GSERM - St. Gallen 2023
Analyzing Panel Data

June 16, 2023

Start with:

$$Y_i^* = \mathbf{X}_i\boldsymbol{\beta} + u_i$$

$$\begin{aligned} Y_i &= 0 \;\; \text{if} \;\; Y_i^* < 0 \\ Y_i &= 1 \;\; \text{if} \;\; Y_i^* \geq 0 \end{aligned}$$

So:

$$\begin{aligned} \Pr(Y_i = 1) &= \Pr(Y_i^* \geq 0) \\ &= \Pr(\mathbf{X}_i\boldsymbol{\beta} + u_i \geq 0) \\ &= \Pr(u_i \geq -\mathbf{X}_i\beta) \\ &= \Pr(u_i \leq \mathbf{X}_i\beta) \\ &= \int_{-\infty}^{\mathbf{X}_i\boldsymbol{\beta}} f(u)du \end{aligned}$$

"Standard logistic" PDF:

$$\Pr(u) \equiv \lambda(u) = \frac{\exp(u)}{[1 + \exp(u)]^2}$$

CDF:

$$\begin{aligned}\Lambda(u) &= \int \lambda(u) du \\ &= \frac{\exp(u)}{1 + \exp(u)} \\ &= \frac{1}{1 + \exp(-u)}\end{aligned}$$

Logistic $\rightarrow$ "Logit"

$$
\begin{aligned}
\Pr(Y_i = 1) &= \Pr(Y_i^* > 0) \\
&= \Pr(u_i \leq \mathbf{X}_i\boldsymbol{\beta}) \\
&= \Lambda(\mathbf{X}_i\boldsymbol{\beta}) \\
&= \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})}
\end{aligned}
$$

$$
\text{(equivalently)} = \frac{1}{1 + \exp(-\mathbf{X}_i\boldsymbol{\beta})}
$$

Likelihoods

$$L_i = \left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right)^{Y_i}\left[1-\left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right)\right]^{1-Y_i}$$

$$L = \prod_{i=1}^{N}\left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right)^{Y_i}\left[1-\left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right)\right]^{1-Y_i}$$

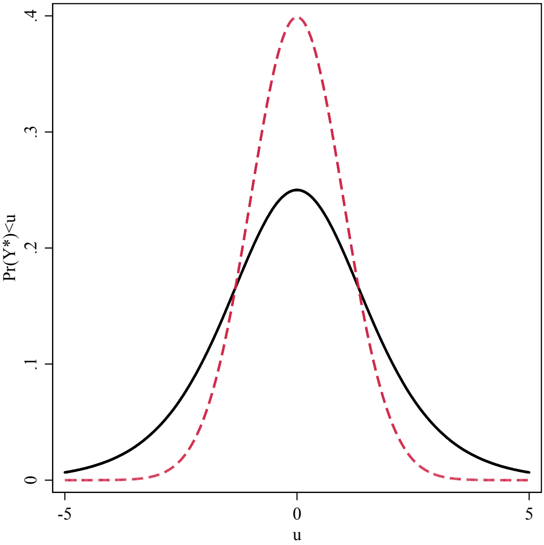
$$\begin{aligned}\ln L &= \sum_{i=1}^{N} Y_i \ln\left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right) + \\ &\quad (1-Y_i)\ln\left[1-\left(\frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i\boldsymbol{\beta})}\right)\right]\end{aligned}$$

Probit...

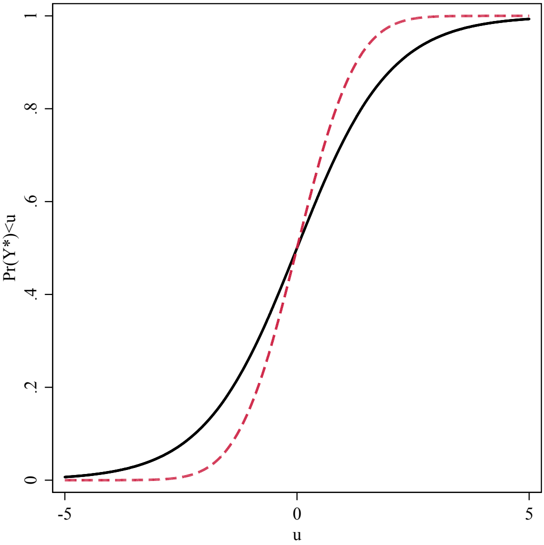
$$\Pr(u) \equiv \phi(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

$$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$$

Standard Normal and Logistic PDFs



Standard Normal and Logistic CDFs



$$\begin{aligned}\Pr(Y_i = 1) &= \Phi(\mathbf{X}_i\boldsymbol{\beta}) \\ &= \int_{-\infty}^{\mathbf{X}_i\boldsymbol{\beta}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\mathbf{X}_i\boldsymbol{\beta})^2}{2}\right) d\mathbf{X}_i\boldsymbol{\beta}\end{aligned}$$

$$L = \prod_{i=1}^{N} [\Phi(\mathbf{X}_i\boldsymbol{\beta})]^{Y_i} [1 - \Phi(\mathbf{X}_i\boldsymbol{\beta})]^{(1-Y_i)}$$

$$\ln L = \sum_{i=1}^{N} Y_i \ln\Phi(\mathbf{X}_i\boldsymbol{\beta}) + (1 - Y_i) \ln[1 - \Phi(\mathbf{X}_i\boldsymbol{\beta})]$$

Panel / TSCS: What Can Go Wrong?

Suppose:

$$
\begin{aligned}
X_{it} &= \rho_X \mathbf{X}_{it-1} + \nu_{it} \\
u_{it} &= \rho_u u_{it-1} + \epsilon_{it}
\end{aligned}
$$

For high values of ρ, logit/probit:

- $\hat{\beta}$s are consistent, but s.e.s are biased, inefficient (Poirier and Ruud 1988);
- $\rightarrow$ underestimate Var(β) by up to 50 percent (Beck and Katz 1997).

Unit Effects, Fixed

One-way unit effects:

$$Y_{it} = f(\mathbf{X}_{it}\beta + \alpha_i + u_{it})$$

for logit only, so:

$$\Pr(Y_{it} = 1) = \frac{\exp(\mathbf{X}_{it}\beta + \alpha_i)}{1 + \exp(\mathbf{X}_{it}\beta + \alpha_i)} \equiv \Lambda(\mathbf{X}_{it}\beta + \alpha_i)$$

<u>Incidental Parameters</u>:

- Nonlinearity $\rightarrow$ inconsistency in both $\hat{\alpha}$s <u>and</u> $\hat{\boldsymbol{\beta}}$.
- Anderson:

$$L^U = \prod_{i=1}^{N}\prod_{t=1}^{T} \Lambda(\mathbf{X}_{it} + \alpha_i)^{Y_{it}}[1 - \Lambda(\mathbf{X}_{it} + \alpha_i)]^{1-Y_{it}}$$

- Chamberlain:

$$L^C = \prod_{i=1}^{N} \Pr\left(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, ...Y_{iT} = y_{iT} \mid \sum_{t=1}^{T} Y_{it}\right)$$

Fixed-Effects (continued)

Intuition: Suppose we have $T = 2$. That means that:

- $\Pr(Y_{i1} = 0 \text{ and } Y_{i2} = 0 \mid \sum_T Y_{it} = 0) = 1.0$
- $\Pr(Y_{i1} = 1 \text{ and } Y_{i2} = 1 \mid \sum_T Y_{it} = 2) = 1.0$

and:

$$\Pr\left(Y_{i1} = 0 \text{ and } Y_{i2} = 1 \mid \sum_T Y_{it} = 1\right) = \frac{\Pr(0,1)}{\Pr(0,1) + \Pr(1,0)}$$

with a similar statement for $\Pr(Y_{i1} = 1 \text{ and } Y_{i2} = 0 \mid \sum_T Y_{it} = 1)$.

