2017

Population Movement in Japan: A Hierarchical Bayesian Approach

Elliot S. Oblander
University of Pennsylvania

Population Movement in Japan: A Hierarchical Bayesian Approach

Abstract
As Japan's population ages, the shifting age distribution threatens to destabilize economic and social conditions. Exacerbating this issue is increasing urbanization that leaves vulnerable demographics isolated in more rural regions. To make meaningful statements about the future of Japan's demographic distribution, it is necessary to analyze population movement within the country. To this end, we perform a descriptive analysis examining the net immigration rates into each prefecture of Japan from other prefectures over the course of 2004 to 2013. In particular, we propose a Bayesian regression model of net immigration rates which incorporates eects of census variables, latent differences between prefectures, and anomalous shocks in wake of the 2011 Tohoku earthquake and subsequent nuclear meltdown. We use two-component spike-and-slab priors on regression coefficients that allow for selective shrinkage of parameters. We further propose a framework for predicting from the model and demonstrate that it provides accurate predictions even for years for which covariate values are not known. Our model is seen to give robust predictions of immigration rates, while also yielding valuable insights about the potential factors influencing migration between regions of Japan.

Keywords
demography; immigration rates; penalized regression; spike-and-slab priors; mixed-effects modeling

Disciplines
Business

Population Movement in Japan:
A Hierarchical Bayesian Approach

Elliot S. Oblander[1]

An Undergraduate Thesis submitted in partial fulfillment

of the requirements for the Wharton Research Scholars Program

Faculty Advisor:

Shane T. Jensen

Associate Professor of Statistics

The Wharton School of the University of Pennsylvania

May 2017

[1]Elliot Oblander is a candidate for the Bachelor of Science in Economics at the Wharton School of the University of Pennsylvania, concentrating in Statistics and Actuarial Science (email: Elliot.Oblander.wh17@wharton.upenn.edu, web: http://elliotshinoblander.com).

Abstract

Population Movement in Japan: A Hierarchical Bayesian Approach

As Japan's population ages, the shifting age distribution threatens to destabilize economic and social conditions. Exacerbating this issue is increasing urbanization that leaves vulnerable demographics isolated in more rural regions. To make meaningful statements about the future of Japan's demographic distribution, it is necessary to analyze population movement within the country. To this end, we perform a descriptive analysis examining the net immigration rates into each prefecture of Japan from other prefectures over the course of 2004 to 2013. In particular, we propose a Bayesian regression model of net immigration rates which incorporates effects of census variables, latent differences between prefectures, and anomalous shocks in wake of the 2011 Tōhoku earthquake and subsequent nuclear meltdown. We use two-component spike-and-slab priors on regression coefficients that allow for selective shrinkage of parameters. We further propose a framework for predicting from the model and demonstrate that it provides accurate predictions even for years for which covariate values are not known. Our model is seen to give robust predictions of immigration rates, while also yielding valuable insights about the potential factors influencing migration between regions of Japan.

Keywords: demography; immigration rates; penalized regression; spike-and-slab priors; mixed-effects modeling

1 Introduction: Japan Today

Japan today faces unique social, demographic, and economic issues. Economic growth has stalled over the last 20 years (Japan's "lost decades"), while the birth rate has declined far below the replacement rate, causing rapid aging and depopulation of Japanese society. Women's participation in the labor force remains low, and there is a massive shortage of labor. These trends form a vicious cycle: day cares struggle to maintain adequate levels of staffing, and so young mothers are discouraged from working and having children by the inaccessibility of day care facilities (Osaki, 2016); in turn, the low birth rate and low female labor force participation rate will continue to exacerbate labor shortages in the future.

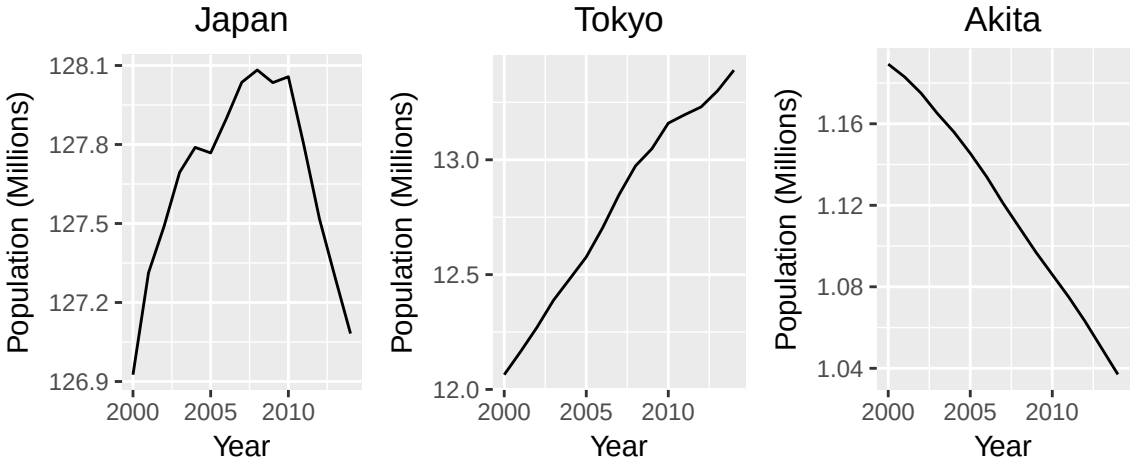
In addition to the low birth rate, over 25% of Japan's population was at least the age of 65 in 2013, and this is expected to increase to nearly 40% by 2060 (The Economist, 2014). As a result, nursing care services have not been able to keep pace with the increase in elderly population (Aoki, 2016). As the ratio of retired to working-age population in Japan continues to grow, the labor shortage will continue to get worse.

In the coming decades, Japan's demographics will continue to undergo massive changes, with the aggregate population expected to decline by over 30% in the next 45 years (The Economist, 2014). This will necessarily spur social, political, and economic changes as well.

However, digging deeper into the aggregate trends present in Japan's demographics, we can see from regional census data (Statistics Bureau of Japan, 2016) that there is substantial variation in these trends. Figure 1 shows the population of Japan and two illustrative prefectures since the year 2000. We can see that population growth has stopped, and since 2010 the population has declined by over a million. This has resulted in the total population being about the same in 2014 as it was in 2000. However, this trend is not consistent in Japan as a whole. For instance, the highly urbanized Tokyo Prefecture saw an 11% increase in population; on the other hand, the very rural Akita Prefecture saw a 13% decline.

Additionally, we can see that there is significant heterogeneity in the population density of each prefecture of Japan. As of 2013, the overall average population density in Japan

Figure 1: Population Trends for Japan and Select Prefectures



was 1,042 residents per inhabitable square kilometer, with the least dense prefecture being Hokkaido (the northernmost large island of Japan) at 245 and the densest being Tokyo Metropolis at 9,554. Additionally, the age distribution of the population varies varies greatly by region. Figure 2 gives a map color-coded by population density as of 2013. Meanwhile, Figure 3 shows the breakdown of Japan's population by 5-year age group as of 2013, both for Japan as a whole (the bars) and by prefecture (the lines).

From Figure 3, we see that the prefectures track the overall age distribution closely, but some prefectures have much different distributions, especially with respect to the 80+ age group. Since this age group has the greatest variation, the spatial distribution of proportion 80+ population as of 2013 is illustrated in Figure 4.

Depending on the prefecture, the proportion of the population that is over 80 years old can vary from 5.2% (in Saitama) to 11.4% (in Shimane), a rather striking difference. A possible driver of this is that younger people from rural areas may tend to move to cities (for work, education, etc.), while older people may elect to stay in their rural residence due to limited mobility and other factors.

What is evident from these illustrative figures is that regional variations are key to understanding Japan's demographic crisis, which looks very different between highly urbanized areas (such as Tokyo and Kyoto) and rural areas (such as Akita and Aomori). In particu-

Figure 2: Population per Inhabitable Square Kilometer by Prefecture in 2013

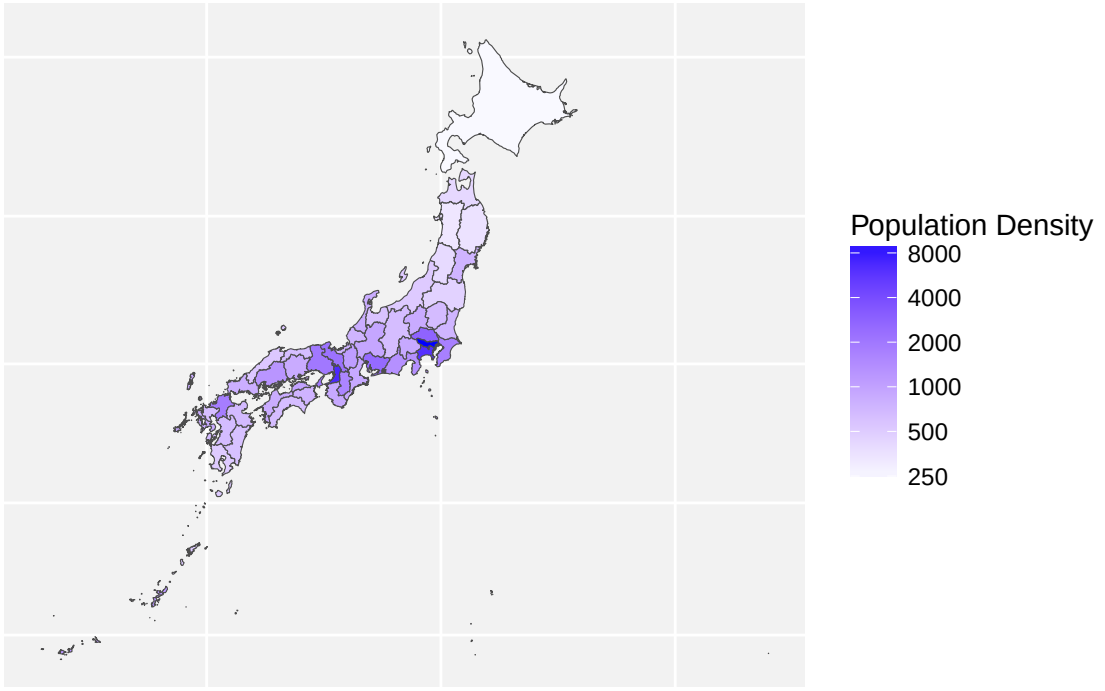


Figure 3: Age Distribution Overall and by Prefecture

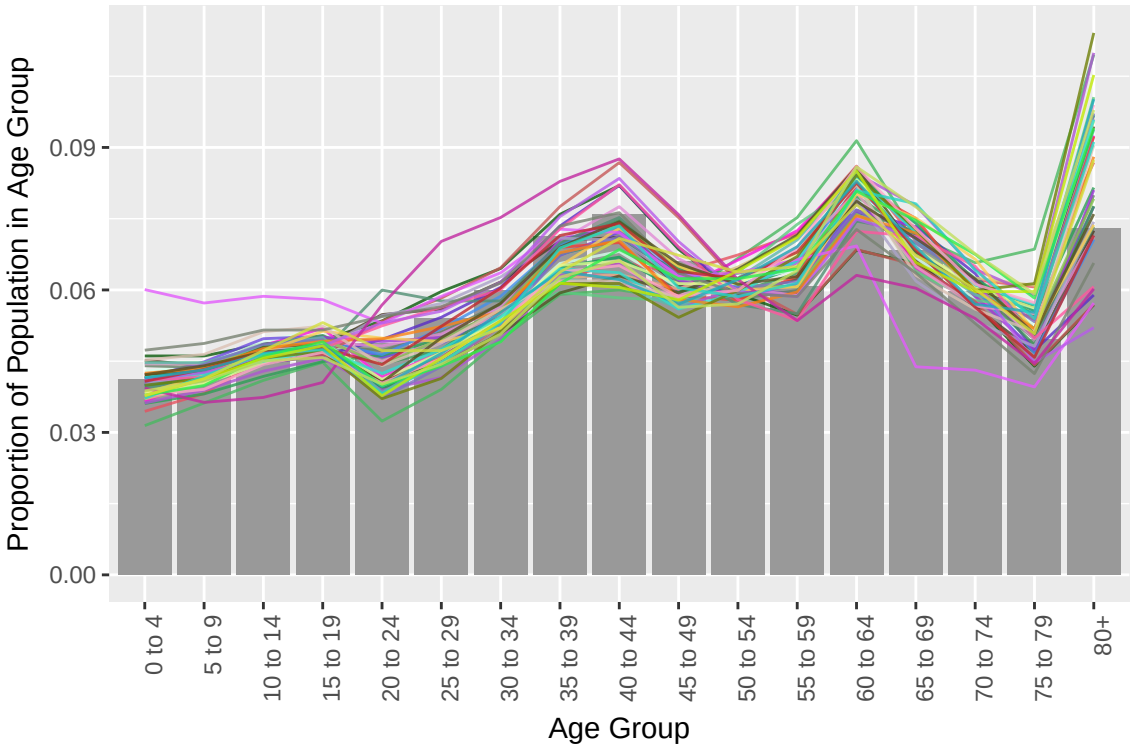
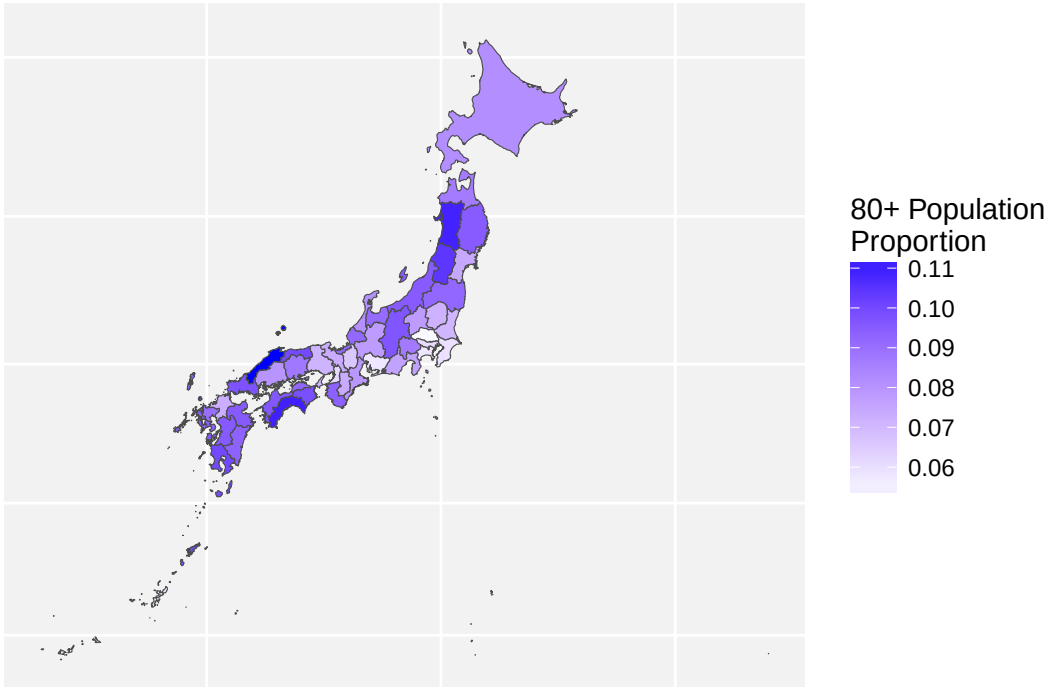


