Measuring the Political Sophistication of Voters in the Netherlands and the United States

Christopher N. Lawrence

Department of Political Science
Saint Louis University

November 2006

Overview

- What is political sophistication?
- How should we measure political sophistication?
- If we use survey questions, what questions should we use?

What is political sophistication?

- Bob Luskin: "the extent to which [a person's personal belief system] is large, wide-ranging, and highly constrained."
- Me (perhaps following Zaller and Krosnick): *the capacity of citizens to understand, process, and utilize new political information*.
- Commonly conflated with *political knowledge*—although I would argue that these are distinct concepts.
- Also known as *political expertise*.

A classic quote

Under various guises, expertise and/or knowledge have long been a concern of political scientists.

"The democratic citizen is expected to be well informed about political affairs. He is supposed to know what the issues are, what their history is, what the relevant facts are, what alternatives are proposed, what the party stands for, what the likely consequences are. By such standards the voter falls short."

Berelson, Lazarsfeld, and McPhee, *Voting* (1954: 308)

Measuring political sophistication

Since political scientists first recognized the importance of political sophistication, there has been debate over measurement:

- The levels of conceptualization (*The American Voter*; Converse and Luskin's "active use" measure): do citizens think in ideological terms?
- Ideological constraint (Converse; Jackson and Marcus; "schema theory"): does the voter's personal belief system "hang together," or is it randomly arranged? (nonattitudes?)
- "Recognition and understanding" (Converse; Luskin): do voters recognize and understand ideological labels?
- Differentiation (Luskin; Zaller): can voters make distinctions between party/candidate issue positions?
- Information-holding/knowledge (Delli Carpini and Keeter)
- Interviewer evaluation (ANES)

Comparing differentiation and knowledge

This project looks at the use of both Luskin-style "differentiation" and political knowledge items included in various surveys of the mass public.

To do this, we need to look at how each type of item performs as an indicator of sophistication more broadly. How can we do this?

Getting a score

In a traditional multiple choice test:

$$\text{score} = \sum_{i=1}^{n} c_i$$

In other words, we simply add up the number of correct answers to get the score.

Thus a simple approach to measuring sophistication would be to add up the number of knowledge items that people get right. But this doesn't indicate how good each question is—all it does is give us a score for each respondent.

Item-response theory models

A promising approach to more in-depth analysis of questions comes from the family of *item-response theory* latent variable models.

These models were originally developed for standardized testing in the fields of educational psychology and test development—psychologists refer to these models of underlying (unobserved or latent) ability as *psychometric* models.

IRT models in political science

In political science, IRT models have mostly been used for spatial models of roll-call voting and Supreme Court decision-making; Poole and Rosenthal's NOMINATE is a special case, while "purer" IRT models have been used by Clinton, Jackman, and Rivers (for roll-calls) and Martin and Quinn (for Supreme Court voting).

However, there has been some application to political knowledge and sophistication: Delli Carpini and Keeter (1996) use them in their book on political knowledge, while Levendusky and Jackman had a working paper circa 2003, contemporaneous with my dissertation research, introducing IRT models as well.

The IRT model

As we saw before, in a traditional multiple choice test:

$$\text{score} = \sum_{i=1}^{n} c_i$$

The IRT model allows us to also determine the *difficulty* of each question and the question's *discrimination*—how well the item separates low-scoring and high-scoring respondents from each other.

The scores are called the *abilities* of the respondents.

The IRT model (continued)

In the IRT model, the probability that the observed response to question *i* by respondent *j* is correct is given by

$$z_{ij} = -\alpha_i + \beta_i \theta_j + \epsilon_{ij}$$

where α is the difficulty of the question, β is the discrimination parameter for the question, and θ is the respondent's ability—for our purposes, level of sophistication.

In other words, whether or not a respondent got a particular question right is determined by his or her ability θ_j, the difficulty of the question α_i, and the question's discrimination β_i. Of course, it is also subject to measurement error (ϵ_{ij}).

The functional form

The z_{ij} aren't observed, so we must treat this like a probit:

$$\Pr(c_{ij} = 1|\theta_j) = \Phi(-\alpha_i + \beta_i\theta_j)$$

All of these parameters—α_i, β_i, and θ_j—are unknown. Using traditional approaches like maximum-likelihood estimation, this would be impossible to solve because of the large number of parameters.

Identifying the IRT model

With sufficient identifying conditions—namely, that both α and β are distributed normally, that the respondent abilities θ_j are independent and distributed standard normal, and constraining one of the β_i to be positive—the model is tractable.

The end result gives us estimates of the respondent abilities, which may be useful for second-stage analyses, as well as the difficulties and the discrimination parameters for each item (question). Estimation is readily available using Martin and Quinn's `MCMCpack` for R.

Benefits of IRT

There are a number of key advantages of using IRT models over a naïve summated scale:

- The contribution of each item is adjusted based on its difficulty and ability to discriminate, rather than equal weights being assumed.
- The respondent abilities are true interval variables rather than integer counts, which may be useful in second-stage estimation.
- Random measurement error is accounted for in the model.
- If used with MCMC, missing data are handled gracefully.

Of course, the key disadvantage is that finding a solution to the IRT model is more complex than generating a summated scale!

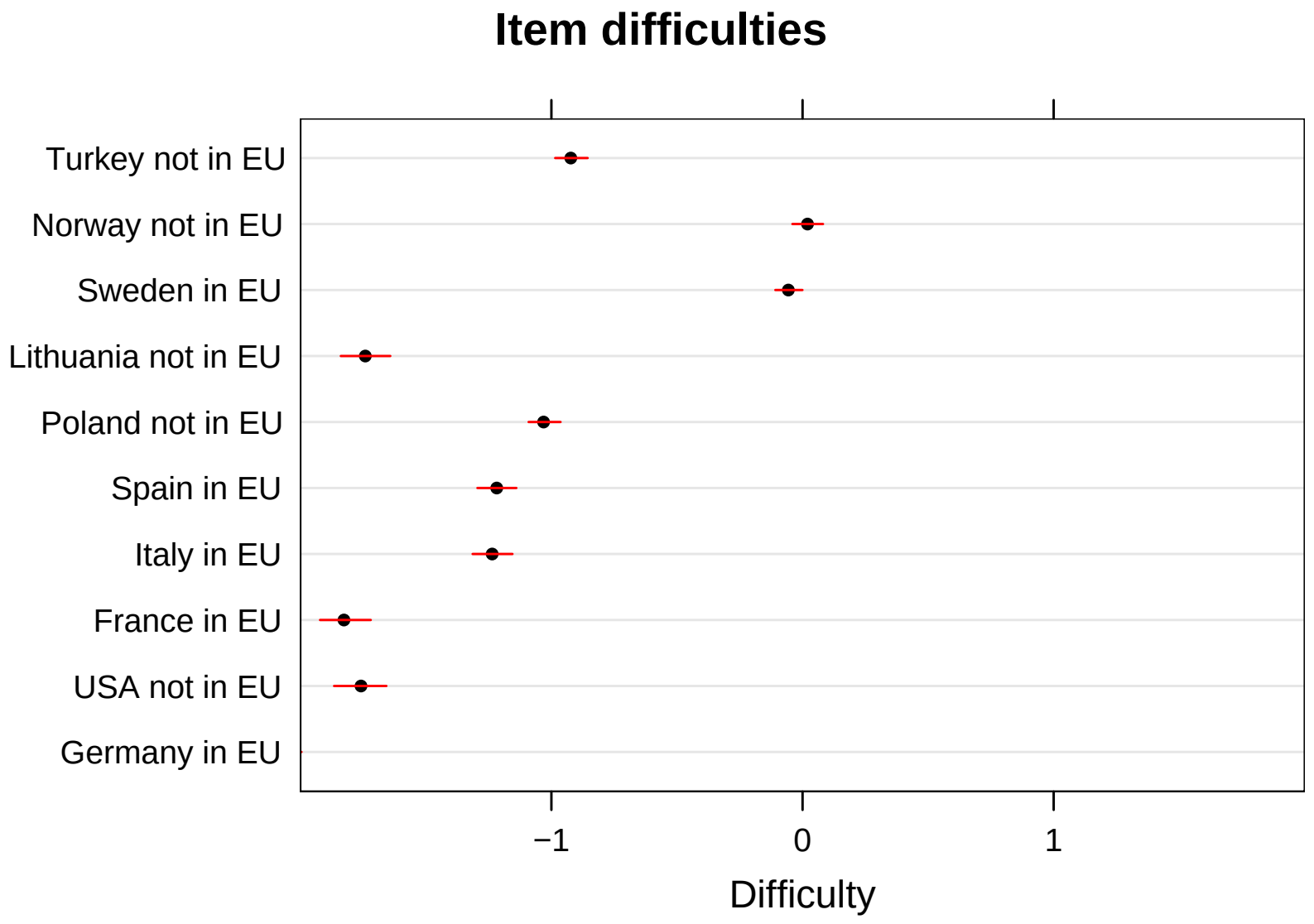
An application: DPES

The 1998 Dutch Parliamentary Election Study (DPES) included a battery of items suitable for this analysis:

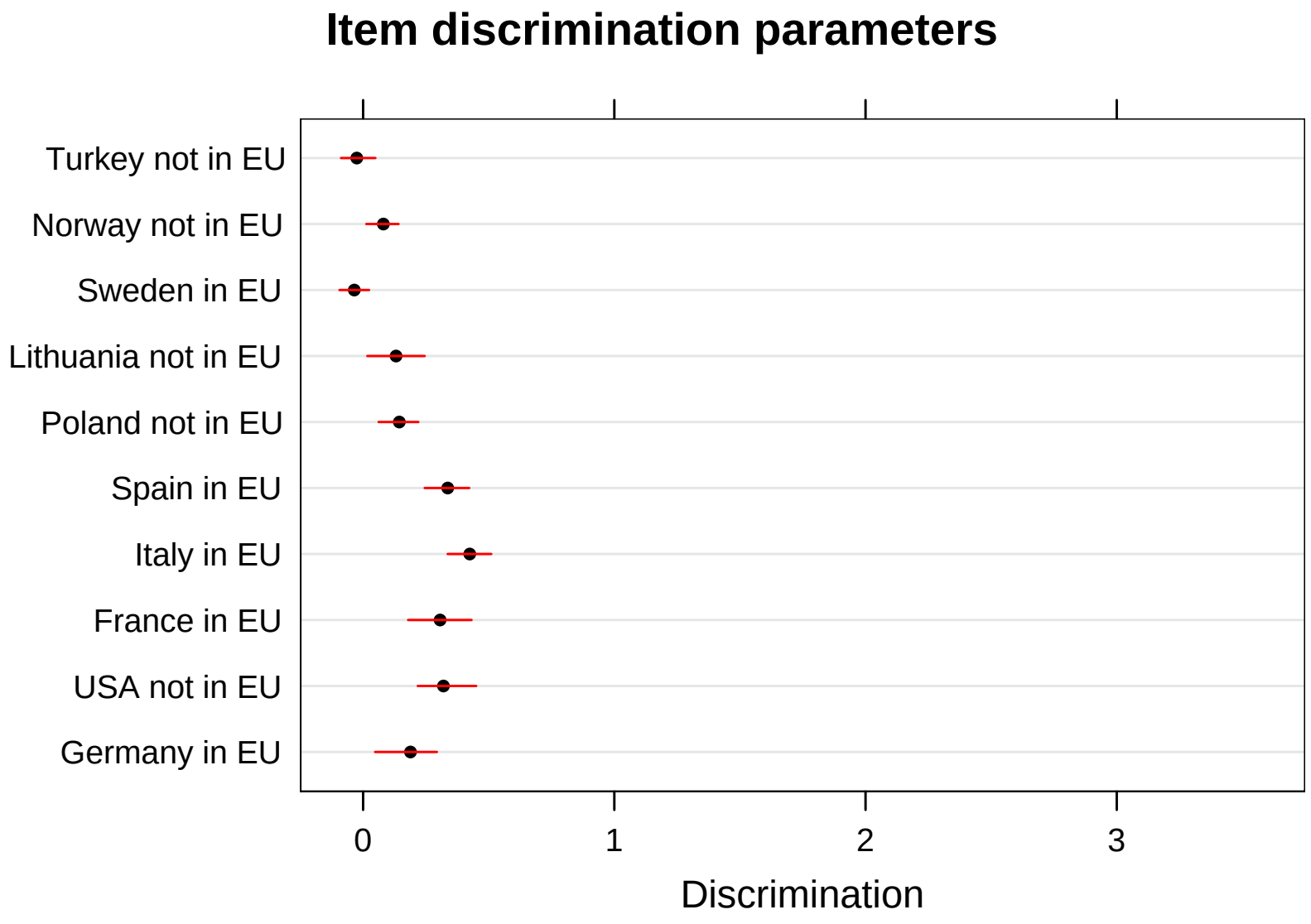
- Knowledge of EU membership status of various nations.
- Knowledge of name, party, and position of four Dutch political figures.
- Knowledge of governing coalition members (and non-members).
- Knowledge of the relative strength of major parties in the Dutch parliament.
- Identification of relative positions of main parties on five major issues. (Differentiation measure.)

The following graphs show the relative performance of items within each of these groups.

