A Bayesian Multilevel Modeling Approach to Time Series Cross-Sectional Data

Boris Shor*, Joseph Bafumi†, Luke Keele‡, David Park§

April 9, 2005

*Harris School of Public Policy Studies, University of Chicago, and Center for the Study of Democratic Politics, Woodrow Wilson School, Princeton University; boris@bshor.com

†Department of Political Science, Dartmouth College; jb878@columbia.edu

‡Department of Political Science, Ohio State University; luke.keele@politics.ox.ac.uk

§Department of Political Science, Washington University; dpark@artsci.wustl.edu

TSCS Data

TSCS Data

- Allows study of cross-sectional variation

TSCS Data

- Allows study of cross-sectional variation
- But also incorporation of dynamics

TSCS Data

- Allows study of cross-sectional variation
- But also incorporation of dynamics
- In conjunction, powerful tests of theory

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.
  - ⋆ Dynamics Not Easily Included.

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.
  - ⋆ Dynamics Not Easily Included.
  - ⋆ Costly in terms of degrees of freedom.

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.
  - ⋆ Dynamics Not Easily Included.
  - ⋆ Costly in terms of degrees of freedom.
- Random Effects via MLE

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.
  - ⋆ Dynamics Not Easily Included.
  - ⋆ Costly in terms of degrees of freedom.
- Random Effects via MLE
  - ⋆ Unit Effects Can't Be Correlated with X variables.

Complicated Error Structures

No Best Solution to Statistical Modeling of TSCS Data

- OLS with PCSEs
  - ⋆ Prone to Unit Effects.
  - ⋆ Asymptotic in T
- FE's
  - ⋆ Can't Include Time Invariant Covariates.
  - ⋆ Dynamics Not Easily Included.
  - ⋆ Costly in terms of degrees of freedom.
- Random Effects via MLE
  - ⋆ Unit Effects Can't Be Correlated with X variables.

TSCS Data Generating Process

Standard TSCS DGP:

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \tag{1}$$

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \tag{1}$$

- We generate
  - ⋆ Heteroscedastic errors.

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \tag{1}$$

- We generate
  - ⋆ Heteroscedastic errors.
  - ⋆ Contemporaneous correlation.

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \quad (1)$$

- We generate
  - ⋆ Heteroscedastic errors.
  - ⋆ Contemporaneous correlation.
  - ⋆ Unit effects of various sizes.

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \tag{1}$$

- We generate
  - ⋆ Heteroscedastic errors.
  - ⋆ Contemporaneous correlation.
  - ⋆ Unit effects of various sizes.
  - ⋆ No serial correlation yet.

TSCS Data Generating Process

Standard TSCS DGP:

$$y_{jt} = \alpha_j + \beta x_{jt} + \epsilon_{jt} \tag{1}$$

- We generate
  - ⋆ Heteroscedastic errors.
  - ⋆ Contemporaneous correlation.
  - ⋆ Unit effects of various sizes.
  - ⋆ No serial correlation yet.
    - ∗ Through the error
    - ∗ Through a lag
  - ⋆ No correlation between unit effects and $\boldsymbol{X}$.

Multilevel Modeling

- Popularized in fields like education research and biostatistics

Multilevel Modeling

- Popularized in fields like education research and biostatistics

- Increasing use in political science.
  - ⋆ Gelman and King (1993), Steenbergen and Jones (2002)
  - ⋆ Midwest Conference papers

Multilevel Modeling

- Popularized in fields like education research and biostatistics
- Increasing use in political science.
  - ⋆ Gelman and King (1993), Steenbergen and Jones (2002)
  - ⋆ Midwest Conference papers
- Rare applications to TSCS data.

Multilevel Modeling

- Popularized in fields like education research and biostatistics
- Increasing use in political science.
  - ⋆ Gelman and King (1993), Steenbergen and Jones (2002)
  - ⋆ Midwest Conference papers
- Rare applications to TSCS data.
  - ⋆ Exception is Western (1998)

Multilevel Modeling

- Popularized in fields like education research and biostatistics
- Increasing use in political science.
  - ⋆ Gelman and King (1993), Steenbergen and Jones (2002)
  - ⋆ Midwest Conference papers
- Rare applications to TSCS data.
  - ⋆ Exception is Western (1998)
- Missing Monte Carlo analysis of BML under various conditions