Figure 4: Proportion of Population Over the Age of 80 by Prefecture



lar, Japan is becoming increasingly urbanized as younger people move from rural areas into urban areas, while birth rates in urban areas are much lower than in rural areas (perhaps partially due to the inaccessibility of day care services, as mentioned above). This leaves a large number of childless young people in urban areas, with shrinking super-aged populations in rural areas. Hence, understanding the spatial aspects of Japan's demographic changes is key to understanding the future of Japan's population and how policies can be constructed to ensure the continued thriving of both rural and urban communities around the country.

Clearly, the future of Japanese demographics and its implications for socioeconomic livelihood are of great interest to a variety of Japan-specialist audiences. While current demographic issues such as depopulation are well-known and well-discussed in Japanese academic and political circles, relatively few analyses have specifically looked at regional variations in these triends from a quantitatively rigorous perspective. We believe that much can be learned through this disaggregated approach—policies and initiatives to help Japan prepare for its future must not only consider population changes in the country as a whole, but

also regional changes. Only then can policy be effectively tailored to account for the unique characteristics of each region within Japan.

Thus, in this project, we seek to move beyond aggregate trends and contribute to discussion around Japan's demographic changes by performing a descriptive analysis of population movement in Japan and providing a framework for obtaining robust predictions of future population movement. In particular, we construct a regression model to analyze the net immigration rates into each prefecture of Japan from other prefectures over the course of 2004 to 2013. Through this analysis, we hope to highlight how aggregate population trends vary regionally due to population movement within Japan, so that policymakers can better understand Japan's demographic crisis and derive meaningful predictions and insights about future population movement.

2 Literature

Thus far, as alluded to in Section 1, relatively little English academic literature has performed quantitatively rigorous analysis of Japan's recent demographic trends. However, there is a small but growing body of literature on urban analysis that helps pave the way for our analysis.

Much of the existing literature arises from spatial econometrics, and much recent work is focused on the rapid development of cities in China and India. For instance, Zhu et al. (2015) perform a sophisticated analysis of the development of urban recreational business districts in Beijing using unsupervised learning methods, while Liu and Guo (2013) take an econometric approach and use vector autoregressive time series models to analyze the causal relationships between urban features and economic factors, aggregated to the city level, for many cities in China. Meanwhile, Harari (2016) analyzes the compactness of the shapes of developing cities in India as they expand, and how this affects the economic livelihood of the city in question. While these analyses tend to examine economic outcomes, they can easily be generalized

to demographic outcomes. And, whereas these papers are primarily focused on inference, others have taken approaches that can directly inform inform business and policy decisions, such as Johann et al. (2014), which takes an operations research perspective combined with a social impact perspective to determine the optimal locations of supermarkets that could reduce the extent of food deserts in Philadelphia.

Narrowing our focus to literature on Japan, we find that much of the literature is geared towards forecasting energy demand and optimizing energy planning, a sensible focus after the nuclear meltdown caused by the 2011 Thoku earthquake. For instance, see Wang et al. (2014), who analyze the optimal placement of renewable energy sources in Fukushima (the prefecture of the nuclear meltdown). A number of papers, such as Ohtsuka et al. (2010), are focused on the forecasting of electricity demand across regions of Japan.

Another focus of the literature is on transportation infrastructure. Studies such as Koike et al. (2009) and Matsunaka et al. (2013) examine the existence and development of infrastructures (the former on expressway development and the latter on urban public transportation) and transit times, and how these relate to economic and demographic factors. Additionally, some literature such as Sun et al. (2014) and Nishida et al. (2014) focuses on informatization of Japan and how the variability in adoption of modern technology affects regional economies across prefectures in Japan.

Lastly, a handful of papers are similar in scope to the literature that we saw for other countries. For instance, Kakamu et al. (2008) use spatial modeling to examine the relationships between crime levels in different regions of Japan, while Takagi et al. (2012) perform a similar analysis to Roh and Lee (2013) on the relationship between social capital and crime victimization through hierarchical spatial regressions. One unique and very relevant study is by Zhang (2014), who performs an in-depth spatiotemporal analysis of the migration patterns of Chinese nationals living in Japan.

While these analyses are somewhat disjoint, they provide a sense of the current scope and extent of literature on this topic. Our project is most similar in scope to Zhang (2014), but

focuses specifically on total net population movement between prefectures of Japan rather than on migration of foreign nationals into Japan.

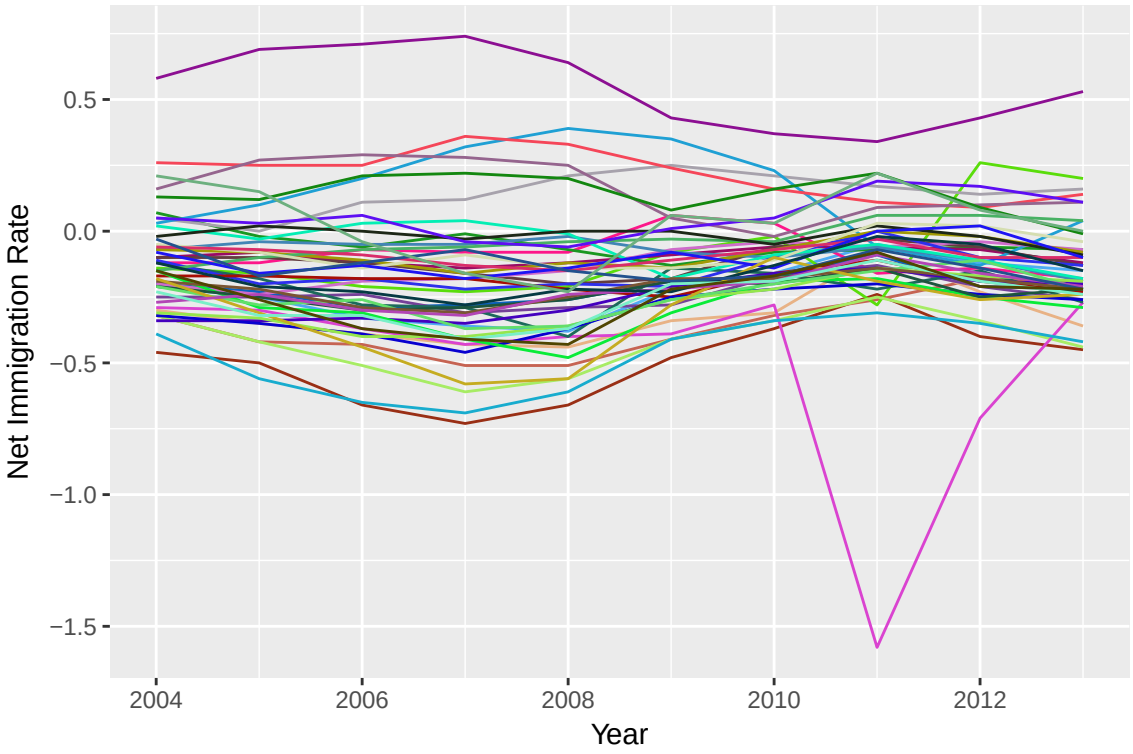
3 Census and Migration Data

The bulk of the data for this project comes from the Statistics Bureau of Japan through their online portal e-Stat (Statistics Bureau of Japan, 2016), and unless otherwise stated, data mentioned in this paper is either retrieved from the Statistics Bureau of Japan or derived directly from their retrieved data. This resource has a wide variety of publicly available data collected by various government bureaus. Of particular interest are two types of data: regional census data and migration data. The regional census data includes many statistics on demographics, geography, economy, housing, welfare usage, education, and health for each of Japan's 47 prefectures. The migration data gives the net migration rates into each prefecture from other prefectures. We restrict our consideration of census variables to those that are interpretable, non-missing, and non-redundant. Several derived variables are included as well. After extensive narrowing down of our data to variables that meet these criteria, we retain 91 variables that serve as indexes of various regional demographic and economic factors. Additionally, for creating spatial visualizations of data and model outputs, we match the data up with political boundary shapefiles from the Global Administrative Areas database (GADM Database, 2016).

For our model, we use data on the net immigration rates of each prefecture of Japan over the course of ten years. For each year from 2004 to 2013 and for each of the 47 prefectures in Japan, we have the net rate of immigration from other prefectures per 100,000 population; for instance, in 2004, Hokkaido had a net migration rate of -0.21, indicating that for every 100,000 residents of Hokkaido, 0.21 moved out of the prefecture into other prefectures (net of the residents who moved into the prefecture from elsewhere in Japan).

Additionally, as mentioned above we have 91 explanatory variables such as demographic/census

Figure 5: Net Immigration Rates Over Time by Prefecture

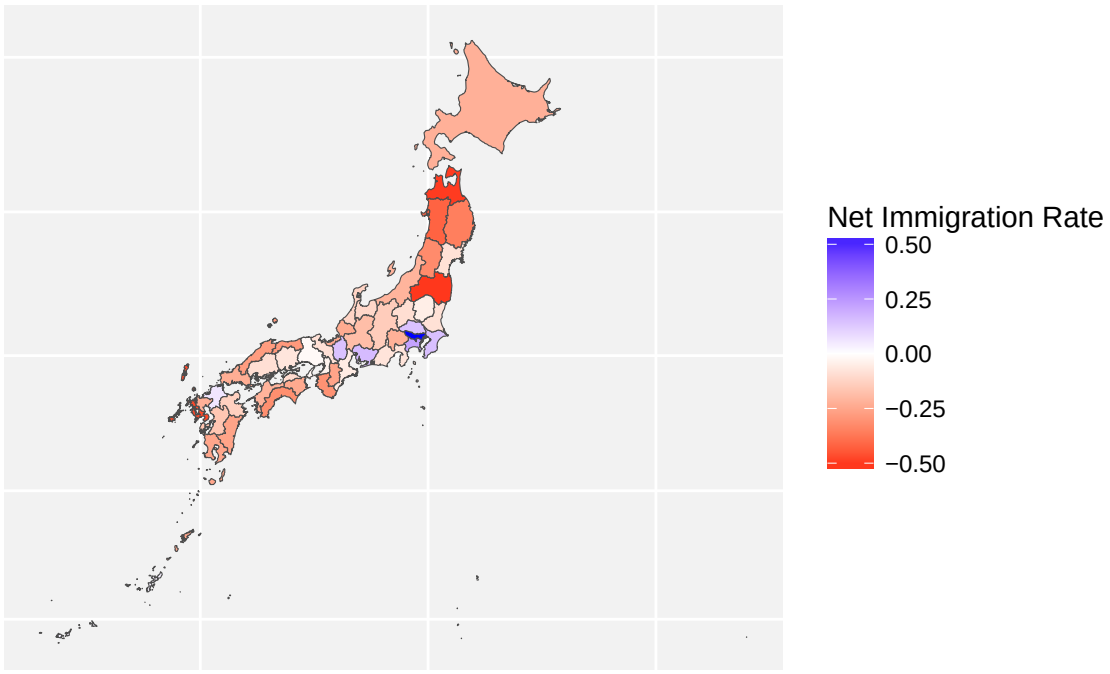


data, density of public resources like schools and parks, and government spending on welfare and other programs. For our analysis, we normalize the explanatory variables to ensure that the posterior coefficients are not affected by the scale of the variables.

