International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

SOCIOECONOMIC EFFECTS ON ALL-CAUSE MORTALITY: EVIDENCE FROM THE AMERICAN RURAL WEST

Mitch Kunce

*Douglas**Mitchell*** Econometric Consulting Laramie, WY USA

DOI: 10.46609/IJSSER.2021.v06i10.042 URL: https://doi.org/10.46609/IJSSER.2021.v06i10.042

ABSTRACT

This study seeks to decompose the variation in place-based all-cause mortality by focusing on within-state (Wyoming, U.S.) county-level socioeconomic and structural factors. A deeper understanding of the significant role these factors play may foster policy solutions that can better address mortality disparity. Results from several two-way error component empirical specifications finds no evidence supporting the contention that observable socioeconomic factors matter in explaining variation in mortality within the state of Wyoming. Evidence from an all county panel spanning over 11 years suggests that differences in all-cause mortality rates can be explained by the asymmetric and growing share of the population aged 65 and above (the greying of many parts of the state) and certain latent county and time-specific effects. Implications of the investigation call for a shift in research focus from the aggregate to individual-level data controlling for key socioeconomic resources.

Keywords: Wyoming mortality rates, Rural mortality disparities, Socioeconomic effects

JEL Classification: I15, J11, N32, O18

1. Introduction

Over the past three decades, place-based mortality patterns have shifted in the United States. The historic 'urban mortality penalty' has evolved to greater mortality disparities found in rural America. For example, in 1980 the difference between rural and urban age-standardized mortality (rural minus urban) was around -1 death per 100,000 population (Cosby et al 2008). With time, this wedge has changed signs and increased to +135 by 2016 (Cosby et al 2019). Recent analyses indicate that the rural-urban difference is broad based − occurring in all U.S. census divisions (Gong et al 2019) and across all leading causes of death (Cosby et al 2019). It is important to note that the growing difference is not the result of rising overall mortality in rural American − rather the rural disparity is the consequence of overall urban mortality declining at a much faster rate (James 2014).

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Because this rural disparity is a recent discovery (Cosby et al 2008), definitive causes continue to be assessed. Early findings suggest that socioeconomic forces may play a major role in determining differences in place-based mortality. Important socioeconomic determinants including education, income, poverty and unemployment have been found to be significantly associated with mortality variation (see Spencer et al 2018 and Cosby et al 2019 for extensive reviews). Other factors associated with rural residence such as race, heath care access and insurance coverage have also explained differences in place-based mortality (Gong et al 2019). Based on the current inventory of literature − higher percentages of non-white populations, higher percentages of poverty, higher rates of unemployment, lower levels of education attainment, lower levels of income and deficiencies in health care access are associated with higher mortality.

Interestingly, socioeconomic effects on variations in all-cause mortality *within* smaller jurisdictions (e.g. U.S. states) has received little attention. This void in the literature is puzzling in part because unpacking within state differences could assist best placed state-level policy efforts addressing rural mortality disparities. In this paper we extend the analysis by examining socioeconomic variation in mortality within the least populated state in the union, Wyoming. Examining Wyoming, with its diverse and asymmetric natural resource based economy, provides the perfect natural experiment aimed to disentangle veritable rural mortality disparity. This research debits the literature inventory by utilizing a balanced panel of time-series, cross-section data from 2010 - 2020 (T = 11) and across all 23 counties (N = 23) within the state. Exploiting the richness of this panel, allowances are made for any unobservable county and year effects which could otherwise bias results based on single cross-sections or time-series (Kunce 2021).

Results from several two-way error component specifications finds no evidence supporting the contention that observable socioeconomic factors matter in explaining variation in mortality within the state of Wyoming. Evidence from the all county panel suggests that differences in all-cause mortality rates can be explained by the asymmetric and growing share of the population aged 65 and above (the greying of many parts of the state) and certain latent county and time-specific effects. The balance of this examination is divided into four sections. Section 2 describes the data, provides sources and presents descriptive statistics. Additional attention is paid to one key dependent variable, all-cause age-standardized mortality rates. Section 3 presents the empirical model and discusses the complex econometric issues. Section 4 interprets the empirical results with conclusions and implications drawn in section 5.

2. Data

Figure 1 shows the age-standardized death rate trends for the state of Wyoming and two counties – Niobrara, population 2,467 least in 2020 and Laramie, population 100,512 most in 2020. The Laramie county trend appears relatively flat to slightly increasing, hovering around the county 11 year mean of 7.7 deaths per 1,000 population. Niobrara county's trend presents erratic, mostly driven by the low population denominator. Niobrara county's 11 year mean is 7.3 per 1,000 population. The state of Wyoming's 11 year overall mean is 7.6 deaths per 1,000 population.