Points:

- Fixed effects = no estimates for $\boldsymbol{\beta}_b$
- Interpretation: per logit, but $\mid \hat{\alpha}_i$.
- BTSCS in IR: Green et al. (2001) v. B&K (2001).

Random Effects

Model is:

$$
\begin{aligned}
Y_{it}^* &= \mathbf{X}_{it}\beta + u_{it} \\
Y_{it} &= 0 \text{ if } Y_{it}^* \leq 0 \text{ ;} \\
&= 1 \text{ if } Y_{it}^* > 0
\end{aligned}
$$

with:

$$u_{it} = \alpha_i + \eta_{it}$$

with $\eta_{it} \sim$ i.i.d. N(0,1), and $\alpha_i \sim N(0, \sigma_\alpha^2)$. This implies:

$$\text{Var}(u_{it}) = 1 + \sigma_\alpha^2$$

and so:

$$\text{Corr}(u_{it}, u_{is},\ t \neq s) \equiv \rho = \frac{\sigma_\alpha^2}{1 + \sigma_\alpha^2}$$

which means that we can write $\sigma_\alpha^2 = \left(\frac{\rho}{1-\rho}\right)$.

Random Effects Variants

Probit:

$$
\begin{aligned}
L_i &= \text{Prob}(Y_{i1} = y_{i1},\ Y_{i2} = y_{i2},\ ...Y_{iT} = y_{iT}) \\
&= \int_{-\infty}^{X_{i1}\beta} \int_{-\infty}^{X_{i2}\beta} ... \int_{-\infty}^{X_{iT}\beta} \phi(u_{i1}, u_{i2}...u_{iT}) du_{iT}...du_{i2}\ du_{i1}
\end{aligned}
$$

Logit:

$$
\begin{aligned}
L_i &= \text{Prob}(Y_{i1} = y_{i1},\ Y_{i2} = y_{i2},\ ...Y_{iT} = y_{iT}) \\
&= \int_{-\infty}^{X_{i1}\beta} \int_{-\infty}^{X_{i2}\beta} ... \int_{-\infty}^{X_{iT}\beta} \lambda(u_{i1}, u_{i2}...u_{iT}) du_{iT}...du_{i2}\ du_{i1}
\end{aligned}
$$

Solution?

$$
\phi(u_{i1}, u_{i2}, ...u_{iT}) = \int_{-\infty}^{\infty} \phi(u_{i1}, u_{i2}, ...u_{iT} \mid \alpha_i)\phi(\alpha_i) d\alpha_i
$$

- $\hat{\rho} =$ proportion of the variance due to the α_is.
- Implementation: Gauss-Hermite quadrature or MCMC.
- Best with N large and T small.
- Critically requires $\text{Cov}(\mathbf{X}, \alpha) = 0$ (see notes re: Chamberlain's CRE Estimator).

Unit Effects in Practice - Some Simulations

Start with:

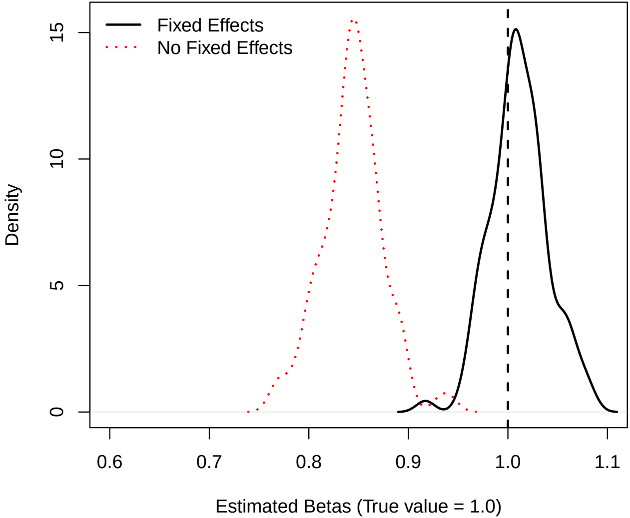
$$\begin{aligned} Y_{it}^* &= 0 + (1 \times X_{it}) + (1 \times D_{it}) + (1 \times \alpha_i) + u_{it} \\ Y_{it} \in \{0,1\} &= f(Y_{it}^*) \end{aligned}$$

where:

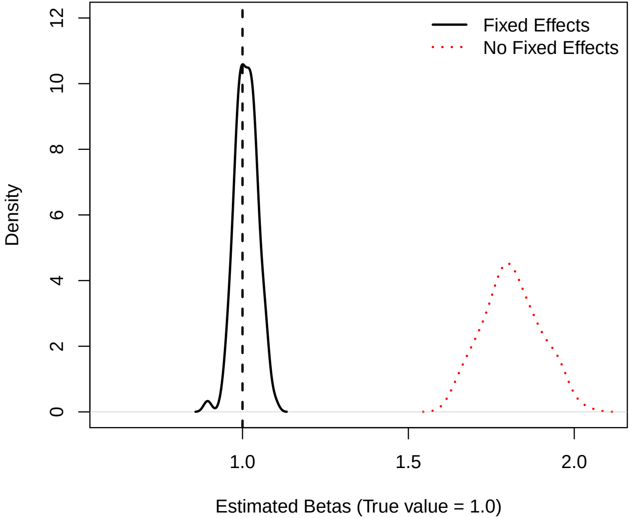
- $\alpha_i \sim N(0,1)$
- $X_{it} \sim N(0, \sigma_X^2)$
- $D_{it} \in \{0,1\}$
- $\text{Cov}(X_{it}, \alpha_i) = \{0, 0.69\}$
- $\text{Cov}(D_{it}, \alpha_i) = 0$
- $f(\cdot) = \{\text{logit, probit}\}$ (as appropriate)

and $N = T = 100$.

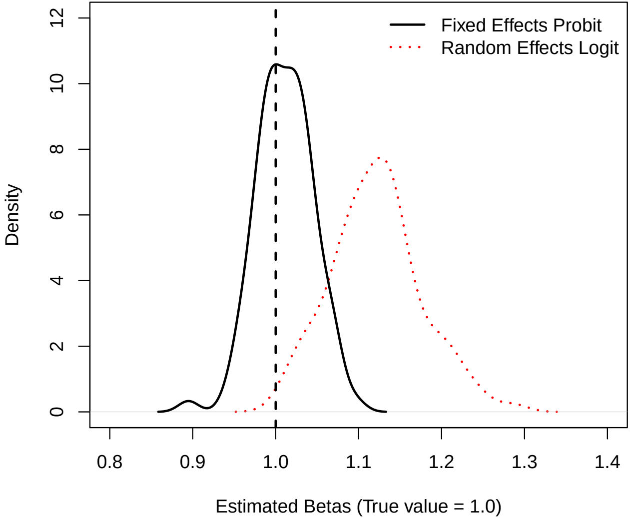
Logit $\hat{\beta}_X$s for $\mathrm{Cov}(X_{it}, \alpha_i) = 0$



Logit $\hat{\boldsymbol{\beta}}_X$s for $\text{Cov}(X_{it}, \alpha_i) \approx 0.69$



Logit $\hat{\boldsymbol{\beta}}_X$s for $\text{Cov}(X_{it}, \alpha_i) \approx 0.69$



R

- `pglm` (panel GLMs) (maximum likelihood + quadrature)
- `bife` (fixed-effects logit / probit only)
- `glmer` (general mixed-effects models, including RE)
- `glmmML` (via Gauss-Hermite quadrature)
- `MCMCpack` (`MCMChlogit`)
- Various user-generated functions (e.g., here).

