Appendix

for

"Ambiguous Platforms and Correlated Preferences: Experimental Evidence"

by Juha Tolvanen, James Tremewan and Alexander K. Wagner

MODEL: PROOFS

In this section, we provide the proofs of Proposition 1 and 2 for the model in the main text.

Proof of Proposition 1. We prove the proposition with backward induction. Consider first the voters' problem. Suppose that all non-centrist candidates are expected to choose $A = \{-1, 1\}$ while all centrist candidates are expected to choose $A = \{0\}$. On the equilibrium path, it is clearly weakly dominant for a centrist voter to vote for $\{0\}$ in all subgames where that is offered. In all other subgames, both off and on the equilibrium path, the centrists are indifferent over all of the alternatives and hence it is compatible with equilibrium for them to randomize with equal probability over any two non-centrist alternatives offered to them.

Consider next a rightist voter, $\tau = 1$. Consider first the only non-trivial pair of platforms on the equilibrium path where $A_1 = \{-1, 1\}$ and $A_2 = \{0\}$. Since, conditional on winning, it is strictly dominant for Candidate 1 to implement the policy that matches her type, the expected utility for the rightist voter from voting for A_1 is then

$$\begin{aligned} & \mathbb{P}(\tau_1 = 1 \mid \tau = 1, \tau_1 \in \{-1, 1\}) \times 1 + \mathbb{P}(\tau_1 = 1 \mid \tau = 1, \tau_1 \in \{-1, 1\}) \times 0 \\ = \;& \mathbb{P}(\tau_1 = 1 \mid \tau = 1, \tau_1 \in \{-1, 1\}). \end{aligned} \tag{1}$$

Strictly speaking, the above is true only when the voter behaves non-strategically in the sense that she does not condition on the event of being pivotal. Here the voter is never pivotal on the equilibrium path; see Tolvanen (2020) for a version of the model that allows for pivotality considerations. However, any refinement that allows for mistakes from other voters will result in the same calculation as below, since the pivotality event will be independent from the state and the type of the ambiguous candidate. In other words, there is nothing to be learned from conditioning on being pivotal.

By the conditional independence of the types conditional on the state,

$$\begin{aligned} &\mathbb{P}(\tau_1 = 1 \mid \tau = 1, \tau_1 \in \{-1, 1\}) \\ = &\sum_{i \in \{L,R\}} \mathbb{P}(\tau_1 = 1 \mid \tau_1 \in \{-1, 1\}, \omega = i)\mathbb{P}(\omega = i \mid \tau = 1, \tau_1 \in \{-1, 1\}). \end{aligned} \tag{2}$$

We will solve for each of the terms above separately. First, by the definition of a conditional probability,

$$\begin{aligned} &\mathbb{P}(\tau_1 = 1 \mid \tau_1 \in \{-1, 1\}, \omega = 1) \\ = &\frac{\mathbb{P}(\{\tau_1 = 1\} \cap \{\tau_1 \in \{-1, 1\}\} \mid \omega = R)}{\mathbb{P}(\tau_1 \in \{-1, 1\} \mid \omega = R)} \\ = &\frac{p}{p + q} \end{aligned} \tag{3}$$

Analogously,

$$\mathbb{P}(\tau_1 = 1 \mid \tau_1 \in \{-1, 1\}, \omega = 1) = \frac{q}{p + q}. \tag{4}$$

The remaining two terms can be calculated using Bayes' Law:

$$\begin{aligned} &\mathbb{P}(\omega = R \mid \tau = 1, \tau_1 \in \{-1, 1\}) \\ = &\frac{\mathbb{P}(\tau = 1 \mid \omega = R)\mathbb{P}(\tau_1 \in \{-1, 1\} \mid \omega = R)\mathbb{P}(\omega = R)}{\sum_{i \in \{L,R\}} \mathbb{P}(\tau = 1 \mid \omega = i)\mathbb{P}(\tau_1 \in \{-1, 1\} \mid \omega = i)\mathbb{P}(\omega = i)} \\ = &\frac{p \times \frac{1}{2}}{p \times \frac{1}{2} + q \times \frac{1}{2}} = \frac{p}{p + q} \end{aligned} \tag{5}$$

where the first equality follows from Bayes' Rule and the conditional independence of the voters' and candidate's type conditional on the state and the second follows as $\mathbb{P}(\omega = R) = \mathbb{P}(\omega = L) = \frac{1}{2}$ and $\mathbb{P}(\tau_1 \in \{-1, 1\} \mid \omega = R) = \mathbb{P}(\tau_1 \in \{-1, 1\} \mid \omega = L)$. An identical argument yields

$$\mathbb{P}(\omega = L \mid \tau = 1, \tau_1 \in \{-1, 1\}) = \frac{q}{p + q}. \tag{6}$$

Now, substituting equations (3)-(6) into (2) yields:

$$\mathbb{P}(\tau_1 = 1 \mid \tau = 1, \tau_1 \in \{-1, 1\}) = \frac{p^2 + q^2}{(p+q)^2}.$$

Therefore, given the equilibrium hypothesis that all non-centrist candidates submit $A = \{-1, 1\}$, a voter of type $\tau = 1$ will best respond by voting for that candidate, if and only if

$$\frac{p^2 + q^2}{(p+q)^2} \geq u_0.$$

By symmetry, the same will hold for a voter whose type is $\tau = -1$. Off the equilibrium path, a voter whose type is $\tau = 1$ will find it weakly dominant to vote for platform $A = \{1\}$ over any other platform, for platform $A = \{0\}$ over platform $A = \{-1\}$. Furthermore, since the utility from platform $A = \{0\}$ is higher than from $A = \{-1\}$, the same voter will best respond by voting for $A = \{-1, 1\}$ over $A = \{-1\}$ when $u_0 \geq \frac{p^2+q^2}{(p+q)^2}$.