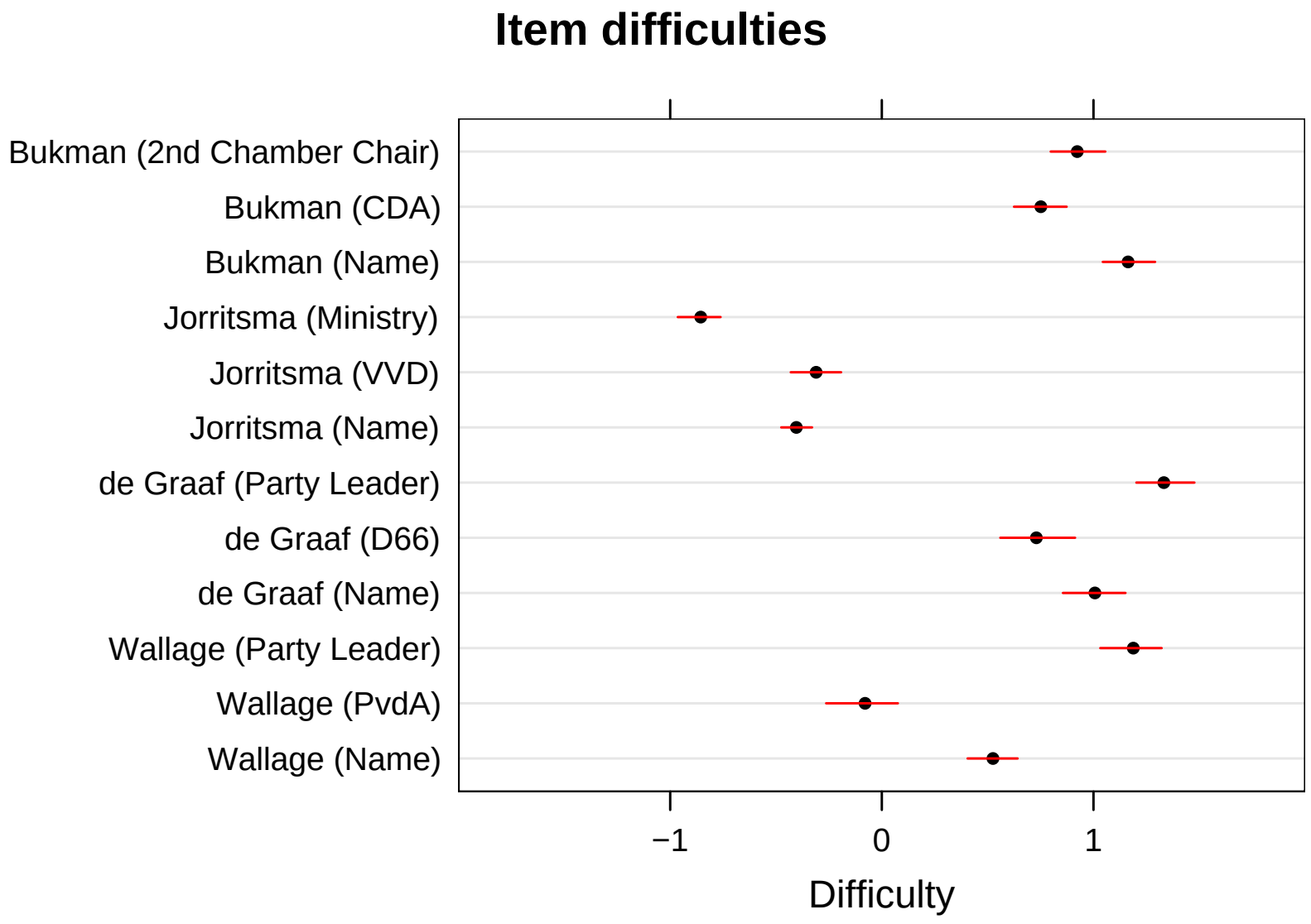
EU membership items



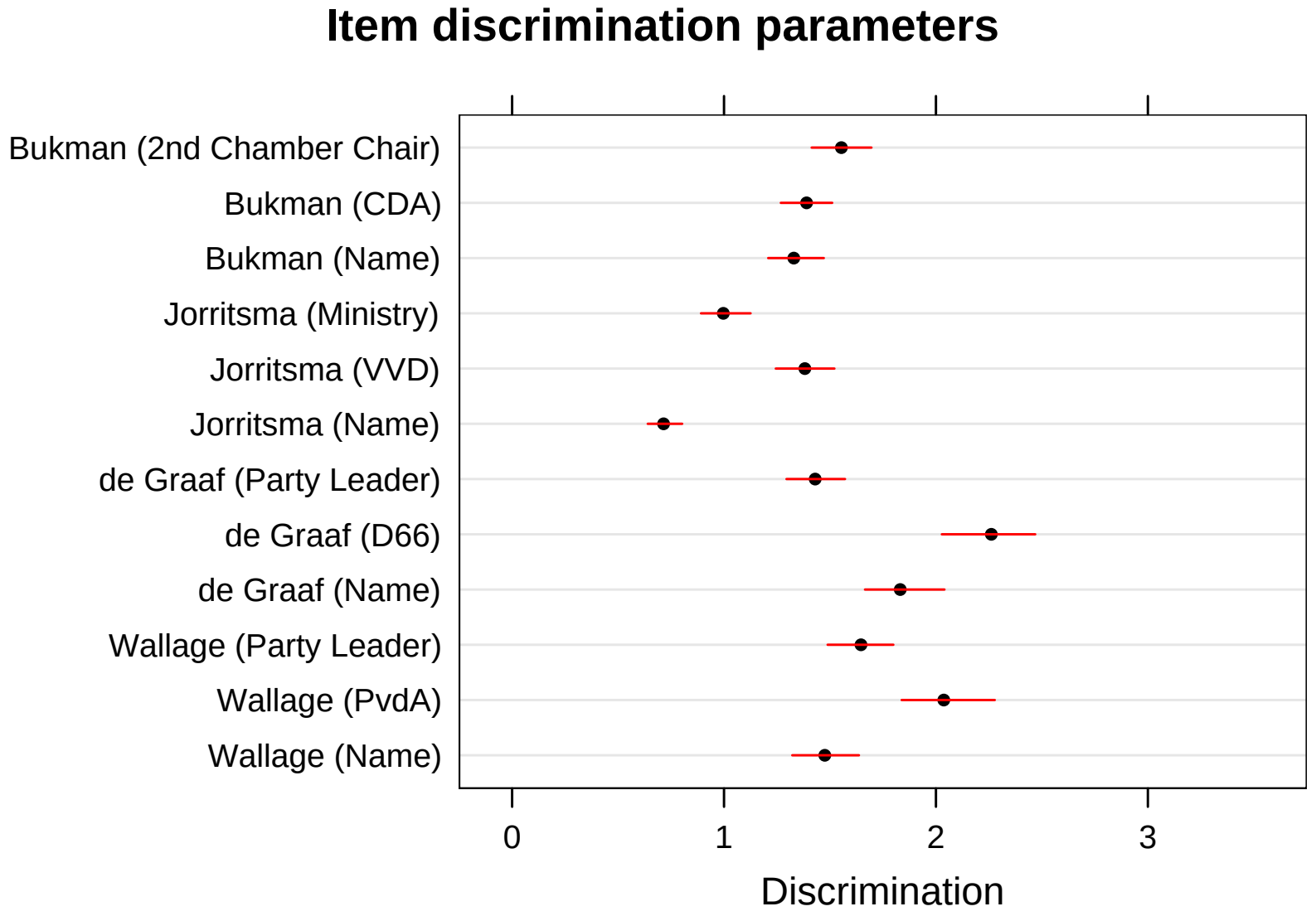
EU membership items



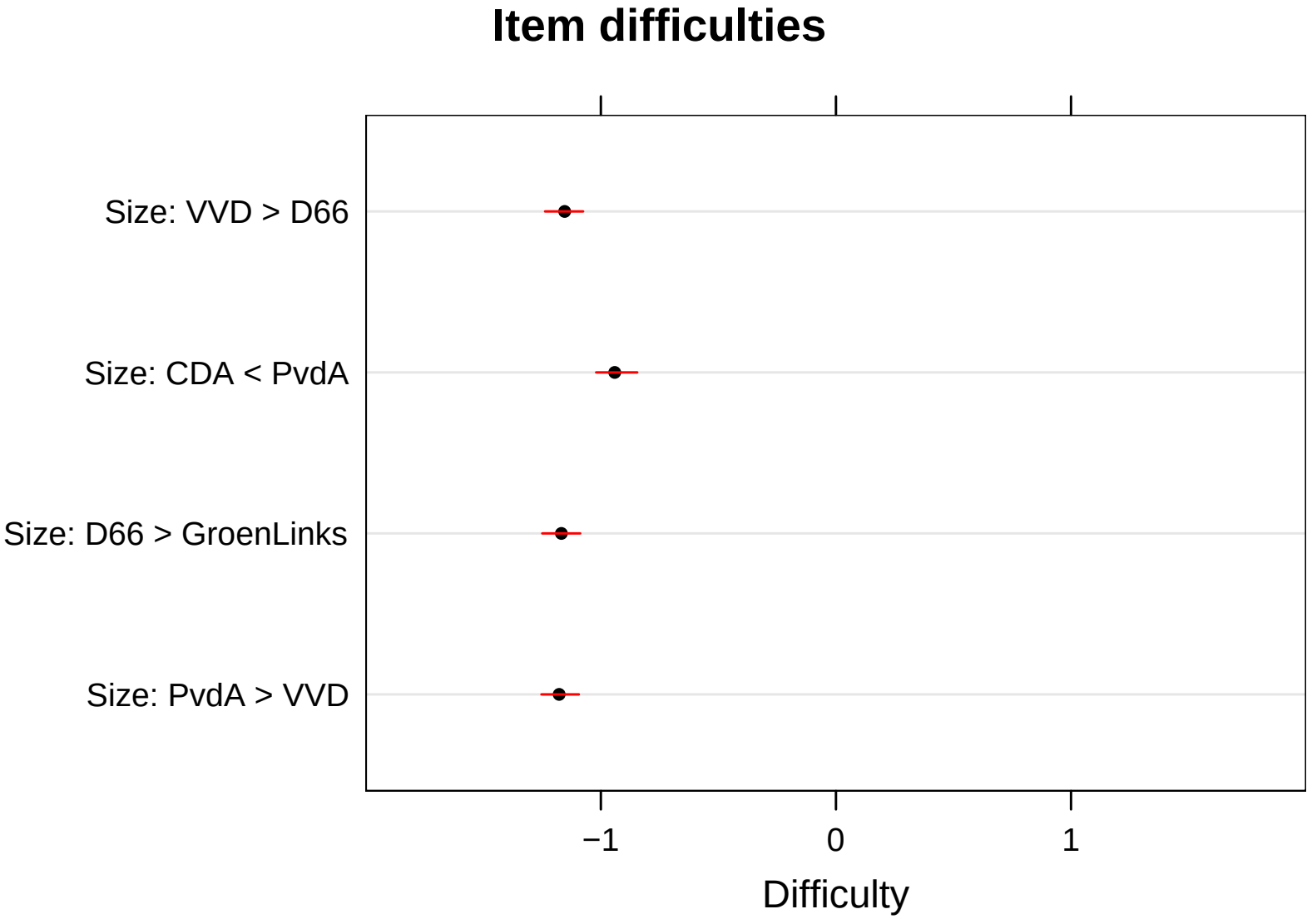
Party leader items



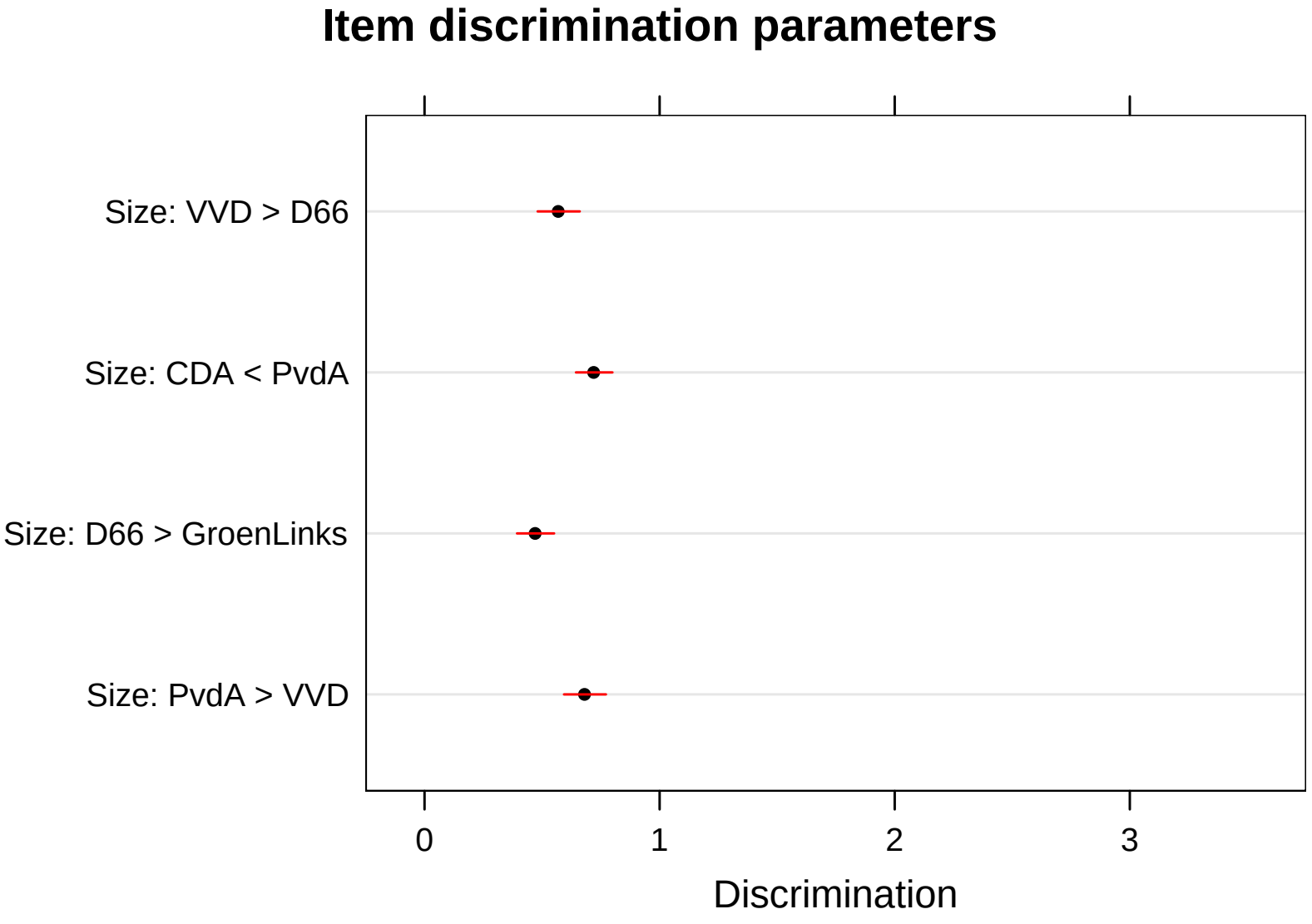
Party leader items



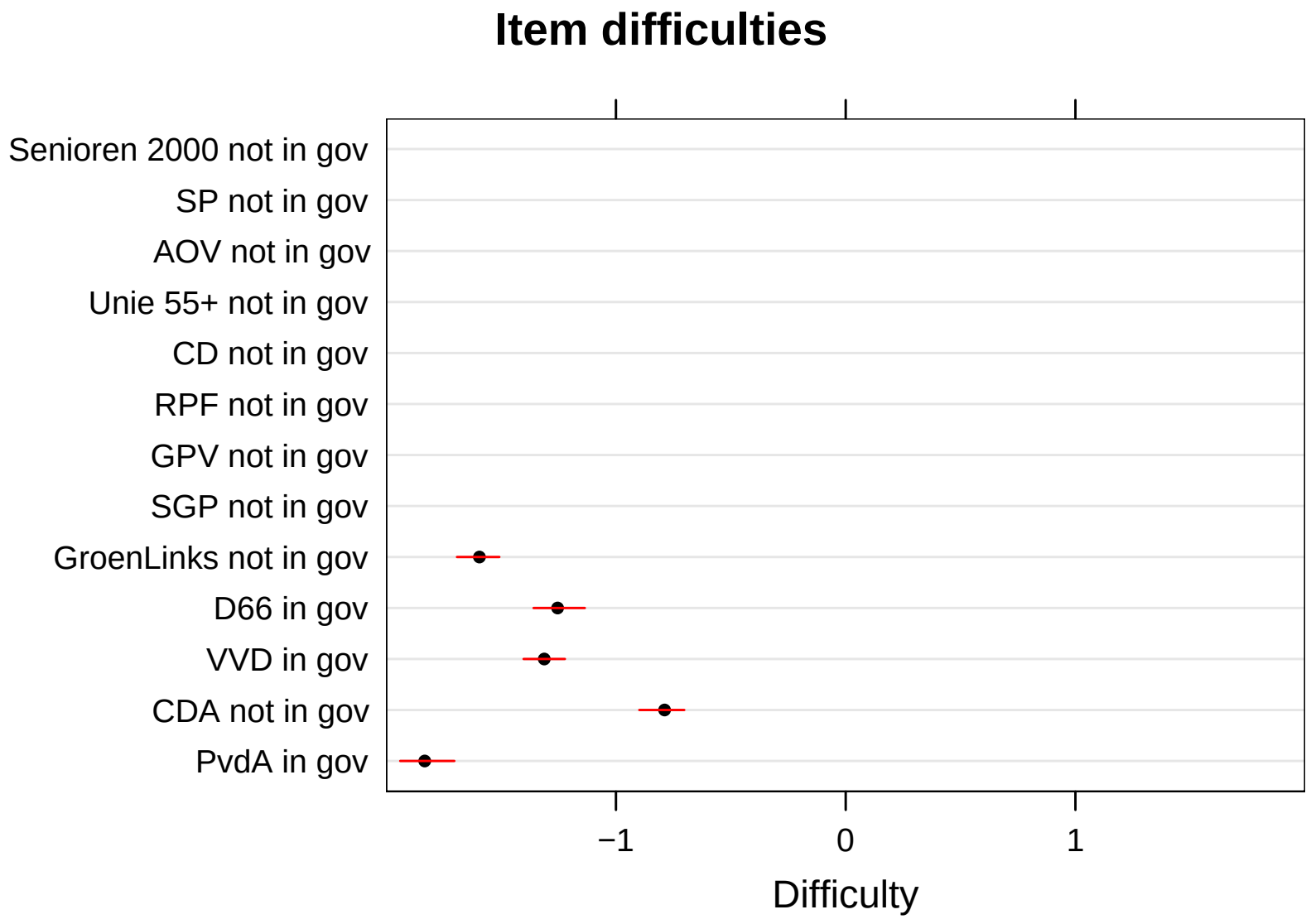
Party size ID items



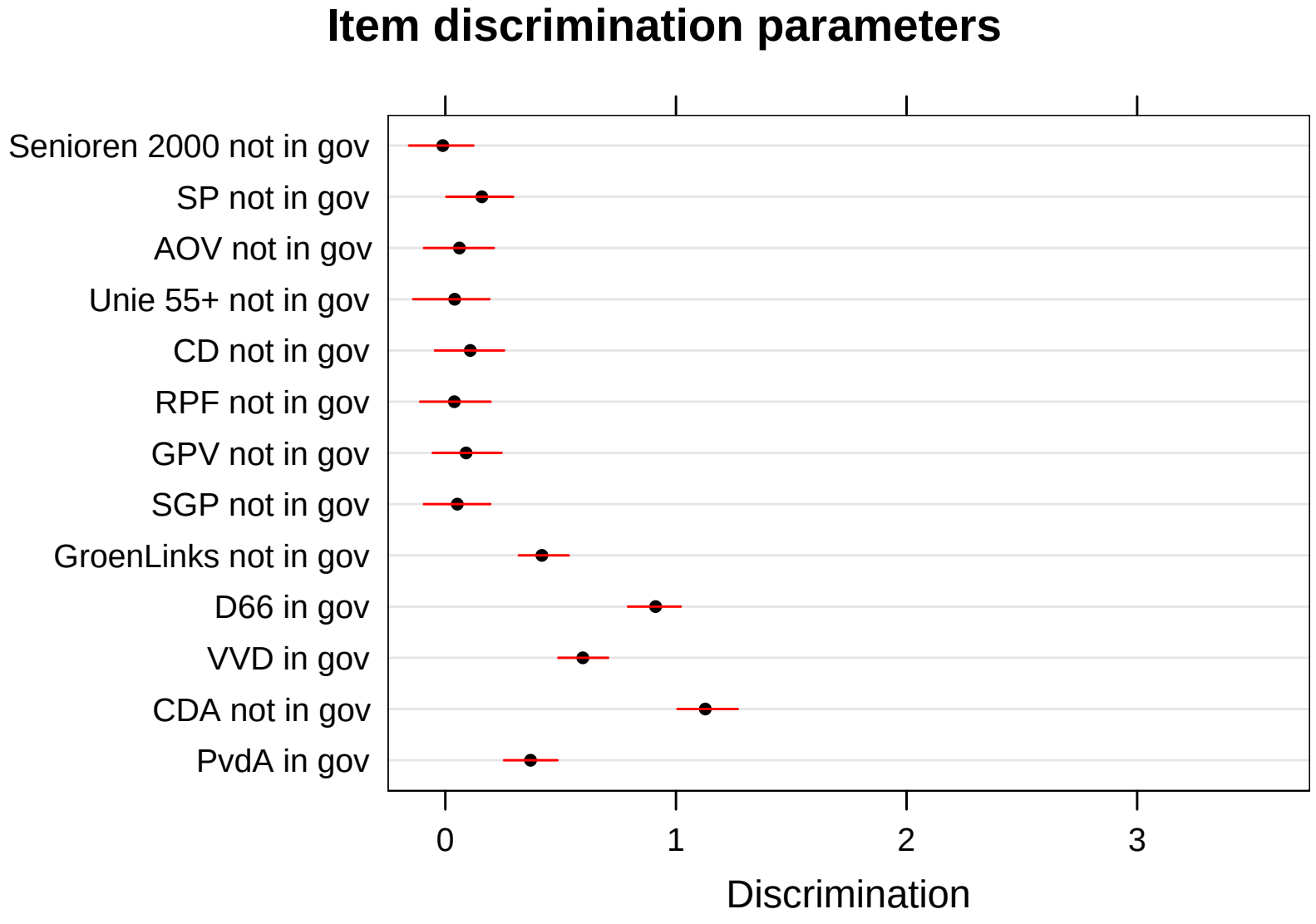
