Hypothesis Registration: Structural Predictions for the 2013 World Schools Debating Championships

Tom Gole* and Simon Quinn†

January 26, 2013

Abstract

This document registers testable predictions about committee voting behavior in the forthcoming 2013 World Schools Debating Championships. The document will be registered in the Hypothesis Registry of the Abdul Latif Jameel Poverty Action Lab (J-PAL). Testable predictions are recorded in Table 5 on page 13, and in Figures 1 to Figure 18 (pages 14 to 31). We will submit this document to the Hypothesis Registry today (26 January 2013), and we ask the Registry to publish on or after 6 February 2013 (the day after the Championships concludes).

1 Introduction: The randomized field experiment

The World Schools Debating Championships are an annual debate tournament between high school students. Debaters are drawn from around the world to represent their countries; each nation is entitled to one team in the competition. The Championships are the premiere international debate tournament for school students. This document registers predictions for judge voting behavior in the Preliminary Rounds of the 2013 Championships, to be held in Antalya, Turkey, later this month. Our predictions are formed on the basis of structural estimates using data from the previous three Championships, held in 2010, 2011 and 2012. This document outlines the structure of the Championships (including, critically, the method by which judges will be randomly assigned to committees), reports in-sample predictions for the data from 2010, 2011 and 2012, and makes out-of-sample predictions for 2013.

*Department of Economics, Harvard University; tgole@fas.harvard.edu.
†Department of Economics, University of Oxford; simon.quinn@economics.ox.ac.uk.
Hypothesis Registration: The 2013 World Schools Debating Championships

Each debate pitches one national team against another; teams are randomly assigned to argue either for or against a controversial idea. The Championships comprise both Preliminary Rounds and Finals Rounds. In the Preliminary Rounds, each nation competes against eight randomly-drawn opponents. These eight debates occur across four days: Rounds 1 and 2 on the first day, Rounds 3 and 4 on the second day, and so on. The top 16 teams then progress to the Finals Rounds, a series of five knock-out debates culminating in the Grand Final. Our analysis and predictions focuses exclusively on data from the Preliminary Rounds.

Judging committees: The winner of each debate is determined by a committee of three judges. Together, this committee is required to decide which team has argued more persuasively. Judges assess the debate separately, assigning points to speakers based on the categories of ‘style’, ‘content’ and ‘strategy’. A comprehensive explanation of these categories is available at www.schoolsdebate.com. Each judge is required to complete a ballot, in which he or she records speaker points and decides the winner of the debate; judges may not award a tie. The debate is won by whichever team wins two or three of the judges; committee outcomes can therefore be either ‘unanimous’ (3-0) or ‘split’ (2-1).

Judges are not allowed to communicate with each other (or with the competitors) until after making their decisions.\(^1\) Having made their decisions, the three judges then leave the room to confer; judges may not change their decisions after leaving the room. Having discussed the debate together, the committee returns to the room; one judge announces the committee’s result, and gives a brief justification for the committee’s decision. Teams and their coaches are then encouraged to speak separately with the judges; at this point, there is a strong emphasis on constructive feedback.

\(^1\) Judges are seated apart. There is no evidence of judges trying to ‘cheat’ by looking at each other’s notes; indeed, there are strong norms at the Championships against such behavior. Judges are also discouraged from allowing their facial expressions or body language to indicate their views on the debate.
In 2013, as in 2010, 2011 and 2012, judges will be assigned to committees randomly (using a computer), and judges will know this.\(^2\) This assignment will be subject to several constraints, designed to improve the ‘balance’ of the randomization. Most importantly, each committee comprised one ‘class 1’ judge (most experienced/competent), one ‘class 2’ judge and one ‘class 3’ judge (least experienced/competent). These classes will be assigned subjectively by the tournament organizers, to ensure a balance of judging experience across different committees. Second, each committee will include at least one man and one woman, to the extent possible.\(^3\) Third, we will limit cases of judges seeing the same team more than once in the same tournament, and no judge will be allowed to assess his or her own team. Fourth, because the Preliminary Rounds will be divided between different venues, we will often need to assign pairs of judges to committees together in two debates on the same day.

