Quantitatively Predictive Logical Models and Applications in Political Science:

Connections with Statistics and Machine Learning[1]

Yuki Atsusaka
atsusaka.org

Provost's Fellow of the Politics of Race and Ethnicity
Dartmouth College

Spring Methodology Reading Group [Latest Slides]
American University
2/1/2023

[1] For feedback, I thank Hikaru Yamagishi.

Simply Put..., Logical Models Want to Do This

Physics theories

- $\mathbf{F} = m\mathbf{a}$ — Netwon's second law
- $E = mc^2$ — Einstein's special relativity
- $S = k \ln W$ — Baltzmann equation

⇝ Building rockets

⇝ Go to Moon, Mars, Jupiter

Logical models

- $N_s = (MS)^{1/6}$ — Effective # of seat-winning parties
- $\mathbb{P}(W) = \Phi((MC)^{1/2} - 50)$ — Prob. of minority electoral success
- $C_{i,k} = \left(\frac{(1 - \sum_{j=1}^{k-1} C_{i,j}) C_{i,k-1}}{L_i - k + 1} \right)^{1/2}$ — Within-list vote shares in OLPR

⇝ Building political institutions

⇝ Shape competition, party system, representation

Goal of This Talk

- What are quantitatively predictive logical models?
    1. Definition
    2. Historical context
    3. How to build logical models

- Discussion (Q&A)
    - A1. Open research questions
    - A2. Minority representation (Atsusaka, 2021)
    - A3. Candidate vote shares in open-list PR (Atsusaka & Cantú, 2023)

Novel approach in this talk

Explain **logical models** through the lens of **statistics** and **machine learning**

Takeaway Points

- Logical models make predictions **without a black box** based on theory in the **absence of data**
- Logical models help us **accumulate** scientific knowledge

What Are Logical Models?

Origin: **Physics-oriented** approach in electoral systems
Can be dubbed: Taagepera-Shugart framework of **theory building**

Figure: Rein Taagepera (left) and Matthew Shugart (right)

What Are Logical Models?

Definition

Logical models are a class of mathematical models based on deductive logic that seek to find law-like patterns in politics.

Motivation

- Discover universal "laws" in politics $\rightsquigarrow$ # parties
- Use laws for "electoral engineering" $\rightsquigarrow$ changing # seats $\longrightarrow$ # parties

Features

- **Model**: Mathematical ($\leftrightarrow$ verbal theory)
- **Logical**: Explaining why ($\leftrightarrow$ machine learning)
- **Predictive**: *Before* looking at data ($\leftrightarrow$ regression)
- **Quantitatively**: *Exactly how much* ($\leftrightarrow$ formal model/game theory)

Why? (Historical Context)

"Age of Regression" (late 1980s to 1990s)

- **Model**: Mathematical (same)
- **Empirical**: Need data to estimate $\rightsquigarrow (X^T X)^{-1} X^T y$
- **Descriptive**: Need estimated coefficients $\rightsquigarrow \widehat{Y} = f(\widehat{\beta}_1)$
- **Qualitatively**: Only directional hypothesis $\rightsquigarrow$ X increases Y or $\beta_1 = 0$

Alternative approaches

- Beck et al. (2000) $\rightsquigarrow$ Machine learning
- Gill (2002) $\rightsquigarrow$ Bayesian statistics
- King & Zeng (2004) $\rightsquigarrow$ Causal inference
- Granato & Scioli (2004) $\rightsquigarrow$ EITM (formal model + regression)
- Achen (2005) $\rightsquigarrow$ Control variables
- Druckman et al. (2006) $\rightsquigarrow$ Experiments

- Taagepera (2007) *Predicting Party Sizes: The Logic of Simple Electoral Systems*
- Taagepera (2008) *Making Social Sciences More Scientific: The Need for Predictive Models*

Historical Development

Pre-History

- Taagepera (1976) *TPS* — physics-inspired political science
- Taagepera (1978) *SSR* — duration and size of empires
- Laakso & Taagepera (1979) *CPS* — effective number of parties
- Taagepera & Shugart (1989) *Seats from Votes* — generalized Duverger's rule
- Shugart & Carey (1992) *Presidents and Assemblies* — number of presidents
- Taagepera & Shugart (1993) *APSR* — generalized Duverger's rule