Party size ID items



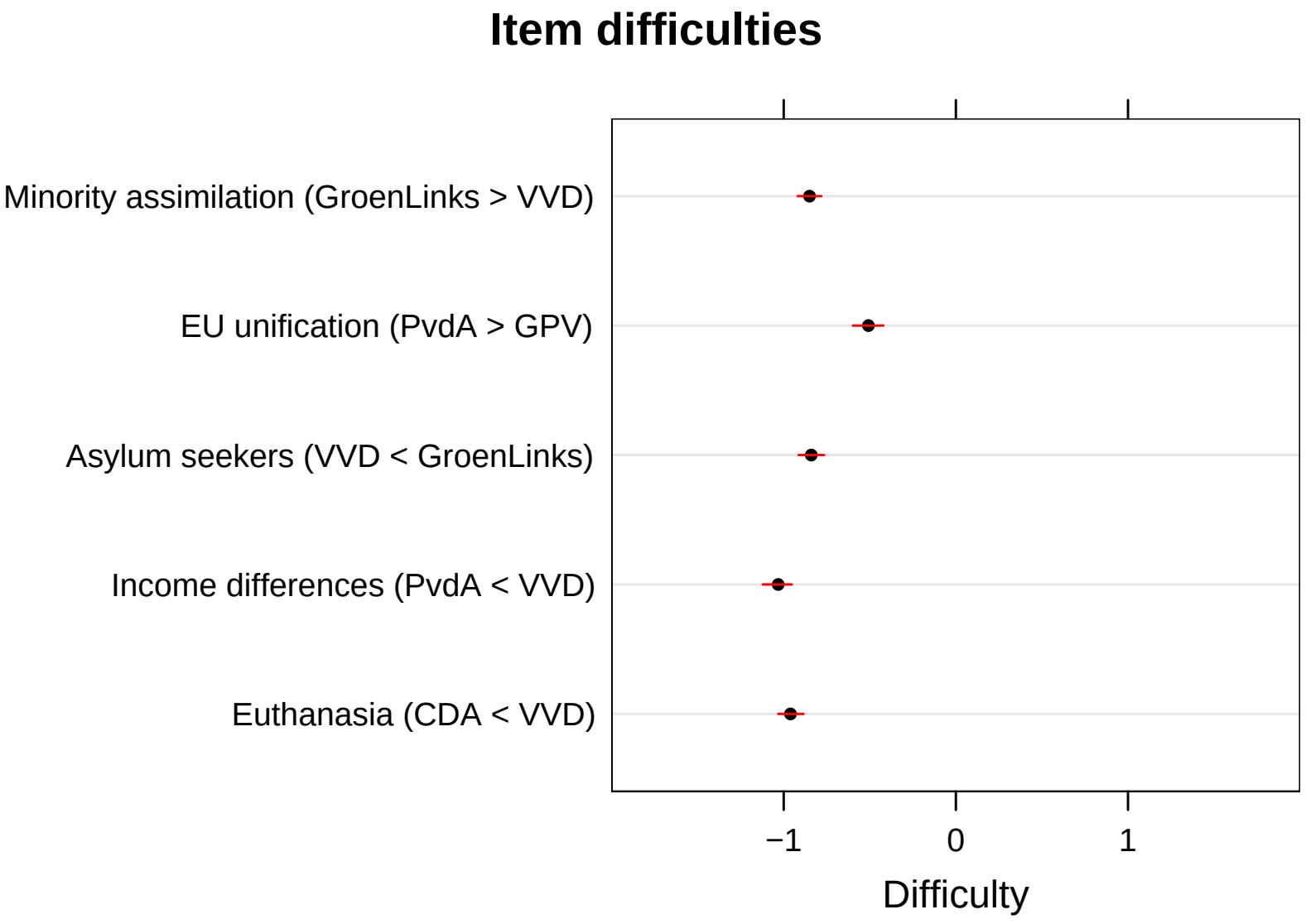
Coalition membership items



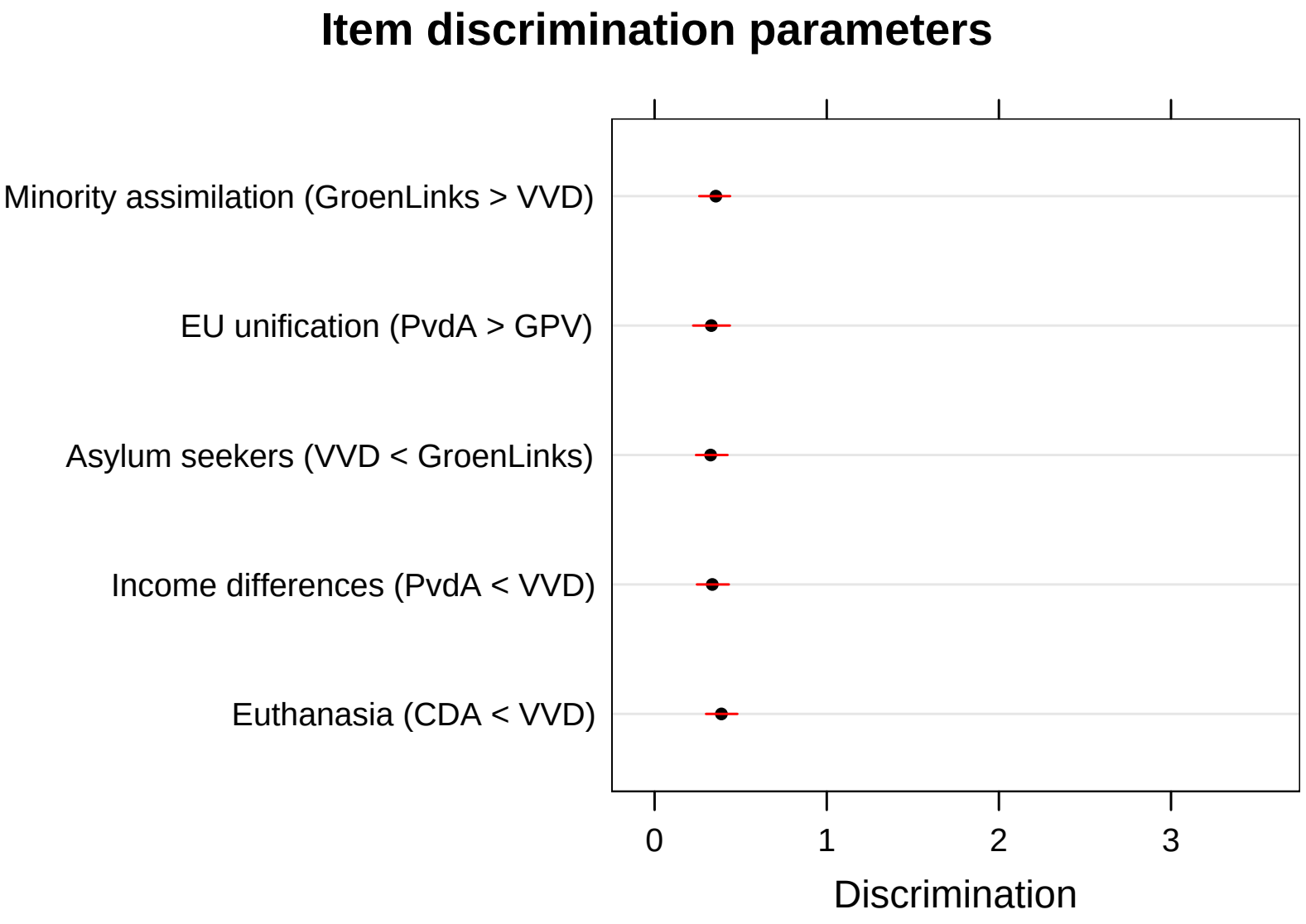
Coalition membership items



Issue placement items



Issue placement items



Validation of the Measure

The respondent abilities were validated against other measures in the model:

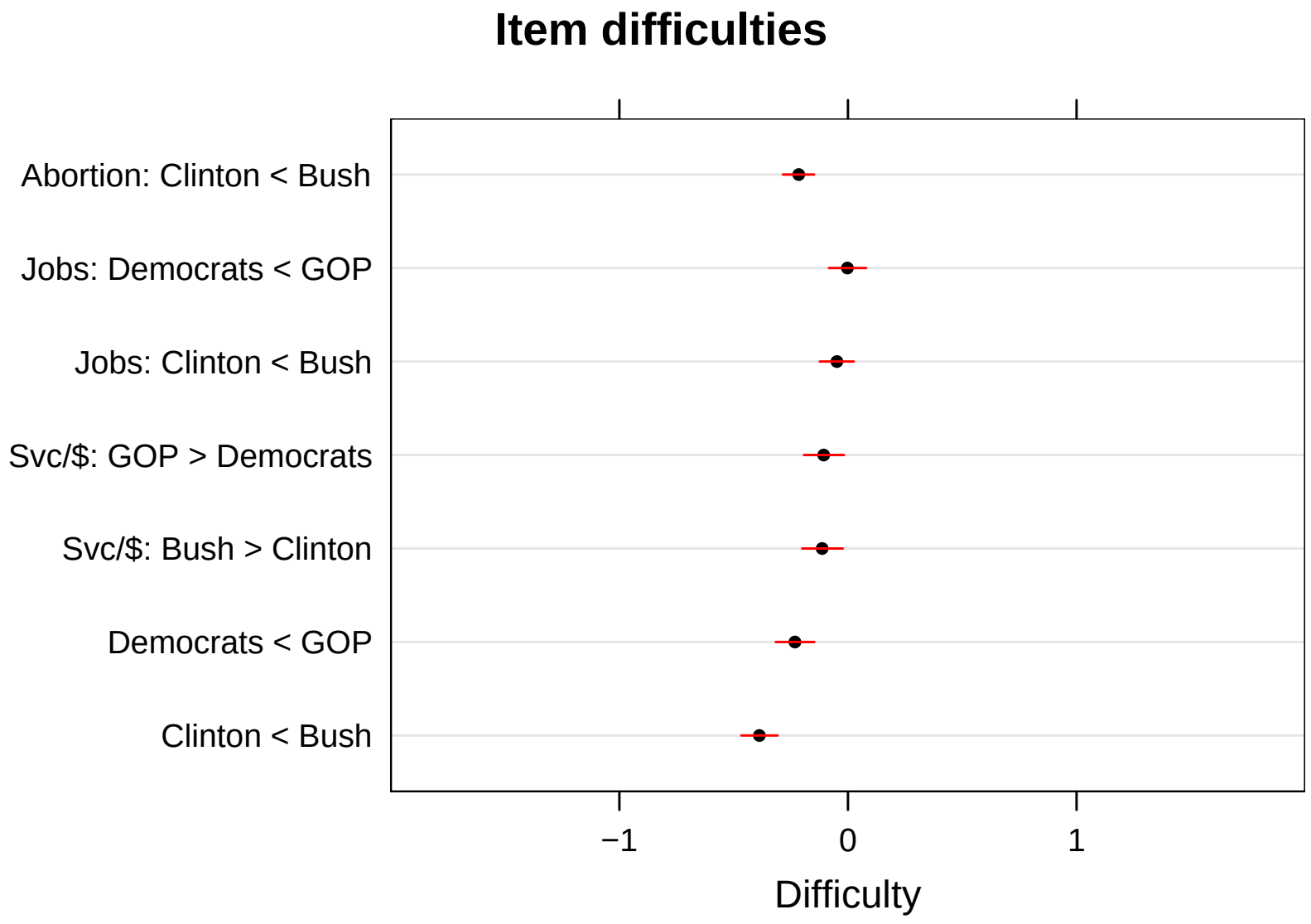
- Correlation with simple knowledge scale based on photo IDs: $r = 0.95$.
- Correlation with knowledge scale based on number of completely correct IDs: $r = 0.85$.
- Correlation with respondent's self-reported political interest scale: $r = 0.47$.
- Correlation with respondent's self-reported civic participation scale: $r = 0.29$.
- Correlation with respondent's level of educational attainment: $r = 0.34$.

Another application: ANES

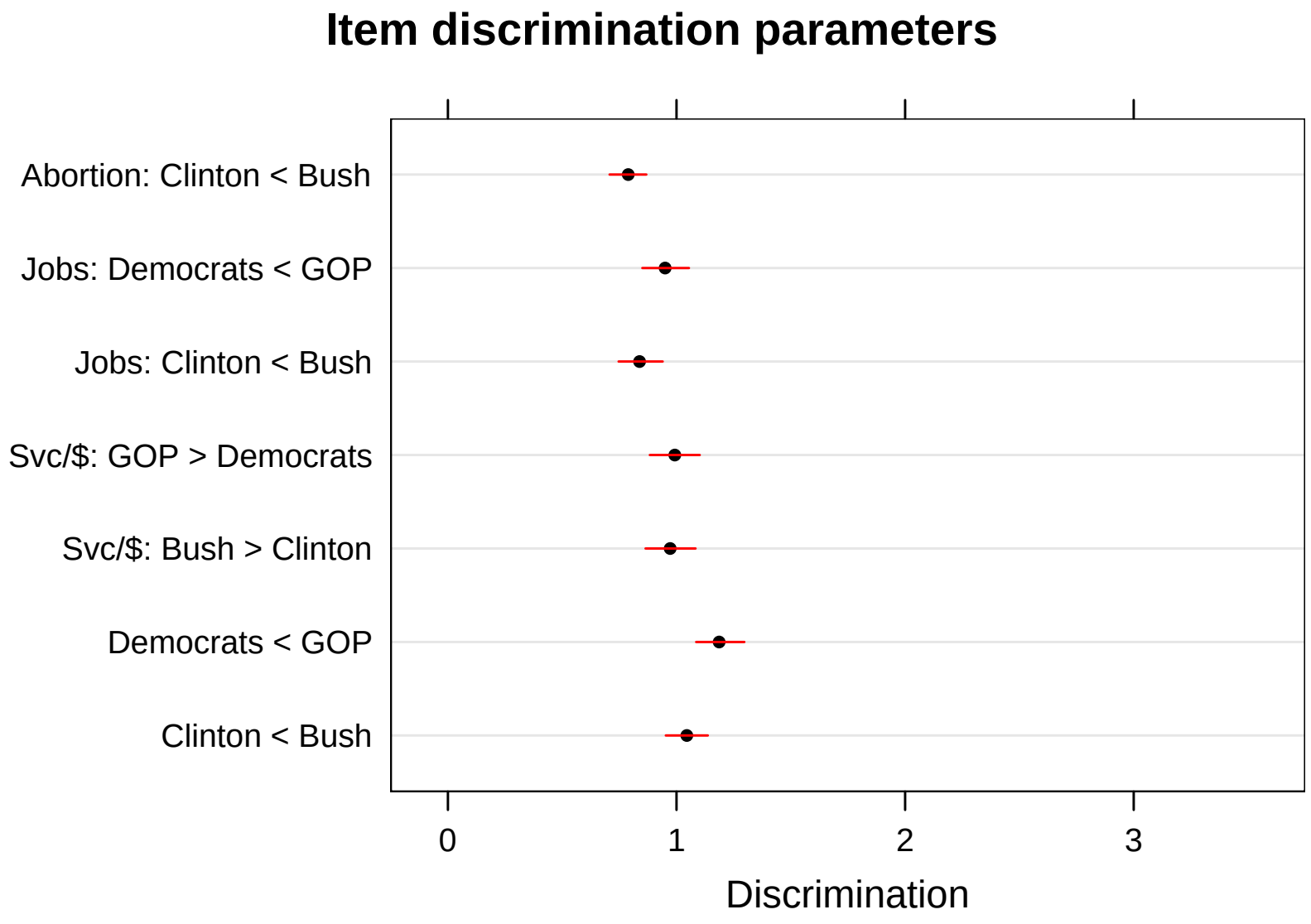
Recent editions of the American National Election Studies also provide a wealth of potential knowledge items:

- Knowledge of key political figures.
- Knowledge of largest party in each chamber of Congress.
- Knowledge of biographical details of presidential and vice-presidential candidates. (2000)
- Placement of parties and candidates on political issues. (Differentiation.)
- Placement of parties and candidates on a liberal-conservative scale. (Differentiation.)