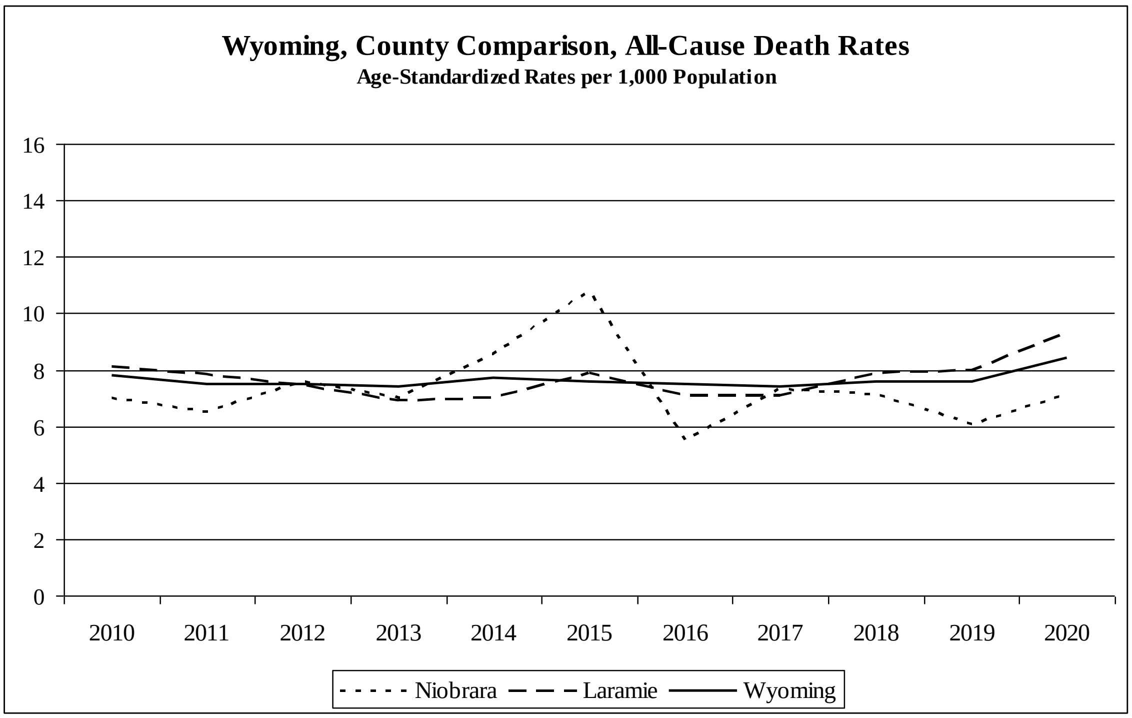

Fig. 1. Wyoming, County Annual Death Rate Comparison, 2010-2020

Means, over the 11 year panel, of annual age-standardized death rates exhibit considerable variation across counties. Highest rates are found in Hot Springs 9.6 and Freemont 9.1, whereas the lowest rates appear in Teton 4.4 and Sublette 5.6. Figure 2 provides a map of the state showing each county, seat, 2020 census population and the sample 11 year mean death rate per 1,000 population.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

