Model Comparison and Simulation for Hierarchical Models: Analyzing Rural-Urban Migration in Thailand

Filiz Garip[1]
Bruce Western
Harvard University

December 2008

[1]Department of Sociology, Harvard University, 33 Kirkland St, Cambridge, MA 02138. Nang Rong Survey is a collaborative effort between investigators at the University of North Carolina, Carolina Population Center, and investigators at the Institute for Population and Social Research (IPSR), Mahidol University, Salaya, Thailand. It is partially funded by Grant R01- HD25482 from the National Institute of Child Health and Human Development to the Carolina Population Center, University of North Carolina at Chapel Hill. Persons interested in obtaining data files from the Nang Rong Survey Project should contact The Nang Rong Survey Project, Carolina Population Center, 123 W. Franklin Street, CB 8120, Chapel Hill, NC 27516-3997 (nangrong@unc.edu).

Abstract

Sociological research often examines the effects of social context with hierarchical models. In these applications, individuals are nested in social contexts—like school classes, neighborhoods or villages—whose effects are thought to shape individual outcomes. Although applications of hierarchical models are common in sociology, analysis usually focuses on inference for fixed parameters. Researchers seldom study model fit or examine aggregate patterns of variation implied by model parameters. We present an analysis of Thai migration data, in which survey respondents are nested within villages and report annual migration information. We study a variety of hierarchical models, investigating model fit with DIC and posterior predictive statistics. We also describe a simulation to study how different initial distributions of migration across villages produce increasing inter-village inequality in migration.

Sociologists often argue that social context matters. Features of the social context, not just the characteristics of individuals, help produce aggregate outcomes like the distribution of economic rewards, or paths of development. Multilevel designs where individuals are nested within social contexts provide a strong design for observing both contextual effects and the aggregate outcomes those effects might produce.

We present an analysis of migration in rural Thailand, in which survey respondents are nested within villages, providing annual reports on migration for the 1980s and 1990s. Rural to urban migration has propelled economic development as rural migrants remit their earnings back to their villages and return with news of economic opportunities for friends and family members. Though our data describe thousands of individual migration decisions, our interest focuses on aggregate differences across villages. The rural northeast of Thailand varies tremendously in the degree to which villages are integrated into the urban economies further south. The evolution of inequality in migration across villages is thus important for our understanding patterns of poverty and development in the rural areas of countries experiencing rapid growth.

Hierarchical models provide a valuable tool for studying multilevel sociological data like the Thai migration surveys (Mason, Wong, and Entwisle 1983; Western 1999). In sociology and demography, panel surveys of individuals and households, survey data from many countries, pooled time series data from U.S. states and cities, have all been analyzed with hierarchical models (DiPrete and Forristal 1994). Sometimes, sociological applications have studied the heterogeneity of parameters across units, though more commonly hierarchical models offer a way to account for clustering in inferences about fixed parameters. In these cases, random effects are a nuisance, integrated out for correct inference.

Hierarchical models are common in sociology, but applied research often neglects two important topics. First, sociological analysis of hierarchical models rarely provides a detailed examination of model fit. In our analysis of the Thai migration data we study the fit of several alternative models by comparing DIC and posterior predictive statistics. Model fit is an important applied topic because often in sociology, theory will be indifferent to alternative specifications of random effects. The structure of random effects may also have important implications for substantive conclusions. In particular, substantively important aggregate outcomes that are not directly modeled—like inequality in a response across units or response variable quantiles—may be sensitive to the specification of random effects. A second limitation of applied sociological research with hierarchical models is that these aggregate implications of model estimates typically go unexamined. Our analysis of rural-urban migration in Thailand examines several hierarchical models. In our analysis, MCMC computation for hierarchical models provides a convenient framework for studying aggregate patterns of variation by simulating migration given different hypothetical distributions of covariates.

Introducing the Thai Migration Data

The Thai migration data are based on the Nang Rong surveys[1] of men and women aged 13 to 41, from 22 villages in the Nang Rong district of Northeastern Thailand (Curran, Garip, Chung, and Tangchonlatip 2005). We combine data from two waves (1994 and 2000) of the life history survey. The 1994

[1] Nang Rong surveys are conducted by the University of North Carolina and Mahidol University in Thailand. The data and information about the surveys are available at http://www.cpc.unc.edu/projects/nangrong.

wave begins with men and women aged 13 to 35 in 1994, and asks about respondents' migration experiences since the age of 13. This design is replicated in 2000: men and women aged 18 to 41 are asked about their migration behavior starting at the age of 13. Some respondents were living away from the village at the time of the survey, and they were followed up and interviewed [2]. We merge these data with household censuses conducted in 1984, 1994 and 2000 to obtain household and village characteristics. The resulting data contain information on migration of 6,768 respondents nested within 22 villages over a 16-year time period from 1984 to 2000 ($N = 93,914$).

Our interest focuses on how the level of migration in a village might subsequently promote more migration among individuals. Figure 1 shows the distribution of village migration rates, $\bar{y}_{jt} = \sum_i y_{ijt}/n_{jt}$, from 1984 to 2000. The survey data are retrospective, and the age distributions vary over time. The figure displays the migration rates for 18-25 year-old men and women, the age group that we observe every year. Migration rates generally increase until 1996. In 1984, around a quarter of young residents in Nang Rong left their district for at least two months. By 1996, the migration rate for the region had increased to about 50 percent. In 1996, the Asian financial crisis precipitated recession in Thailand. Migration rates declined over the next four years. In some villages, migration declines were particularly steep, with migration rates falling to around 10 percent. Trends for a high-migration and low-migration village are also shown in the plot. These trends share some common features, like the increase in migration in the first decade, and the

[2]Related project manuscripts report that the success at finding migrants was relatively high (Rindfuss, Kaneda, Chattopadhyay, and Sethaput 2007). On average, about 44% of the migrants were successfully interviewed at some point in the six months following the village surveys.

