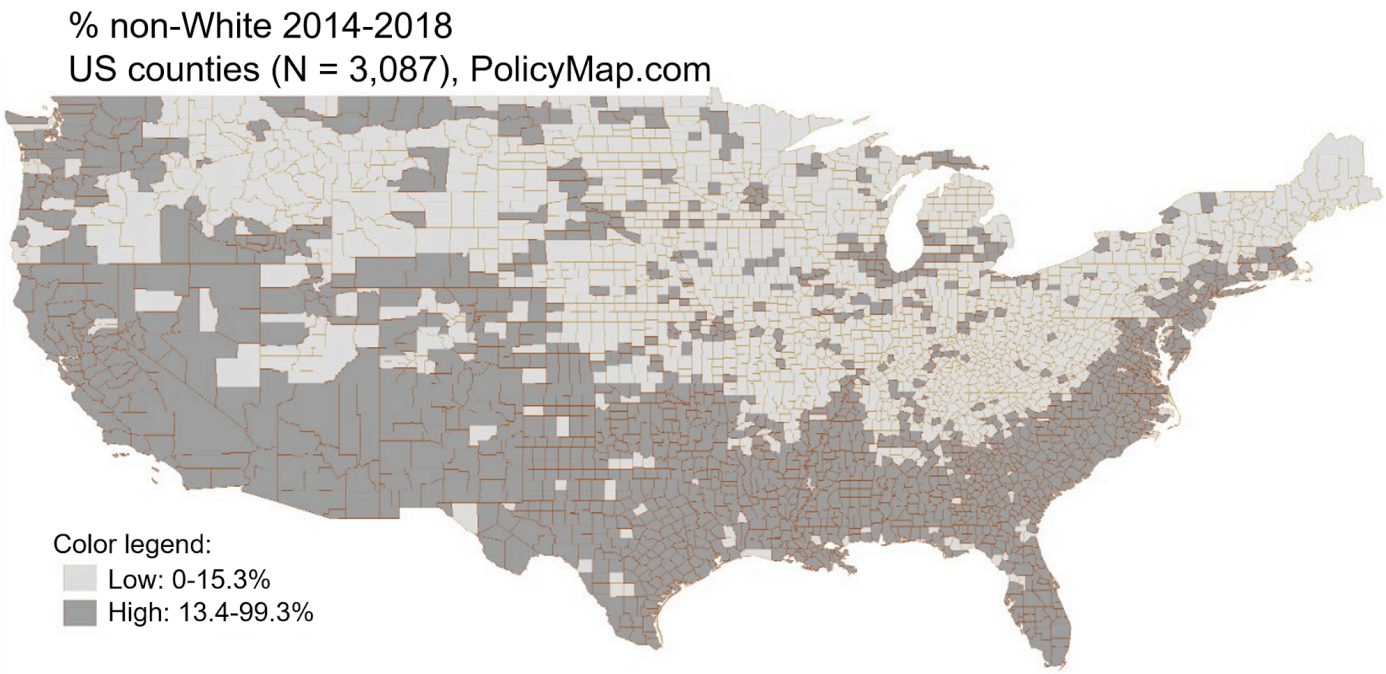
Appendix accompanying the paper "Life Expectancy Disparities between White and non-White U.S. residents"

Content:

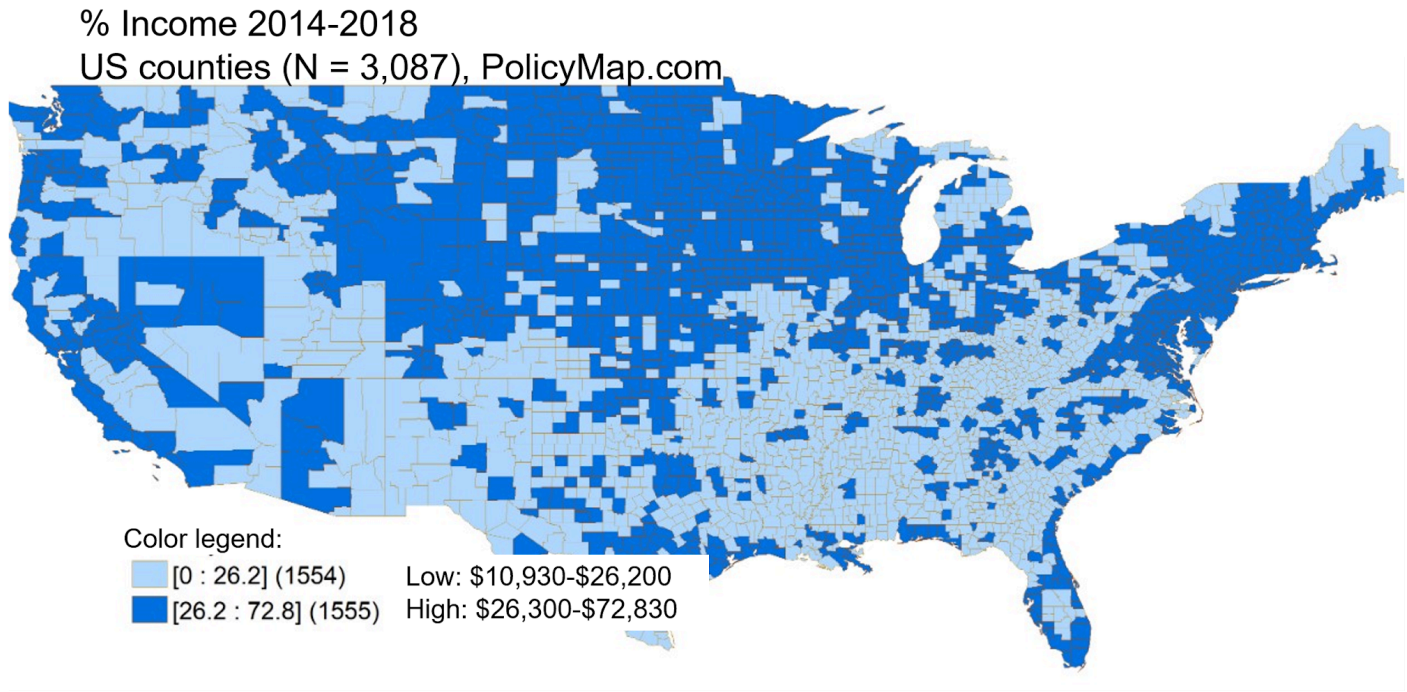
1. Maps of % non-White (10%s), income ($1,000s), and Life Expectancy (years), all dichotomized to better indicate potential relations.
2. Technical notes and quotes
3. Naïve/a-spatial and spatial lag regressions - Stata
4. Anova-type derivation of ICCs for the main variables
5. Naïve Pearson correlations between main variables across three regional levels
6. Geographically Weighted Regression GWR
7. Indirect effects lfex1015stct <- inckstcy pnwhtstct
   a. Heatmaps
   b. Stata gsem analysis
8. Regression coefficients for the % non-White to Life Expectancy regional variables, at census tract, county and state levels, from naïve/a-spatial and proper/spatial models
9. ICCs for (census tracts)counties clustering