**Pre-tournament rankings:** Teams are ranked before each tournament. On the basis of this ranking, a random draw determines each team’s position in the draw. Pre-tournament ranking is necessary so that each team is drawn against opponents of a range of different qualities; *i.e.* so that a team does not face a disproportionate number of very strong teams in its Preliminary Rounds, nor a disproportionate number of weaker teams. Teams are ranked on the basis of their performance in the Preliminary Rounds of the three previous tournaments. These rankings are public information. The pre-tournament rankings for WSDC 2013 are available online at [www.schoolsdebate.com/years/2013/2013.pre.rankings.pdf.](http://www.schoolsdebate.com/years/2013/2013.pre.rankings.pdf)

---

\(^2\) The computer code is written in Stata, and is available on request.

\(^3\) In the past three tournaments, we were required to relax this constraint four times: in 2010, we allowed one all-male committee, in 2011, we allowed two all-female committees, and in 2012, we allowed one all-male committee.
2 Model, estimates and predictions

We model each committee as an independent Bayesian game between three players. For each committee, we denote judge class by $i \in \{1, 2, 3\}$. Each player $i$ receives a signal $x_i$, and then chooses whether to vote for the favorite ($a_i = 1$) or against ($a_i = 0$). Player $i$ receives utility from two mechanisms: (i) from voting for the team that (s)he prefers (where the strength of that preference is determined by the signal $x_i$), and (ii) from agreeing with player $j$ and/or with player $k$. We treat these mechanisms as additively separable. Note that, for example, $\delta_{ij}$ measures the utility gain for judge $i$ from voting with judge $j$, and that $\delta_i$ measures the gain from agreeing with both judges $j$ and $k$.

\[
U_i(a_i; a_j, a_k, x_i) = \begin{cases} 
  x_i + \delta_i & \text{if } a_i = 1, a_j = 1, a_k = 1; \\
  x_i + \delta_{ij} & \text{if } a_i = 1, a_j = 1, a_k = 0; \\
  x_i + \delta_{ik} & \text{if } a_i = 1, a_j = 0, a_k = 1; \\
  x_i & \text{if } a_i = 1, a_j = 0, a_k = 0; \\
  0 & \text{if } a_i = 0, a_j = 1, a_k = 1; \\
  \delta_{ij} & \text{if } a_i = 0, a_j = 1, a_k = 0; \\
  \delta_{ik} & \text{if } a_i = 0, a_j = 0, a_k = 1; \\
  \delta_i & \text{if } a_i = 0, a_j = 0, a_k = 0. 
\end{cases} 
\]  

(1)

We assume that each judge weakly prefers agreement over dissent: $\delta_{ij}, \delta_{ik}, \delta_i \geq 0$.

For each committee, the distribution of signals is trivariate normal (where we assume positive correlations, $\rho_{12}, \rho_{13}, \rho_{23} > 0$):\(^4\)

\[
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \sim \mathcal{N}\left( \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix} \right). 
\]  

(2)

The signal $x_i$ therefore plays a dual role: it directly affects the relative utility of voting for the favorite, and it determines the conditional expectation of the other judges’ signals:

\(^4\) Of course, as with any trivariate normal, we must also assume a positive definite covariance matrix; this implies the further restrictions $\rho_{12}, \rho_{13}, \rho_{23} < 1$ and $1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12} \cdot \rho_{13} \cdot \rho_{23} > 0$. Note that the restriction $\text{Var}(x_1) = \text{Var}(x_2) = \text{Var}(x_3) = 1$ is made without loss of generality; if we were to parameterize these variances, all of our results would simply rescale by those new parameters. This is the same reasoning that justifies the same restriction for the trivariate probit.
\[
\begin{pmatrix} x_j \\ x_k \end{pmatrix} \mid x_i \sim \mathcal{N}\left(\begin{pmatrix} \mu_j + \rho_{ij} \cdot (x_i - \mu_i) \\ \mu_k + \rho_{ik} \cdot (x_i - \mu_i) \end{pmatrix}, \begin{pmatrix} 1 - \rho_{ij}^2 & \rho_{jk} - \rho_{ij} \cdot \rho_{ik} \\ \rho_{jk} - \rho_{ij} \cdot \rho_{ik} & 1 - \rho_{ik}^2 \end{pmatrix}\right). \]

