Milestone 6.5

Ali Crump

4/3/2020

Abstract

Zoorob (2019) shows that geography and fentanyl exposure explain much of the variation in increased overdose mortality rates between 2011 and 2017. This paper succeeds in replicating all of the figures and tables in Zoorob's research with the exception of Table 2B, the total estimated deaths attributable to fentanyl by each model. The estimates from table in the original publication more closely match official mortality statistics than the ones replicated in this paper. In addition to replicating Zoorob's work, the extension of this paper aims to improve the two-stage least squares regression analysis by performing a *TBD* alternative strategy to analyze the instrumental variables. In my analysis I find *X*, which matters because *Y*.

Introduction

Zoorob uses two models; Model 1 shows that fentanyl exposure has a positive association with mortality rates, and Model 2 tries to estimate the causal effect of fentanyl exposure on mortality rates. Zoorob runs a least squares regression for the first model. The model predicts overdose mortality as a function of fentanyl exposure. Fentanyl exposure takes into account the state, year, an error term, and the natural logarithm of the number of test results containing fentanyl:

$Fentanyl_{ij} = \log\left(\frac{S_{ij}}{P_{ij}} + 1\right)$

Model 1 below is an ordinary least squares equation where α_i is state i and η_j is year j The standard errors are two-way clustered by state and year and includes population weights (Zoorob (2019)).

$Overdose_{ij} = \alpha_i + \eta_j + \beta_1 Fentanyl_{ij} + \epsilon_{ij}$

The second model uses a two-stage least squares regression:

$\widehat{Fentanyl}_{ij} = \alpha_i + \eta_j + \beta_1(Longitude_i \cdot Year_j) + \epsilon_{ij} Overdose_{ij} = \alpha_i + \eta_j + \beta_2 \widehat{Fentanyl}_{ij} + \epsilon_{ij}$

Findings in the paper show that much of the variation in the increased overdose mortality is explained by fentanyl exposure, and that fentanyl deaths are highly correlated with geography, as the epicenter of the overdose crisis has shifted towards the eastern U.S. They also found that longitude is better able to explain levels of overdose mortality over time. States east of the Mississippi River tend to have greater fentanyl exposure and sharper increases in overdose deaths than states west of the Mississippi River (Zoorob (2019)). Zoorob also uses both models to estimate the number or overdose deaths attributable to fentanyl and claims that they are broadly consistent with official mortality statistics.

Zoorob obtained the data used for his analysis through a Freedom of Information Act request. The data consist of state test results for drug seizures between 2011 and 2016, which he filters for test results containing fentanyl. Zoorob also uses age-adjusted mortality data from the National Center for Health Statistics. All the data used contain state and year information, and he uses state-annual populatons to calcaulte mortality rates relative to a state's population in a particular year. The data and code that Zoorob used in his paper is available on the Harvard Dataverse. To conduct my replication, I used R. More information on this project can be found on my Github repository.[1]

[1] Github repository

What did I do?

What did I find?

