The objective of this study is to identify important family- and community-level factors that affect whether Guatemalan children are immunized. A nationally representative sample of 5160 mothers, between 15 and 44 years old were interviewed.

The Variables

Level 1 (children)

- immun: dummy variable for child being immunized, the response variable.
- kid2p: child at least 2 years old at the time of the interview.

Level 2 (mothers)

- mom: identifier for mother
- Ethnicity (dummy variables with 'Latino' as reference category)
  - indNoSpa: mother is indigenous, not Spanish speaking
  - indSpa: mother is indigenous, Spanish speaking
- Mother's eduation (dummy variables with 'no education' as reference category)
  - monEdPri: mother has primary education
  - monEdSec: mother has secondary education
- Husband's education (dummy variables with 'no education' as reference category)
  - husEdPri: husband has primary education
  - husEdSec: husband has secondary education
  - husEdDK: husband's education is not known

Level 3 (communities)

- cluster: identifier for communities
- rural: dummy variable for community being rural
- pcInd81: percentage of population that was indigenous in 1981

```
proc print data=guatemala (obs=10);
run;
```

Obs	kid	mom	cluster	immun	kid2p	mom25p	order23	order46	order7p	indNoSpa	indSpa
1	2	2	1	1	1	0	0	0	0	0	0
2	269	185	36	0	1	0	1	0	0	0	0
3	272	186	36	0	1	0	0	0	0	0	0
4	273	187	36	0	1	0	1	0	0	0	0
5	274	188	36	0	1	1	0	1	0	0	0
6	275	188	36	1	1	0	1	0	0	0	0
7	276	189	36	1	1	0	1	0	0	0	0
8	277	190	36	1	0	1	0	0	1	0	0
9	278	190	36	1	1	1	0	1	0	0	0
10	280	191	36	1	1	1	0	1	0	0	0

Obs	momEdPri	momEdSec	husEdPri	husEdSec	husEdDK	momWork	rural	pcInd81
1	0	1	0	1	0	0	0	0.10750
2	1	0	1	0	0	1	0	0.04373
3	1	0	0	1	0	1	0	0.04373
4	1	0	1	0	0	1	0	0.04373
5	1	0	0	0	1	1	0	0.04373
6	1	0	0	0	1	1	0	0.04373

Obs	momEdPri	momEdSec	husEdPri	husEdSec	husEdDK	momWork	rural	pcInd81
7	0	1	1	0	0	1	0	0.04373
8	1	0	1	0	0	1	0	0.04373
9	1	0	1	0	0	1	0	0.04373
10	1	0	0	1	0	1	0	0.04373

First, we'll take a look at how many observations there are per mother and per community

```
proc freq data=guatemala noprint;
tables mom/  out=freq_ID;
run;

title "Frequency of mother-level observations";
proc freq data=freq_ID;
tables COUNT;
run;
```

Frequency Count				
COUNT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1063	66.65	1063	66.65
2	500	31.35	1563	97.99
3	32	2.01	1595	100.00

```
proc freq data=guatemala noprint;
tables cluster/  out=freq_ID;
run;

title "Frequency of community-level observations";
proc freq data=freq_ID;
tables COUNT;
run;
```

Frequency Count				
COUNT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	2	1.24	2	1.24
2	7	4.35	9	5.59
3	7	4.35	16	9.94
4	6	3.73	22	13.66
5	6	3.73	28	17.39
6	7	4.35	35	21.74
7	6	3.73	41	25.47
8	8	4.97	49	30.43
9	11	6.83	60	37.27
10	13	8.07	73	45.34
11	13	8.07	86	53.42
12	5	3.11	91	56.52
13	2	1.24	93	57.76
14	5	3.11	98	60.87
15	6	3.73	104	64.60
16	7	4.35	111	68.94
17	6	3.73	117	72.67
18	4	2.48	121	75.16
19	2	1.24	123	76.40
20	3	1.86	126	78.26
21	6	3.73	132	81.99
22	3	1.86	135	83.85
23	5	3.11	140	86.96
24	5	3.11	145	90.06
25	5	3.11	150	93.17
26	2	1.24	152	94.41
27	2	1.24	154	95.65
29	1	0.62	155	96.27
30	2	1.24	157	97.52
32	1	0.62	158	98.14
34	1	0.62	159	98.76
50	1	0.62	160	99.38
55	1	0.62	161	100.00