Stata

- `xtprobit`, `xtlogit`, `xtcloglog`
- Plus `xttrans` (transition probabilities), `quadchk` (quadrature checking), `xtrho` / `xtrhoi` (estimation of within-unit covariances)

Example: WDI "Plus"

Data from the WDI plus POLITY and the UCDP:

- `ISO3` - The country's International Standards Organization (ISO) three-letter identification code.
- `Year` - The year that row of data applies to.
- `CivilWar` - Civil conflict indicator: 1 if there was a civil conflict in that country in that year; 0 otherwise. From UCDP.
- `OnsetCount` - The sum of new conflict episodes in that country / year. From UCDP.
- `LandArea` - Land area (sq. km).
- `PopMillions` - Popluation (in millions).
- `PopGrowth` - Population Growth (percent).
- `UrbanPopulation` - Urban Population (percent of total).
- `GDPPerCapita` - GDP per capita (constant 2010 $US).
- `GDPPerCapGrowth` - GDP Per Capita Growth (percent annual).
- `PostColdWar` - 1 if `Year` $>$ 1989, 0 otherwise.
- `POLITY` - The POLITY score of democracy/autocracy. Scaled so that 0 = most autocratic, 10 = most democratic.

$N = 216$, $\bar{T} = 61$, NT varies (due to missingness).

```
> describe(DF,skew=FALSE)

                vars     n      mean         sd    min        max       range       se
ISO3*              1 13607    108.49      62.35   1.00      216.0      215.00     0.53
Year*              2 13607     32.00      18.18   1.00       63.0       62.00     0.16
country*           3 13545    108.00      62.07   1.00      215.0      214.00     0.53
CivilWar           4  9052      0.13       0.34   0.00        1.0        1.00     0.00
OnsetCount         5  9394      0.05       0.24   0.00        4.0        4.00     0.00
LandArea           6 12728 611322.13 1764229.22   2.03 16389950.0 16389947.97 15637.77
PopMillions        7 13300     24.92     104.04   0.00     1412.4     1412.36     0.90
UrbanPopulation    8 13268     51.55      25.74   2.08      100.0       97.92     0.22
GDPPerCapita       9  9843  11874.12   18895.82 144.03   204190.2   204046.13   190.46
GDPPerCapGrowth   10  9818      1.93       6.17 -64.43      140.5      204.91     0.06
PostColdWar       11 13545      0.52       0.50   0.00        1.0        1.00     0.00
POLITY            12  8279      5.55       3.71   0.00       10.0       10.00     0.04
POLITYSquared     13  8279     44.57      40.24   0.00      100.0      100.00     0.44
```

Pooled Logit

```
> Logit<-glm(CivilWar~log(LandArea)+log(PopMillions)+
+              UrbanPopulation+log(GDPPerCapita)+
+              GDPPerCapGrowth+PostColdWar+POLITY+
+              POLITYSquared,data=DF,family="binomial")


> summary(Logit)

Call:
glm(formula = CivilWar ~ log(LandArea) + log(PopMillions) + UrbanPopulation +
    log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar + POLITY +
    POLITYSquared, family = "binomial", data = DF)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)       -1.49782    0.51881   -2.89  0.00389 **
log(LandArea)      0.01617    0.03242    0.50  0.61792
log(PopMillions)   0.65816    0.03675   17.91  < 2e-16 ***
UrbanPopulation    0.00792    0.00331    2.39  0.01668 *
log(GDPPerCapita) -0.43195    0.06004   -7.19  6.3e-13 ***
GDPPerCapGrowth   -0.04159    0.00649   -6.40  1.5e-10 ***
PostColdWar       -0.29512    0.08563   -3.45  0.00057 ***
POLITY             0.68401    0.06105   11.20  < 2e-16 ***
POLITYSquared     -0.06648    0.00578  -11.51  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5840.7  on 6998  degrees of freedom
Residual deviance: 4639.8  on 6990  degrees of freedom
  (6608 observations deleted due to missingness)
AIC: 4658

Number of Fisher Scoring iterations: 6
```

Fixed Effects

```
> FELogit<-bife(CivilWar~log(LandArea)+log(PopMillions)+
+              UrbanPopulation+log(GDPPerCapita)+
+              GDPPerCapGrowth+PostColdWar+POLITY+
+              POLITYSquared|ISO3,data=DF,model="logit")


> summary(FELogit)
binomial - logit link

CivilWar ~ log(LandArea) + log(PopMillions) + UrbanPopulation +
    log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar + POLITY +
    POLITYSquared | ISO3

Estimates:
                  Estimate Std. error z value Pr(> |z|)
log(LandArea)     -13.76753    8.17528   -1.68     0.092 .
log(PopMillions)    0.68167    0.29453    2.31     0.021 *
UrbanPopulation     0.01736    0.01242    1.40     0.162
log(GDPPerCapita)  -0.32466    0.17392   -1.87     0.062 .
GDPPerCapGrowth    -0.05224    0.00844   -6.19   6.0e-10 ***
PostColdWar        -0.22301    0.17875   -1.25     0.212
POLITY              0.71218    0.09359    7.61   2.8e-14 ***
POLITYSquared      -0.07382    0.00890   -8.29   < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

residual deviance= 2838,
null deviance= 4419,
n= 3970, N= 83

( 6608 observation(s) deleted due to missingness )
( 3029 observation(s) deleted due to perfect classification )

Number of Fisher Scoring Iterations: 6

Average individual fixed effect= 171.8
```

Alternative Fixed Effects (using `feglm`)

```
> FELogit2<-feglm(CivilWar~log(LandArea)+log(PopMillions)+UrbanPopulation+log(GDPPerCapita)+
+                   GDPPerCapGrowth+PostColdWar+POLITY+POLITYSquared|ISO3,data=DF,family="binomial")

NOTES: 6,608 observations removed because of NA values (LHS: 4,555, RHS: 6,608).
       77 fixed-effects (3,029 observations) removed because of only 0 (or only 1) outcomes.

> FELogit2

GLM estimation, family = binomial, Dep. Var.: CivilWar
Observations: 3,970
Fixed-effects: ISO3: 83
Standard-errors: Clustered (ISO3)

                   Estimate Std. Error t value    Pr(>|t|)
log(LandArea)     -13.76929    9.21285 -1.4946 0.135025561
log(PopMillions)    0.68167    0.75849  0.8987 0.368802391
UrbanPopulation     0.01736    0.03675  0.4724 0.636653205
log(GDPPerCapita)  -0.32466    0.41321 -0.7857 0.432038386
GDPPerCapGrowth    -0.05224    0.01279 -4.0830 0.000044449 ***
PostColdWar        -0.22301    0.48194 -0.4627 0.643556228
POLITY              0.71218    0.24823  2.8690 0.004117503 **
POLITYSquared      -0.07382    0.02448 -3.0161 0.002560480 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Log-Likelihood: -1,419.0   Adj. Pseudo R2: 0.317077
           BIC:  3,592.1     Squared Cor.: 0.401272
```

```
> RELogit<-pglm(CivilWar~log(LandArea)+log(PopMillions)+
+                 UrbanPopulation+log(GDPPerCapita)+
+                 GDPPerCapGrowth+PostColdWar+POLITY+
+                 POLITYSquared|ISO3,data=DF,family=binomial,
+                 effect="individual",model="random")
```

```
> summary(RELogit)
--------------------------------------------
Maximum Likelihood estimation
Newton-Raphson maximisation, 3 iterations
Return code 3: Last step could not find a value above the current.
Boundary of parameter space?
Consider switching to a more robust optimisation method temporarily.
Log-Likelihood: -1659
10  free parameters
Estimates:
                   Estimate Std. error t value Pr(> t)
(Intercept)        0.8269498  0.6811125    1.21    0.22
log(LandArea)      0.0000976  0.0479679    0.00    1.00
log(PopMillions)   0.6302824  0.1045877    6.03 1.7e-09 ***
UrbanPopulation   -0.0011367  0.0010206   -1.11    0.27
log(GDPPerCapita) -0.7120370  0.0806616   -8.83 < 2e-16 ***
GDPPerCapGrowth   -0.0499556  0.0076386   -6.54 6.2e-11 ***
PostColdWar       -0.0071774  0.1213515   -0.06    0.95
POLITY             0.8713968  0.0964073    9.04 < 2e-16 ***
POLITYSquared     -0.0949833  0.0097413   -9.75 < 2e-16 ***
sigma              2.3422224  0.0878870   26.65 < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
--------------------------------------------
```

Nice Table...

