9-1-2013

Modeling Area-Level Health Rankings

Charles J. Courtemanche
Georgia State University, ccourtemanche@gsu.edu

Samir Soneji
Dartmouth College, samir.soneji@dartmouth.edu

Rusty Tchernis
Georgia State University, rtchernis@gsu.edu

W. J. Usery Workplace Research Group Paper Series

Working Paper 2013-9-1
September 2013

Modeling Area-Level Health Rankings

Charles Courtemanche
Georgia State University

Samir Soneji
Dartmouth College

Rusty Tchernis
Georgia State University

Andrew Young School
of Policy Studies

Modeling Area-Level Health Rankings*

Charles Courtemanche† Samir Soneji‡ Rusty Tchernis§

September 11, 2013

Abstract

We propose a Bayesian factor analysis model to rank the health of localities. Mortality and morbidity variables empirically contribute to the resulting rank, and population and spatial correlation are incorporated into a measure of uncertainty. We use county-level data from Texas and Wisconsin to compare our approach to conventional rankings that assign deterministic factor weights and ignore uncertainty. Greater discrepancies in rankings emerge for Texas than Wisconsin since the differences between the empirically-derived and deterministic weights are more substantial. Uncertainty is evident in both states but becomes especially large in Texas after incorporating noise from imputing its considerable missing data.

Keywords: County, rank, health, factor analysis, Bayesian
JEL Codes: I14, C11

*We thank Barry Hirsch, Dann Millimet, Patrick Remington, Chris Ruhm, Harold Sox, and audiences at the Georgia Health Econ Research Day, Eastern Economics Conference, Usery Workplace Research Group, and Universidad del Pacifico for helpful feedback.

†Department of Economics, Andrew Young School of Policy Studies, Georgia State University and National Bureau of Economic Research. P.O. Box 3992, Atlanta, GA 30302. Phone: 404-413-0082. Fax: 404-413-0145. Email: ccourtemanche@gsu.edu.

‡Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth College, and the Norris Cotton Cancer Center. Norris Cotton Cancer Center, One Medical Center Drive, Lebanon NH, 03756. Phone: 603-653-3908. Fax: 603-653-0820. Email: samir.soneji@dartmouth.edu.

§Corresponding author. Department of Economics, Andrew Young School of Policy Studies, Georgia State University, National Bureau of Economic Research, and Institute for the Study of Labor (IZA). P.O. Box 3992, Atlanta, GA 30302. Phone: 404-413-0154. Fax: 404-413-0145. Email: rtchernis@gsu.edu.

1 Introduction

Researchers consider a broad range of health, social, economic, and environmental determinants when assessing population health and identifying areas of greatest need (Institute of Medicine, 2011; Kindig et al., 2008). Population health assessments are often presented to policymakers and communities as ranks, given their ubiquity and ease of interpretation (Booske et al., 2010; Kanarek et al., 2011; Erwin et al., 2011). Local area rankings can motivate stakeholders in lagging communities to design and promote local public health interventions. Such rankings can also assist policymakers with resource allocation decisions, which can be especially critical in times of declining local, state, and federal revenues. Understanding one's community's relative health status can help local officials assess the importance of public health initiatives relative to other priorities. At the state and federal levels, knowledge of the least healthy communities can assist with funding decisions regarding local interventions and demonstration projects (Remington and Booske, 2011).

Despite the potential usefulness of local area-level health rankings, two important difficulties arise in credibly assessing them. The first is the lack of a single comprehensive observable measure of health. This necessitates the use of some weighting procedure that combines available health-related variables into a single summary measure. The second is the need to account for uncertainty arising from sources such as sampling error and missing data. Observable attributes of health are often not available at the population level and must instead be estimated using samples that can become small as the geographic area narrows. The amount of uncertainty can therefore be considerable at local levels such as the county. Moreover, data are often missing entirely for all or some components of health in certain localities, so inherently noisy procedures for imputing missing data are necessary in order to produce comprehensive rankings.

The single most prominent local area-level health rankings are the County Health Rankings (CHRs), produced by the University of Wisconsin Population Health Institute (UWPHI) and begun in 2010. The CHRs address the difficulties inherent in local area-level health rankings by making strong and sometimes unrealistic assumptions. These rankings acknowledge the multi-faceted nature of health, but fix subjectively-assessed deterministic weights of each component in contributing to the overall health measure rather than allowing the weights to be empirically derived. The CHRs also do not account at all for uncertainty, despite their use of sample data and an imputation process for missing data. It is therefore not possible to assess whether reported differences in counties' rankings are statistically meaningful.[1]

This paper develops an alternative method for ranking county health that addresses the issues of factor weighting and uncertainty through the use of a Bayesian hierarchal model for factor analysis. We treat health as a latent variable that depends on observable factors related to mortality and morbidity. The model empirically derives factor weights and measures uncertainty, which is inversely related to population size and accounts for spatial covariance. Incorporating information from neighboring areas improves precision, which can be especially helpful in areas with small populations that would otherwise have high levels of uncertainty. These features of the model follow Hogan and Tchernis (2004), who ranked census tracts' levels of material deprivation in Rhode Island. We build on the Hogan and Tchernis (2004) framework by also including an iterative procedure to impute missing data, which is important in our context because a full array of health information is often not available for small localities. We incorporate uncertainty from the imputation process into the overall measure of uncertainty.

We apply our method to the UWPHI's county-level data for Texas (TX) and Wisconsin (WI). Though our model could be utilized for any state, we choose these two because TX is the state with the most counties while WI's rankings served as the UWPHI's template for the CHRs. The results demonstrate the importance of addressing both of the key issues of factor weighting and uncertainty. First, the data-derived factor weights are more similar

[1]Subsequent to the initial version of our paper, Athens et al. (2013) provide a preliminary attempt to account for uncertainty in the rankings coming from sampling error. They do not, however, address the issues of determnistic factor weights or missing data.

to the UWPHI's deterministic weights for WI than TX, making the difference between our rankings and the CHRs much greater for TX. This is also the case for our identification of least healthy counties and suburban-urban health disparities, two issues of particular interest to policymakers. Second, allowing for uncertainty reveals a number of instances where apparent geographic disparities are misleading. When the measure of uncertainty depends only on population and spatial covariance, the mean rank of counties should be at least 25 ranks apart in TX (21 ranks in WI) for them to differ with 90% confidence. When the uncertainty from imputing missing data is also considered, the overall amount of uncertainty rises drastically in TX, where many counties have at least some missing information. Indeed, it is not possible to reach clear conclusions for most of the counties in TX because of the prevalence of missing data, although some of the least healthy counties can still be identified.

While we focus on county health rankings, our methods could be useful in a variety of other contexts since they can summarize latent variables related to observed manifest variables on any level of aggregation. Factor correlation need not be spatial; correlation could also arise across observational units with, for instance, similar demographic, economic, or historic characteristics. One possible application of our method could be to rank hospitals' quality of care, where quality of care is a latent variable with observable manifestations such as in-hospital mortality rates that are correlated across hospitals of similar sizes. As another example, overall national economic performance could be modeled as a latent variable that manifests itself through measureable statistics such as consumption and investment, which may be correlated based on population demographics.

2 Methods

2.1 The Model

We incorporate a latent variable framework in which mortality and morbidity variables are observed manifestations of the latent construct–health (Figure 1, left panel). The underlying assumption of this framework is that health is not directly observed, but rather is manifested through a number of measurable variables. Latent variable frameworks have been widely adopted to assess quality of life (McAuley et al., 2006), investigate geographic patterns of disease and their relationship to behavioral and social risk factors (Best and Hansell, 2009), and measure health inequality across population subgroups (Murray et al., 1999; van Doorslaer and Jones, 2003). We consider the same mortality and morbidity variables as the CHRs to facilitate comparison. In contrast to our latent variable framework, the CHRs utilize a deterministic framework in which the construct health outcomes explicitly consist of a weighted combination of mortality and morbidity variables (Figure 1, right panel).

Before we present our methodology, it is useful to understand how the UWPHI calculates the existing CHRs. UWPHI calculates an overall health outcomes score based on standardized mortality and morbidity variables and their corresponding deterministic weights (Booske et al., 2010). UWPHI first transforms the value of each mortality and morbidity variable into its corresponding z-score based on the distribution of values within the state. Next, the z-scores are multiplied by their corresponding deterministic weight. Finally, UWPHI sums over the weighted z-scores to create a final score for each county. This score is the basis for the county health rankings.

We utilize a factor analysis model with spatially correlated factors to estimate the distribution of ranks for counties within a state. The traditional factor analysis model (Bartholomew and Knott, 1999) explains the variability in observed variables Y_{ij} for county

i in the following way:

$$Y_{ij} = \mu_j + \lambda_j \delta_i + e_{ij},$$

where μ_j is the average of variable j across counties, the factor $\delta_i \sim N(0,1)$ represents latent health level for county $i = 1, \ldots, n$, λ_j is the factor loading for variable $j = 1, \ldots, J$ which represents the covariance between the latent health and the observed variables, and $e_{ij} \sim N(0, \sigma_j^2)$ are the idiosyncratic error terms. The model assumes that all of the observed variables are influenced by the underlying latent health factor represented by δ_i. The model is identified by decomposing the covariance matrix of the variables within the county into the correlation represented by the factors since the error terms, e_{ij}, are uncorrelated, $\sigma_{jk} = 0 \; \forall \; j \neq k$.

Stacking over the variables within the county we can rewrite the model in vector notation as follows:

$$Y_i = \mu + \lambda \delta_i + e_i,$$

where λ is a vector of stacked λ_j and $Var(e_i) = \Sigma = diag\{\sigma_j^2\}$. Finally, stacking over the counties we can write the model in hierarchical form:

$$\begin{aligned} Y|\delta &\sim N(\mu + \Lambda\delta, I_n \otimes \Sigma) \\ \delta &\sim N(0, I_n) \end{aligned} \tag{1}$$

where $\Lambda = I_n \otimes \lambda$.