As seen in Figure 5, the net immigration rates are centered around 0 (as they should be, since the immigration rates will net to near 0 across all prefectures). A huge outlier is seen in 2011—this is Fukushima, which experienced a nuclear meltdown after the 2011 Tōhoku Earthquake that drove many residents out of Fukushima and nearby prefectures into other regions due to fear of radiation.

We can also visualize the spatial variation of immigration rates. Figure 6 gives a map of Japan, with each prefecture color-coded by its average net immigration rate over the 10-year data window. We see net positive migration rates to the Tokyo and Kyoto areas, and negative migration rates for almost all other regions. This is consistent with overall population changes by prefectures in this time period (although differences in age distributions

Figure 6: Average Net Immigration Rates, 2004-2013



contribute to the overall population changes in addition to migration).

In the next section, we discuss formulation of a regression model that is able to capture these trends based on both obversable and latent variation between prefecture characteristics, and which is able to assess the anomalous effects in 2011 after the Tōhoku earthquake.

4 Model Formulation

To model the net migration patterns, we posit a regression model which has random intercepts and random effects for the effect of the 2011 earthquake. Given the small number of repeat observations for each prefecture, the covariate effects are left as homogeneous fixed effects. Denoting y_{it} as the net migration rate and X_{it} as the explanatory variables for prefecture i in year t, the model formulation is as follows:

$$y_{it} = \alpha_i + \gamma_i \cdot \delta_{t,2011} + \beta^T X_{it} + \epsilon_{it}$$

$$\epsilon_{it} \sim N(0, \sigma^2)$$

where $\delta_{a,b}$ is the Kronecker delta function. Thus α_i is the intercept of prefecture i, γ_i is the shock to prefecture i in 2011, and β is the p-vector of regression coefficients for the explanatory variables (which is assumed to be homogeneous across prefectures).

The residual standard deviation is given an improper flat prior:

$$p(\sigma) \propto 1$$

The prior distribution for the intercepts is assumed to be normal with unknown mean and variance, to allow for overdispersion in migration propensities due to latent differences between prefectures. Improper flat priors are placed on the mean and standard deviation of α_i. That is:

$$p(\alpha_i|\alpha_0, \sigma_\alpha^2) = N(\alpha_0, \sigma_\alpha^2)$$

$$p(\alpha_0, \sigma_\alpha) \propto 1$$

For the γ_is, we expect a priori that only some prefectures, particularly Fukushima and the surrounding area, will have γ_is that are relatively large in magnitude; meanwhile, most prefectures are expected to have been impacted minimally by the earthquake and nuclear meltdown. Similarly, for the β_ks, we expect that some covariates will have primarily spurious correlations with the outcomes, whereas others will substantially impact outcomes. Thus, we wish to employ a prior specification that can enforce shrinkage on parameters. However, commonly used penalized estimation procedures such as the LASSO produce bias by penalizing all parameters equally (Hastie et al., 2009). To avoid this, we require a prior specification that induces *selective shrinkage*, wherein coefficients are differentially penalized based on their relative degree of influence: less influetial coefficients are penalized more heavily to prevent overfitting, while coefficients demonstrated to be highly influential are penalized less to reduce the bias on their estimates.

To this effect, we employ a spike-and-slab prior (Ishwaran and Rao, 2005), which is a 2-component finite mixture of normal distributions centered at 0, one with low variance (the "spike") and one with higher variance (the "slab"). The proportion w of mass in the higher-variance component (also known as the complexity parameter), gives the prior probability mass placed on the slab. The prior variance of the spike component shrinks the γ_is closer to 0, while the variance of the slab component enforces less shrinkage. The relative variance and membership probabilities of the two components are left as parameters of the model, to be estimated simultaneously with the coefficients.

We assume the standard deviation of the spike is some proportion $v \in [0, 1]$ of the standard deviation of the slab, and that some proportion $w \in [0, 1]$ of the prior mass is concentrated on the slab. Uniform priors are placed on v, and w, and the standard deviation of the slab is given a weakly informative prior to prevent boundary solutions. Separate hyperparameters are estimated for the β_k and γ_i distributions. That is:

$$p(\gamma_i|w_\gamma, \sigma^2_\gamma) = w_\gamma N(0, \sigma^2_\gamma) + (1 - w_\gamma)N(0, (v_\gamma \sigma_\gamma)^2)$$

$$p(v_\gamma, w_\gamma) = 1$$

$$p(\sigma^2_\gamma) = \text{Gamma}^{-1}(0.00001, 0.00001)$$

$$p(\beta_k|w_\beta, \sigma^2_\beta) = w_\beta N(0, \sigma^2_\beta) + (1 - w_\beta)N(0, (v_\beta \sigma_\beta)^2)$$

$$p(v_\beta, w_\beta) = 1$$

$$p(\sigma^2_\beta) = \text{Gamma}^{-1}(0.00001, 0.00001)$$

where $\text{Gamma}^{-1}(a, b)$ represents the inverse-gamma distribution. Note that, when $w = 1$ and/or $v = 1$, this reduces to a regular normal prior specification as used for α_i. This completes the prior and likelihood specifications.

5 Model Estimation and Selection

The model is implemented in a fully Bayesian manner using a Hamiltonian Markov Chain Monte Carlo (H-MCMC) algorithm. Posterior sampling is performed using the No U-Turn Sampler (NUTS) algorithm, as implemented in STAN (Carpenter et al., 2017; Hoffman and Gelman, 2014). NUTS has several advantages over more traditional MCMC algorithms such as Gibbs sampling and Metropolis-Hastings. Namely, it utilizes information on the gradient of the likelihood to reduce autocorrelation and speed up convergence; additionally, it allows for more flexible model specifications as it does not require conditional conjugacy.

To assess which model specification to use for further analysis, we test the in-sample and out-of-sample fit of several nested model specifications. This is vital to performing a robust analysis, since beyond fitting the data well, we also want to test the generalizability of the model beyond the training data to ensure we are not overfitting. As such, the following nested model specifications are tested:

1. Fixed Intercept Model: an intercept-only model with one fixed intercept across all prefectures.
2. Random Intercept Model: an intercept-only model allowing for intercepts to vary across prefectures.
3. Fixed Intercept Regression Model: a regression model with one fixed intercept across all prefectures.
4. Random Intercept Regression Model: a regression model allowing for intercepts to vary across prefectures.
5. Full Model, Normal Prior: a regression model including all the features of Model 4, also allowing for random shocks in 2011. Normal priors are placed on both the regression coefficients and the shock coefficients.
6. Full Model, Spike-and-Slab Prior 1: a model with the same features as Model 5, but utilizing a spike-and-slab prior on the 2011 shock coefficients.
7. Full Model, Spike-and-Slab Prior 2: a model with the same features as Model 5, but

utilizing spike-and-slab priors on both the 2011 shocks and the regression coefficients.

Each of the models can be seen as a simplified version of Model 7 with some parameters fixed. In particular:

- Model 7 reduces to Model 6 when $w_\beta = 1$ and/or $v_\beta = 1$.
- Model 6 reduces to Model 5 when $w_\gamma = 1$ and/or $v_\gamma = 1$.
- Model 5 reduces to Model 4 when $\gamma = 0$.
- Model 4 reduces to Model 3 when $\sigma_\alpha = 0$.
- Model 4 reduces to Model 2 when $\beta = 0$.
- Model 3 reduces to Model 1 when $\beta = 0$.

The specifications for each of the nested models are analogous to the full model, and each of the models tested is also estimated in a fully Bayesian manner using the NUTS algorithm. The in-sample fit is calculated using the Deviance Information Criterion (DIC), a hierarchical Bayes' analog of the Akaike Information Criterion (AIC), as specified by Spiegelhalter et al. (2002). The implied effective model dimension based on deviance is also calculated.

The DIC is calculated as:

$$DIC = -2E_\theta[\log(p(y|\theta))] + p_D$$

where θ is the vector of model parameters and p_D is the effective dimension of the model, which is calculated as:

$$p_D = -2\left(E_\theta[\log(p(y|\theta))] - \log(p(y|E[\theta]))\right)$$

Since the DIC is a simple function of the data and the posterior distribution of the model parameters, it is also convenient to calculate: once posterior samples are obtained, it is trivial to aggregate the samples to calculate DIC without needing any additional computationally intensive sampling procedures.

The out-of-sample fit is approximated through 10-fold cross-validation. The folds are

Table 1: Comparison of Model Fits

Model Name	Dimension	DIC	RMSE (Y)	MAD (Y)	RMSE (P)	MAD (P)
1. Fixed Intercept	2	-78.5	0.222	0.158	0.225	0.160
2. Random Intercept	46	-688.9	0.117	0.081	0.225	0.160
3. Fixed Intercept Regression	76	-945.9	0.111	0.077	0.186	0.133
4. Random Intercept Regression	89	-1020.4	0.102	0.070	0.176	0.121
5. Full, Normal Prior	127	-1137.7	0.099	0.066	0.162	0.111
6. Full, Spike-and-Slab Prior 1	117	-1140.8	0.102	0.067	0.161	0.109
7. Full, Spike-and-Slab Prior 2	115	-1144.5	0.103	0.067	0.168	0.112

constructed in 2 ways: year-based and prefecture-based. In the former case, the model is trained on 9 years of data in the dataset and used to predict the values for the remaining year. In the latter case, the prefectures are split at random into 10 groups, and the model is trained on 9 groups and used to predict the values for the remaining group (the same group partitions are used for all models to reduce stochastic error in fit metric calculations). Root Mean Squared Error (RMSE) and Mean Absolute Deviation (MAD) are then calculated for the two sets of predictions. For the models incorporating shocks in 2011, the shocks are included in the model when 2011 is one of the training years and are left off when 2011 is the holdout year (since there would be no observations to identify the effects).

The metrics are reported in Table 1. "Dimension" denotes the effective size of the model as implied by its deviance, while "(Y)" and "(P)" denote whether an error metric was calculated using year-based or prefecture-based holdouts, respectively.

We see that each added layer of complexity on the model uniformly improves the DIC, and usually improves out-of-sample fit metrics as well. Unsurprisingly, models allowing for random intercepts and 2011 shocks tend to do much better on year-based holdouts than prefecture-based holdouts, as they are able to take advantage of each prefecture's observed outcome in other years, whereas the prefecture-based holdouts are a "cold start" that require simply drawing values of α_i and/or γ_i from their respective prior distributions without incorporating any observed outcomes for that prefecture.

Perhaps more surprising, however, is that the prefecture-based holdout performance actually improves substantially when random intercepts and shocks are introduced to the model.

This seems counterintuitive due to the "cold start" mentioned above—the model is not able to take into account past information on the prefectures to be predicted. However, since the model allows for latent differences in baseline immigration rates, it helps reduces the tendency of the model to overfit the covariate coefficients. Hence, despite the random effects adding many parameters while adding no new information that can be used for prediction, there is still a reduction in overfitting that results in improved out-of-sample forecasts.

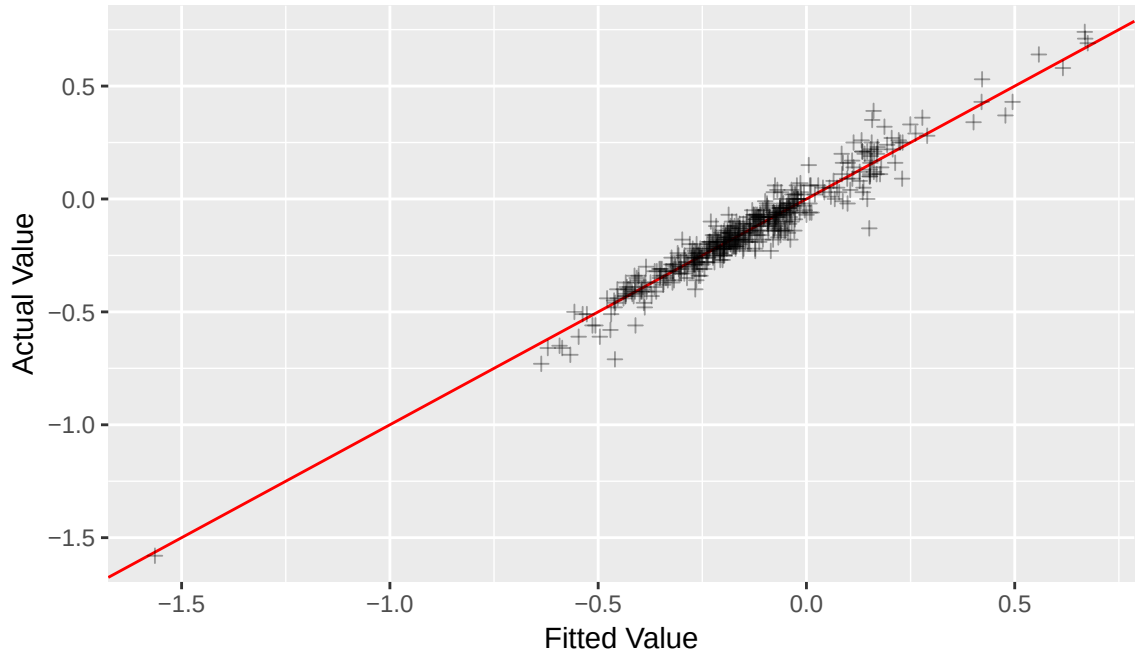
Where we stop seeing uniform improvement is in comparing prior specifications of the full model. It is interesting to note that, despite having more complex formulations, the spike-and-slab models have fewer effective dimensions than the normal prior model, since they allow for selective shrinkage. The three full models have almost identical out-of-sample performance, and while the deviance on the spike-and-slab models are very slightly worse, the reduced effective dimension offsets this and results in a better overall DIC.