Models of Civil War

	Logit	FE Logit	FEs+Robust	RE Logit
Intercept	-1.50^{*}			0.83
	(0.52)			(0.68)
ln(Land Area)	0.02	-13.77	-13.77	0.00
	(0.03)	(8.18)	(9.21)	(0.05)
ln(Population)	0.66^{*}	0.68^{*}	0.68	0.63^{*}
	(0.04)	(0.29)	(0.76)	(0.10)
Urban Population	0.01^{*}	0.02	0.02	-0.00
	(0.00)	(0.01)	(0.04)	(0.00)
ln(GDP Per Capita)	-0.43^{*}	-0.32	-0.32	-0.71^{*}
	(0.06)	(0.17)	(0.41)	(0.08)
GDP Growth	-0.04^{*}	-0.05^{*}	-0.05^{*}	-0.05^{*}
	(0.01)	(0.01)	(0.01)	(0.01)
Post-Cold War	-0.30^{*}	-0.22	-0.22	-0.01
	(0.09)	(0.18)	(0.48)	(0.12)
POLITY	0.68^{*}	0.71^{*}	0.71^{*}	0.87^{*}
	(0.06)	(0.09)	(0.25)	(0.10)
POLITY Squared	-0.07^{*}	-0.07^{*}	-0.07^{*}	-0.09^{*}
	(0.01)	(0.01)	(0.02)	(0.01)
Estimated $\hat{\sigma}$				2.34^{*}
				(0.09)
AIC	4657.80			3337.27
BIC	4719.48			
Log Likelihood	-2319.90	-1419.03	-1419.03	-1658.63
Deviance	4639.80	2838.06	2838.06	
Num. obs.	6999	3970	3970	
Num. groups: ISO3			83	
Pseudo R^2			0.32	

$^{*}p < 0.05$

Models For Event Counts

Properties:

- Discrete / integer-values
- Non-negative
- "Cumulative"

Motivation:

$$\text{Arrival Rate} = \lambda$$

$$\Pr(\text{Event})_{t,t+h} = \lambda h$$

$$\Pr(\text{No Event})_{t,t+h} = 1 - \lambda h$$

$$\begin{aligned}\Pr(Y_t = y) &= \frac{\exp(-\lambda h)\lambda h^y}{y!} \\ &= \frac{\exp(-\lambda)\lambda^y}{y!}\end{aligned}$$

Poisson: Assumptions and Motivations

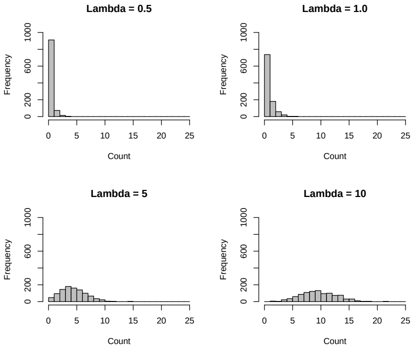
- No Simultaneous Events
- Constant Arrival Rate
- Independent Event Arrivals

Another motivation: For M independent Bernoulli trials with (sufficiently small) probability of success π and where $M\pi \equiv \lambda > 0$,

$$
\begin{aligned}
\Pr(Y_i = y) &= \lim_{M \to \infty} \left[\binom{M}{y} \left(\frac{\lambda}{M} \right)^y \left(1 - \frac{\lambda}{M} \right)^{M-y} \right] \\
&= \frac{\lambda^y \exp(-\lambda)}{y!}
\end{aligned}
$$

Poisson: Characteristics

- Discrete
- $\mathrm{E}(Y) = \mathrm{Var}(Y) = \lambda$
- Is not preserved under affine transformations...
- For $X \sim \text{Poisson}(\lambda_X)$ and $Y \sim \text{Poisson}(\lambda_Y)$, $Z = X + Y \sim \text{Poisson}(\lambda_{X+Y})$ <u>iff</u> X and Y are <u>independent</u> but
- ...same is not true for differences.
- $\lambda \to \infty \iff Y \sim N$



Suppose

$$\mathrm{E}(Y_i) \equiv \lambda_i = \exp(\mathbf{X}_i\boldsymbol{\beta})$$

then

$$\Pr(Y_i = y|\mathbf{X}_i, \boldsymbol{\beta}) = \frac{\exp[-\exp(\mathbf{X}_i\boldsymbol{\beta})][\exp(\mathbf{X}_i\boldsymbol{\beta})]^y}{y!}$$

with likelihood:

$$L = \prod_{i=1}^{N} \frac{\exp[-\exp(\mathbf{X}_i\boldsymbol{\beta})][\exp(\mathbf{X}_i\boldsymbol{\beta})]^{Y_i}}{Y_i!}$$

and log-likelihood:

$$\ln L = \sum_{i=1}^{N} [-\exp(\mathbf{X}_i\boldsymbol{\beta}) + Y_i\mathbf{X}_i\boldsymbol{\beta} - \ln(Y_i!)]$$

The Poisson model:

$$Y_{it} \sim \text{Poisson}(\mu_{it} = \alpha_i \lambda_{it})$$

with $\lambda_{it} = \exp(\mathbf{X}_{it}\boldsymbol{\beta})$ implies:

$$\begin{aligned}
\text{E}(Y_{it} \mid \mathbf{X}_{it}, \alpha_i) &= \mu_{it} \\
&= \alpha_i \exp(\mathbf{X}_{it}\boldsymbol{\beta}) \\
&= \exp(\delta_i + \mathbf{X}_{it}\boldsymbol{\beta})
\end{aligned}$$

where $\delta_i = \ln(\alpha_i)$.

<u>Fixed-Effects Poisson</u>:

- ...has no "incidental parameters" problem (see e.g. Cameron and Trivedi, pp. 281-2)
- This means "brute force" approach works
- Fitted via `glmmML` in R, `xtpoisson` (and `xtnbreg`) in Stata

Random-Effects Models

The Poisson with random effects is:

$$\begin{aligned}
\Pr(Y_{i1} = y_{i1}, ... Y_{iT} = y_{iT}) &= \int_0^\infty \Pr(Y_{i1} = y_{i1}, ... Y_{iT} = y_{iT}) f(\alpha_i) d\alpha_i \\
&= \int_0^\infty \left[\prod_{t=1}^{T} \Pr(Y_{it} \mid \alpha_i) \right] f(\alpha_i) d\alpha_i
\end{aligned}$$

- Simplest to assume $\alpha_i \sim \Gamma(\theta)$
- Yields a model with $\mathrm{E}(Y_{it}) = \lambda_{it}$ and $\mathrm{Var}(Y_{it}) = \lambda_{it} + \frac{\lambda_{it}^2}{\theta}$
- Via `glmmML` or `glmer` in R, or `xtpois, re` in Stata
- $\exists$ random effects negative binomial too...