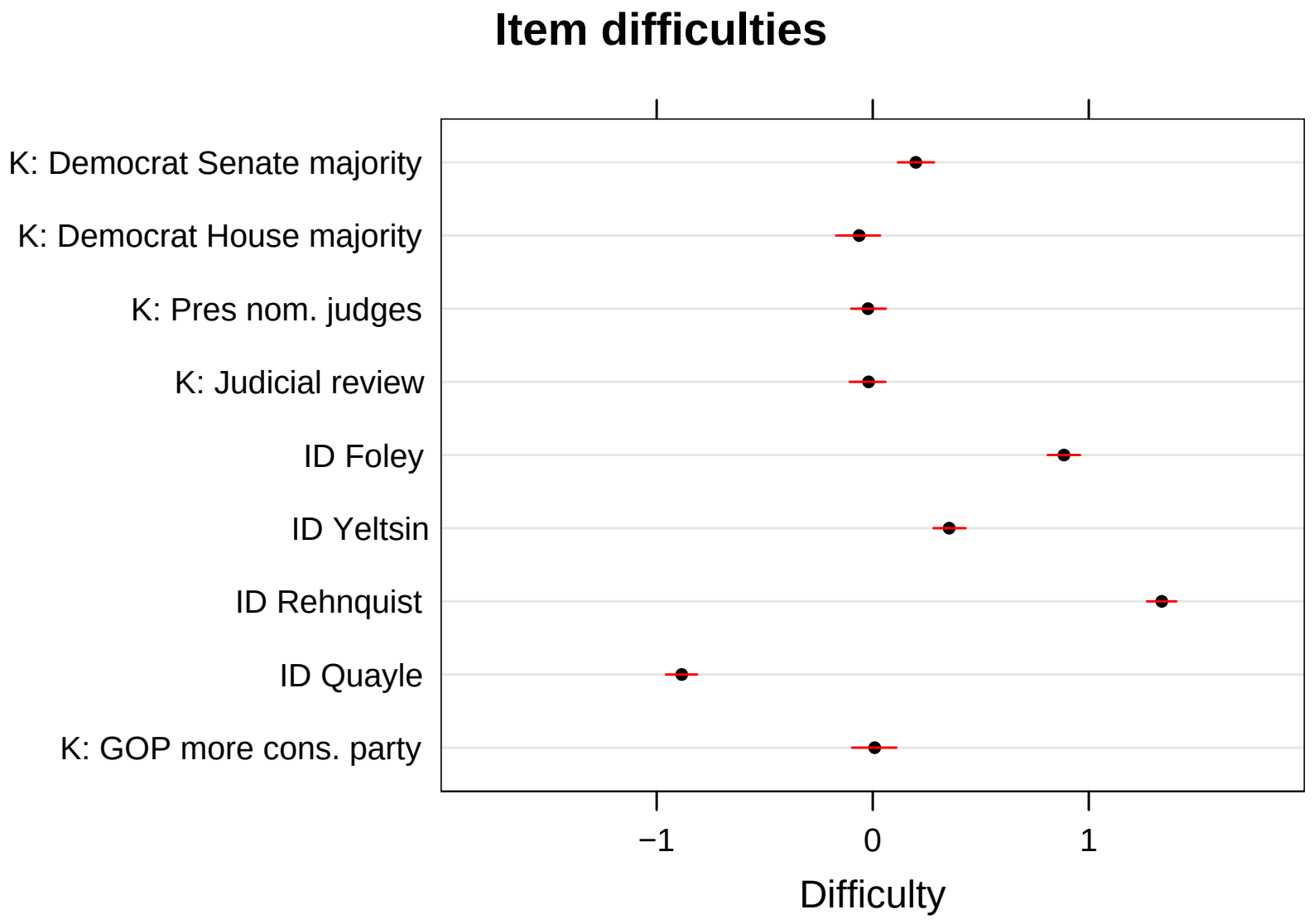
1992 party/candidate placement items



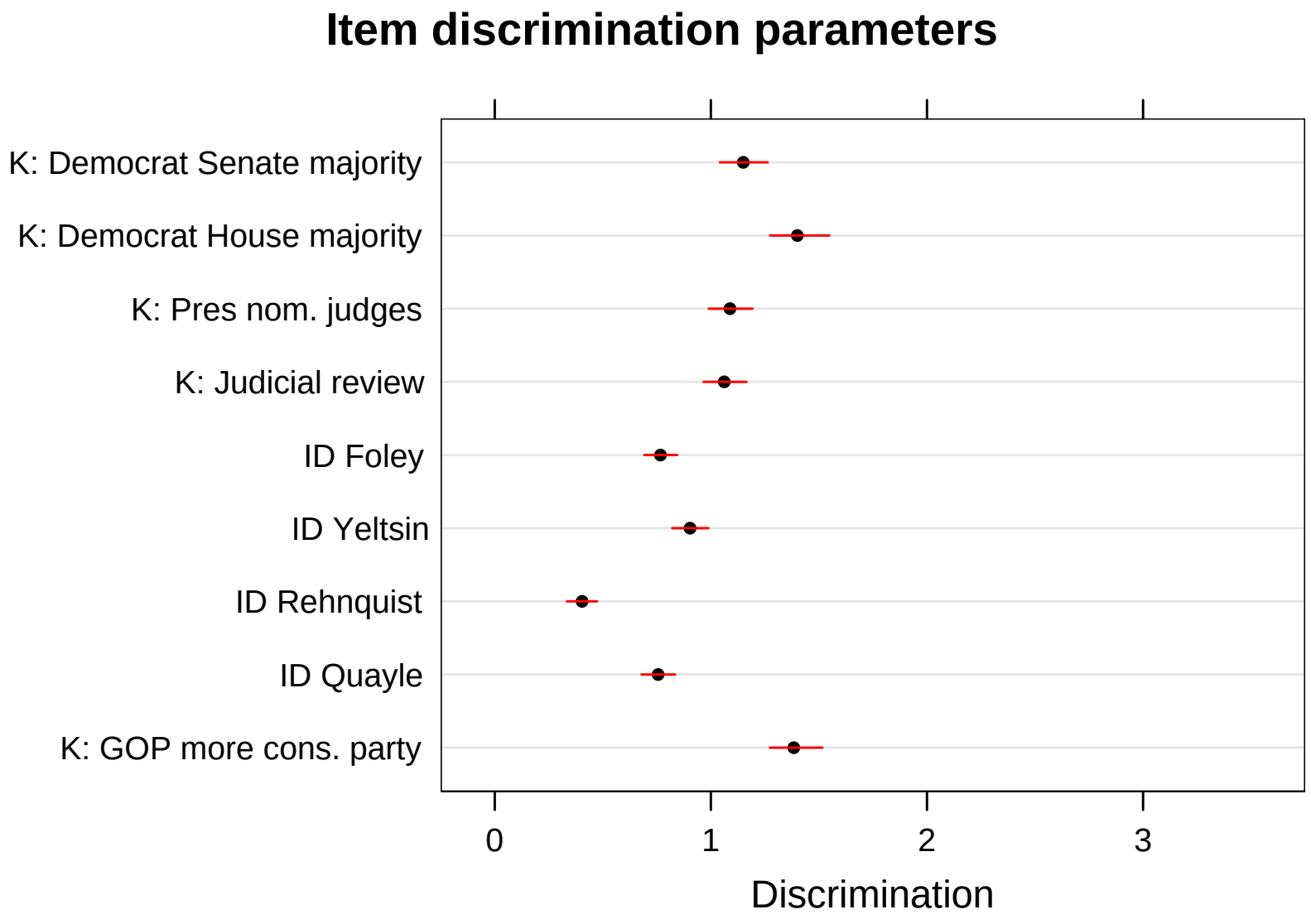
1992 party/candidate placement items



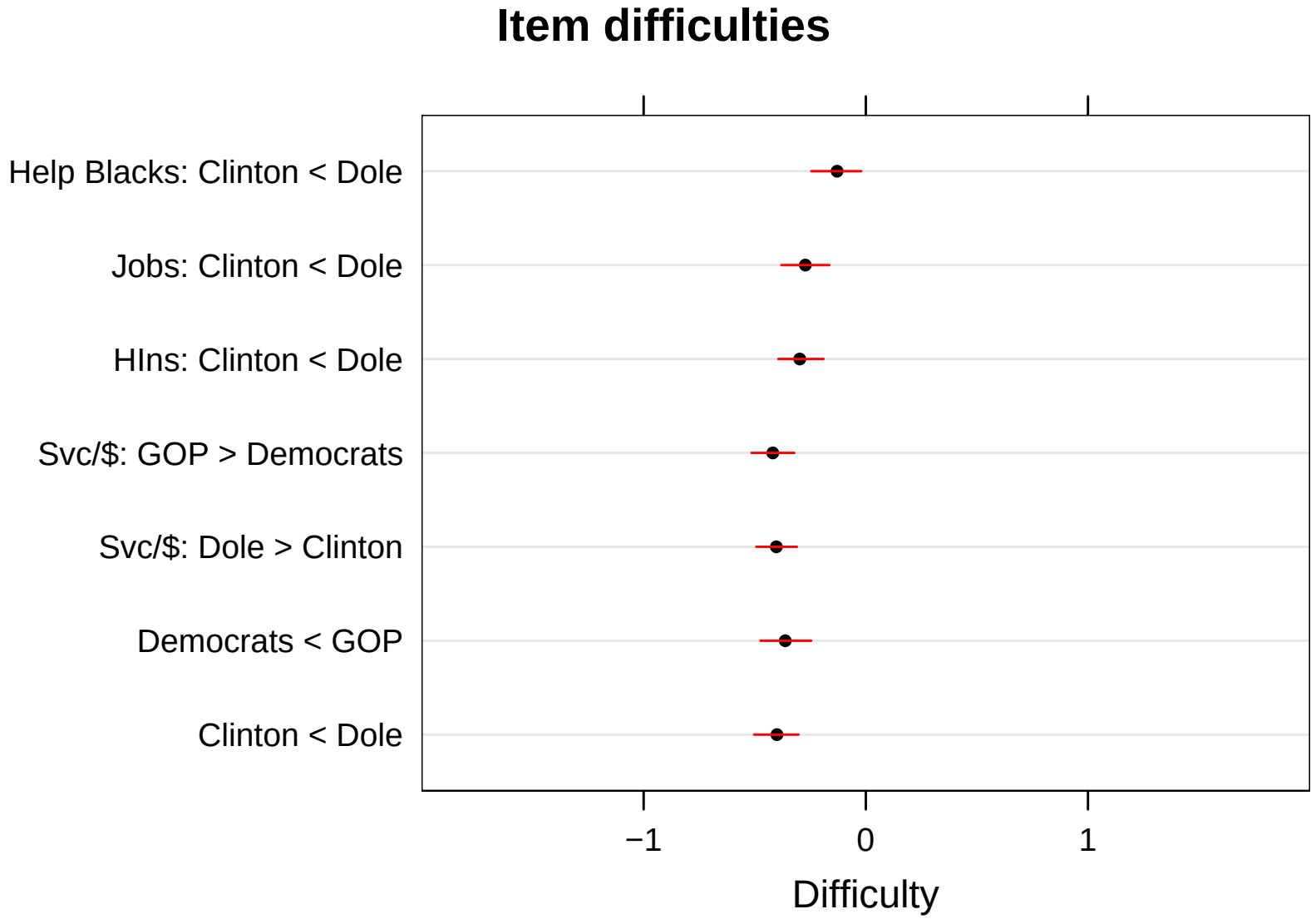
1992 knowledge items



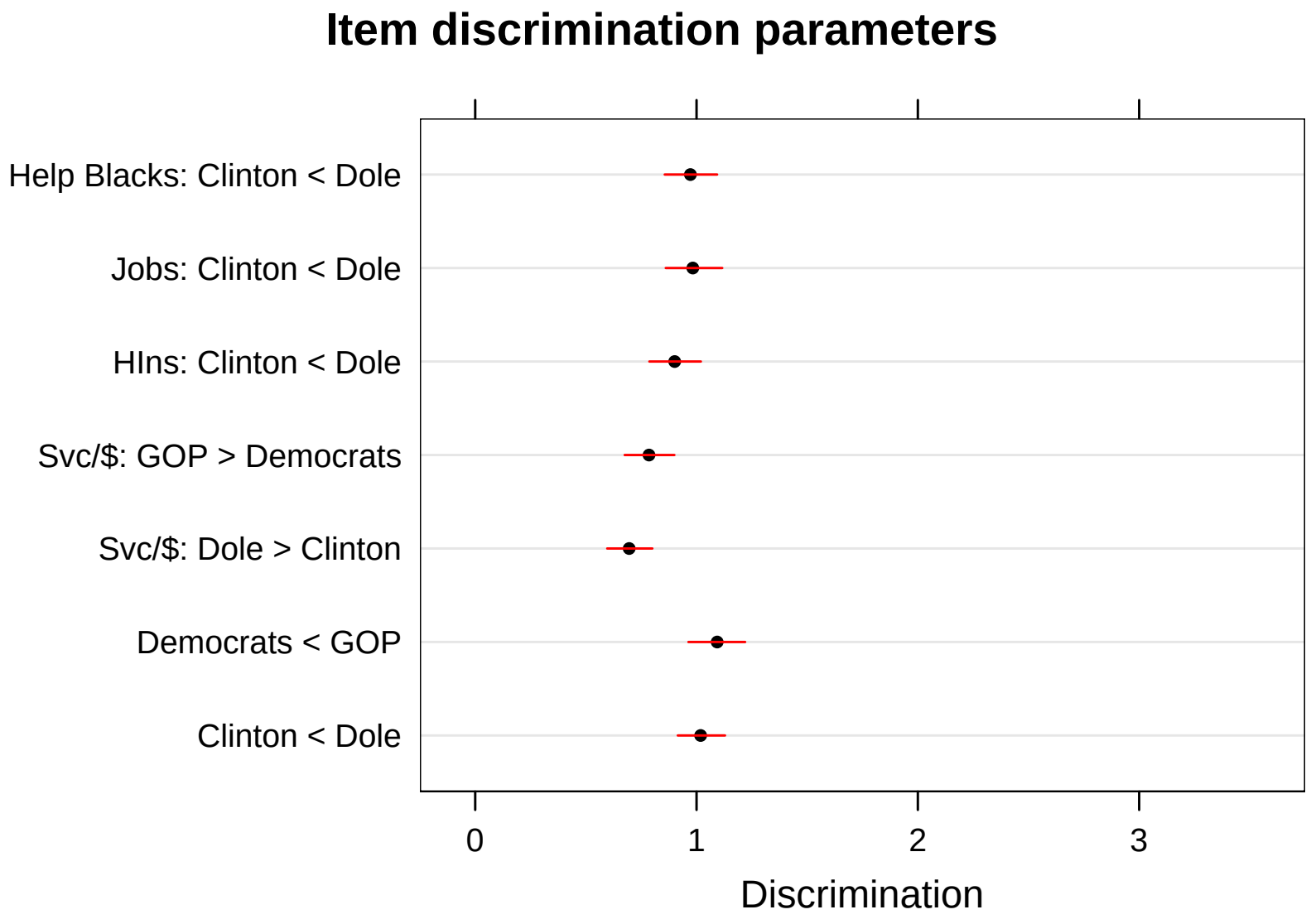
1992 knowledge items



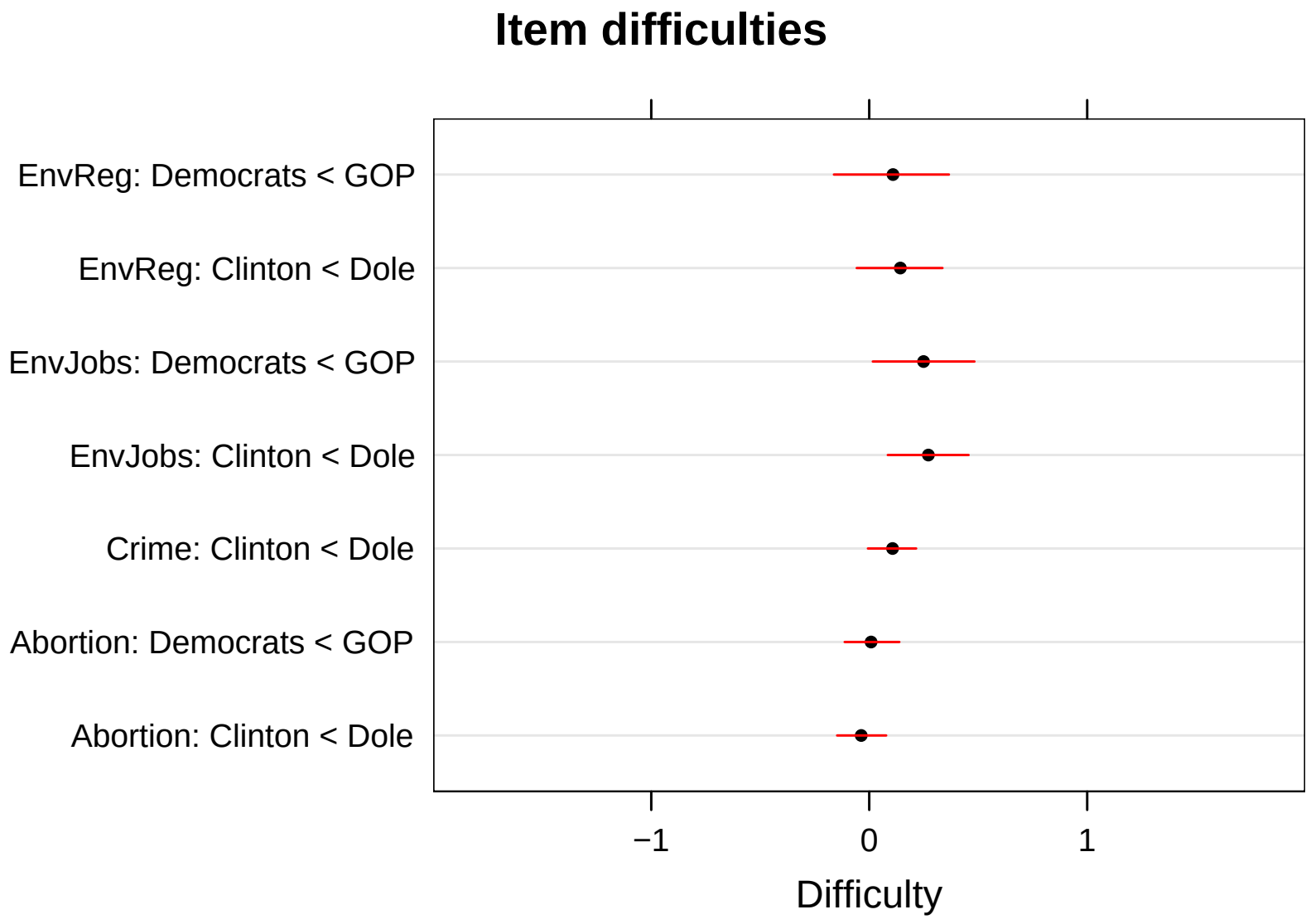
1996 party/candidate placement items (group 1)