The next step is to introduce the population sizes in the variance of both the error terms and the factors. The assumption is that error terms and the factors in more populous counties have smaller variance, which is reasonable since the sample sizes used to compute the underlying factors are smaller in less populous counties as long as sampling is random (i.e. small counties are not oversampled).[2] We define $M = diag\{m_i\}$, where m_i is population

[2] One could also incorporate sample sizes rather than population sizes into the variance; we use population for simplicity because the factors come from different data sources and therefore have different sample sizes.

of county i and specify the new model as follows:

$$\begin{aligned} Y|\delta &\sim N(\mu + \Lambda\delta, M^{-1} \otimes \Sigma) \\ \delta &\sim N(0, M^{-1}) \end{aligned} \tag{2}$$

In this specification the variances are inversely proportional to the county population sizes. In sum, our model accounts for stochastic uncertainty as well as uncertainty from sampling error and the fact that the factor loadings are estimated rather than known with certainty.

2.2 Spatial Correlation

The last step in the development of our model introduces spatial dependence of the factors. So far our model assumes independence across geographic areas. However, it is reasonable to consider spatial spillovers that can affect area health measures. Thus we introduce spatial dependence of the factors via the spatial correlation matrix Ψ. The model can be rewritten in the following way.

$$\begin{aligned} Y|\delta &\sim N(\mu + \Lambda\delta, M^{-1} \otimes \Sigma) \\ \delta &\sim N(0, M^{-1/2}\Psi M^{-1/2}) \end{aligned} \tag{3}$$

We work with the conditional autoregressive (CAR) specification (Besag, 1974; Sun et al., 1999) which produces a tractable relationship between the conditional and the marginal specifications. Hogan and Tchernis (2004) show that this specification performs well relative to several alternatives. We start from specifying the conditional relationship between the factor for county i and other counties in the neighborhood of i, R_i, and define the neighborhood

As long as population size and sample size are proportional, this distinction does not matter.

as the set of counties adjacent to i

$$\delta_i|\{\delta_j : j \in R_i\} \sim N\Big(\sum_{j\in R_i}\beta_{ij}\delta_j, v/\alpha_i\Big).$$

As discussed in Hogan and Tchernis (2004), for identification as part of the factor analysis model we can only identify one parameter and thus restrict $\beta_{ij} = \omega$, and $v/\alpha_i = 1$. This specification models the conditional mean of the distribution of the factors as a weighted average of the factors from neighboring counties, with higher values of ω representing stronger spatial dependence. This conditional specification results in the marginal distribution of $\delta \sim N(0, (I - \omega R)^{-1})$, where $R_{ij} = 1$ if a county j is adjacent to county i and $R_{ii} = 0$. Thus $\Psi = (I - \omega R)^{-1}$ is a full matrix inducing the correlation between variables between counties. This specification induces a restriction on the support of ω to be between the reciprocals of the smallest and the largest eigenvalues of R.

2.3 Estimation

The model is estimated using MCMC methods (Chib and Greenberg, 1996). We use Gibbs Sampler (Gelfand and Smith, 1990) with one Metropolis-Hastings (Chib and Greenberg, 1995) step to obtain draws from ω. While the exact conditional distributions are summarized in Hogan and Tchernis (2004), we will discuss here how we produce the samples from the posterior distribution of health ranks. At each iteration of the sampler we rank the posterior means of factors, resulting in one sample from the posterior distribution of ranks.

2.4 Missing Data

In our application some of the values of the manifest variables Y_{ij} are unobserved and thus need to be imputed. UWPHI dealt with this issue simply by replacing missing covariates with their corresponding state-level means. This approach is potentially problematic for two reasons. First, it ignores the uncertainty inherent in the imputation process. Second, imputing missing data with state averages may lead to biased rankings if counties with missing data are systematically more or less healthy than the average county. For example, the rank of a county missing data on all but one factor will approach the state average. In practice, such a county may be rural and may therefore may be less healthy than average (e.g. Weden et al., 2011).

Accordingly, we consider two different approaches to missing data. Our baseline model simply replaces the missing factors(s) with ordinary least squares (OLS) predictions based on the other (non-missing) factors. This solves the problem of counties with missing data automatically being drawn towards the middle, but it does not address the issue of uncertainty.[3] Our second imputation approach is therefore to sample from the distribution of missing values conditional on the parameters of the model using (3) at each iteration of the sampler.[4] This incorporates the uncertainty of predicting missing values as part of the estimation, similarly to multiple imputations procedure.[5]

3 Data

We implement the model using UWPHI data applied to the year 2011 county health outcome rankings in TX and WI. We selected WI and TX as our two example states because the former served as the focus of the precursor to the CHRs, the Wisconsin County Health Rankings (Peppard et al., 2003 and 2008). The latter is a populous state, includes several of the largest metropolitan areas in the US, and has the largest number of counties – our unit of analysis. We used Matlab Version R2009a for all statistical analyses.

[3] We also tried simply replicating the UWPHI's state averages approach, and the results (available upon request) were very similar to those obtained using the OLS imputation. This suggests the biggest problem with the state averages imputation in this context is its neglect of uncertainty, not its introduction of systematic bias.

[4] See Hogan and Tchernis (2004) for details about the sampling algorithm.

[5] Note that it is possible that even some of the non-missing data could be incorrectly reported; our model does not account for uncertainty from such measurement error.

Mortality and morbidity data were downloaded on June 1, 2012, from the website maintained by UWPHI: www.countyhealthrankings.org. The mortality variable used in the CHRs is the years of potential life lost before age 75 years ("premature death"), estimated by UWPHI using 2005-2007 life table data from the National Center for Health Statistics (NCHS). The morbidity variables used are: [1] the percent of adults reporting fair or poor health ("self-reported health"), [2] the mean number of physically unhealthy days per month for adults ("physical unhealthy days"), [3] the mean number of mentally unhealthy days per month for adults ("mental unhealthy days"), and [4] the percent of live births with birthweight < 2500 grams ("low birthweight"). The first three morbidity variables were estimated by UWPHI using 2003-2009 data from the Behavioral Risk Factor Surveillance System and the fourth morbidity variable was estimated by UWPHI using 2001-2007 birth certificate data from NCHS.

UWPHI did not rank 31 out of 254 counties in TX because they were missing at least four of the five variables. 116 of the remaining 223 counties had missing data on at least one variable, and UWPHI applied their aforementioned imputation procedure to these counties. UWPHI ranked all 72 counties in WI, only 2 of which had missing data on any variables. To facilitate comparability, we focus only on the counties ranked by UWPHI. Our sample therefore consists of 223 TX counties and 72 WI counties.

4 Results

4.1 Factor Weights

We estimate the model using data for TX and WI separately. Appendix Table A1 reports the means, standard errors, and 95% probability intervals (PIs) of the posterior distributions of all of the model's parameters. We focus our discussion, however, on the results of most interest, beginning with the estimated factor weights.

Table 1 compares the deterministic UWPHI CHR weights to our normalized square correlations for the mortality and morbidity variables. Normalized square correlation represents the proportion of the variance in the variable that is explained by the factors and, therefore, is comparable to the weights used by UWPHI.[6] Normalized square correlations differ between TX and WI. Additionally, our squared correlations differ from the CHR weights, which are applied uniformly across all states. For example, we estimate the squared correlation of the mean number of physically unhealthy days per month to be 0.41 for TX (95% CI, 0.34 to 0.48) and 0.21 for WI (95% CI, 0.11 to 0.31), whereas UWPHI sets the weight of this variable to 0.10 for all states. The fact that the difference between our squared correlations and the UWPHI weights is greater for TX than WI suggests that our rankings are likely to be less similar to the CHRs for TX than WI. This difference also illustrates the broader point that assuming the factor weights to be constant across states may be problematic. For instance, physical health might be more closely tied to overall health than mental health in one state, perhaps due to differences in demographic, economic, or cultural charactersitics, while the reverse may be true elsewhere.[7]

[Table 1 About Here]

4.2 Mean Rankings

For each county, we compute the posterior distribution of its health outcomes rank, including its mean and 95% probability interval (PI). Appendix Tables A2 and A3 show our full county rankings for TX and WI, along with the corresponding CHRs. In Figure 2 we summarize this information in a way that illustrates the differences between our rankings and the CHRs as

[6] In addition, we estimated regression coefficients by regressing the mortality and morbidity measures on the posterior means of factors. These regression coefficients are directly comparable to the UWPHI weights. The information in these regression coefficients was nearly identical to squared correlations and they are available upon request from the authors.

[7] Of course, similar logic suggests that in some contexts weights might differ within states, e.g. at the Metropolitan Statistical Area level. In our view, it makes sense to keep the weights constant within the geographic level for which the rankings are reported, which in this case is the state.

well as the extent of uncertainty in the rankings. Specifically, we plot the middle 95% of the posterior distribution of ranks (horizontal line) and the mean of the posterior distribution (solid circle) relative to each county's UWPHI rank. These initial estimations use the OLS imputations for missing data.

[Figure 2 About Here]

We begin by discussing the differences between our (mean) rankings and the CHRs. Greater distances between solid circles and the 45 degree line indicate greater differences between our rankings and the CHRs. A solid circle to the right of the 45 degree line indicates our method ranks the county worse than the CHRs, and vice versa. Comparing the panels for TX and WI, we observe that there are more disagreements in TX than in WI - the intervals in the WI panel are much closer to the 45 degree line. The correlations between the UWPHI ranks and our ranks are 0.65 for TX (95% CI, 0.60 to 0.71) and 0.89 for WI (95% CI, 0.81 to 0.94). This is exactly as expected given the fact that our factor weights differ more substantially from the CHR weights for TX than WI, as discussed in Section 4.1.

4.3 Uncertainty

We next turn to the results regarding uncertainty in the rankings. The horizontal lines from Figure 2 suggest considerable uncertainty in both states. The 95% PIs range from 0 to 69 ranks wide in TX, with a mean width of 14.5 ranks and median width of 11 ranks. In WI, the 95% PIs range from 1 to 38 ranks wide, while the mean and median widths are 12 and 12 ranks.