1. Maps of % non-White 2014-2018, income 2014-2018 ($1,000s), and Life Expectancy 2010-2015 (years), all dichotomized to better indicate potential relations

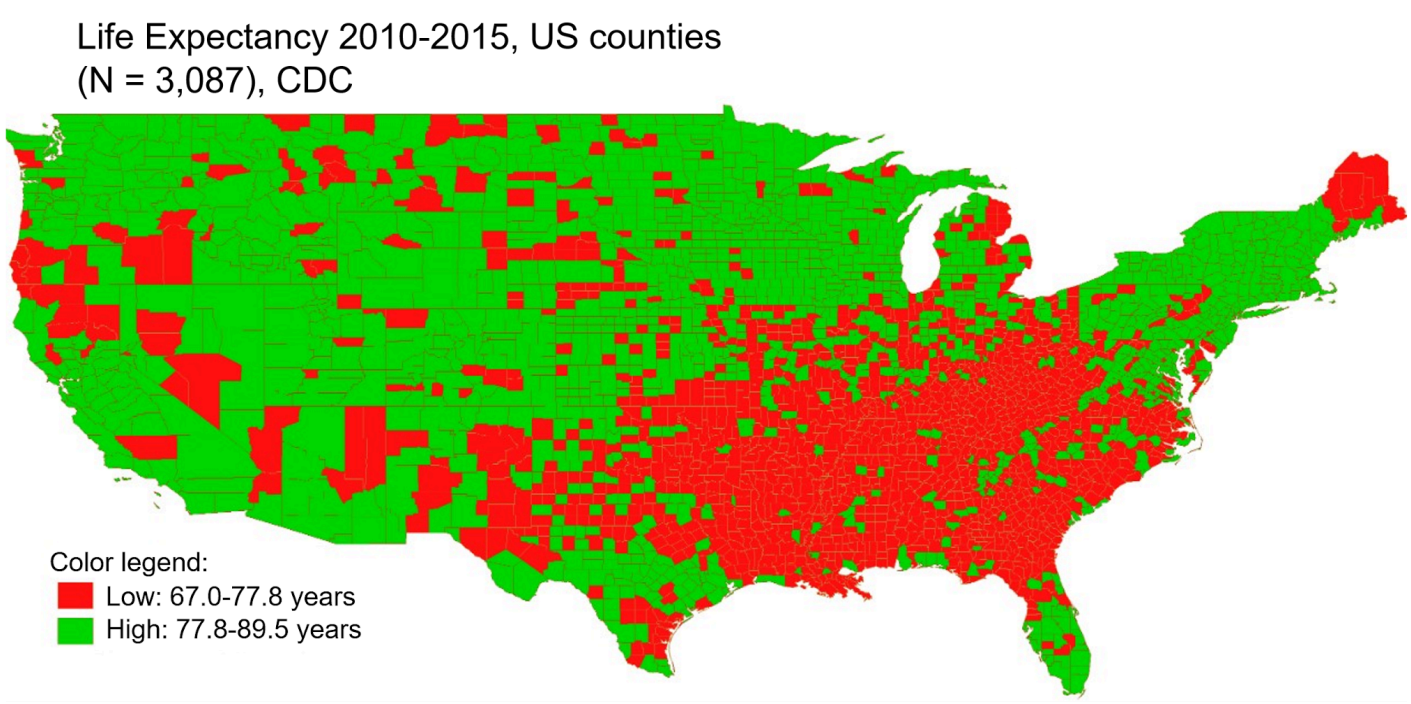
1. % non-White 2014-2018



2. Income ($1,000s) 2014-2018



3. Life Expectancy (years) 2010-2015



GO BACK UP

2. Technical notes and quotes

0.1. On nonindependence and 'auto'-correlation:

0.1.i "Note that in the binomial distribution the trials are assumed to be independent and, therefore, the X_i are independent. This assumption is not needed in order to find the mean." (Conover, 1999)p. 40. Calculating the mean when individual region values are not independent *would* however require one to adjust the p values in the E(Y) = n·p formula, to allow them to vary across regions, to become e.g. larger than the nominal 1/n for regions with similar values: the probability increases once we know/see the value of a neighboring region.

0.1.ii. "The coefficient of intraclass correlation [R] is the measure of the relative homogeneity of the scores within the classes in relation to the total variation among all the scores in the table. Thus maximal positive correlation exists when all the intraclass scores are identical and the scores differ only from class to class. As the relative heterogeneity of the intraclass scores increases, the computed value of R will decrease; maximal negative correlation exists when the heterogeneity of the intraclass scores is maximal and all the class means are the same." (Haggard, 1958), p. 5-6.

0.1.iii. There are several methods for analyzing of spatial (or georeferenced (Vallejos et al., 2020)) data (a taxonomy is in (Baltagi et al., 2016); a simple spatial correction model is the Cliff and Ord (1973 (Cliff & Ord, 1973)) 'spatial autoregressive' model (which is informed by Whittle's two-dimensional linear autoregression (Whittle, 1954), see (Kelejian & Prucha, 2004). Moreover, the spatial dimension can be utilized as an ordering criterion either in a sociometric spatial manner, i.e. a region connected to its natural adjacent neighboring regions, or continuously, by marking a region's location as latitude and longitude (a third dimension would rarely be needed analytically); geographically weighted regressions is a common analytical solution in such contexts (Brunsdon et al., 1996).