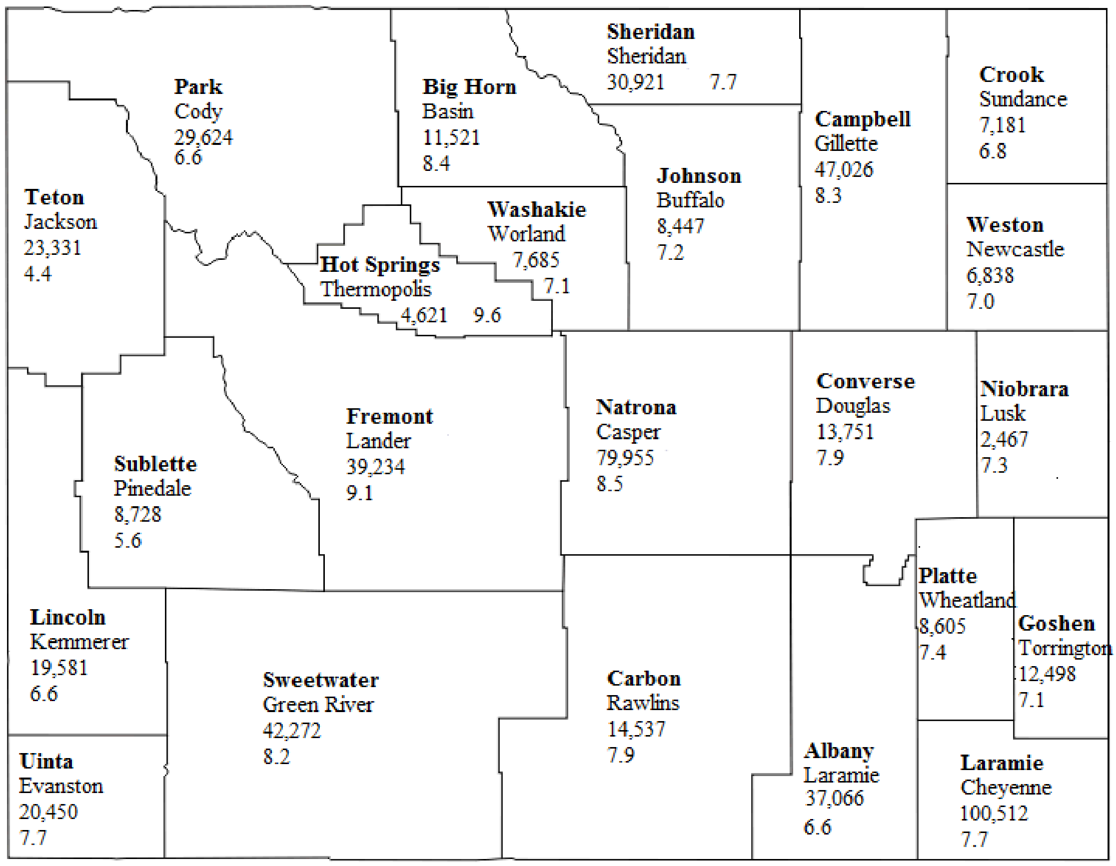

Fig. 2. Wyoming County Map

Table 1 describes, provides data sources and shows descriptive statistics for all variables examined. The selection of regressors closely follows established specifications in the literature (see Spencer et al 2018, Cosby et al 2019 and Gong et al 2019). The variable BEDS is intended as a proxy for a county's healthcare infrastructure. Mindful of the multicollinearity issues inherent in aggregate data, attention is paid to the orthogonality of covariates. A correlation matrix for the right-hand-side (RHS) is provided in Table 2. Moreover, variance inflation factors (VIF) are estimated for each regressor and shown in the far-right column of Table 2. Multicollinearity, as a matter of degree, does not appear problematic with this suite of regressors.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Table 1. Data Description, Sources and Descriptive Statistics

DEATHRATE. All-cause, jurisdiction of occurrence, age-standardized rate per 1,000 total county population. National Center for Health Statistics, National Vital Statistics System, 2010-2020. 2020 population data, including age group breakdowns, from the current census detail, U.S. Bureau of the Census. Mean 7.41, STD 1.42. Total County Deaths Mean 207, STD 205; Population Mean 25,143, STD 23,475; Crude Death Rate Mean 8.93, STD 2.77.
NONWHITE. Percent of the total county population classified as non-white. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 5.89, STD 4.38.
EDUCATION. Percent of the total county population with a bachelor's degree or higher. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 24.36, STD 9.12.
POVERTY. Percent of the total county population whose income is below the national poverty level. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 11.54, STD 4.24.
INCOME. County household median income in 1,000s of 2020 dollars. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 61.99, STD 11.96.
UNEMPLOYMENT. County annual unemployment rate. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of Labor Statistics, 2010-2020. Mean 4.41, STD 0.91.
BEDS. County licensed hospital bed capacity per capita. Wyoming County Profiles, State of Wyoming Economic Analysis Division, 2010-2020. Mean 2.69, STD 1.77.
UNINSURED. Percent of total county population with no health insurance coverage. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 13.52, STD 2.76.
65PLUS. Percent of total county population aged 65 and above. Wyoming County Profiles, State of Wyoming Economic Analysis Division and U.S. Bureau of the Census, 2010-2020. Mean 17.04 , STD 4.73.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Table 2. Correlation Matrix and Variance Inflation Factors