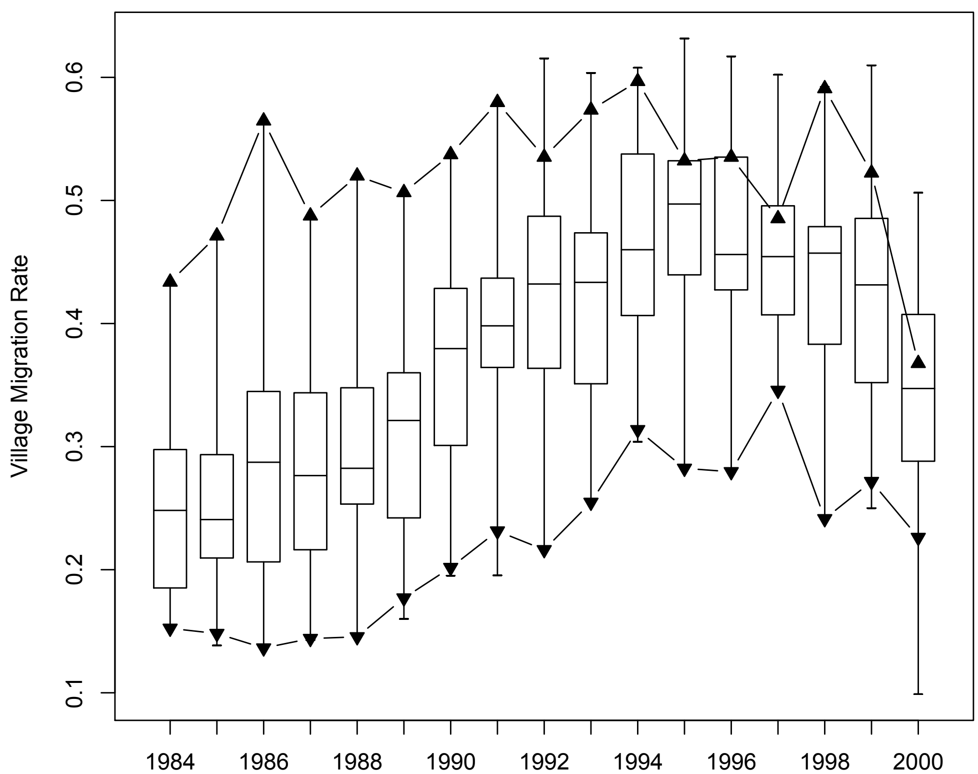

Figure 1: Boxplots of annual village migration rates, men and women aged 18 to 25, Nang Rong, Thailand, 1984 to 2000. Migration rates for villages with the largest and smallest migration rates in 1984 are shown by the trend lines.

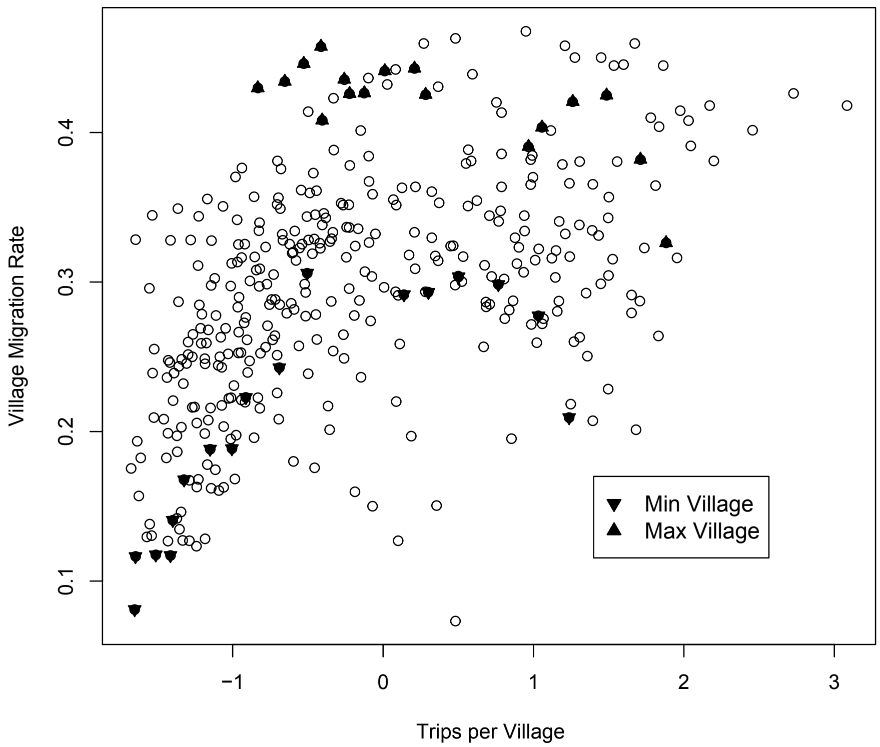

Figure 2: Scatterplot of village trips and village migration rates, Nang Rong, Thailand, 1984 to 2000. Villages with the smallest and largest migration rates in 1984 are indicated separately.

decline from 1996.

Part of our substantive interest focuses on how the accumulation of migration experiences within villages is associated with an individual's likelihood of migration. Migration for an individual may become more likely if they live in a village in which many others have migrated. This phenomenon, called the cumulative causation of migration, occurs because prior migration generates resources or influence that make individuals more likely to migrate (Massey

1990). Extensive empirical evidence documents how past migration becomes a primary engine for future migration flows, eventually diminishing the importance of alternative explanations (Garip 2008; Massey and Espinosa 1997; Massey and Zenteno 1999).

