own cluster, or repre-

[2]This is the conjugate prior for a multivariate normal distribution where mean and variance are unknown.

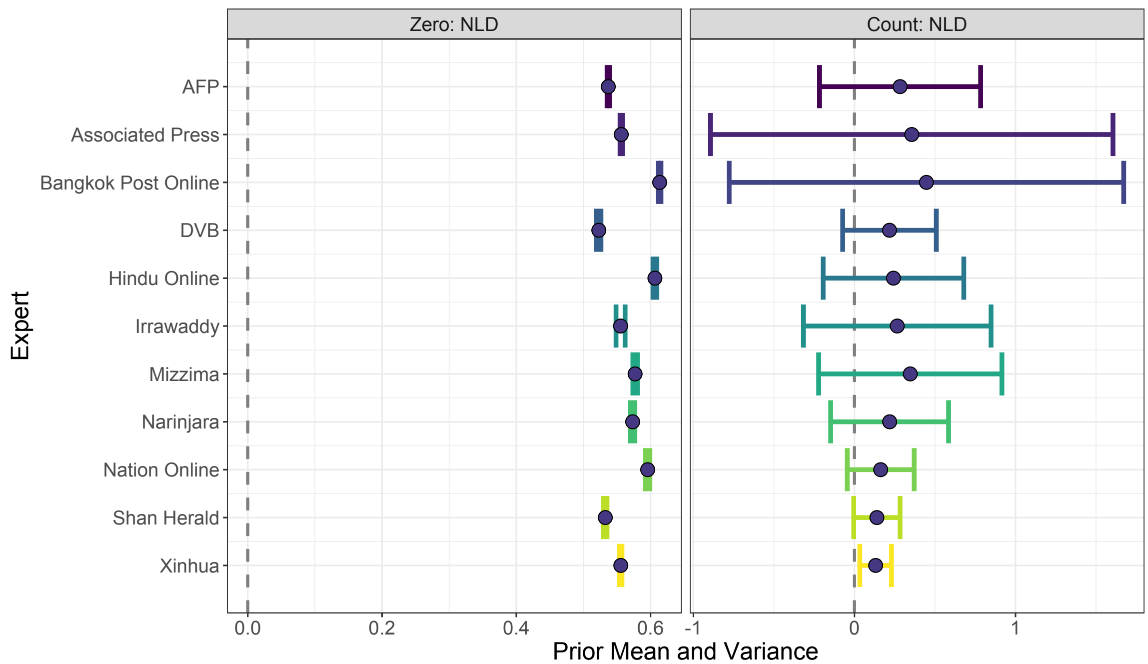


Figure 4.8: Example Prior Means and Variances: party = "NLD"

sent their own school of thought, and only to aggregate when their perspectives are sufficiently similar. These elicited means and variances serve as the basis for the coefficients β for both the zero (logit) and count sides of the ZINB model. The Dirichlet Process, as previously discussed, will facilitate the inclusion of the divergent perspectives reflected in each of these news sources while allowing the priors used in the ultimate analysis to be more precise and overcome technical challenges related to quasi-perfect separation.

$$
\begin{aligned}
y_i &= \begin{cases} y_i^* & \text{if } z_i = 1 \\ 0 & \text{if } z_i = 0 \end{cases} \\
y_i^* &= \mathrm{NegBinom}\left(r, \eta_i\right) \\
\Lambda^{-1}\left(\eta_i\right) &= \beta_{count}^{\top} x_i^{(count)} \\
z_i &\sim \mathrm{Bernoulli}\left(p_i\right) \\
\Lambda^{-1}\left(p_i\right) &= \beta_{zero}^{\top} x_i^{(zero)} \\
\begin{pmatrix} \beta_{zero} \\ \beta_{count} \end{pmatrix} &\sim N\left(\mu_0, \Sigma_0\right) \\
\mu_0, \Sigma_0 &\sim \mathcal{NIW}\left(\theta, \lambda, \Psi, \nu\right) \\
\mu_j &\sim N\left(\mu_j^*, T\right) \\
\Sigma_j &\sim \mathcal{IW}\left(\phi \Sigma_j^*, \phi\right) \\
\left(\mu_j^*, \Sigma_j^*\right) &\sim \mathcal{DP}\left(\mathcal{NIW}\left(\theta, \lambda, \Psi, \nu\right), \alpha\right)
\end{aligned}
$$

Figure 4.9 and Figure 4.10 illustrate the cluster assignment for each of the included newspapers in the analysis. Clustering was performed separately on each side of the model to overcome poor mixing; as is evident in these figures, sources are assigned to quite different clusters depending on whether the zero or count side model parameters are considered. To the extent that this elicitation process accurately reflects the beliefs of these newspaper sources with respect to legislative participation in this period, the Dirichlet clustering is useful in aggregating the zero-side priors in particular because leveraging these effectively will be especially important in overcoming quasi-perfect separation. In both instances of clustering, government-owned *New Light of Myanmar* is relegated to a cluster by itself. Within the zero-side clustering, the particular nature of ethnic media from Myanmar supports *Narinjara* and *Shan Herald* being allocated to clusters unto themselves, while the potential shared beliefs of *Shan Herald* and *Xinhua* may reflect the shared considerations of Shan State (which borders Yunnan Province) and China. Likewise, the common clustering of various opposition and foreign news sources makes sense given their shared English-language audience and the general preference of the international community for the opposition National League for Democracy.

The difference in clustering between the zero and count sides of the model may reflect differing practices with respect to sentiment versus word count. Sentiment is likely more distinct across news sources given word choice and editorial position, whereas word counts for a given topic may not have as much variation conditional on the mentioning of words to constitute a topic. That is, news sources, when choosing to cover, for example, the role of ethnic minority parties, may use relatively similar rates of words pertaining to that topic in order to define their articles within that topic, while the other words used in the articles that more clearly indicate sentiment may vary significantly.

Figure 4.11 below shows the clustered prior means and variances for each covariate level in the analysis, for each side of the model. Note that the log variance is displayed for ease of presentation, and that because no news sources contained words relevant to the base category for ethnicity (Burman), a default prior is used for that covariate level (with a mean of 0 and a variance of 10).