We can use the measures of uncertainty to ask how far apart the ranks of two counties should be to give a researcher reasonable confidence that they are different (e.g., 90% or 95% confident). To answer this question we calculate the percentage overlap in the posterior distribution of ranks between two counties that are k units apart in their mean rank. Consider

Harris County, TX with mean rank equal to 66. The mean rank of Blanco County, TX equals k = 5 ranks higher at 71; 2.2% of the posterior distribution of Harris County overlaps with the posterior distribution of Blanco County (Figure 3, left panel). In TX, there are 216 unique pairs of counties with mean rank k = 5 units apart and we calculate the percentage overlap for each pair. The median percentage overlap among the 216 pairs of counties is 12.8%.

[Figure 3 About Here]

As the difference between mean ranks, k, increases, the median percentage overlap decreases. For example, as k increases from 1 to 5 to 10, the median overlap in the posterior distributions equals 40.8%, 12.8%, and 1.3%, respectively (Figure 3, middle panel). When k exceeds 25, 90% of county pairs have 1% or less overlap. Similarly, when k exceeds 30, 95% of county pairs have 1% or less overlap. In WI, as k increases from 1 to 5 to 10, the median overlap in the posterior distributions equals 41.7%, 13.5%, and 0.5%, respectively (Figure 3, right panel). When k exceeds 21, 90% of county pairs have 1% or less overlap and when k exceeds 23, 95% of county pairs have 1% or less overlap. Thus, to be reasonably confident, say 90%, that two counties are different with respect to their health outcomes ranking, the distance between the mean of their health rank distribution should be approximately 25 counties apart in TX and 21 counties apart in WI.

4.4 Least Healthy Counties

A key purpose of the CHRs is identifying the least healthy counties to more effectively mobilize action to improve the health of these populations (Friedman et al., 2003; Kindig and Stoddart, 2003; McDowell et al., 2004). We therefore next examine how closely the least healthy counties identified by our rankings correlate to those from the CHRs. The vertical and horizontal lines in Figure 2 represent the 80th percentile (178th out of 223 ranked counties in TX and 62nd out of 72 in WI) separating the least healthy quintile of counties. According to our model, the least healthy counties will be those counties in which the mean of their rank distribution lies to the right of the vertical grey line. According to the CHRs, the least healthy counties will be those with ranks above the horizontal grey line. Again, the differences are much more pronounced for TX than WI. In TX, 26 counties are classified as least healthy by both models, 19 are classified as least healthy by our model and not UWPHI, and 19 are classified as least healthy by UWPHI and not our model. In WI, 13 counties are classified as least healthy by both models, 3 are classified as least healthy by our model and not UWPHI, and 2 are classified as least healthy by UWPHI and not our model. Among the counties classified as least healthy either by our rankings or the CHRs, the correlation between the CHRs and our mean ranks was -0.06 for TX (95% CI, -0.17 to 0.14) and 0.68 for WI (95% CI, 0.34 to 0.85).

It is also useful to examine the degree of uncertainty in our identification of the least healthy counties, which we can do by presenting the probability of each county being in the least healthy quintile. Figure 4 presents county-level maps of TX and WI of this probability. According to the CHRs, counties either are or are not in this quintile. In contrast, each county in our model has a probability of being in the least healthy quintile, which is reflected in increasing shades of gray. In TX, we observe a large concentration of unhealthy counties in East Texas. The probability of being in the least healthy quintile is 1.00 for Austin County, 0.63 for Colorado County, and 0.59 for Polk County. In other words, the entire posterior distribution of rankings for Austin County and 59% of the posterior distribution of rankings for Polk County lie to the right of 80th percentile. In WI, the probability is high for several of the northernmost counties (e.g., 0.96 for Douglas County) and several counties in the Milwaukee area (e.g., 1.00 for Milwaukee County and 0.54 for Kenosha County).

[Figure 4 About Here]

4.5 Metropolitan Statistical Areas

Another useful feature of county health rankings is to identify broader patterns in disparities, such as whether the urban core of metropolitan areas typically has worse health than outlying suburban areas. We next illustrate the potential of our method to illuminate such disparities by examining a metropolitan area in both of our sample states: the Dallas-Ft. Worth-Arlington, TX Metropolitan Statistical Area (MSA) and the Milwaukee-Racine-Waukesha, WI Combined Statistical Area (CSA) (Office of Management and Budget, 2009). Measuring differences in health rankings within metropolitan areas addresses disparities defined by both demographic characteristics and geographic location (Phillips and McLeroy, 2004) and the physical, social, and economic environments that shape these differences. Weden et al. (2011) investigated evidence of rural health disadvantage and found differences in mortality and morbidity among rural, suburban, and urban areas. For simplicity we continue to use the factors weights estimated earlier for TX and WI, though in principle these could be re-estimated using only the subset of counties within each MSA.

Figure 5 plots the posterior distribution of health outcome ranks for each county of the Dallas MSA, showing the median of each posterior distribution as a vertical line, the mean as a solid circle, and the corresponding UWPHI rank as an open triangle. We observe a suburban-urban gradient in health rankings. Counties north of Dallas County (e.g., Denton, Collin, and Rockwell) rank healthier than the central urban county and they rank among the healthiest in TX. In contrast, several counties south of Dallas County (e.g. Johnson, Ellis, and Kaufman) rank less healthy than the central urban county and its surrounding northern counties. Our results are quite different than those using the CHRs, as the CHRs of only three of the twelve counties in the MSA are within our 95% PIs. We observe more agreement between our ranks and the CHRs among the healthiest counties and more disagreement among the least healthy counties. For example, we rank Johnson County at the 85th percentile (95% PI, 82nd to 86th percentile) compared to the 36th percentile by UW-

PHI. Consequently, the suburban-urban health gradient is more pronounced in our rankings.

[Figure 5 About Here]

Figure 6 shows the results for the Milwaukee CSA. A suburban-urban gradient in health again emerges, with the central county of the urban core – Milwaukee – having the worst health in the CSA and close to the worst health in the state. The northern and western suburban counties of Waukesha, Washington, and Ozaukee have the best health in the CSA and are among the healthiest counties in the state. The southern counties of Walworth and Racine rank relatively poorly but better than Milwaukee County, while the outermost suburbs of Jefferson and Dodge rank in the middle. As with the overall rankings and least healthy counties, our results are much more similar to the CHRs in WI than TX. The CHRs of all but one of the Milwaukee CSA counties are within our 95% PIs, and the conclusions reached about suburban-urban disparities are the same.

[Figure 6 About Here]

4.6 Missing Data

All our rankings thus far use OLS imputations for missing data. This approach does not account for uncertainty about the imputed values, perhaps leading to PIs that are too narrow and estimates of minimum ranking differences necessary for statistically significance that are too small. This is likely especially problematic in TX, which – as discussed in Section 3 – has much more missing data than WI. Figure 7 therefore explores the sensitivity of our results to the use of the iterative imputation procedure. The left column ("Naive Imputation") reproduces the results using the OLS imputations from Figure 2, while the middle column ("Posterior Imputation") uses the iterative procedure. Recall that our main sample includes all counties with valid data for at least one of the five factors. Another useful comparison is therefore to see how the results change using a more restrictive set of counties. To that end, the right column ("Restricted Imputation") drops counties with more than one missing factor. This reduces the number of counties from 223 to 152 in TX, but does not affect the number of counties in WI.

[Figure 7 About Here]

The bottom half of Figure 7 shows that incorporating uncertainty from imputing missing data has essentially no effect on the results for WI. This is because only two of WI's 72 counties have any missing data, and those two counties are missing only one factor. Since no counties are missing more than one factor, the second and third figures are exactly the same.

The top half of Figure 7 shows that the results are much different for TX. Keeping all counties with at least one non-missing factor in the sample ("Posterior Imputation"), we see that incorporating uncertainty from imputing missing data drastically increases the overall amount of uncertainty. The 95% PIs for the 116 counties with at least one missing factor generally expand to include almost the entire range of possible ranks. Even the PIs for counties with no missing data widen considerably in most cases because of the shared component in the uncertainty measure. When we drop the counties with two or more missing factors ("Restricted Imputation"), the level of uncertainty drops relative to the previous "Posterior Imputation" graph, but remains considerably greater than the baseline case with the naive imputation approach. Of the 152 remaining counties, 71 have 95% PIs that reach into the "least healthy quintile" range and 47 of these extend into the "least healthy decile" range.

The TX results illustrate the broader point that when the amount of missing data is extensive, it is not possible to obtain clear conclusions for most counties after accounting for all sources of uncertainty. Attempts to produce comprehensive county health rankings for such states are therefore likely to produce a large number of misleading results. However, this does not necessarily mean that *no* useful results can come from such analyses, only that caution should be used when assessing which results are reliable and which could simply be a product of uncertainty. For instance, even in the full-sample TX graph with the iterative imputation process, the 95% PIs for 10 counties lie entirely to the right of the 80th percentile line. These 10 counties can therefore confidently be identified as among the least healthy in TX despite the fact that clear conclusions cannot be reached for most other TX counties.

5 Conclusion

This paper implements a Bayesian hierarchal model for factor analysis to rank the health of localities, where health is a latent variable that depends on observable factors related to mortality and morbidity. Our model improves on the previous UWPHI CHRs by using a data-driven process to determine factor weights and including a measure of uncertainty that incorporates population, spatial covariance, and noise from imputing missing data. Applying our method to county-level data from TX and WI reveals the importance of these improvements. We show that our data-derived factor weights differ substantially from the deterministic CHR weights in TX but less so in WI. Consequently, our results for overall rankings, least healthy counties, and suburban-urban health gradients are much more similar to the CHRs for WI than for TX. We also document considerable uncertainty in both states. When the measure of uncertainty only reflects population and spatial covariance, the mean rank of counties should generally be at least 25 ranks apart in TX (21 ranks in WI) to be meaningfully different. When we also include the uncertainty inherent in imputation, the overall level of uncertainty rises sharply in TX since the state has considerable missing data. It becomes impossible to reach clear conclusions for most counties, although some of the least healthy counties can still be identified.