We study the effect of social context, by a constructing a "village trips" variable that records the number of trips taken in a village in the years preceding the current year. This variable is standardized to have mean zero and unit variance. A scatterplot of village trips and annual village migration rates for the 1984 to 2000 period is shown in Figure 2. In any given year, villages with the highest migration rates, have a history of high levels of migration. This pattern is not surprising, but it remains an open empirical question whether a village's history of migration is associated with an individual's likelihood of migration, after accounting for their own history of migration, their family's migration history, and other covariates.

To study the effect of village trips for these multilevel data we write several hierarchical logistic regression models. For respondent i $(1 = 1, \ldots, n_{tj})$ in village j $(j = 1, \ldots, 22)$ in year t $(t = 1984, \ldots, 2000)$, y_{ijt} denotes the binary migration outcome, that equals 1 if respondent travels away from the village for more than two months in the year, and 0 otherwise. Individual and village-level covariates are collected in vectors, $\boldsymbol{x_{ijt}}$ and $\boldsymbol{z_{jt}}$. In each of the following logistic regressions, y_{ijt} conditional on fixed and random effects collected in the vector $\boldsymbol{\theta}$, is assumed to be Bernoulli, $P(y|\boldsymbol{\theta}) = p^y(1-p)^{1-y}$, with expectation, $E(y) = p$ and likelihood, $L(\boldsymbol{\theta}; \boldsymbol{y}) = \prod P(y_{ijt}|\boldsymbol{\theta})$.

If we consider only the panel aspect of the data design, we can fit a respondent-level random effect, α_i, to allow for the correlation of observations for the same respondent, yielding the logistic regression:

$$\text{logit}(p_{ijt}) = \alpha_i + \delta_t + \boldsymbol{\beta}_1'\boldsymbol{x_{ijt}} + \boldsymbol{\beta}_2'\boldsymbol{z_{jt}}, \tag{1}$$

where $\boldsymbol{x_{ijt}}$ and $\boldsymbol{z_{jt}}$ represent vectors of individual and village-level covariates, with corresponding fixed effects $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. This specification also includes a time effect, δ_t that captures the common trend in migration across villages. The two levels of clustering, by respondent and village, could be modeled with separate effects, where a village effect captures a migration propensity that is common to all residents of the same village:

$$\text{logit}(p_{ijt}) = \alpha_i + \gamma_j + \delta_t + \boldsymbol{\beta}_1' \boldsymbol{x_{ijt}} + \boldsymbol{\beta}_2' \boldsymbol{z_{jt}} \tag{2}$$

Finally, heterogeneity in village effects over time can be captured with a village-by-year effect, γ_{jt}:

$$\text{logit}(p_{ijt}) = \alpha_i + \gamma_{jt} + \delta_t + \boldsymbol{\beta}_1' \boldsymbol{x_{ijt}} + \boldsymbol{\beta}_2' \boldsymbol{z_{jt}} \tag{3}$$

Given the observed variability in migration trends, this last model seems most realistic. This model is shown as a directed acyclical graph in Figure 3. The parameters, μ and σ^2, are the means and variances of the hyper-distributions from which the random effects are drawn. Boxes and ovals denote covariates and parameters respectively. Full arrows indicate probabilistic dependencies whereas broken arrows are deterministic relationships. The clustered structure of the data (individuals within villages for each year) is denoted by stacked sheets. In this figure, the year sheet is dotted indicating that year-specific effects will induce correlations among observations from same time point, though individuals and villages are not nested within years.

The full Bayesian specification requires hyper-distributions for the random effects, and proper priors for their hyper-parameters. In our analysis, the random effects for our three models are each given a normal distribution. The means are given diffuse normal prior distributions. The precisions (the inverse variances) are given gamma distributions. The priors, displayed

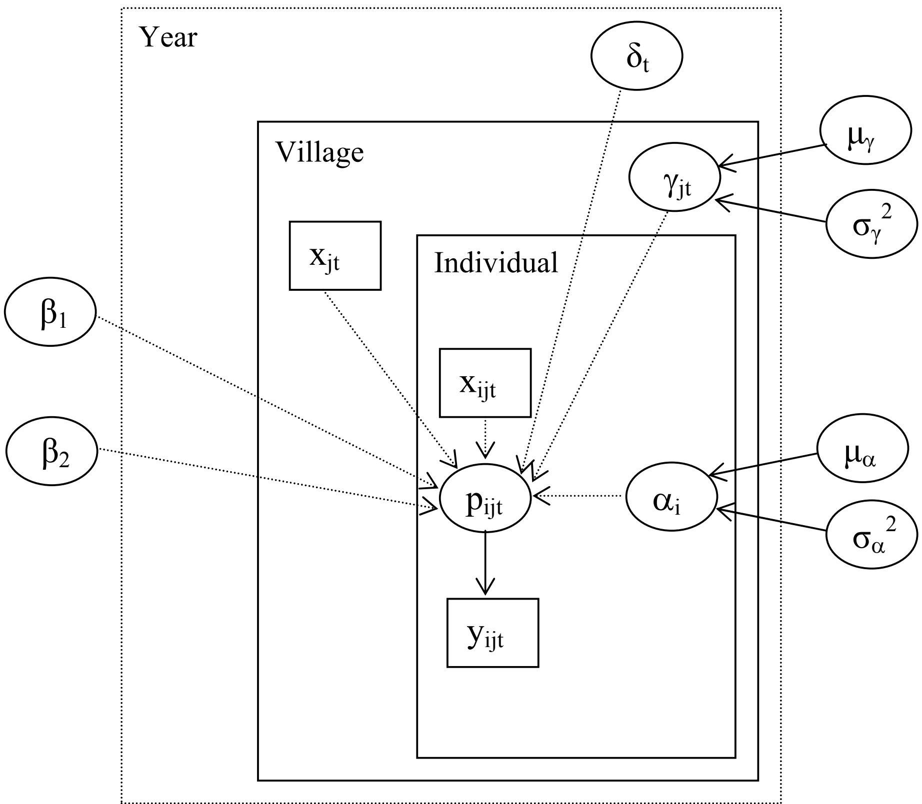

Figure 3: Three-level logit model with individual, village and year effects on individual migration outcome

