A Nonlinear Hierarchical Model for Estimating Prevalence Rates with Small Samples

Xiao-Li Meng*, Margarita Alegria†, Chih-nan Chen‡, Jingchen Liu*

Abstract

Estimating prevalence rates with small weighted samples, especially for rare diseases is a challenging task. We encountered such a situation in the recent National Latino and Asian American Study (NLAAS) on mental health. Due to small sizes of the weighted samples in various age groups, the standard designed-based estimators are highly variable. Bayesian hierarchical modeling offers a more workable approach by incorporating our knowledge on the smoothness of the prevalence rates as a function of age. The non-linear nature of this function, however, presents some intricate modeling issues such as the sensitivity to the link function (for converting a rate parameter onto the real line). In this paper we report our findings and some strategies we adopted to combat such problems.

1 Background and NLAAS

In the last four decades, the United States has experienced an unprecedented wave of immigration, primarily from Latin America and Asia, which presents considerable challenges for health care delivery systems. Unfortunately, the problems in health-care delivery for immigrants are compounded by incomplete data on these populations. National prevalence estimates of psychiatric disorders for the 41 million people of Latino ancestry living in the United States remain elusive because studies fail to disaggregate them by national origin or nativity groups or to consider the heterogeneity between and within Latino groups.

The National Latino and Asian Study (NLAAS) is a nationally representative survey of household residents (ages 18 and older) in the non-institutionalized population of the coterminous United States. Data were collected between May 2002 and November 2003. A total of 4864 individuals, including Latinos, Asians, and whites, were interviewed. Among them, a total of 2554 English and Spanish-speaking Latinos, divided into four strata (Puerto Rican, Cuban, Mexican, and all Other Latinos), comprised the final Latino sample with a response rate of 75.5%. The sample includes an NLAAS Core, designed to be nationally representative of all Latino origin groups regardless of geographic patterns; and NLAAS-HD supplements, designed to oversample geographic areas with moderate to high density (HD) of Latino households. Weighting reflecting the joint probability of selection from the pooled Core and HD samples provides sample-based coverage of the national Latino population.

The NLAAS weighted sample is similar to the 2000 Census in gender, age, education, marital status and geographical distribution, but different in nativity and household income, with more U.S. immigrants and lower income respondents in the NLAAS sample. This discrepancy may be due to, among others, Census undercounting of immigrants, non-inclusion of undocumented workers, lack of fully bilingual interviewers of Latino ethnicity conducting Census interviews, or sample recruitment differences of participants.

*Harvard University
†Cambridge Health Alliance and Harvard Medical School
‡Boston University

2 Goal of Study

In order to compare the prevalence of psychiatric disorders across different ethnic groups, one of the most important variables to control for is age. A conventional way is to estimate the prevalence rate for each age group, and average them according to the census age proportion, that is, to compare the prevalence rate as if all ethnic groups have the same age distribution as the whole population in the country. While a more ideal and informative comparison would be by age groups, in this paper we focus on the age-aggregated comparison mostly because of its common use in current psychiatric literature. The Bayesian method we adopted is particularly useful for making the more detailed comparisons by age groups, precisely because they provide more reliable estimates of age-specific prevalence rates than traditional survey methods can, for reasons we discuss below.

To reliably estimate prevalence rate within each age group, we need to deal with the serious problem of small sample sizes, compounded by the problem of very variable survey weights, which lead to even smaller "effective sample size." That is, often we need to deal with age-groups in which sample sizes vary anywhere from zero to twenty. Standard survey estimators, such as weighted means with jackknife variance estimates, are known to yield very noisy point and interval estimates (or there is no valid estimate if there is no sample in an age-group). For example, Figure 1 shows the observed rates for Cuban male. The rate for adjacent groups jumped up from 9.5% (age 30-34) to 24% (age 35-39), and then fall down to 0% (age 40-44). While underlying rates do vary with age, it is difficult to explain such large fluctuations other than that they are due to sampling errors resulting from small samples and weights with large variations.

In our simulation studies to check the reliability of the traditional methods, we found that such methods not only lead to estimates with very large variance (as it should be given the size of the data) but also unacceptable confidence coverage for resulting interval estimators. The problem is particularly serious for those psychiatric disorders with low prevalence rates (e.g., less than 5%), where we found that a nominal 95% confidence interval may actually have as low as about 50% actual coverages (see Section 5), especially when the rates are very low, say 2%. This is mainly because of the serious skewness in the distribution of the estimator, which makes the large-sample normal approximations underlying the standard methods completely inadequate. Further evidences of the inadequacy of standard estimators for NLAAS studies can be found in Section 5 and Alegria et. al. (2004).