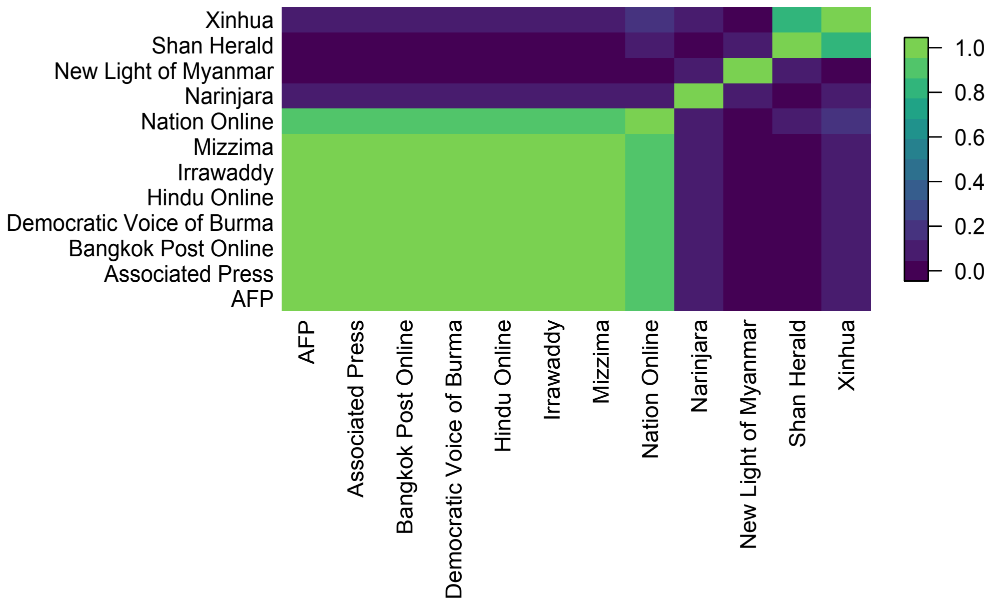


Figure 4.9: Source Clustering: Logit (Zero) Side

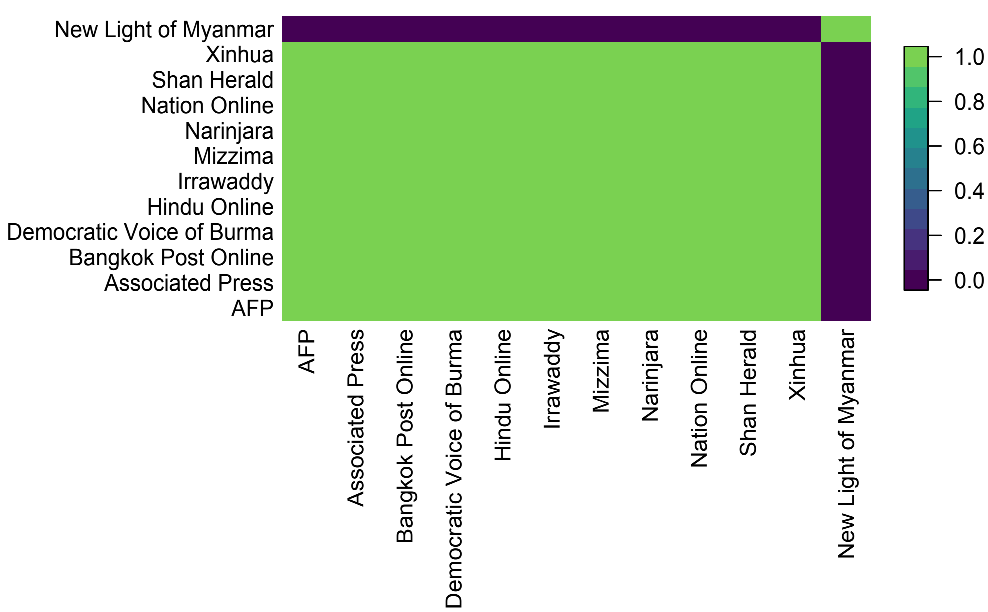


Figure 4.10: Source Clustering: Count Side

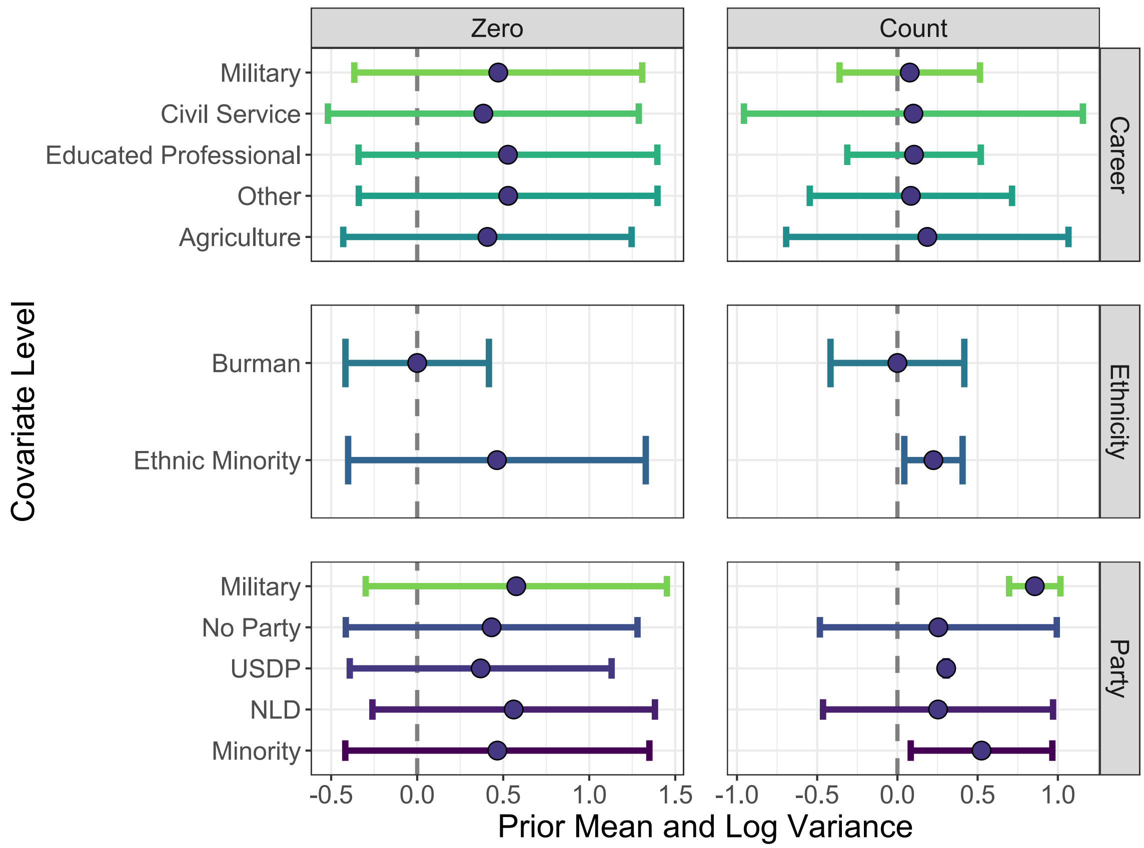


Figure 4.11: Clustered Priors by Covariate Level

4.5 Revised Results: Elicited Priors

Using priors elicited through this text-as-data approach and clustered using a Dirichlet Process, I reanalyze the ZINB model with the Myanmar parliamentary data. For computational reasons, clusters are assigned separately on the logit and count sides of the model (that is, theoretically, sources might align with differing "schools of thought" about the what leads MPs to ask questions versus what leads them to ask a certain number of questions). Notably, the revised priors used reflect equal weighting of each cluster or school of thought. Later analyses could include a hyperprior that weights sources differently depending on the perspective of the researcher. The revised results are presented in the figures below.

The logit side results reflect the most pronounced change as a result of using more specific priors relative to the weakly informative ones used in the initial analysis. Without the interference of quasi-perfect separation, the effect of each covariate is more evident. In particular, and as before, those with a prior career in civil service or in an educated profession are much more likely to ask questions, as are those with a prior career in agriculture. Whereas those with careers in educated professions and agriculture may be reflecting their strong policy preferences, those with past careers in civil service may be reflecting their career ambitions within the new regime via participation. Ethnic minority representatives are also more likely to speak, which aligns with the very recent observations of Myanmar researchers who suggest that because ethnic minority parties are smaller and more cohesive, party discipline and policy positions are clearer and more influential. This is more weakly demonstrated by the posterior distribution for ethnic parties. By contrast, being from the then-ruling USDP or the NLD does not make you more or less likely to ask questions, contrary to expectations that those representatives would be motivated by career ambitions or policy, respectively.

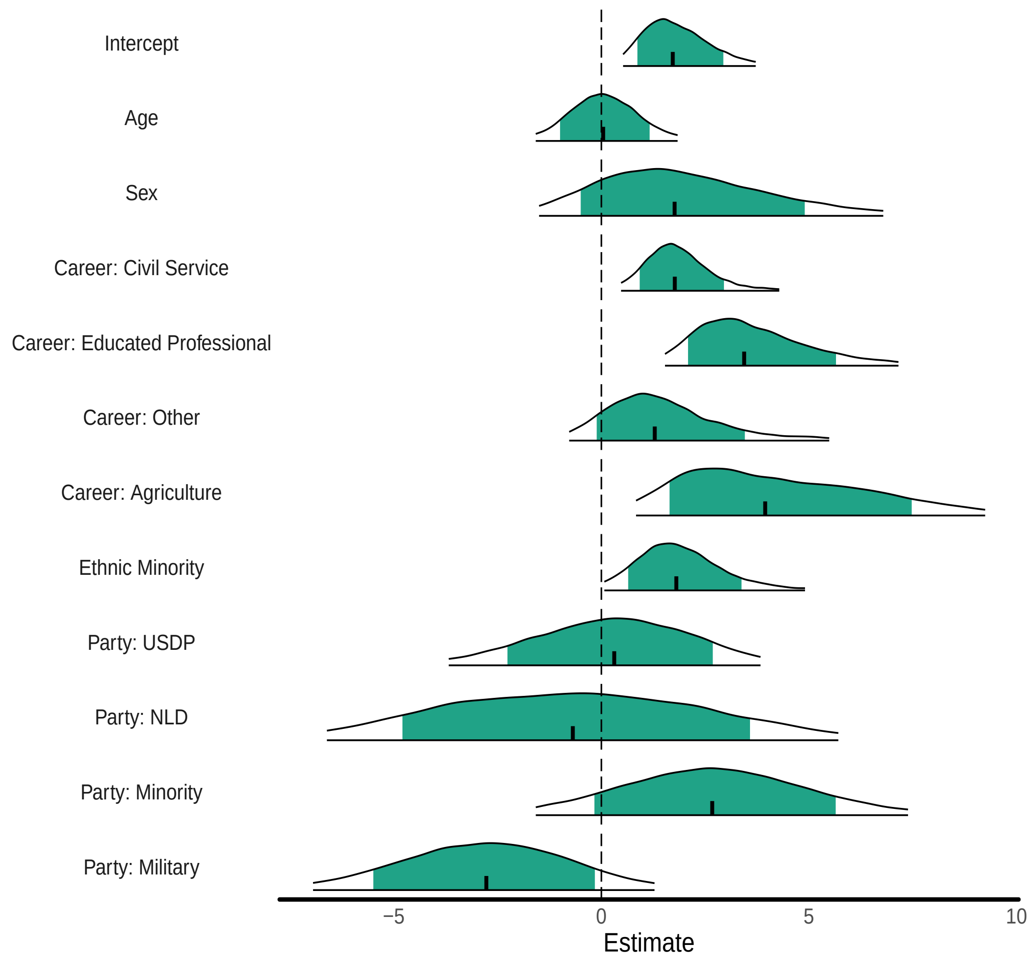


Figure 4.12: Logit Side Posterior Distributions with Revised Priors

The count side of the model demonstrates far fewer changes, which does suggest that these more informative priors do not interfere too extensively with the trends evident in the data itself. This is also expected because variance in word count is much greater than variance in sentiment across articles by source. Removing the interference from quasi-perfect separation, however, does allow us to more confidently conclude that, having spoken, members of the NLD are much less likely to ask a greater number of questions. This has significant implications for whether the main opposition party could hope to meaningfully represent their constituen-

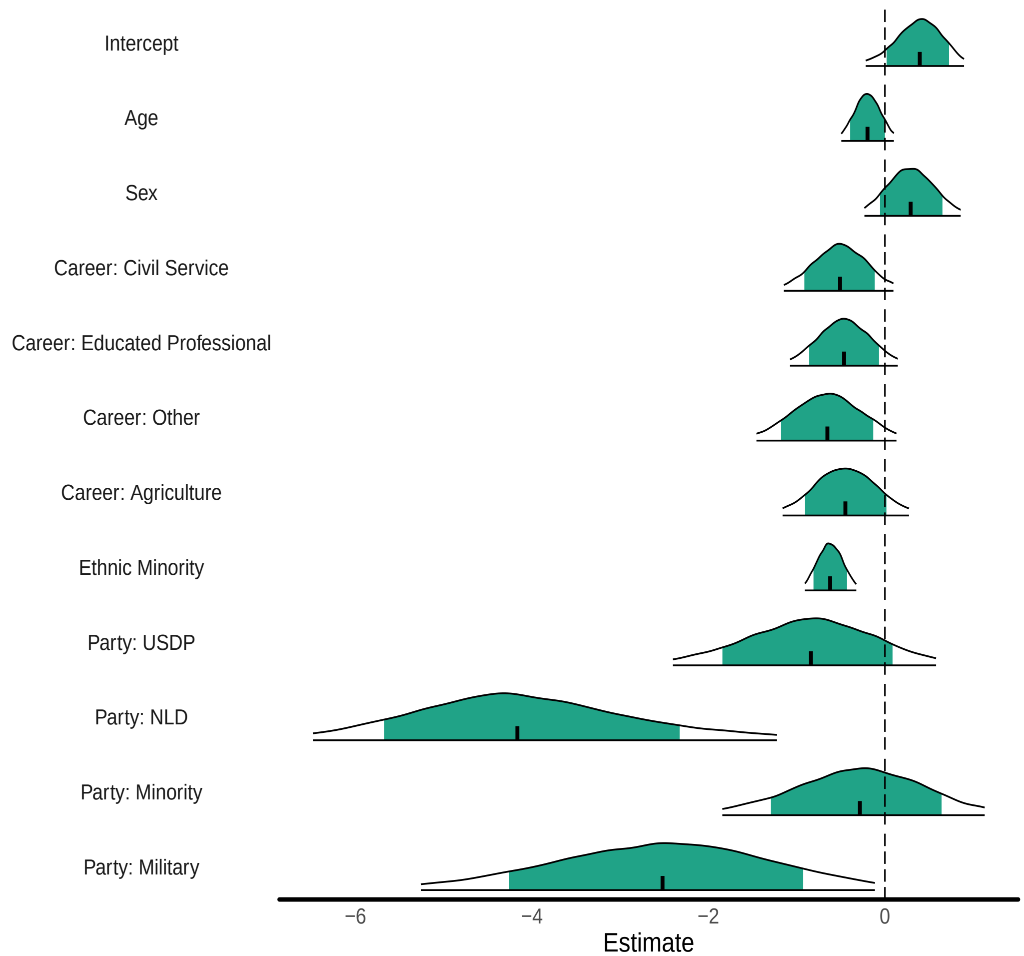


Figure 4.13: Count Side Posterior Distributions with Revised Priors

cies. Likewise, while ethnic minorities are more likely to speak, they are also less likely to ask many questions having asked one at all, which raises concerns about the openness of the forum to participation by minority legislators.

These results provide evidence that MPs in Myanmar's Pyithu Hluttaw do not necessarily participate because of their career ambitions with respect to the ruling party; ethnic minority party members demonstrate a higher probability of participation after all, while both USDP and NLD members largely remain silent. These results, as well as the greater probability of participation from those in educated professions and agriculture instead suggests that policy incentives may undergird participation motives. This is consistent with a lower volume of questioning even among these subsets, because not every minister visiting the parliament has a purview pertaining to policies of interest. That NLD, USDP, and military membership indicate both a lower probability of participation and a lower volume of participation contingent on having participated is not consistent with the notion that regime party-based co-optation dominates parliamentary participation in the case of Myanmar. That is, it does not necessarily appear that USDP MPs are clambering to draw the attention of the USDP leadership by, for example, disproportionately participating. It also does not appear from these results that only those MPs with a vested interest in moving up within the regime (again, perhaps USDP-affiliated MPs or those in the civil service) disproportionately silenced themselves to demonstrate loyalty.