Now we'll run a three level random intercept model

```
title "Three-level random intercept model";
proc glimmix data=guatemala noitprint NOCLPRINT method=LAPLACE;
class mom cluster;
model immun=kid2p mom25p order23 order46 order7p indNoSpa indSpa momEdPri momEdSec
husEdPri husEdSec husEdDK momWork rural pcInd81/d=bin link=logit solution;
random intercept/subject=cluster G;
random intercept/subject=mom(cluster) G;
covtest 'var(cluster) = 0'          0  .;
covtest 'var(mom(cluster)) = 0' .  0;
run;
quit;
```

Model Information	
Data Set	WORK.GUATEMALA
Response Variable	immun
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	cluster
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	2159
Number of Observations Used	2159

Dimensions	
G-side Cov. Parameters	2
Columns in X	16
Columns in Z per Subject	38
Subjects (Blocks in V)	161
Max Obs per Subject	55

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	18
Lower Boundaries	2
Upper Boundaries	0
Fixed Effects	Not Profiled
Starting From	GLM estimates

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2711.40
AIC (smaller is better)	2747.40
AICC (smaller is better)	2747.72
BIC (smaller is better)	2802.87
CAIC (smaller is better)	2820.87
HQIC (smaller is better)	2769.92

Fit Statistics for Conditional Distribution	
-2 log L(immun | r. effects)	1757.11
Pearson Chi-Square	1190.26
Pearson Chi-Square / DF	0.55

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
Intercept	cluster	0.5199	0.1474
Intercept	mom(cluster)	1.2878	0.3658

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-0.9466	0.3388	159	-2.79	0.0059
kid2p	1.2816	0.1601	558	8.00	<.0001
mom25p	-0.1285	0.1655	558	-0.78	0.4377
order23	-0.1383	0.1738	558	-0.80	0.4264
order46	0.1743	0.2138	558	0.81	0.4154
order7p	0.2894	0.2673	558	1.08	0.2793
indNoSpa	-0.1127	0.3379	558	-0.33	0.7390
indSpa	-0.03441	0.2515	558	-0.14	0.8912
momEdPri	0.2954	0.1526	558	1.94	0.0533
momEdSec	0.3020	0.3347	558	0.90	0.3673
husEdPri	0.3949	0.1594	558	2.48	0.0135
husEdSec	0.3680	0.2852	558	1.29	0.1975
husEdDK	0.01427	0.2457	558	0.06	0.9537

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
momWork	0.2705	0.1392	558	1.94	0.0524
rural	-0.6494	0.2100	558	-3.09	0.0021
pcInd81	-0.8578	0.3460	558	-2.48	0.0135

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
kid2p	1	558	64.08	<.0001
mom25p	1	558	0.60	0.4377
order23	1	558	0.63	0.4264
order46	1	558	0.66	0.4154
order7p	1	558	1.17	0.2793
indNoSpa	1	558	0.11	0.7390
indSpa	1	558	0.02	0.8912
momEdPri	1	558	3.75	0.0533
momEdSec	1	558	0.81	0.3673
husEdPri	1	558	6.14	0.0135
husEdSec	1	558	1.66	0.1975
husEdDK	1	558	0.00	0.9537
momWork	1	558	3.78	0.0524
rural	1	558	9.57	0.0021
pcInd81	1	558	6.15	0.0135

Tests of Covariance Parameters Based on the Likelihood					
Label	DF	-2 Log Like	ChiSq	Pr > ChiSq	Note
var(cluster) = 0	1	2745.73	34.33	<.0001	MI
var(mom(cluster)) = 0	1	2744.74	33.34	<.0001	MI