Now that we have indentified a strategy profile that is a best response for all voters we can turn to candidate behavior. Centrist candidates have a weakly dominant strategy to always play $A = \{0\}$. Consider then Candidate 1 when her type is $\tau_1 = 1$. On the equilibrium path, by following the suggested equilibrium strategy, they get their favorite policy implemented in all cases except when the other candidate has type $\tau_2 = -1$ *and* that candidate wins the equal randomization by all of the indifferent voters. In other words, the expected payoff from choosing $A = \{-1, 1\}$ is

$$1 - \frac{1}{2}\mathbb{P}(\tau_2 = -1 \mid \tau_1 = 1)$$

In the case where the opponent has type $\tau_2 = -1$, deviating to platform $\{0\}$ is not helpful, as all of the non-centrist voters will vote against her and she will lose with certainty and in all of the other eventualities she was getting her favorite outcome following the suggested equilibrium strategy. In other words, given the voters' strategy profile this deviation loses in all realizations of the state and the opponent's type. Deviating to $A = \{1\}$ loses to a centrist and wins against $A = \{-1, 1\}$ at most when $\omega = 1$ (depending on centrists' randomization). In this case the only time when the candidate strictly benefits from the deviation is when $\tau_2 = -1$ and they would lose the equal randomization by the voters. In other words, an upper bound for the payoff from this deviation is:

$$\begin{aligned} & \mathbb{P}(\tau_2 = 0 \mid \tau_1 = 1)u_0 + \mathbb{P}(\tau_2 = 1 \mid \tau_1 = 1) + \mathbb{P}(\tau_2 = -1 \mid \tau_1 = 1)\mathbb{P}(\omega = 1 \mid \tau_1 = 1, \tau_2 = -1) \\ = \; & \mathbb{P}(\tau_2 = 0 \mid \tau_1 = 1)u_0 + \mathbb{P}(\tau_2 = 1 \mid \tau_1 = 1) + \frac{1}{2}\mathbb{P}(\tau_2 = -1 \mid \tau_1 = 1) \end{aligned}$$

Where the equation follows since, by symmetry, both states are equally likely conditional on $\tau_1 = 1$ and $\tau_2 = -1$. Taking the difference of the expected payoff from following the equilibrium strategy and this deviation yields

$$\begin{aligned} & 1 - \mathbb{P}(\tau_2 = -1 \mid \tau_1 = 1) - \mathbb{P}(\tau_2 = 0 \mid \tau_1 = 1)u_0 - \mathbb{P}(\tau_2 = 1 \mid \tau_1 = 1) \\ > \; & 1 - \mathbb{P}(\tau_2 = -1 \mid \tau_1 = 1) - \mathbb{P}(\tau_2 = 0 \mid \tau_1 = 1) - \mathbb{P}(\tau_2 = 1 \mid \tau_1 = 1) = 0 \end{aligned}$$

where the inequality follows as $u_0 < 1$. In other words, the deviation never pays off. An symmetrical argument holds for a candidate whose type is $\tau_1 = 1 - 1$ proving the Proposition. □

Proof of Proposition 2. We will show that when $u_0 \geq \frac{p^2+q^2}{(p+q)^2}$ then truthful platforms are part of a perfect Bayesian equilibrium. We again start by analyzing voters' behavior but now we assume that candidates are expected to commit to a platform matching their own type. Consider first the possible combinations of equilibrium path platforms. If a voter of type $\tau = 1$ is faced with a choice between a platform $\{1\}$ and any other equilibrium path platform (i.e. $\{0\}$ or $\{-1\}$) it is weakly dominant for her to vote for $\{1\}$. Similarly, when faced with a choice between $\{0\}$ and $\{-1\}$, it is weakly dominant for her to vote for $\{0\}$. As a symmetric argument holds for a voter of type $\tau = -1$, we know that non-centrist voters best respond to equilibrium path platforms by choosing their favorite platform over everything else and by choosing the centrist platform over the opposing platform. Similarly, centrists best respond by choosing the centrist platform over both non-centrist platforms and are free to randomize as they want over two non-centrist platforms. To consider the most adversarial case, we assume that the centrists always vote for $\{-1\}$. In other words, higher u_0 are compatible with the equilibrium if centrists split their vote.

Consider next voters observing the off-equilibrium platform $A = \{-1, 1\}$. As this is off the equilibrium path we can assume that all voters believe that only rightist candidates choose this platform.

Then all types of voters are best responding if they vote according to their strategies above but treating $\{-1, 1\}$ as the platform $\{1\}$.