Recent coverage of Myanmar's parliament since the NLD assumed political leadership underscores this point. Rather, while the USDP-led parliament made significant headway on legal reforms and featured relatively robust participation during sessions, the parliament under the NLD has seen significantly less engagement and has been marred by NLD-imposed constraints on member activities (The Economist 2017). As *The Economist* reported, "[Current MPs'] silence is not the result of intimidation by the representatives of the armed forces, who occupy a quarter of the seats in the Hluttaw under the constitution drafted by the military regime. ... Rather, NLD lawmakers are muzzled by their own leaders. No NLD parliamentarian has ever voted against the party line. Members only ask questions that have been vetted by NLD bosses" (The Economist 2017). This democratically elected parliament under the leadership of the democratic "opposition," then, appears in many ways much more authoritarian than its predecessor. Party-based co-optation need not be limited to nondemocratic institutional settings, certainly, but the theoretical expectation that silence in an authoritarian parliament arises primarily from a concern with pleasing the authoritarian elite appears less applicable to an NLD-led set of institutions.

4.6 Validation: Survey-Based Elicitation

To provide a preliminary validation of the text-based elicited priors used in this study, I also conducted an "elite" survey using a roulette elicitation technique as described by the Sheffield Elicitation Framework (SHELF). In this framework, respondents are asked to construct a prior distribution using ten "chips" that each represent a 10% probability value (O'Hagan and Oakley 2016). Respondents in this survey evaluated a series of descriptions of hypothetical members of parliament with that vary MP characteristics (ethnicity, party, occupation), and place chips in "bins" corresponding to the number of times they expect this hypothetical type of MP would have asked a question in the course of the 2-year parliamentary term. An example prompt is presented below. For this survey, a group of 56 experts were hand-chosen based upon their scholarship on Myanmar politics, participation in international conferences (such as the biannual Burma Update hosted at Australian National University), and their publication record. These experts comprise scholars and practitioners from a number of countries, most of whom have attained some level of higher education. This survey was conducted exclusively in English, and all responses are anonymous. Experts 1 and 2, as labeled in the figures below, are both male and white, while expert 3 identifies as female and Asian. All three experts are between the ages of 30–39. Both male experts hold a Ph.D., while the female expert holds a master's degree.

While approximately one-third of potential respondents attempted the sur-

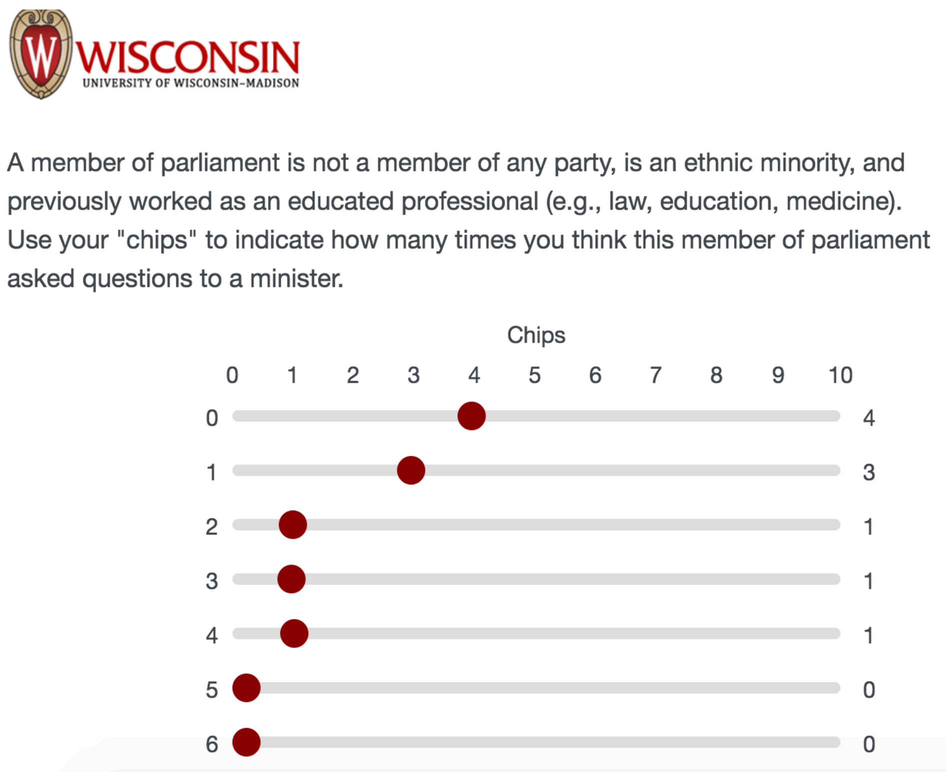


Figure 4.14: Elicitation Survey Preview

vey, only three experts completed the full battery of questions. Figures 4.15, 4.16, and 4.17 below illustrate the chip allocations of each of these experts with respect to each covariate level, forming their prior distributions.

As is evident from each of these figures, expert 2's priors are fairly diffuse, while experts 1 and 3 in general have more concentrated priors. Most critically, these elicited priors illustrate the diversity of opinions that can arise when conducting elicitation in general. While the more narrow priors place significant emphasis on non-speaking (i.e., a large number of chips in bins corresponding to "0" times speaking or a small number of times speaking), the priors for party identification in particular demonstrate some degree of bimodality, as indicated by the ZINB process.

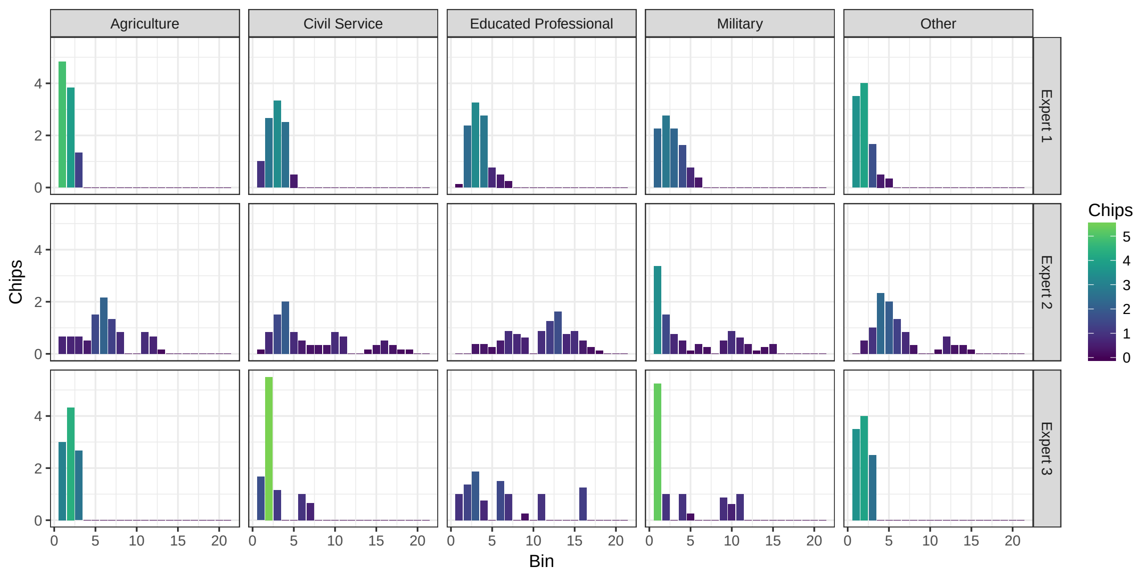


Figure 4.15: Validation: Survey-Based Expert Elicitation for Career Covariates

Figures 4.18, 4.19, and 4.20 compare these survey-based priors from three experts with the priors elicited through the newspaper process. In order to visualize this comparison, I transformed the binned chip counts from the survey elicitation process into means and variances for the zero and count sides of the model. Specifically, the reported means and variances for each covariate level in the zero side of the model reflect the mean proportion of chips in the zero bin versus any of the count bins. Means are taken across all combinations of covariate levels in the survey questions. That is, chip counts for each bin are averaged across all questions that include a particular covariate level. For example, Figure 4.14 asks about a member of parliament who is not a member of any party, but also includes other characteristics (ethnic minority, educated professional). The chip counts in each of the bins for this question would be averaged with the other questions about members of parliament who are not members of any party (i.e., who are Burman, who have careers in the civil service, etc.).

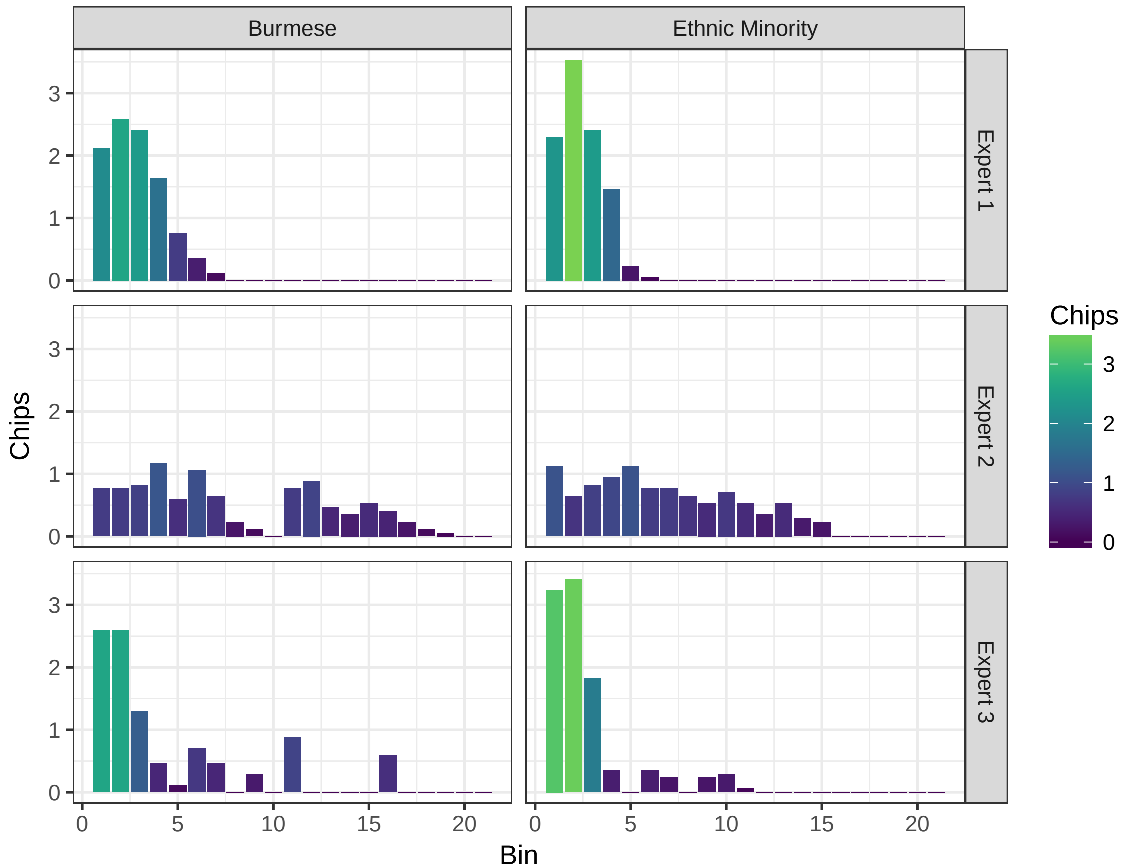


Figure 4.16: Validation: Survey-Based Expert Elicitation for Ethnicity Covariates

The means and variances for the count side of the model reflect this mean of all bin counts with the relevant characteristics as well, and reflect the mean of the untruncated overall distribution (that is, including the zero bin). The inclusion of the zero bin (zero questions asked) in the overall proportions of chips on the count side of the model reflects the untruncated nature of the ZINB model, where "sometimes" zeros exist alongside "always zeros." This process to construct means and variances for the count side of the model, however, provides different results than the means and variances measured with a truncated (only > 0 counts) distribution. In general, the process of collapsing the distribution provided by the surveyed experts as bin counts into a set of means and variances reflective of two separate sides