(3)

Player \(i\) must choose a best response \(a^*_i(x_i)\); we limit attention to cutoff strategies: \(a^*_i(x_i) = 1(x_i \geq x^*_i)\), where \(x^*_i\) denotes the cutoff and \(1(\cdot)\) the indicator function. Player \(i\) must be indifferent between \(a_i = 0\) and \(a_i = 1\) if \(x_i = x^*_i\); that is,

\[
x^*_i = [\Pr(a_j = 0, a_k = 0 \mid x_i = x^*_i) - \Pr(a_j = 1, a_k = 1 \mid x_i = x^*_i)] \cdot \delta_i + \\
[\Pr(a_j = 0, a_k = 1 \mid x_i = x^*_i) - \Pr(a_j = 1, a_k = 0 \mid x_i = x^*_i)] \cdot (\delta_{ij} - \delta_{ik}).
\]

(4)

**Proposition 1 (Conditional State Monotonicity)** Without loss of generality, denote \(\delta_{ij} \geq \delta_{ik}\). Then, in order for player \(i\) to have a unique cutoff \(x^*_i\), it is sufficient that:

\[
(\delta_{ij} - \delta_{ik} - \delta_i) < \frac{\left(1 - \rho_{ik}^2\right)}{\rho_{ik}} \cdot \sqrt{\frac{2\pi \cdot (1 - \rho_{ij}^2)}{1 - \rho_{ij}^2 - \rho_{ik}^2 + 2\rho_{jk} \cdot \rho_{ij} \cdot \rho_{ik}}}.
\]

(5)

**Proof:** Proofs are omitted from this document, but will be presented in an academic paper in due course.

**Proposition 2 (Unique Equilibrium)** It is sufficient for the existence of a unique equilibrium that, for each judge \(i\),

\[
\max \{\delta_i, |\delta_{ij} - \delta_{ik}|\} < \sqrt{\frac{\pi}{2(1 - \omega_{jk}^2)}} \cdot \left(\sqrt{\frac{1 - \rho_{ij}}{1 + \rho_{ij}}} + \sqrt{\frac{1 - \rho_{ik}}{1 + \rho_{ik}}}\right)^{-1},
\]

(6)

where \(\omega_{jk} = \frac{\rho_{jk} - \rho_{ij} \cdot \rho_{ik}}{\sqrt{(1 - \rho_{ij}^2) \cdot (1 - \rho_{ik}^2)}}\).

(7)

\(^5\) Note that the restriction that the covariance matrix is positive definite is sufficient to ensure that the righthand side of equation 5 evaluates to a real number.
2.1 Structural implementation

Parameterization: We estimate common values for $\rho_{12}$, $\rho_{13}$ and $\rho_{23}$ across all committees. For each committee $c$, we denote the difference in pre-tournament rankings by $R_c > 0$. We allow this ranking difference to shift judges’ signal means; for flexibility, we adopt a quadratic specification, and allow a different relationship for each judge class:\(^6\)

\[
\mu_{1c} = \beta_1 \cdot R_c + \gamma_1 \cdot R_c^2; \\
\mu_{2c} = \beta_2 \cdot R_c + \gamma_2 \cdot R_c^2; \\
\mu_{3c} = \beta_3 \cdot R_c + \gamma_3 \cdot R_c^2.
\]

We use the dummy $D_{ic}$ to denote that judge $i$ on committee $c$ dissented in the previous round. We allow this dummy to shift each judge’s preference for agreement. We allow different classes of judges to be differentially affected by previous dissent, and we impose that each judge is indifferent between agreeing with one peer and agreeing with two:

\[
\delta_{1c} = \delta_{12c} = \delta_{13c} = \delta_1 \cdot D_{1c}; \\
\delta_{2c} = \delta_{21c} = \delta_{23c} = \delta_2 \cdot D_{2c}; \\
\delta_{3c} = \delta_{31c} = \delta_{32c} = \delta_3 \cdot D_{3c}.
\]