0.2. On adjusting in linear regression

0.2.i. "The slope in the multiple regression differs from that in the simple regression by including a correction that accounts for the influence of the additional variable […] The general result for a three-variable regression in which x_1 is a constant term is $b_{y2\cdot3} = (b_{y2} - b_{y3}b_{32})(1 - r^2_{23})$ (3-11)" (Greene, 2003) p. 23 – x_1 is the 'first predictor, a vector or column of 1's, such that b_{y1} is the intercept. This bias formula' can be derived visually in a few steps using Sewall Wright's tracing rule. We show in Appendix a table of correlations between the main three variables and the lag of life expectancy, to illustrate the procedure. The census tract naïve correlation between % Minority and life expectancy is $b_{\text{LifeExp \%Min}} = -.230$, whereas $b_{\text{LifeExp \%Min.LagLifeExp}} = -.178$, while county level ones are $b'_{\text{LifeExp \%Min}} = -.165$, whereas $b'_{\text{LifeExp \%Min.LagLifeExp}} = -.145$.

0.2.ii. "$P(y \mid x^{hat}) = \Sigma_z P(y|x,z)\cdot P(z)$ (3.19) The summation in (3.19) represents the standard formula obtained under adjustment for Z; " p. 80 & "The phrases "hold F fixed" and "control for F," used by both philosophers (e.g., Eells 1991 (Eells & Sober, 1983)) and statisticians (e.g., Pratt and Schlaifer 1988 (Pratt & Schlaifer, 1988)), connote external interventions and may therefore be misleading. In statistical analysis, all one can do is simulate "holding F fixed" by considering cases with equal values of F – that is, "conditioning" on F and – an operation that I will call "adjusting for F." (Pearl, 2009) p. 175

0.3. On adjusting for spatial lag

A3.i. "The magnitude and significance of the spatial lag coefficient is conventionally interpreted as evidence for contagion or diffusion of the dependent variable, in this case heart disease decline. While heart disease itself is not 'contagious', the norms and behaviors (e.g. smoking, dietary habits, physical activity) that drive it could be (Ball, Jeffery, Abbott, McNaughton, & Crawford, 2010; Luke & Stamatakis, 2012). Beyond the individual-level diffusion of behavior, it is also possible that the spatial scale of social, economic, and health service contexts of local areas are not aligned with county boundaries, producing an apparent spillover effect across adjacent counties (Voss et al., 2006)." (Kramer et al., 2017)

Technical notes references:

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., et al. (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA,* 315, 1750-1766.

Conover, W.J. (1999). *Practical nonparametric statistics*. New York: Willey & Sons.

Eells, E., & Sober, E. (1983). Probabilistic causality and the question of transitivity. *Philosophy of Science,* 50, 35-57.

Freedman, D.A. (2005). Linear Statistical Models for Causation: A Critical Review https://sci-hub.st/https://onlinelibrary.wiley.com/doi/abs/10.1002/0470013192.bsa598. *Encyclopedia of statistics in behavioral science*.

Greene, W.H. (2003). *Econometric analysis* *https://www.ctanujit.org/uploads/2/5/3/9/25393293/_econometric_analysis_by_greence.pdf*: Pearson Education India.

Haggard, E.A. (1958). Intraclass correlation and the analysis of variance https://drive.google.com/file/d/15sqL7oOhYtLar-iUScwg6r7PG0sB2l8_/view?usp=share_link.

Jha, P., Ramasundarahettige, C., Landsman, V., Rostron, B., Thun, M., Anderson, R.N., et al. (2013). 21st-century hazards of smoking and benefits of cessation in the United States. *New England Journal of Medicine,* 368, 341-350.

Kenny, D.A. (1988). The analysis of data from two-person relationships. In S. Duck, D.F. Hay, S.E. Hobfoll, W. Ickes, & B.M. Montgomery (Eds.), *Handbook of personal relationships: Theory, research and interventions* pp. 57-77). Oxford, England: John Wiley & Sons.

Kramer, M.R., Black, N.C., Matthews, S.A., & James, S.A. (2017). The legacy of slavery and contemporary declines in heart disease mortality in the US South. *SSM-population health,* 3, 609-617.

Modig, K., Rau, R., & Ahlbom, A. (2020). Life expectancy: what does it measure? *BMJ Open,* 10, e035932.

Pearl, J. (2009). *Causality: models, reasoning, and inference* *https://drive.google.com/file/d/0BzBnj233_MlwZFV3dGFXZHJTZTg/view?usp=sharing&resourcekey=0-mJQJR4iK2_X5Hi5dBx4acw*: Cambridge university press.

Pratt, J.W., & Schlaifer, R. (1988). On the interpretation and observation of laws. *Journal of Econometrics,* 39, 23-52.

Wilkinson, R.G. (1992). Income distribution and life expectancy. *BMJ: British Medical Journal,* 304, 165.

Yule, G.U. (1899). An investigation into the causes of changes in pauperism in England, chiefly during the last two intercensal decades (Part I) https://sci-hub.st/https://www.jstor.org/stable/2979889. *Journal of the Royal Statistical Society,* 62, 249-295.

GO BACK UP

3. Naïve/a-spatial and spatial lag regressions - Stata

```
. sem lfex1015stct <- pnwhtstct  // state lglfex
Structural equation model                               Number of obs = 49
Estimation method: ml
Log likelihood = -290.39913
------------------------------------------------------------------------------
                 |                 OIM
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015stct   |
        pnwhtstct |  -.0224077    .014918    -1.50   0.133    -.0516464     .006831
            _cons |   79.30008   .4833988   164.05   0.000     78.35264    80.24753
-----------------+------------------------------------------------------------
var(e.lfex1015stct)|   2.291846   .4630228                      1.54247     3.40529
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                  Prob > chi2 = .

. sem lfex1015stct <- pnwhtstct lglfex // state
Structural equation model                               Number of obs = 49
Estimation method: ml
Log likelihood = -346.75214
------------------------------------------------------------------------------
                 |                 OIM
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015stct   |
        pnwhtstct |   .0107523   .0105282     1.02   0.307    -.0098827    .0313873
           lglfex |   .9789099   .1199531     8.16   0.000     .7438062    1.214014
            _cons |   1.342231   9.557936     0.14   0.888    -17.39098    20.07544
-----------------+------------------------------------------------------------
```

```
var(e.lfex1015stct)|   .9714721    .196267                      .6538252    1.443441
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                Prob > chi2 = .
```