	NONWHITE	EDUC	POVERTY	INCOME	UNEMP	BEDS	UNINS	VIF
NONWHITE	1.00							**1.39**
EDUC	0.07	1.00						**1.12**
POVERTY	0.27	0.12	1.00					**2.01**
INCOME	-0.12	0.10	-0.64	1.00				**2.31**
UNEMP	0.23	-0.11	-0.18	0.14	1.00			**1.24**
BEDS	0.14	-0.13	0.33	-0.55	-0.17	1.00		**1.56**
UNINS	0.35	-0.12	-0.04	0.02	0.28	0.14	1.00	**1.27**
65PLUS	-0.11	-0.17	0.03	-0.61	-0.08	0.33	-0.08	**2.58**

The initial dependent variable examined is a county's annual age-standardized all-cause mortality rate per 1,000 population. Age-standardized (adjusted) rates (R') are commonly used in the ecological mortality literature to compare relative indexes across groups and over time. The National Center for Health Statistics (NCHS) computes the standardized rates by weighting all-cause death rates (R_j) as follows,

$$R' = \sum_j \frac{P_{sj}}{P_s} R_j \qquad (1)$$

where P_{sj} is the standard county population for age group j and P_s is the county total standard population (all ages combined). Age-standardized rates examined herein are based on the accepted 2000 U.S. standard population statistics (National Vital Statistics Reports 2017). Appendix A details the 2000 standard population weights for Laramie County as an example. Age categorized death counts form the numerator of R_j and are based on death certificate data received and coded by the NCHS. According to the NCHS, death certificate counts deliver the most complete and accurate picture of lives lost in the U.S. Death related data used for this paper was accessed from the NCHS on September 30, 2021.

3. Econometrics

The general model estimated becomes,

$$Y_{it} = \alpha + X_{it}\beta + \eta_{it} \quad i = 1, \ldots, N; \; t = 1, \ldots, T \qquad (2)$$

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

where Y_{it} is the dependent variable, α is a scalar intercept, X_{it} are observable socioeconomic variables that vary across counties i and over time t, β is a vector of estimated coefficients and η_{it} denotes the overall error term. The error term is comprised of three components,

$$\eta_{it} = \mu_i + \lambda_t + \nu_{it}, \tag{3}$$

where μ_i denotes the unobservable county specific effects, λ_t represents the unobservable year specific effects and ν_{it} is the remainder stochastic disturbance. The component μ_i is time-invariant and will account for county specific effects (characteristics) not included in the right-hand-side. Similarly, λ_t is county-invariant and will account for any time effects not included in the regression. Examples of what the component λ_t controls for herein are medical technological advancement and/or contagious disease outbreaks effecting all counties. The remainder disturbance ν_{it} varies with counties and time and denotes the usual error term.

Generally, two specifications of equation (2) are considered. Fixed effects treats μ_i and λ_t as fixed yet unknown constants differing across counties and over time. This specification is easily estimated by including county and year dummy variables in the right-hand-side (Least Squares LSDV estimator). If N and/or T are large, conserving precious degrees of freedom, estimates are generally obtained by transforming the data into deviations from respective group means ('within' estimator). Alternatively, random effects assumes that μ_i and λ_t are random, distributed independently across counties and over time. Estimates of this specification are based on transformations of the data into deviations from weighted respective group means where the weights are based on the variances of the components in equation (3), N and T (Feasible GLS estimator). The potential correlation of μ_i and λ_t with the variables in X_{it} is a primary consideration. If these correlations are present, random effects estimation yields biased and inconsistent estimates of β and the variances of μ_i , λ_t and ν_{it} . By transforming the data, into deviations from the basic group means, the fixed effects estimator is not impacted by this lack of orthogonality but is not fully efficient since it ignores variation across counties and perhaps time periods. The choice of estimator generally rests on statistical considerations and hypothesis testing. Hausman (1978) outlines a specification test of the null hypothesis of orthogonality between the latent effects and X_{it}.

The random effects specification requires exogeneity of all regressors and the components in equation (3). Conversely, the fixed effects model allows for endogeneity of all the regressors and μ_i , λ_t . In order to avoid this all or nothing choice of exogeneity, Hausman and Taylor (1981) (HT) propose a third specification for estimating equation (2) where the RHS is split into two sets of variables, those assumed uncorrelated (exogenous) with μ_i , λ_t and ν_{it}, and those

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

correlated (endogenous) with μ_i and λ_t , but not v_{it}. The uncorrelated set identified serves two functions, (i) using mean deviations, unbiased estimates of the respective elements of β are produced, and (ii) the exogenous set and group means provide valid instruments for the unbiased and efficient estimation of β. The two sets of variables need not rely solely on a priori assumptions, correlation hypothesis can be tested.