Birth of Logical Models

- **Taagepera (2007)** *Predicting Party Sizes* — # seat-winning parties
- **Taagepera (2008)** *Making Social Sciences* — # electoral volatility, assembly size
- Taagepera & Sikk (2010) *PP* — cabinet duration
- Bergman et al. (2013) *ES* — intra-list vote shares
- Taagepera et al. (2014) *ES* — voter turnout
- Li & Shugart (2016) *ES* — # seat-winning parties
- Shugart & Taagepera (2017) *Votes from Seats* — proportionality + more

Expansion of Logical Models

- Taagepera & Nemčok (2021) *PP* — interest group plurality
- Atsusaka (2022) *APSR* — minority representation
- Atsusaka & Cantú (2023) *Working Paper* — vote shares in OLPR
- Atsusaka & Yamagishi (2023) *Working Paper* — # parties in MMM

What Do Logical Models Do?

- Logical Model as an Algorithm
  - Input: **Measurable** political outcome
    - Outcome with universally recognized/natural units
    - Counts, years, proportions ($\circ$) $\leftrightarrow$ feelings, latent scores ($\times$)
  - Output: **Average** outcome value conditional on institutional variables

- Three Steps in Logical Model Building
  1. Take an outcome of interest Y
  2. Find its minimal and maximum possible values (y_0,y_1) based logic or theory
  3. Take the geometric mean of the two logical bounds (as an estimator of $\mathbb{E}[Y|x] = (y_0 \times y_1)^{1/2}$)

Example: Seat-Product Model (Taagepera 2007)

What is the average # of seat-winning parties in a country?

seat-winning parties in a district with M seats (i.e., district magnitude)

- Minimal # seat-winning parties $\rightsquigarrow 1$ (lowest fragmentation)
- Maximal # seat-winning parties $\rightsquigarrow M$ (highest fragmentation)
- Logical bound for # parties $\rightsquigarrow \mathbb{E}[N_{district}] \in [1, M]$
- Geometric mean $\rightsquigarrow (1 \times M)^{1/2} = M^{1/2}$
- Logical model $\rightsquigarrow \mathbb{E}[N_{district}] = M^{1/2}$ * Prior studies: $N_{district} = M^{1/2}$

seat-winning parties in a country with S seats and average district magnitude M

- Minimal # parties $\rightsquigarrow M^{1/2}$ (no variation across districts, $S > M$)
- Maximal # parties $\rightsquigarrow S^{1/2}$ (nationwide district, $S = M$)
- Logical bound for # parties $\rightsquigarrow \mathbb{E}[N_{national}] \in [M^{1/2}, S^{1/2}]$
- Geometric mean $\rightsquigarrow (M^{1/2} \times S^{1/2})^{1/2} = ((MS)^{1/2})^{1/2} = (MS)^{1/4}$
- Logical model $\rightsquigarrow \mathbb{E}[N_{national}] = (MS)^{1/4}$

seat-winning parties in the U.S. House:
$M \approx 1, S = 435 \rightsquigarrow \mathbb{E}[N_{national}] = (1 \times 435)^{1/4} = 4.56$

Example: Seat-Product Model (Taagepera 2007)

Shugart & Taagepera (2017): Effective Number of Seat-Winning Parties

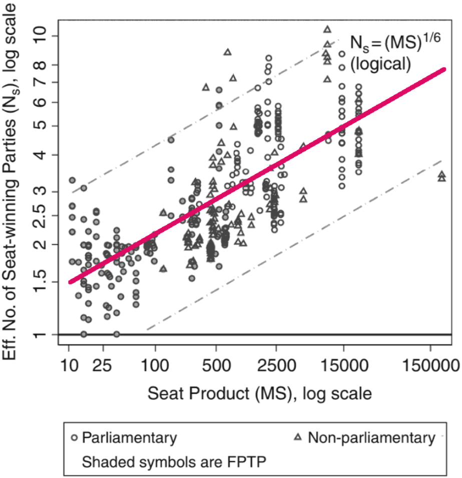


FIGURE 7.1 Relation of the nationwide effective number of seat-winning parties (N_S) to the seat product (MS)