MI: P-value based on a mixture of chi-squares.

```
title "Three-level random intercept model with a subset of covariates";
proc glimmix data=guatemala noitprint NOCLPRINT method=LAPLACE;
class mom cluster;
model immun=kid2p rural pcInd81/d=bin link=logit solution;
random intercept/subject=cluster G;
random intercept/subject=mom(cluster) G;
covtest 'var(cluster) = 0'          0   .;
```

```
covtest 'var(mom(cluster)) = 0' .  0;
run;
quit;
```

Model Information	
Data Set	WORK.GUATEMALA
Response Variable	immun
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	cluster
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	2159
Number of Observations Used	2159

Dimensions	
G-side Cov. Parameters	2
Columns in X	4
Columns in Z per Subject	38
Subjects (Blocks in V)	161
Max Obs per Subject	55

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	6
Lower Boundaries	2
Upper Boundaries	0
Fixed Effects	Not Profiled
Starting From	GLM estimates

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2733.16
AIC (smaller is better)	2745.16
AICC (smaller is better)	2745.19
BIC (smaller is better)	2763.64
CAIC (smaller is better)	2769.64
HQIC (smaller is better)	2752.66

Fit Statistics for Conditional Distribution	
-2 log L(immun | r. effects)	1807.74
Pearson Chi-Square	1226.00
Pearson Chi-Square / DF	0.57

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
Intercept	cluster	0.5236	0.1421
Intercept	mom(cluster)	1.1903	0.3398

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-0.1796	0.2179	159	-0.82	0.4110
kid2p	1.2469	0.1558	562	8.01	<.0001
rural	-0.7778	0.1953	562	-3.98	<.0001
pcInd81	-1.1860	0.2432	562	-4.88	<.0001

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
kid2p	1	562	64.09	<.0001
rural	1	562	15.86	<.0001
pcInd81	1	562	23.78	<.0001

Tests of Covariance Parameters Based on the Likelihood					
Label	DF	-2 Log Like	ChiSq	Pr > ChiSq	Note
var(cluster) = 0	1	2772.23	39.08	<.0001	MI
var(mom(cluster)) = 0	1	2764.18	31.02	<.0001	MI

MI: P-value based on a mixture of chi-squares.

```
title "Using random coefficients";
proc glimmix data=guatemala noitprint NOCLPRINT method=LAPLACE;
class mom cluster;
model immun=kid2p rural pcInd81/d=bin link=logit solution;
random intercept kid2p/subject=cluster type=UN G;
random intercept /subject=mom(cluster) G;
run;
quit;
```

Model Information	
Data Set	WORK.GUATEMALA
Response Variable	immun
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	cluster
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	2159
Number of Observations Used	2159

Dimensions	
G-side Cov. Parameters	4
Columns in X	4
Columns in Z per Subject	39
Subjects (Blocks in V)	161
Max Obs per Subject	55

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	8
Lower Boundaries	3
Upper Boundaries	0
Fixed Effects	Not Profiled
Starting From	GLM estimates

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2729.96
AIC (smaller is better)	2745.96
AICC (smaller is better)	2746.03
BIC (smaller is better)	2770.61
CAIC (smaller is better)	2778.61
HQIC (smaller is better)	2755.97

Fit Statistics for Conditional Distribution	
-2 log L(immun | r. effects)	1774.33
Pearson Chi-Square	1191.38
Pearson Chi-Square / DF	0.55

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
UN(1,1)	cluster	0.8664	0.4061
UN(2,1)	cluster	-0.4142	0.3548
UN(2,2)	cluster	0.5118	0.3724
Intercept	mom(cluster)	1.1763	0.3467

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-0.2513	0.2405	159	-1.04	0.2976
kid2p	1.3252	0.1872	141	7.08	<.0001
rural	-0.7816	0.1974	447	-3.96	<.0001
pcInd81	-1.1933	0.2459	447	-4.85	<.0001