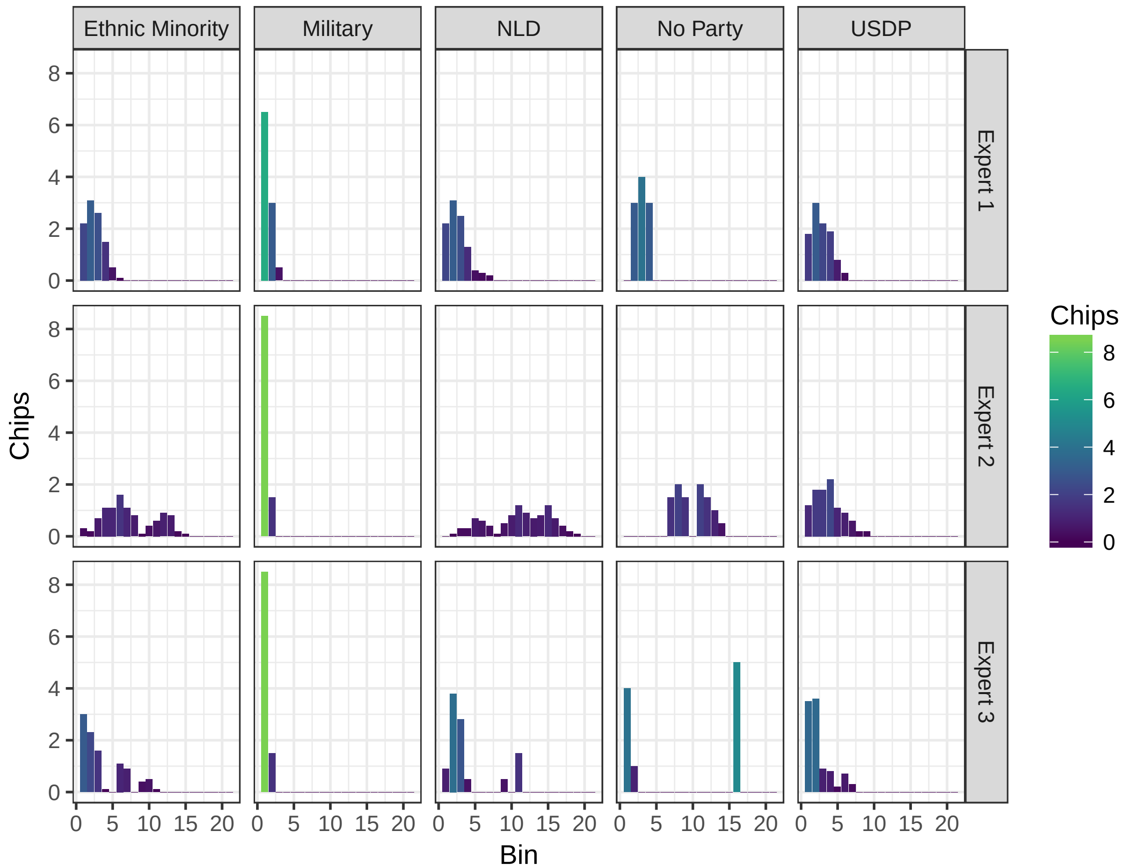


Figure 4.17: Validation: Survey-Based Expert Elicitation for Party Covariates

of the model should be evaluated loosely.

The comparison of the three sets of survey-based elicitation means and variances with the clustered means and variances elicited from newspapers shows that while the priors elicited from newspapers do differ from these survey-based priors, in general they tend not to be more extreme than any given individual surveyed expert's opinion. A few aberrations bear mentioning, however. First, note that although "other" careers were described within the question text as "less well-represented careers" including media and the arts, expert 2 reported in survey feedback that they interpreted the career value "Other" to mean "former political prisoner." This presumably differs from how the other experts interpreted this prompt, and certainly differs from the newspaper-based elicitation process. Second, note again that the means and variances for the "Burman" ethnicity as elicited through the newspaper process reflect a default prior (mean 0, variance 10) because no elicited information was available. Finally, note that where the newspaper-based elicitation means and variances differ significantly from those of the three surveyed experts, the reported values tend to be lower, particularly for the count side of the model. This may indicate that refinements to the word count basis for elicitation in that side of the model could improve estimates. That said, evaluating the newspaper elicitation on the basis of these three surveyed experts privileges the expert's responses, which as noted can vary significantly. This raises additional questions concerning prior validation to be considered elsewhere, but a different approach to these differing priors would be to incorporate the three survey-based elicited priors into the clustering process with the newspaper-based experts and observe both whether the three single experts share "schools of thought" with those reported in the newspaper sample, as well as whether their inclusion meaningfully alters the posterior estimates.