Our framework is flexible enough to allow numerous variations in future research on health rankings. Most obviously, our model could be estimated in other states using the same five mortality and morbidity variables. It would also be straightforward to empirically produce alternative rankings using additional health-related variables beyond these five. Indeed, numerous other measures of health outcomes exist and are routinely measured at the population level (Kindig, 2007). Our latent variable framework could incorporate these additional manifestations of health and empirically derive their relationship on health outcomes without requiring subjective expert opinion on variable weights. Moreover, our method could easily be used to produce rankings at geographic levels besides the county, such as the state or MSA. The existing state-level America's Health Rankings (United Health Foundation, 2010) suffer from the same limitations as the CHRs – although sampling error and missing data are probably less problematic at larger geographic levels. Our ranking methodology could even be extended to international comparisons of health system performance (Anderson and Markovich, 2010; Squires, 2011) or human development (United Nations Development Programme, 2011). An additional application could focus on ranking health and healthcare across hospitals, hospital referral regions, and physician hospital networks to identify optimal and suboptimal areas.

We also offer methodological innovations that could be useful in various contexts besides health. While our model relies heavily on the framework developed by Hogan and Tchernis (2004) to rank levels of material deprivation, we add to their framework an explicit procedure for imputing missing data and incorporating noise from the imputation process into the overall measure of uncertainty. Missing or incomplete data at local levels such as the county is certainly not a phenomenon that is unique to health data. Moreover, Hogan and Tchernis (2004) only studied one state – Rhode Island. By including multiple states we are able to demonstrate the importance of allowing the estimated factor weights to vary in each state rather than simply using national data to determine universal weights. It seems reasonable to suspect this would be the case in non-health contexts as well.

We acknowledge several limitations in this study that provide directions for future research. First, we (and the CHRs) only rank counties with at least one morbidity or mortality variable measured. Counties without any measured variables, which are often the smallest, may also be among the most disadvantaged and least healthy. Second, several morbidity measures are based on self-reported physical and mental status, which may be subject to differential recall bias. Next, while rankings are useful to compare the health of populations, they do not convey absolute differences in health. A county may improve in health ranking even though its population became less healthy if the health of other counties declines faster over time. Depending on the research or policy goal, progress may be better measured using absolute measures of population health. Whether we measure population health on an absolute or relative scale, assessing the corresponding statistical uncertainty enables us to evaluate meaningful differences among counties. An additional caveat to our research is that it only sheds light on whether one county is healthier than another; it does not address why this is the case. Additional research is needed to understand how medical, behavioral, social, physical, and biological determinants interact to produce health and perpetuate disparities (Kawachi et al., 2002; Stoddart, 1995). Finally, ranking health at the county level masks potentially substantial heterogeneity within counties. Future work could apply our method to ranking zip codes or census tracts within counties, exploiting spatial covariance to mitigate the considerable uncertainty from ranking at such a narrow level.

In short, we provide several useful insights for researchers and policymakers interested in local area health rankings. First, we demonstrate the importance of allowing the data to determine the weights of the different health variables when constructing an overall health measure. Second, we show that these weights can be quite different in different parts of the country. Next, our results illustrate the significance of appropriately accounting for uncertainty. Even differences in rankings that appear sizeable could still be misleading, especially if the localities ranked are small or if some of the data are imputed. Finally, our finding that missing data can drastically limit the ability to draw clear conclusions underscores the need to invest in comprehensive data if comprehensive locality health rankings are desired.

References

[1] Anderson, G.F. and Markovich, P. (2010). Multinational Comparisons of Health Systems Data. The Commonwealth Fund.

[2] Athens, J.K., Catlin, B.B., Remington, P.L., and Gangnon, R.E. (2013). Using Empirical Bayes Methods to Rank Counties on Population Health Measures. Preventing Chronic Disease, 10, Available http://www.cdc.gov/pcd/issues/2013/pdf/13_0028.pdf.

[3] Bartholomew, D. and Knott, M. (1999). *Latent Variable Models and Factor Analysis.* New York: Oxford University Press.

[4] Best, N. and Hansell, A.L. (2009). Geographic variations in risk: adjusting for unmeasured confounders through joint modeling of multiple diseases. *Epidemiology*, 20(3), 400–410.

[5] Booske, B., Rohan, A., Kindig, D., and Remington, P. (2010). Grading and reporting health and health disparities. *Preventing Chronic Disease*, 7(1), A16.

[6] Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4), 327–335.

[7] Chib, S. and Greenberg, E. (1996). Markov Chain Monte Carlo Simulation Methods in Econometrics. *Econometric Theory*, 12, 409–31

[8] Erwin, P.C., Myers, C., Myers, G., and Daugherty, L. (2011). State responses to America's Health Rankings: The search for meaning, utility, and value. *Journal of Public Health Management and Practice*, 17(5), 406–412.

[9] Friedman, D.J. and Starfield, B. (2003). Models of population health: Their value for US public health practice, policy, and research. *American Journal of Public Health*, 93(3), 366–369.

[10] Gelfand, A.E. and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398–409.

[11] Hogan, J. W. and Tchernis, R. (2004). Bayesian factor analysis for spatially correlated data, with application to summarizing area-level material deprivation from census data. *Journal of the American Statistical Association*, 99(466), 314–324.

[12] Institute of Medicine (2011). *For the Public's Health: The Role of Measurement in Action and Accountability.* Washington, DC: The National Academies Press.

[13] Kanarek, N., Tsai, H.-L., and Stanley, J. (2011). Health rankings of the largest US counties using the Community Health Status Indicators Peer Strata and Database. *Journal of Public Health Management and Practice*, 17(5), 401–405.

[14] Kawachi, I., Subramanian, S. V., and Almeida-Filho, N. (2002). A glossary for health inequalities. *Journal of Epidemiology and Community Health*, 56(9), 647–652.

[15] Kindig, D. (2007). Understanding population health terminology. *Milbank Quarterly*, 85(1), 139–161.

[16] Kindig, D., Asada, Y., and Booske, B. (2008). A population health framework for setting national and state health goals. *Journal of the American Medical Association*, 299(17), 2081–2083.

[17] Kindig, D. and Stoddart, G. (2003). What Is population health? *American Journal of Public Health*, 93(3), 380–383.

[18] McAuley, E., Konopack, J.F., Motl, R.W., Morris, K.S., Doerksen, S.E., and Rosengren, K.R. (2006). Physical activity and quality of life in older adults: influence of health status and self-efficacy. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, 31(1), 99–103.

[19] McDowell, I., Spasoff, R.A., and Kristjansson, B. (2004). On the classification of population health measurements. *American Journal of Public Health*, 94(3), 388–393.

[20] Murray, C.J., Gakidou, E.E., and Frenk, J. (1999). Health inequalities and social group differences: what should we measure? *Bulletin of the World Health Organization*, 77(7), 537–543.

[21] Office of Management and Budget. (2009). Update of Statistical Area Definitions and Guidance on Their Uses.

[22] Peppard, P.E., Kindig, D.A., Dranger, E., Jovaag, A., and Remington, P.L. (2008). Ranking community health status to stimulate discussion of local public health issues: The Wisconsin County Health Rankings. *American Journal of Public Health,* 98(2), 209–212.

[23] Peppard, P., Kindig, D., Riemer, A., Dranger, E., and Remington, P. (2003). Wisconsin County Health Rankings, 2003. Wisconsin Public Health and Health Policy Institute.

[24] Phillips, C. D. and McLeroy, K. R. (2004). Health in rural America: Remembering the importance of place. *American Journal of Public Health*, 94(10), 1661–1663.

[25] Population Health Institute. (2011). County Health Rankings. University of Wisconsin.

[26] Remington, P.L. and Booske, B.C. (2011). Measuring the health of communities – how and why? Journal of Public Health Management and Practice, 17(5), 397-400.

[27] Squires, D. (2011). The U.S. Helath System in Perspective: A Comparison of Twelve Industrialized Nations. The Commonwealth Fund.

[28] Stoddart, G. (1995). The challenge of producing health in modern economies. Centre for Health Economics and Policy Analysis Working Paper Series No. 1995-15, Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada.

[29] United Health Foundation. (2010). America's Health Rankings. United Health Foundation, American Public Health Association, Partnership for Prevention.

[30] United Nations Development Programme. (2011). Human Development Report 2011. United Nations.

[31] Van Doorslaer, E. and Jones, A.M. (2003). Inequalities in self-reported health: validation of a new approach to measurement. *Journal of Health Economics*, 22(1), 61–87.

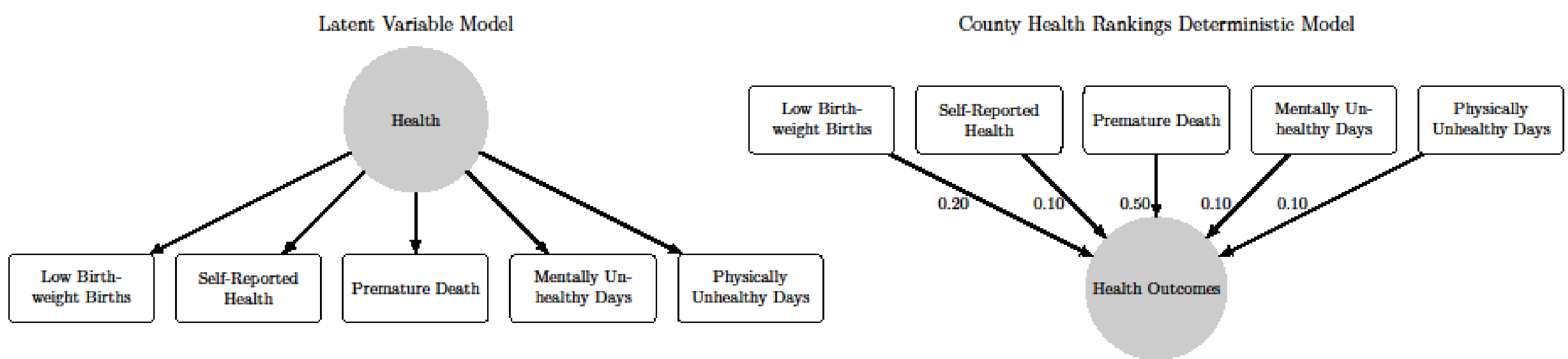
[32] Weden, M.M., Bird, C.E., Escarce, J.J., and Lurie, N. (2011). Neighborhood archetypes for population health research: Is there no place like home? *Health and Place*, 17(1), 289 – 299.