R:

- Tobit = `censReg` (in **censReg**)
- Poisson (random effects) = `glmmML` in **glmmML** or `glmer` in **lme4**
- Poisson (fixed effects) = `glmmML` or "brute force"

Stata:

- Tobit = `xttobit` (`re` only)
- Poisson / negative binomial = `xtpoisson`, `xtnbreg` (both with `fe`, `re` options)

Conflict Onsets: Pooled Poisson

```
> xtabs(~DF$OnsetCount)
DF$OnsetCount
   0    1    2    3    4
8981  375   30    7    1

> Poisson<-glm(OnsetCount~log(LandArea)+log(PopMillions)+UrbanPopulation+log(GDPPerCapita)+
+              GDPPerCapGrowth+PostColdWar+POLITY+POLITYSquared,data=DF,family="poisson")

> summary(Poisson)

Call:
glm(formula = OnsetCount ~ log(LandArea) + log(PopMillions) +
    UrbanPopulation + log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar +
    POLITY + POLITYSquared, family = "poisson", data = DF)

Coefficients:
                  Estimate Std. Error z value    Pr(>|z|)
(Intercept)       -2.46627    0.71420   -3.45     0.00055 ***
log(LandArea)      0.07460    0.04698    1.59     0.11232
log(PopMillions)   0.42366    0.04575    9.26     < 2e-16 ***
UrbanPopulation    0.00612    0.00469    1.31     0.19129
log(GDPPerCapita) -0.42730    0.07996   -5.34 0.000000091 ***
GDPPerCapGrowth   -0.03720    0.00661   -5.62 0.000000019 ***
PostColdWar        0.26711    0.12019    2.22     0.02626 *
POLITY             0.32677    0.08290    3.94 0.000080877 ***
POLITYSquared     -0.03607    0.00793   -4.55 0.000005383 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2387.1  on 6998  degrees of freedom
Residual deviance: 1946.9  on 6990  degrees of freedom
  (6608 observations deleted due to missingness)
AIC: 2699

Number of Fisher Scoring iterations: 6
```

```
> FEPoisson<-pglm(OnsetCount~log(LandArea)+log(PopMillions)+
+                UrbanPopulation+log(GDPPerCapita)+
+                GDPPerCapGrowth+PostColdWar+POLITY+
+                POLITYSquared,data=DF,family="poisson",
+                effect="individual",model="within")


> summary(FEPoisson)
--------------------------------------------
Maximum Likelihood estimation
Newton-Raphson maximisation, 3 iterations
Return code 8: successive function values within relative tolerance limit (reltol)
Log-Likelihood: -1018
8  free parameters
Estimates:
                 Estimate Std. error t value    Pr(> t)
log(LandArea)    -2.81096    2.86598   -0.98    0.32669
log(PopMillions)  0.63667    0.31900    2.00    0.04595 *
UrbanPopulation  -0.04563    0.01341   -3.40    0.00067 ***
log(GDPPerCapita) -0.10200    0.14542   -0.70    0.48306
GDPPerCapGrowth  -0.02865    0.00682   -4.20 0.00002673 ***
PostColdWar       0.47487    0.19574    2.43    0.01526 *
POLITY            0.52050    0.10801    4.82 0.00000144 ***
POLITYSquared    -0.05323    0.01062   -5.01 0.00000054 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
--------------------------------------------
```

Equivalent Fixed Effects Poisson (using `feglm`)

```
> FEPoisson<-feglm(OnsetCount~log(LandArea)+log(PopMillions)+UrbanPopulation+
+                  log(GDPPerCapita)+GDPPerCapGrowth+PostColdWar+POLITY+
+                  POLITYSquared|ISO3,data=DF,family="poisson")

NOTES: 6,608 observations removed because of NA values (LHS: 4,213, RHS: 6,608).
       67 fixed-effects (2,502 observations) removed because of only 0 outcomes.

> summary(FEPoisson,cluster="ISO3")

GLM estimation, family = poisson, Dep. Var.: OnsetCount
Observations: 4,497
Fixed-effects: ISO3: 93
Standard-errors: Clustered (ISO3)
                  Estimate Std. Error t value     Pr(>|t|)
log(LandArea)     -2.81096   3.679400 -0.7640 0.4448843443
log(PopMillions)   0.63667   0.343155  1.8553 0.0635490951 .
UrbanPopulation   -0.04563   0.019612 -2.3265 0.0199895986 *
log(GDPPerCapita) -0.10200   0.153751 -0.6634 0.5070840537
GDPPerCapGrowth   -0.02865   0.006643 -4.3133 0.0000160819 ***
PostColdWar        0.47487   0.297378  1.5969 0.1102958347
POLITY             0.52050   0.111801  4.6556 0.0000032305 ***
POLITYSquared     -0.05323   0.011664 -4.5632 0.0000050385 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Log-Likelihood: -1,154.2   Adj. Pseudo R2: 0.095024
           BIC:  3,158.0     Squared Cor.: 0.16378
```

Random Effects Poisson

```
> REPoisson<-pglm(OnsetCount~log(LandArea)+log(PopMillions)+UrbanPopulation+
+                  log(GDPPerCapita)+GDPPerCapGrowth+PostColdWar+POLITY+
+                  POLITYSquared,data=DF,family="poisson",effect="individual",
+                  model="random")

> summary(REPoisson)

--------------------------------------------
Maximum Likelihood estimation
Newton-Raphson maximisation, 4 iterations
Return code 8: successive function values within relative tolerance limit (reltol)
Log-Likelihood: -1289
10  free parameters
Estimates:
                 Estimate Std. error t value     Pr(> t)
(Intercept)      -3.69707    1.04333   -3.54     0.00039 ***
log(LandArea)     0.05669    0.07318    0.77     0.43856
log(PopMillions)  0.44479    0.08006    5.56 0.000000028 ***
UrbanPopulation  -0.00563    0.00639   -0.88     0.37854
log(GDPPerCapita) -0.19570   0.10268   -1.91     0.05666 .
GDPPerCapGrowth  -0.03401    0.00680   -5.00 0.000000574 ***
PostColdWar       0.28947    0.12897    2.24     0.02480 *
POLITY            0.47485    0.09588    4.95 0.000000732 ***
POLITYSquared    -0.05263    0.00929   -5.66 0.000000015 ***
sigma             1.69967    0.41207    4.12 0.000037125 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
--------------------------------------------
```

Negative Binomial: Fixed Effects

```
> FENegBin2<-fenegbin(OnsetCount~log(LandArea)+log(PopMillions)+
+                  UrbanPopulation+log(GDPPerCapita)+
+                  GDPPerCapGrowth+PostColdWar+POLITY+
+                  POLITYSquared|ISO3,data=DF)

NOTES: 6,608 observations removed because of NA values (LHS: 4,213, RHS: 6,608).
       67 fixed-effects (2,502 observations) removed because of only 0 outcomes.

Very high value of theta (10000). There is no sign of overdispersion, you may
consider a Poisson model.

Warning message:
[femlm]: The information matrix is singular: presence of collinearity.
```

Table!

Panel Event Count Models

	Poisson	FE Poisson	RE Poisson
Intercept	-2.47^*		-3.70^*
	(0.71)		(1.04)
ln(Land Area)	0.07	-2.81	0.06
	(0.05)	(2.87)	(0.07)
ln(Population)	0.42^*	0.64^*	0.44^*
	(0.05)	(0.32)	(0.08)
Urban Population	0.01	-0.05^*	-0.01
	(0.00)	(0.01)	(0.01)
ln(GDP Per Capita)	-0.43^*	-0.10	-0.20
	(0.08)	(0.15)	(0.10)
GDP Growth	-0.04^*	-0.03^*	-0.03^*
	(0.01)	(0.01)	(0.01)
Post-Cold War	0.27^*	0.47^*	0.29^*
	(0.12)	(0.20)	(0.13)
POLITY	0.33^*	0.52^*	0.47^*
	(0.08)	(0.11)	(0.10)
POLITY Squared	-0.04^*	-0.05^*	-0.05^*
	(0.01)	(0.01)	(0.01)
Estimated $\hat{\sigma}$			1.70^*
			(0.41)
AIC	2699.12	2052.26	2598.61
BIC	2760.80		
Log Likelihood	-1340.56	-1018.13	-1289.31
Deviance	1946.94		
Num. obs.	6999		