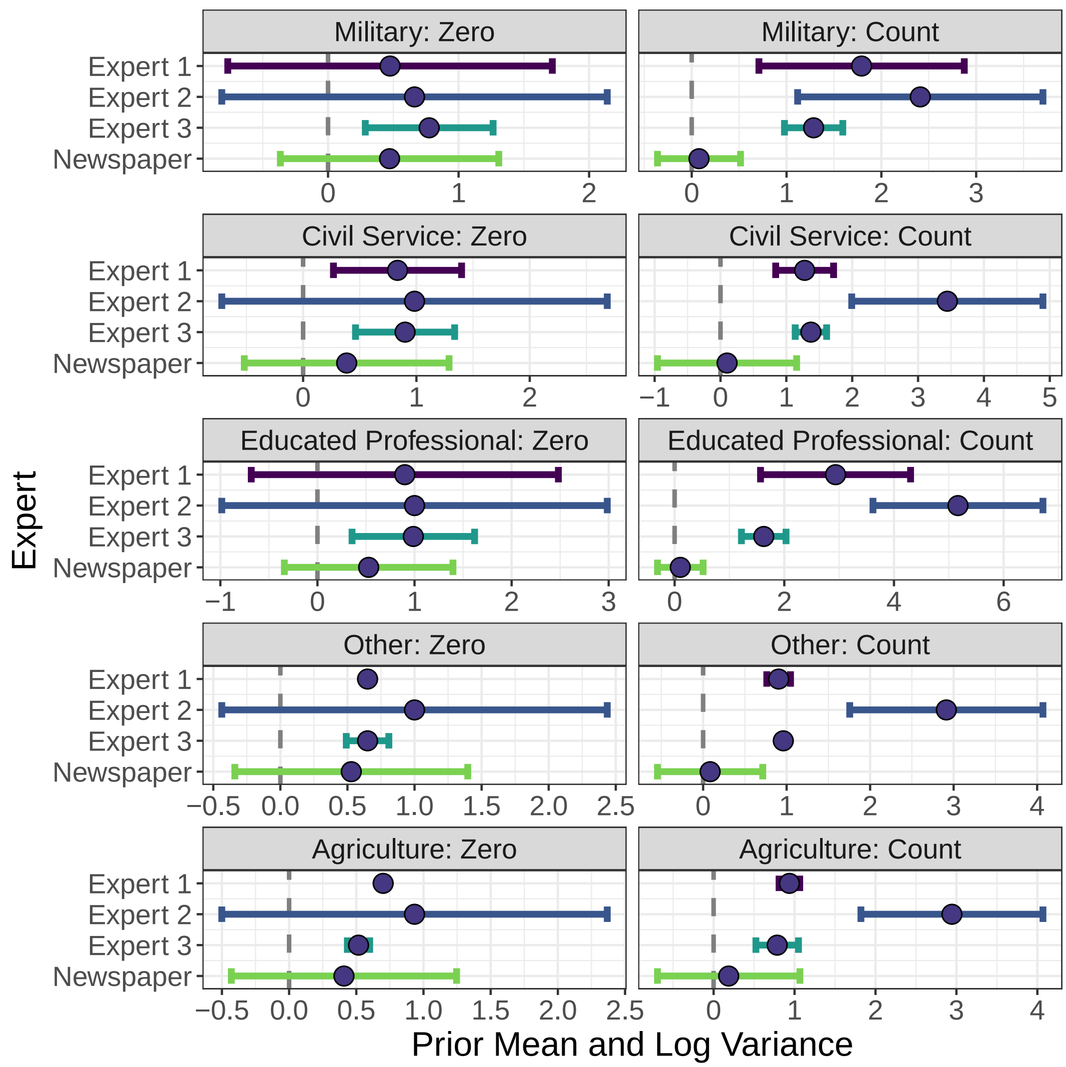


Figure 4.18: Survey vs. Newspaper Elicitation for Career Covariates

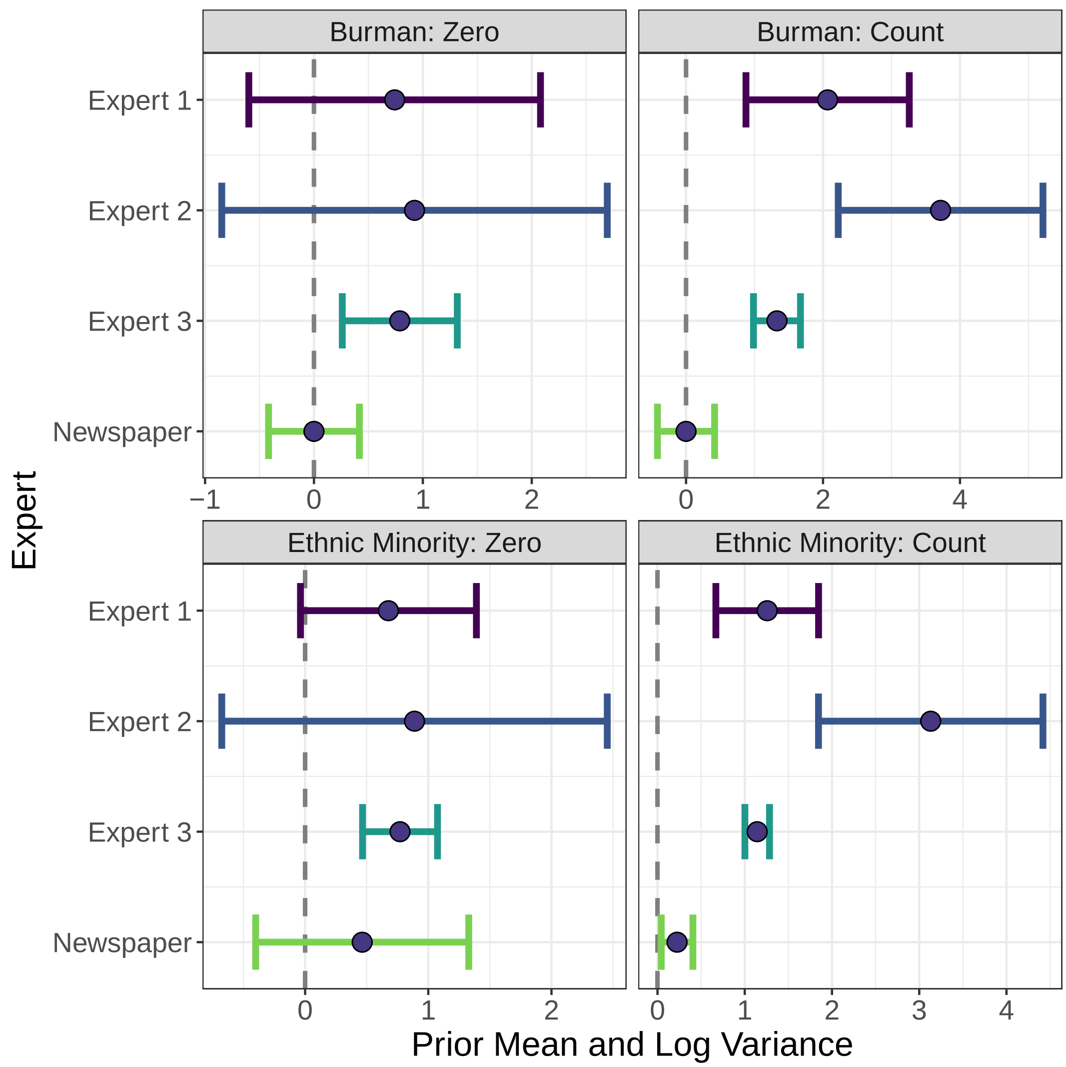


Figure 4.19: Survey vs. Newspaper Elicitation for Ethnicity Covariates

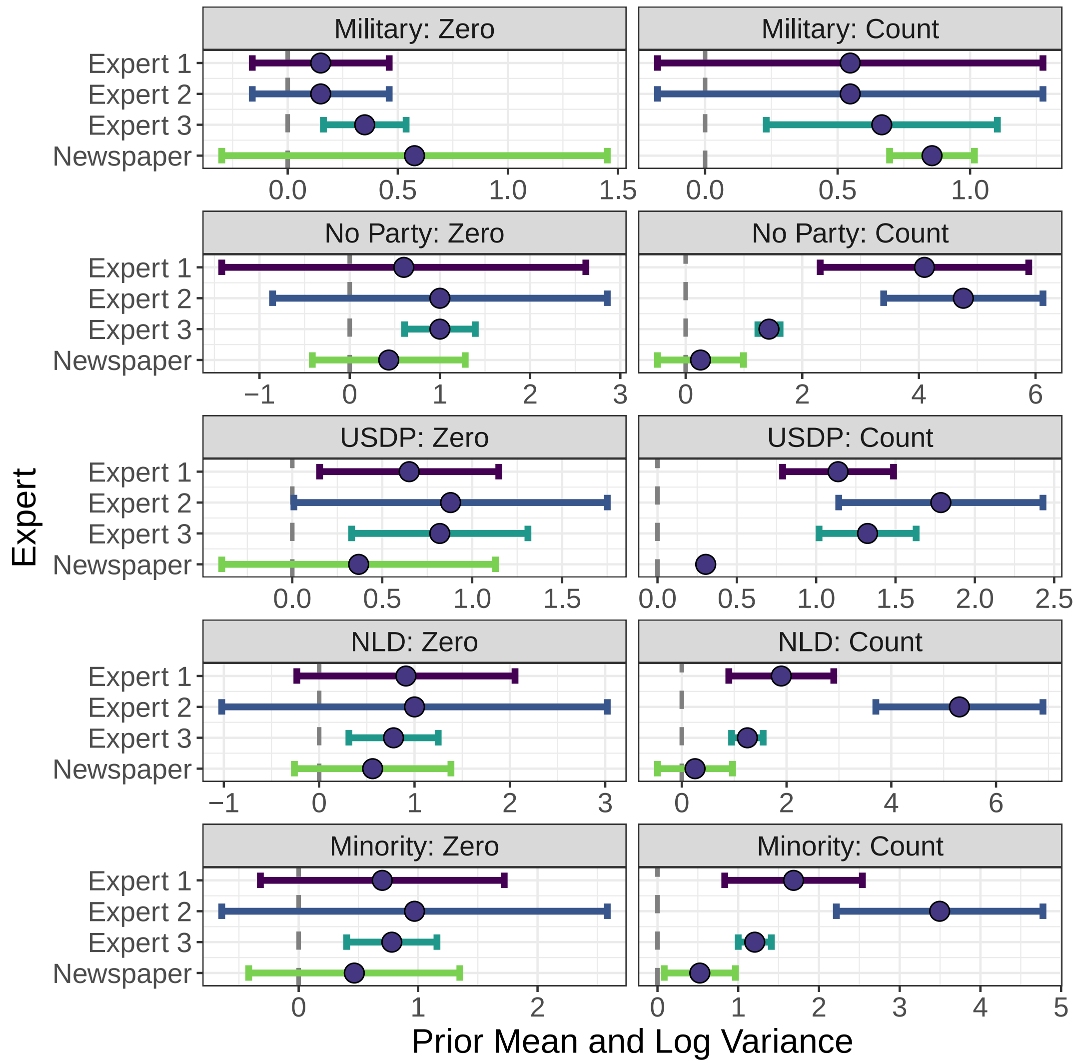


Figure 4.20: Survey vs. Newspaper Elicitation for Party Covariates

4.7 Extensions & Refinements

This empirical analysis aims to serve as the basis for more complete models and understandings of legislative behavior under autocracy. As previously indicated, a complete picture of legislative participation and its implications requires further analysis of participation by issue area, an evaluation of question content as well as count, the inclusion of or accounting for behavior of multiple legislators simultaneously, the inclusion of additional controls for legislator characteristics, and the direct measurement of co-optation and policy preference strength, potentially through an instrumental variables analysis. Below I discuss future directions and extensions for this project to take, in order to provide a full account of legislative participation in authoritarian settings.

4.7.1 Theoretical Extensions

A first clear avenue for extension is in the conception of timing and independence of behavior. Currently because observations are treated independently, all legislators are implicitly thought to adopt a position simultaneously. This is useful for understanding some sorts of legislative behavior more than others. For example, if voting is secret or simultaneous, this timing would give a sense of how individuals with a given distribution of preferences would act independently. A more accurate to life extension of this model would attempt to treat the participation of legislators as sequential. That is, it would evaluate how and at what point certain kinds of legislators are likely to enter given that someone else has entered before them at some particular policy point. This might also allow for coordination between legislators that is currently excluded by empirical accounts of participation.

Alternatively, this participation could be interpreted and analyzed in a networked pattern rather than interpreted sequentially. With network considerations in place, the analysis could focus more directly on which individuals signaled the preferences of parties or the regime for either participation or silence. Likewise, an MP with expertise in the field of education might defer to a colleague with more years of experience, and choose not to speak precisely because their colleague spoke on the issue of most importance to them. These issues are currently not accounted for in the model, but a network analysis could at least offer insights about whether nodes cluster according only to party, or also correspond to career history or other characteristics.

4.7.2 Empirical Extension: Survey Validation

This paper has so far presented one new approach—text-based analysis—to generating more specific and informative priors. Eliciting priors from experts or those knowledgeable about Myanmar politics provides a principled basis for choosing the informative prior to use, and text analysis broadens the set of "experts" who could be consulted. Still, there are several avenues for improving this elicitation procedure. To supplement the text-based elicitation procedure and also to validate its contributions in the previous analysis, I am conducting surveys to directly elicit priors from Myanmar "experts." The "elite" survey that served as the basis for the validation above represents one possible avenue for further extension and validation of this text-based work: additional experts could be identified and provide answers in a roulette-based framework. The low response rate in the validation survey, however, indicates that future surveys require more specific targeting and a design that lends itself to more expedient completion.

4.7.3 Extensions to Text-Based Elicitation

In addition to pursuing survey-based validation of the text-based elicitation approach undertaken here, there are multiple ways in which the current text-based elicitation technique could be refined. First, the corpus could be expanded and refined, either by including more news sources or by consulting alternative textual sources such as published ethnographies. Second, the variance measure could be further refined, either by developing a dictionary of probability or uncertainty terms (e.g., "likely," "uncertain," "probable," "definitely," etc.) in order to evaluate certainty directly and calibrate the variance measure, or by mapping uncertainty terms onto coverage of regular, probabilistic events in these news sources, such as weather (where, e.g., a 30% chance of rain described as "likely" provides a clear mapping from words to an uncertainty measure). Finally, alternatives or supplements to sentiment analysis could improve the accuracy of the prior mean measure. For example, word counts could be used on the sentiment as well as count sides of the model, or sentiment analysis could be changed to incorporate a different tokenization that would evaluate phrases or sentences for sentiment rather than words. Likewise, a topic analysis could provide insights into the relationships among words that contribute to particular sentiment levels, and alternative sentiment dictionaries besides Bing Liu's could be used to test the sensitivity of the current logit-side measures.

Additional data and disaggregation of categories could also add to this analysis. Further extensions of this project will seek to incorporate each house of parliament (Amyotha Hluttaw, Pyithu Hluttaw, Pyidaungsu Hluttaw, and regional and state parliaments) as well as a greater amount of biographical data about the MPs (including number of children and their careers, parents' education and career experience, educational background, etc.). Likewise, the data from the Open

Myanmar Institute (OMI) includes both questions and proposals for each of the 10 sessions of parliament held since 2011. Incorporating these additional sessions might allow for the factor categories to be separated further without cost to separation, and could encompass textual analysis of the content of questions in addition to their counts. In addition, these data might include participation by military MPs that would facilitate the investigation of yet another source of variation. Specifically, this model in conjunction with these data could be used to assess the balance of power in civil-military relations and treat the military both as a self-interested actor or set of actors and as a politically pivotal player, rather than simply a tool at the autocratic regime's disposal (Acemoglu, Ticchi, and Vindigni 2010, 4). In particular, this balance could be more rigorously assessed, comparing legislative participation in circumstances with the military as a veto player, and further dovetailing to examine conditions under which the military assumes that role (e.g., if the civilian government strongly relies on the military for repression as in Svolik (2012, 124–125)).

Finally, future extensions will also seek to disaggregate participation by policy area in order to allow for greater variation in policy preferences by individual and even within parties. These additions should offer better insights into the bargaining processes and co-optation at work in authoritarian legislatures, while also providing some better evidence of how parties are differentiated, and how the role of the military in the legislature differs from elected members of parties.

4.8 Conclusion

This project provides both substantive and methodological contributions. Investigating legislative behavior under autocracy in the case of Myanmar, I have argued for a more nuanced perspective on the condition of co-optation, disaggregating between co-opted members of parliament who might participate subject to career or policy considerations. This substantive intervention refines the observations of previous work to suggest that non-participation is a multi-faceted behavioral choice, distinct from the magnitude of participation. Although further extensions of this work are necessary to fully specify the sources of co-optation (whether via parties or networks) and its multiple manifestations, the evidence presented here from Myanmar indicates that defining non- or low participation in an authoritarian institution as equivalent to acquiescence is at best a limited assessment.

Methodologically, this project puts forth a novel way of incorporated elicited prior information from textual sources, which addresses both a practical problem of quasi-perfect separation in sparse data contexts, as well as a theoretical concern about how best to incorporate additional sources of knowledge and information into the study of authoritarian regimes. The use of newspaper sources as "experts" for elicitation in this project, furthermore, is suggestive of the possibilities for opening the sphere of expertise beyond academics—a process that is facilitated in particular by the Dirichlet clustering application that allows for a diverse array of perspectives to be aggregated for statistical analysis.

4.9 Appendix

4.9.1 Natural Language Processing in Burmese

Text analysis in this paper is conducted in English in part because of the significant challenges that arise in natural language processing for the Burmese language. Burmese is more synthetic than a typical analytical language like English, which means that significant and different meanings can reside in smaller phrases or morphemes (word particles) that might otherwise need to be communicated with full sentences. This raises concerns in attempting, for example, to assign sentiment to particular word particles. Likewise, written text can be either literary or colloquial, which increases the challenges to identifying words with a limited dictionary. Burmese also lacks spaces between words, which creates computational challenges for natural language processing, which typically leverages space delimiters to ease word identification. Finally, Burmese also has no agreed-upon input method (some sources using unicode text and others not), which makes text analysis font-specific. For example, depending on the input method, characters could be read either as ိ ု ခ င ် or concatenated as ခိုင်. While international consortia for natural language processing for less privileged languages are attempting to resolve some of these challenges for languages like Burmese, these efforts are still very preliminary.