Table 1: UWPHI Deterministic Weights and Normalized Square Correlations

	UWPHI	Texas		Wisconsin	
Health Outcomes	ω	ρ^2	95% CI	ρ^2	95% CI
Premature Death	0.50	0.14	(0.09,0.19)	0.27	(0.17,0.38)
Self-Reported Health Status	0.10	0.24	(0.20,0.29)	0.21	(0.12,0.30)
Physically Unhealthy Days	0.10	0.41	(0.34,0.48)	0.21	(0.11,0.31)
Mentally Unhealthy Days	0.10	0.15	(0.10,0.20)	0.17	(0.08,0.25)
Low Birthweight Births	0.20	0.06	(0.02,0.10)	0.15	(0.05,0.24)

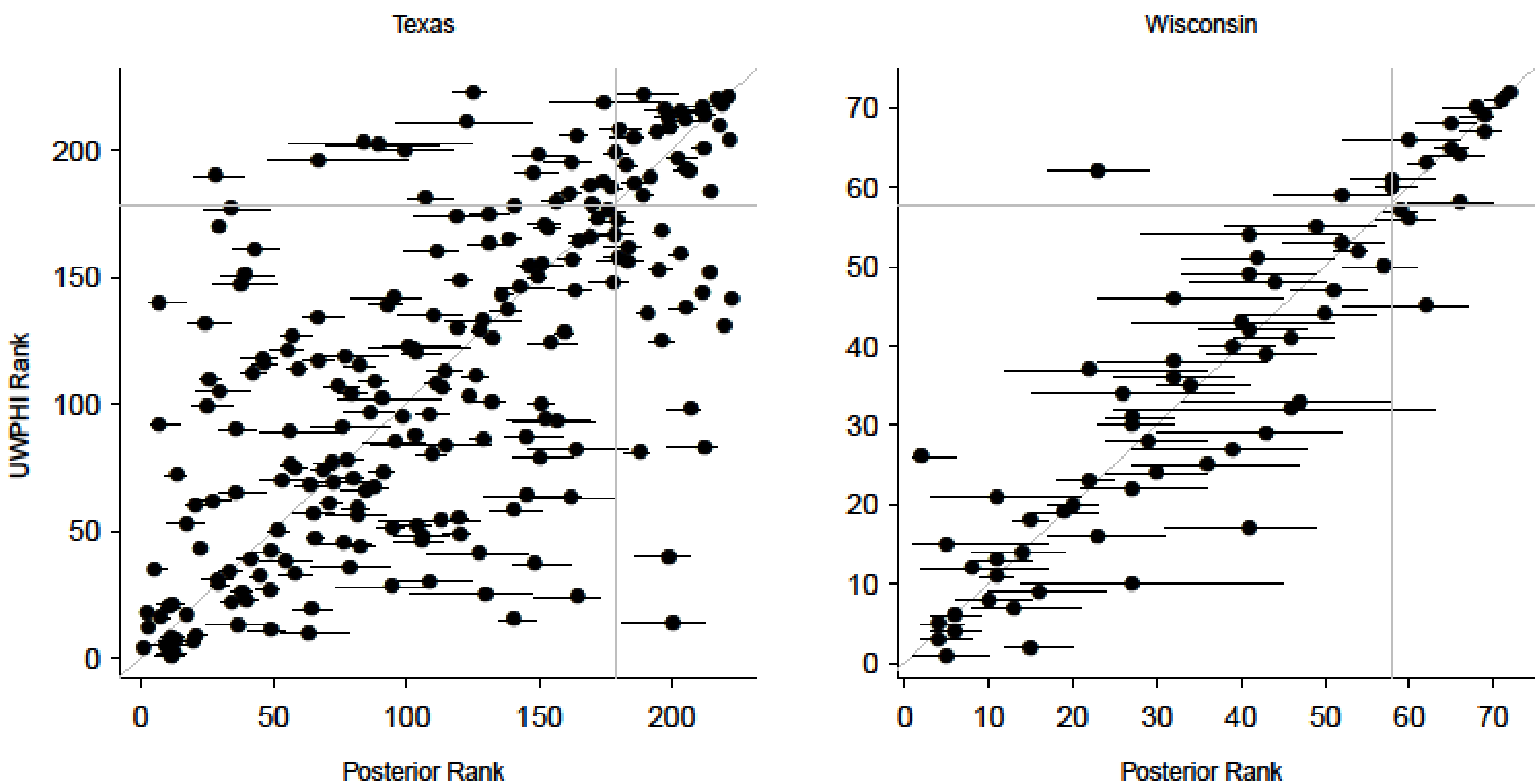
Note: UWPHI=University of Wisconsin Population Health Institute; ω=weight; ρ^2=squared correlation; CI=confidence interval.

Figure 1: Latent Construct & Deterministic Models of Health Outcomes



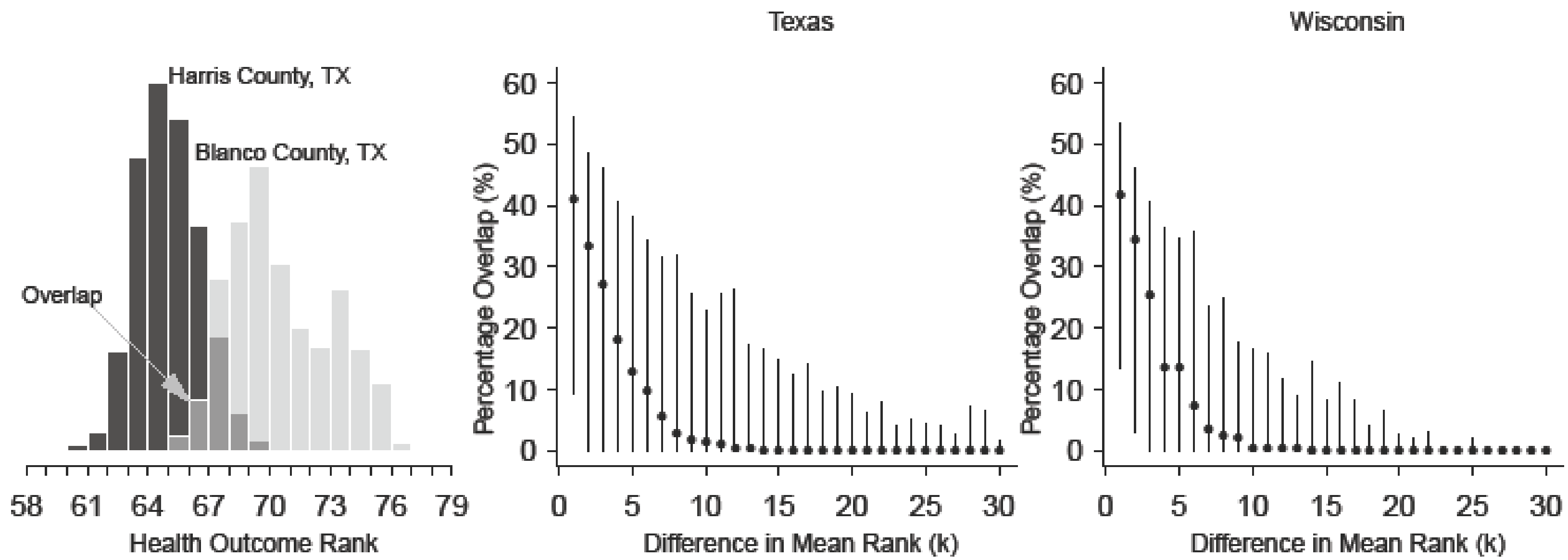
The left panel shows our latent variable framework in which observed morbidity and mortality variables are manifestations of the latent construct - health. The right panel shows the County Health Rankings model of health outcomes, which explicitly consists of four morbidity variables and one mortality variable that combine to form health outcomes. The weights assigned to each variable are shown beside each arrow.