We then show that candidates best-respond to this voters' strategy profile by posting truthful platforms. Notice first, that given the voters' profile, a centrist platform will always win the election if it is available, because it will gain the support of the centrists and at least one of the non-centrist voter groups. Hence its vote share will always be at least $1 - p - q + q = 1 - p > \frac{1}{2}$. Consequently, centrist candidates get their maximum utility by playing $\{0\}$ and hence cannot benefit from deviations. Consider next a non-centrist Candidate 1 of type $\tau_1 = 1$. Given the off-path behavior specified for the voters above, deviating to $\{-1, 1\}$ will yield the same utility as committing to $\{1\}$. Deviating to $\{-1\}$ is dominated by $\{0\}$. Hence, it is enough to show that the deviation to $\{0\}$ is not beneficial. This deviation will win the election with certainty and the expected utility from it is hence u_0. If the candidate plays the truthful strategy instead, she loses whenever the opponent is type $\tau_2 = 0$ or $\tau_2 = -1$. Otherwise she gets her favorite policy. It follows that the payoff from playing the conjectured equilibrium strategy is

$$\begin{aligned}
&\mathbb{P}(\tau_2 = 1 \mid \tau_1 = 1) + \mathbb{P}(\tau_2 = 0 \mid \tau_1 = 1)u_0 \\
=\ & \frac{\sum_{i\in\{R,L\}} \mathbb{P}(\omega = i)\mathbb{P}(\{\tau_2 = 1\} \cap \{\tau_1 = 1\} \mid \omega = i)}{\sum_{i\in\{R,L\}} \mathbb{P}(\omega = i)\mathbb{P}(\tau_1 = 1 \mid \omega = i)} \\
&+ \frac{\sum_{i\in\{R,L\}} \mathbb{P}(\omega = i)\mathbb{P}(\{\tau_2 = 0\} \cap \{\tau_1 = 1\} \mid \omega = i)}{\sum_{i\in\{R,L\}} \mathbb{P}(\omega = i)\mathbb{P}(\tau_1 = 1 \mid \omega = i)} u_0 \\
=\ & \frac{\frac{1}{2}(p^2 + q^2)}{\frac{1}{2}(p + q)} + \frac{\frac{1}{2}(1 - p - q)(p + q)}{\frac{1}{2}(p + q)} u_0 \\
=\ & \frac{p^2 + q^2}{p + q} + (1 - p - q)u_0
\end{aligned}$$

Hence the deviation is not beneficial as long as

$$\frac{p^2 + q^2}{p + q} + (1 - p - q)u_0 \geq u_0 \Leftrightarrow u_0 \leq \frac{p^2 + q^2}{(p + q)^2}. \tag{7}$$

By symmetry, the same arguments hold for a candidate of type $\{-1\}$. Notice also that if this inequality does not hold *and* the centrists have a strict preference for the left-wing policy, then the truthful equilibrium will not exist, because right-wing politician will rather deviate to $\{0\}$. □

Proof that a fully separating equilibrium may not exist. Since a non-centrist candidate lacks information about the realized state of the world, she runs the risk of being a minority non-centrist and losing the election to a majority one. The opportunity cost of this outcome is relatively high when u_0 is large. A good insurance against this outcome is to deviate from her truthful platform to $A = \{0\}$ and guarantee herself u_0. The right-hand side term $\frac{p^2+q^2}{(p+q)^2}$ in the proposition above is simply the probability with which the candidate's favorite policy wins conditional on the other candidate being a non-centrist. This happens in two possible ways, either the candidate herself is a majority non-centrist or both candidates are minority non-centrists. It is easy to verify that, conditional on the opposing candidate being a non-centrist, the losing alternative where the candidate herself is a minority non-centrist and the opponent is a majority non-centrist happens with probability $\frac{pq}{(p+q)^2}$. Given Remark 1, if both candidates are expected to commit to singleton platforms,[19] non-centrist policies matter only when both candidates are non-centrists. Conditional on both being non-centrists, committing to one's favorite policy yields the expected utility of $\frac{p^2+q^2}{(p+q)^2} \times 1 + \frac{pq}{(p+q)^2} \times 0$. Hence, there is an incentive to deviate to the centrist position, if and only if this quantity is less than u_0. The equilibria for the case when $u_0 > \frac{p^2+q^2}{(p+q)^2}$ are not relevant given our parametrization of the experimental game. □

MODEL: ROBUSTNESS

In this section, we discuss the robustness of the model along a number of dimensions.

Asymmetric centrist preferences. Tolvanen (2020) shows that the results do not change qualitatively if the centrist prefers one of the non-centrist policies over the other (while still preferring the centrist policies over everything else). If centrists have such preferences, Proposition 1 in the main text stays valid as long as $p + q \leq \frac{2}{3}$ (see Proposition 4 and Lemma 2 in Tolvanen 2020). The intuition here is that if there are too many non-centrists, the non-centrist who is preferred by the centrist may benefit from deviating to her favorite platform, because this will guarantee her the victory when running against $\{-1, 1\}$. We prove our Proposition 2 without assuming any knife-edge randomization from the centrists. Consequently, the equilibrium will exist for a larger range of u_0 if exactly half of the non-centrists can

[19] In other words, when deviations to $A = \{-1, 1\}$ are punished with unfavorable off-equilibrium beliefs.

be expected to vote for each platform whenever they are indifferent between platforms. This, however, seems like a very particular case.