4. Results

One of the key motivations behind pooling a time series of cross-sections is to broaden the data base in order to obtain the best and most reliable parameter estimates. The general restriction is the pooled model with the same slope parameters across counties and over time. These structural tests share roots with Chow (1960) and test the null hypothesis of equal slope coefficients. Table 3 shows F-tests for pooling across counties and over time. Though not statistically imperious, both fail to reject the null at the < 5% level.

Table 3. Hypotheses tests

Pooling slopes across counties	F(176,69) = 1.40 (p 0.055)
Pooling slopes over time	F(80,165) = 1.34 (p 0.064)
Within Sample White Test	10.11 (p 0.182)
LM Test RE	94.58 (p 0.000)
Two-way county & year effects	
vs. pooled OLS	F(33,213) = 4.74 (p 0.000)
Pooled Durbin-Watson	1.74
Pesaran CD	1.02

Careful testing depicted in Table 3 fails to reject the null hypothesis of homoscedastic disturbances indicated by the pooled 'within' sample White statistic of 10.11. Two tests denoted confirm county and year heterogeneity and verify the importance of controlling for unobservable county and year effects. The Lagrange multiplier test statistic of 94.58 distributed χ^2 with 2 degrees of freedom rejects the null hypothesis, $H_0 : \sigma_\mu^2 = \sigma_\lambda^2 = 0$. Second, the test statistic F(33, 213) = 4.74 is sufficient to reject the null hypothesis of county and year homogeneity at the < 1% level. The Durbin-Watson statistic of 1.74 lies between the critical value limits of the test. This inconclusive region arises because the sequence of residuals is influenced by the movement of the covariates in the regression − not just serial correlation of the errors (Bartels and Goodhew 1981). Lastly, the Pesaran CD test statistic of 1.02 fails to reject the null hypothesis of cross-sectional (spatial) independence at the < 5% level (Pesaran 2020).

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Results from the pooled ordinary least squares (OLS) and two-way error components estimators are presented in Table 4. The pooled OLS estimates in the upper left column are easily challenged due to the sizable LM test statistic and F-test rejection of homogeneity shown in Table 3. The random effects estimates in the upper right column suffer from the endogeneity defect described above. The Hausman null, $H_0 : E(\mu_i, \lambda_t | X_{it}) = 0$, of latent effects orthogonality is rejected at the < 10% level (exact p value of 0.052). The random effects estimates lean toward misspecification. The unbiased yet less efficient two-way fixed effects estimates (lower left column) offer no support to the social causes hypothesis that aggregate socioeconomic factors matter in explaining the variation in within-state mortality rates. The inconspicuous explanatory power of the two-way fixed effects construct (R^2 = 0.63) rests solely on the μ_i and λ_t. County and year latent effects, captured by the fixed error components, hold the explanation to more than 60% of the variation in Wyoming county level age-standardized mortality rates. A note of interpretive caution, fixed effects estimation places great demands on the data. For example, μ_i capture any between county variation leaving only within county variation to be picked up by regressors.

Table 4. Pooled OLS and Two-Way Error Components estimates

Variable (t)	**Pooled OLS**	**Two-Way RE**
Constant (t)	8.71 (7.53)***	7.96 (5.78)***
NONWHITE (t)	0.08 (3.48)***	0.08 (2.65)***
EDUCATION (t)	-0.07 (-7.29)***	-0.07 (-5.07)***
POVERTY (t)	0.02 (0.82)	0.02 (0.62)
INCOME (t)	-0.01 (-0.92)	-0.003 (-0.19)
UNEMPLOYMENT (t)	0.23 (2.62)***	0.14 (1.48)
BEDS (t)	0.03 (0.55)	0.08 (0.97)
UNINSURED (t)	-0.05 (-1.51)	-0.02 (-0.50)
R^2	0.36	a
LM Test (p value)	-	94.58 (0.00)
Variable (t)	**Two-Way FE**	**Two-Way HT**
Constant (t)	2.11 (0.62)	8.12 (1.78)*
NONWHITE (t)	0.11 (0.54)	0.12 (1.34)
EDUCATION (t)	0.01 (0.30)	-0.06 (-0.81)
POVERTY (t)	0.04 (0.76)	0.03 (0.70)
INCOME (t)	0.04 (1.38)	0.005 (0.22)

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

UNEMPLOYMENT (t)	0.19 (1.61)	0.11 (0.92)
BEDS (t)	-0.12 (-0.37)	0.04 (0.25)
UNINSURED (t)	0.07 (1.40)	-0.13 (-0.65)
R^2	0.63	a
Hausman Test (p value)	13.97 (0.052)	2.54 (0.47)

[a] No precise counterpart to R^2 in these constructs. Pseudo calculations of R^2 are not 'fit' measures in the same sense as in ordinary least squares.
***, **, * significance at the 1%, 5%, 10% level. Observations 253.