Figure 2: UWPHI Rank, Mean Posterior Ranks, and 95% Probability Intervals



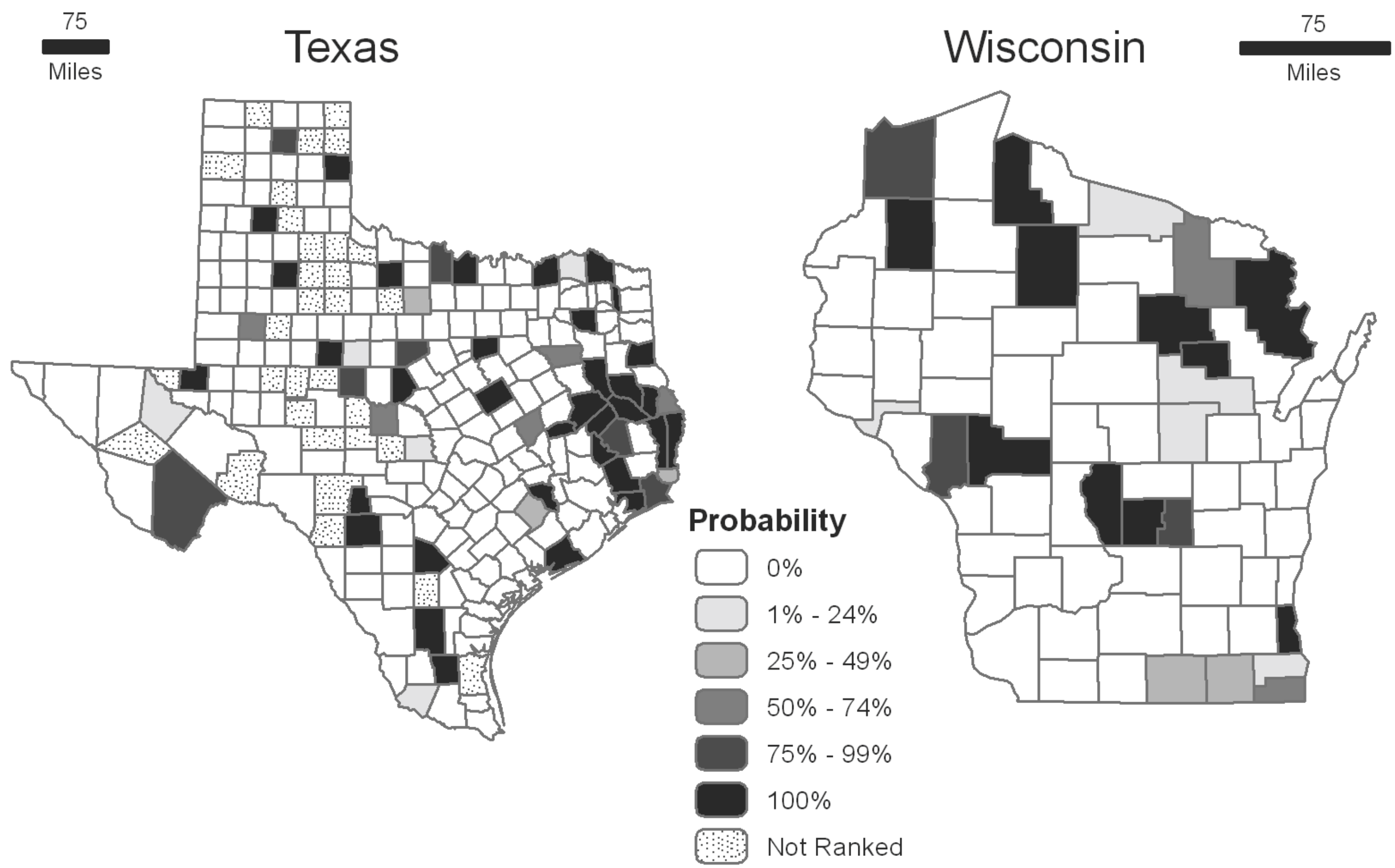
The left (right) panel shows the posterior rank and UWPHI rank for each county in TX (WI). The 95% probability interval of the posterior distribution is denoted by a horizontal line and mean posterior rank is denoted by a solid circle. The gray horizontal and vertical lines represent the 80th percentile of ranks and the 45 degree line represents equality between the UWPHI and posterior ranks.

Figure 3: Health Outcome Rankings Posterior Distribution & Overlap in Posterior Distribution of Ranks Between Counties



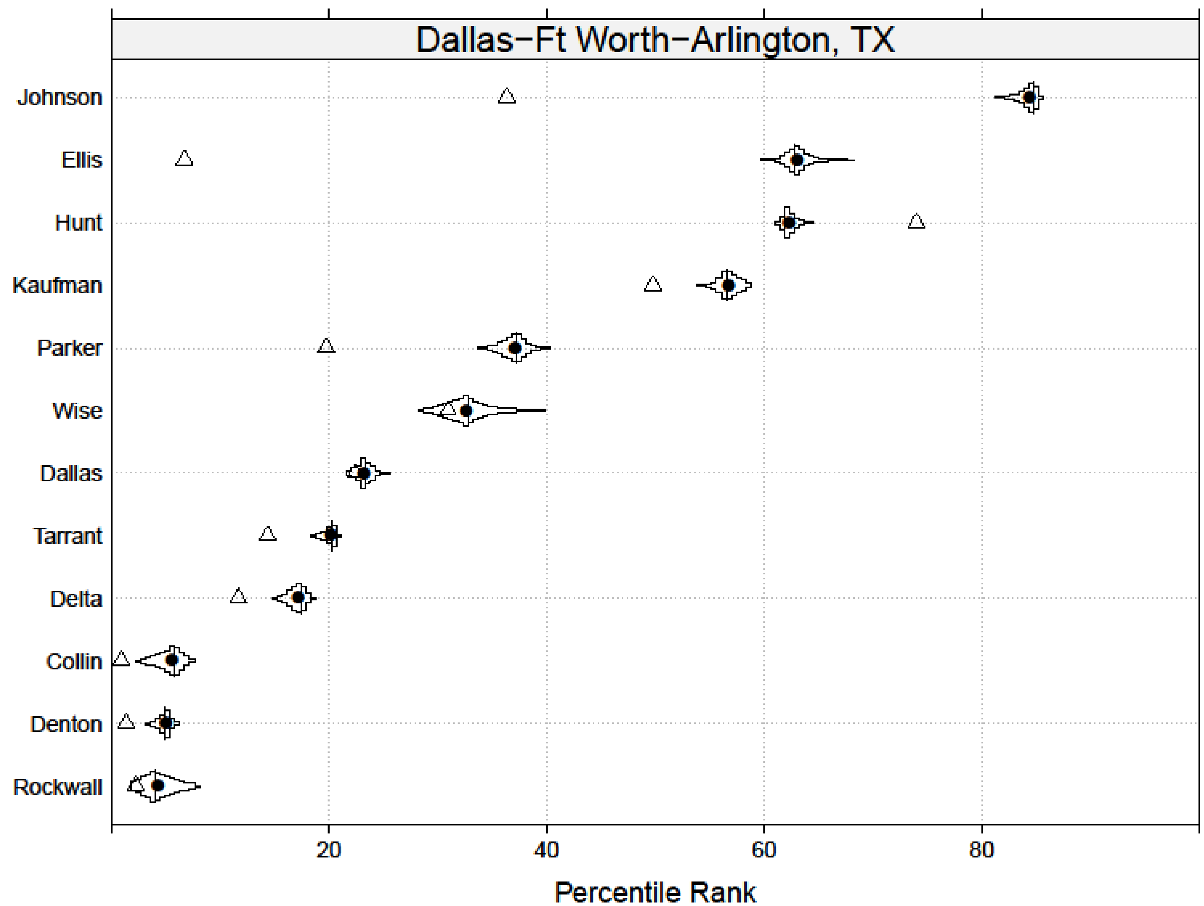
The left panel shows the posterior distribution of health outcomes rankings for Harris County, TX (mean rank 66) and Blanco County, TX (mean rank 71). The overlap between the two posterior distributions shown in dark grey. The middle (right) panel shows the percentage overlap in the posterior distribution of ranks between counties that are k units apart in mean rank for Texas (Wisconsin). The 95% probability interval of the percentage overlap is denoted by a vertical line and median percentage overlap is denoted by a solid circle.

Figure 4: Probability of Being in Least Healthy Quintile, TX and WI



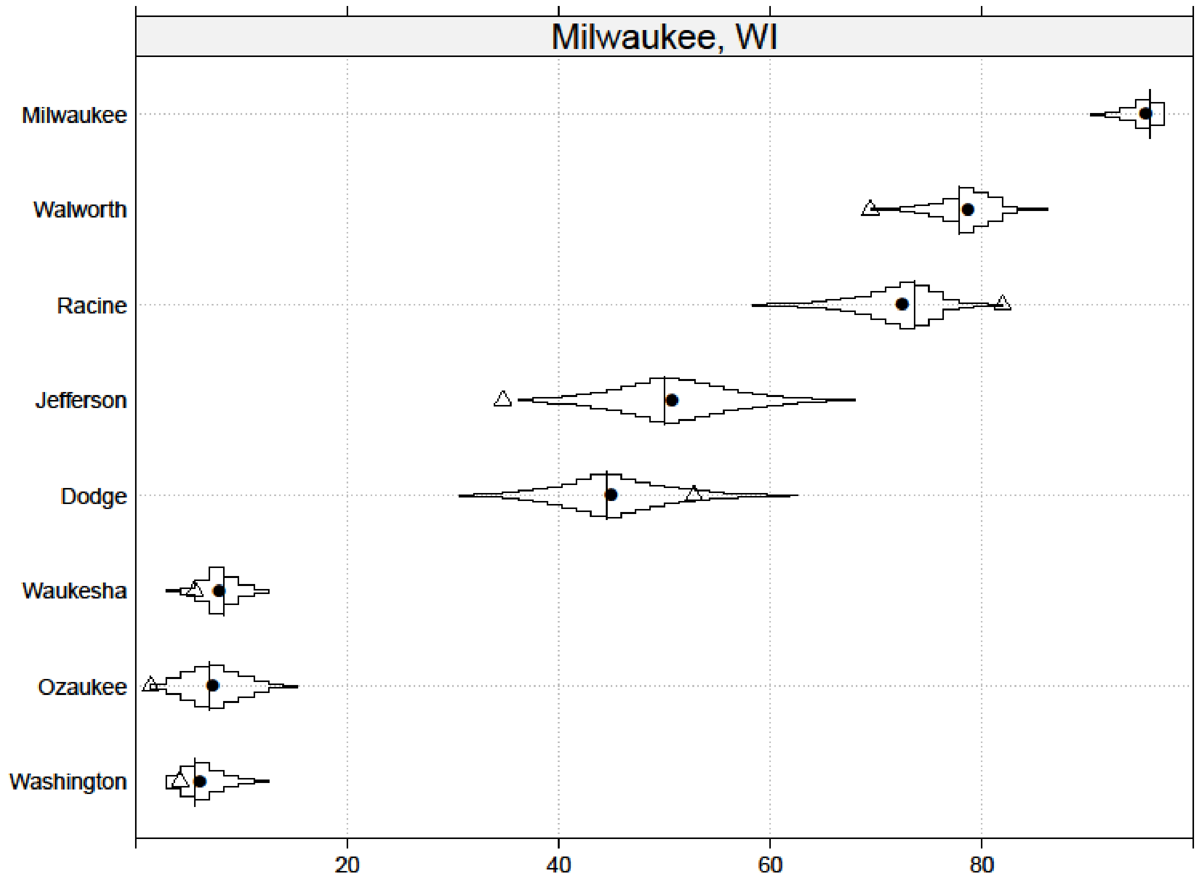
The left (right) panel shows the probability of being in the least healthy quintile under our model for Texas (Wisconsin). Unranked counties dotted.