Thus, the spike-and-slab prior sacrifices almost nothing in terms of fit, while being more parsimonious. Additionally, it is more compelling from a generative standpoint: as mentioned in the model formulation, we a priori expect some covariates to have minimal impact on migration rates, and for some (if not most) prefectures to experience only small abnormal effects due to the Fukushima nuclear meltdown. Hence we expect the inferences from the model that allows for selective shrinkage to be more robust. For these reasons we choose to use Model 7 for all further analyses.

6 Regression Results

The following section discusses the results of the model. All inferences are based on 200,000 posterior samples from Model 7 as defined in the previous section.

Figure 7: Posterior Mean Fitted Values versus Actual Values ($R^2 = 93.6\%$)



6.1 Fitted Values and Residuals

First, we check the posterior mean fitted values of each datapoint and compare this to observed migration rates. The fitted value plot is given in Figure 7.

Clearly, there is a strong correspondence between the fitted and actual values. Additionally, as seen in Figure 8, the residual plot and histogram (with a kernel density estimate superimposed) do not show any particular evidence of heteroskedasticity or skewness.

However, this fit is not particularly impressive given how highly parameterized the model is. Thus, we also assess out-of-sample fit based on the cross-validation procedure described in the previous section. This tests the generalizability of the parameters to predicting outcomes for prefectures and/or years not in the training sample. The fitted value scatterplots are given in Figure 9.

The correspondence is remarkly strong, even out-of-sample. The prefecture holdout performance is worse than the year holdout, as it is not able to utilize other years' data to identify the prefecture-specific intercepts (and so the intercepts must be drawn directly from

Figure 8: In-Sample Residual Diagnostic Plots

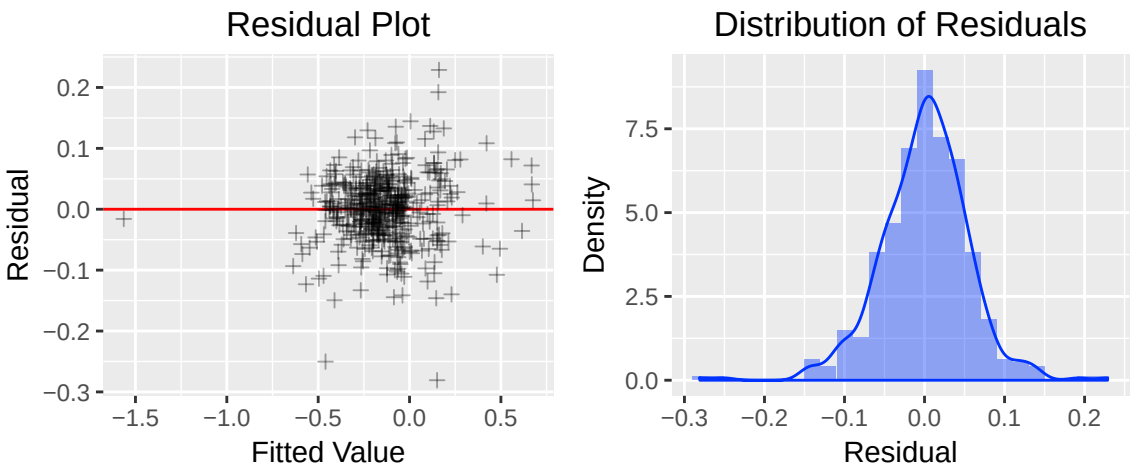


Figure 9: Out-of-Sample Fitted Value Plots

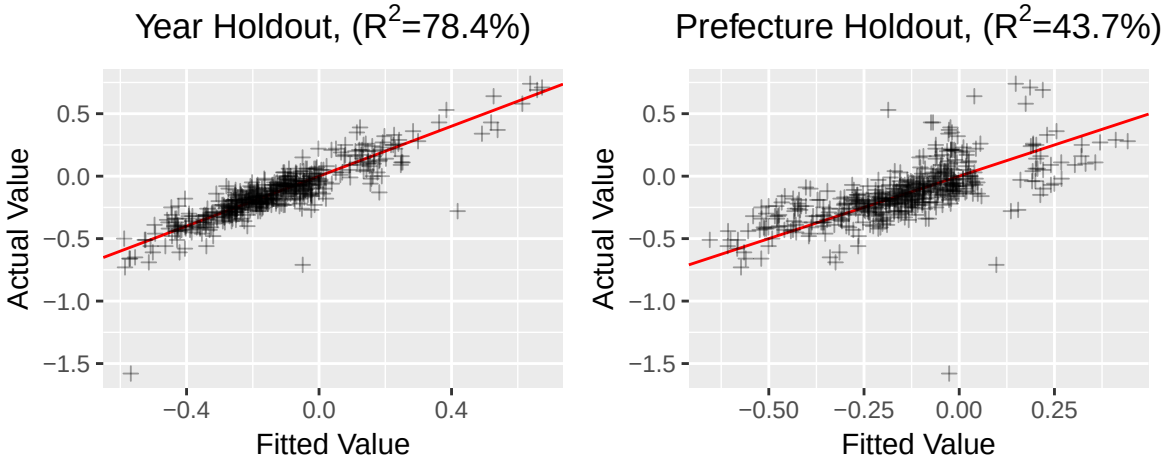
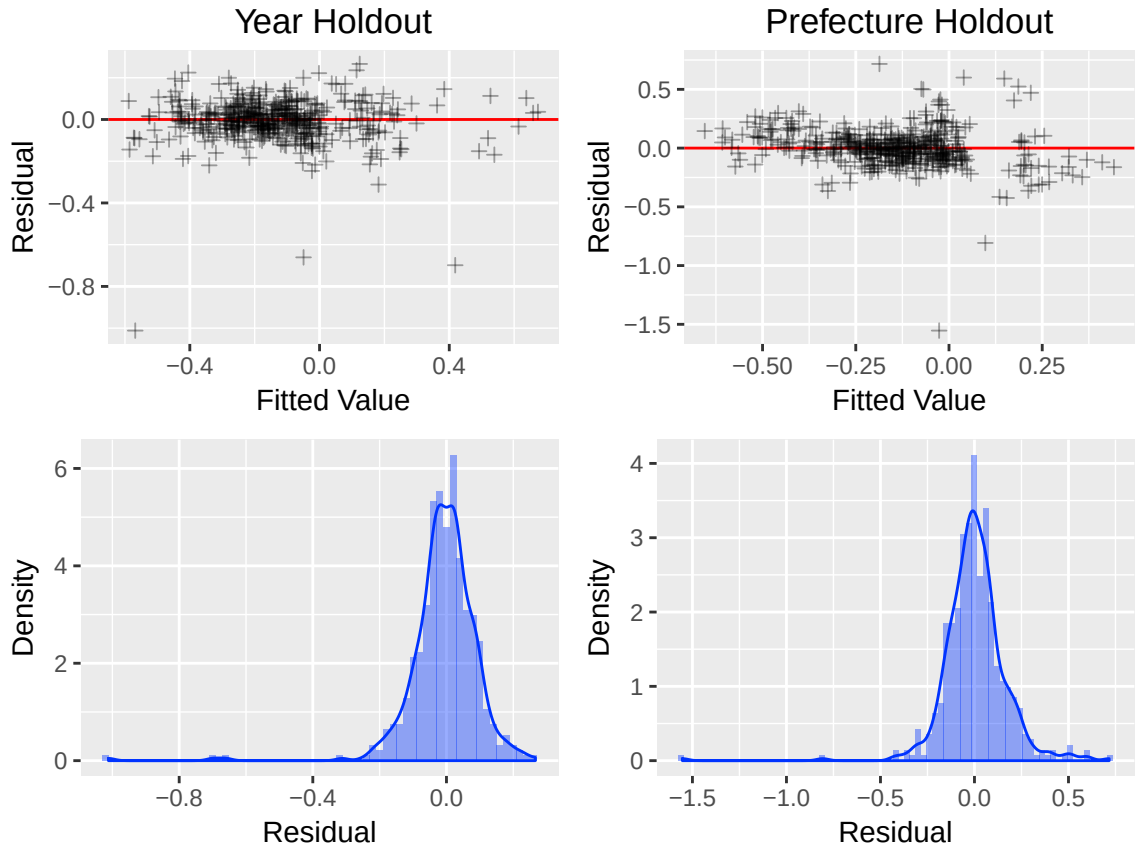


Figure 10: Out-of-Sample Residual Diagnostic Plots



the prior). Nonetheless, the R^2 is over 40%, validating that the covariates have substantial predictive power even out-of-sample. The year-based holdout, which is able to take advantage other years' observations to identify intercepts, gets an out-of-sample R^2 of nearly 80%.

Both models are not able to predict the 2011 shocks out-of-sample (since there are no repeat observations), which introduces some extreme outliers and resulting skewness in the residuals, but other than that the residuals are fairly well-behaved as well, as seen in Figure 10.

Based on these results, we can be reasonably confident in the model's ability to predict out-of-sample. In the next sections we discuss what insights can be gained from the parameter estimates of the model.

6.2 Parameter Estimates

Next, we discuss the posterior distributions of the model parameters. The posterior mean and 95% credible interval of each hyperparameter are given in Table 2. The subsections hereafter will discuss and interpret these hyperparameters and their corresponding prefecture- and covariate-level parameters.

Table 2: Estimated Hyperparameter Values

Parameter	Posterior Mean	95% Credible Interval
σ	0.064	(0.059, 0.069)
α_0	-0.141	(-0.165, -0.116)
σ_α	0.080	(0.056, 0.111)
σ_β	0.037	(0.021, 0.067)
v_β	0.218	(0.022, 0.552)
w_β	0.359	(0.084, 0.736)
σ_γ	0.793	(0.248, 2.633)
v_γ	0.145	(0.027, 0.342)
w_γ	0.079	(0.009, 0.245)

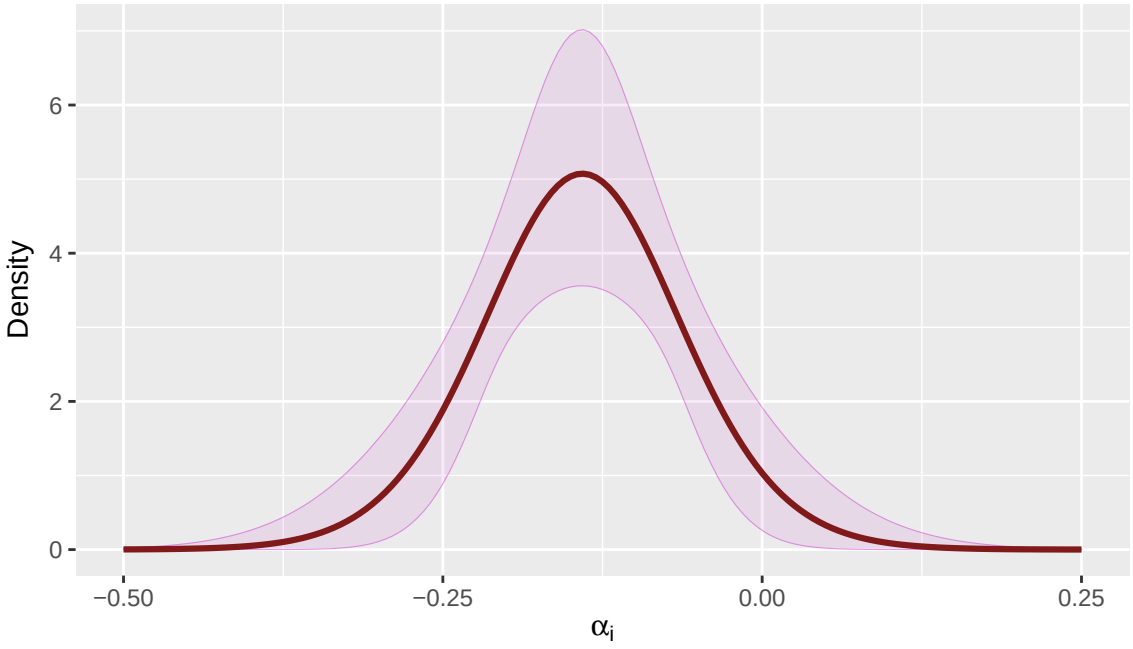
6.2.1 Residual Standard Deviation

The residual standard deviation, σ, is quite well-identified, with a posterior mean of 0.064 and a 95% credible interval of $(0.059, 0.069)$. The sample standard deviation of y is 0.2219, indicating that the regression model captures roughly 91.8% of the variation in migration rates (slightly lower than the R^2 based on posterior means, since it adjusts for uncertainty in parameter estimates).