Chapter 5

Conclusion

5.1 Arguments & Contributions

This dissertation explores the benefits of elicited-priors approaches to quantitative analysis, and refinements to methods to implement elicited priors, through three related essays. Chapter 2 examines literature, primarily from scientific fields, implementing elicited-priors approaches and details the ways in which the common routine for aggregating elicited priors in these works—averaging—does not lend itself well to Bayesian applications in the social sciences. This chapter proposes a new method of aggregation for elicited priors based upon a Dirichlet Process. Dirichlet clustering allows for the identification of latent "schools of thought" among elicited priors from several experts who have differing points of view, and aggregates these schools of thought accordingly. As the chapter shows, this approach better incorporates divergent viewpoints without allowing extreme positions to overly skew prior distributions. Chapter 2 focuses primarily on a technology to utilize and aggregate priors once they have been elicited, whether this takes place via reading previous literature, or via interviews or focus groups with experts, as is traditional.

Chapter 3 builds on this expansion of the sources from which to elicit, questioning the underlying concept of "expertise" that supposedly defines ideal sources for priors. After examining the psychology and management science literature related to expertise, and the companion literature on election forecasting from political science, Chapter 3 utilizes the 2016 U.S. national election as a form of validation for the Dirichlet clustering approach proposed in Chapter 2. This project interrogates the distinction of "expert" in order to better understand the conditions under which elicitation would produce more accurate results with a large sample of respondents relative to a smaller sample with domain expertise or experience. The results offer hope that elicitation need not be uniformly costly because even elicitation without additional information via covariates is sufficiently accurate to, for example, provide adequate bounds for overcoming practical challenges like quasi-perfect separation. Likewise, the results demonstrate that despite divergence in estimates across subgroups of experts in each of the mass and elite samples, the Dirichlet clustering process proposed in Chapter 2 provides an accurate overall assessment of potential vote share.

Chapter 4 extends the clustering approach of Chapter 2, and the emphasis on expanding the notion of expertise from Chapter 3, with an application to sparse data from an authoritarian context: legislative participation in Myanmar. This chapter not only uses the Dirichlet clustering aggregation of elicited priors introduced in Chapter 2, it also demonstrates their application in a more complex modeling context, using a zero-inflated negative binomial. In addition, although priors are elicited from human experts via an online survey for the purposes of validation, Chapter 4 introduces a text-based method of elicitation focusing on newspaper sources covering Myanmar's politics. Beyond providing substantive insights into the nature of participation in Myanmar's lower house of parliament, this paper expands the notion of valid sources from which to elicit priors.

Each of these contributions—articulating an aggregation technique for elicited priors, illustrating a text-based approach to elicitation and clustering, and evaluating and validating expertise within elicited priors—requires further refinement and extension in future projects. The following section details several possible next steps in each of these domains, and more broadly related to elicited-priors approaches.

5.2 Extensions & Next Steps

For each of the projects undertaken in this dissertation, several extensions can improve their accuracy and applicability. In this section, I detail a series of possible steps to improve these chapters, and conclude with a discussion of additional directions for research on elicited priors in political science.

5.2.1 Chapter 2 Extensions

While the application of the Dirichlet Process clustering approach to Jackman and Western (1994) in Chapter 2 serves to illustrate the method with a tangible and accessible example in which data constraints are binding, additional vignettes will better illustrate the applicability of this method to a wide range of social science problems. First and foremost, a series of simulations using both OLS and probit will demonstrate the efficacy of the Dirichlet clustering process for both continuous and discrete modeling cases. Another vignette featuring a contemporary question of general interest in political science can further underscore the flexibility of the approach. For example, an application to civil war or nuclear conflict could leverage the fact that data are often sparse or incomplete in these settings, but significant expertise exists from which to derive elicited priors to bound analysis. In general, incorporating additional applications might provide an opportunity to underscore the broader philosophical point that elicited priors can support the progression of science and acquisition of knowledge by showing how previously published results can concretely and transparently influence future analyses.

In addition to illustrating further applications of the approach, a more comprehensive articulation of the scope conditions and limitations of the Dirichlet-based approach would improve its applicability by researchers. Some of the scope conditions will become clear in the course of simulations and further comparing the clustering method with pooling in addition to averaging. While in general questions with adequate data suggest fewer benefits to a labor-intensive elicited-priors approach, better understanding the distinction between when an elicited-priors approach might be beneficial or necessary and when it might have only a marginal impact also requires further discussion of the process of selecting experts. The quality and quantity of experts from whom one can elicit priors impacts the clustering process as well as the ultimate priors. The method articulated in Chapter 2 assumes a good-faith effort by the researcher to cull citations from relevant literature or identify plausible, relevant channels through which to identify and include experts for elicitation. Like any methodology, there is no guarantee that its implementation will always reflect good-faith intentions, but laying out specific principles for practice with this approach in particular would be useful for guiding the expert selection process. In addition to generating a set of "best practices" in that domain, describing the outcome in terms of the clustering process when good-faith practices have *not* been used will allow for more transparent and critical research.

Furthermore, developing a formalization of the clustering process would aid in better articulating the incentives of experts and researchers in the elicitation process, and evaluating the quality of information a researcher is likely to receive through elicitation, and how that might differ across experts and clusters. Formalization would reinforce the process of identifying the aforementioned indications of a flawed expert-selection process as well by offering insights into how this might change the nature of the information elicited.

Finally, as previously mentioned, the analysis in Chapter 2 assumes a preexisting elicitation protocol and focuses instead on how best to aggregate priors once elicited. Additional work should evaluate whether particular elicitation methods improve the performance of aggregated priors in later analysis.

5.2.2 Chapter 3 Extensions

Chapter 3 seeks to more critically evaluate who counts as an "expert" and what types of "expertise" are best incorporated into elicited-priors analysis. While this chapter usefully employs the 2016 U.S. election to validate elicited priors in an effort to interrogate the distinction between elites and masses, the extent to which these findings apply in non-U.S. or non-developed contexts is unclear. In general, further work is needed to investigate whether and to what extent a definition of "expert" developed with respect to one set of data usefully applies when investigating other questions.

In addition to including results about the U.S. House of Representatives election, Chapter 3 can be refined in several ways, such as comparing and pooling both mass and elite samples within the clustering framework; evaluating samples against averaged or pooled estimates; better accounting for contagion across bins in the roulette setup; and dynamically updating errors and error variance. The results of these additional analyses may provide useful insights into screening procedures, for example, that could help to identify better "experts" from whom to elicit priors in advance and thereby decrease the costs of elicitation. Beyond providing an indication about the impact of sample size when using survey-based techniques, for example, learning that in general more educated individuals perform better would allow for the more effective selection of respondents up front. Learning that survey response time or political knowledge strongly correlates with more accurate priors, on the other hand, might indicate additional considerations for survey design even beyond the U.S. context.

Likewise, Chapter 3 highlights a tradeoff or a tension between area-specific domain expertise and more general theoretical information that a given expert or group of experts might bring to bear on a particular question. For example, one issue with the idea of identifying "superforecasters" is that their accuracy related to one question or set of questions may not imply accuracy related to other questions. In the realm of comparative politics or international relations, this problem develops an additional facet. Suppose one wants to evaluate factors influencing economic growth in Southeast Asia. Should the group of experts one recruits be those with knowledge particular to, say, the Singaporean experience, or the economic history of Southeast Asia; with knowledge of economic processes in general but not necessarily area expertise; or some mix of the two? Similarly, while the surveys used in Chapter 3 identify individuals in a nationally representative sample and individuals with academic credentials, one could also imagine targeting those with experiential expertise—pollsters, political operatives, elected officials, etc. Additional work to understand these overlapping, intersecting, and divergent areas of expertise is critical to expert recruitment and elicitation.

5.2.3 Chapter 4 Extensions

Unlike the more illustrative approach in Chapter 2, Chapter 4 seeks to make both substantive and methodological contributions. First, the chapter applies the Dirichlet clustering process in the complex modeling context of a zero-inflated negative binomial. Moreover, this process is not only undertaken to illustrate the capacity of a clustering approach for elicited priors to overcome technical challenges such as quasi-perfect separation, but rather is meant to investigate a question of substantive importance: what generates variation in legislative participation under autocracy?

On this substantive dimension, the paper offers a few preliminary insights that would benefit from extension. Co-optation in the Myanmar context appears more nuanced than simply participation or non-participation, and incorporating additional information—more data about recent parliamentary sessions, bill sponsorship and committee membership, question content in addition to count—would better illustrate the ways in which party institutions, career incentives, or policy preferences govern legislative behavior.

Second, Chapter 4 seeks to provide at least a preliminary approach to principled text-based elicitation. Whereas the prior "elicitation" undertaken in Jackman and Western (1994) involves the specification of prior means and variances on the basis of previous literature, it does not use the textual structure of that literature or its explicit specification of prior parameters in order to conduct elicitation. At the same time, however, the text-based approach discussed in Chapter 3 is specific to the newspaper sources and nature of the legislative participation project. Future extensions should seek to generalize on this approach and interrogate how it might work for other *types* of textual sources, such as ethnographies, memoirs, first-hand accounts, academic publications, etc.

Third, Chapter 4 includes a brief section on validation tailored to the newspaper sources used. This validation reflects the priors of only three "elite" experts who responded to an online survey, but incorporating additional expertise would facilitate comparing clustered human expert priors, for example, with those from the newspaper elicitation process. I have already initiated the process of conducting a "mass" elicitation survey, to be launched in 2017–2018, for the Burmese context accordingly. To increase accessibility for those with less formal education and regardless of language ability (the survey will be available in both English and Burmese), the mass survey adopts a conjoint analysis technique from Hainmueller, Hopkins, and Yamamoto (2014). Respondents will evaluate pairs of hypothetical MPs, where characteristics vary on the same dimensions as in the elite survey discussed in Chapter 3 (ethnicity, party, occupation). In response to these example MP profiles, respondents will answer

1. with what probability (%) they believe MP 1 asked any questions,
2. with what probability (%) they believe MP 2 asked any questions,
3. if both MPs were to ask questions, who would ask more questions, and
4. how much more, in percent terms, would that MP ask.

To construct MP profiles for this survey, I began with prompts based on all possible combinations of all possible levels of each MP characteristic factor variable, as was used in the elite survey in Chapter 4. That is, party (none, USDP, NLD, minority party, military), ethnicity (Burman, minority), and career (military, civil service, educated professional, agriculture, less commonly represented career (e.g., arts, tourism, media)) each vary. After creating an exhaustive set of combinations, I eliminated any combinations that did not occur in the 2011–2013 MP participation data, creating 34 total possible variations. Respondents to the elite survey evaluated each of these 34 possibilities. Respondents to the mass survey will evaluate a random subset.

In order to increase response rate relative to the elite elicitation survey and previous attempts at mass elicitation for the Burmese context, I intend to leverage the high Facebook user presence in Myanmar society, advertising the survey as a promoted link in Facebook feeds of Myanmar profiles through Facebook ads. This broader base for survey response will provide a design somewhat analogous to that pursued in Chapter 3, and will provide a wider array of priors against which to validate the text-based approach in Chapter 4.

5.2.4 Additional Projects

In addition to each of these extensions to the specific interventions made throughout Chapters 2, 3, and 4, there are several areas of inquiry within the domain of elicited priors that build on the work of this dissertation. Beyond simply concern for aggregating elicited priors, these pertain to prior validation and alternative elicitation frameworks.

Alternative methods for validation beyond using mass survey responses as in Chapter 4 will provide a more solid basis for adapting an elicited-priors approach to a variety of social science questions. For example, one might evaluate text-based elicited priors from academic sources or previously published works against survey- or focus-group-based elicitation with the authors directly. Alternatively, subjects in a survey-based design could be randomized into a roulette style elicitation or a group directed to write a narrative on a particular topic, in order to refine the process for eliciting prior moments from text sources.