Connection with Partial Identification

- QOI: conditional mean $\mathbb{E}[Y|x]$
- Y has bounded range, $Y \in [y_0, y_1]$
- Data are partially observed ($z = 1$) and missing ($z = 0$)

$$\mathbb{E}[Y|x] = \underbrace{\mathbb{E}[Y|x, z = 1]}_{\text{Observed Data}} \mathbb{P}(z = 1|x) + \underbrace{\mathbb{E}[Y|x, z = 0]}_{\textbf{Missing Data}} \mathbb{P}(z = 0|x) \quad (1)$$

- Identification region

$$H\{\mathbb{E}[Y|x]\} = \Big[\underbrace{\mathbb{E}[Y|x, z = 1]\mathbb{P}(z = 1|x) + y_0\mathbb{P}(z = 0|x)}_{\textbf{Lower bound}}, \underbrace{\mathbb{E}[Y|x, z = 1]\mathbb{P}(z = 1|x) + y_1\mathbb{P}(z = 0|x)}_{\textbf{Upper bound}}\Big] \quad (2)$$

- When we have no data: $\mathbb{P}(z = 0|x) = 1$

$$H\{\mathbb{E}[Y|x]\} = [y_0, y_1] \quad (3)$$

- For our model:

$$H\{\mathbb{E}[N_{district}|M]\} = [1, M] \quad (4)$$

Logical Models, Statistics, Machine Learning

- Logical models push us hard to think about
  - What we can say about the outcome of interest in the "absence of data"
  - *Evaluating* our theoretical predictions
  - Using our models to solve *practical electoral engineering questions*
    - District magnitude → party fragmentation
    - % minority voters → minority representation
    - List size → intraparty competition

- New interpretations
  - Logical models as partial identification for **theory** building (↔ empirical analysis)
  - Geometric mean as an **estimator** of $\mathbb{E}[Y|x]$
  - Logical models offer **out-of-sample** predictions

Conclusion

✓ Logical models make predictions **without a black box** based on theory

✓ Logical models help us **accumulate** scientific knowledge

Open Research Questions

- Theoretical properties of logical models
    - Why $\mathbb{E}[\cdot]$ for **theory building**?
    - Causal explanation beyond logical constrain (e.g., strategic voting)
- Statistical properties of logical models
    - Why geometric mean for **estimation**?
    - How can we account for uncertainty?
- Accessing predictive accuracy of logical models
    - What are best practices for **validation**?
    - Integrating theory with empirics (e.g., calibration)

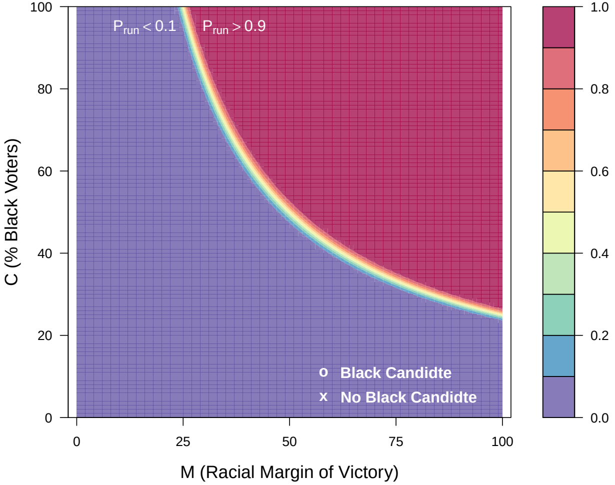
Application 1

Atsusaka (2021): What is $\mathbb{P}(\text{Minority Runs}|\text{District})$?

- Model building
  - Clairvoyant: $\mathbb{P}(\text{Minority runs}_t) = I(V_t^M - V_t^W > 0) = I(\text{Minority wins})$
  - Realistic: $\mathbb{P}(\text{Minority runs}_t) = \Phi(\underbrace{V_t^M - V_t^W}_{\text{QOI to model}}) = \mathbb{P}(\text{Minority wins})$
  - Bound I: $C' = V_{t-1}^M - V_{t-1}^W$ (Same as the last time)
  - Bound II: $M' = V_t^{M*} - V_t^{W*}$ (Perfect racially polarized voting)
  - Identification range: $\mathbb{E}[V_t^M - V_t^W] \in [C', M']$ **theoretical* bound
  - Logical model: $\mathbb{E}[V_t^M - V_t^W] = (M'C')^{1/2}$
  - Adjustment: $M = M' + 50$, $C = C' + 50$ *for geometric mean
  - Adjustment: $\mathbb{E}[V_t^M - V_t^W] = (MC)^{1/2} - 50$ *centering around zero
  - Full model: $\mathbb{P}(\text{Minority runs}_t) = \Phi((MC)^{1/2} - 50)$
  - Full model: $\mathbb{P}(\text{Minority wins}_t) = \Phi((MC)^{1/2} - 50)$
- Model evaluation
  - 90% (running) & 95% (winning) accuracy via *out-of-sample* ePCP (Herron 1999)
  - Sometimes better than estimated models