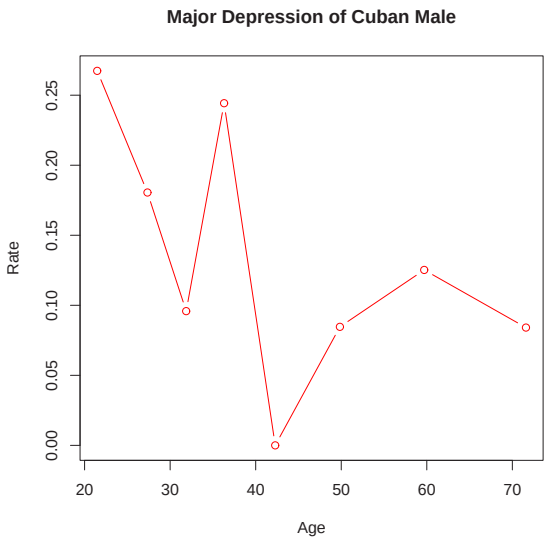

Figure 1: Major Depression for Cuban Male

3 Bayesian Modeling

3.1 A Binomial-like Likelihood Approximation

To combat such a problem, we adopt a nonlinear hierarchical modeling approach (e.g., Gelman et. al., 2002; Gelman and Meng, 2004), a method for dealing with small-sample estimation. The method allows us

to impose reliable prior knowledge to compensate for the large survey variability due to small size and survey conditions (e.g., large variable weights and nonresponse). In our current study, we assume that the logit of the prevalence rate is a quadratic curve as a function of age, based on common observations that the rates tend to increase with age but then "die off" for high age groups. A known interpretation for this "die off" phenomenon is that psychiatric disorders (e.g., major depressions) are often very good predictors for mortality.

In particular, we divide the sample into 8 groups using the census categories, by age 18-24, 25-29, 30-34, 35-39, 40-44, 45-54, 55-64, and above 65 (including 65). For each group, we calculated the weighted mean of the responses, denoted by $\bar{y}_i$, where $i = 1, \ldots, 8$. Let μ_i denote the true prevalence rate of group i . To deal with the complex issue of weighting and survey design, we adopt an approximate likelihood modelling for $\bar{y}_i$ as

$$\mathrm{p}(\bar{y}_i|\mu_i) \propto \mu_i^{\tilde{n}_i\bar{y}_i} \cdot (1-\mu_i)^{\tilde{n}_i(1-\bar{y}_i)}, \tag{1}$$

where $\tilde{n}_i = (\sum_j w_{ij})/\sum_j w_{ij}^2$ approximates the effective sample size, and w_{ij} is the weight of the j^{th} sample in the i^{th} group. Note that in the simple case of independent equal probability sampling, all the weights are identical, $\tilde{n}_i$ will be the same as the real sample size and $\tilde{n}_i\bar{y}_i$ follows exactly Binomial($\tilde{n}_i$,μ_i), such that, (1) will be exact for the special case.

Accepting the approximated likelihood (1) for μ_i, the next step is to put a non-linear regression model to link μ_i to the age variable. Our strategy is to first transform μ onto the real line via a link function G, as routinely done with GLM. We then model the transformed rate, $\xi_i = G(\mu_i)$, to follow a normal model, with the mean a quadratic curve of the age:

$$\xi_i|\beta \ \sim N(\beta_0 + \beta_1 a_i + \beta_2 a_i^2, \tau^2), \tag{2}$$

where $\beta = (\beta_0, \beta_1, \beta_2)^\top$, and a_i is the average age of the i^{th} group. We emphasize that the use of averaged age a_i is rather an ad-hoc approach, which also highlights the sensitivity of the results to our choices of the age groups. Similar to our use of the approximated likelihood (1), we adopt this strategy primarily for simplicity in modeling and for ease of interpretation to researchers in psychiatric and related studies, where the notion of using Bayesian method is still a new one.

Under independent prior on β and τ^2, the resulting posterior distribution is,

$$\begin{aligned} & p(\xi, \beta, \tau^2|y) \propto p(\beta)p(\tau^2)\frac{1}{\tau} \\ & \prod_{i=1}^{8} \Big\{ \exp\big\{\frac{-1}{2\tau^2}(\xi_i - \beta_0 - \beta_1 a_i - \beta_2 a_i^2)^2\big\} \\ & \mu_i^{\tilde{n}_i\bar{y}_i} \cdot (1-\mu_i)^{\tilde{n}_i(1-\bar{y}_i)} \Big\}, \end{aligned} \tag{3}$$

where $\mu_i = G^{-1}(\xi)$.

3.2 Choice of The Link Function

Unlike the common situations with GLM, where the choices of the link function are often not crucial, for our current application, our results are sensitive to the choice of G both because of the small sample sizes and the very low rates we face for some psychiatric disorders. To illustrate this, Figure 2 plots three common link functions: logit, complementary log-log, and Normal inverse CDF (probit).