The Results

- Bayesian multilevel model is equally as unbiased as common estimators

The Results

- Bayesian multilevel model is equally as unbiased as common estimators
- Considerably more efficient

The Results

- Bayesian multilevel model is equally as unbiased as common estimators
- Considerably more efficient
- Reports higher quality standard errors

The Results

- Bayesian multilevel model is equally as unbiased as common estimators
- Considerably more efficient
- Reports higher quality standard errors
- And has additional advantages

The Results

- Bayesian multilevel model is equally as unbiased as common estimators
- Considerably more efficient
- Reports higher quality standard errors
- And has additional advantages

Advantages of Multilevel Models for TSCS Data: Structure

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)
  - ⋆ Unit level (state)

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)
  - ⋆ Unit level (state)
  - ⋆ Time level (year)

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)
  - ⋆ Unit level (state)
  - ⋆ Time level (year)
  - ⋆ Additional levels as needed - regions, districts, etc.

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)
  - ⋆ Unit level (state)
  - ⋆ Time level (year)
  - ⋆ Additional levels as needed - regions, districts, etc.

Advantages of Multilevel Models for TSCS Data: Structure

- Explicit modeling of data structure
  - ⋆ Predictors at multiple levels of analysis
  - ⋆ Individual level (say, state-year)
  - ⋆ Unit level (state)
  - ⋆ Time level (year)
  - ⋆ Additional levels as needed - regions, districts, etc.

Advantages of Multilevel Models for TSCS Data: Pooling

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling
  - ⋆ Complete pooling ignores unit heterogeneity.

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling
  - ⋆ Complete pooling ignores unit heterogeneity.
  - ⋆ No pooling design exaggerates the incomparability of units, whose heterogeneity is treated as maximally important.

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling
  - ⋆ Complete pooling ignores unit heterogeneity.
  - ⋆ No pooling design exaggerates the incomparability of units, whose heterogeneity is treated as maximally important.
  - ⋆ The compromise is partial pooling, which allows for unit effects to be estimated conditionally on the data and parameters at all levels.

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling
  - ⋆ Complete pooling ignores unit heterogeneity.
  - ⋆ No pooling design exaggerates the incomparability of units, whose heterogeneity is treated as maximally important.
  - ⋆ The compromise is partial pooling, which allows for unit effects to be estimated conditionally on the data and parameters at all levels.
    - ∗ When sample sizes are small, estimates of unit effects benefit from "shrinkage"

Advantages of Multilevel Models for TSCS Data: Pooling

- More reasonable approach to pooling
  - ⋆ Complete pooling ignores unit heterogeneity.
  - ⋆ No pooling design exaggerates the incomparability of units, whose heterogeneity is treated as maximally important.
  - ⋆ The compromise is partial pooling, which allows for unit effects to be estimated conditionally on the data and parameters at all levels.
    - ∗ When sample sizes are small, estimates of unit effects benefit from "shrinkage"
    - ∗ Better estimates of uncertainty

Bayesian Inference and TSCS Data 1

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled
- TSCS observations in N and T are typically low

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled
- TSCS observations in N and T are typically low
  - ⋆ Classical estimators rely on asymptotic theory

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled
- TSCS observations in N and T are typically low
  - ⋆ Classical estimators rely on asymptotic theory
  - ⋆ While Bayesian estimators benefit from more data ...

Bayesian Inference and TSCS Data 1

- TSCS units and years fixed, not sampled
  - ⋆ In Bayesian inference, data is fixed, parameters are sampled
- TSCS observations in N and T are typically low
  - ⋆ Classical estimators rely on asymptotic theory
  - ⋆ While Bayesian estimators benefit from more data ...
  - ⋆ They do not require attaining some threshold value for N or T

Bayesian Inference and TSCS Data 2

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - ⋆ "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)
  - Can estimate all unit indicators (varying intercepts) together

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)
  - Can estimate all unit indicators (varying intercepts) together
  - Sluggish or time-invariant predictors not dropped in varying intercepts model

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - ⋆ "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)
  - ⋆ Can estimate all unit indicators (varying intercepts) together
  - ⋆ Sluggish or time-invariant predictors not dropped in varying intercepts model

- Improved model fit

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - ⋆ "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)
  - ⋆ Can estimate all unit indicators (varying intercepts) together
  - ⋆ Sluggish or time-invariant predictors not dropped in varying intercepts model
- Improved model fit
  - ⋆ Inclusion of contextual effects and predictors