Model	Random Effects	Prior Distributions
(1)	$\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$	$\mu_\alpha \sim N(0, 10^6)$
		$\sigma_\alpha^{-2} \sim Ga(.001, .001)$
(2)	$\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$	$\mu_\alpha \sim N(0, 10^6)$
		$\tau_\alpha^{-2} \sim Ga(.001, .001)$
	$\gamma_j \sim N(\mu_\gamma, \sigma_\gamma^2)$	$\mu_\gamma \sim N(0, 10^6)$
		$\sigma_\gamma^{-2} \sim Ga(.001, .001)$
(3)	$\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$	$\mu_\alpha \sim N(0, 10^6)$
		$\sigma_\alpha^{-2} \sim Ga(.001, .001)$
	$\gamma_{jt} \sim N(\mu_\gamma, \sigma_\gamma^2)$	$\mu_\gamma \sim N(0, 10^6)$
		$\sigma_\gamma^{-2} \sim Ga(.001, .001)$

Table 1: Hyper-distributions and prior distributions for hierarchical logistic regression models of Thai migration.

in Table 1 are intended to be uninformative so the sample data dominates estimation of the hyper-parameters. In some applications, precisions have been sensitive to the prior parameters of the gamma distribution. We experimented with several alternative priors and obtained essentially the same results as those reported here.

Regression Results

We can easily explore the model fit and run simulation experiments with draws from the posterior obtained by MCMC simulation. The results below are based on 10,000 iterations from parallel chains, after a burn-in of 2,500 iterations. Figure 4 displays the output from the Gibbs sampler for selected parameters from the village-year model, equation (3). The left panel shows trace plots of a single chain. The right panel shows the posterior densities estimated using a kernel smoother. The first row of Figure 4 shows the village trip coefficient. This effect is relatively slow-mixing compared to other coefficients and the variance parameters. Still, convergence diagnostics, including

Table 2: Logistic regression coefficients (standard errors) for hierarchical models of migration, Nang Rong, Thailand, 1984 to 2000.

	Individual		Village		Village-Year	
Village trips	.648	(.058)	.643	(.074)	.664	(.068)
Household trips	.113	(.020)	.117	(.021)	.113	(.021)
Ind. trips	1.457	(.022)	1.455	(.023)	1.462	(.022)
Age	-.248	(.041)	-.235	(.042)	-.243	(.041)
Male	.255	(.066)	.270	(.076)	.265	(.070)
Married	-1.181	(.037)	-1.179	(.037)	-1.186	(.037)
Education	.758	(.031)	.773	(.031)	.762	(.031)
Land	-.056	(.018)	-.061	(.019)	-.052	(.019)
σ_i	2.600	(.036)	2.576	(.037)	2.607	(.037)
σ_v	-	-	.398	(.078)	-	
σ_{vy}	-	-	-	-	.190	(.019)
DIC	61797.200		61784.300		61679.300	
p_D	4948.770		4938.940		5095.610	

Note: $N = 93,914$ for 6,768 individual respondents in 22 villages. Deviance is the average deviance evaluated over all posterior draws.

Gelman and Rubin (1992)'s, indicate convergence for all parameters.

Posterior means and standard deviations for the regression coefficients are reported in Table 2. These results show the positive association of the village history of migration with an individual's migration decision in a given year. A standard deviation difference in the trips per village nearly doubles odds of migration for an individual ($e^{.65} \approx 1.9$). A household's and individual's history of migration are also strongly associated with migration. All these effects are consistent across model specifications. Unsurprisingly, individual trips is estimated to have the strongest effect on individual migration. Less expected, however, is the relatively strong effect of the village level of migration. Covariate effects are also similar across models. Men, the unmarried, and the more educated are all somewhat more likely to migrate.

Most of the point estimates for the coefficients are insensitive to alternative specifications of the random effects, still some models may fit the data