$^*p < 0.05$

Wrap-Up: Some Useful Packages

- `pglm`
  - Workhorse package for panel (FE, RE, BE) GLMs
  - Binary + ordered logit/probit, Poisson / negative binomial
  - Discussed + used extensively in Croissant and Millo (2018) *Panel Data Econometrics with R*
  - The one thing it won't (apparently) do is fixed-effects, binary-response models...
- `fixest`
  - Fast / efficient fitting of FE models
  - Fits linear models, logit, Poisson, and negative binomial
  - Includes easy coefficient plots & tables; simple multi-threading; built-in "robust" S.E.s
- `alpaca`
  - Fast / efficient fitting of GLMs with high-dimensional fixed effects
  - *Includes bias correction for incidental parameters after binary-response models*
  - Also includes useful panel data simulation routines + average partial effects

Generalized Estimating Equations

Linear-normal model is:

$$Y_i = \mu_i + u_i$$

with:

$$\mu_i = \mathbf{X}_i \boldsymbol{\beta}.$$

Generalize:

$$g(\mu_i) = \mathbf{X}_i \boldsymbol{\beta}$$

and:

$$Y_i \sim \text{i.i.d.}\, F[\mu_i, \mathbf{V}_i].$$

GLM Estimation

"Score" equations:

$$\mathbf{U}(\beta) = \sum_{i=1}^{N} \mathbf{D}_i' \mathbf{V}_i^{-1} [Y_i - \mu_i] = \mathbf{0}.$$

with:

- $\mathbf{D}_i = \frac{\partial \mu_i}{\partial \beta}$,
- $\mathbf{V}_i = \frac{h(\mu_i)}{\phi}$, and
- $(Y_i - \mu_i) \approx$ a "residual."
- Known as "quasi-likelihood" (e.g. Wedderburn 1974 *Biometrika*).

Now suppose:

$$Y_{it} = \mu_{it} + u_{it}$$

where

- $i \in \{1, ... N\}$ are i.i.d. "units,"
- $t \in \{1, ... T\}$, $T > 1$ are "time points,"
- we want $g(\mu_{it}) = \mathbf{X}_{it}\boldsymbol{\beta}$.

Key issue: Accounting for (conditional) dependence in Y over time.

Full joint distributions over T are hard. But...

Define:

$$\underset{T\times T}{\mathbf{R}_i(\boldsymbol{\alpha})} = \begin{pmatrix} 1.0 & \alpha_{12} & \cdots & \alpha_{1,T} \\ \alpha_{21} & 1.0 & \cdots & \alpha_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{T,1} & \cdots & \alpha_{T,T-1} & 1.0 \end{pmatrix},$$

$\rightarrow$ "working correlation" matrix.

- Completely defined by $\boldsymbol{\alpha}$,
- Structure specified by the analyst.

Liang and Zeger (1986): We can decompose the variance of Y_{it} as:

$$\mathbf{V}_i = \text{diag}(\mathbf{V}_i^{\frac{1}{2}})\,\mathbf{R}_i(\boldsymbol{\alpha})\,\text{diag}(\mathbf{V}_i^{\frac{1}{2}})$$

With a standard GLM assumption about the mean and variance, this is:

$$\mathbf{V}_i = \frac{(\mathbf{A}_i^{\frac{1}{2}})\,\mathbf{R}_i(\boldsymbol{\alpha})\,(\mathbf{A}_i^{\frac{1}{2}})}{\phi}$$

where

$$\underset{T\times T}{\mathbf{A}_i} = \begin{pmatrix} h(\mu_{i1}) & 0 & \cdots & 0 \\ 0 & h(\mu_{i2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & h(\mu_{iT}) \end{pmatrix}$$

$\mathbf{V}_i = \text{Var}(Y_{it}|\mathbf{X}_{it}, \beta)$ has two parts:

- $\mathbf{A}_i =$ <u>unit-level</u> variation,
- $\mathbf{R}_i(\alpha) =$ within-unit <u>temporal</u> variation.

Specifying $\mathbf{R}_i(\boldsymbol{\alpha})$

Independent: $$\mathbf{R}_i(\boldsymbol{\alpha}) = \begin{pmatrix} 1.0 & 0 & \cdots & 0 \\ 0 & 1.0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1.0 \end{pmatrix}$$

- Assumes no within-unit temporal correlation.
- Equivalent to GLM on pooled data.

Exchangeable: $$\mathbf{R}_i(\boldsymbol{\alpha}) = \begin{pmatrix} 1.0 & \alpha & \cdots & \alpha \\ \alpha & 1.0 & \cdots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \cdots & \alpha & 1.0 \end{pmatrix}$$

- One free parameter in $\mathbf{R}_i(\boldsymbol{\alpha})$ ($\alpha_{ts} = \alpha \ \forall \ t \neq s$)
- Temporal correlation within units is constant across time points.
- Akin (in some respects) to a random-effects model...

Specifying $\mathbf{R}_i(\boldsymbol{\alpha})$

AR(p) (e.g., AR(1)):

$$\mathbf{R}_i(\boldsymbol{\alpha}) = \begin{pmatrix} 1.0 & \alpha & \alpha^2 & \cdots & \alpha^{T-1} \\ \alpha & 1.0 & \alpha & \cdots & \alpha^{T-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha^{T-1} & \cdots & \alpha^2 & \alpha & 1.0 \end{pmatrix}$$

- One free parameter in $\mathbf{R}_i(\boldsymbol{\alpha})$ ($\alpha_{ts} = \alpha^{|t-s|} \; \forall \; t \neq s$).
- Conditional within-unit correlation an exponential function of the lag.

Stationary(p):

$$\mathbf{R}_i(\boldsymbol{\alpha}) = \begin{pmatrix} 1.0 & \alpha_1 & \cdots & \alpha_p & 0 & \cdots & 0 \\ \alpha_1 & 1.0 & \alpha_1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \alpha_p & \cdots & \alpha_1 & 1.0 \end{pmatrix}$$

- AKA "banded," or "*p*-dependent."
- $p \leq T - 1$ free parameters in $\mathbf{R}_i(\boldsymbol{\alpha})$.
- Conditional within-unit correlation an exponential function of the lag, up to lag p, and zero thereafter.

Specifying $\mathbf{R}_i(\boldsymbol{\alpha})$

Unstructured: $$\mathbf{R}_i(\boldsymbol{\alpha}) = \begin{pmatrix} 1.0 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1,T-1} \\ \alpha_{12} & 1.0 & \alpha_{23} & \cdots & \alpha_{2,T-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{1,T-1} & \alpha_{2,T-1} & \cdots & \alpha_{T-1,T-1} & 1.0 \end{pmatrix}$$

- $\frac{T(T-1)}{2}$ free parameters in $\mathbf{R}_i(\boldsymbol{\alpha})$.
- Conditional within-unit correlation is completely data-dependent.

Score equations:

$$\boldsymbol{U}_{GEE}(\boldsymbol{\beta}_{GEE}) = \sum_{i=1}^{N} \mathbf{D}_i' \left[\frac{(\mathbf{A}_i^{\frac{1}{2}})\,\mathbf{R}_i(\boldsymbol{\alpha})\,(\mathbf{A}_i^{\frac{1}{2}})}{\phi} \right]^{-1} [Y_i - \mu_i] = \mathbf{0}$$

Two-step estimation:

- For fixed values of $\boldsymbol{\alpha}_s$ and ϕ_s at iteration s, use Newton scoring to estimate $\hat{\boldsymbol{\beta}}_s$,
- Use $\hat{\boldsymbol{\beta}}_s$ to calculate standardized residuals $(Y_i - \hat{\mu}_i)_s$, from which consistent estimates of $\boldsymbol{\alpha}_{s+1}$ and ϕ_{s+1} can be estimated.