GO BACK UP

4. Anova-type derivation of ICCs for the main variables

The data needs to be turned vertical/long, matched by state, with each state's neighbors on successive rows, with an n of 218=267-49 like:

state2dig	cntslfneighb	tot_neighslf	neighbor	st2dothers	inc1519k1	lf20exp1	inc1519k2	lf20exp2
CT	1	4	CT		78.444	78.4	78.444	78.4
CT	2	4	NY	CT	78.444	78.4	68.486	77.7
CT	3	4	MA	CT	78.444	78.4	81.215	79
CT	4	4	RI	CT	78.444	78.4	67.167	78.2

```
1.
gen pmin1519dv= pmin1519_2-pmin1519_1
lab var pmin1519dv  "neighbor-self pminrty15_19st deviations"
loneway  pmin1519dv st2dothers //pminrty15_19st we need to EXCLUDE SELF however to match the
Moran's ICC logic  n=
> 218=267-49

 One-way analysis of variance for pmin1519dv: neighbor-self pminrty15_19st dev

                                        Number of obs =       218
                                            R-squared =    0.4726

    Source                  SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between st2dothers     24627.093      48    513.06443      3.16     0.0000
Within st2dothers      27482.488     169    162.61827
------------------------------------------------------------------------
Total                   52109.58     217    240.13632

         Intraclass       Asy.
         correlation      S.E.       [95% conf. interval]
         -------------------------------------------------
            0.32693     0.07626       0.17746     0.47639

         Estimated SD of st2dothers effect       8.887485
         Estimated SD within st2dothers          12.75219
         Est. reliability of a st2dothers mean    0.68305
              (evaluated at n=4.44)
2.
 loneway   lf20expdv st2dothers  //lf20exp2

  One-way analysis of variance for lf20expdv: neighbor-self lglfex20 deviation

                                        Number of obs =       218
                                            R-squared =    0.5244

    Source                  SS         df      MS            F     Prob > F
------------------------------------------------------------------------
Between st2dothers      400.4189      48    8.3420605      3.88     0.0000
Within st2dothers      363.12072     169    2.1486433
------------------------------------------------------------------------
Total                  763.53962     217    3.5186158

         Intraclass       Asy.
         correlation      S.E.       [95% conf. interval]
         -------------------------------------------------
            0.39382     0.07643       0.24403     0.54361

         Estimated SD of st2dothers effect       1.181499
         Estimated SD within st2dothers          1.465825
```

```
         Est. reliability of a st2dothers mean    0.74243
              (evaluated at n=4.44)
3.
 loneway  inc1519kdv st2dothers //inc1519k2

 One-way analysis of variance for inc1519kdv: neighbor-self inck1519st deviati

                                          Number of obs =          218
                                              R-squared =       0.5556

    Source                  SS         df       MS            F     Prob > F
------------------------------------------------------------------------------
Between st2dothers       15373.95      48     320.29062       4.40      0.0000
Within st2dothers       12296.316     169     72.759269
------------------------------------------------------------------------------
Total                   27670.266     217     127.51275

         Intraclass       Asy.
         correlation      S.E.       [95% conf. interval]
         ---------------------------------------------
            0.43400     0.07571       0.28561     0.58240

         Estimated SD of st2dothers effect       7.469359
         Estimated SD within st2dothers          8.529904
         Est. reliability of a st2dothers mean    0.77283
              (evaluated at n=4.44)
```

GO BACK UP

5. Naïve Pearson correlations between main variables across three regional levels

	LfExp	LagInc	Inc	LagPov	Pov	LagUnp	Unp	LagNHs	NHs
LfExp.Stat	**0.766**								
LfExp.Cnty	**0.351**								
LfExp.CsTr	**0.452**								
LagInc.Stat	**0.780**	**0.502**							
LagInc.Cnty	**0.301**	**0.508**							
LagInc.CsTr	**0.419**	**0.527**							
Inc.Stat	**0.496**	**0.655**	**0.526**						
Inc.Cnty	**0.253**	**0.565**	**0.737**						
Inc.CsTr	**0.320**	**0.569**	**0.830**						
LagPov.Stat	**-0.893**	**-0.600**	**-0.864**	**-0.428**					
LagPov.Cnty	**-0.323**	**-0.546**	**-0.790**	**-0.588**					
LagPov.CsTr	**-0.472**	**-0.562**	**-0.647**	**-0.543**					
Pov.Stat	**-0.615**	**-0.786**	**-0.420**	**-0.722**	**0.522**				
Pov.Cnty	**-0.259**	**-0.578**	**-0.573**	**-0.711**	**0.690**				
Pov.CsTr	**-0.344**	**-0.599**	**-0.501**	**-0.595**	**0.717**				
LagUnp.Stat	*-0.087*	*-0.028*	*0.162*	*0.246*	*0.197*	*0.027*			
LagUnp.Cnty	**-0.182**	**-0.465**	**-0.514**	**-0.373**	**0.765**	**0.552**			
LagUnp.CsTr	**-0.383**	**-0.456**	**-0.477**	**-0.397**	**0.701**	**0.524**			
Unp.Stat	*-0.065*	*-0.152*	*0.193*	*0.131*	*0.025*	**0.298**	**0.398**		
Unp.Cnty	**-0.151**	**-0.454**	**-0.356**	**-0.463**	**0.525**	**0.650**	**0.637**		
Unp.CsTr	**-0.261**	**-0.440**	**-0.345**	**-0.401**	**0.492**	**0.580**	**0.586**		
LagNHs.Stat	**-0.840**	**-0.688**	**-0.605**	**-0.397**	**0.873**	**0.656**	**0.352**	**0.288**	
LagNHs.Cnty	**-0.259**	**-0.468**	**-0.648**	**-0.475**	**0.752**	**0.559**	**0.604**	**0.432**	
LagNHs.CsTr	**-0.293**	**-0.369**	**-0.623**	**-0.532**	**0.664**	**0.510**	**0.491**	**0.355**	
NHs.Stat	**-0.718**	**-0.760**	**-0.495**	**-0.619**	**0.705**	**0.847**	**0.296**	**0.322**	**0.827**
NHs.Cnty	**-0.212**	**-0.491**	**-0.494**	**-0.625**	**0.583**	**0.629**	**0.466**	**0.460**	**0.746**
NHs.CsTr	**-0.230**	**-0.419**	**-0.528**	**-0.586**	**0.554**	**0.603**	**0.411**	**0.399**	**0.819**