These alternative validation methods dovetail with a second broad domain for further investigation: elicitation framework design. Even the roulette design for elicitation implemented in this dissertation was designed primarily for use in focus-group settings, rather than in survey settings, let alone online surveys. Evaluating whether answers are more consistent in each of these elicitation contexts, using the same elicitation instrument, is a crucial first step toward understanding how to make elicitation a more useful and accessible method. While survey techniques for elicitation might broaden the set of "experts" from whom one can elicit priors and the set of contexts or questions in which elicitation can occur, additional care is required to design appropriate tools for sensitive contexts. For example, while the analysis in Chapter 4 effectively utilizes text-based elicitation to circumvent some of the challenges with identifying and eliciting priors from experts in an authoritarian context, authoritarian and less developed countries are those that often have sparse data, for which an elicited-priors approach is particularly useful. Further articulating whether having even a select group of experts, non-randomly chosen or non-representative in nature, is preferable to no expertise, or how one might compensate for non-ideal survey design will facilitate the application of elicited-priors approaches in these difficult contexts.

Bibliography

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni. 2010. "A Theory of Military Dictatorships." *American Economic Journal: Macroeconomics* 2 (1): 1–42.

Ahlquist, John S. 2017. "Labor Unions, Political Representation, and Economic Inequality." *Annual Review of Political Science* 20.

Albert, Isabelle, Sophie Donnet, Chantal Guihenneuc-Jouyaux, Samantha Low-Choy, Kerrie Mengersen, and Judith Rousseau. 2012. "Combining Expert Opinions in Prior Elicitation." *Bayesian Analysis* 7 (3): 503–531.

Aldrich, John H. 1995. *Why Parties?: The Origin and Transformation of Party Politics in America.* Chicago: University of Chicago Press.

ANES. 2015. "User's Guide and Codebook for the ANES 2012 Time Series Study." University of Michigan and Stanford University. `http://www.electionstudies.org/studypages/anes_timeseries_2012/anes_timeseries_2012_userguidecodebook.pdf`.

Art, David. 2012. "Review: What Do We Know About Authoritarianism After Ten Years?" *Comparative Politics* 44 (3): 351–373.

Associated Press. 2016. "U.S. Elections: Ad Spending." `http://elections.ap.org/content/ad-spending`.

Atanasov, Pavel, Phillip Rescober, Eric Stone, Samuel A. Swift, Emile Servan-Schreiber, Philip Tetlock, Lyle Ungar, and Barbara Mellers. 2016. "Distilling the Wisdom of Crowds: Prediction MMarket vs. Prediction Polls." *Management Science* 63 (3): 691–706.

Bakker, Ryan. 2009. "Re-measuring left-right: A comparison of SEM and Bayesian measurement models for extracting left-right party placements." *Electoral Studies* 28 (3): 413–421.

Bayarri, M., and James Berger. 2004. "The Interplay Between Bayesian and Frequentist Analysis." *Statistical Science* 19:58–80.

Behrens, Martin, Kerstin Hamann, and Richard W. Hurd. 2004. "Conceptualizing Labor Union Revitalization." In *Varieties of Unionism: Strategies for Union Revitalization in a Globalizing Economy,* edited by C. Frege and J. Kelly, 11–29. New York: Oxford University Press.

Berg, Joyce E., Forrest D. Nelson, and Thomas A. Rietz. 2008. "Prediction Market Accuracy in the Long Run." *International Journal of Forecasting* 24:285–300.

Berger, James. 2004. "The Case for Objective Bayesian Analysis." *Bayesian Analysis* 1 (1): 1–17. `https://people.eecs.berkeley.edu/~jordan/sail/readings/berger-2006.pdf`.

Berger, James O., and J.M. Bernardo. 1992. "On the Development of Reference Priors." *Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting* 4:35–49.

Bernardo, J.M. 1979. "Reference Prior Distributions for Bayesian Inference." *Journal of the Royal Statistical Society Series B–Methodological* 41:113–128.

Besley, Timothy. 2006. *Principled Agents?: The Political Economy of Good Government.* New York: Oxford University Press.

Blei, David, Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3:993–1022.

Boix, Carles, and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions and Power-Sharing in Dictatorships." *Journal of Politics* 75:300–316.

Borrel, Monique J. 2004. "Industrial Conflict, Mass Demonstrations, and Economic and Political Change in Postwar France: An Econometric Model." *Global, Area, and International Archive (GAIA):* 1–85. `https://escholarship.org/uc/item/5nq0p78c`.

Boydstun, Amber, Nadia Brown, Kim Yi Dionne, Samara Klar, Yanna Krupnikov, Melissa Michelson, Kathleen Searles, and Christina Wolbrecht. 2017. "Women Also Know Stuff." `http://womenalsoknowstuff.com/`.

Budescu, David V., and Eva Chen. 2014. "Identifying Expertise to Extract the Wisdom of Crowds." *Management Science* 61 (2): 267–280.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival.* Cambridge, MA: MIT Press.

Bureau of Labor Statistics. 2016. *Regional and State Employment and Unemployment (Monthly) News Release.* United States Department of Labor. `https://www.bls.gov/news.release/archives/laus_11182016.htm`.

Campbell, James E. 2008. "Evaluating U.S. Presidential Election Forecasts and Forecasting Equations." *International Journal of Forecasting* 24:259–271.

———. 2014. "Issues in Presidential Election Forecasting: Election Margins, Incumbency, and Model Credibility." *PS: Perspectives on Politics* April:301–303.

Campbell, James E., and Michael S. Lewis-Beck. 2008. "US Presidential Election Forecasting: An Introduction." *International Journal of Forecasting* 24:189–192.

Carroll, Royce, Jeff Lewis, James Lo, Nolan McCarty, Keith Poole, and Howard Rosenthal. 2015. "DW-NOMINATE Scores With Bootstrapped Standard Errors." `http://voteview.com/dwnomin.htm`.

Daponte, Beth Osborne, Joseph B. Kadane, and Lara J. Wolfson. 1997. "Bayesian Demography: Projecting the Iraqi Kurdish Population, 1977–1990." *Journal of the American Statistical Association* 92 (440): 1256–1267.

Datta, G., and R. Mukerjee. 2004. *Probability Matching Priors: Higher Order Asymptotics.* New York: Springer.

Datta, G.S., and M. Ghosh. 1991. "Asymptotic Optimality of Hierarchical Bayes Estimators and Predictors." *Journal of Statistical Planning and Inference* 29 (3): 229–243.

Dey, Dipak K., and Junfeng Liu. 2007. "A Quantitative Study of Quantile Based Direct Prior Elicitation from Expert Opinion." *Bayesian Analysis* 2 (1): 137–166.

Dey, D.K., and L.R. Birmiwal. 1994. "Robust Bayesian Analysis Using Divergence Measures." *Statistics and Probability Letters* 20 (4): 287–294.

Dhami, Mandeep K., David R. Mandel, Barbara A. Mellers, and Philip E. Tetlock. 2015. "Improving Intelligence Analysis with Decision Science." *Perspectives on Psychological Science* 10 (6): 753–757. `http://journals.sagepub.com/doi/pdf/10.1177/1745691615598511`.

Evans, Jocelyn, and Gilles Ivaldi. 2010. "Comparing Forecast Models of Radical Right Voting in Four European Countries (1973-2008)." *International Journal of Forecasting* 26:82–97.

Farber, Henry S., and Bruce Western. 2001. "Accounting for the Decline of Unions in the Private Sector, 1973-1998 *." *Journal of Labor Research* 22 (3): 459–485. `http://EconPapers.repec.org/RePEc:tra:jlabre:v:22:y:2001:i:3:p:459-485`.

Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship.* New York: Cambridge University Press.

Gandhi, Jennifer, and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40 (11): 1279–1301.

Garthwaite, Paul H., Joseph B. Kadane, and Anthony O'Hagan. 2005. "Statistical Methods for Eliciting Probability Distributions." *Journal of the American Statistical Association* 100 (470): 680–700.

Gehlbach, Scott, Konstantin Sonin, and Milan W. Svolik. 2016. "Formal Models of Nondemocratic Politics." *Annual Review of Political Science* 19:565–584.

Gelman, Andrew. 2002. "Prior Distribution." In *Encyclopedia of Environmetrics,* edited by Abdel H. El-Shaarawi and Walter W. Piegorsch, 3:1634–1637. Chichester: John Wiley & Sons.

———. 2009. "Bayes, Jeffreys, Prior Distributions, and the Philosophy of Statistics." *Statistical Science* 24 (2): 176–178.

Gelman, Andrew, and Christian Hennig. 2017. "Beyond Subjective and Objective Statistics." *Journal of the Royal Statistical Society* (forthcoming).

Gelman, Andrew, Aleks Jakulin, Maria Grazia Pittau, and Yu-Sing Su. 2008. "A Weakly Informative Default Prior Distribution for Logistic and Other Regression Models." *The Annals of Applied Statistics* 2 (4): 1360–1383.

Gelman, Andrew, and Gary King. 1993. "Why are American Presidential Election Campaign Polls so Variable When Votes are so Predictable?" *British Journal of Political Science* 23 (4): 409–451.

Gill, Jeff, and Lee D. Walker. 2005. "Elicited priors for Bayesian model specifications in political science research." *Journal of Politics* 67 (3): 841–872.

Gill, Jeff, and Christopher Witko. 2013. "Bayesian Analytical Methods: A Methodological Prescription for Public Administration." *Journal of Public Administration Research and Theory* 23 (2): 457–494.

Graefe, Andreas. 2015. "German Election Forecasting: Comparing and Combining Methods for 2013." *German Politics* 24 (2): 195–204.

Graefe, Andreas, J. Scott Armstrong, Jr. Randall J. Jones, and Alfred G. Cuzán. 2014. "Combining Forecasts: An Application to Elections." *International Journal of Forecasting* 30:43–54.

Graefe, Andreas, Helmut Küchenhoff, Veronika Stierle, and Bernhard Riedl. 2015. "Limitations of Ensemble Bayesian Model Averaging for Forecasting Social Science Problems." *International Journal of Forecasting* 31:943–951.

Greene, Jay P. 1993. "Forewarned before Forecast: Presidential Election Forecasting Models and the 1992 Election." *PS: Political Science and Politics* 26 (1): 17–21.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22:1–30.

Hall, Peter A., and David Soskice. 2001. *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage.* Oxford University Press.

Hayek, Friedrich A. 1945. "The Use of Knowledge in Society." *American Economic Review* 35 (4): 519–530.

Hillygus, D. Sunshine. 2011. "The Evolution of Election Polling in the United States." *Public Opinion Quarterly* 75 (5): 962–981.

Huberty, Mark. 2015. "Can We Vote with Our Tweet? On the Perennial Difficulty of Election Forecasting with Social Media." *International Journal of Forecasting* 31:992–1007.

Jackman, Simon, and Gary N. Marks. 1994. "Forecasting Australian Elections: 1993, and All That." *Australian Journal of Political Science* 29 (2): 277–291.

Jackman, Simon, and Bruce Western. 1994. "Bayesian Inference for Comparative Research." *American Political Science Review* 88 (2): 412–423.

Jeffreys, H. 1961. *Theory of Probability.* London: Oxford University Press.

Jensen, Nathan M., Edmund Malesky, and Stephen Weymouth. 2013. "Unbundling the Relationship between Authoritarian Legislatures and Political Risk." *British Journal of Political Science* 44:655–684.