The primary reason why those link functions lead to different estimate is the behavior of the functions at low probability areas, since most of the average prevalence rates are below 15%, which is to the left of the vertical dashed line. Recall that $\xi = G(\mu) \ \sim X\beta + N(0, \tau^2)$. The smaller τ^2 is, the more $\hat{\xi}$ is pooled towards the regression line $X\beta$. As we can see, logit has the longest negative tail among the three. More specifically, $\mathrm{logit}([0.005, .15]) = [-5.3, -1.7]$, $\Phi^{-1}([0.005, .15]) = [-2.6, -1.0]$ and $-\log(-\log([0.005, .15])) = [-1.7, -0.64]$. This implies that for the same τ^2, the pooling is the most significant for logit and the least for complementary log-log (see Section 6.1 for a discussion on the pooling effect of the Bayesian modeling).

A visually appealing way to investigate the effect of pooling is to inspect the smoothness of the resulting

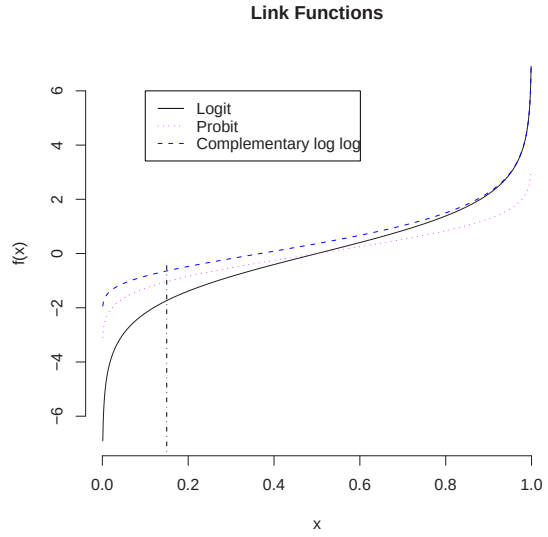

Figure 2: Link Functions

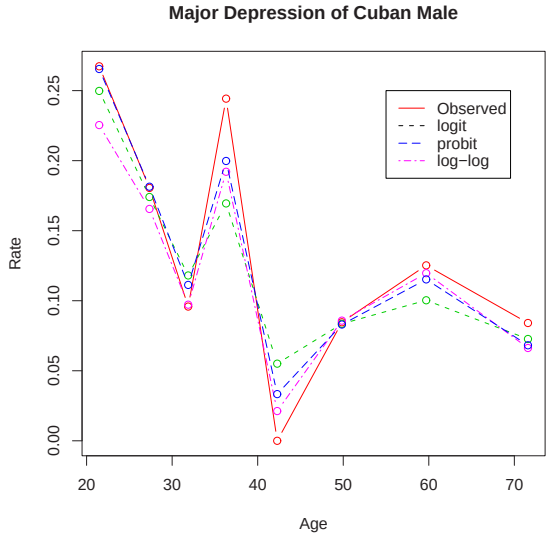

Figure 3: Major Depression Cuban Male

Bayesian estimates as a function of the age. Figure 3 plots these curves under the three link functions for Cuban male, which we have seen in Figure 1. Evidently, the curve from the logit link is most smooth. For this reason and for its easy interpretability and common acceptance in psychiatric studies, our main results are based on the logit link. But we emphasize that the sensitivity to the choice of the link function is an issue that should be recognized. One important fact in our choice should be the amount of smoothness we want to impose on our age curves.

3.3 Choice of Prior

It can be shown that with the likelihood constructed above, the most popular noninformative prior, either $p(\beta) \propto c$ or $p(\tau) \propto 1/\tau$ will lead to an improper posterior distribution. Accordingly, we need to choose a proper prior for both β and τ (or τ^2). Our simulation study shows that the posterior distribution is not very sensitive to the choice of $p(\tau^2)$, as long as it does not put too much mass at or near 0, so we just adopt a conjugate prior, the inverse of χ^2 with three degree of freedom, which has expectation 1 and infinite variance. We emphasis that, however, this insensitivity holds only for a given link function, because, as we discussed in Section 3.2, the pooling effects for different link functions are different with the same value of τ^2.

The prior of β we adopted is a tri-variate normal, such that the prior expectation of $\mu = \sum_{i=1}^{8} p_i \mu_i$ (where p_i is the proportion of the i age group according to the 2000 Census) roughly matches our prior knowledge of the overall prevalence rate, and that its prior variance is relatively large. Figure 4 shows the prior distributions of the average rates as a function of the prior mean of β_0 and a scale factor c (see below), with the prior means of β_1 and β_2 always set to zero. Plots in each column share the same c, and in each row share the same β_0.