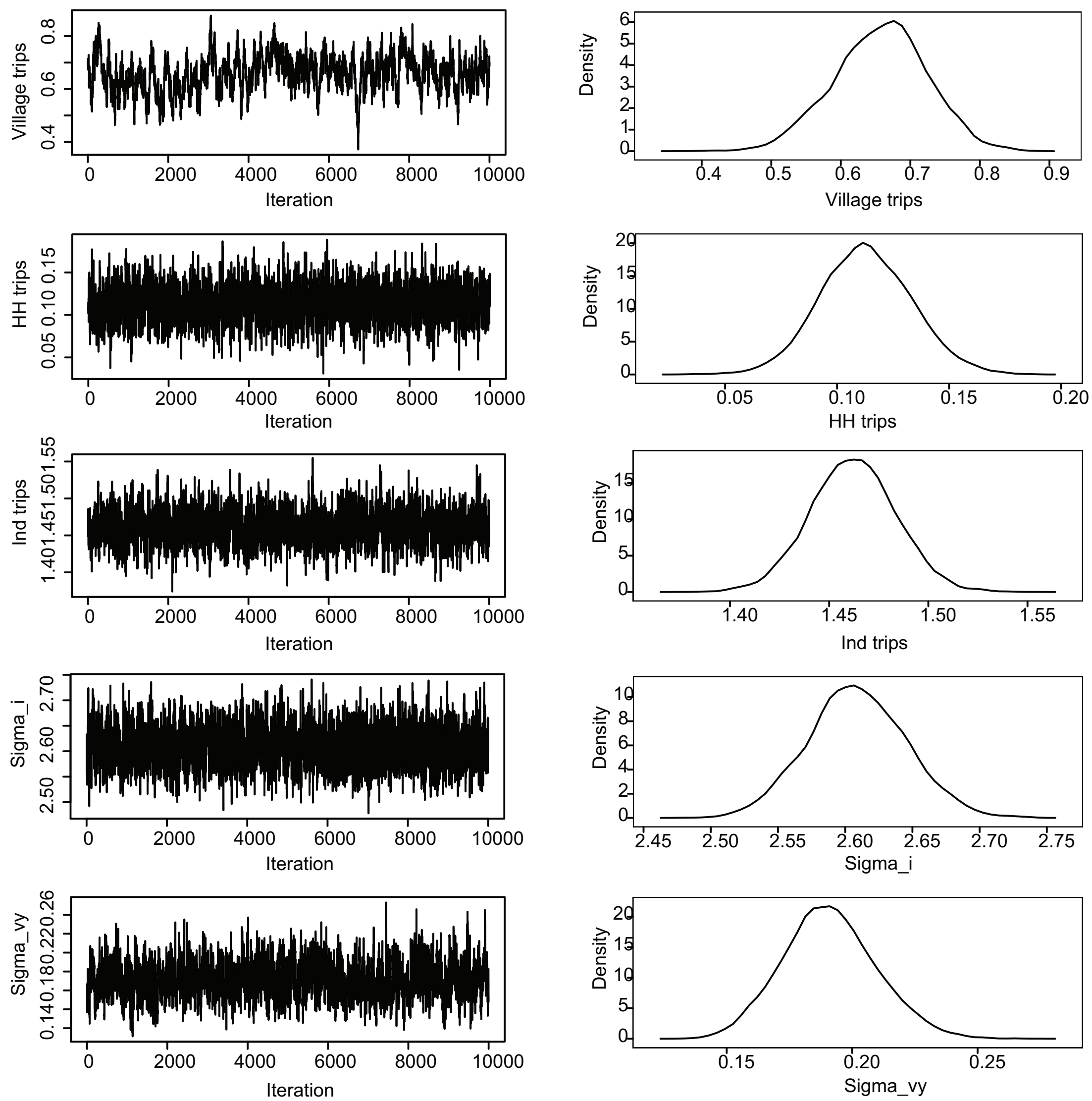

Figure 4: Trace and density plots for parameters from the village-year model.

better than others. The DIC statistic, proposed by Spiegelhalter, Nicola Best, and van der Linde (2002), is readily calculated from MCMC output. The DIC is based on the usual deviance statistic, $D(\boldsymbol{y}, \boldsymbol{\theta}) = -2 \log L(\boldsymbol{\theta}; \boldsymbol{y})$, evaluated at the simulated values of the parameters. Like the deviance, better fitting models have lower DIC statistics. DIC statistics are virtually the same for the individual and village random effects models. The DIC statistic for the village-year model, which includes random effects for each village in each year, is about 100 points lower.

A component of the DIC statistic, the p_D, is given by the difference between the posterior mean deviance and the deviance evaluated at the posterior mean and has been proposed as a measure of the effective number of parameters of a Bayesian model. The village-year model is the parametrically most complex and this is reflected in the relatively high p_D statistic. The village year model includes an additional 352 random effect over the village model, an effective addition of 156 new parameters according to the p_D.

Posterior Predictive Checks

The DIC statistic is an omnibus measure of fit and the p_D can yield odd results in some applications. An alternative approach, tailored to the substantive objectives of the research, examines model predictions for quantities of key substantive interest (Gelman, Men, and Stern 1996). The posterior predictive distribution is the distribution of future data, $\tilde{\boldsymbol{y}}$, integrating over the posterior parameter distribution for a given model:

$$p(\tilde{\boldsymbol{y}}|\boldsymbol{y}) = \int p(\tilde{\boldsymbol{y}}|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta}.$$

To study the posterior predictive distribution the researcher must define a test statistic which can be calculated from the observed data. Because we are interested in the inequality in migration across villages and over time, we define the test statistic in year t as,

$$R_t = \frac{\max(\bar{y}_{jt})}{\min(\bar{y}_{jt})},$$

the ratio of the largest to the smallest annual village migration rate. A well-fitting model should yield posterior predictions that track the observed trend in village inequality in migration.

Figure 5 compares the observed trend in R_t to the 95 percent posterior predictive confidence interval for R_t under the individual model that include only respondent-level random effects. The predictive distribution generally captures the U-shaped trend in inequality in village migration rates. In most years, the observed level of inequality falls within the predictive interval indicating that the data are not extreme under the model. Several of the most extreme observations, however, fall well outside the predictive interval.

The village model adds time-invariant random effects for each village to the individual model that includes only respondent random effects. Figure 6 shows the posterior predictive interval for the village model. Adding village-level random effects does little to improve the model's fit to longitudinal patterns of inequality in village migration rates. As for the individual model, several extreme values at the ends of the time series are poorly predicted under the village model.