Bold: p < .05; italics p > .05; Another State %Unemp. Variable, built off the census tract values, showed significant correlations at state level; the models afterwards were not converging however.
GO BACK UP

6. Geographically Weighted Regression GWR

```
. gwr lfex1015ct pnwhtcstr, east(longeastw) north(latnorths) test
```

NOT different than naïve/a-spatial model

```
Global Model
      Source |       SS           df       MS      Number of obs   =    60,943
-------------+----------------------------------   F(1, 60941)     =   3241.86
       Model |   47129.369         1   47129.369   Prob > F        =    0.0000
    Residual |    885945.3    60,941  14.5377546   R-squared       =    0.0505
-------------+----------------------------------   Adj R-squared   =    0.0505
       Total |  933074.669    60,942  15.3108639   Root MSE        =    3.8128
------------------------------------------------------------------------------
  lfex1015ct | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   pnwhtcstr |  -.0304475   .0005348   -56.94   0.000    -.0314956   -.0293994
       _cons |   79.49284   .0244332  3253.48   0.000     79.44496    79.54073
------------------------------------------------------------------------------
no observations
r(2000);
```

```
. sem (lgctlfex pnwhtcstr -> lfex1015ct)  if contig49 == 1   // census tracts     spatial
(4540 observations with missing values excluded)
```

```
Endogenous variables
  Observed: lfex1015ct

Exogenous variables
  Observed: lgctlfex pnwhtcstr

Fitting target model:
Iteration 0:  Log likelihood = -628590.55
Iteration 1:  Log likelihood = -628590.55

Structural equation model                               Number of obs = 60,602
Estimation method: ml

Log likelihood = -628590.55

------------------------------------------------------------------------------
                 |                 OIM
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015ct     |
       lgctlfex |      0.361      0.003   118.57   0.000        0.355       0.367
      pnwhtcstr |     -0.022      0.000   -45.83   0.000       -0.023      -0.021
          _cons |     50.928      0.242   210.45   0.000       50.454      51.403
-----------------+------------------------------------------------------------
var(e.lfex1015ct)|     11.760      0.068                       11.629      11.894
------------------------------------------------------------------------------
```

```
. sem (pnwhtcstr -> lfex1015ct) if contig49 == 1   // census tracts       naïve/a-spatial
(4533 observations with missing values excluded)
Endogenous variables
  Observed: lfex1015ct
Exogenous variables
  Observed: pnwhtcstr
Fitting target model:
Iteration 0:  Log likelihood = -456751.91
Iteration 1:  Log likelihood = -456751.91
Structural equation model                               Number of obs = 60,609
Estimation method: ml
Log likelihood = -456751.91
------------------------------------------------------------------------------
                 |                 OIM
```

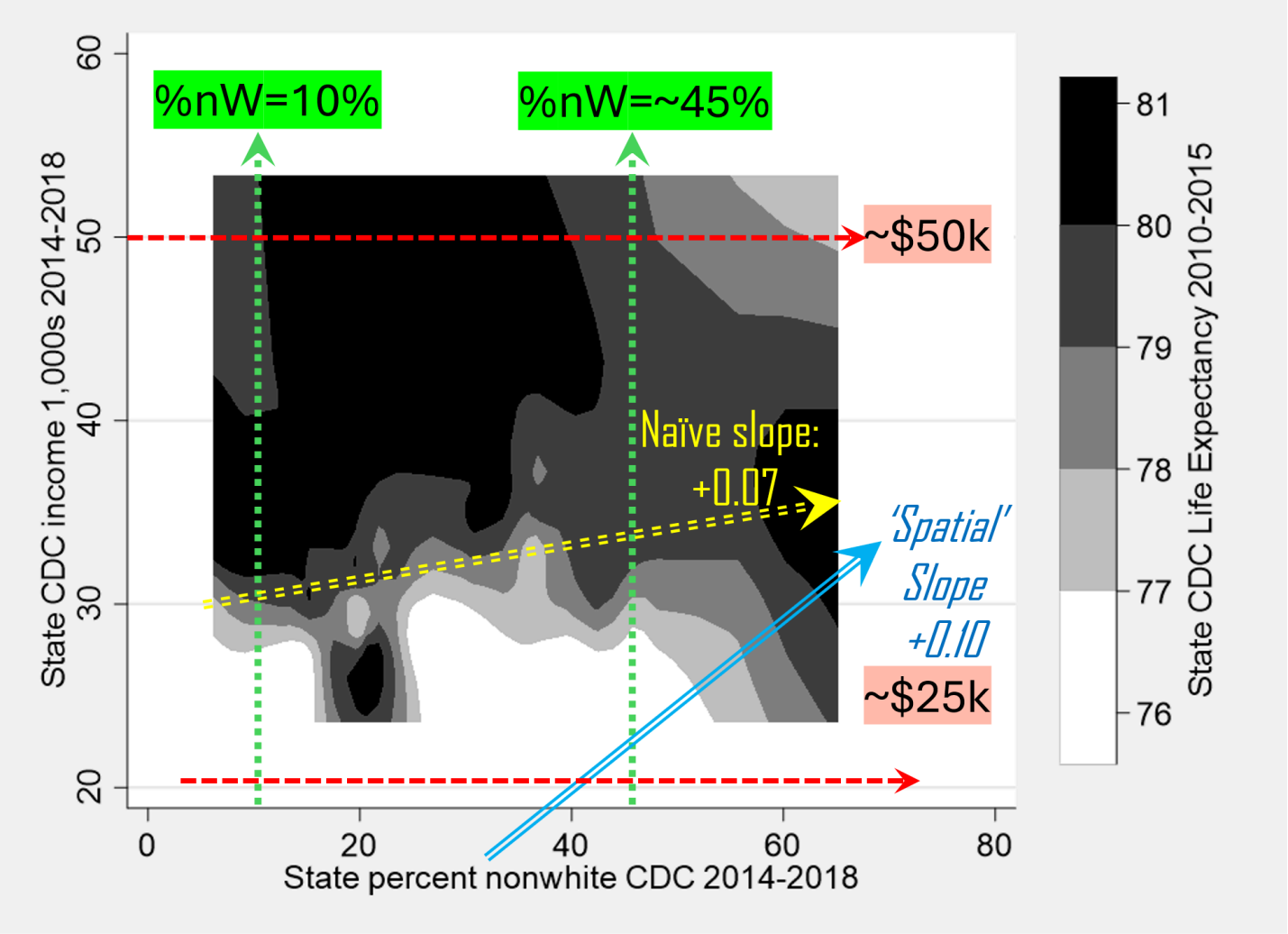
```
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------------
Structural       |
  lfex1015ct     |
       pnwhtcstr |     -0.031       0.001   -58.23   0.000       -0.032      -0.030
           _cons |     79.502       0.024  3255.02   0.000       79.455      79.550
-----------------+------------------------------------------------------------------
var(e.lfex1015ct)|     14.488       0.083                       14.325      14.652
------------------------------------------------------------------------------
```