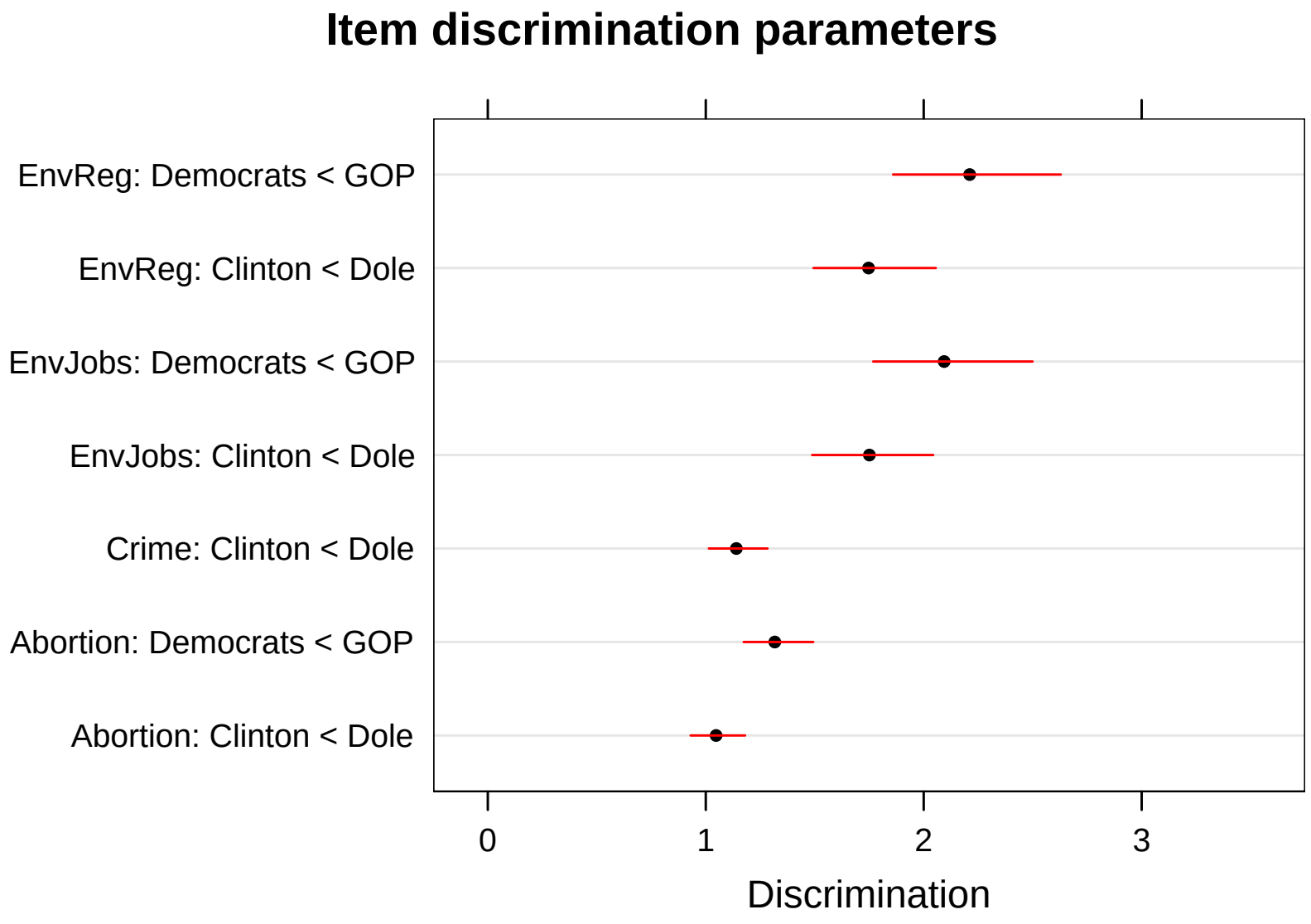
1996 party/candidate placement items (group 1)



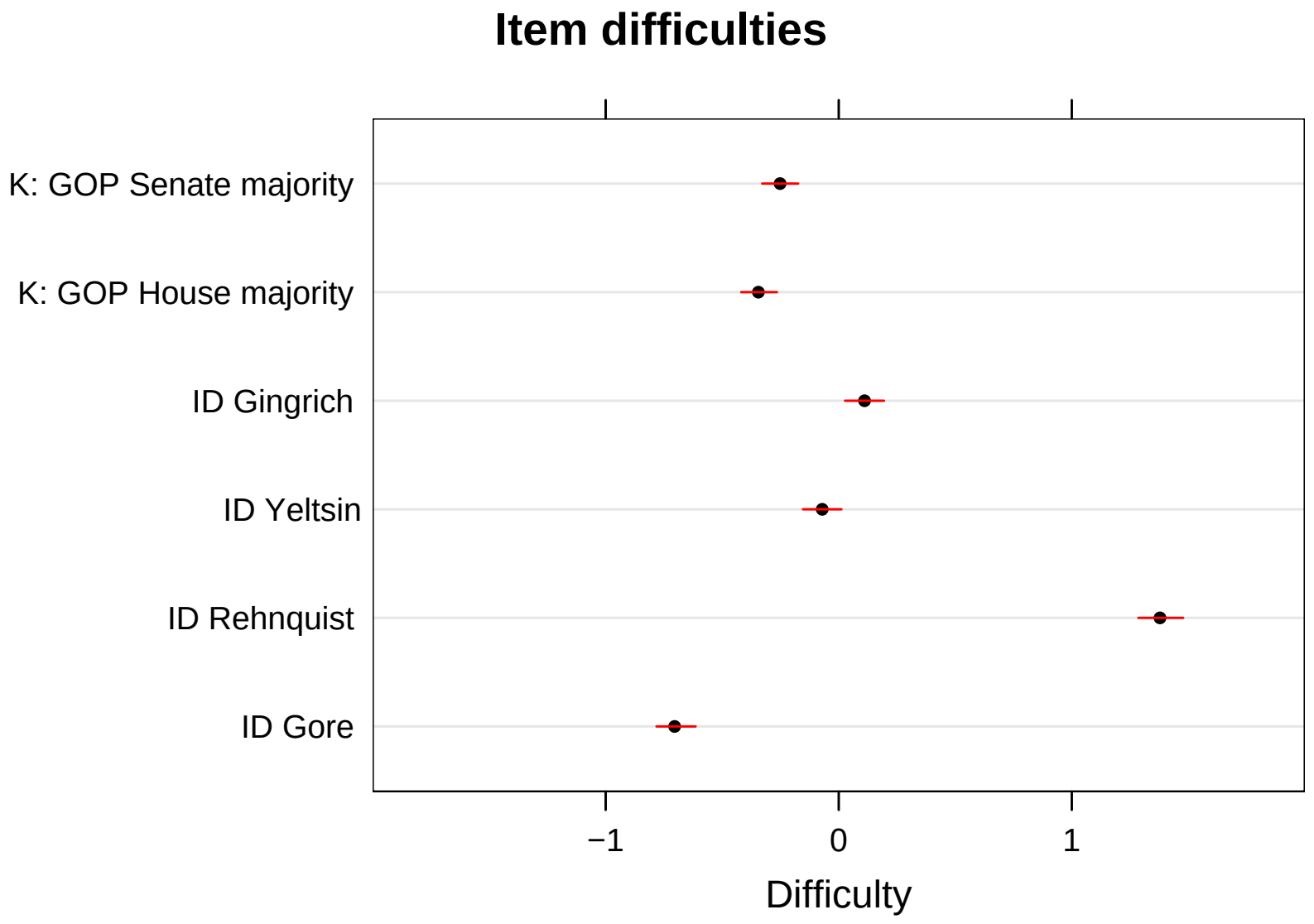
1996 party/candidate placement items (group 2)



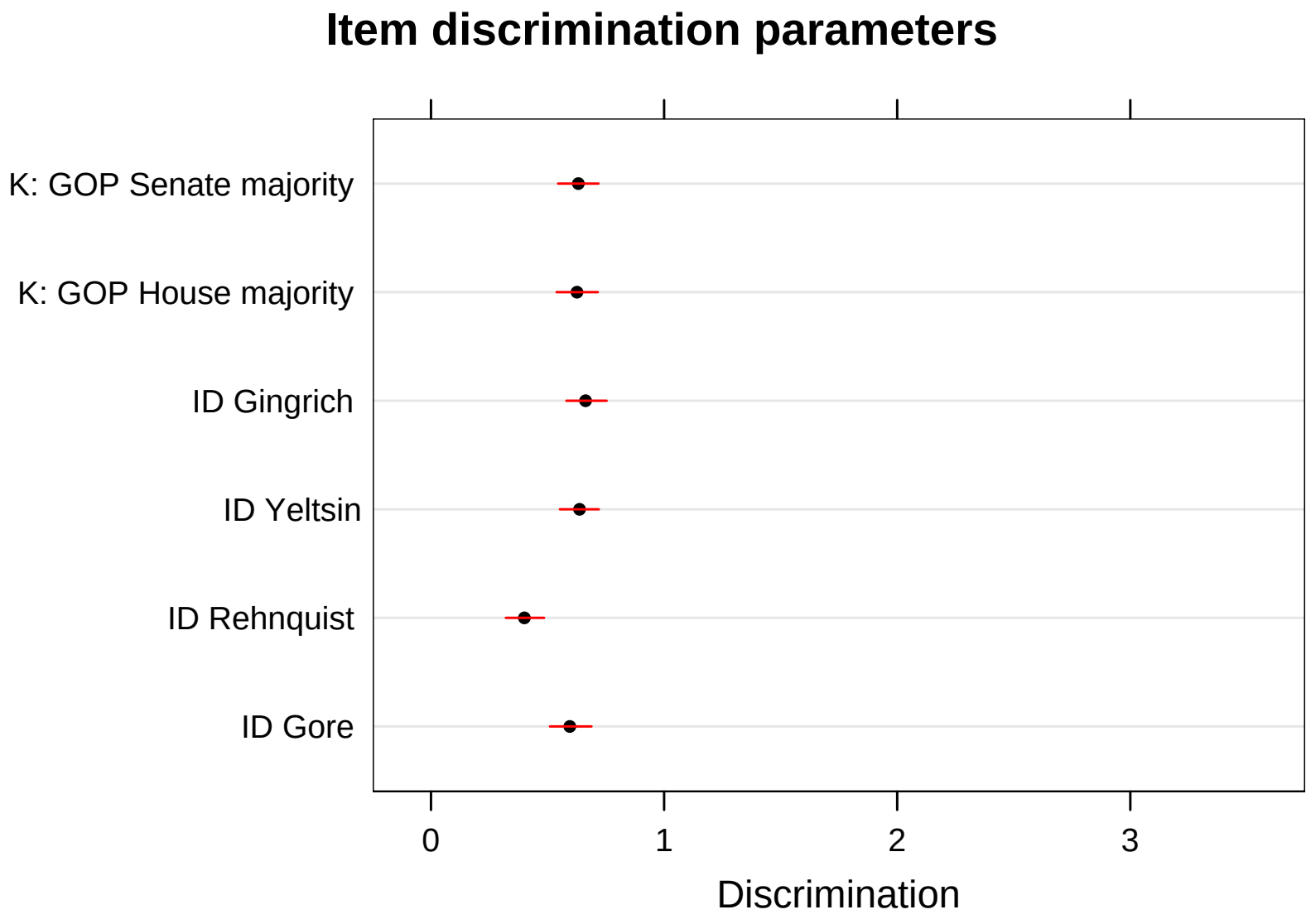
1996 party/candidate placement items (group 2)