Bayesian Inference and TSCS Data 2

- Collinearity is less problematic
  - ⋆ "...because Bayesian pooling borrows strength from the whole sample, linear dependencies in the data within [units] present no obstacles to estimating [group-level] intercepts." (Western 1998)
  - ⋆ Can estimate all unit indicators (varying intercepts) together
  - ⋆ Sluggish or time-invariant predictors not dropped in varying intercepts model

- Improved model fit
  - ⋆ Inclusion of contextual effects and predictors

Multilevel Model

$$y_{jt} = \alpha_j + \gamma_t + \beta x_{jt} + \epsilon_{jt} \tag{2}$$

where $\epsilon_{jt} \sim N(0, \sigma^2_{\epsilon(j)})$.

Multilevel Model

$$y_{jt} = \alpha_j + \gamma_t + \beta x_{jt} + \epsilon_{jt} \tag{2}$$

where $\epsilon_{jt} \sim N(0, \sigma^2_{\epsilon(j)})$.

$$\alpha_j = \alpha_0 + \eta_j, \tag{3}$$

where $\eta_j \sim N(0, \sigma^2_\eta)$.

Multilevel Model

$$y_{jt} = \alpha_j + \gamma_t + \beta x_{jt} + \epsilon_{jt} \tag{2}$$

where $\epsilon_{jt} \sim N(0, \sigma^2_{\epsilon(j)})$.

$$\alpha_j = \alpha_0 + \eta_j, \tag{3}$$

where $\eta_j \sim N(0, \sigma^2_{\eta})$.

$$\gamma_t = \gamma_0 + \nu_t, \tag{4}$$

where $\nu_t \sim N(0, \sigma^2_{\nu})$.

Multilevel Model

$$y_{jt} = \alpha_j + \gamma_t + \beta x_{jt} + \epsilon_{jt} \tag{2}$$

where $\epsilon_{jt} \sim N(0, \sigma^2_{\epsilon(j)})$.

$$\alpha_j = \alpha_0 + \eta_j, \tag{3}$$

where $\eta_j \sim N(0, \sigma^2_{\eta})$.

$$\gamma_t = \gamma_0 + \nu_t, \tag{4}$$

where $\nu_t \sim N(0, \sigma^2_{\nu})$.

The Contenders

- OLS + regular standard errors

The Contenders

- OLS + regular standard errors
- OLS + PCSEs (Beck-Katz 1995, etc.)

The Contenders

- OLS + regular standard errors
- OLS + PCSEs (Beck-Katz 1995, etc.)
- FE + FE standard errors

The Contenders

- OLS + regular standard errors
- OLS + PCSEs (Beck-Katz 1995, etc.)
- FE + FE standard errors
- FE + robust standard errors (Arrelano 1987)

The Contenders

- OLS + regular standard errors
- OLS + PCSEs (Beck-Katz 1995, etc.)
- FE + FE standard errors
- FE + robust standard errors (Arrelano 1987)
- Bayesian multilevel model

The Contenders

- OLS + regular standard errors
- OLS + PCSEs (Beck-Katz 1995, etc.)
- FE + FE standard errors
- FE + robust standard errors (Arrelano 1987)
- Bayesian multilevel model

The Monte Carlo Analysis

- Ran DGP for **Nsim** simulations (250) with true β of 10

The Monte Carlo Analysis

- Ran DGP for **Nsim** simulations (250) with true β of 10

The Monte Carlo Analysis

- Ran DGP for **Nsim** simulations (250) with true β of 10
- We varied N, T, heteroscedasticity, contemporaneous correlation, and the scale of the unit effects

The Monte Carlo Analysis

- Ran DGP for **Nsim** simulations (250) with true β of 10
- We varied N, T, heteroscedasticity, contemporaneous correlation, and the scale of the unit effects
- Ran estimators for **Nsim** data sets

Bias

How biased are these estimators on average?

$$\text{Bias} = 100 \cdot (\bar{\beta}/\beta_{true}) \tag{5}$$

Bias

	N	T	Cor	Het	A.Sc	OLS Bias	FE Bias	BML Bias
1	15	5	0.25	0.3	5	101	100	100
2	15	20	0.25	0.3	5	100	100	100
3	15	50	0.25	0.3	5	100	100	100
4	5	20	0.25	0.3	5	101	100	100
5	50	20	0.25	0.3	5	100	100	100
6	500	20	0.25	0.3	5	100	100	100
7	15	20	0.25	0.0	5	100	100	100
8	15	20	0.25	0.6	5	100	100	100
9	15	20	0.25	0.9	5	100	100	100
10	15	20	0.01	0.3	5	100	100	100
11	15	20	0.50	0.3	5	100	100	100
12	15	20	0.75	0.3	5	100	100	100
13	15	20	0.25	0.3	1	100	100	100
14	15	20	0.25	0.3	10	100	100	100
15	15	20	0.25	0.3	25	100	100	100

Betas

So βs are unbiased on average, but what about their spread?