Liang & Zeger (1986):

$$\hat{\boldsymbol{\beta}}_{GEE} \underset{N\to\infty}{\sim} \boldsymbol{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma}).$$

For $\hat{\boldsymbol{\Sigma}}$, two options:

$$\hat{\boldsymbol{\Sigma}}_{\text{Model}} = N \left(\sum_{i=1}^{N} \hat{\boldsymbol{D}}_i' \hat{\boldsymbol{V}}_i^{-1} \hat{\boldsymbol{D}}_i \right)$$

$$\hat{\boldsymbol{\Sigma}}_{\text{Robust}} = N \left(\sum_{i=1}^{N} \hat{\boldsymbol{D}}_i' \hat{\boldsymbol{V}}_i^{-1} \hat{\boldsymbol{D}}_i \right)^{-1} \left(\sum_{i=1}^{N} \hat{\boldsymbol{D}}_i' \hat{\boldsymbol{V}}_i^{-1} \hat{\boldsymbol{S}}_i \hat{\boldsymbol{V}}_i^{-1} \hat{\boldsymbol{D}}_i \right) \left(\sum_{i=1}^{N} \hat{\boldsymbol{D}}_i' \hat{\boldsymbol{V}}_i^{-1} \hat{\boldsymbol{D}}_i \right)^{-1}$$

where $\hat{\boldsymbol{S}}_i = (Y_i - \hat{\mu}_i)(Y_i - \hat{\mu}_i)'$.

Inference (aka, magic!)

- $\hat{\Sigma}_{\text{Model}}$
  - Requires that $\mathbf{R}_i(\boldsymbol{\alpha})$ be "correct" for consistency.
  - Is slightly more efficient than $\hat{\boldsymbol{\Sigma}}_{\text{Robust}}$ if so.

- $\hat{\Sigma}_{\text{Robust}}$
  - Is consistent *even if* $\mathbf{R}_i(\boldsymbol{\alpha})$ *is misspecified*.
  - Is slightly less efficient than $\hat{\boldsymbol{\Sigma}}_{\text{Model}}$ if $\mathbf{R}_i(\boldsymbol{\alpha})$ is correct.

Moral: Use $\hat{\Sigma}_{\textbf{Robust}}$.

GEEs:

- Are a straightforward variation on GLMs, and so
- Can be applied to a range of data types (continuous, binary, count, proportions, etc.),
- Yield robustly consistent point estimates of βs,
- Account for within-unit correlation in an informed way, but also
- Yield consistent inferences even if that correlation is misspecified.

Practical Issues: Model Interpretation

- In general, GEEs = GLMs.
- GEEs are *marginal* models, so:
  - $\hat{\beta}$s have an interpretation as average / total effects.
  - Estimates / effect sizes generally be smaller than conditional (e.g. fixed/random) effects models.
  - E.g., for logit, $\hat{\boldsymbol{\beta}}_M \approx \frac{\hat{\boldsymbol{\beta}}_C}{\sqrt{1+0.35\sigma_\eta^2}}$, where $\sigma_\eta^2 > 0$ is the variance of the unit effects.
  - See (e.g.) Gardiner et al. (2009) or Koper and Manseau (2009) for expositions.

Practical Issues: Specifying $\mathbf{R}_i(\boldsymbol{\alpha})$

- Has been called "more art than science."
- Pointers:
  - Choose based on *substance* of the problem.
  - Remember that $\mathbf{R}_i(\boldsymbol{\alpha})$ is conditional on $\mathbf{X}$, $\hat{\boldsymbol{\beta}}$.
  - Consider unstructured when T is small and N large.
  - Try different ones, and compare.
- In general, it shouldn't matter terribly much...

GEEs: Software

Software	Command(s)/Package(s)
R	gee / geepack / geeM / multgeeB / orth / repolr
Stata	xtgee / xtlogit / xtprobit / xtpois / etc.
SAS	genmod (w/ repeated)

GEEs: Software Tips

- Generally follow GLMs (specify "family" + "link")
- Certain combinations not possible/recommended
- Estimation: Fisher scoring, MLE, etc. (MCMC?)

From the `geepack` manual:

Warning

Use "unstructured" correlation structure only with great care. (It may cause R to crash).

Civil War Redux... GEE: Independence

```
> GEE.ind<-geeglm(CivilWar~log(LandArea)+log(PopMillions)+UrbanPopulation+
+               log(GDPPerCapita)+GDPPerCapGrowth+PostColdWar+POLITY+
+               POLITYSquared,data=DF,id=ISO3,family="binomial",
+               corstr="independence")

> summary(GEE.ind)

Call:
geeglm(formula = CivilWar ~ log(LandArea) + log(PopMillions) +
    UrbanPopulation + log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar +
    POLITY + POLITYSquared, family = "binomial", data = DF, id = ISO3,
    corstr = "independence")

 Coefficients:
                  Estimate  Std.err  Wald Pr(>|W|)
(Intercept)       -1.49782  2.05999  0.53   0.4672
log(LandArea)      0.01617  0.12454  0.02   0.8967
log(PopMillions)   0.65816  0.15550 17.92 0.000023 ***
UrbanPopulation    0.00792  0.01428  0.31   0.5794
log(GDPPerCapita) -0.43195  0.25246  2.93   0.0871 .
GDPPerCapGrowth   -0.04159  0.01310 10.08   0.0015 **
PostColdWar       -0.29512  0.26142  1.27   0.2589
POLITY             0.68401  0.21059 10.55   0.0012 **
POLITYSquared     -0.06648  0.01936 11.79   0.0006 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation structure = independence
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.805   0.313
Number of clusters:   160  Maximum cluster size: 57
```

GEE: Exchangeable

```
> GEE.exc<-geeglm(CivilWar~log(LandArea)+log(PopMillions)+UrbanPopulation+
+                   log(GDPPerCapita)+GDPPerCapGrowth+PostColdWar+POLITY+
+                   POLITYSquared,data=DF,id=ISO3,family="binomial",
+                 corstr="exchangeable")

> summary(GEE.exc)

Call:
geeglm(formula = CivilWar ~ log(LandArea) + log(PopMillions) +
    UrbanPopulation + log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar +
    POLITY + POLITYSquared, family = "binomial", data = DF, id = ISO3,
    corstr = "exchangeable")

 Coefficients:
                  Estimate  Std.err  Wald Pr(>|W|)
(Intercept)       -2.71412  2.03596  1.78  0.18250
log(LandArea)      0.02935  0.15493  0.04  0.84975
log(PopMillions)   0.55753  0.16214 11.82  0.00058 ***
UrbanPopulation    0.00488  0.01164  0.18  0.67488
log(GDPPerCapita) -0.20698  0.17283  1.43  0.23108
GDPPerCapGrowth   -0.03678  0.00925 15.83 0.000069 ***
PostColdWar       -0.14425  0.23432  0.38  0.53816
POLITY             0.55323  0.17081 10.49  0.00120 **
POLITYSquared     -0.05639  0.01669 11.42  0.00073 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation structure = exchangeable
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.731   0.177
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.343   0.112
Number of clusters:   160  Maximum cluster size: 57
```

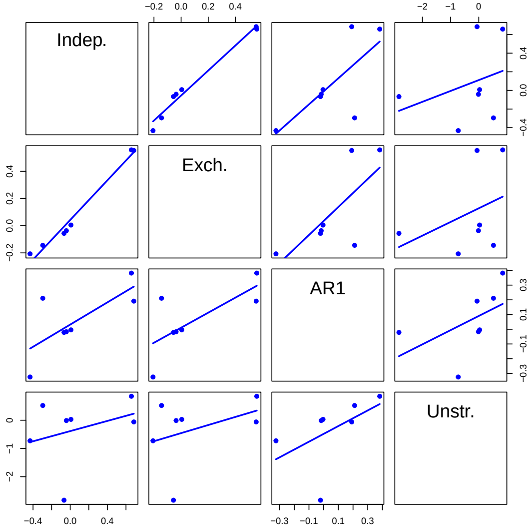
GEE: AR(1)

```
> GEE.ar1<-geeglm(CivilWar~log(LandArea)+log(PopMillions)+UrbanPopulation+
+                   log(GDPPerCapita)+GDPPerCapGrowth+PostColdWar+POLITY+
+                   POLITYSquared,data=DF,id=ISO3,family="binomial",
+                   corstr="ar1")

> summary(GEE.ar1)

Call:
geeglm(formula = CivilWar ~ log(LandArea) + log(PopMillions) +
    UrbanPopulation + log(GDPPerCapita) + GDPPerCapGrowth + PostColdWar +
    POLITY + POLITYSquared, family = "binomial", data = DF, id = ISO3,
    corstr = "ar1")

 Coefficients:
                  Estimate  Std.err Wald Pr(>|W|)
(Intercept)       -1.39416  3.04574 0.21    0.647
log(LandArea)      0.08486  0.21368 0.16    0.691
log(PopMillions)   0.38180  0.21907 3.04    0.081 .
UrbanPopulation   -0.00424  0.01753 0.06    0.809
log(GDPPerCapita) -0.32445  0.26143 1.54    0.215
GDPPerCapGrowth   -0.01668  0.00790 4.45    0.035 *
PostColdWar        0.21084  0.24470 0.74    0.389
POLITY             0.19120  0.12532 2.33    0.127
POLITYSquared     -0.02143  0.01328 2.60    0.107
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation structure = ar1
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.816    0.37
  Link = identity

Estimated Correlation Parameters:
      Estimate Std.err
alpha    0.922  0.0386
Number of clusters:   160  Maximum cluster size: 57
```