6.2.2 Intercept Distribution

The parameters α_0 and σ_α in tandem determine the distribution of latent baseline migration propensities. The posterior mean $E[\sigma_\alpha|y] = 0.080$ indicates that there is substantial variation across prefectures in immigration propensities that is due to latent factors that cannot be

Figure 11: Implied Density of α_i



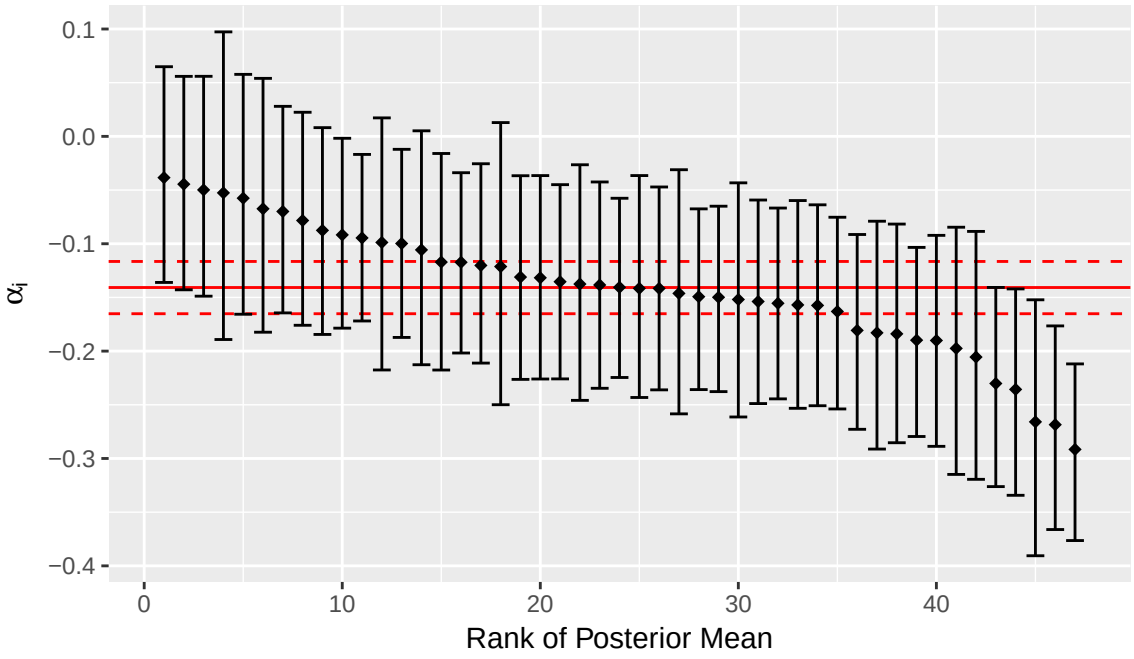
explained by the covariates. The mean and 95% interval of the implied mixing distribution density is given in Figure 11.

Thus, while there is some uncertainty in the mean and variance of the intercepts, the posterior parameters suggest that almost all prefectures have a baseline immigration propensity between -0.5 and 0.25. This is corroborated by the posterior distributions of each of the individual α_i parameters. Figure 12 gives the posterior mean and 95% credible interval for each of the α_is, sorted by their posterior means. The mean and 95% credible interval of α_0 is also plotted in red to show the uncertainty in the overall mean.

We can see from Figure 12 that there is a fair amount of variation in baseline immigration propensities which is not explained by observable variables. Allowing for this sort of overdispersion by prefectures helps improve model fit both in-sample and out-of-sample, as was seen in 1.

Figure 13 gives a map of Japan, color-coded by the posterior mean values of α_i. Red represents a baseline below the average across all prefectures, while blue represents a value above. While there are possible explanations for some of these deviations (for instance,

Figure 12: Posterior Distributions of α_is



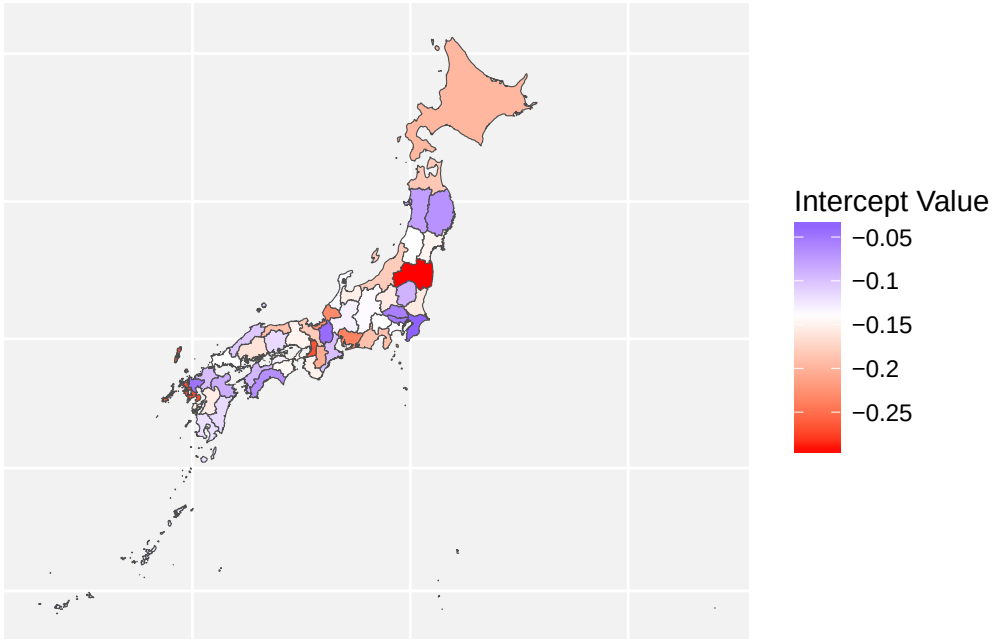
Fukushima's below-average intercept could be due to continuing effects of the Fukushima meltdown in 2012 and 2013), the magnitude of the deviations is fairly small so we refrain from adding any additional parameters to capture these possible explanations.

6.2.3 2011 Shock Distribution

Next, we discuss the distribution of contemporaneous effects experienced in 2011. As seen is Table 2, the spike-and-slab parameters (particularly the vs and ws) are only weakly identified. This is likely due to there being compensatory effects between the parameters—for instance, higher w_γ indicates less overall penalization, but this could be offset by increased penalization from a lower v_γ, resulting in a similar likelihood.

$E[w_\gamma|y] = 0.0793$ indicates that only a small portion of prefectures belong to the "slab" component of the distribution, and $E[v_\gamma|y] = 0.145$ indicates that the "spike" component has substantially lower variance than the slab. However, given the substantial uncertainty in the posterior parameter estimates for the spike-and-slab parameters, our ability to interpret the parameters directly is limited. So, we also consider the shape of the mixing distribution

Figure 13: Map of Baseline Net Immigration Rates



implied by the parameters, and the uncertainty thereof. The mean implied density and 95% credible interval of the shock distribution is given in Figure 14.

The inferred distribution is seen to be stiffly peaked at 0 with a fast dropoff, followed by a long tail, which is consistent with our a priori expectations (that most prefectures will exhibit minimal effects from the meltdown, but some will have very large effects). The exact height of the peak and steepness of drop-off, however, are fairly uncertain. Nonetheless, this sort of high-peaked, heavy-tailed distribution would not be adequately captured by a typical normal distribution, validating the need for this more sophisticated prior specification.

Figure 15 gives the posterior means and 95% credible intervals ordered by posterior mean. We see that a few prefectures have negative shocks, with Fukushima experiencing by far the strongest negative effect, while most prefectures have slight positive shocks.

The selective shrinkage afforded by the spike-and-slab prior allows some prefectures to have shocks far away from 0 without too much shrinkage, while the prefectures with small shocks are penalized more heavily. A typical normal prior penalizes all datapoints equally, and so does not allow for differential shrinkage like this. To illustrate this, we compare the