Equations 11 – 13 show two important limitations of this estimation method. First, the current experimental context does not allow us to identify $\delta_i$, $\delta_{ij}$ and $\delta_{ik}$ separately; this is because the exogenous variation (past dissent) operates at the level of the individual judge. We could use the present structural methodology to separately identify $\delta_i$, $\delta_{ij}$ and $\delta_{ik}$, if we observed some exogenous shock operating at the level of the relationship between judges $i$ and $j$. Second, we use the structural model to identify the additional preference for coordination driven by past dissent, but we do not seek to identify the preference for coordination generally. That is, if

\(^6\) Equations 8 – 10 imply that, in the hypothetical case that two teams were equally matched ($R_c = 0$), then $\mu_{c1} = \mu_{c2} = \mu_{c3} = 0$. This is exactly as we would expect and require.
Hypothesis Registration: The 2013 World Schools Debating Championships

\[ D_{ic} = 0, \text{ we normalize } \delta_{ic} = 0; \text{ if } D_{ic} = 0, \text{ any preference for coordination is already proxied by the function } \mu_{ic}(R_c). \]

**Constraints:** We constrain the estimation so that \( \rho_{12}, \rho_{13}, \rho_{23} \in [0.01, 0.99], \) and so that the covariance matrix is positive definite.\(^7\) We impose \( \delta_1, \delta_2, \delta_3 \geq 0. \) Together, these constraints ensure conditional state monotonicity (Proposition 1). Additionally, we impose the single-equilibrium condition of Proposition 2.

**Identification:**

**Proposition 3 (GLOBAL IDENTIFICATION OF THE THREE-PLAYER PROBIT GAME)** Assume that the conditions in Proposition 1 and Proposition 2 hold, and that \( R_c \) takes at least two unique values. Then the structural model is globally identified.

**Estimation method:** The proof of identification (omitted here) will rely upon a subset of cases on \( (D_{1c}, D_{2c}, D_{3c}). \) But, to estimate efficiently, we use all of our data. Specifically, we use Maximum Likelihood with a nested fixed-point approach. Denote the stacked parameter vector as \( \theta. \) Then, for some candidate value for \( \theta, \) the inner loop solves the game for each committee \( c \) in the dataset,\(^8\) given the ranking data: \( (x^{*}_{1c}(\theta; R_c), x^{*}_{2c}(\theta; R_c), x^{*}_{3c}(\theta; R_c)) \). We then calculate the log-likelihood \( \ell(\theta) \) using a standard triprobit structure (where we approximate the \( cdf \) of the trivariate normal using the method of Genz (2004, ‘Numerical Computation of Rectangular Bivariate and Trivariate Normal and \( t \) Probabilities’)).\(^9\) The outer loop updates

\(^7\) That is, we additionally impose \( 1 - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2 + 2\rho_{12} \cdot \rho_{13} \cdot \rho_{23} > 0. \)

\(^8\) Note that this aspect of the problem — and the calculation of the log-likelihood — is trivially parallelizable. We exploit this to improve dramatically the speed of estimation.

\(^9\) Formally, we calculate the log-likelihood for committee \( c \) as:

\[
\ell_c(\theta; a_{1c}, a_{2c}, a_{3c}) = \ln \Phi_3 \left[ (2a_{1c} - 1) \cdot (\beta_1 \cdot R_c + \gamma_1 \cdot R^2_c - x^*_1(\theta; R_c)), \\
(2a_{2c} - 1) \cdot (\beta_2 \cdot R_c + \gamma_2 \cdot R^2_c - x^*_2(\theta; R_c)) \\
(2a_{3c} - 1) \cdot (\beta_3 \cdot R_c + \gamma_3 \cdot R^2_c - x^*_3(\theta; R_c)) \right], \\
\rho_{12}, \rho_{23}, \rho_{13} > 0, \]

where \( \Phi_3(\cdot) \) refers to the \( cdf \) of the standard trivariate normal. We treat draws of \( (x_1, x_2, x_3) \) as independent across committees, so the sample log-likelihood is \( \ell(\theta) = \sum_{c=1}^{603} \ell_c(\theta). \)
using a Sequential Quadratic Program. We calculate p-values for our parameter estimates using standard Likelihood Ratio tests.