If we are inclined to assume that certain variables in X_{it} are uncorrelated with the latent effects, Hausman and Taylor (1981) outlines an estimator that produces consistent *and* efficient estimates of β. In order to identify variable sets, iterative two-way random effects regressions were performed − varying 7 sets of variables by dropping the regressor indicated in the first column of Table 5 with the resulting Hausman statistic in the second column. For example, the fifth row depicts the resulting test statistic when the UNEMPLOYMENT variable is dropped from the right-hand-side. Note that the Hausman test statistic reduces to 6.97 from 13.97. The UNEMPLOYMENT variable appears to be a significant 'correlation contributor'. With 2020 data included in the sample, it is not surprising that the UNEMPLOYMENT variable suffers from endogeneity. Following this logic, set X_1 (uncorrelated) includes NONWHITE, POVERTY, INCOME and BEDS with set X_2 (correlated) containing EDUCATION, UNEMPLOYMENT and UNINSURED.

Table 5. Latent Effect Correlation tests*

	χ^2_6
NONWHITE	13.38
EDUCATION	8.43
POVERTY	12.09
INCOME	11.46
UNEMPLOYMENT	6.97
BEDS	11.70
UNINSURED	7.19

*All RHS variables $\chi^2_7 = 13.97$

With the variable sets identified, *LIMDEP*® Version 11 provides a Hausman-Taylor estimator for the one-way random effects model. In order to estimate a comparable two-way specification, T - 1 time dummies are included in variable set X_1 (see Wyhowski 1994). Results are presented in the lower right column of Table 4 above. Again, none of the observable covariates or

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

instruments (IVs) test significant at any conventional p level. County and year specific effects again provide the majority of the model's explanatory power. A Hausman test based on the difference between the HT and FE estimator yields, $\chi_3^2 = 2.54$, which fails to reject the null hypothesis of orthogonality. The chi-squared degrees of freedom equal to 3 indicates the variable sets are properly identified. The HT use of within model instruments appears appropriate and estimates test unbiased, consistent and are more efficient than the FE counterparts.

Table 6 shows year-specific effect estimates for 3 unbiased specifications − 'within' fixed effects, least squares dummy variable (LSDV) fixed effects and the Hausman-Taylor IV random effects model. In 2020, death certificate counts captured more public attention in Wyoming, arguably, than any mortality data in recent history.

Table 6. Year Specific Effect estimates

	'Within' Estimator		**LSDV Estimator**
	Year Effects (t)		**Year Effects (t)**
Scalar Constant	2.11 (0.62)		
2010	0.10 (0.43)	**Albany 2010**	1.20 (1.77)*
2011	-0.14 (-0.65)	**2011**	-0.24 (-0.87)
2012	-0.04 (-0.18)	**2012**	-0.14 (-0.48)
2013	-0.29 (-1.66)*	**2013**	-0.39 (-1.65)*
2014	0.08 (0.42)	**2014**	-0.01 (-0.05)
2015	0.07 (0.34)	**2015**	-0.03 (-0.10)
2016	-0.40 (-1.98)**	**2016**	-0.50 (-1.66)*
2017	-0.15 (-0.76)	**2017**	-0.25 (-0.74)
2018	0.08 (0.39)	**2018**	-0.01 (-0.04)
2019	0.11 (1.47)	**2019**	0.01 (0.03)
2020	0.58 (1.53)	**2020**	0.48 (1.22)
	Hausman-Taylor		
	Year Effects (t)		
Scalar Constant 2010	8.12 (1.78)*		
2011	-0.27 (-0.96)		
2012	-0.19 (-0.68)		
2013	-0.49 (-1.72)*		

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

2014	-0.13 (-0.44)		
2015	-0.29 (-0.75)		
2016	-0.78 (-1.86)*		
2017	-0.73 (-1.35)		
2018	-0.56 (-0.93)		
2019	0.09 (0.84)		
2020	0.43 (1.36)		

***, **, * significance at the 1%, 5%, 10% level. Observations 253.

Interestingly, after controlling for age, population, observable and unobservable Wyoming county characteristics and all county-invariant year effects, the 2020 specific effect is statistically insignificant at conventional probability levels for all specifications. It appears the heightened public attention within the state was centered on the anecdotal.