Specifically, our choice of prior is as follows,

$$p(\tau^2) \propto \tau^{-5} \exp\{-\frac{1}{2\tau^2}\}, \tag{4}$$

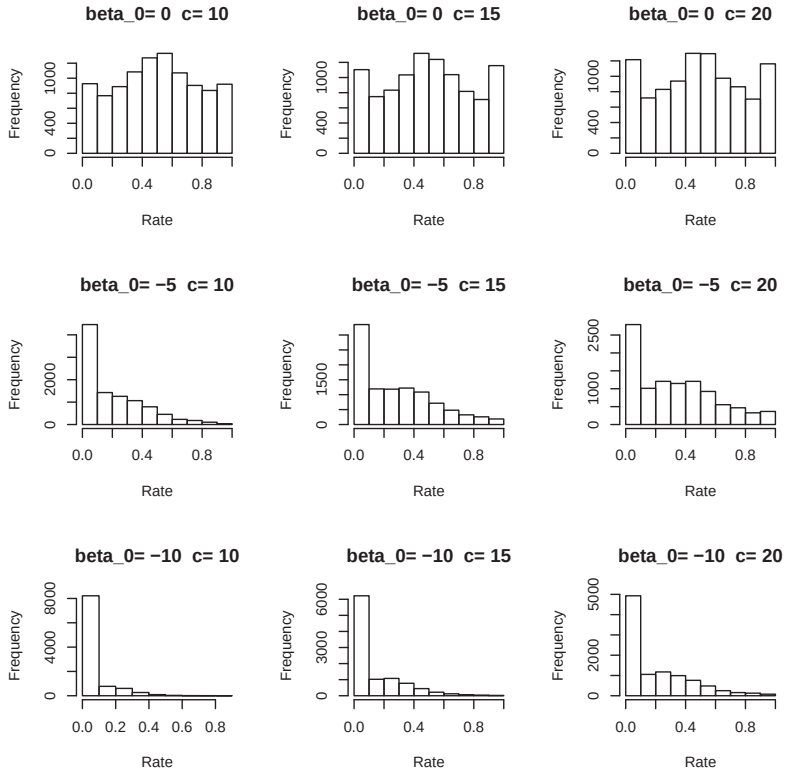

Figure 4: Prior Distribution of Average Rate

$$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \sim N\left(\begin{pmatrix} -10 \\ 0 \\ 0 \end{pmatrix}, \Sigma\right), \tag{5}$$

where $\Sigma = (X^\top X)^{-1}c^2$, $X = \begin{pmatrix} 1, & a_1 & a_1^2 \\ 1, & a_2 & a_2^2 \\ \cdots & & \\ 1, & a_8 & a_8^2 \end{pmatrix}$, and

$$c = \begin{cases} 10 & \text{for any depressive disorder,} \\ & \text{any substance disorder} \\ & \text{any anxiety disorder} \\ & \text{and any psychiatric disorder,} \\ 15 & \text{for major depression} \\ 20 & \text{other disorders.} \end{cases}$$

The choices of the scale c factor here are based on the findings in Figure 4, so is the choice of $\beta_0 = -10$.

4 Bayesian Computation

Because it is impossible to calculate statistics of the posterior distribution analytically, we used a standard Gibbs sampler to sample from the posterior distribution by the following algorithm. Starting from some arbitrary point $(\xi^{(0)}, \beta^{(0)}, \tau^{(0)})$, and given the output from the t^{th} iteration, $(\xi^{(t)}, \beta^{(t)}, \tau^{(t)})$, we performed the following steps, at the $(t+1)^{th}$ iteration,

1. Draw $\beta(t+1)$ from $p(\beta|\xi^{(t)}, \tau^{(t)})$, which is $N(\zeta, \tilde{\Sigma})$, where $\tilde{\Sigma} = (\sigma^{-1} + \frac{1}{[\tau^{(t)}]^2}X^\top X)^{-1}$, and

$$\zeta = \tilde{\Sigma}\Big[\Sigma^{-1}\begin{pmatrix} -10 \\ 0 \\ 0 \end{pmatrix} + \tfrac{1}{[\tau^{(t)}]^2} X^\top \xi^{(t)}\Big];$$

2. Draw $[\tau^{(t+1)}]^2$ from $p(\tau^2|\beta^{(t+1)}, \mu^{(t)})$, which is

$$\frac{[\xi^{(t)}]^\top (1 - X(X^\top X)^{-1} X^\top)\xi^{(t)} + 1}{\chi^2_8};$$

3. Use a Metropolis algorithm to update ξ, that is, we propose $\tilde{\xi}$ from $N(\xi^{(t)}, \lambda \mathrm{I})$, and set

$$\xi^{(t+1)} = \begin{cases} \tilde{\xi}, & \text{with } p = \\ & \min(1, \frac{p(\tilde{\xi}, \beta^{(t+1)}, [\tau^{(t+1)}]^2|y)}{p(\xi^{(t)}, \beta^{(t+1)}, [\tau^{(t+1)}]^2|y)}) \\ \xi^{(t)}, & \text{with } 1-p\ . \end{cases}$$

We ran 5 chains which started from random positions and use the Gelman-Rubin statistic $\hat{R}$ to monitor the convergence of the Markov chains. All $\hat{R}$'s reached 1.1 after 50000 iterations (with the first 25000 samples discarded for burn-in). We also performed various graphical diagnostics to ensure the proper convergence of our MCMC chains.