Finally, the village-year model adds a random effect for each village in each year. The posterior predictive distribution in this case covers the observed trend in inequality in all years but one. The flexibility of the village-year model is reflected in the relatively wide predictive distribution displayed in Figure 7. Accounting for yearly differences in village effects adds significantly to predictive uncertainty about possible migration rates. As conse-

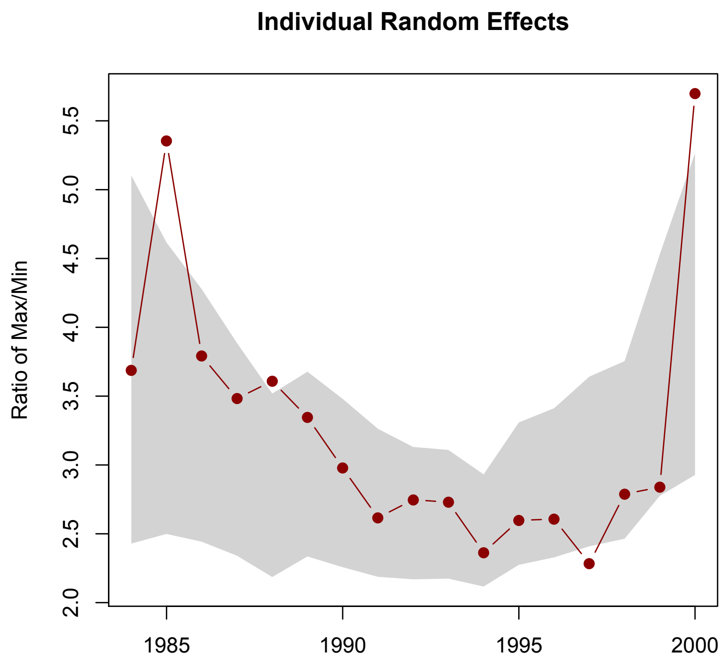

Figure 5: Inequality in village migration, R_t, and the 95 percent confidence region for the posterior predictive distribution of the individual model.

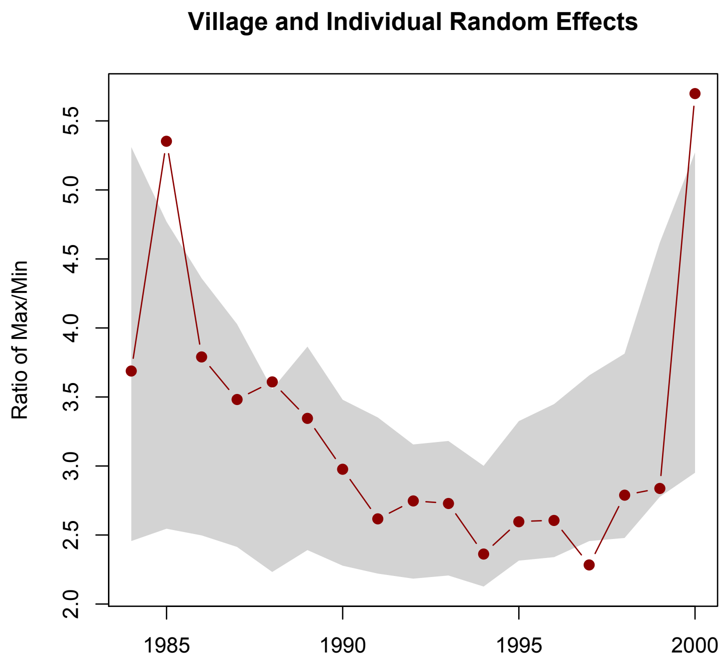

Figure 6: Inequality in village migration, R_t, and the 95 percent confidence region for the posterior predictive distribution of the village model.

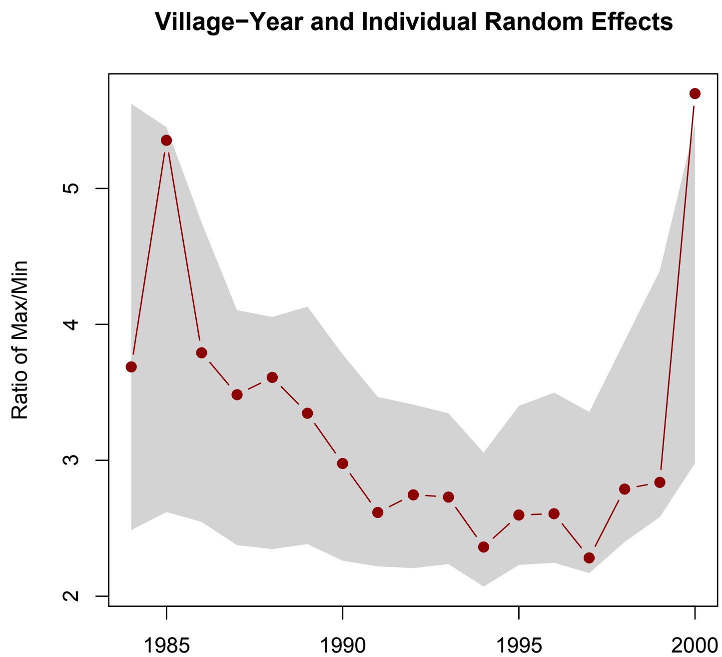

Figure 7: Inequality in village migration, R_t, and the 95 percent confidence region for the posterior predictive distribution of the village-year model.

quence, however, the observed trend in inequality is relatively likely under the village-year model.

Exploring Model Implications with Simulation

The posterior predictive check allows us to study the fit of the model, but like most applications in sociology, we have not yet examined the implications of model estimates for understanding aggregate patterns. We explore the implications of the estimated model for inequality in village migration rates using simulations. Coefficient estimates show the strong effect of village trips on individuals' migration probabilities. Those living in villages with a high number of prior trips are more likely to migrate. In those villages, more trips accumulate over time further increasing the likelihood of migration. This phenomenon, called the cumulative causation of migration, suggests a dynamic mechanism of stratification in migration patterns across villages (Massey 1990).