A review of an earlier version of this paper suggests that the 20 year old age adjustment imposed on the dependent variable may be masking a vital and timely explanatory factor of mortality variation. The reviewer cites a recent population turning point analysis published by the U.S. Census Bureau (see Vespa et al 2020). This Current Population Report focuses on key demographic changes expected into the year 2060. Most notably, the percent of the U.S. population aged 65 and above is expanding and is expected to be roughly one forth by 2060. Moreover, the report projects that by 2034, those aged 65 and above will outnumber children (those aged 18 and below) for the first time in U.S. history. With the population rapidly aging and deaths expected to rise substantially, it is projected that immigration will become the primary source of future population growth – overtaking natural (births over deaths) increases.

Pivoting this national analysis to Wyoming, Table A1 in Appendix A shows that Laramie county has experienced a 48% increase in the aged 65 and above population demographic from the 2000 NCHS standard which roughly mirrors the overall increase for the state. In response to the review, the entire empirical method outlined above was recalculated with a new dependent variable, all-cause crude mortality rates (deaths/(population/1,000); and one additional covariate, the percent of a county's total population aged 65 and above (65PLUS). Table 1 above provides descriptive statistics for the new variables. Table 7 depicts the alternative two-way error component results.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Table 7. Alternative Two-Way Error Component results

Variable (t)	Two-Way RE	Two-Way FE	Two-Way HT
Constant (t)	7.88 (3.19)***	5.72 (1.14)	3.42 (1.76)*
NONWHITE (t)	0.08 (2.03)**	-0.005 (-0.02)	0.04 (0.26)
EDUCATION (t)	-0.07 (-4.10)***	0.004 (0.09)	0.06 (0.46)
POVERTY (t)	-0.06 (-1.46)	-0.001 (-0.02)	-0.04 (-0.81)
INCOME (t)	-0.05 (-2.36)**	0.01 (0.37)	0.007 (0.18)
UNEMPLOYMENT (t)	0.03 (0.25)	0.16 (1.26)	0.13 (0.94)
BEDS (t)	0.06 (0.57)	-0.52 (-1.47)	0.06 (0.23)
UNINSURED (t)	-0.005 (-0.13)	0.03 (0.71)	0.22 (0.63)
65PLUS (t)	0.35 (7.86)***	0.26 (1.69)*	0.41 (2.60)***
2020 EFFECT	-	0.20 (0.41)	0.40 (0.59)
R^2	a	0.89	a
LM Test (p value)	174.54 (0.00)		
F(33,212) (p value)		6.52 (0.00)	
Hausman Test (p value)	19.86 (0.01)		3.97 (0.26)

[a] No precise counterpart to R^2 in these constructs.

***, **, * significance at the 1%, 5%, 10% level. Observations 253.

The Hausman test statistic of 19.86 provides strong evidence of misspecification in the random effects construct. The inefficient fixed effects results shown explain 89% of the variation in crude mortality rates with the new covariate 65PLUS significant at the < 10% level. The inconspicuous explanatory power of the two-way fixed effects construct continues to rest primarily on the μ_i and λ_t. The Hausman-Taylor IV random effects estimates test unbiased and efficient given $\chi_3^2 = 3.97$. The additional variable 65PLUS is significant at the < 1% level. The coefficient estimate of 0.41 indicates that, for the average county, a 1% increase in the aged 65 and above population share increases deaths by roughly 10 per year. Recall that mean deaths for a Wyoming county is 207 annually (see Table 1 above). None of the county level socioeconomic or structural variables are significant at conventional levels. The 2020 year specific effect is again insignificant at conventional p levels. The reviewer's concern that the mistimed age-adjustment to the left-hand-side was masking one crucial explanation for the variation in spatial mortality appears warranted and is supported herein.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

5. Conclusion

Within the least populated state in the union, Wyoming, the contention that observable socioeconomic factors matter in explaining variation in mortality is not supported. Evidence from an all county panel spanning over 11 years suggests that differences in all-cause mortality rates can be explained by the asymmetric and growing share of the population aged 65 and above and certain latent county and time-specific effects. The difficult task of disentangling these latent effects provides a compelling future research avenue. Moreover, results herein bolster those found in Gong et al (2019) where Wyoming was one of three states established as exceptions to the authors' broader analyses and results.