Figure 14: Implied Density of γ_i

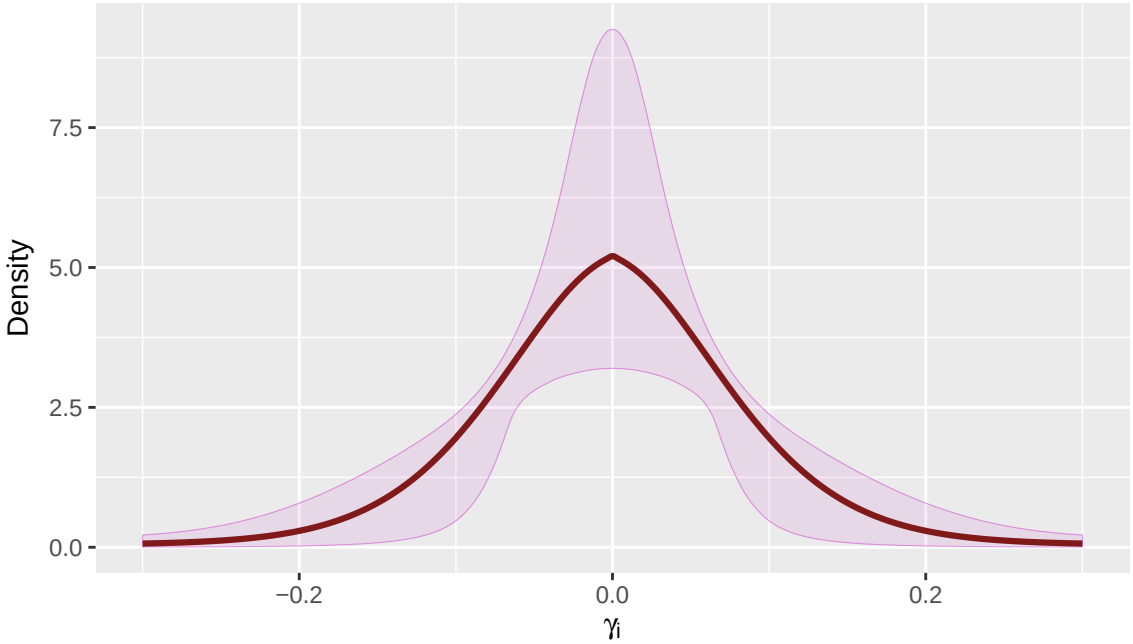


Figure 15: Posterior Distributions of γ_is

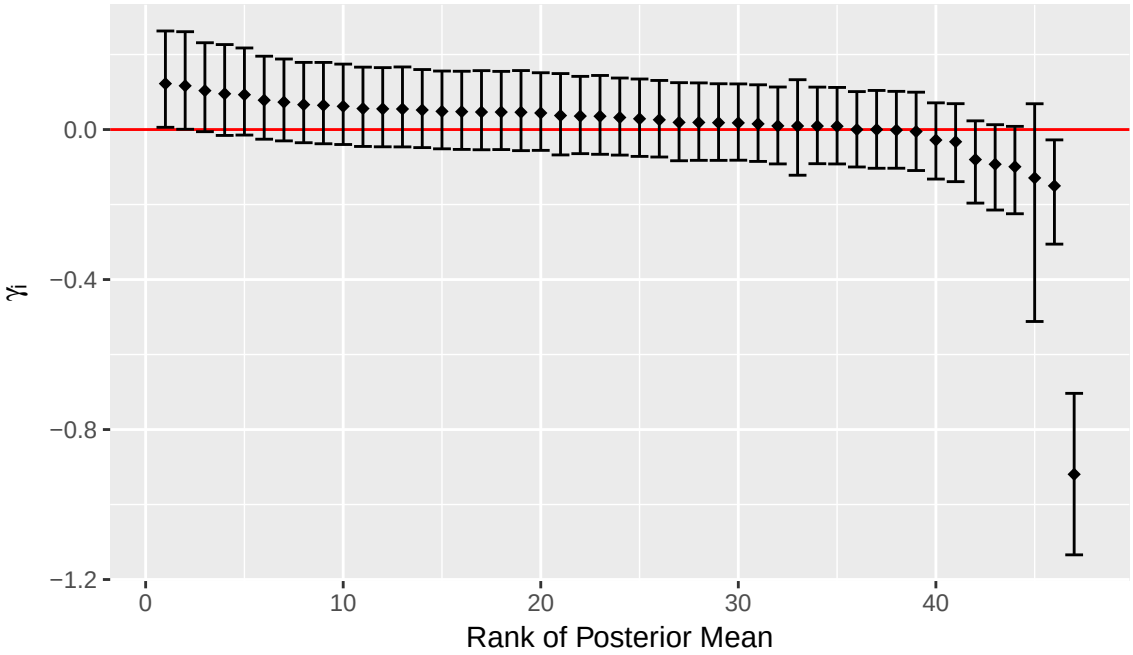
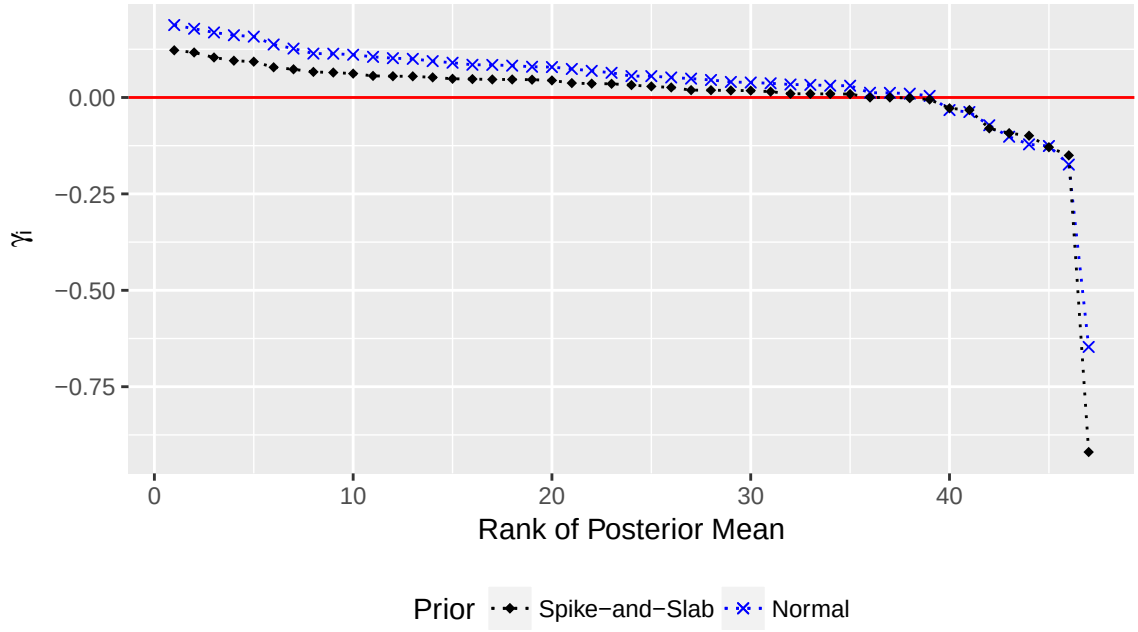


Figure 16: Estimates of γ_i by Prior Specification



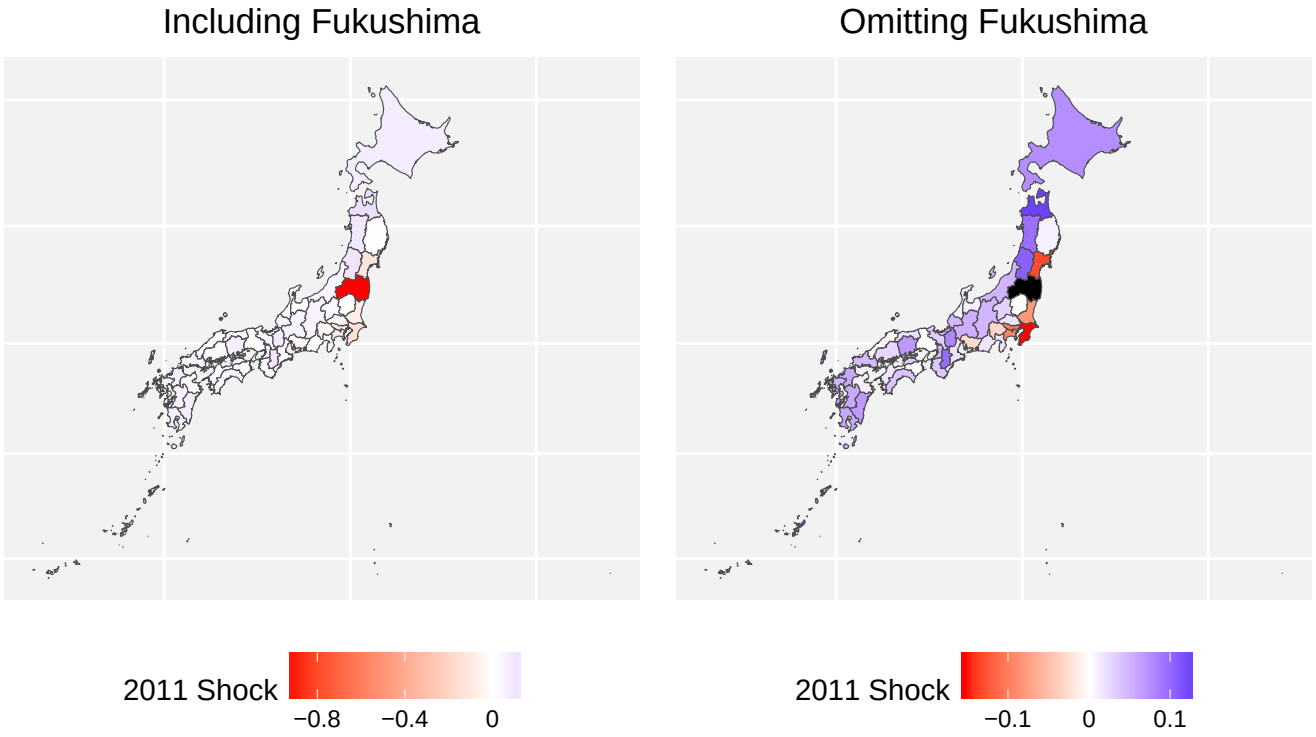
distribution of γ_is based on a normal prior (Model 5) versus a spike-and-slab prior (Model 7) in Figure 16.

We see that, because the normal prior cannot apply selective shrinkage, it must compromise by substantially shrinking Fukushima towards the origin, while shrinking other points less compared to the spike-and-slab prior.

In Figure 17, as with the α_is, we give a map of Japan color-coded by γ_i. Since Fukushima's coefficient is on such a different scale from the other coefficients, we also give a map with Fukushima omitted from the color scale to allow for clearer visual differentiation between the other prefectures.

It is interesting to note that the two prefectures immediately North and South of Fukushima have essentially neutral effects, but the Tokyo area further to the South has negative effects. For the surrounding regions not critically impacted by the earthquake, it is likely that people choose to move due to fear of radiation. For the immediately neighboring prefectures, this effect is presumably offset by physically displaced people from Fukushima evacuating to the neighboring regions. Most other prefectures, especially the ones to the West of Fukushima,

Figure 17: Maps of 2011 Shocks



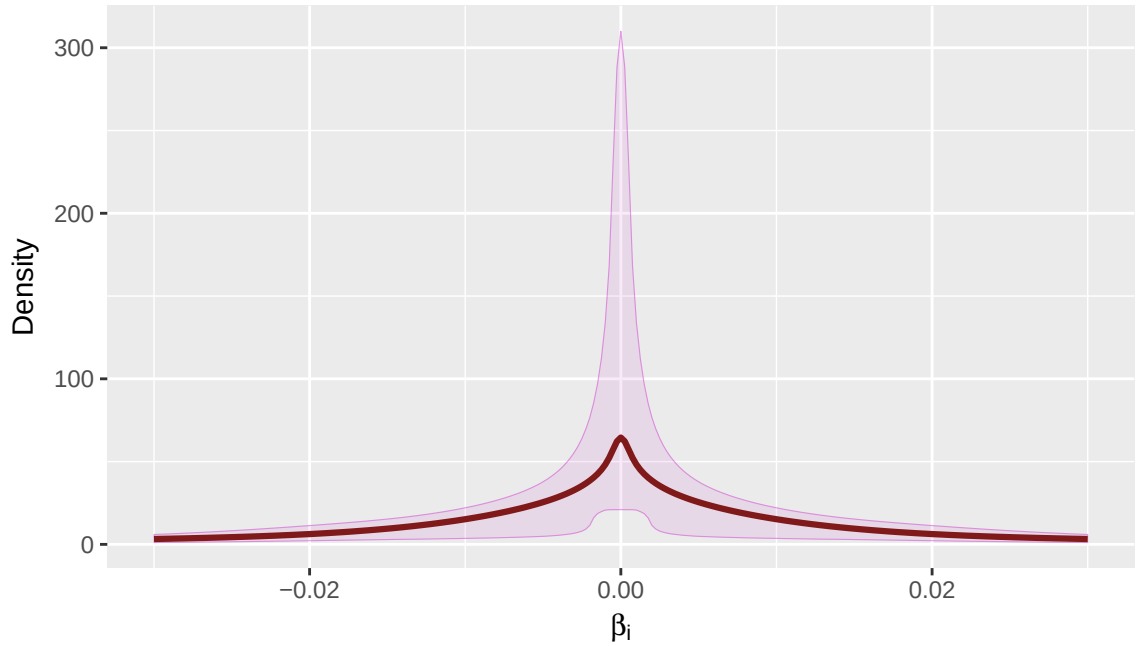
experience positive effects, which is consistent with records of where Fukushima evacuees have moved (Fukushima on the Globe, 2014).

Of course, since the γ_is accommodate any anomalous shock for 2011, some of these effects may be due to other contemporaneous events that occurred in 2011 besides the Tōhoku earthquake and Fukushima meltdown. Nonetheless, the spatial distribution of shocks are sensible and generally consistent with what we would expect to see a priori based on Japan's geography and population distribution.

6.2.4 Regression Coefficient Distribution

Lastly, we discuss the β coefficients and the inferences to be made from the model. First, considering the hyperprior parameters: the posterior mean of σ_β is 0.037, indicating a fairly tight distribution of regression coefficients. The w_β and v_β also indicate a substantial amount of shrinkage, but are difficult to interpret directly due to their rather wide credible intervals.

Figure 18: Implied Density of β_k



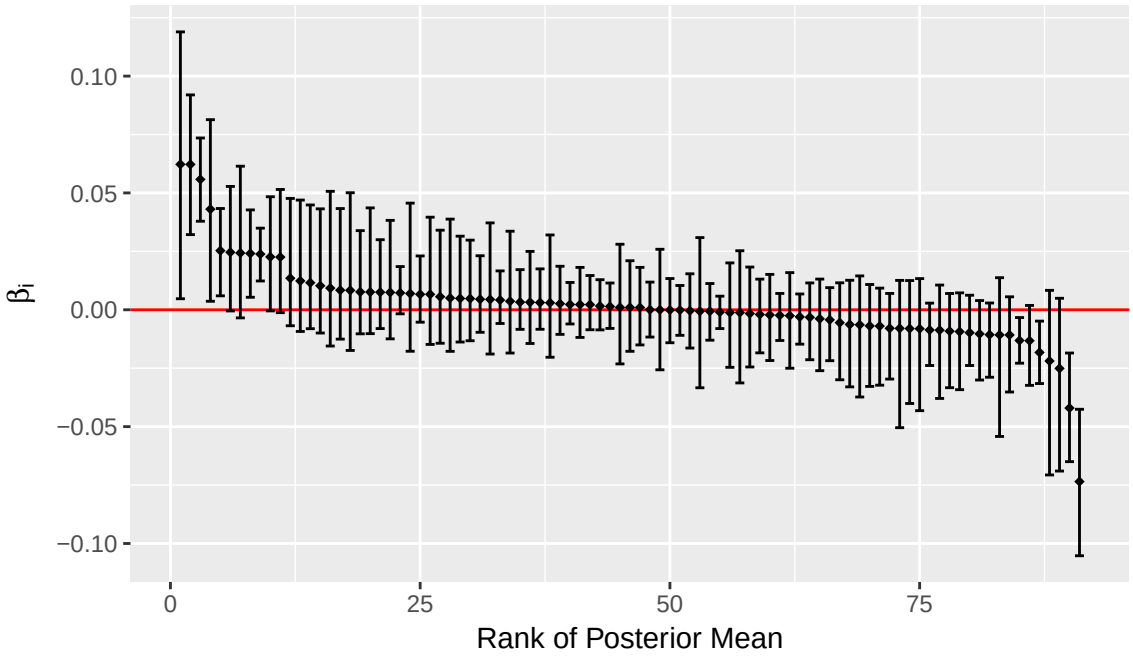
As such, we again focus on the implied shape of the distribution of β_ks. Figure 18 gives the posterior mean and 95% credible interval of the implied density of regression coefficients.