Larger action space. Allowing for the whole power set does not change the qualitative results as shown in Tolvanen (2020). The equilibria presented here will survive the added platforms when deviation are discourgaed with suitably adversarial off-equilibrium beliefs. Furthermore, centrists will never want to join forces with either of the non-centrists and run with $\{-1, 0\}$ or $\{0, 1\}$, because a deviation to $\{0\}$ will always win the election. Last, a complete babbling equilibrium where everyone plays $\{-1, 0, 1\}$ exists for a range of high values of u_0. This is the only additional equilibrium that is not outcome equivalent with the equilibria in this paper. However, that equilibirum is not a perfect sequential equilibrium in the sense of Grossman and Perry (1986), because both types of non-centrists lose compared to our ambiguous equilibrium while the centrists prefer the full babbling equilibrium to it. Consequently, the refinement requires expecting both non-centrist types to deviate to $\{-1, 1\}$ with equal probability and the centrist to never play it. The correlation between the voters and candidates then implies that, if that platform is played, both non-centrist types of voters prefer voting for it over $\{-1, 0, 1\}$.

Slighly office motivated candidates. All of the results are robust to introducing small office motivations for candidates. For example, in the ambiguous equilibrium, centrists currently win only when faced with another centrist. Hence, the benefit a centrist enjoys from holding office and implementing one of the non-centrist policies with certainty should not exceed the expected payoff from facing another centrist and implementing the centrist policy. Because the first scenario is strictly worse than the second with zero office motivation, one can find a range of payoffs from holding office where the ambiguous equilibrium still survives. Even when all centrists have an incentive to deviate, the equilibrium will survive (with a different bound on u_0) as long as the centrists will randomize close to equally between the two non-centrist policies if elected. This would guarantee that the inclusion of centrists among the candidates committing to $A = \{-1, 1\}$ does not destroy the correlation between the implemented policy and the voter's type. Similarly, the honest equilibria will survive as long as $u_0 + o(p + q + \frac{1}{2}(1 - p - q)) \leq \frac{p^2+q^2}{(p+q)^2}$ where o is the payoff from holding the office. This guarantees that non-centrists do not have an incentive to deviate to the centrist platform instead of playing their preferred platform. Here, $p + q$ is the probability of running against another non-centrist and hence winning the office certainly with the deviation. The residual probability corresponds to the chance of running against a non-centrist, in which case the deviation wins the office with probability $1/2$.

MODEL EXTENSION: ASYMMETRIC IMPACT OF SOPHISTICATION

In this section, we propose a simple model extension which can explain the experimentally observed asymmetric impact of our sophistication measures (correlation awareness and strategic reasoning) on behavior of candidates and voters. Recall that the rational choice model, presented in the main text, rests on the agents' ability to reason strategically and to understand the correlated preference structure in the population. In this extension, we add boundedly-rational information processing of agents to the model by assuming that agents exhibit some degree of false consensus bias or projection bias (cf. Ross et al. 1977; Jensen 2009). This bias is widely documented in experimental and empirical work, and implies, in our context, that an agent overestimates the probability that others share the same preferences type. We add the assumption of false consensus bias in this extension as it is a reasonable approach to generate the experimentally observed asymmetric effect of our sophistication measures on behavior of candidates and voters. More generally, our approach illustrates how perceptual or cognitive biases can directly contribute to the effect correlation awareness has on behavior in the fully rational model presented in the main text.

Maintaining the general setup of the rational choice model in the main text, we now assume that all agents exhibit a false consensus bias regarding the underlying preference distribution in the population. Specifically, if agent i has preference type τ_i and the rational posterior about agent j's type matching i's type, given agent i's information $\mathcal{I}_i$, is $\mathbb{P}(\tau_j = \tau_i \mid \mathcal{I}_i) = \alpha$, then agent i believes that $\tau_j = \tau_i$ with probability $\alpha + (1 - \alpha)\beta$ where $0 < \beta < 1$. β represents the degree of false consensus bias, with $\beta = 0$ for unbiased individuals, and $\beta = 1$ for those who assume that everybody is of their own type.

We also assume that none of the agents understands that others suffer from this bias. This seems reasonable, since understanding that others suffer from a bias should allow players to realize their own bias. Furthermore, we assume that this bias does not affect conjectures about the distribution of preference types in the whole population and applies only to random draws from the population. Assume further that all agents can be either correlation aware or correlation neglecting. Correlation aware agents are able to perform Bayesian updating correctly while correlation neglecting agents treat all random variables as independent. We denote agent i's correlation awareness as $CA_i \in \{1, 0\}$ where 1 denotes correlation awareness and 0 correlation neglecting.

Regarding strategic reasoning, we develop a level-k type model. Rather than level-k players best responding to level-(k-1), we assume that they can perform k iterated deletions of *weakly* dominated strategies and believe that other players are only capable of k-1 deletions. As is standard in level-k models, we assume that level 0 agents randomize uniformly over their actions. This implies the following strategy spaces for different types of voters and candidates in the KC treatment.

Level-1 voters: Since, one candidate is committed to the centrist platform, the only weakly non-dominated strategy for centrists is to vote for a centrist. For non-centrists, the opposing non-centrist platform is weakly dominated by voting for the centrist.