5 A Simulation Study

As a simple demonstration of the usefulness of the Bayesian method, we performed a simulation study to compare the Bayesian results with the standard design-based estimates. To avoid any potential complication with the choice of the design-based variance estimates, we performed a simple random sampling (SRS), treating the NLAAS sample as the population. For our Bayes method, we used posterior means as the point estimates and central 95% posterior intervals as the interval estimates. We randomly selected subsamples, by SRS, from the total Latino sample, which is of size 2554. We applied both our hierarchical model and SRS estimator to the subsample. After subsampling many times, we compare the bias, efficiency (variance estimate), 95% interval length, and the actual frequentist's coverage of the interval estimators. Our simulation results are based on 500 subsamplings and are shown in Table 1 and 2 for different disorders. We chose these four disorders because they cover the range of typical rates we see in practice, which vary from 1% to 30%. For large subsample sizes, namely the subsample size is 500, our estimators and SRS estimators produce similar biases and interval coverage, but the Bayesian estimates in general have slightly smaller variances and hence shorter intervals, though the improvements are minor.

For sample size 100, occasionally the Bayesian intevals err on being slightly too short, in contrast to the SRS intervals which err on being slightly too long. The only exception is the bulimia for which both methods have significant low coverage about 80%, though the Bayes intervals are only about 65% on average of the length compared to the SRS intervals. When the sample size dropped down to 50, while the performance of both methods deteriorates, the inference from the Bayesian method is still acceptable. In contrast, for bulimia, the inaccurate SRS estimation of variance leads to unacceptably low coverage at about only 51%, yet at the same time the average interval length is still about 40% longer than the one from the Bayesian interval, which has almost 93% coverage. This seemingly paradoxical phenomenon, that is, longer intervals having less coverage, is due to the grave inefficiency in the SRS estimators with small sample sizes. This further demonstrates that the Bayes approach is more reliable than the standard survey estimator for small sample sizes, which is exactly the problem we face with NLAAS.

6 Empirical Findings

6.1 The Pooling Effect

A good way to visualize the pooling effect of the Bayesian approach is to plot the Bayesian estimates against both the raw data and the mean curve from the posterior distribution of μ_i as well as the regression curve. We use the sample mean of $\text{logit}^{-1}(\beta_0 + \beta_1 a_i + \beta_2 a_i^2)$ from the posterior distribution, as the estimated regression curve at age group i, where a_i is the average age in that group. This curve estimates

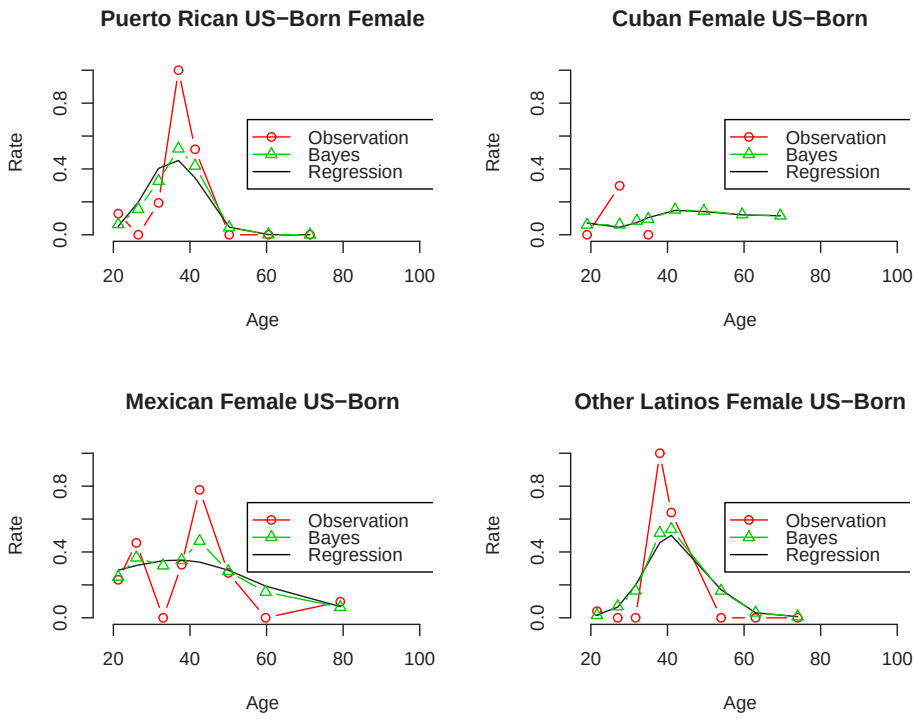

Figure 5: Graphical Diagnostics for Major Depression Prevalence Rates

how the rate varies with the age if our model forces the rate to be exactly as a quadratic function of the age, that is, by forcing $\tau = 0$. Our Bayesian model is much more flexible than this "forced" regression model by allowing the true rate to deviate from the quadratic curve. In other words, the quadratic curve is used to model a *general trend* as how the rate varies with age.