This distribution is seen to be even more steeply peaked than the distribution of γ_is, but also has more uncertainty in how steep the peak is. This distribution indicates heavy shrinkage towards 0 for both the spike and the slab, with substantial uncertainty in how probabilities should be allocated between the the mixture components. As with the distribution of γ_is, a normal distribution would not be flexible enough to accommodate this shape.

Next we look at the posterior distributions of the individual coefficients. Figure 19 gives the posterior mean and 95% credible interval of each of the 91 regression parameters, ordered by posterior mean. Despite the heavy penalization, 11 of the covariates are significant at the 5% level. So, although the model opts for heavy penalization on covariate coefficients, several of them are still important to model fit (as corroborated by the good holdout performance of the model).

To contextualize this distribution, Figure 20 provides a comparison between the coefficient

Figure 19: Posterior Distributions of β_ks

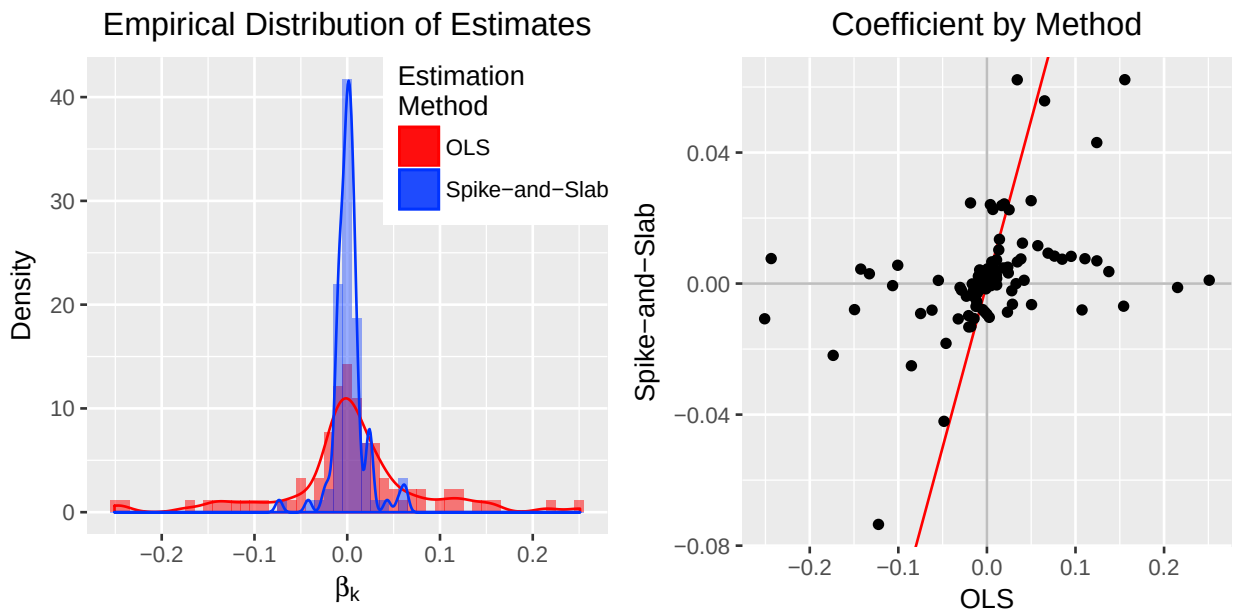


estimates based on the spike-and-slab prior to those based on traditional non-penalized estimation. The panel on the left gives the empirical densities of the distribution of regression parameters with randoms effects penalization (using the posterior means under the spike-and-slab prior) and fixed effects without penalization (using ordinary least squares to estimate the intercepts and coefficients). The panel on the right compares directly the estimates from the two methods, with the red line demarcating what would make a perfect 1:1 correspondence between estimates.

On the left panel, we see that the distribution of penalized coefficients is much lighter-tailed, demonstrating the enormous shrinkage that results from the spike-and-slab prior. The sample variance of the posterior means is 95% lower than the sample variance of the OLS coefficients, indicating heavy penalization. On the right panel, we can see directly the effects of selective shrinkage: most coefficients are pulled sharply towards 0 compared to their OLS estimates, but some are shrunk much less, and a few even have larger coefficients under the spike-and-slab estimation procedure than under OLS.

The predictors which are significant at the 5% level, as well as their respective slopes and

Figure 20: Estimates of β_k by Estimation Method



credible intervals, are presented in Table 3. For scale interpretability, the coefficients have been transformed back to the unnormalized scale. Since there are a substantial number of covariates, typically a multiple hypothesis testing correction would be required to keep the false discovery rate low. However, since the estimation scheme implicitly shrinks coefficients towards zero in a selective manner, this reduces the need to make such corrections. As such, we opt to use a regular 5% significance threshold.

The interpretation of these coefficients is, for the most part, fairly straightforward and intuitive. We see that, for instance, people tend to immigrate more to higher-income prefectures and tend to immigrate less to prefectures where the consumer price index is increasing relatively quickly. Particularly notable is that 4 of the 11 selected variables are related to labor and employment—suggesting that migration patterns are strongly tied to job availability and labor force dynamics. Unsurprisingly, we see that people tend to move to prefectures where there is greater job availability.

However, beyond the inferences obtained from model parameters, we are also interested in predicting future net immigration rates and evaluating how these predictions vary under

Table 3: Estimates of Significant Regression Coefficients

Variable ID	Posterior Mean	95% Credible Interval	Posterior P-Value
#D02206	0.018	(0.001, 0.033)	0.0221
#D0310801	-0.012	(-0.021, -0.003)	0.0053
#D0310901	0.033	(0.008, 0.056)	0.0065
#D0330403	217.225	(112.327, 321.040)	0.0001
#D0330503	53.424	(4.523, 100.958)	0.0239
#F03103	0.202	(0.137, 0.266)	0.0000
#F03302	-0.006	(-0.008, -0.003)	0.0000
#F03402	0.005	(0.001, 0.009)	0.0083
#H06310	-0.241	(-0.373, -0.106)	0.0003
#L04102	-0.016	(-0.027, -0.004)	0.0059
#L04302	0.013	(0.007, 0.018)	0.0000

The descriptions of these covariates are as follows:

- **#D02206**: Taxable income (100,000 yen per taxpayer)
- **#D0310801**: Government expenditure on health (% of administrative spending)
- **#D0310901**: Government expenditure on labor (% of administrative spending)
- **#D0330403**: Elderly welfare costs (100,000 yen per capita)
- **#D0330503**: Child welfare costs (100,000 yen per capita)
- **#F03103**: Ratio of job openings to job seekers
- **#F03302**: High school graduates getting a job outside of the prefecture (%)
- **#F03402**: Unemployment rate of new college graduates
- **#H06310**: Mobile phone subscriptions per capita
- **#L04102**: Annual change in Consumer Price Index, less imputed rent (%)
- **#L04302**: Annual change in residential land prices (%)

counterfactual conditions. Hence, in Section 7, we provide a framework for predicting future outcomes from this model and give illustrative examples.

7 Predictive Analysis

7.1 Covariate Extrapolation

Prediction is an important potential use of this model, but is difficult because covariate values outside of the 2004-2013 range are not known. Thus, in order to predict future outcomes from the model, it is necessary to model the covariate values themselves.

To obtain baseline predictions for the model, we predict future covariate values using a simple linear time trend:

$$X_{itk} \sim N(\xi_{ik} + \eta_{ik}t, \tau_k^2)$$

where X_{itk} is the value for prefecture i in year t of covariate k.

We then use priors on ξ_{ik} and η_{ik} analogous to the prior used for the main model. The parameters are assumed independent across covariates, allowing for each set of parameters to uniquely characterize the time trends for each covariate; a priori there is no reason to believe that time trends for different covariates (e.g. job openings versus mobile phone subscribers) would come from the same distribution.

Specifically, we allow the intercepts for each prefecture to follow a normal hyperprior, and place flat priors on the corresponding mean and variance:

$$p(\xi_{ik}|\xi_{0k}, \tau_{\xi k}^2) = N(\xi_{0k}, \tau_{\xi k}^2)$$

$$p(\xi_{0k}, \tau_{\xi k}) \propto 1$$

Next, for time trends, we would a priori expect some prefectures to have fairly stable

covariate levels, while others may be more volatile; hence, we use a spike-and-slab prior as used for regression coefficients in the full model:

$$p(\eta_{ik}|w_{\eta k}, \tau^2_{\eta k}) = w_{\eta k} N(0, \tau^2_{\eta k}) + (1 - w_{\eta k}) N(0, (v_{\eta k}\tau_{\eta k})^2)$$

$$p(v_{\eta k}, w_{\eta k}) = 1$$

$$p(\tau^2_{\eta k}) = \text{Gamma}^{-1}(0.00001, 0.00001)$$

Using this model specification, we estimate the model for each covariate individually and simulate posterior predictive values of the covariates out to 2020. Using the sampled predictive distributions of the covariates, we combine these with the posterior samples from the full model to get predicted net immigration rates for each prefecture from 2014 to 2020.

7.2 Prediction and Validation

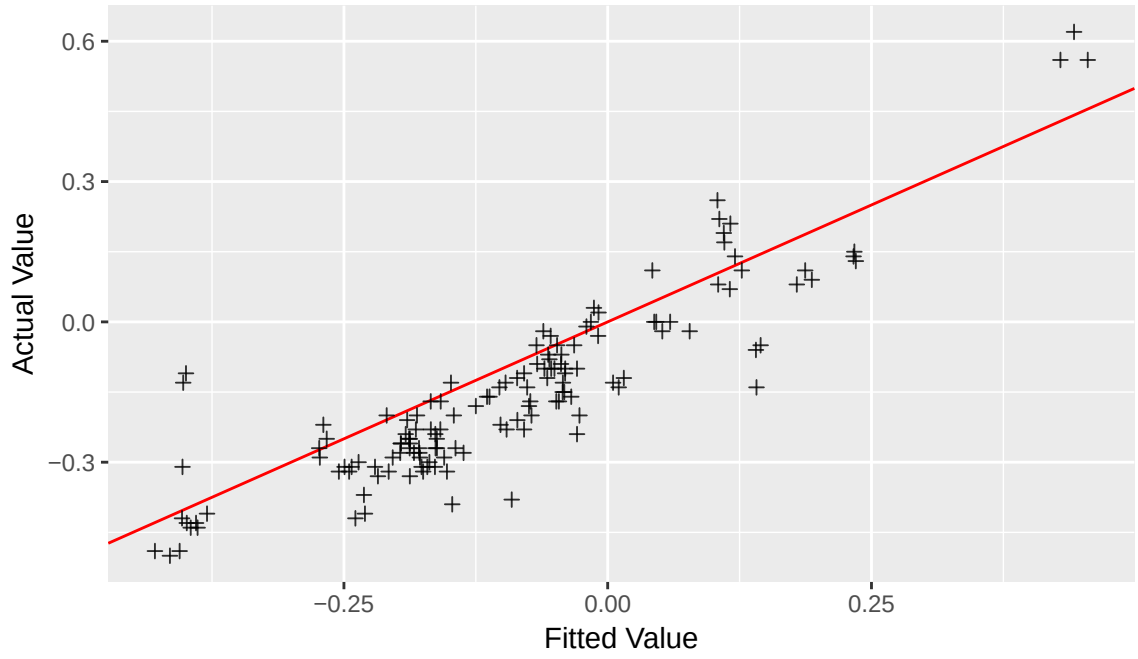
Using the posterior predictive distribution of extrapolated covariate values for 2014 to 2020, we predict the net immigration rate for each prefecture for each year in the prediction window.

While we do not yet have the true covariate values for 2014-2016, we are able to obtain the net immigration rates for each prefecture for these years from the Statistics Bureau of Japan (2016). This provides an opportunity to measure the predictive validity of the model beyond the cross-validation procedure used in Section 5, and tests the ability of the model to generalize to cases where the covariates are not known and must be extrapolated.

Figure 21 gives the fitted value plot for the net immigration rates in 2014 to 2016. The R^2 is almost identical to that of the year-based holdout in Section 6, where covariate values were known. As such, it appears that there is minimal loss in predictive validity when extrapolating covariate values over a short horizon.

However, the fitted value plot also shows signs of the model systematically overpredicting the prefectures with low fitted values—possibly a bias introduced by the covariate extrapo-

Figure 21: Fitted versus Actual Values for 2014-2016 ($R^2 = 80.0\%$)



lation procedure. Nonetheless, the predictions are quite robust, especially for such a short time series. Additionally, this framework is simple to generalize to more sophisticated extrapolation methods.

Next, we consider the predicted net immigration rate for the whole holdout window up to 2020. The predictions are plotted in Figure 22, with only the posterior mean shown to avoid overplotting. We see that the mean predictions for all prefectures are fairly stable, and most are trending slightly upwards, while some stay flat or trend downwards.

Beyond these posterior mean predictions, one of the principal benefits to using Bayesian models is the ability to obtain credible intervals of estimates and predictions. So, beyond considering just the mean predicted trends, we also provide the 95% posterior credible interval of predicted values for two select prefectures (Tokyo and Fukushima) in Figure 23.