Level-1 candidates: Proposing the opposing non-centrist platform is weakly dominated. Furthermore, if one wins with the ambiguous platform, choosing the opposing policy is strictly dominated.

Level-2 voters: Nothing is weakly dominated given the removals in the previous round.

Level-2 candidates: Proposing one's favorite platform will never win, since neither the opposing non-centrists nor the centrists will vote for it by level 1. However, if all non-centrists vote for the ambiguous platform, it *will* win. Hence, the non-centrist's honest platform is weakly dominated. The same holds for the centrist platform, since the expected payoff from it is the same as from the honest non-centrist platform. Hence, non-centrists who are level 2 or higher will play the ambiguous platform.

Level-3 voters: The level 2 removal for candidates above implies that level 3 voters must believe that *all* non-centrists play the ambiguous platform. In other words, the prior probability (without knowing one's own type) that a candidate running with an ambiguous platform is a leftist must be 50%.

We assume that all voters are at least level 1. This is validated by the observation that agents in our game almost never played weakly dominated strategies (see Tables 3 and 4). We denote the strategic reasoning ability of agent i by $SR_i \in \{1, 2, 3\}$ where each level corresponds to the levels of strategic reasoning above. In line with the level-k literature, we assume agents of level k to believe others to be level k-1. Specifically this implies that level 1 agents best respond to other agents uniformly randomizing over their actions.

Consider the known centrist (KC) treatment of the experiment. We solve for the optimal actions of all CA and SR-types of voters and candidates backwards. We start with the behavior of voters. Consider first a correlation unaware ($CA = 0$) level 1 ($SR = 1$) voter whose preference type is $\tau_i = 1$. Since this voter is only able to play weakly non-dominated actions, they will vote for a candidate who commits to $a = 1$ over the known centrist, and will vote for the known centrist over a candidate who commits to -1. Now, consider the situation in which the voter is faced with a candidate who chooses the ambiguous platform $\{-1, 1\}$. Since the voter is level 1, they believe that the ambiguous candidate will randomize equally between the two policies if they win the election. Consequently, the voter's expected payoff from voting for this candidate is 0.5×1, whereas the incumbent centrist offers them a certain payoff of u_0. Hence, the level 1 voter will vote for the known centrist as long as $u_0 > 0.5$. The exactly same logic holds for level 1 voters who understand correlations ($CA = 1$).

Consider next a voter whose type is ($\tau_i = 1, SR = 2, CA = 0$). The only interesting case is still when the unknown candidate chooses the ambiguous platform. The difference to the case where the voter had type $SR = 1$ is that now the voter understands that if the ambiguous candidate is a leftists and they win, they will implement the leftist policy and vice versa. However, both non-centrist types of level 1 candidates have multiple surviving actions and we have not constrained the belief formation of level 2 voters over conditional on seeing these actions. Notice however, that as long as any asymmetry between the two non-centrist voter types in this updating is captured by the false consensus bias β, it is enough to consider the case where the voter believes that both non-centrist types play the ambiguous platform with the same probability. In this case the expected utility from voting for the ambiguous candidate is $0.5 + \beta$. Hence the voter will vote for the ambiguous candidate only if $0.5 + \beta > u_0$.

The case for ($\tau_i = 1, SR = 2, CA = 1$) is fairly similar to above. If we still assume that any asymmetry in updating between the two non-centrist voter types is captured by β, then the only difference to the case above is that these voters will correctly update their beliefs after observing $\{-1, 1\}$. In this case the expected utility from voting for the ambiguous candidate is $\frac{p^2+q^2}{(p+q)^2} + \beta$. Hence the voter will vote for the ambiguous candidate only if $\frac{p^2+q^2}{(p+q)^2} + \beta > u_0$.

Notice, however, that both types of $SR = 2$ voters may have been expecting the honest platform from the candidates. In this case their "off-equilibrium" beliefs can be more adversarial than this. For example, all voters of type $SR = 2$ may believe that only leftists play $\{-1, 1\}$ in which case only leftists with $SR = 2$ will vote for the ambiguous platform no matter what their CA type is.

The case of $(\tau_i = 1, SR = 3, CA = 0)$ is highly similar to $(\tau_i, SR = 2, CA = 0)$. The only difference is that these voters expect all of the non-centrist candidates to play the ambiguous platform. Hence, there is no room for "off-equilibrium" beliefs and these voters will vote for the ambiguous candidate, if and only if $0.5 + \beta > u_0$.

The remaining type $(\tau_i = 1, SR = 3, CA = 1)$ behaves like the rational type in the main section except that they are influenced by the false consensus bias. Hence, they vote for the ambiguous candidate, if and only if $\frac{p^2+q^2}{(p+q)^2} + \beta > u_0$.

To summarize the reasoning above, voters who are level 3 or higher and correlation aware ($CA = 1$) have the strongest incentive to vote for the ambiguous candidate. Depending on how level 2 voters who are correlation aware form their beliefs, they may have the same incentives but at least some of them may also hold adversarial beliefs when faced with $\{-1, 1\}$. Similarly, the next highest incentives are for voters with $SR = 3$ and $CA = 0$, while the incentives for voters with $SR = 2$ and $CA = 0$ depend on their belief formation. Level 1 voters (whether correlation aware or not) face the lowest incentives to vote for the ambiguous candidates because they regard the ambiguous platform as a 50/50 gamble over the two non-centrist policies.