As a result, the Bayesian estimate can be viewed as an appropriately balanced "compromise" between the observed rate, that is, the weighted sample means, and the fitted value from the curve, as illustrated in Figure 5. Due to the small sample sizes and large variation of the sampling weights, the observed prevalence rates fluctuate very much as age changed. The Bayes estimates (the triangle curve), pooled the observed weighted mean (point curve) towards the regression curve and stabilized the estimates. For age groups where no sample is observed (Cuban Female in Figure 5), Bayes estimates also gives estimates, although it is close to the prior mean. Also, the estimated regression curve does seem to capture the trend of how the rate changes with age.

Also from Figure 5, we see that the pooling down of the higher rates are usually more than the pooling up of the lower rates. This is partly because of our binomial-like likelihood approximation, since the sample variance is smaller at lower rates than higher rates (but less than 50%); and partly because of the concavity of link functions at the range of prevalence rate (1% - 30%). The derivative, $G'(\mu)$ at lower rates is always larger than at higher rates (when less than 50%). From the identity $d\mu = \frac{d\xi}{G'(\mu)}$, the same amount of change in ξ will lead to smaller change in

μ when μ is small than when μ is large. This implies the pooling is more significant in the original scale for higher rates.

6.2 Sample Analysis Results

As an illustration of the results from our analysis, Table 3 and 4 presents traditional and Bayesian lifetime prevalence estimates for a number of psychiatric disorders, adjusted for age and gender. The results in Table 3 and 4 shows that whenever subpopulation sizes are large (e.g., for Mexican), the traditional and Bayesian methods provide essentially identical results. For small subgroups, the Bayesian prevalence estimates are likely to be more reliable, as the vast literatures on Bayesian small-area estimates demonstrated (e.g. Ghosh et.al. 1998, Long 1999, Nandram & Choi 2002).

Our results indicate that major depressive episode disorder, social phobia, and alcohol abuse disorder are the most prevalent lifetime psychiatric disorders for all Latinos in the U.S. Overall, Mexicans, Cubans, and Other Latinos did not differ in lifetime rates of specific psychiatric disorders, except Cubans who present lower prevalence estimates of lifetime substance disorder than the other groups. Puerto Ricans had significantly higher lifetime prevalence estimates than the other groups for post traumatic stress disorder, any anxiety disorder, and any psychiatric disorder but not for any depressive disorder. Further studies, of course, are very much needed to check how sensitive are these results to our Bayesian modeling assumptions.

References

[1] Alegria, M., Takeuchi, D., Canino, G., Duan, N., Shrout, P., Meng, X.L., Vega, W., Zane, N., Vila, D., Woo, M., Vera, M., Guarnaccia, P., Aguilar-Gaxiola, S., Sue, S., Escobar, J., Lin, K-M, Gong, F. (2004). Considering Context, Place, and Culture: The National Latino and Asian American. *International Journal of Methods in Psychiatric Research*, **13**, 208-220

[2] Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. (2003). *Bayesian Data Analysis*, CRC Press.

[3] Gelman, A., Meng, X.L. (2004). *Applied Bayesian Modeling and Causal Inference from Incomplete Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*, Wiley, John & Sons, Incorporated.

[4] Ghosh, M., Natarajan, K., Stroud, T.W.F., Carlin, B.P. (1998). Generalized Linear Models for Small-Area Estimation, *Journal of the American Statistical Association*, Vol. 93, No. 441, pp. 273-282.

[5] Longford N.T. (1999). Multivariate Shrinkage Estimation of Small Area Means and Proportions, *Journal of the Royal Statistical Society.* Series A (Statistics in Society) Vol. 162, No. 2 , pp. 227-245.

[6] Nandram B., Choi J.W. (2002), Hierarchical Bayesian Nonresponse Models for Binary Data From Small Areas With Uncertainty About Ignorability, *Journal of the American Statistical Association*, Volume 97, No. 458, pp. 381-388(8).

Sample Size	Any Disorder: 30.70%(truth)						Major Depression: 15.66%(truth)					
	500		100		50		500		100		50	
	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS
Mean	30.51	30.55	30.01	30.20	30.14	30.32	15.54	15.60	14.98	15.26	14.95	15.27
MSE	0.04	0.04	0.21	0.19	0.43	0.39	0.02	0.02	0.13	0.12	0.28	0.26
VAR	0.04	0.03	0.20	0.19	0.43	0.39	0.02	0.02	0.12	0.12	0.27	0.26
Coverage	96.40	97.00	94.40	96.60	93.00	97.20	96.00	96.00	93.80	97.80	90.40	97.40
Interval Length	8.00	8.09	17.39	18.12	23.99	25.86	6.28	6.39	13.45	14.43	18.32	20.97

All the numbers are in 10^{-2} scale.