Betas

So βs are unbiased on average, but what about their spread?

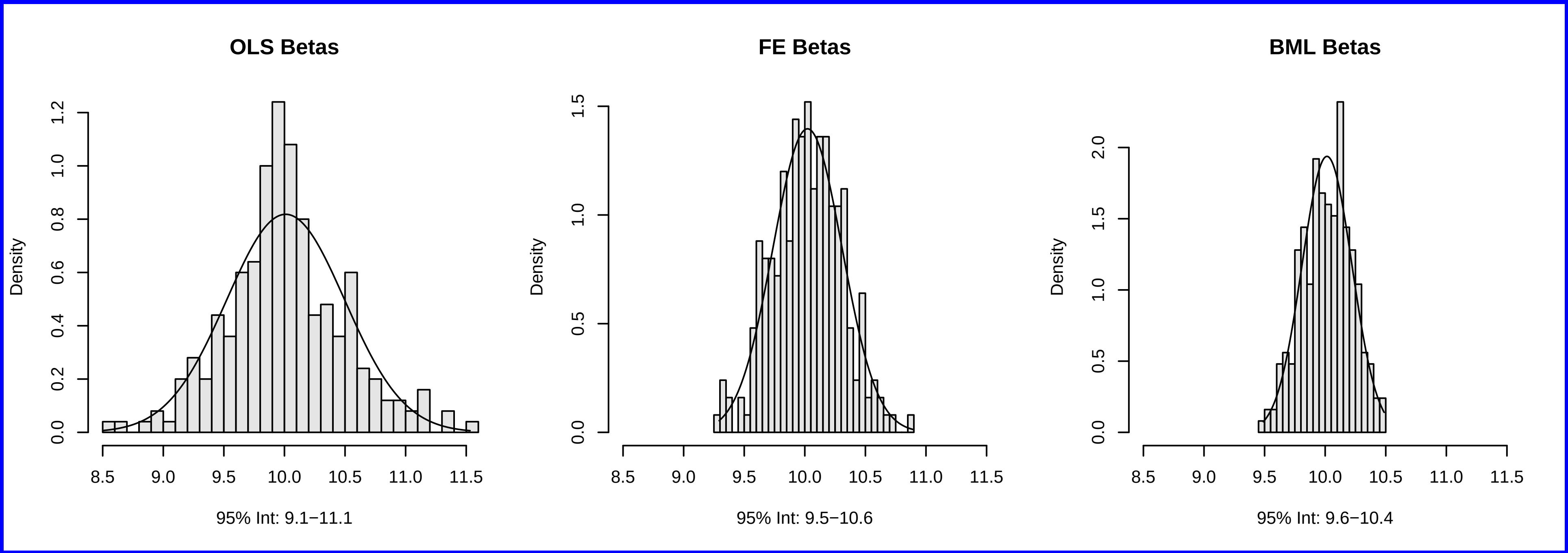

Figure 1: *Histogram of estimated βs for Ordinary Least Squares, Fixed Effects, and Bayesian multilevel models. N(sims) $=$ 250, N $=$ 15, T $=$ 20, Cor $=$ 0.25, Het $=$ 0.3, A.Scale $=$ 5.*

RMSE/Efficiency

$$\text{RMSE} = \sqrt{\sum_{l=1}^{nsims} (\beta^{(l)} - \beta_{true})^2} \quad (6)$$

RMSE/Efficiency

$$\text{RMSE} = \sqrt{\sum_{l=1}^{nsims} (\beta^{(l)} - \beta_{true})^2} \tag{6}$$

$$\text{Relative Efficiency} = 100 \cdot \frac{\text{RMSE}_{OLS|FE}}{\text{RMSE}_{BML}} \tag{7}$$