2.2 Structural estimates

We obtain the following structural estimates.

Table 1: Structural estimates: Basic specification

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>ESTIMATE</th>
<th>$\ell_r$</th>
<th>LR</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{12}$</td>
<td>0.725</td>
<td>-902.433</td>
<td>115.467</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\rho_{13}$</td>
<td>0.526</td>
<td>-871.012</td>
<td>52.625</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>0.542</td>
<td>-873.090</td>
<td>56.781</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.064</td>
<td>-869.736</td>
<td>50.074</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-6.912e-4</td>
<td>-847.440</td>
<td>5.480</td>
<td>0.019**</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.045</td>
<td>-855.134</td>
<td>20.868</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.084e-4</td>
<td>-844.735</td>
<td>0.071</td>
<td>0.791</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.035</td>
<td>-852.373</td>
<td>15.346</td>
<td>0.000***</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>0.448e-4</td>
<td>-844.712</td>
<td>0.025</td>
<td>0.875</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.000</td>
<td>-844.699</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>0.000</td>
<td>-844.699</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>$\delta_3$</td>
<td>0.879</td>
<td>-845.427</td>
<td>1.455</td>
<td>0.228</td>
</tr>
</tbody>
</table>

LOG-LIKELIHOOD ($\ell_u$) -844.699

Confidence: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We calculate p-values using the the $\chi^2$ distribution with one degree of freedom. Note that, for the covariance terms, we test $H_0 : \rho_{12} = 0.01$, $H_0 : \rho_{13} = 0.01$ and $H_0 : \rho_{23} = 0.01$; we take this approach because we use 0.01 as the lower bound on the covariance terms throughout.
2.3 In-sample structural predictions

We use a simulation method to make in-sample structural predictions. As explained, each tournament involves random assignment of teams, balanced so that each team faces opponents of a variety of different strengths. For each tournament, we take the actual structure of team matches; this allows us to condition on the actual $R_c$ used in each tournament, in the correct order. We form judging committees, drawing randomly without replacement in each round. We solve the game numerically for each committee in each tournament round, where $R_c$ is given by the actual $R_c$ in the data, and $(D_{1c}, D_{2c}, D_{3c})$ is given by each simulated judge’s performance in the previous simulated debate. We then simulate each tournament 1000 times, to generate a distribution of summary statistics for judge performance. (For the 2010 tournament, we truncated so that there are only 48 teams in each round; we did this by dropping the worst-ranked teams in each round.)

Tables 2, 3 and 4 (pages 10, 11 and 12) report means and percentiles for these summary statistics, grouped by judge class. Against each statistic, we report the statistic from the actual data, and the corresponding percentile drawn from the Empirical CDF of the simulated data. Together, we interpret these tables as showing that the model has a reasonably good in-sample fit.

2.4 Out-of-sample structural predictions

We use the same method to make out-of-sample predictions for the forthcoming 2013 Championships, where the structure of team matches is the actual structure to be used in the tournament. We provide summary statistics in Table 5 (page 13). We show the Empirical CDF for each simulated summary statistic in Figures 1 to 18 (pages 14 to 31); these CDFs will be used to calculate the percentiles for the actual data observed in the tournament.
### Table 2: Structural predictions: 2010 (in-sample)

<table>
<thead>
<tr>
<th>STATISTIC (PREDICTED PROBABILITY)</th>
<th>PREDICTED</th>
<th>ACTUAL DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10th perc.</td>
<td>25th perc.</td>
</tr>
<tr>
<td>Class 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>6.3%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.1%</td>
<td>7.2%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>72.4%</td>
<td>74.5%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>60.0%</td>
<td>68.8%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>72.4%</td>
<td>74.6%</td>
</tr>
<tr>
<td>Class 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>5.7%</td>
<td>6.8%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.0%</td>
<td>7.1%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>70.8%</td>
<td>72.9%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>60.0%</td>
<td>66.7%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>70.7%</td>
<td>72.5%</td>
</tr>
<tr>
<td>Class 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>12.0%</td>
<td>13.5%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>4.2%</td>
<td>7.7%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>12.0%</td>
<td>13.5%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>67.7%</td>
<td>69.8%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>65.4%</td>
<td>70.8%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>66.7%</td>
<td>68.7%</td>
</tr>
</tbody>
</table>