The model in the main text already highlights why the multiple non-dominated equilibria in the baseline (BL) treatment might lead voters to support ambiguous platforms less than in the known centrist (KC) treatment even when agents are fully rational. The model above highlights how the treatment difference across levels of strategic sophistication can be generated by two forces: First, independent of correlation awareness, $SR = 1$ voters are not affected by the false consensus effect while the others are and hence vote less for the ambiguous platforms than others. Second, again independent of correlation awareness, $SR = 3$ voters have only 1 valid conjecture about the types who play the ambiguous platform and in what proportions, while the other types are more free to choose different beliefs and hence $SR = 3$ voters potentially vote more for ambiguous platforms. Notice also that, if strategic sophistication helps voters to find the truthful equilibrium in the BL treatment, this will strengthen the result, as strategically sophisticated subjects have more reasons to vote against the ambiguous candidate in the BL treatment compared to the strategically unsophisticated voters. The within treatment difference, although statistically insignificant, goes in this direction, as shown in Figure 6(b)).

Consider now the candidate's problem and suppose that $u_0 \geq 0.5$, i.e. that all agents are relatively risk-averse. All level 1 candidates will regard both non-dominated platforms equally good, since they believe voters to randomize evenly between the two candidates. Notice then that candidates who are level 2 or higher but who are unaware of the correlation ($CA = 0$), base their weak dominance argument on voters playing a strategy that yields them $0.5 < u_0$. Consequently, a candidate who is unaware of the correlation ($CA = 0$), conjectures that even the ambiguous platform will be losing and hence must posit a conjecture about the voters making a mistake when voting to win. For correlation aware ($CA = 1$) candidates of level 2 or higher, the ambiguous equilibrium exists whenever $\frac{p^2+q^2}{(p+q)^2} > u_0$. In this case, they understand that voters can have a strict preference for voting for them and hence there is a clear way of breaking the tie between the two platforms.

Notice also that if we allow the false consensus bias to extend to conjectures about the distribution of preferences in the population, a level 2 candidate who is unaware of the correlation and believes that $u_0 \geq 0.5$ for others, expects $0.5(p + q) + \beta$ votes when running with the honest platform, while expecting 0 votes when running with the ambiguous platform (since they think that all voters prefer the centrist platform, as $u_0 \geq 0.5$, and they do not project their own bias on other voters). With strong enough a bias, the honest platform may even seem winning. Even if $\beta = 0$, they may conjecture that fewer voters need to make a mistake in favor of the honest platform, compared to what is needed if they run with the ambiguous platform.

To summarize, the arguments above highlight two things. First, the iterative deletion of weakly dominated strategies stops earlier for candidates compared to voters. Hence, the game is arguably strategically simpler for candidates compared to the voters, explaining why the treatment effects we observe in our experiment are sensitive to strategic sophistication only for voters. Second, candidates who are unaware of the correlation structure will strictly speaking be indifferent between the honest and

TABLE A.1. Platform choice by candidate type and treatment

	Centrist		Rightist		Leftist	
Platform choice	BL	KC	BL	KC	BL	KC
Centrist	0.93	0.96	0.10	0.14	0.08	0.16
Rightist	0.02	0.02	0.61	0.55	0.01	0.02
Leftist	0.02	0.01	0.01	0.02	0.61	0.55
Ambiguous	0.03	0.01	0.28	0.29	0.30	0.27
Observations	608	679	496	420	496	501

Notes: Proportion of candidates of given preference type choosing given platform in each treatment (BL and KC).

the ambiguous platform, since neither will win. Only understanding correlations allows them to break this tie, explaining the importance of correlation awareness for candidates in our experimental results.

EXPERIMENT: ADDITIONAL RESULTS

Behavior by preference types

In the main text, we pooled non-centrist types, i.e. -1 (or white) and 1 (or black), and referred to them simply as non-centrist types in the analysis because they are pure labelling; see Table 3 for candidates and Table 4 for voters. The disaggregated data from candidates and voters of different non-centrist types shows that there is no labelling effect. Choice proportions of the centrist and non-centrist platform vary by at most one or two percentage points between 'leftists' and 'rightists', see Tables A.1 and A.2 respectively. For both candidates and voters, Wilcoxon rank sum tests based on session-level averages fail to find statistically significant differences in either treatment between the proportions who choose the centrist platform, their preferred non-centrist platform, the opposite non-centrist platform, or the ambiguous platform.

Probit regressions

Candidate behavior. To substantiate our non-parametric analysis of the treatment effect and the effects of the individual-level sophistication controls, we also performed a regression analysis. Table

TABLE A.2. Votes for platform 2 by voter type and treatment

	Centrist		Rightist		Leftist	
Given platform 1/2	BL	KC	BL	KC	BL	KC
Center/Right	0.04	0.02	0.90	0.90	0.09	0.10
Center/Left	0.05	0.03	0.08	0.10	0.87	0.86
Center/Ambiguous	0.05	0.04	0.53	0.66	0.55	0.60
Right/White	0.46	–	0.06	–	0.94	–
Right/Ambiguous	0.57	–	0.09	–	0.92	–
Left/Ambiguous	0.64	–	0.92	–	0.09	–
Observations	960	960	715	690	725	750

Notes: Proportion of voters of given preference type voting for platform 2 for each combination of platform competition and treatment (BL and KC).