GO BACK UP

7. Indirect effects lfex1015stct <- inckstcy pnwhtstct

a. Heatmaps

Because of sample size of counties and census tracts, only heatmaps for state-level spatial variables are shown, for illustration.




```
twoway contour lfex1015stct inckstcy pnwhtstct   , crule(linear) scolor(white) ecolor(black)
scheme(s2mono) // Y = lfex1015stct M = inckstcy X = pnwhtstct
```

The 'contour graph' provides a way to examine the entire data patterns of Y(M(X)) relations, nonlinearly too. Here, we examine LifeExp.(Income (%nonWhite)) or
State CDC Life Expectancy 2010-2015 (State CDC income 1,000s 2014-2018 (State percent nonwhite CDC 2014-2018))

X->Y:

- At **low** (fixed!) levels of M=Income, low values ~= &25K, with increasing %nonWhite we see an increase in LifeExp. up to (some) ~20%nW, then a decrease to ~27%nW, stability until ~45%nW, then an increase from some ~55%nW to ~65%nW

- At **higher** (fixed!) levels of M=Income, low values ~= &50K, with increasing %nonWhite we see an increase in LifeExp. up to (some) ~35%nW, then a decrease to ~65%nW.
This pattern indicates (at kleast some) interaction/moderation/effect modification processes.
M->Y:
- Correspondingly, at **low** (fixed!) levels of X=%nonWhite, ~10%, increasing Income leads to longer LifeExp., up to some ~$40k, then a slight decline in LifeExp.
- At **higher** (fixed!) levels of X=%nonWhite, ~47%, increasing Income leads monotonically to longer LifeExp. from ~$25k to some ~$32k, with then a plateau for higher income.
Linear modeling:
If we impose a linearity Income(%nonWhite) assumption, the linear slope (+.10, blue double line) leads to a mediation/indirect effect relation %nonWhite -> Income -> LifeExp: with increasing %nonWhite values, income increases (state-level relations are not as meaningful, however, this does not replicate at county and census tract levels), and when that happens, LifeExp increases too.
- Note that a naïve model would lead to a different %nonWhite -> Income slope than the spatial lag model, shown with interrupted double blue line: this process is not as easily interpretable (the linear slope itself is statistically null too, z = 1.42)

b. Stata gsem analysis

```
. gsem (inckstcy <- pnwhtstct lginck) (lfex1015stct <- inckstcy pnwhtstct lglfex)  if cntcstrst ==1
& contig49 == 1

Iteration 0:  Log likelihood = -202.59165
Iteration 1:  Log likelihood = -202.59165

Generalized structural equation model                     Number of obs = 49

Response: inckstcy
Family:   Gaussian
Link:     Identity

Response: lfex1015stct
Family:   Gaussian
Link:     Identity

Log likelihood = -202.59165

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
inckstcy     |
    pnwhtstct |   .1244338   .0382529     3.25   0.001     .0494595    .1994082
       lginck |   1.015627   .1492799     6.80   0.000     .7230434     1.30821
        _cons |  -3.920341   5.144315    -0.76   0.446    -14.00301    6.162331
-------------+----------------------------------------------------------------
lfex1015stct |
     inckstcy |   .0525945   .0323941     1.62   0.104    -.0108968    .1160858
    pnwhtstct |   .0019849   .0115907     0.17   0.864    -.0207325    .0247024
       lglfex |   .8352339   .1465785     5.70   0.000     .5479454    1.122522
        _cons |   11.20181   11.11614     1.01   0.314    -10.58542    32.98905
-------------+----------------------------------------------------------------
var(e.inckstcy)|   14.50705   2.930867                      9.76361    21.55499
var(e.lfex1015stct)|   .9218784   .1862476                      .6204474    1.369753
------------------------------------------------------------------------------
. *       Indirect Effect
. nlcom _b[inckstcy:pnwhtstct]*_b[lfex1015stct:inckstcy]

       _nl_1: _b[inckstcy:pnwhtstct]*_b[lfex1015stct:inckstcy]

------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .0065445   .0045051     1.45   0.146    -.0022853    .0153744
------------------------------------------------------------------------------

. *       Total Effect
. nlcom _b[lfex1015stct:pnwhtstct] + _b[inckstcy:pnwhtstct]*_b[lfex1015stct:inckstcy]

       _nl_1: _b[lfex1015stct:pnwhtstct] + _b[inckstcy:pnwhtstct]*_b[lfex1015stct:inckstcy]
```

```
------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .0085295   .0105408     0.81   0.418      -.01213     .029189
------------------------------------------------------------------------------
```