Table 1: Comparing point and interval estimates for Any Disorder and Major Depression

Sample Size	Social Phobia: 7.64%(truth)						Bulimia: 1.68%(truth)					
	500		100		50		500		100		50	
	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS	Bayes	SRS
Mean	7.52	7.63	7.04	7.51	6.98	7.63	1.63	1.70	1.51	1.62	1.57	1.56
MSE	0.01	0.01	0.07	0.06	0.13	0.14	<0.01	<0.01	0.01	0.01	0.02	0.03
VAR	0.01	0.01	0.06	0.06	0.12	0.14	<0.01	<0.01	0.01	0.01	0.02	0.03
Coverage	97.20	97.00	93.20	98.00	90.00	95.80	95.00	97.00	82.60	80.00	92.80	51.00
Interval Length	4.56	4.71	9.40	11.00	12.58	16.64	2.14	2.41	4.06	6.11	5.36	7.50

All the numbers are in 10^{-2} scale.

Table 2: Comparing point and interval estimates for Social Phobia and Bulimia

Disorder		Puerto Rican		Cuban		Mexican	
Major Depressive Episode	a	18.5	(14.7-22.9)	16.1	(13.4-19.2)	13.0	(11.6-14.5)
	b	15.9	(12.8 -19.1)	17.5	(13.3 -22.1)	12.7	(10.4 -15)
Dysthymia	a	4.7	(3.4-6.4)	3.1	(1.7-5.7)	1.8	(1-3.1)
	b	3.7	(2.2 -5.2)	3.8	(2.2 -5.5)	1.7	(0.9 -2.6)
Any Depressive Disorder	a	18.7	(14.8-23.4)	16.5	(13.7-19.8)	13.0	(11.6-14.5)
	b	16.8	(13.7 -20)	18.6	(14.3 -23)	12.9	(10.6 -15.1)
Agoraphobia without panic	a	3.3	(1.5-7.3)	2.5	(1.1-5.7)	3.0	(1.9-4.8)
	b	3.3	(1.8 -4.9)	2.4	(1 -3.9)	2.9	(1.8 -4)
Panic Disorder	a	4.9	(2.9-8.1)	1.8	(1.1-2.7)	2.7	(1.7-4.1)
	b	4.9	(3 -6.8)	2.4	(1.2 -3.6)	2.6	(1.5 -3.8)
GAD	a	7.1	(4.5-11.1)	6.7	(4.7-9.5)	3.6	(2.4-5.2)
	b	6.0	(4 -8.1)	5.8	(3.9 -8)	4.1	(2.9 -5.3)
Social Phobia	a	10.4	(6.8-15.8)	8.1	(5.4-12)	7.4	(5.6-9.9)
	b	9.1	(6.7 -11.5)	8.4	(5.5 -11.3)	6.9	(5.3 -8.6)
PTSD	a	8.0	(5.2-12.1)	5.4	(3-9.7)	4.4	(3.2-5.9)
	b	7.1	(4.8 -9.4)	4.7	(2.8 -6.8)	3.9	(2.6 -5.1)
Any Anxiety Disorder	a	21.6	(16.1-28.4)	16.1	(13.6-19)	15.0	(13-17.3)
	b	20.8	(17.4 -24.5)	16.8	(13.2 -20.5)	15.1	(12.8 -17.5)
Alcohol Dependence	a	5.9	(3.7-9.3)	2.7	(1.4-5.2)	4.7	(3.5-6.2)
	b	5.0	(3 -7)	2.8	(1.4 -4.5)	5.0	(3.5 -6.4)
Alcohol Abuse	a	8.7	(5.5-13.4)	3.4	(2-5.8)	6.1	(4.1-9)
	b	7.0	(4.7 -9.6)	4.0	(2 -6.2)	5.6	(4.1 -7.3)
Alcohol Dependence/Abuse	a	14.6	(10.1-20.7)	6.2	(3.9-9.5)	10.8	(8.4-13.8)
	b	11.8	(8.9 -14.7)	6.0	(3.9 -8.4)	10.6	(8.7 -12.7)
Drug Dependence	a	4.0	(2.2-7.3)	2.6	(1.3-5.2)	2.1	(1.3-3.6)
	b	3.7	(2.1 -5.6)	2.2	(0.9 -3.7)	3.0	(1.8 -4.2)
Drug Abuse	a	4.6	(2.9-7.1)	0.8	(0.2-3.3)	3.7	(2.6-5.3)
	b	4.2	(2.4 -6.1)	1.5	(0.2 -2.9)	3.7	(2.5 -5)
Drug Dependence/Abuse	a	8.6	(6-12.1)	3.4	(2.1-5.5)	5.9	(4.4-7.8)
	b	7.6	(5.4 -10)	2.9	(1.4 -4.6)	6.2	(4.6 -7.9)
Any Substance Disorder	a	15.4	(10.9-21.4)	6.7	(4.7-9.5)	11.3	(8.4-15.1)
	b	14.4	(11.2 -17.4)	8.1	(5.4 -11.1)	12.0	(9.8 -14.3)
Bulimia	a	2.4	(1.2-4.4)	2.2	(0.9-5.3)	1.3	(0.6-2.5)
	b	2.3	(1 -3.8)	2.2	(1 -3.7)	1.7	(0.7 -2.7)
Anorexia	a	0.0	$(. - .)^{\dagger\dagger}$	0.1	(0-0.6)	0.0	$(. - .)^{\dagger\dagger}$
	b	0.7	(0.1 -1.7)	0.8	(0.1 -1.8)	0.5	(0 -1)
Any Disorder	a	39.2	(33.3-45.5)	31.1	(27.6-35)	29.2	(25.8-32.9)
	b	36.0	(32 -40)	31.4	(26.7 -36.4)	29.5	(26.6 -32.6)