RMSE/Efficiency

	N	T	Cor	Het	A.Sc	OLS RMSE	FE RMSE	BML RMSE	OLS/BML Eff	FE/BM Eff
1	15	5	0.25	0.3	5	1.33	0.73	0.57	233	128
2	15	20	0.25	0.3	5	0.49	0.29	0.21	233	138
3	15	50	0.25	0.3	5	0.25	0.17	0.13	192	131
4	5	20	0.25	0.3	5	1.06	0.61	0.38	279	161
5	50	20	0.25	0.3	5	0.25	0.16	0.11	227	145
6	500	20	0.25	0.3	5	0.18	0.11	0.08	225	138
7	15	20	0.25	0.0	5	0.77	0.52	0.37	208	141
8	15	20	0.25	0.6	5	0.33	0.13	0.10	330	130
9	15	20	0.25	0.9	5	0.27	0.03	0.02	1350	150
10	15	20	0.01	0.3	5	0.45	0.30	0.22	205	136
11	15	20	0.50	0.3	5	0.51	0.28	0.17	300	165
12	15	20	0.75	0.3	5	0.54	0.29	0.13	415	223
13	15	20	0.25	0.3	1	0.40	0.29	0.21	190	138
14	15	20	0.25	0.3	10	0.68	0.29	0.21	324	138
15	15	20	0.25	0.3	25	1.42	0.29	0.20	710	145