Jones, Randall J., Jr. 2008. "The State of Presidential Election Forecasting: The 2004 Experience." *International Journal of Forecasting* 24:310–321.

Kadane, Joseph B., and Lara J. Wolfson. 1998. "Experiences in Elicitation." *Journal of the Royal Statistical Society Series D (The Statistician)* 47 (1): 3–19.

Karvetski, C.W., K.C. Olson, D.R. Mandel, and C.R. Twardy. 2013. "Probabilistic Coherence Weighting for Optimizing Expert Forecasts." *Decision Analysis* 10:305–326.

Kendall, Chad, Tommaso Nannicini, and Francesco Trebbi. 2015. "How Do Voters Respond to Information? Evidence from a Randomized Campaign." *American Economic Review* 105 (1): 322–353.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions, 3rd Edition.* Chicago: University of Chicago Press.

Lee, Cheol-Sung. 2005. "International Migration, Deindustrialization and Union Decline in 16 Affluent OECD Countries, 1962-1997." *Social Forces* 84 (1): 71–88.

Lele, S.R., and A. Das. 2000. "Elicited data and incorporation of expert opinion for statistical inference in spatial studies." *Mathematical Geology* 32 (4): 465–487.

Lewis-Beck, Michael S. 2010. "Editorial: European Election Forecasting: An Introduction." *International Journal of Forecasting* 26:9–10.

Lewis-Beck, Michael S., and Mary Stegmaier. 2014. "US Presidential Election Forecasting." *PS: Perspectives on Politics* April:284–288.

Lewis-Beck, Michael S., and Charles Tien. 1996. "The Future in Forecasting: Prospective Presidential Models." *American Politics Quarterly* 24 (4): 468–491.

———. 1999. "Voters as Forecasters: A Micromodel of Election Prediction." *International Journal of Forecasting* 15:175–184.

Lin, C.H., Y.C. Lin, and P. Raghubir. 2004. "The interaction between order of elicitation and event controllability on the self-positivity bias." In *Advances in Consumer Research, Volume XXXI,* edited by B.E. Kahn and M.F. Luce, 31:523–529.

Linderman, Scott, Matthew Johnson, and Ryan P. Adams. 2015. "Dependent Multinomial Models Made Easy: Stick-Breaking with the Polya-gamma Augmentation." In *Advances in Neural Information Processing Systems 28 (NIPS 2015).* https://papers.nips.cc/paper/5660-dependent-multinomial-models-made-easy-stick-breaking-with-the-polya-gamma-augmentation.

Linzer, Drew. 2013. "Dynamic Bayesian Forecasting of Presidential Elections in the States." *Journal of the American Statistical Association* 108 (501): 124–134.

———. 2014. "The Future of Election Forecasting: More Data, Better Technology." *PS: Perspectives on Politics* April:326–328.

Lipset, Seymour Martin, and Ivan Katchanovski. n.d. "Corporatism, Left Government and Union Density in Advanced Western Countries." :%20https://www.researchgate.net/publication/259982791.

Liu, Bing. 2004. "Sentiment Lexicon." https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon.

Lock, Kari, and Andrew Gelman. 2010. "Bayesian Combination of State Polls and Election Forecasts." *Political Analysis* 18:337–348.

Magaloni, Beatriz. 2008. "Credible Power-Sharing and the Longevity of Authoritarian Rule." *Comparative Political Studies* 41 (4/5): 715–741.

Magaloni, Beatriz, and Ruth Kricheli. 2010. "Political Order and One-Party Rule." *Annual Review of Political Science* 13 (1): 123–143.

Malesky, Edmund, and Paul Schuler. 2010. "Nodding or Needling: Analyzing Delegate Responsiveness in an Authoritarian Parliament." *American Political Science Review* 104 (3): 482–502.

Manion, Melanie. 2013. *Authoritarian Parochialism: Local Congressional Representation in China.* Technical report 128. CDDRL Working Paper.

Mayo, Deborah G., and Aris Spanos. 2010. "Introduction and Background." In *Error and Inference: Recent Exchanges on Experimental Reasoning, Reliability, and the Objectivity and Rationality of Science,* edited by Deborah G. Mayo and Aris Spanos, 1–27. New York: Cambridge University Press.

Mellers, Barbara, Eric Stone, Terry Murray, Angela Minster, Nick Rohrbaugh, Michael Bishop, Eva Chen, et al. 2015. "Identifying and Cultivating Superforecasters as a Method of Improving Probabilistic Predictions." *Perspectives on Psychological Science* 10 (3): 267–281. http://journals.sagepub.com/doi/abs/10.1177/1745691615577794.

Morris, Rachel K., Gemma L. Malin, Elisabeth Quinlan-Jones, Lee J. Middleton, Karla Hemming, Danielle Burke, Jane P. Daniels, Khalid S. Khan, Jon Deeks, and Mark D. Kilby. 2013. "Percutaneous vesicoamniotic shunting versus conservative management for fetal lower urinary tract obstruction PLUTO: a randomised trial." *Lancet* 382 (9903): 1496–1506.

Morse, Yonatan L. 2012. "The Era of Electoral Authoritarianism." *World Politics* 64 (1): 161–198.

Moses, Jonathan, and Torbjørn Knutsen. 2012. *Ways of Knowing: Competing Methodologies in Social and Political Research.* New York: Palgrave Macmillan.

Murr, Andreas E. 2015. "The Wisdom of Crowds: Applying Condorcet's Jury Theorem to Forecasting US Presidential Elections." *International Journal of Forecasting* 31:916–929.

Murr, Andreas Erwin. 2011. ""Wisdom of Crowds"? A Decentralized Election Forecasting Model that Uses Citizens' Local Expectations." *Electoral Studies* 30:771–783.

Neal, Radford M. 2000. "Markov Chain Sampling MMethod for Dirichlet Process Mixture Models." *Journal of Computational and Graphical Statistics* 9 (2): 249–265.

New York Times. 2017. "Presidential Election Results: Donald J. Trump Wins." `https://www.nytimes.com/elections/results/president`.

Nichols, Tom. 2017. *The Death of Expertise: The Campaign Against Established Knowledge and Why it Matters.* Oxford: Oxford University Press.

———. *The Death of Expertise,* accessed. `http://thefederalist.com/2014/01/17/the-death-of-expertise/`.

O'Hagan, Anthony, C.E. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley, and T. Rakow. 2006. *Uncertain Judgments: Eliciting Experts' Probabilities.* Chichester: John Wiley & Sons, Ltd.

O'Hagan, Tony, and Jeremy Oakley. 2016. "Sheffield Elicitation Framework (SHELF) v. 3.0." `http://www.tonyohagan.co.uk/shelf/SHELF3.html`.

O'Leary, Rebecca A., Samantha Low Choy, Justine V. Murray, Mary Kynn, Robert Denham, Tara G. Martin, and Kerrie Mengersen. 2009. "Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby Petrogale penicillata." *Environmentrics* 20 (4): 379–398.

Polson, Nicholas, James G. Scott, and Jesse Windle. n.d. "Bayesian Inference for Logistic Models Using Pólya-Gamma Latent Variables." `https://arxiv.org/pdf/1205.0310.pdf`.

Reuter, Ora John, and Graeme B. Robertson. 2014. "Legislatures, Cooptation, and Social Protest in Contemporary Authoritarian Regimes." *Journal of Politics* 77 (1): 235–248.

Rigdon, Steven E., Sheldon H. Jacobson, Wendy K. Tam Cho, Edward Sewell, and Christopher J. Rigdom. 2009. "A Bayesian Prediction Model for the U.S. Presidential Election." *American Politics Research* 37 (4): 700–724.

Satopaa, V.A., S.T. Jensen, B.A. Mellers, P.E. Tetlock, and L. Ungar. 2014. "Aggregation in Time-Series: Dynamic Hierarchical Modeling of Sparse Expert Beliefs." *The Annals of Applied Statistics* 8:1256–1280.

Savage, Leonard J. 1971. "Elicitation of Personal Probabilities and Expectations." *Journal of the American Statistical Association* 66 (336): 783–801.

Shuford, Emir, and Thomas A. Brown. 1975. "Elicitation of Personal Probabilities and Their Assessment." *Instructional Science* 4:137–188.

Sivaganesan, S. 1993. "Range of the Posterior Probability of an Interval for Priors with Unimodality Preserving Contaminations." *Annals of the Institute of Statistical Mathematics* 45 (1): 187–199.

Sloman, Steven, and Philip Fernbach. 2017. *The Knowledge Illusion: Why We Never Think Alone.* New York: Riverhead Books.

Small, Mitchell J. 2008. "Methods for assessing uncertainty in fundamental assumptions and associated models for cancer risk assessment." *Risk Analysis* 28 (5): 1289–1307.

Spetzler, Carl S., and Carl-Axel S. Staël von Holstein. 1975. "Probability Encoding in Decision Analysis." *Management Science* 22:340–358.

Svolik, Milan W. 2012. *The Politics of Authoritarian Rule.* New York: Cambridge University Press.

Tetlock, Philip. 2017. "Good Judgment Project." `https://www.goodjudgment.com`.

Tetlock, Philip, and Barbara Mellers. 2014. "Judging Political Judgement," *Proceedings of the National Academy of Sciences of the United States of America* 111 (32): 11574–11575.

The Economist. 2017. "Democracy has muzzled Myanmar's parliament." June. Accessed. `http://www.economist.com/news/asia/21722862-aung-san-suu-kyi-does-not-seem-mind-democracy-has-muzzled-myanmars-parliament?fsrc=scn%2Ffb%2Fte%2Fbl%2Fed%2Fanotalkingshopdemocracyhasmuzzledmyanmarsparliament`.

Visser, Penny S., Jon A. Krosnick, Jesse Marquette, and Michael Curtin. 1996. "Mail SSurvey for Election Forecasting: An Evaluation of the *Columbus Dispatch* Poll." *Public Opinion Quarterly* 60:181–227.

Wallace, C.S., and D.L. Dowe. 1999. "Minimum Message Length and Kolmogorov Complexity." *Computer Journal* 42 (4): 270–283.

Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman. 2015. "Forecasting Elections with Non-Representative Polls." *International Journal of Forecasting* 31:980–991.

Wang, Zhongming, and Hui Zhou. 2009. "A General Method of Prior Elicitation in Bayesian Reliability Analysis." In *Proceedings of the 2009 8th International Conference on Reliability, Maintainability and Safety, Vols. I and II,* edited by R. Kang, H.Z. Huang, M.J. Zuo, and Q. Miao, 415–419.

Wasserman, David. *The Cook Political Report,* January. Accessed. `http://cookpolitical.com/story/10174`.

Western, Bruce. 1994. "Institutional Mechanisms for UnionizUnion in Sixteen OECD Countries: An Analysis of Social Survey Data." *Social Forces* 73 (2): 497–519.

———. 2002. "Institutions and the Labor Market." In *The New Institutionalism in Sociology*, edited by Mary C. Brinton and Victor Nee, 224–243. Stanford: Stanford University Press.

Wheeler, David C., Antonio Paez, Jamie Spinney, and Lance A. Waller. 2014. "A Bayesian approach to hedonic price analysis." *Papers in Regional Science* 93 (3): 663–683.

Wright, Joseph, and Abel Escribà-Folch. 2011. "Authoritarian Institutions and Regime Survival: Transitions to Democracy and Subsequent Autocracy." *British Journal of Political Science* 42:283–309.

Zaffalon, M. 2005. "Credible classification for environmental problems." *Environmental Modelling and Software* 20 (8): 1003–1012.

Zorn, Christopher J. 1996. "Evaluating Zero-Inflated and Hurdle Poisson Specifications." *Midwest Political Science Associations* 18–20 April:1–16.