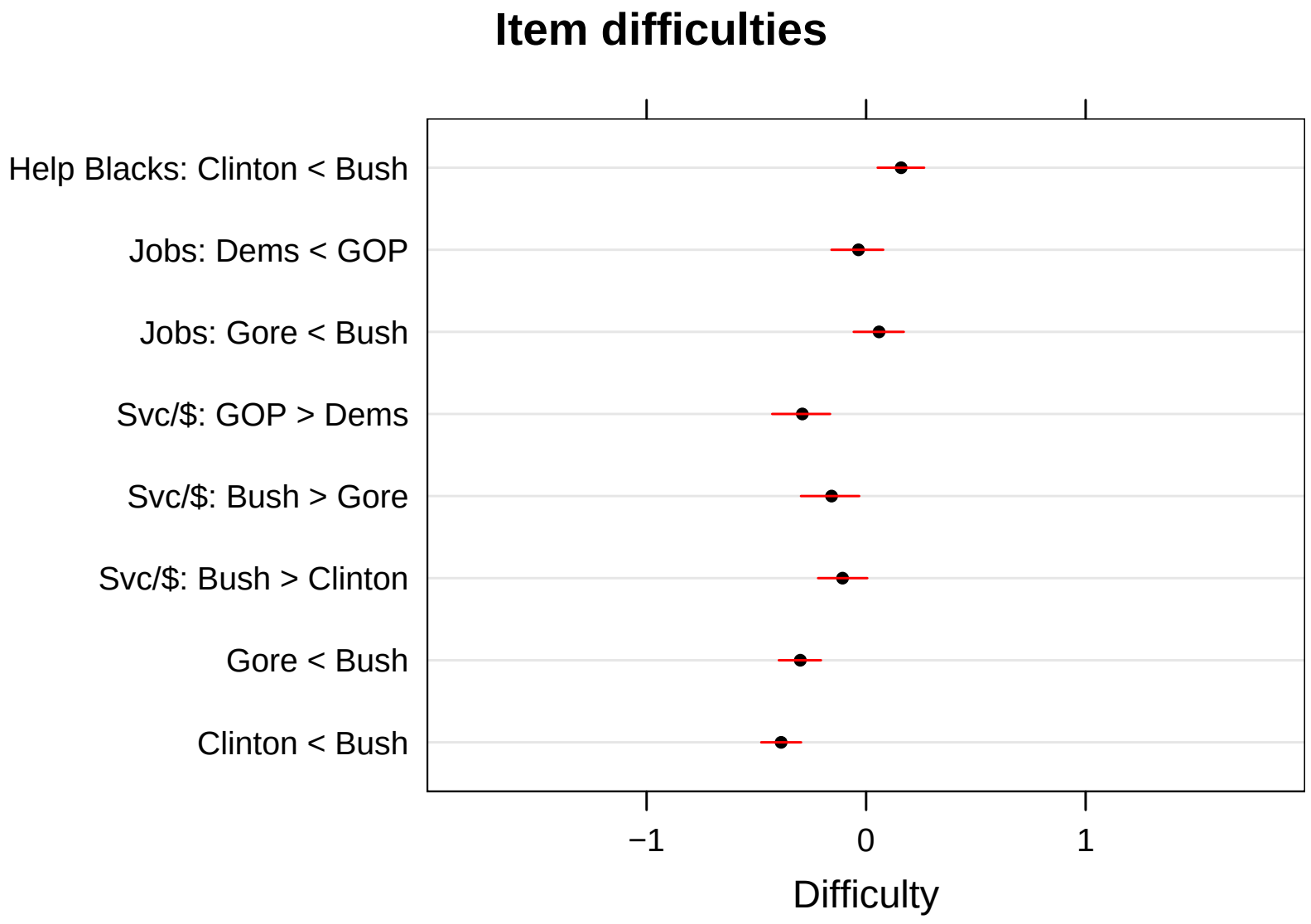
1996 knowledge items



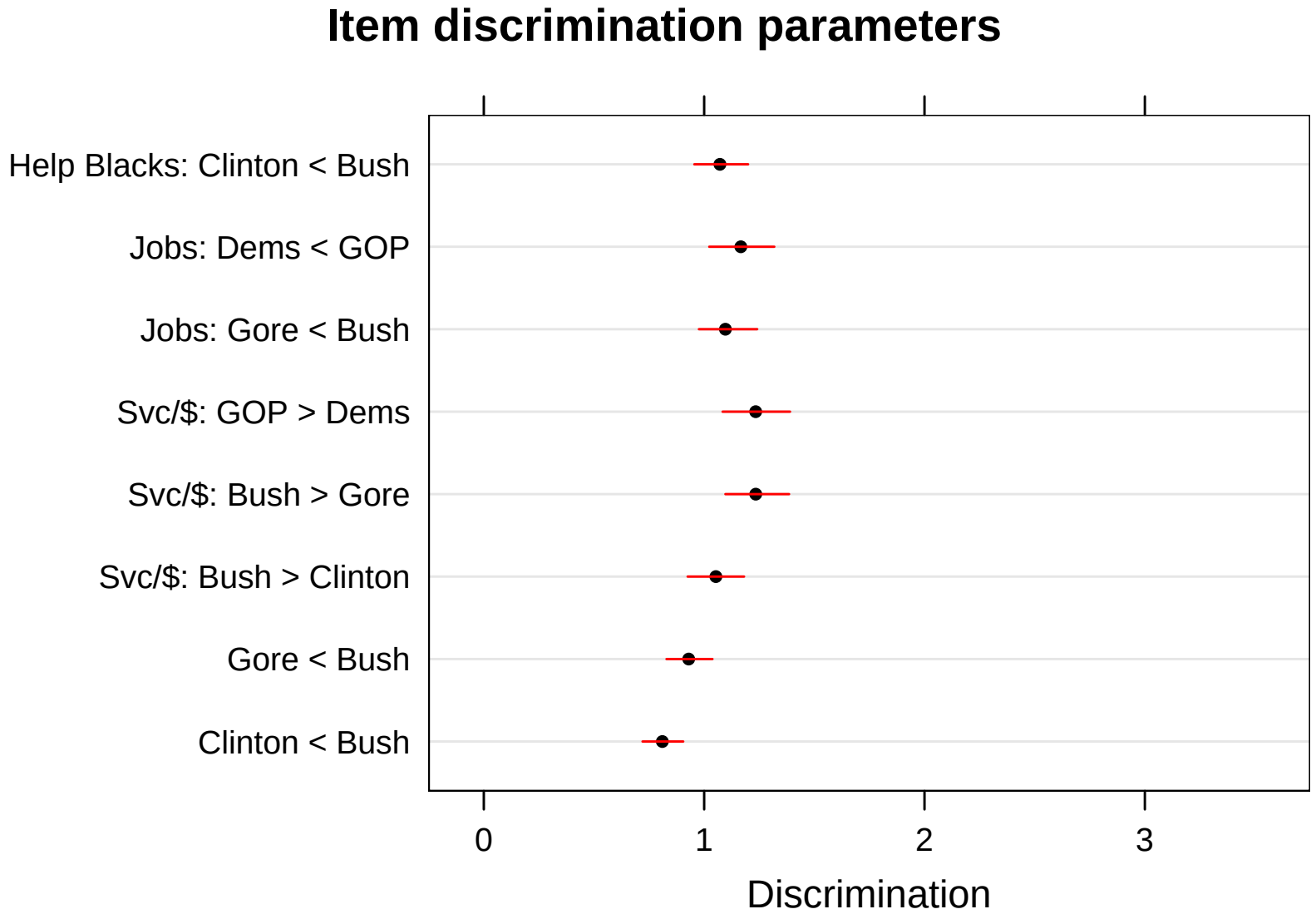
1996 knowledge items



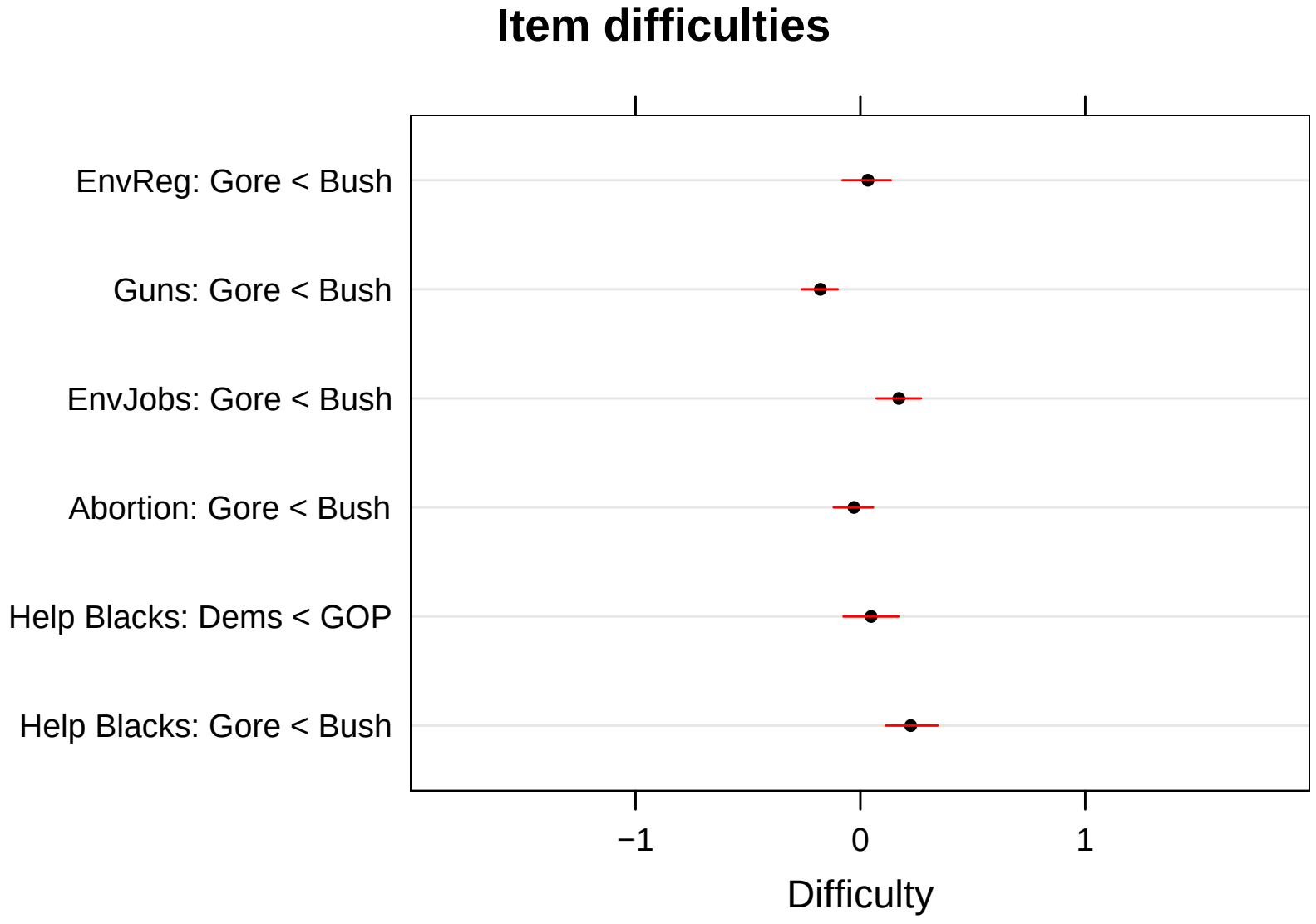
2000 party/candidate placement items (group 1)



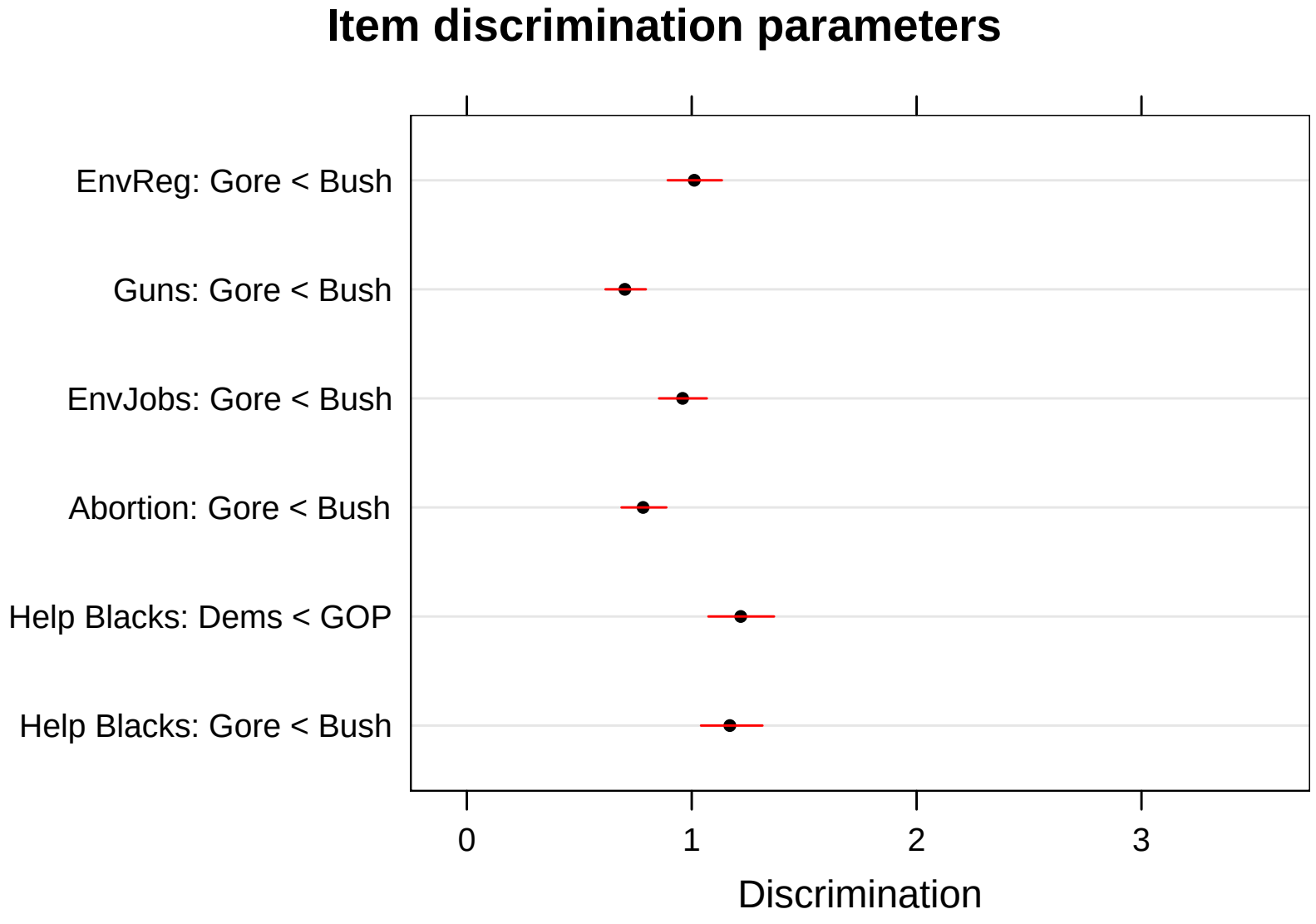
2000 party/candidate placement items (group 1)



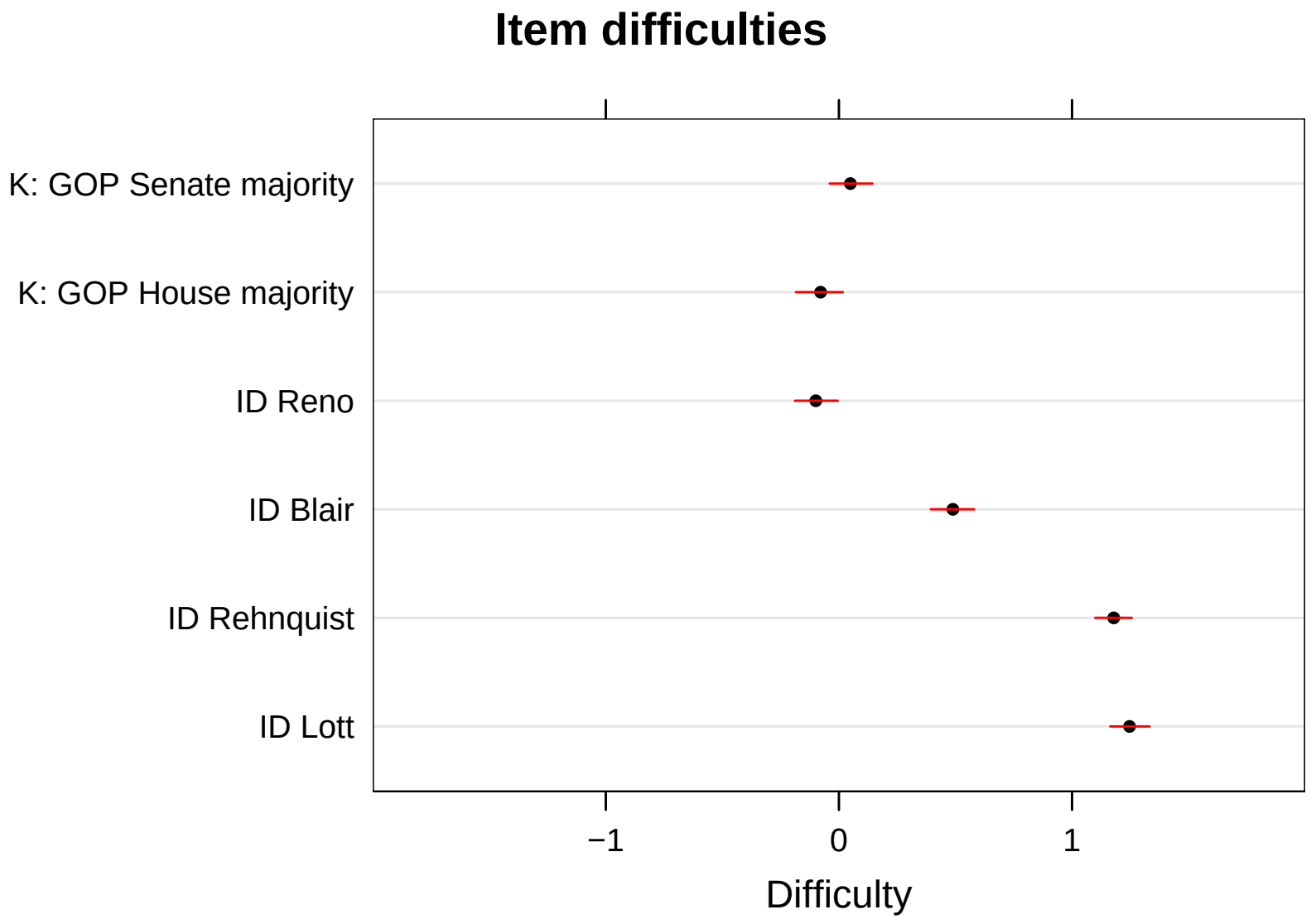
2000 party/candidate placement items (group 2)