Table 3: Lifetime Prevalence Rate for Latinos

a – Design-based estimates
b – Bayes estimates
† All the numbers are percentages.
†† Design-based method cannot provide estimates because the sample size is zero.

Disorders		Other Latinos		Total Latinos	
Major Depressive Episode	a	13.4	(11.4-15.8)	13.8	(12.6-15.1)
	b	13.9	(11.1 -16.6)	13.6	(12.1 -15.1)
Dysthymia	a	2.2	(1.1-4.5)	2.3	(1.7-3)
	b	2.1	(0.9 -3.4)	2.1	(1.5 -2.7)
Any Depressive Disorder	a	14.1	(11.7-16.9)	14.0	(12.8-15.4)
	b	14.7	(12 -17.5)	14.1	(12.5 -15.6)
Agoraphobia without panic	a	1.5	(0.9-2.7)	2.6	(1.9-3.6)
	b	2.0	(0.9 -3.2)	2.7	(1.9 -3.4)
Panic Disorder	a	2.4	(1.4-3.8)	2.8	(2.2-3.7)
	b	2.7	(1.4 -4.1)	2.9	(2.1 -3.7)
GAD	a	4.3	(2.8-6.6)	4.2	(3.4-5.2)
	b	3.5	(2.1 -4.9)	4.2	(3.4 -5)
Social Phobia	a	6.9	(4.6-10.1)	7.4	(6.1-9)
	b	6.9	(4.9 -9)	7.2	(6.1 -8.4)
PTSD	a	3.3	(2.2-4.9)	4.4	(3.6-5.4)
	b	3.6	(2.1 -5.2)	4.2	(3.3 -5.1)
Any Anxiety Disorder	a	14.0	(11.1-17.4)	15.3	(13.6-17.1)
	b	14.5	(11.7 -17.4)	15.6	(13.9 -17.2)
Alcohol Dependence	a	3.6	(2.3-5.6)	4.4	(3.4-5.6)
	b	3.7	(2.2 -5.5)	4.5	(3.5 -5.5)
Alcohol Abuse	a	6.2	(3.9-9.8)	6.3	(4.7-8.4)
	b	5.8	(4 -7.8)	5.8	(4.7 -6.9)
Alcohol Dependence/Abuse	a	9.8	(6.8-14)	10.7	(8.5-13.2)
	b	9.6	(7.2 -12)	10.2	(8.9 -11.6)
Drug Dependence	a	1.2	(0.5-2.7)	2.0	(1.5-2.8)
	b	1.5	(0.5 -2.6)	2.6	(1.8 -3.4)
Drug Abuse	a	4.5	(2.8-7.2)	4.0	(3-5.3)
	b	4.1	(2.5 -5.9)	3.8	(2.9 -4.7)
Drug Dependence/Abuse	a	5.7	(3.9-8.3)	6.0	(4.7-7.6)
	b	5.2	(3.3 -7.2)	5.9	(4.8 -7)
Any Substance Disorder	a	10.3	(7.2-14.6)	11.2	(8.8-14.2)
	b	11.0	(8.5 -13.6)	11.8	(10.3 -13.3)
Bulimia	a	2.1	(0.9-4.6)	1.7	(1.1-2.5)
	b	2.3	(1.2 -3.5)	1.9	(1.2 -2.6)
Anorexia	a	0.2	(0-1.4)	0.1	(0-0.4)
	b	0.7	(0.1 -1.4)	0.6	(0.2 -0.9)
Any Disorder	a	28.0	(23.2-33.3)	29.8	(26.9-32.9)
	b	28.7	(24.9 -32.4)	30.0	(27.9 -32.1)

Table 4: Confidence Interval Comparison for Major Depression (ctd)

a – Design-based estimates

b – Bayes estimates

† All the numbers are percentages.