GO BACK UP

8. Regression coefficients for the % non-White to Life Expectancy regional variables, at census tract, county and state levels, from naïve/a-spatial and proper/spatial models

```
. sem (pnwhtcstr -> lfex1015ct) if contig49 == 1  // census tracts        naïve/a-spatial
(4533 observations with missing values excluded)

Endogenous variables
  Observed: lfex1015ct

Exogenous variables
  Observed: pnwhtcstr

Fitting target model:
Iteration 0:  Log likelihood = -456751.91
Iteration 1:  Log likelihood = -456751.91

Structural equation model                               Number of obs = 60,609
Estimation method: ml

Log likelihood = -456751.91

------------------------------------------------------------------------------
                 |                 OIM
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015ct     |
       pnwhtcstr |  -.0312258   .0005362   -58.23   0.000    -.0322768   -.0301749
           _cons |   79.50243   .0244246  3255.02   0.000     79.45455    79.5503
-----------------+------------------------------------------------------------
var(e.lfex1015ct)|    14.4876   .0832229                      14.3254    14.65164
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                    Prob > chi2 = .

    Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015ct     |
       pnwhtcstr |  -.2301878   .0037954   -60.65   0.000    -.2376267    -.222749
```

```
. sem (lgctlfex pnwhtcstr -> lfex1015ct)  if contig49 == 1  // census tracts      spatial
(4540 observations with missing values excluded)

Endogenous variables
  Observed: lfex1015ct

Exogenous variables
  Observed: lgctlfex pnwhtcstr

Fitting target model:
Iteration 0:  Log likelihood = -628590.55
Iteration 1:  Log likelihood = -628590.55

Structural equation model                               Number of obs = 60,602
Estimation method: ml

Log likelihood = -628590.55

------------------------------------------------------------------------------
```

```
                |                 OIM
                | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
----------------+----------------------------------------------------------------
Structural      |
  lfex1015ct    |
       lgctlfex |   .3606257   .0030415   118.57   0.000     .3546645     .366587
      pnwhtcstr |  -.0224033   .0004888   -45.83   0.000    -.0233614   -.0214452
          _cons |   50.92825   .2419928   210.45   0.000     50.45396    51.40255
----------------+----------------------------------------------------------------
var(e.lfex1015ct)|   11.76046   .0675609                      11.62878    11.89362
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                   Prob > chi2 = .
    Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+----------------------------------------------------------------
Structural       |
  lfex1015ct     |
       lgctlfex |    .427257   .0031272   136.63   0.000     .4211278    .4333863
      pnwhtcstr |  -.1651474   .0035364   -46.70   0.000    -.1720786   -.1582162
```

```
. sem ( pnwhtcyct -> lfex1015cyct )  if cntcstrcnty  ==1 & contig49 == 1  // county naïve/a-spatial
(27 observations with missing values excluded)

Endogenous variables
  Observed: lfex1015cyct

Exogenous variables
  Observed: pnwhtcyct

Fitting target model:
Iteration 0:  Log likelihood = -20635.512
Iteration 1:  Log likelihood = -20635.512

Structural equation model                       Number of obs = 3,060
Estimation method: ml

Log likelihood = -20635.512

------------------------------------------------------------------------------
                |                 OIM
                | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
----------------+----------------------------------------------------------------
Structural      |
  lfex1015cyct  |
       pnwhtcyct |  -.0223835   .0024228    -9.24   0.000    -.0271321   -.0176349
          _cons |   78.31897   .0716737  1092.72   0.000     78.17849    78.45944
----------------+----------------------------------------------------------------
var(e.lfex1015cyct)|   6.659227   .1702465                      6.333772    7.001405
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                   Prob > chi2 = .
    Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+----------------------------------------------------------------
Structural       |
  lfex1015cyct   |
       pnwhtcyct |  -.1647317   .0174673    -9.43   0.000    -.1989669   -.1304965
```

```
. sem (lgcylfex pnwhtcyct -> lfex1015cyct )    if cntcstrcnty  ==1 & contig49 == 1  // county
spatial
(31 observations with missing values excluded)

Endogenous variables
  Observed: lfex1015cyct

Exogenous variables
  Observed: lgcylfex pnwhtcyct

Fitting target model:
Iteration 0:  Log likelihood = -28611.732
Iteration 1:  Log likelihood = -28611.732
```

```
Structural equation model                                   Number of obs = 3,056
Estimation method: ml

Log likelihood = -28611.732

------------------------------------------------------------------------------
                   |                 OIM
                   | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
Structural         |
  lfex1015cyct     |
          lgcylfex |   .2496336   .0123865    20.15   0.000     .2253565    .2739108
         pnwhtcyct |  -.0181607   .0022829    -7.95   0.000    -.0226352   -.0136862
             _cons |   58.92447   .9643151    61.10   0.000     57.03445    60.81449
-------------------+----------------------------------------------------------------
var(e.lfex1015cyct)|   5.863192   .1499935                      5.576459    6.164669
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                     Prob > chi2 = .

      Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
Structural         |
  lfex1015cyct     |
          lgcylfex |   .3392273   .0153978    22.03   0.000     .3090482    .3694064
         pnwhtcyct |  -.1338978   .0166168    -8.06   0.000     -.166466   -.1013295
```

```
. sem ( pnwhtstct ->  lfex1015stct)   if cntcstrst ==1 & contig49 == 1  // state naïve/a-spatial
```

```
Endogenous variables
  Observed: lfex1015stct

Exogenous variables
  Observed: pnwhtstct

Fitting target model:
Iteration 0:  Log likelihood = -290.39913
Iteration 1:  Log likelihood = -290.39913

Structural equation model                                   Number of obs = 49
Estimation method: ml

Log likelihood = -290.39913

------------------------------------------------------------------------------
                   |                 OIM
                   | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
Structural         |
  lfex1015stct     |
         pnwhtstct |  -.0224077    .014918    -1.50   0.133    -.0516464     .006831
             _cons |   79.30008   .4833988   164.05   0.000     78.35264    80.24753
-------------------+----------------------------------------------------------------
var(e.lfex1015stct)|   2.291846   .4630228                      1.54247     3.40529
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                     Prob > chi2 = .
      Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
Structural         |
  lfex1015stct     |
         pnwhtstct |  -.2098045   .1350576    -1.55   0.120    -.4745126    .0549036
```

```
. sem (lglfex pnwhtstct ->  lfex1015stct)    if cntcstrst ==1 & contig49 == 1  // state spatial
```

```
Endogenous variables
  Observed: lfex1015stct

Exogenous variables
```

```
  Observed: lglfex pnwhtstct

Fitting target model:
Iteration 0:  Log likelihood = -346.75214
Iteration 1:  Log likelihood = -346.75214

Structural equation model                                   Number of obs = 49
Estimation method: ml

Log likelihood = -346.75214

------------------------------------------------------------------------------
                 |                 OIM
                 | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-----------------+------------------------------------------------------------
Structural       |
  lfex1015stct   |
           lglfex |   .9789099   .1199531     8.16   0.000     .7438062    1.214014
        pnwhtstct |   .0107523   .0105282     1.02   0.307    -.0098827    .0313873
            _cons |   1.342231   9.557936     0.14   0.888    -17.39098    20.07544
-----------------+------------------------------------------------------------
var(e.lfex1015stct)|   .9714721    .196267                      .6538252    1.443441
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(0) = 0.00                     Prob > chi2 = .

      Standardized | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------
Structural         |
  lfex1015stct     |
           lglfex |   .8044597   .0572806    14.04   0.000     .6921919    .9167276
        pnwhtstct |   .1006741   .0980638     1.03   0.305    -.0915274    .2928756
```

