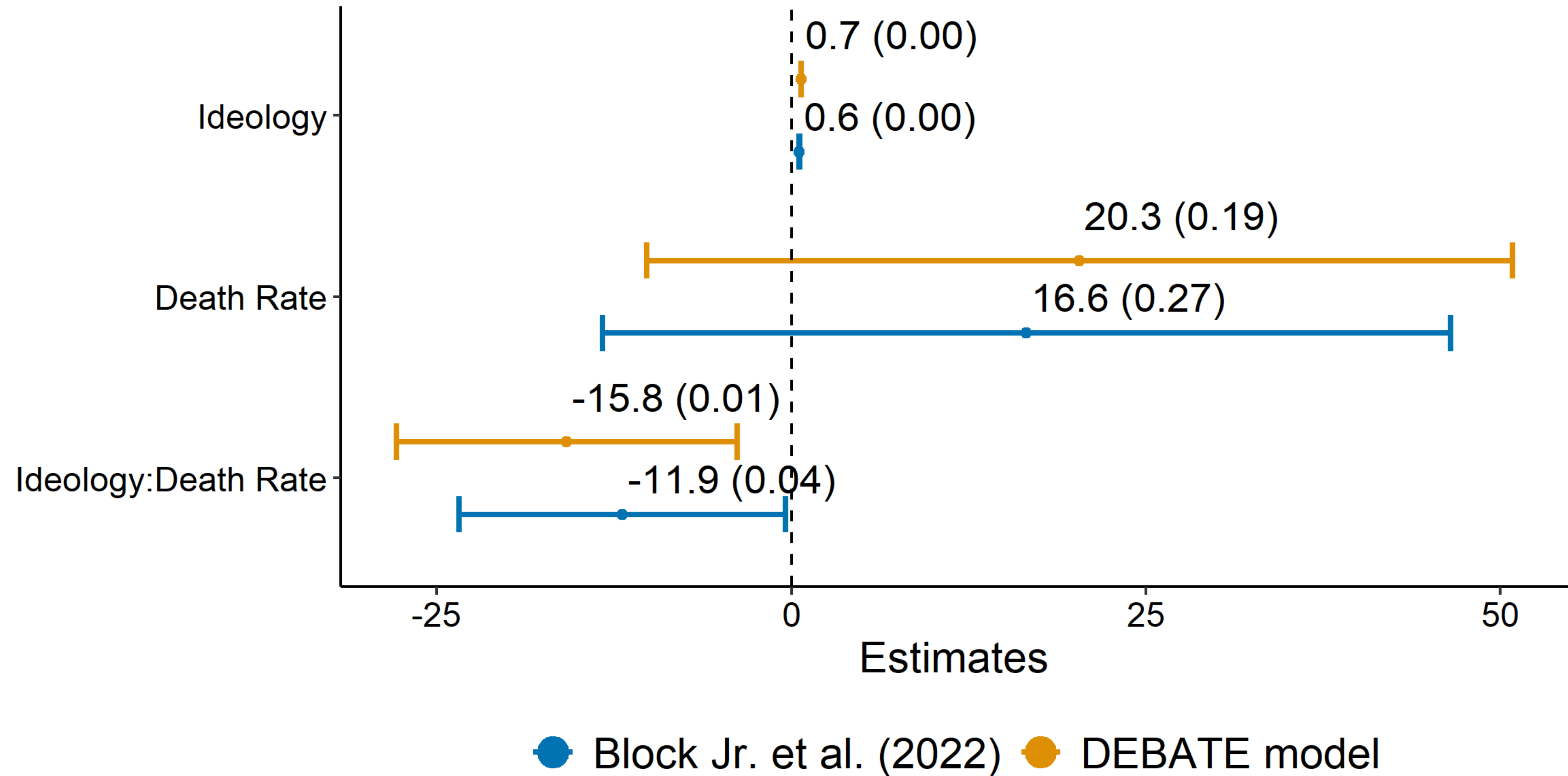# Midterm Review & Estimands

POLS 602
Fall 2025

Dr. Mike Burnham
Texas A&M Political Science

Forest Plots

Connecting Theory to Math

$$P(y_{ij} = 1 | \alpha_j, \beta_i, \gamma, \theta_i, \phi_j) = \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma \|\theta_i - \phi_j\|^2\right).$$