A.3 reports the results of a series of probit regressions on the probability of non-centrist candidates choosing ambiguous platforms. Model (1) includes only a treatment dummy for the KC treatment, but the coefficient is not significant (as in non-parametric tests on raw data level). Adding individual-level controls, model (2) includes a dummy for our individual-level variables of correlation awareness and strategic reasoning elicited in the additional tasks after the voting part of the experiment. It also includes participants' willingness to take risks and socio-demographic controls.

Addressing the main behavioral prediction (Hypothesis 1), model (3) includes both sophistication measures and its treatment interactions. We find that a large and significant increase in the probability of running on an ambiguous platform by candidates who understand correlations but not for our measure of strategic reasoning. Model (4), our preferred specification, adds the full set of controls and again the treatment interaction with correlation awareness is significant ($p = 0.012$) and supports our candidate results of the non-parametric analysis in the main text. Although the sign of being a strategic reasoner is positive as expected, both in model (3) and (4), it is not significantly different from zero. The interaction with the treatment dummy does also not significantly influence the likelihood of the ambiguous platform being played. Regarding controls in model (4), the risk coefficient suggests that those who are more willing to take risks run slightly less on ambiguous platforms ($p = 0.044$). Even though this effect may seem counter-intuitive, closer inspection indicates that it is driven by non-centrist candidates who are playing the other non-centrist singleton platform (a loosing strategy) with a high probability, and these candidates are also the ones who are willing to take on high risks

TABLE A.3. Non-centrist candidates choosing ambiguous platform

	(1)	(2)	(3)	(4)
KC (treatment)	-0.042	-0.079	-0.238	-0.253
	(0.167)	(0.147)	(0.207)	(0.176)
Correlation awareness		0.238	0.028	-0.174
		(0.184)	(0.250)	(0.239)
Strategic reasoning		-0.004	0.075	0.088
		(0.156)	(0.177)	(0.210)
KC × correlation awareness			0.708**	0.819**
			(0.358)	(0.328)
KC × Strategic reasoning			-0.120	-0.209
			(0.276)	(0.282)
Risk taking		-0.054*		-0.056**
		(0.028)		(0.028)
Female		-0.466***		-0.511***
		(0.174)		(0.164)
Age		-0.008		-0.007
		(0.009)		(0.009)
Constant	-0.544***	0.119	-0.583***	0.215
	(0.117)	(0.368)	(0.158)	(0.363)

Notes: Probit panel regressions reporting coefficients. Standard errors clustered at the session level in parentheses. Dependent variable in all models is non-centrist candidate choosing ambiguous platform. Number of observations in each model is 1913. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

according to our risk attitude question. The average answer to the risk question, with higher numbers indicating a higher inclination to take risks, is not significantly different for those who chose the centrist platform (5.54), their own preferred platform (5.80), or the ambiguous platform (5.47), but much higher for those who chose the other non-centrist platform (8.17). Lastly, females are significantly less likely to run on an ambiguous platform than males ($p = 0.002$), whereas age does not impact the probability of running on an ambiguous platform in the electoral competition for candidates.

Voters behavior. To corroborate results from the non-parametric analysis of voter behavior in the main text we provide the results of a series of probit regressions on the probability of voting for an ambiguous platform over a centrist one in Table A.4. The baseline specification in model (1) includes only a dummy for treatment KC. In line with findings in the above non-parametric tests on the unrestricted voter choice data, the treatment coefficient is positive and significant ($p = 0.011$). Model (2) adds a set of individual-level controls including a dummy for correlation awareness and strategic reasoning, our measures of voter sophistication. Adding these controls leads to a slightly

TABLE A.4. Non-centrist voters voting for ambiguous platform

	(1)	(2)	(3)	(4)	(5)
KC (treatment)	0.225**	0.246***	-0.068	-0.050	-0.132
	(0.089)	(0.093)	(0.195)	(0.182)	(0.478)
Correlation awareness		0.042	-0.032	0.024	0.020
		(0.186)	(0.243)	(0.256)	(0.257)
Strategic reasoning		0.157	-0.135	-0.141	-0.139
		(0.130)	(0.122)	(0.111)	(0.110)
KC × Correlation awareness			0.110	0.079	0.085
			(0.348)	(0.307)	(0.320)
KC × Strategic reasoning			0.617***	0.618***	0.619***
			(0.235)	(0.223)	(0.224)
Risk taking		0.064**		0.067**	0.061
		(0.032)		(0.031)	(0.037)
KC × Risk taking					0.015
					(0.065)
Female		-0.011		0.016	0.018
		(0.159)		(0.149)	(0.147)
Age		-0.016*		-0.013	-0.013
		(0.010)		(0.010)	(0.096)
Constant	0.101	0.079	0.171	0.116	0.148
	(0.071)	(0.428)	(0.134)	(0.428)	(0.431)