GO BACK UP

9. ICCs for (census tracts)counties clustering

```
 . mixed lfex1015ct || cntcstrcnty:  if contig49 == 1

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -167571.25
Iteration 1:  Log likelihood = -167570.12
Iteration 2:  Log likelihood =  -167570.1
Iteration 3:  Log likelihood =  -167570.1
Computing standard errors ...
Mixed-effects ML regression                     Number of obs     = 60,609
Group variable: cntcstrcnty                     Number of groups  =  1,763
                                                Obs per group:
                                                              min =      1
                                                              avg =   34.4
                                                              max =  2,939
                                                Wald chi2(0)      =      .
Log likelihood =  -167570.1                     Prob > chi2       =      .
------------------------------------------------------------------------------
  lfex1015ct | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _cons |   79.22592   .0349474  2267.01   0.000     79.15742    79.29442
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
cntcstrcnty: Identity        |
                  var(_cons) |   .3424435   .0329969      .2835108    .4136265
-----------------------------+------------------------------------------------
               var(Residual) |   14.64356   .0843731      14.47912    14.80987
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 2186.00       Prob >= chibar2 = 0.0000
```

```
end of do-file

. mixed lfex1015ct lgctlfex || cntcstrcnty:  if contig49 == 1   // spatial CsTr ICC

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -161117.66
Iteration 1:  Log likelihood = -161114.34
Iteration 2:  Log likelihood = -161114.07
Iteration 3:  Log likelihood = -161114.07

Computing standard errors ...

Mixed-effects ML regression                     Number of obs     =     60,602
Group variable: cntcstrcnty                     Number of groups  =      1,763
                                                Obs per group:
                                                              min =          1
                                                              avg =       34.4
                                                              max =      2,938
                                                Wald chi2(1)      =  14377.06
Log likelihood = -161114.07                     Prob > chi2       =    0.0000

------------------------------------------------------------------------------
  lfex1015ct | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    lgctlfex |   .3662846   .0030548   119.90   0.000     .3602973    .3722719
       _cons |   50.19082   .2431229   206.44   0.000     49.71431    50.66733
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
cntcstrcnty: Identity        |
                  var(_cons) |   .1713497   .0187208      .1383202    .2122663
-----------------------------+------------------------------------------------
               var(Residual) |   11.86456   .0683213      11.73141    11.99923
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 1186.05       Prob >= chibar2 = 0.0000

.   estat icc

Residual intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                 cntcstrcnty |   .0142365   .0015367      .0115184    .0175846
------------------------------------------------------------------------------

.   estat icc
Intraclass correlation
------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                 cntcstrcnty |   .0228509   .0021577       .018983    .0274848
------------------------------------------------------------------------------
```

```
. mixed lfex1015ct lgctlfex || cntcstrcnty:  if contig49 == 1   // spatial CsTr ICC

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -161117.66
Iteration 1:  Log likelihood = -161114.34
Iteration 2:  Log likelihood = -161114.07
Iteration 3:  Log likelihood = -161114.07
```

```
Computing standard errors ...
Mixed-effects ML regression                     Number of obs     =     60,602
Group variable: cntcstrcnty                     Number of groups  =      1,763
                                                Obs per group:
                                                              min =          1
                                                              avg =       34.4
                                                              max =      2,938
                                                Wald chi2(1)      =  14377.06
Log likelihood = -161114.07                     Prob > chi2       =     0.0000
------------------------------------------------------------------------------
  lfex1015ct | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    lgctlfex |   .3662846   .0030548   119.90   0.000     .3602973    .3722719
       _cons |   50.19082   .2431229   206.44   0.000     49.71431    50.66733
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
cntcstrcnty: Identity        |
                  var(_cons) |   .1713497   .0187208      .1383202    .2122663
-----------------------------+------------------------------------------------
               var(Residual) |   11.86456   .0683213      11.73141    11.99923
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 1186.05       Prob >= chibar2 = 0.0000

.   estat icc
Residual intraclass correlation
------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                 cntcstrcnty |   .0142365   .0015367      .0115184    .0175846
------------------------------------------------------------------------------

. loneway  lfex1015ct cntcstrcnty  if contig49 == 1
 One-way analysis of variance for lfex1015ct: CsTr CDC Life Expectancy 2010-20
                                          Number of obs =       60,609
                                              R-squared =       0.0633
    Source                SS         df      MS            F     Prob > F
-------------------------------------------------------------------------
Between cntcstrcnty    58720.512  1,762    33.326057      2.26     0.0000
Within cntcstrcnty     868488.12 58,846     14.75866
-------------------------------------------------------------------------
Total                  927208.63 60,608    15.298453
         Intraclass       Asy.
         correlation      S.E.       [95% conf. interval]
         ---------------------------------------------
            0.03575     0.00591       0.02418     0.04733

         Estimated SD of cntcstrcnty effect      .7397613
         Estimated SD within cntcstrcnty           3.8417
         Est. reliability of a cntcstrcnty mean   0.55714
              (evaluated at n=33.93)
```

GO BACK UP