Suppose that each Twitter user $i \in \{1, \ldots, n\}$ is presented with a choice between following or not following another target user $j \in \{1, \ldots, m\}$, where j is a political actor who has a Twitter account.[4] Let $y_{ij} = 1$ if user i decides to follow user j, and $y_{ij} = 0$ otherwise. For the reasons explained above, I expect this decision to be a function of the squared Euclidean distance in the latent ideological dimension[5] between user i and j: $-\gamma \|\theta_i - \phi_j\|^2$, where $\theta_i \in \mathbb{R}$ is the ideal point of Twitter user i, $\phi_j \in \mathbb{R}$ is the ideal point of Twitter user j, and γ is a normalizing constant.

To this core model, I add two additional parameters, α_j and β_i. The former measures the popularity of user j. This parameter accounts for the fact that some political accounts are more likely to be followed, due to the higher profile of the politicians behind them (e.g., we would expect the probability of following @BarackObama to be higher than the probability of following a random member of the US Congress) or for other reasons (politicians who "tweet" more often are more likely to be highly visible and therefore also to have more followers). The latter measures the level of political interest of each user i. Similarly, this parameter accounts for the differences in

the number of political accounts each user i decides to follow, which could be related to the overall number of Twitter users they follow, or their overall level of interest in politics.

The probability that user i follows a political account j is then formulated as a logit model:

$$P(y_{ij} = 1 | \alpha_j, \beta_i, \gamma, \theta_i, \phi_j) = \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma \|\theta_i - \phi_j\|^2\right). \tag{1}$$

What is your Estimand?

See: Lundberg et al. 2021

Hypothesizing in the model

- Much of social science can be described as "hypothesis about regression coefficients"

- Research goals are defined *inside* a statistical model

- Results:

  - HARKing

  - Bad controls

  - Unstated assumptions and uninformative models

  - Undefined research goals

What is your estimand?

- Estimand: The precise quantity we want to estimate

- Theoretical estimand: The unit specific quantity we want to estimate, and the target population we want to aggregate across.

- Empirical estimand: An observable quantity linked to a theoretical estimand by assumptions about the relationship between observed and unobserved data.

What is your estimand?

- Estimand: The precise quantity we want to estimate

- Theoretical estimand: The unit specific quantity we want to estimate, and the target population we want to aggregate across.

- Empirical estimand: An observable quantity linked to a theoretical estimand by assumptions about the relationship between observed and unobserved data.

Theoretical estimand

By argument

What is your estimand?

- Estimand: The precise quantity we want to estimate

- Theoretical estimand: The unit specific quantity we want to estimate, and the target population we want to aggregate across.

- Empirical estimand: An observable quantity linked to a theoretical estimand by assumptions about the relationship between observed and unobserved data.

```
┌─────────────┐            ┌─────────────┐
│ Theoretical │ ─────────▶ │  Empirical  │
│  estimand   │            │  estimand   │
└─────────────┘            └─────────────┘
  By argument                By assumption
```

What is your estimand?

- Estimand: The precise quantity we want to estimate

- Theoretical estimand: The unit specific quantity we want to estimate, and the target population we want to aggregate across.

- Empirical estimand: An observable quantity linked to a theoretical estimand by assumptions about the relationship between observed and unobserved data.

```
┌──────────────┐        ┌──────────────┐        ┌──────────────┐
│  Theoretical │ ─────► │  Empirical   │ ─────► │  Estimation  │
│   estimand   │        │   estimand   │        │   strategy   │
└──────────────┘        └──────────────┘        └──────────────┘
  By argument             By assumption            By data
```