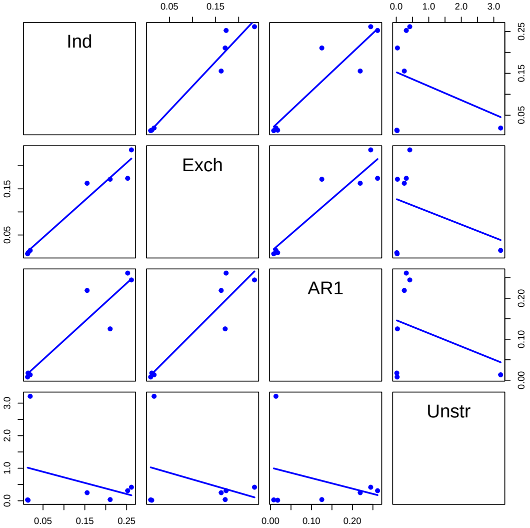
GEE: Unstructured (2013-2017)

```
> GEE.unstr<-geeglm(CivilWar~log(LandArea)+log(PopMillions)+UrbanPopulation+log(GDPPerCapita)+GDPPerCapGrowth+POLITY+
+                    POLITYSquared,data=DF5,id=ISO3,family="binomial",corstr="unstructured")

> summary(GEE.unstr)

Call:
geeglm(formula = CivilWar ~ log(LandArea) + log(PopMillions) +
    UrbanPopulation + log(GDPPerCapita) + GDPPerCapGrowth + POLITY +
    POLITYSquared, family = "binomial", data = DF5, id = ISO3,
    corstr = "unstructured")

 Coefficients:
                  Estimate Std.err  Wald Pr(>|W|)
(Intercept)        -2.8329  3.2113  0.78  0.37769
log(LandArea)       0.1657  0.1925  0.74  0.38927
log(PopMillions)    0.8493  0.2479 11.74  0.00061 ***
UrbanPopulation     0.0302  0.0175  2.98  0.08441 .
log(GDPPerCapita)  -0.7275  0.3095  5.53  0.01873 *
GDPPerCapGrowth    -0.0111  0.0277  0.16  0.68822
POLITY              0.5221  0.4166  1.57  0.21011
POLITYSquared      -0.0603  0.0366  2.71  0.09978 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation structure = unstructured
Estimated Scale Parameters:

            Estimate Std.err
(Intercept)    0.678   0.808
  Link = identity

Estimated Correlation Parameters:
          Estimate Std.err
alpha.1:2    0.395   0.489
alpha.1:3    0.408   0.508
alpha.1:4    0.352   0.441
alpha.1:5    0.325   0.405
alpha.2:3    0.757   0.862
alpha.2:4    0.290   0.367
alpha.2:5    0.518   0.593
alpha.3:4    0.400   0.509
alpha.3:5    0.742   0.861
alpha.4:5    0.436   0.547
Number of clusters:   159  Maximum cluster size: 5
```

Comparing $\hat{\beta}$s



Comparing $\widehat{\text{s.e.}}\text{s}$



GEEs are:

- Robust
- Flexible
- Extensible beyond panel/TSCS context

Appendix: Discrete-Time Survival Models

Survival models:

- ...are models for *time-to-event data*.
- ...have their roots in biostats/epidemiology, plus engineering, sociology, economics.
- Examples...
    - Political careers, confirmation durations, position-taking, bill cosponsorship, campaign contributions, policy innovation/adoption, etc.
    - Cabinet/government durations, length of civil wars, coalition durability, etc.
    - War duration, peace duration, alliance longevity, length of trade agreements, etc.
    - Strike durations, work careers (including promotions, firings, etc.), criminal careers, marriage and child-bearing behavior, etc.

Time-To-Event Data

Characteristics:

- *Discrete* events (i.e., not continuous),
- Take place over *time*,
- May not (or never) experience the event (i.e., possibility of *censoring*).

Terminology:

$$
\begin{aligned}
Y_i &= \text{the duration until the event occurs,} \\
Z_i &= \text{the duration until the observation is ``censored''} \\
T_i &= \min\{Y_i, Z_i\}, \\
C_i &= 0 \text{ if observation } i \text{ is } \textit{censored}\text{, 1 if it is not.}
\end{aligned}
$$

Density:

$$f(t) = \Pr(T_i = t)$$

CDF:

$$\Pr(T_i \leq t) \equiv F(t) = \int_0^t f(t)\,dt$$

Survival function:

$$\begin{aligned} \Pr(T_i \geq t) \equiv S(t) &= 1 - F(t) \\ &= 1 - \int_0^t f(t)\,dt \end{aligned}$$

Hazard:

$$\begin{aligned} \Pr(T_i = t | T_i \geq t) \equiv h(t) &= \frac{f(t)}{S(t)} \\ &= \frac{f(t)}{1 - \int_0^t f(t)\,dt} \end{aligned}$$

Grouped-Data Survival Approaches

Model:

$$\Pr(C_{it} = 1) = f(\mathbf{X}_{it}\beta)$$

Advantages:

- Easily estimated, interpreted and understood
- Natural interpretations:
  - $\hat{\beta}_0 \approx$ "baseline hazard"
  - Covariates shift this up or down.
- Can incorporate data on time-varying covariates
- Lots of software

Potential Disadvantages:

- Requires time-varying data
- Must deal with time dependence explicitly

Temporal Issues in Grouped-Data Models

(Implicit) "Baseline" hazard:

$$h_0(t) = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)}$$

$\longrightarrow$ No temporal dependence / "flat" hazard

Time trend:

$$\Pr(Y_{it} = 1) = f(\mathbf{X}_{it}\beta + \gamma T_{it})$$

- $\hat{\gamma} > 0 \rightarrow$ rising hazard
- $\hat{\gamma} < 0 \rightarrow$ declining hazard
- $\hat{\gamma} = 0 \rightarrow$ "flat" (exponential) hazard

Variants/extensions: Polynomials...

$$\Pr(Y_{it} = 1) = f(\mathbf{X}_{it}\beta + \gamma_1 T_{it} + \gamma_2 T_{it}^2 + \gamma_3 T_{it}^3 + ...)$$

Temporal Issues in Grouped-Data Models

"Time dummies":

$$\Pr(Y_{it} = 1) = f[\mathbf{X}_{it}\beta + \alpha_1 I(T_{i1}) + \alpha_2 I(T_{i2}) + ... + \alpha_{t_{\max}} I(T_{it_{\max}})]$$

$\longrightarrow$ Beck, Katz, and Tucker's (1998) cubic splines; might also use:

- Fractional polynomials
- Smoothed duration
- Loess/lowess fits
- Other splines (B-splines, P-splines, natural splines, etc.)