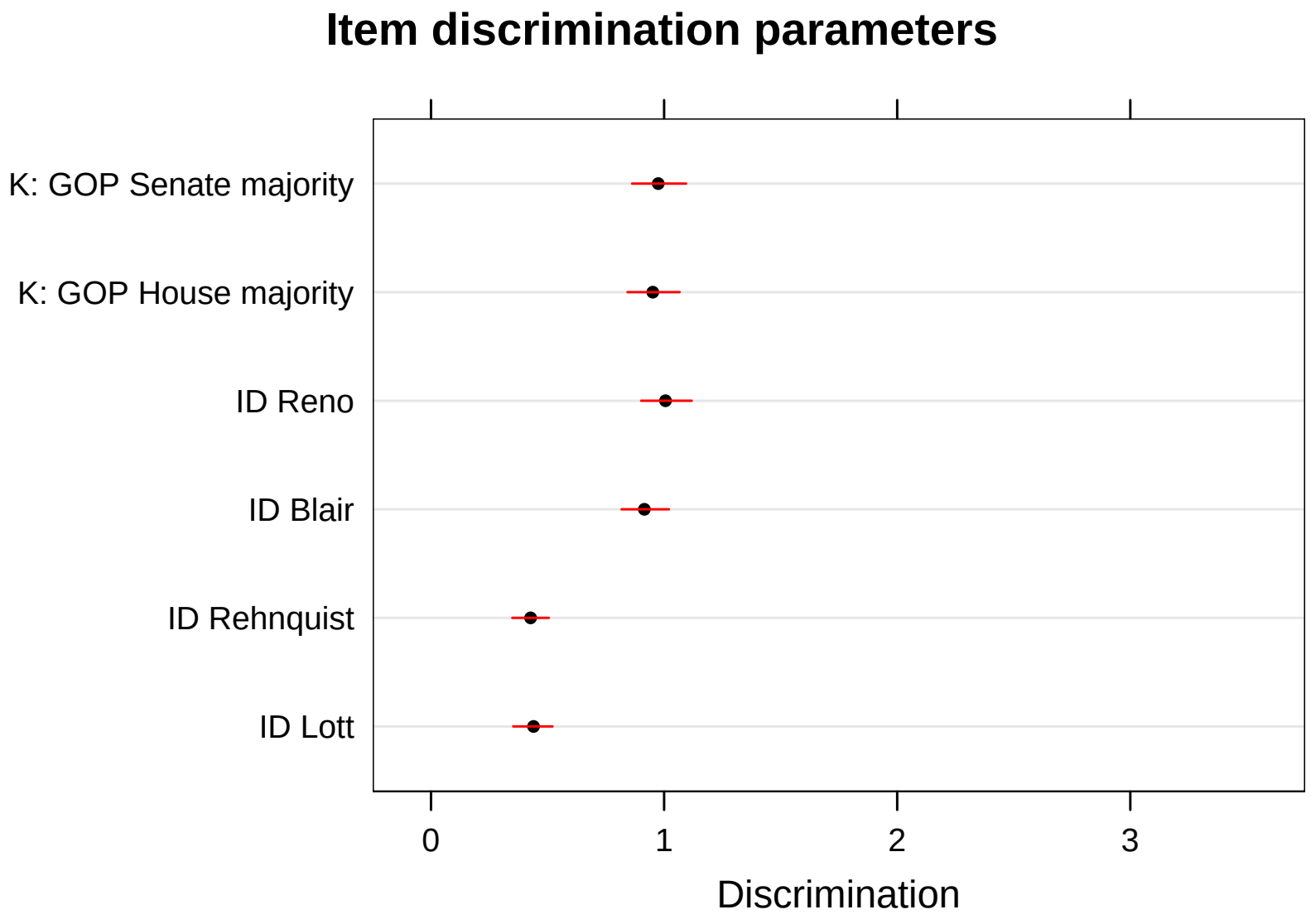
2000 party/candidate placement items (group 2)



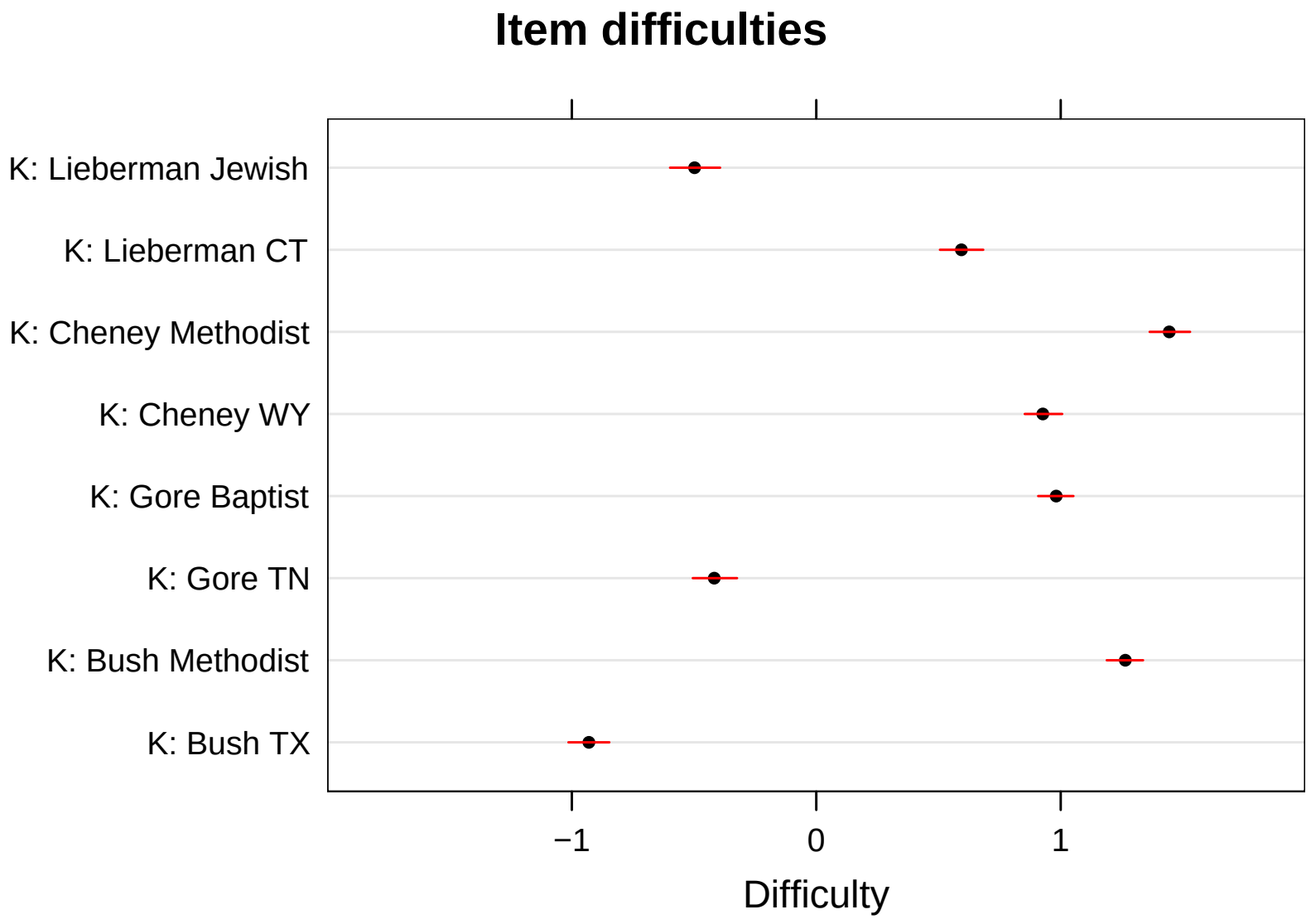
2000 knowledge items



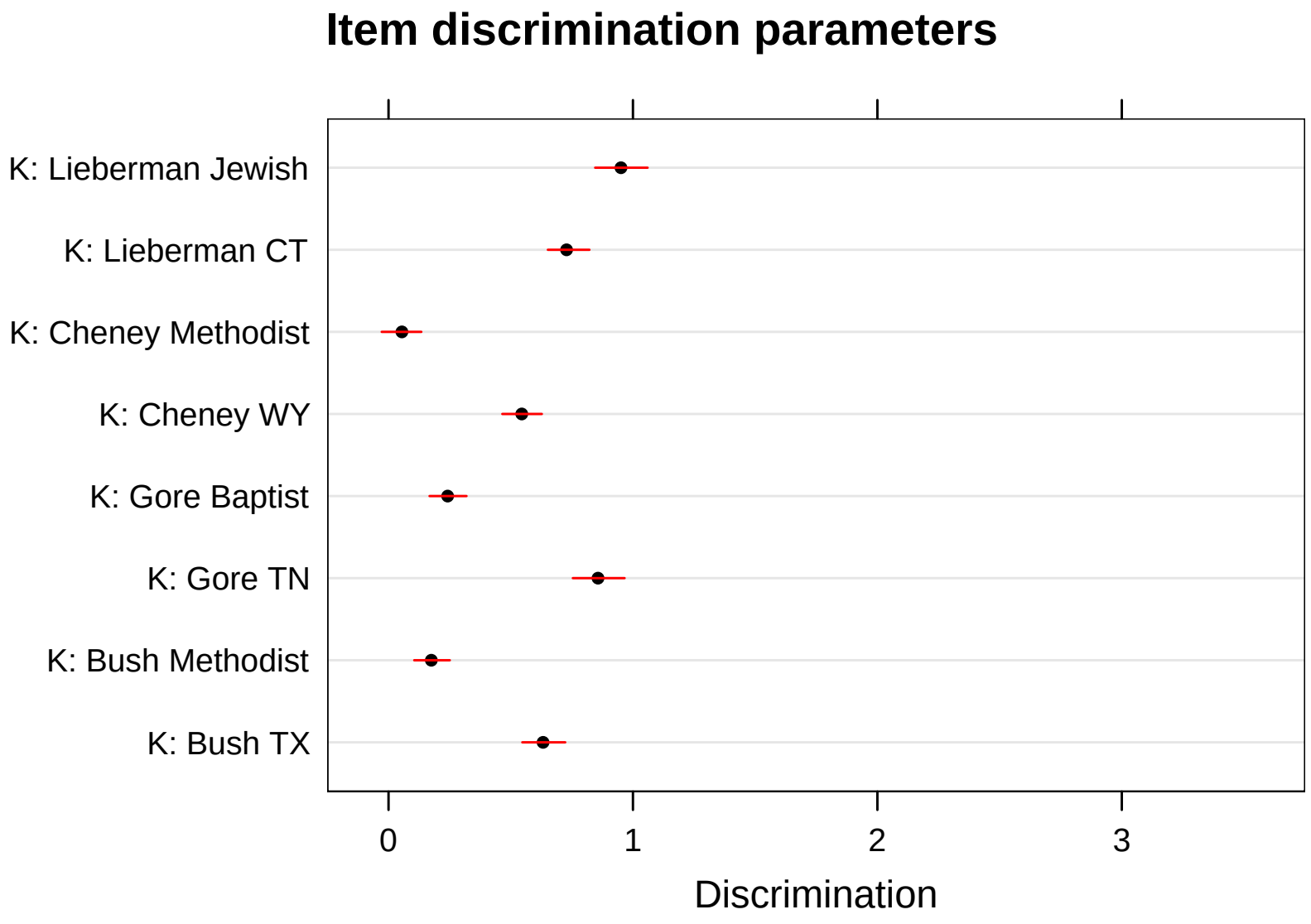
2000 knowledge items



2000 candidate biographical items



2000 candidate biographical items



Findings

- Knowledge items appeared to outperform party placement items in the Netherlands, at least in 1998.
- In the U.S., both knowledge items and party/candidate placement items appeared to perform similarly in all three years examined. (But note weak performance of Supreme Court and congressional leader IDs.)
- Most candidate biographical data questions in 2000 did not perform well (particularly religion), perhaps due to low public awareness and low salience.

Future extensions

- Additional years (2002, 2005 DPES; 2004 ANES) and countries (Britain, Canada, . . .).
- Should consider the possibility of multidimensionality.
- Importance of general versus domain-specific political knowledge.
- Need to account for the known error in the estimated abilities when used in second-stage analysis, rather than using point estimates. Quinn and Martin argue it is not problematic–but we could produce better estimates of the effects of the abilities if we account for the known error.
- Incorporating ideological measures like RU (recognition/understanding) and AU (active use) into the analysis.