Tokyo is predicted to have a gradual slowing of immigration in the coming years, although there is substantial uncertainty in this trend. Even at the higher bound of the credible interval, Tokyo is not predicted to surpass the peak in immigration that it experienced in 2007. Fukushima, meanwhile, is expected to stabilize post-disaster and settle at a slightly

Figure 22: Posterior Mean Predictions Through 2020

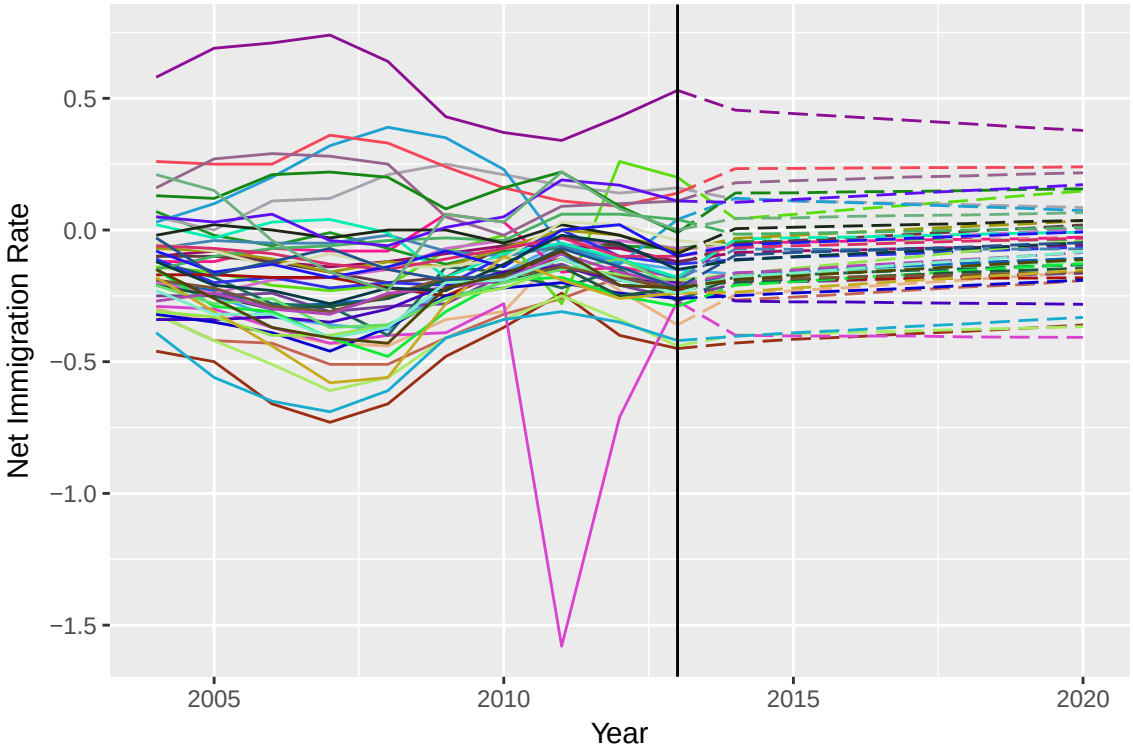
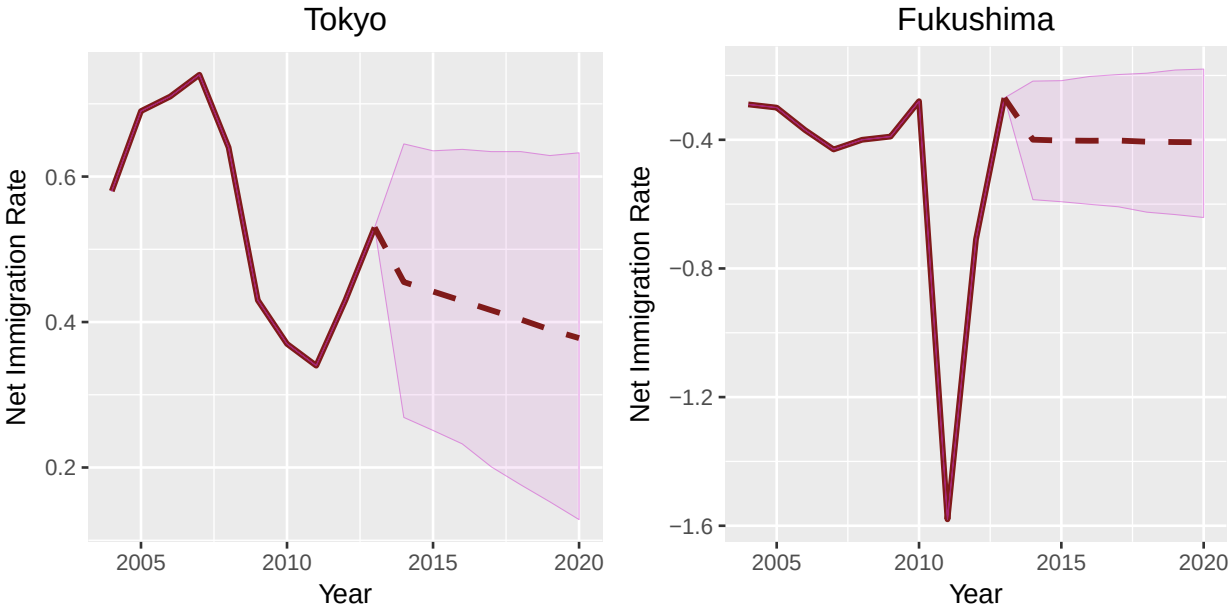


Figure 23: Posterior Mean and Credible Interval of Predictions for Select Prefectures



negative net immigration rate.

These illustrative examples directly use the extrapolated covariate values using our proposed framework, but the methods used here can easily be generalized to other extrapolation methods as well as to counterfactual conditions. For instance, to see the effect of an exogenous shock to one or more covariates, it is simple to perturb the extrapolated covariate values to observe the resulting change in predictions.

8 Conclusions

In this analysis, we have constructed a hierarchical Bayesian regression model to analyze the annual net immigration rates for each of Japan's prefectures from 2004 to 2013. The model posits immigration rates to be explained by a combination of census and economic covariates in addition to latent differences between prefectures. The model additionally allows for contemporaneous shocks in 2011 to account for the abnormalities in migration patterns that arose from the Fukushima nuclear meltdown. The model gives good predictions both in-sample and out-of-sample; and, through the use of prior specifications on the shock coefficients and regression coefficients that enforce *selective shrinkage*, we are able to differentiate between factors that are strongly related to observed migration patterns and those which are negligibly related.

Additionally, our model allows for prediction and counterfactual analysis through our proposed covariate extrapolation framework. Based on recent immigration rates in years for which no covariates were available, we see that our model maintains its predictive validity even in the absence of exact covariate values.

It is important to note that these results are purely descriptive: causal inferences are beyond the scope of this analysis, as there is a clear endogeneity problem with observational data on macroeconomic variables. For instance, the regression coefficients suggest that there is a positive relationship between residential land prices and immigration rates, but this

is presumably due to increased immigration causing an increase in real estate prices (as opposed to the other way around). Similarly, some of the other covariate coefficients also have counterintuitive signs that make interpretation difficult.

Nonetheless, descriptive analysis is useful in that (1) it allows us to narrow down a large set of variables to a handful that are empirically seen to be closely related to migration patterns; and (2) it allows for prediction and counterfactual analysis of future migration outcomes, assuming future values of covariates are determined exogenously.

This is of interest to public policy researchers and practitioners alike. Identification of observational associations are essential to determining what policy variables are important to consider for further analysis: experimental research and econometric methods for mimicking experimantal designs can be used to help establish causal relationships and infer treatment effects, but these methods are much more resource-intensive and less amenable to large numbers of inputs. As such, this analysis helps narrow down the scope of inputs that are worth further investigation. Additionally, even if endogeneity precludes us from performing an operational analysis (e.g. analyzing how to optimally allocate government expenditure to stabilize the spatial population distribution), the ability to predict immigration rates (and uncertainty thereof, thanks to the Bayesian formulation) is still very useful for policymakers, both so that they can prepare appropriately for population inflow/outflow and so that they can perform post hoc assessment of the effectiveness of policy decisions (e.g. analyzing how much immigration rate differed from predictions after a policy change, and how statistically significant that difference is). Thus, this analysis helps provide a descriptive and predictive tool for analyzing migration patterns that can help inform policy decisions and direct future experimental research.

References

Mizuho Aoki. Nursing care workers hard to find but in demand in aging Japan. *The Japan Times Ltd.*, June 2016.

Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 2017.

Fukushima on the Globe. Situation of the Evacuees, 2014. URL `http://fukushimaontheglobe.com/the-earthquake-and-the-nuclear-accident/situation-of-the-evacuees`.

GADM Database. Global Administrative Areas, 2016. URL `http://gadm.org/`.

Mariaflavia Harari. Cities in Bad Shape: Urban Geometries in India. Working paper, 2016.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer-Verlag New York, 2nd edition, 2009. ISBN 978-0-387-84857-0.

Marek Hlavac. *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Harvard University, Cambridge, USA, 2015. URL `http://CRAN.R-project.org/package=stargazer`. R package version 5.2.

Matthew D. Hoffman and Andrew Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15 (Apr):1593–1623, 2014.

Hemant Ishwaran and J. Sunil Rao. Spike and slab variable selection: Frequentist and Bayesian strategies. *The Annals of Statistics*, 33(2):730–73, 2005.

Holger Johann, Margeret Hall, Steven O. Kimbrough, Nicholas Quintus, and Christof Weinhardt. Service District Optimization: Usage of Facility Location Methods and Geographic Information Systems to Analyze and Optimize Urban Food Retail Distribution. *SSRN Electronic Journal*, 2014.

Kazuhiko Kakamu, Wolfgang Polasek, and Hajime Wago. Spatial interaction of crime incidents in Japan. *Mathematics and Computers in Simulation*, 78(2-3):276–82, 2008.

Atsushi Koike, Lori Tavasszy, and Keisuke Sato. Spatial Equity Analysis on Expressway Network Development in Japan. *Transportation Research Record: Journal of the Transportation Research Board*, 2133:46–55, 2009.

Fengjuan Liu and Shengda Guo. Quantitative Analysis of Interaction among Employment in China's Urban Area, Urbanization and Economic Development. *Information Technology Journal*, 12(23):7796–802, 2013.

Ryoji Matsunaka, Tetsuharu Oba, Dai Nakagawa, Motoya Nagao, and Justin Nawrocki. International comparison of the relationship between urban structure and the service level of urban public transportation—A comprehensive analysis in local cities in Japan, France and Germany. *Transport Policy*, 26-39, 2013.

Tetsushi Nishida, James B. Pick, and Avijit Sarkar. Japans prefectural digital divide: A multivariate and spatial analysis. *Telecommunications Policy*, 38(11):992–1010, 2014.

Yoshihiro Ohtsuka, Takashi Oga, and Kazuhiko Kakamu. Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA approach. *Computational Statistics & Data Analysis*, 54(11):2721–735, 2010.

Tomohiro Osaki. Day care crisis stuck in vicious cycle. *The Japan Times Ltd.*, April 2016.

Sunghoon Roh and Ju-Lak Lee. Social capital and crime: A cross-national multilevel study. *International Journal of Law, Crime and Justice*, 41(1):58–80, 2013.

David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64(4):583–639, 2002.

Stan: A C++ Library for Probability and Sampling, Version 2.10.0. Stan Development Team, 2015a. URL `http://mc-stan.org/`.

Stan Modeling Language User's Guide and Reference Manual, Version 2.10.0. Stan Development Team, 2015b. URL `http://mc-stan.org/`.

Statistics Bureau of Japan. e-Stat. database, 2016. URL `http://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do`.

Chuan Sun, Hao Jiao, and Yun Ren. Regional Informatization and Economic Growth in Japan: An Empirical Study Based on Spatial Econometric Analysis. *Sustainability*, 6(10): 7121–141, 2014.

Daisuke Takagi, Ken'Ichi Ikeda, and Ichiro Kawachi. Neighborhood social capital and crime victimization: Comparison of spatial regression analysis and hierarchical regression analysis. *Social Science & Medicine*, 75(10):1895–902, 2012.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL `https://www.R-project.org/`.

The Economist. The incredible shrinking country. *The Economist Newspaper*, March 2014.

Qianna Wang, Martin M'ikiugu, and Isami Kinoshita. A GIS-Based Approach in Support of Spatial Planning for Renewable Energy: A Case Study of Fukushima, Japan. *Sustainability*, 6(4):2087–117, 2014.

Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009. ISBN 978-0-387-98140-6. URL `http://ggplot2.org`.

Changping Zhang. Spatial-Temporal Analysis on Migration of Chinese Registered in Japan. *Journal of International Migration and Integration*, 16(3):701–16, 2014.

He Zhu, Jiaming Liu, Chen Chen, Jing Lin, and Hui Tao. A spatial-temporal analysis of urban recreational business districts: A case study in Beijing, China. *Journal of Geographical Sciences*, 25(12):1521–536, 2015.