This table reports in-sample predictions and actual data from the 2010 World Schools Debating Championships. Predictions are formed from 1000 simulations, using the structural estimates from Table 1. We use the 2010 tournament structure, truncated so that there are only 48 teams in each round. (We truncate by dropping debates involving the worst-ranked teams.) ‘Predicted’ values report the 10th, 25th, 75th and 90th percentiles of the Empirical CDF of the 1000 simulated tournaments, along with the mean. ‘Actual data’ reports the actual statistic, and the corresponding percentile value from the simulated Empirical CDF. ‘Judge dissents’ reports the unconditional probability of a judge dissenting. ‘Judge dissents if just dissented’ reports the probability of a judge dissenting, conditional on having just dissented in the previous round. ‘Judge dissents if not just dissented’ reports the probability of a judge dissenting, conditional on not having just dissented in the previous round. The same structure applies for statistics reporting ‘votes for the favorite’.
Table 3: **Structural predictions: 2011 (in-sample)**

<table>
<thead>
<tr>
<th>STATISTIC (PREDICTED PROBABILITY)</th>
<th>PREDICTED</th>
<th>ACTUAL DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10th perc.</td>
<td>25th perc.</td>
</tr>
<tr>
<td>Class 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>6.3%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.6%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>71.9%</td>
<td>73.4%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>60.0%</td>
<td>66.7%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>71.8%</td>
<td>73.6%</td>
</tr>
<tr>
<td>Class 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>5.7%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.0%</td>
<td>7.2%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>69.8%</td>
<td>71.4%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>58.6%</td>
<td>66.7%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>69.5%</td>
<td>71.5%</td>
</tr>
<tr>
<td>Class 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>12.5%</td>
<td>13.5%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>4.8%</td>
<td>8.7%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>12.2%</td>
<td>14.0%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>66.7%</td>
<td>68.8%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>64.9%</td>
<td>70.0%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>65.7%</td>
<td>67.7%</td>
</tr>
</tbody>
</table>

This table reports in-sample predictions and actual data from the 2011 World Schools Debating Championships. Predictions are formed from 1000 simulations, using the structural estimates from Table 1. We use the 2011 tournament structure, which had 48 teams in each round. ‘Predicted’ values report the 10th, 25th, 75th and 90th percentiles of the Empirical CDF of the 1000 simulated tournaments, along with the mean. ‘Actual data’ reports the actual statistic, and the corresponding percentile value from the simulated Empirical CDF. ‘Judge dissents’ reports the unconditional probability of a judge dissenting. ‘Judge dissents if just dissented’ reports the probability of a judge dissenting, conditional on having just dissented in the previous round. ‘Judge dissents if not just dissented’ reports the probability of a judge dissenting, conditional on not having just dissented in the previous round. The same structure applies for statistics reporting ‘votes for the favorite’.
Table 4: **Structural predictions: 2012 (in-sample)**

<table>
<thead>
<tr>
<th>STATISTIC (PREDICTED PROBABILITY)</th>
<th>PREDICTED</th>
<th>ACTUAL DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10th perc.</td>
<td>25th perc.</td>
</tr>
<tr>
<td>Class 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>6.3%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.1%</td>
<td>7.7%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>72.4%</td>
<td>74.0%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>60.0%</td>
<td>68.4%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>71.9%</td>
<td>73.9%</td>
</tr>
<tr>
<td>Class 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>6.3%</td>
<td>7.3%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>6.0%</td>
<td>7.2%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>69.8%</td>
<td>71.9%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>58.3%</td>
<td>66.7%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>69.5%</td>
<td>71.8%</td>
</tr>
<tr>
<td>Class 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>judge dissents</td>
<td>12.5%</td>
<td>13.5%</td>
</tr>
<tr>
<td>judge dissents if just dissented</td>
<td>4.5%</td>
<td>8.3%</td>
</tr>
<tr>
<td>judge dissents if not just dissented</td>
<td>12.4%</td>
<td>13.9%</td>
</tr>
<tr>
<td>judge votes for the favorite</td>
<td>66.7%</td>
<td>68.8%</td>
</tr>
<tr>
<td>judge votes for the favorite if just dissented</td>
<td>64.0%</td>
<td>69.6%</td>
</tr>
<tr>
<td>judge votes for the favorite if not just dissented</td>
<td>65.7%</td>
<td>67.7%</td>
</tr>
</tbody>
</table>