There is a vast literature critical of the use of aggregate data to explain heterogeneous individual occurrence (see Stroker 1993 and Holderness 2016 for reviews). Statistical properties and the biases introduced by using aggregated or averaged data have yet to be adequately explained. Aggregation defects likely plague the ecological literature cited in the introduction section above and the examination herein. Practical implications of this paper call for a shift in focus to a smaller unit of analysis. Preference is given to individual-level specifications controlling for socioeconomic factors and individual specific effects. For example, a recent paper by Zhao et al (2020) exploits the Chinese Longitudinal Health Longevity Survey from 2002 to 2014 for older Chinese adults aged 65 and above. The survey sample tracks 28,235 individuals over time and is used to examine the urban-rural disparity in mortality in China. Findings suggest that mortality disparity among older adults in China can be linked to differences in *individual* socioeconomic resources. Locating or building a comparable panel for the U.S. may prove to be a challenging but worthwhile effort.

References

Bartels, R. and Goodhew, J. (1981) The robustness of the Durbin-Watson test. Review of Economics and Statistics 63, pp. 136-139.

Chow, G. (1960) Tests of equality between sets of coefficients in two linear regressions. Econometrica 28, pp. 591-605.

Cosby, A., McDoom-Echebiri, M., James, W., Khandekar, H., Brown, W. and Hanna, H. (2019) Growth and persistence of placed-based mortality in the United States: The rural mortality penalty. American Journal of Public Health 109(1), pp. 155-162.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Cosby, A., Neaves, T., Cossman, R., Cossman, J., James, W. and Feierabend, N. (2008) Preliminary evidence for an emerging nonmetropolitan mortality penalty in the United States. American Journal of Public Health 98(8), pp. 1470–1472.

Gong, G., Phillips, S., Hudson, C., Curti, D. and Phillips, B. (2019) Higher U.S. rural mortality rates linked to socioeconomic status, physician shortages, and lack of health insurance. Health Affairs 38(12), pp. 2003-2010.

Hausman, J. (1978) Specification tests in econometrics. Econometrica 46(6), pp. 1251-1271.

Hausman, J. and Taylor, W. (1981) Panel data and unobservable individual effects. Econometrica 49(6), pp. 1377-1398.

Holderness, C. (2016) Problems using aggregate data to infer individual behavior: Evidence from law, finance, and ownership concentration. Critical Finance Review 5(1), pp. 1-40.

James, W. (2014) All rural places are not created equal: Revisiting the rural mortality penalty in the U.S. American Journal of Public Health 104(11), pp. 2122-2129.

Kunce, M. (2021) The impact of socioeconomic factors on state suicide rates: Revisited. Journal of Statistical and Econometric Methods 10(4), pp. 1-14.

National Vital Statistics Reports. (2017) Volume 66, Number 6.

Pesaran, M. (2020) General diagnostic tests of cross-sectional dependence in panels. Empirical Economics 60, pp. 13-50.

Spencer, J., Wheeler, S., Rotter, J. and Holmes, G. (2018) Decomposing mortality disparities in urban and rural U.S. counties. Health Services Research 53(6), pp. 4310-4331.

Stroker, T. (1993) Empirical approaches to the problem of aggregation over individuals. Journal of Economic Literature 31, pp. 1827-1874.

Vespa, J., Medina, L., and Armstrong, D. (2020) Demographic turning points for the United States. P25-1144. U.S. Census Bureau, Washington, D.C.

Wyhowski, D. (1994) Estimation of a panel data model in the presence of correlation between regressors and a two-way error component. Econometric Theory 10, pp. 130 - 139.

ISSN: 2455-8834

Volume:06, Issue:10 "October 2021"

Wyoming County Profiles. Department of Administration and Information, Economic Analysis Division. State of Wyoming. http://eadiv.state.wy.us/ for examples.

Zhao, Y., Xu, X., Dupre, M., Xie, Q., Qiu, L. and Gu, D. (2020) Individual-level factors attributable to urban-rural disparity in mortality among older adults in China. BMC Public Health 20:1472, pp. 1-11.

Appendix A

Table A1. Laramie County 2000 Age Population Weights

Age Category	Population	Weight		
< 1 year	1,078	1.32%		
1-4 years	4,306	5.28%		
5-14 years	11,983	14.68%		
15-24 years	11,460	14.04%		
25-34 years	11,617	14.24%		
35-44 years	13,277	16.27%		
45-54 years	11,489	14.08%		
55-64 years	7,046	8.63%		
65-74 years	5,013	6.14%		
75-84 years	3,223	3.95%		
85+ years	1,115	1.37%	65 +	11.46%
	81,607			
		2020 Census	65 +	16.96%