Standard Errors

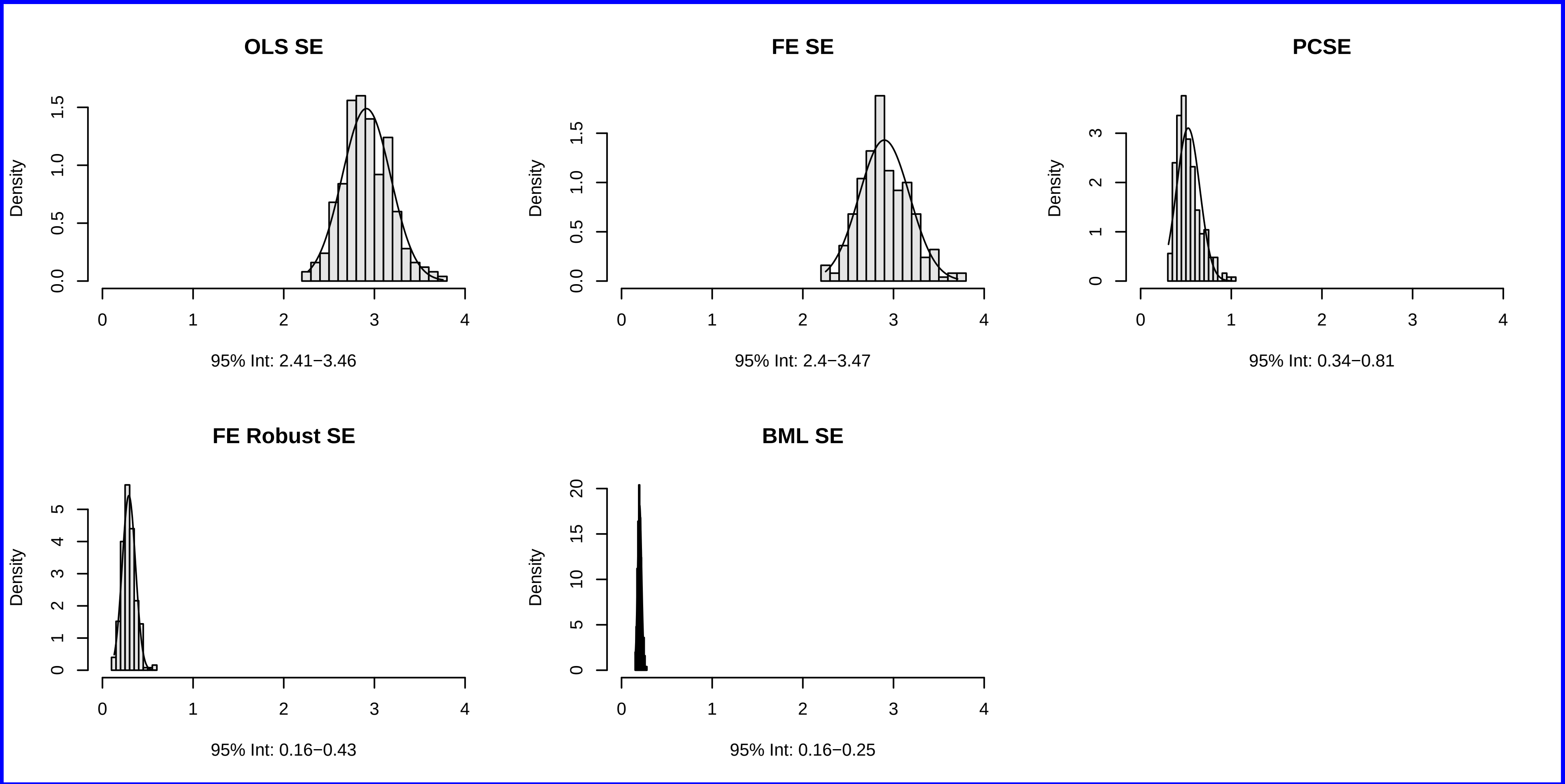

Figure 2: *Histogram of estimated S.E.s for Ordinary Least Squares, Fixed Effects, and Bayesian multilevel models. N(sims) = 250, N = 15, T = 20, Cor = 0.25, Het = 0.3, A.Scale = 5.*

Optimism

- So the BML standard errors are the smallest.

Optimism

- So the BML standard errors are the smallest.
- But is this false confidence? No.

$$\text{Optimism} = 100 \cdot \frac{\sqrt{\sum_{l=1}^{nsims}(\beta^{(l)} - \bar{\beta})^2}}{\sqrt{\sum_{l=1}^{nsims}(s.e.(\beta^{(l)}))^2}} \tag{8}$$

Optimism

	N	T	Cor	Het	A.Sc	OLS Opt	FE Opt	PCSE Opt	FE Robust Opt	BML Opt
1	15	5	0.25	0.3	5	22	12	81	105	108
2	15	20	0.25	0.3	5	17	10	90	95	102
3	15	50	0.25	0.3	5	13	9	85	96	104
4	5	20	0.25	0.3	5	21	12	95	104	92
5	50	20	0.25	0.3	5	16	10	90	101	105
6	500	20	0.25	0.3	5	16	10	95	102	107
7	15	20	0.25	0.0	5	26	18	94	95	102
8	15	20	0.25	0.6	5	11	5	81	95	102
9	15	20	0.25	0.9	5	9	1	73	95	101
10	15	20	0.01	0.3	5	15	10	95	100	102
11	15	20	0.50	0.3	5	18	10	86	94	102
12	15	20	0.75	0.3	5	19	10	82	94	101
13	15	20	0.25	0.3	1	14	10	87	95	102
14	15	20	0.25	0.3	10	23	10	93	95	102
15	15	20	0.25	0.3	25	44	10	95	95	102

Conclusion

Conclusion

- Computationally intensive, but ...

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties
  - ⋆ Group-level predictors to explain group differences

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties
  - ⋆ Group-level predictors to explain group differences
  - ⋆ Additional levels of structure (example: CDs and states)

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties
  - ⋆ Group-level predictors to explain group differences
  - ⋆ Additional levels of structure (example: CDs and states)
  - ⋆ Sluggish predictors can be estimated simultaneously with unit effects

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties
  - ⋆ Group-level predictors to explain group differences
  - ⋆ Additional levels of structure (example: CDs and states)
  - ⋆ Sluggish predictors can be estimated simultaneously with unit effects
  - ⋆ More general

Conclusion

- Computationally intensive, but ...
- Compared with pooled OLS and FE estimators with variety of SE's
  - ⋆ BML is equally unbiased on average
  - ⋆ Considerably more efficient
  - ⋆ Reports higher quality standard errors
- And has additional advantages
  - ⋆ Better small-sample properties
  - ⋆ Group-level predictors to explain group differences
  - ⋆ Additional levels of structure (example: CDs and states)
  - ⋆ Sluggish predictors can be estimated simultaneously with unit effects
  - ⋆ More general

Future Work

Future Work

- Serial correlation

Future Work

- Serial correlation
- Varying slopes

Future Work

- Serial correlation
- Varying slopes
- Increase cross-level correlation between unit effects and X

Future Work

- Serial correlation
- Varying slopes
- Increase cross-level correlation between unit effects and X
- Test against well-known applications of TSCS data

Future Work

- Serial correlation
- Varying slopes
- Increase cross-level correlation between unit effects and X
- Test against well-known applications of TSCS data