Due to cumulative causation, small initial differences in village trips may lead to large inequalities in village migration rates over time (Garip 2008). Our model does not account for the initial distribution of village trips. The observed distribution of village trips in the data is one among many possible configurations. To observe the full extent of the implications of our model for inequality in village migration, we use a simulation exercise.

Keeping the aggregate trips constant, we alter the initial distribution of trips across villages in the data. We simulate the migration patterns from 1985 to 2000 using the following procedure. For each year, we compute individuals' predicted migration probabilities from our estimated model. We randomly assign migrants based on that probability. We then update the cumulative individual, household and village trips, and compute individuals' expected migration probabilities for the next year. We repeat this procedure many times ($N = 1000$), and compute average village migration rates over repetitions. In simulation runs, we take random draws from the MCMC-generated posterior distribution of the parameters to calculate simulate inter-village inequality in migration. By drawing from the whole posterior distribution, simulation results reflect posterior uncertainty about parameters.

The pseudo-algorithm is as follows:

1. Distribute the initial number of village trips, V_{t_0}, across villages $j = 1, \ldots, J$, according to scenario S such that $\sum_{j=1}^{J} v_{jt_0} = V_{t_0}$.

2. Sample parameters, $\hat{\boldsymbol{\beta}}$, from the MCMC-generated posterior distribution.

3. From the fitted model, $\text{logit}(\hat{\mathbf{p}}) = \mathbf{X}\hat{\boldsymbol{\beta}}$, obtain predicted probabilities $\hat{p}_{ijt}$ $\forall i, j$ at time period t.

4. Simulate data $\mathbf{y}^*$ from the fitted model, that is, $y^*_{ijt+1} \sim \text{binomial}(1, \hat{p}_{ijt})$ $\forall i, j$.

5. Update cumulative independent variables (individual, household and village trips), $x_{ijt+1} = x_{ijt} + f(y^*_{ijt+1})$, where $f(\cdot)$ a function transforming migration in $t + 1$ into trips $\forall i, j$ [BW: OK?].

6. Compute predicted probabilities from the fitted model $\text{logit}(\mathbf{p}^*) = \mathbf{X}\boldsymbol{\beta}^*$ using the updated independent variables.

7. Increment time period, $t = t + 1$.

8. Repeat 3-7 T times, that is, generate a path of fitted values for T time periods.

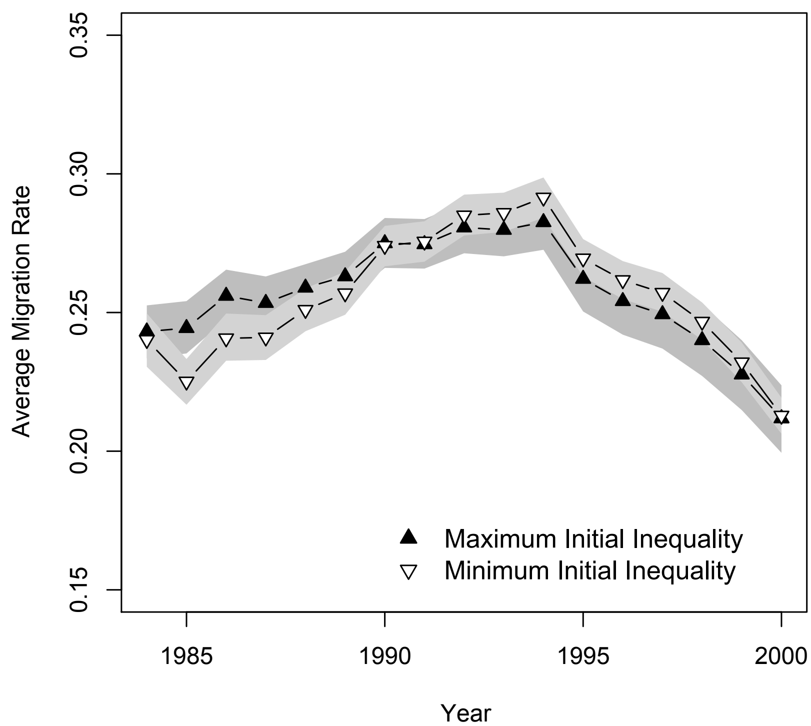

Figure 8: Annual migration rates and the 95 percent confidence region in simulations with maximum and minimum initial inequality in the distribution of village trips.

9. Repeat steps 2-8 M times independently.

10. Compute typical values (e.g., means) of the predicted probabilities over the M replications.

This algorithm is repeated for each scenario S of the initial distribution of village trips.

Figure 8 shows the average migration rate observed in simulations under two scenarios. With minimum initial inequality, we equally distribute the