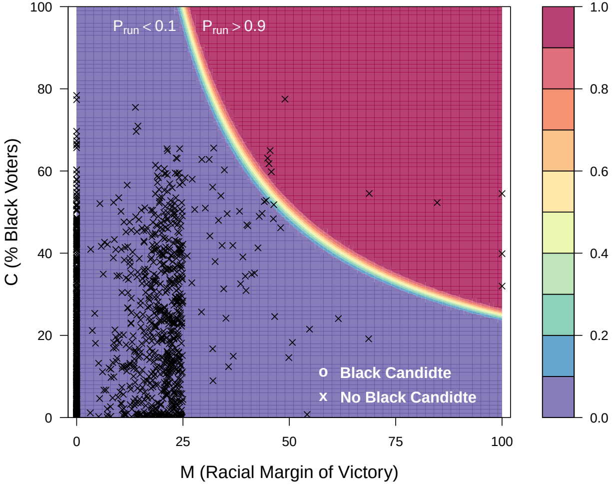
Application 1

Atsusaka (2021): Probability of Minority Candidate Entry



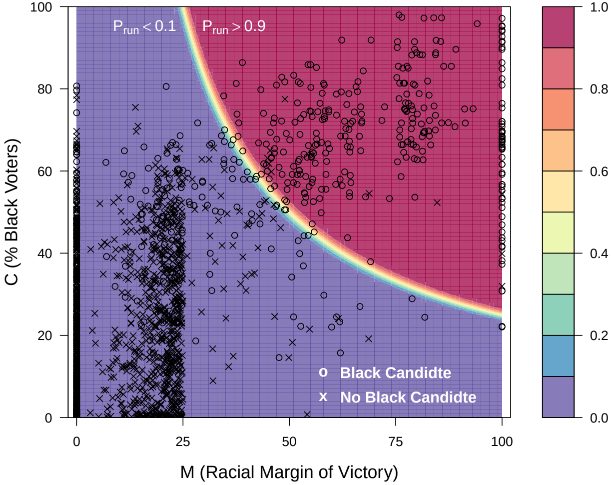
Application 1

Atsusaka (2021): Probability of Minority Candidate Entry



Application 1

Atsusaka (2021): Probability of Minority Candidate Entry



Application 2

Atsusaka & Cantú (2023): Intra-List Vote Shares in Open-List PR

What is $\mathbb{E}[\text{Within-List vote shares}|\text{District}]$?

- Most vote-winning candidate
    - Upper bound: 1
    - Lower bound: $\frac{1}{L_i}$, where L_i = list size for party i
    - Identification range: $\mathbb{E}[C_{i,1}] \in \left[\frac{1}{L_i}, 1\right]$ **logical* bound
    - Logical model: $\mathbb{E}[C_{i,1}] = \left(\frac{1}{L_i}\right)^{1/2}$
- 2nd most vote-winning candidate
    - Upper bound: $C_{i,1}$
    - Lower bound: $\frac{1-C_{i,1}}{1-L_i}$
    - Logical model: $\mathbb{E}[C_{i,2}] = \left(\frac{(1-C_{i,1})C_{i,1}}{L_i-1}\right)^{1/2}$
- k-th most vote-winning candidate
    - $\mathbb{E}[C_{i,k}] = \left(\frac{(1-\sum_{j=1}^{k-1} C_{i,j})C_{i,k-1}}{L_i-k+1}\right)^{1/2}$ $\rightsquigarrow$ entire dist of vote shares in OLPR
- Model evaluation
    - Multiple error metrics
    - Comparison against machine learning models

Application 2

Atsusaka & Cantú (2023): Intra-List Vote Shares in Open-List PR