Figure 5: Posterior Distribution of County Ranks in the Dallas-Ft. Worth-Arlington, TX MSA



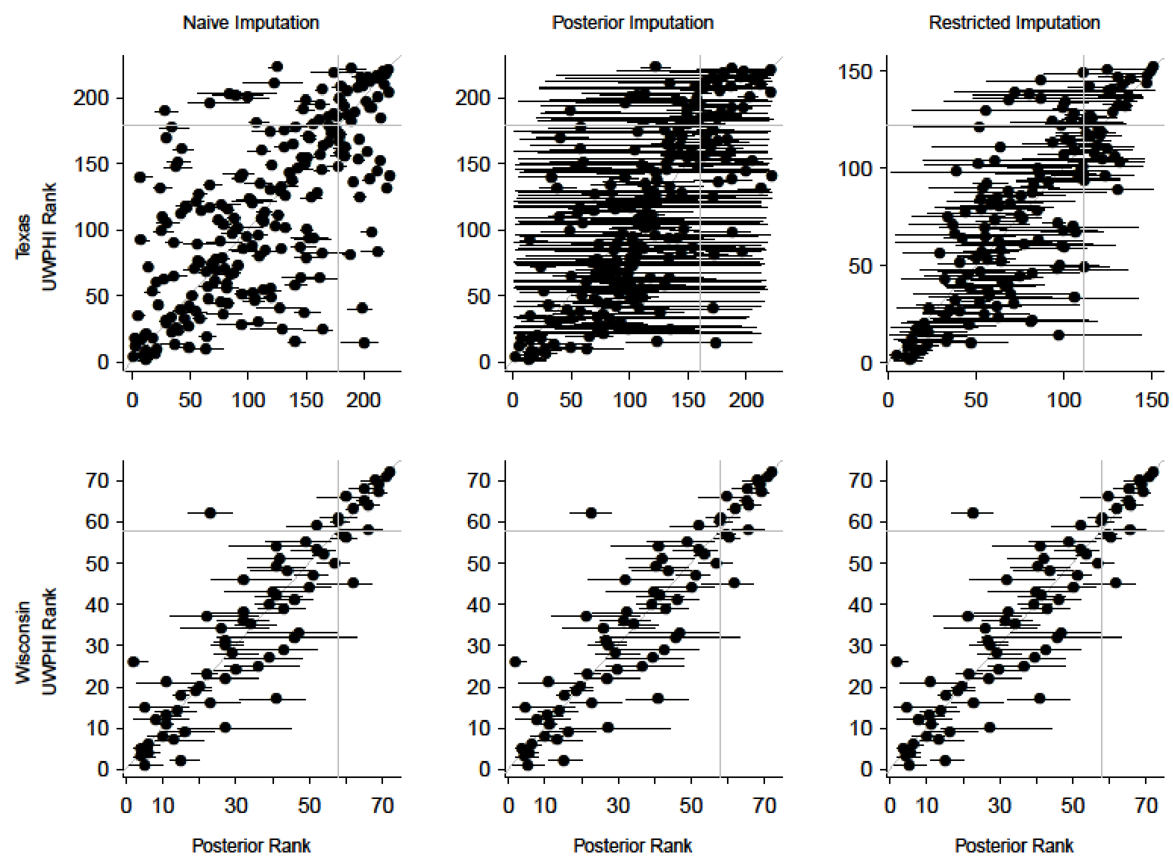
The posterior distribution of health outcome rankings for each county in the MSA. The median of the posterior distribution is shown as a vertical line (|), and the mean as a solid circle (•). The single UWPHI rank for each county is shown as an open triangle (Δ).

Figure 6: Posterior Distribution of County Ranks in the Milwaukee-Racine-Waukesha, WI CSA



The posterior distribution of health outcome rankings for each county in the MSA. The median of the posterior distribution is shown as a vertical line (|), and the mean as a solid circle (●). The single UWPHI rank for each county is shown as an open triangle (∆).

Figure 7: UWPHI Rank, Mean Posterior Ranks, and 95% Probability Intervals with Different Approaches to Missing Data



The top (bottom) panel shows the posterior and UWPHI ranks for each county in TX (WI). The 95% probability interval of the posterior distribution is denoted by a horizontal line and mean posterior rank by a solid circle. The gray horizontal and vertical lines represent the 80th percentile of ranks and the 45° line represents equality between the UWPHI and posterior ranks. The restricted imputation excludes 71 TX counties with more than one missing variable.

Appendix Table A1: Means, Standard Deviations, and 95% PIs of Posterior Distributions of Model Parameters

Parameter	Factor	TX			WI		
		Mean	Std. Dev.	95% PI	Mean	Std. Dev.	95% PI
μ	Premature Death	7.98	0.13	[7.72,8.25]	6.24	0.19	[5.85,6.62]
	Self-Reported Health Status	1.90	0.04	[1.83,1.97]	1.19	0.04	[1.11,1.28]
	Physically Unhealthy Days	3.76	0.07	[3.62,3.88]	3.20	0.10	[3.00,3.39]
	Mentally Unhealthy Days	3.44	0.06	[3.32,3.56]	2.97	0.09	[2.79,3.15]
	Low Birthweight Births	8.09	0.07	[7.95,8.23]	6.35	0.14	[6.07,6.62]
λ	Premature Death	15.11	2.03	[11.20,19.04]	15.26	2.63	[10.45,20.82]
	Self-Reported Health Status	5.18	0.48	[4.27,6.16]	3.20	0.61	[2.04,4.47]
	Physically Unhealthy Days	12.00	0.79	[10.52,13.66]	7.34	1.46	[2.04,4.47]
	Mentally Unhealthy Days	7.41	0.92	[5.66,9.26]	5.95	1.30	[3.57,8.64]
	Low Birthweight Births	5.17	1.13	[3.00,7.48]	8.99	2.19	[4.86,13.41]
σ	Premature Death	580.13	61.85	[468.50,713.54]	151.56	43.41	[77.83,247.67]
	Self-Reported Health Status	27.25	3.10	[21.58,33.71]	11.59	2.53	[7.23,17.03]
	Physically Unhealthy Days	29.08	9.02	[13.51,48.71]	63.17	14.71	[38.48,95.52]
	Mentally Unhealthy Days	121.80	12.76	[98.49,148.96]	60.48	12.21	[40.58,88.22]
	Low Birthweight Births	214.16	21.31	[176.22,259.53]	163.92	33.39	[109.49,240.35]
ω	–	0.04	0.05	[-0.07,0.13]	0.05	0.08	[-0.14,0.18]

Appendix Table A2: County Health Rankings in TX

County Name	Posterior Mean	95% PI	CHR Rank
Anderson	89.74	[70,112]	202
Andrews	25.66	[22,30]	110
Angelina	196.24	[193,199]	168
Aransas	95.20	[79,105]	142
Archer	78.67	[64,94]	36
Atascosa	211.59	[210,213]	144
Austin	214.40	[213,216]	152
Bailey	64.94	[57,73]	57
Bandera	90.82	[74,113]	102
Bastrop	58.02	[55,63]	75
Baylor	212.14	[208,215]	201
Bee	56.13	[45,76]	89
Bell	84.69	[75,90]	66
Bexar	80.17	[75,84]	71
Blanco	71.06	[67,76]	61
Bosque	132.25	[123,137]	101
Bowie	162.33	[152,170]	195
Brazoria	29.17	[26,33]	29
Brazos	13.02	[7,17]	7
Brewster	200.39	[181,212]	14
Brooks	205.22	[200,211]	212
Brown	192.03	[190,194]	189
Burleson	72.09	[65,78]	77
Burnet	10.51	[8,15]	20
Caldwell	41.34	[36,44]	39
Calhoun	6.96	[4,15]	92
Callahan	111.38	[107,114]	108
Cameron	148.32	[140,162]	37
Camp	149.70	[140,163]	198

County Name	Posterior Mean	95% PI	CHR Rank
Carson	53.12	[45,61]	70
Cass	162.48	[157,166]	157
Castro	127.92	[125,130]	129
Chambers	219.89	[219,220]	131
Cherokee	203.22	[199,206]	159
Childress	129.03	[124,132]	86
Clay	198.88	[186,207]	40
Cochran	146.39	[143,152]	154
Coleman	135.74	[134,140]	143
Collin	12.31	[6,17]	2
Collingsworth	66.83	[61,73]	117
Colorado	176.03	[169,182]	176
Comal	11.47	[8,16]	8
Comanche	63.78	[58,70]	68
Cooke	82.37	[77,88]	115
Coryell	140.42	[132,151]	58
Crane	81.60	[70,92]	56
Crockett	131.32	[124,139]	175
Crosby	222.91	[221,223]	204
Culberson	35.96	[29,43]	90
Dallam	153.32	[148,158]	169
Dallas	51.73	[48,56]	50
Dawson	179.84	[174,185]	158
Deaf Smith	76.03	[64,94]	91
Delta	38.26	[34,42]	26
Denton	11.11	[8,14]	3
DeWitt	2.16	[2,3]	18
Dimmit	91.45	[86,95]	73
Donley	120.42	[115,125]	149

Appendix Table A2: County Health Rankings in TX (Continued)

County Name	Posterior Mean	95% PI	CHR Rank
Duval	189.15	[187,193]	182
Eastland	183.68	[174,188]	162
Ector	147.67	[141,157]	191
El Paso	22.43	[21,24]	15
Ellis	140.56	[135,149]	43
Erath	7.55	[6,11]	16
Falls	28.02	[20,39]	190
Fannin	205.36	[199,209]	138
Fayette	6.82	[3,17]	140
Fisher	119.08	[103,129]	174
Floyd	163.60	[157,170]	145
Fort Bend	48.93	[40,54]	11
Franklin	2.89	[2,4]	12
Freestone	46.58	[42,51]	116
Frio	98.72	[96,101]	95
Gaines	59.28	[55,63]	114
Galveston	42.04	[38,47]	112
Garza	149.47	[146,153]	150
Gillespie	1	[1,1]	4
Goliad	49	[42,53]	42
Gonzales	42.77	[35,52]	161
Gray	66.81	[48,101]	196
Grayson	128.80	[115,143]	133
Gregg	107.23	[101,118]	181
Grimes	24.97	[19,35]	99
Guadalupe	64.25	[57,72]	19
Hale	81.66	[57,72]	19
Hall	119.21	[117,122]	130
Hamilton	33.63	[29,38]	34