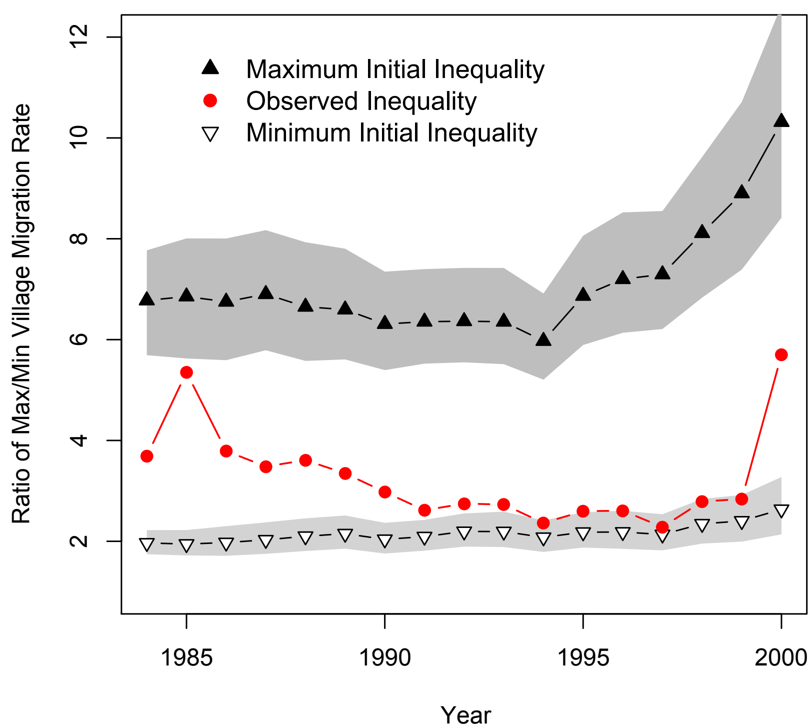

Figure 9: Inequality in village migration, R_t, and the 95 percent confidence region in simulations with maximum and minimum initial inequality in the distribution of village trips.

aggregate number of trips across villages in 1984. With maximum initial inequality, we assign the total number of trips to one randomly-selected village, giving all other villages zero initial trips. Minimum initial inequality case leads to slightly lower average migration until year 1990, and the two scenarios are indistinguishable thereafter.

Figure 9 displays the ratio of the largest to the smallest annual village migration rate, R_t, observed in the data and under the two simulation scenarios. In the minimum initial inequality case, since all villages start at the same point, inequality in village migration rates does not grow over time. In this case, the cumulative mechanism identified in the model does not lead to increasing inequality in village migration. By contrast, with maximum initial inequality, initial inequality increases at a high rate after year 1995. The observed inequality in the data, expectedly, falls between the minimum and maximum inequality cases. The two extreme case scenarios provide upper and lower bounds for the potential inequality outcomes.

Simulation exercise thus links our estimates from the individual-level model to aggregate patterns of inequality between villages. Depending on the initial distribution of village trips, in a period of 16 years, the cumulative mechanism identified in our model could sustain or double inequality in village migration rates.

Conclusion

Hierarchical models are commonly used in sociology, chiefly to study the effects of social context on individual outcomes. In our application, we examined the effects of households and villages on rural-urban migration in northern Thailand. With survey data on individuals at many points in time, individuals also formed contexts for migration decisions in any particular year. In data with this structure, we could specify as many as four hierarchies of random effects—at the individual, household, village, and village-year levels.

The nesting of observations within layers of social context creates data analytic and substantive challenges. From the viewpoint of data analysis, a variety of equally plausible models can be specified to capture the multilevel structure of the data. From a substantive viewpoint, individual outcomes may aggregate to reshape the contexts in which the actions of individuals are determined. Though hierarchical models are common in sociology, the data analytic problem of model comparison, and the substantive problem of the aggregative effects of individual outcomes are often ignored.

Our analysis takes advantages of MCMC methods to fit hierarchical models, compare alternative models, and study the aggregate implications of the models. The problem of model fit was studied with both DIC statistics and posterior predictive checks. Both approaches yielded similar answers. Migration models including individual and village random effects fitted similarly well, but both were inferior to a model that allowed village effects to vary over time. Though the DIC statistic indicated the superiority of the village-year model, posterior predictive check showed that this model better captured the observed trend in inequality in migration across villages.

We conducted a simulation exercise to help aid interpretation of the model parameters. The simulation experiment showed how the initial inequality in patterns of migration across villages influenced inequality in migration 16 years later. Inequality in migration nearly doubled where the initial distribution of migration was highly unequal. Had the initial distribution been equal across village, this distribution would have remained largely unchanged.

In sum, MCMC computation for hierarchical provides an enormously flexible tool for analyzing contextual data. Far beyond the problems of estimation and inferences, posterior simulation with MCMC provides an important basis for data analysis and model interpretation. Though MCMC methods have so far seen relatively little application, they hold enormous promise for the analysis of hierarchical models in sociology.

References

Curran, Sara R., Filiz Garip, Chang Chung, and Kanchana Tangchonlatip. 2005. "Gendered Migrant Social Capital: Evidence from Thailand." *Social Forces* 84:225–255.

DiPrete, Thomas A. and Jerry D. Forristal. 1994. "Multilevel Models: Methods and Substance." *Annual Review of Sociology* 20:331–357.

Garip, Filiz. 2008. "Social Capital and Migration: How Do Similar Resources Lead to Divergent Outcomes?" *Demography* 45:591–617.

Gelman, Andrew, Xiao-Li Men, and Hal Stern. 1996. "Posterior Predictive Assessment of Model Fitness Via Realized Discrepancies." *Statistica Sinica* 6:733–807.

Gelman, Andrew and Donald Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 4:457–511.

Mason, William M., George Y. Wong, and Barbara Entwisle. 1983. "Contextual Analysis through the Multilevel Linear Model." *Sociological Methodology* 14:72–103.

Massey, Douglas S. 1990. "Social Structure, Household Strategies, and the Cumulative Causation of Migration." *Population Index* 56:3–26.

Massey, Douglas S. and Kristin Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102:939–999.

Massey, Douglas S. and Rene Zenteno. 1999. "The Dynamics of Mass Migration." *Proceedings of the National Academy of Sciences* 96:5328–5335.

Rindfuss, Ronald R., Toshiko Kaneda, Arpita Chattopadhyay, and Chanya Sethaput. 2007. "Panel Studies and Migration." *Social Science Research* 36:374–403.

Spiegelhalter, David J., Bradley P. Carlin Nicola Best, and Angelika van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society Series B* 64:583–616.

Western, Bruce. 1999. "Bayesian Analysis for Sociologists: An Introduction." *Sociological Methods and Research* 28:7–34.