Notes: Probit panel regressions reporting coefficients. Standard errors clustered at the session level in parentheses. Dependent variable in all models is non-centrist voter voting for ambiguous platform. Number of observations in each model is 2880. Significance levels: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

larger coefficient of the KC treatment dummy than in the model without controls ($p = 0.008$). The controls for individual-level correlation awareness and strategic reasoning are positive as expected but not statistically significant. The coefficient for a subject's willingness to accept risks is positive and significant ($p = 0.045$). The magnitude of the risk effect is comparable to the one found for candidates above however rather small compared to the treatment effect. We include risk attitude as a potential control in our analysis, although not a key driver in our model, because the early literature on vague or ambiguous policies highlighted risk as a potentially important explanatory variable for voters (cf. Bartels 1986; Tomz and Van Houweling 2009). Interacting risk taking with our treatment variable in model (5), we find no significant evidence for risk taking to impact the higher propensity to vote for ambiguous candidates in the KC treatment, consistent with our model.

Assessing Hypothesis 2, model (3) and (4) include an interaction of the KC treatment with each sophistication control simultaneously. Corroborating our non-parametric tests, we do not find a significant positive effect of the treatment interaction effect with correlation awareness but a strong positive interaction effect of voting for an ambiguous platform in the known centrist treatment (KC) and strategic reasoning ($p = 0.009$ in model 3 and $p = 0.006$ in model 4). Including our set of individual-level controls in model (4) does corroborate the findings of model (3). Overall, we find strong support for both parts of Hypothesis 2 according to the above regression analysis.

Robustness of behavior and sophistication measures

All of the results discussed in the main text, for both candidates and voters, are robust to using alternative measures of correlation awareness and strategic reasoning in our regression analysis (see main text for details of our preferred measures). For instance, using a more stringent definition of correlation awareness (CA) which does not allow for a small mistake in either of the two questions (Q1 and Q2) does not change our main results. Nor does using the measure from Enke and Zimmermann (2019), who use $|Q2 - Q1|$ and exclude extreme outliers in regressions (see their Table 3). Results are also robust to various alternatives to our selected definition of strategic reasoning: replacing the cut-off of 34 with 67 (i.e. dominated vs undominated choices), 51, and 23 (level 2 and above) do not change the results qualitatively. The same is true of replacing the dummy variable with the continuous variable of the subject's guess. Regression tables demonstrating these results are available on request.

From individual behavior to outcomes

We describe here how individual behavior of candidates and voters maps into election outcomes and payoffs. Note that our experiment was explicitly designed to test individual level behavior of both candidates and voter as well as its relation to the model's sophistication assumptions required for ambiguous platforms. One direct implication of the experimental setup with the distribution of voter types in Figure 1 is that non-centrist candidates can win an election with only near-unanimous support from voters of *both* non-centrist types. If all centrist voters vote for a centrist candidate and a centrist platform is proposed against any other platform, 8 out of 9 non-centrists have to vote for the the other platform to beat the centrist candidate. Obviously, this is somewhat an artifact of the small elections that can be ran in the lab while retaining a reasonable correlation between the voters' and the

TABLE A.5. Proportion of platforms in electoral competitions

Election type	BL	(Won)	KC	(Won)
Both Center	0.17		0.49	
Center vs Non-centrist	0.33	0.15	0.34	0.11
Center vs Ambiguous	0.15	0.10	0.17	0.13
Both same Non-centrist	0.08			
Different Non-centrist	0.06			
Non-centrist vs Ambiguous	0.17	0.30		
Both Ambiguous	0.03			
No. of observations	800		1600	

Notes: Won is the proportion of elections won by the second option (where relevant). Unit of observation is an individual choice in the game.

TABLE A.6. Payoffs of non-centrist candidates by policy

Platform choice	BL	Obs.	KC	Obs.
Center	8.47	90	8.00	138
Non-centrist (own)	11.45	604	8.82	506
Non-centrist (other)	9.57	7	8.00	20
Ambiguous	11.32	391	8.87	257

Notes: Numbers are average payoff from chosen platform and respective number of observations.

candidates' types. Given the significant proportion of voters who do not understand correlations or does not reason strategically (see results in main text for exact figure), we expect in accordance with theory to see only a relatively small number of winning ambiguous platforms. Again, note that this near-mechanical consequence of our setup in the experimental setup (due to the relatively small number of citizens that can participate in an experimental election) makes it quite difficult for ambiguous platforms to win elections. Table A.5 summarizes the type of elections by treatment. Observe that 15-17% of the elections comprise competition between a centrist platform an ambiguous platform. Of these, ambiguous platforms win 10% of the time in BL and 13% in KC, the difference being weakly significant (one-sided WRS, $p = 0.094$)

Expressing the consequences in terms of realized payoffs for candidates yields the following picture. Within each treatment, running on either your own preferred platform or on an ambiguous platform brings about a higher payoff than running on the centrist platform as shown in Table A.6. In the KC treatment these are also significantly better than promising the opposite colour to your own ball. All these differences are strongly significant (WSR, $p < 0.01$).

The reason why running on the least preferred platform has a slightly higher mean in the BL treatment than in the KC treatment is because in a couple of cases in BL where both candidates had the more numerous non-centrist type, one ran on the 'wrong' platform while the other ran on the 'correct' platform and the one running on the 'correct' platform won. Given the correlation in the setup, such error-correction is quite likely in BL but cannot happen in treatment KC.