County Name	Posterior Mean	95% PI	CHR Rank
Hansford	34.20	[30,38]	22
Hardeman	46.02	[38,51]	118
Hardin	110.32	[97,121]	135
Harris	65.50	[63,69]	47
Harrison	122.80	[96,147]	211
Hartley	55.14	[49,61]	121
Haskell	88.34	[82,93]	109
Hays	63.33	[53,78]	10
Henderson	179.49	[171,185]	172
Hidalgo	164.56	[148,173]	24
Hill	57.10	[53,64]	127
Hockley	76.88	[65,93]	119
Hood	150.43	[140,163]	79
Hopkins	35.91	[29,47]	65
Houston	199.13	[196,202]	209
Howard	140.85	[135,150]	178
Hudspeth	111.68	[99,119]	160
Hunt	138.82	[136,143]	165
Hutchinson	183.34	[172,189]	156
Jack	132.49	[129,135]	126
Jackson	57.99	[52,64]	33
Jasper	205.44	[202,209]	193
Jefferson	185.61	[177,189]	205
Jim Hogg	142.88	[143,156]	146
Jim Wells	161.52	[156,166]	183
Johnson	188.14	[182,191]	81
Jones	29.76	[21,41]	105
Karnes	113.75	[110,117]	106
Kaufman	126.41	[121,131]	111

Appendix Table A2: County Health Rankings in TX (Continued)

County Name	Posterior Mean	95% PI	CHR Rank
Kendall	20.83	[121,131]	111
Kerr	151.18	[142,159]	155
Kimble	95.81	[84,107]	85
Kleberg	109.67	[104,115]	80
Knox	154.57	[146,164]	124
La Salle	150.86	[146,156]	188
Lamar	174.05	[171,178]	170
Lamb	29.45	[27,32]	55
Lampasas	119.83	[117,123]	100
Lavaca	36.72	[25,47]	13
Lee	129.84	[102,147]	25
Leon	39.34	[30,50]	151
Liberty	212.46	[209,216]	214
Limestone	99.44	[88,118]	200
Live Oak	103.41	[101,106]	88
Llano	161.86	[143,178]	63
Lubbock	66.70	[61,77]	134
Lynn	170.01	[167,173]	179
Madison	212.34	[198,217]	148
Marion	125.34	[118,130]	125
Martin	104.03	[97.5,109]	83
Matagorda	222.71	[222,223]	223
Maverick	20.66	[17,26]	52
McCulloch	178.11	[169,184]	141
McLennan	196.29	[191,201]	60
Medina	106.11	[95,116]	48
Midland	17.38	[15,18]	17
Milam	83.95	[56,125]	203
Mills	11.93	[6,16]	21

County Name	Posterior Mean	95% PI	CHR Rank
Mitchell	88.01	[82,93]	67
Montague	195.18	[191,200]	153
Montgomery	120.28	[114,124]	49
Moore	145.20	[137,159]	87
Morris	197.34	[190,209]	216
Nacogdoches	186.09	[181,193]	187
Navarro	131.27	[123,139]	163
Newton	216.95	[216,218]	220
Nolan	203.33	[198,210]	215
Nueces	123.60	[121,126]	103
Ochiltree	39.67	[34,44]	23
Orange	174.30	[154,195]	219
Palo Pinto	138.16	[132,143]	137
Panola	206.76	[201,209]	192
Parker	82.65	[76,88]	44
Parmer	113.16	[90,128]	54
Pecos	5.04	[4,10]	35
Polk	189.24	[179,202]	222
Potter	164.17	[158,168]	206
Presidio	114.89	[87,132]	84
Rains	165.03	[161,168]	164
Randall	94.54	[74,108]	28
Reagan	27.20	[22,34]	62
Real	202.31	[198,209]	197
Red River	219.00	[218,220]	218
Reeves	177.11	[175,181]	185
Refugio	77.83	[73,84]	78
Robertson	178.87	[174,184]	199
Rockwall	9.44	[4,18]	5

Appendix Table A2: County Health Rankings in TX (Continued)

County Name	Posterior Mean	95% PI	CHR Rank
Runnels	182.94	[179,187]	194
Rusk	37.77	[27,51]	147
Sabine	180.29	[173,188]	208
San Augustine	194.53	[191,201]	207
San Jacinto	218.06	[217,219]	210
San Patricio	86.63	[77,95]	97
San Saba	68.59	[64,73]	74
Scurry	24.30	[18,34]	132
Shackelford	159.71	[156,163]	128
Shelby	34.13	[24,49]	177
Smith	94.84	[89,99]	51
Somervell	56.28	[53,60]	76
Starr	164.04	[146,184]	82
Stephens	74.44	[69,82]	107
Sutton	79.22	[74,85]	104
Swisher	214.76	[213,216]	184
Tarrant	44.89	[41,46]	32
Taylor	169.42	[161,177]	166
Terry	114.61	[108,121]	113
Titus	17.15	[10,24]	53
Tom Green	54.55	[45,64]	38
Travis	19.89	[18,22]	6
Trinity	221.37	[221,222]	221
Tyler	169.41	[166,172]	186
Upshur	156.54	[151,162]	180
Upton	108.66	[104,116]	96
Uvalde	207.13	[198,211]	98
Val Verde	127.71	[108,146]	41
Van Zandt	152.07	[146,158]	171

County Name	Posterior Mean	95% PI	CHR Rank
Victoria	102.74	[86,124]	122
Walker	152.29	[138,169]	94
Waller	100.72	[91,120]	123
Ward	172.17	[169,174]	173
Washington	13.70	[9,17]	72
Webb	108.82	[97,125]	30
Wharton	145.40	[129,166]	64
Wheeler	211.46	[205,216]	217
Wichita	93.09	[86,98]	139
Wilbarger	156.73	[140,171]	93
Willacy	48.71	[43,52]	27
Williamson	11.50	[5,16]	1
Wilson	76.38	[66,87]	45
Winkler	198.64	[194,206]	213
Wise	72.67	[64,86]	69
Wood	190.84	[187,193]	136
Yoakum	105.80	[95,114]	46
Young	178.40	[173,185]	167
Zapata	28.82	[23,35]	31
Zavala	103.57	[98,113]	120

Appendix Table A3: County Health Rankings in WI

County Name	Posterior Mean	95% PI	CHR Rank
Adams	70.84	[69,72]	71
Ashland	65.17	[63,67]	65
Barron	46.01	[39,51]	41
Bayfield	39.34	[35,44]	40
Brown	39.27	[27,48]	27
Buffalo	27.02	[14,45]	10
Burnett	41.21	[35,48]	42
Calumet	13.19	[8,21]	7
Chippewa	27.39	[23,32]	30
Clark	42.96	[36,49]	39
Columbia	21.57	[18,25]	23
Crawford	42.15	[33,51]	51
Dane	16.11	[10,24]	9
Dodge	32.35	[23,43]	38
Door	15.14	[13,17]	18
Douglas	59.48	[57,61]	57
Dunn	9.85	[6,15]	8
Eau Claire	18.56	[15,23]	19
Florence	39.59	[27,51]	43
Fond du Lac	49.94	[40,56]	44
Forest	59.82	[52,66]	66
Grant	14.02	[8,19]	14
Green	11.20	[3,21]	21
Green Lake	61.60	[52,67]	45
Iowa	7.90	[2,17]	12
Iron	21.59	[12,36]	37
Jackson	65.32	[61,68]	68
Jefferson	36.49	[27,47]	25
Juneau	53.66	[51,55]	52

County Name	Posterior Mean	95% PI	CHR Rank
Kenosha	57.95	[53,63]	61
Kewaunee	11.28	[9,13]	11
La Crosse	26.68	[21,36]	22
Lafeyette	26.14	[15,39]	34
Langlade	69.11	[66,71]	67
Lincoln	51.11	[46,55]	47
Manitowoc	34.20	[30,41]	35
Marathon	19.68	[17,23]	20
Marinette	62.00	[60,63]	63
Marquette	68.08	[64,71]	70
Menominee	71.94	[71,72]	72
Milwaukee	68.73	[66,70]	69
Monroe	40.51	[33,49]	49
Oconto	42.56	[30,52]	29
Oneida	41.19	[28,52]	54
Outagamie	10.52	[8,15]	13
Ozaukee	5.21	[1,10]	1
Pepin	46.59	[33,58]	33
Pierce	6.25	[3,9]	6
Polk	26.70	[24,32]	31
Portage	22.72	[17,31]	16
Price	65.57	[60,70]	58
Racine	52.16	[44,58]	59
Richland	31.61	[25,39]	36
Rock	57.65	[56,61]	60
Rusk	32.13	[23,45]	46
Sauk	1.83	[1,6]	26
Sawyer	22.64	[17,29]	62
Shawano	46.30	[25,63]	32

Appendix Table A3: County Health Rankings in WI (Continued)

County Name	Posterior Mean	95% PI	CHR Rank
Sheboygan	40.68	[31,49]	17
St. Croix	15.14	[12,20]	2
Taylor	3.60	[2,7]	5
Trempealeau	60.26	[56,63]	56
Vernon	4.83	[1,17]	15
Vilas	49.04	[38,56]	55
Walworth	56.64	[52,61]	50
Washburn	65.64	[62,69]	64
Washington	4.33	[2,8]	3
Waukesha	5.65	[3,9]	4
Waupaca	51.97	[45,57]	53
Waushara	43.77	[34,50]	48
Winnebago	28.95	[24,36]	28
Wood	29.70	[24,36]	24