This table reports in-sample predictions and actual data from the 2012 World Schools Debating Championships. Predictions are formed from 1000 simulations, using the structural estimates from Table 1. We use the 2012 tournament structure, which had 48 teams in each round. ‘Predicted’ values report the 10th, 25th, 75th and 90th percentiles of the Empirical CDF of the 1000 simulated tournaments, along with the mean. ‘Actual data’ reports the actual statistic, and the corresponding percentile value from the simulated Empirical CDF. ‘Judge dissents’ reports the unconditional probability of a judge dissenting. ‘Judge dissents if just dissented’ reports the probability of a judge dissenting, conditional on having just dissented in the previous round. ‘Judge dissents if not just dissented’ reports the probability of a judge dissenting, conditional on not having just dissented in the previous round. The same structure applies for statistics reporting ‘votes for the favorite’.
This table reports out-of-sample predictions for the 2013 World Schools Debating Championships. Predictions are formed from 1000 simulations, using the structural estimates from Table 1. We use the 2013 tournament structure, which will feature 48 teams in each round. ‘Predicted’ values report the 10th, 25th, 75th and 90th percentiles of the Empirical CDF of the 1000 simulated tournaments, along with the mean. ‘Actual data’ will be used after the tournament to report the actual statistic, and the corresponding percentile value from the simulated Empirical CDF. ‘Judge dissents’ reports the unconditional probability of a judge dissenting. ‘Judge dissents if just dissented’ reports the probability of a judge dissenting, conditional on having just dissented in the previous round. ‘Judge dissents if not just dissented’ reports the probability of a judge dissenting, conditional on not having just dissented in the previous round. The same structure applies for statistics reporting ‘votes for the favorite’.
Empirical CDF: Predicted probability that a judge dissents (Class 1)
Figure 2: Empirical CDF: Predicted probability that a judge dissents, conditional on having just dissented (Class 1)
Figure 3: Empirical CDF: Predicted probability that a judge dissents, conditional on not having just dissented (Class 1)
Hypothesis Registration: The 2013 World Schools Debating Championships

Figure 4: Empirical CDF: Predicted probability that a judge votes for the favorite (Class 1)
Figure 5: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on having just dissented (Class 1)
Hypothesis Registration: The 2013 World Schools Debating Championships

Figure 6: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on not having just dissented

(Class 1)
Figure 7: Empirical CDF: Predicted probability that a judge dissents (Class 2)
Hypothesis Registration: The 2013 World Schools Debating Championships

Figure 8: Empirical CDF: Predicted probability that a judge dissents, conditional on having just dissented (Class 2)
Figure 9: Empirical CDF: Predicted probability that a judge dissents, conditional on not having just dissented (Class 2)
Figure 10: Empirical CDF: Predicted probability that a judge votes for the favorite (Class 2)
Figure 11: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on having just dissented (Class 2)
Figure 12: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on not having just dissented (Class 2)
Figure 13: Empirical CDF: Predicted probability that a judge dissents (Class 3)
Figure 14: Empirical CDF: Predicted probability that a judge dissents, conditional on having just dissented (Class 3)
Figure 15: Empirical CDF: Predicted probability that a judge dissents, conditional on not having just dissented (Class 3)
Figure 16: Empirical CDF: Predicted probability that a judge votes for the favorite (Class 3)
Figure 17: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on having just dissented (Class 3)
Hypothesis Registration: The 2013 World Schools Debating Championships

Figure 18: Empirical CDF: Predicted probability that a judge votes for the favorite, conditional on not having just dissented (Class 3)