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
kid2p	1	141	50.11	<.0001
rural	1	447	15.68	<.0001
pcInd81	1	447	23.55	<.0001

```
 title "Baseline Two-level model";
proc glimmix data=guatemala noitprint  NOCLPRINT
method=LAPLACE;
class mom cluster;
model immun=kid2p/d=bin link=logit solution;
random intercept kid2p/subject=cluster type=UN G;
run;
quit;
```

Model Information	
Data Set	WORK.GUATEMALA
Response Variable	immun
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	cluster
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	2159
Number of Observations Used	2159

Dimensions	
G-side Cov. Parameters	3
Columns in X	2
Columns in Z per Subject	2
Subjects (Blocks in V)	161
Max Obs per Subject	55

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	5
Lower Boundaries	2
Upper Boundaries	0
Fixed Effects	Not Profiled
Starting From	GLM estimates

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2797.19
AIC (smaller is better)	2807.19
AICC (smaller is better)	2807.22
BIC (smaller is better)	2822.60

Fit Statistics	
CAIC (smaller is better)	2827.60
HQIC (smaller is better)	2813.45

Fit Statistics for Conditional Distribution	
-2 log L(immun | r. effects)	2499.99
Pearson Chi-Square	1874.20
Pearson Chi-Square / DF	0.87

Estimated G Matrix			
Effect	Row	Col1	Col2
Intercept	1	1.1384	-0.5345
kid2p	2	-0.5345	0.5315

Covariance Parameter Estimates			
Cov Parm	Subject	Estimate	Standard Error
UN(1,1)	cluster	1.1384	0.4206
UN(2,1)	cluster	-0.5345	0.3413
UN(2,2)	cluster	0.5315	0.3171

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-1.1340	0.1619	160	-7.01	<.0001
kid2p	1.1128	0.1576	141	7.06	<.0001

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
kid2p	1	141	49.85	<.0001

```
 title "Adjusted Two-level model";
proc glimmix data=guatemala noitprint NOCLPRINT method=LAPLACE;
```

```
class mom cluster;
model immun=kid2p rural pcInd81/d=bin link=logit solution;
random intercept kid2p/subject=cluster type=UN G;
run;
quit;
```

Model Information	
Data Set	WORK.GUATEMALA
Response Variable	immun
Response Distribution	Binomial
Link Function	Logit
Variance Function	Default
Variance Matrix Blocked By	cluster
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	2159
Number of Observations Used	2159

Dimensions	
G-side Cov. Parameters	3
Columns in X	4
Columns in Z per Subject	2
Subjects (Blocks in V)	161
Max Obs per Subject	55

Optimization Information	
Optimization Technique	Dual Quasi-Newton
Parameters in Optimization	7
Lower Boundaries	2
Upper Boundaries	0
Fixed Effects	Not Profiled
Starting From	GLM estimates

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2759.07
AIC (smaller is better)	2773.07
AICC (smaller is better)	2773.12
BIC (smaller is better)	2794.64
CAIC (smaller is better)	2801.64
HQIC (smaller is better)	2781.82

Fit Statistics for Conditional Distribution	
-2 log L(immun | r. effects)	2507.19
Pearson Chi-Square	1900.62
Pearson Chi-Square / DF	0.88

Estimated G Matrix			
Effect	**Row**	**Col1**	**Col2**
Intercept	**1**	0.7552	-0.4039
kid2p	**2**	-0.4039	0.4987

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
UN(1,1)	cluster	0.7552	0.3188
UN(2,1)	cluster	-0.4039	0.2860
UN(2,2)	cluster	0.4987	0.3022

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-0.2231	0.2057	159	-1.08	0.2797
kid2p	1.0840	0.1546	141	7.01	<.0001
rural	-0.6337	0.1617	1855	-3.92	<.0001
pcInd81	-0.9534	0.2005	1855	-4.76	<.0001

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
kid2p	1	141	49.17	<.0001
rural	1	1855	15.37	<.0001
pcInd81	1	1855	22.62	<.0001