Literature Review

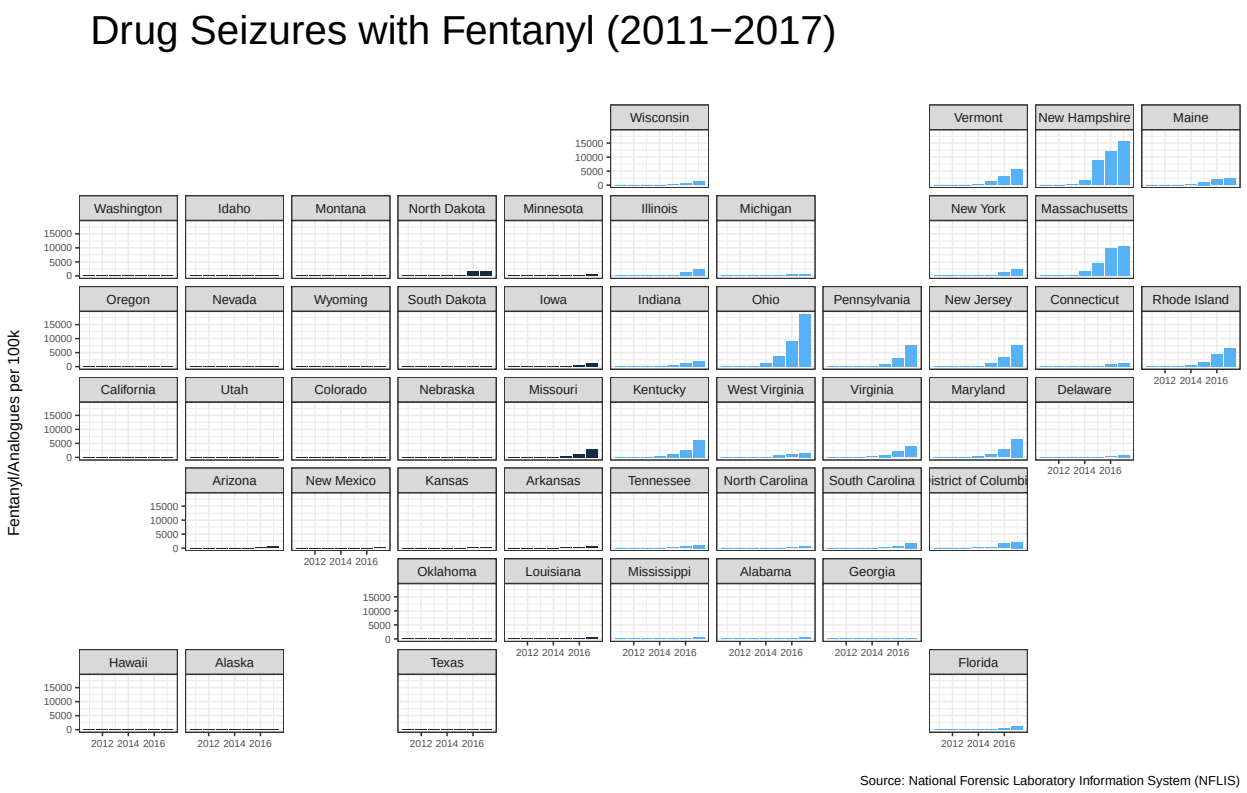
relevant literature in the paper

The number of drug overdose deaths in the United states has rapidly increased since 2014. However, the opiod epidemic did not affect all regions of the U.S. equally; according to the CDC, almost all states west of the Mississippi River did not see an increase while those to the west did. While Dasgupta et. al argue social and economic factors play a role in one's susceptibility to opiod addiction and overdose, Zoorob claims that the geographical patterns point to drug supply also playing a primary role in the epidemic.

relevant literature since the paper published

Research by Barocas et al. on the effect of opiod use among patients with endocarditis cites Zoorob's paper and also finds a decreased risk of overdose associated with the West and South compared to the Northeast (Joshua A. Barocas (2020)). Although in 2018, drug overdose deaths actually decreased by 4.6% from 2017 in the United States, fentanyl deaths continued to rise ((*Drug Overdose Deaths*, n.d.))(Abby Goodnough (2019)). (Still looking for literature on this)

Appendix



tables 2a 2b

table 2 fentr = model 1 fent_r(fit) = model2

% latex table generated in R 3.6.1 by xtable 1.8-4 package % Fri Apr 17 18:39:31 2020

Table 1:

	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
fent_r	4.508***	
	(0.635)	
'fent_r(fit)'		5.443***
		(0.653)
Observations	357	357
R^2	0.928	0.923
Adjusted R^2	0.914	0.908
Residual Std. Error (df = 299)	5,372.861	5,545.678
Note:	*p<0.1; **p<0.05; ***p<0.01	

	Model 1 Deaths	Model 2 Deaths
2011	2580	3115
2012	2659	3210
2013	3723	4495
2014	9973	12041
2015	17367	20969
2016	26491	31985
2017	34176	41263

supplementary table a1 in appendix same as table above but omitting alaska and hawaii

Table 2:

	Dependent variable:	
	age_adjusted_rate	
	(1)	(2)
fent_r	4.546***	
	(0.644)	
'fent_r(fit)'		5.573***
		(0.721)
Observations	343	343
R^2	0.928	0.923
Adjusted R^2	0.915	0.908
Residual Std. Error (df = 287)	5,455.543	5,665.766
Note:	*p<0.1; **p<0.05; ***p<0.01	

table 1 and correlates

longitude r squared

```
## [1] "2011: 0.000133604247072976"
## [1] "2011: (NO Hawaii/Alaska) 0.0204213094278517"
## [1] "2012: 0.00170481159966044"
## [1] "2012: (NO Hawaii/Alaska) 0.0273578839355274"
## [1] "2013: 0.0240861774892379"
## [1] "2013: (NO Hawaii/Alaska) 0.00464511710739798"
## [1] "2014: 0.291779582731855"
## [1] "2014: (NO Hawaii/Alaska) 0.339755720111154"
## [1] "2015: 0.324391935198356"
## [1] "2015: (NO Hawaii/Alaska) 0.447996991934456"
## [1] "2016: 0.463180674808103"
## [1] "2016: (NO Hawaii/Alaska) 0.538614954758827"
## [1] "2017: 0.554045298467256"
## [1] "2017: (NO Hawaii/Alaska) 0.573811061894831"
```

latitude r squared

```
## [1] "2011: 0.022522919018938"
## [1] "2011: (NO Hawaii/Alaska) 0.0229708939591238"
## [1] "2012: 0.0403295747755529"
## [1] "2012: (NO Hawaii/Alaska) 0.0265512094159112"
## [1] "2013: 0.0404112441498816"
## [1] "2013: (NO Hawaii/Alaska) 0.0378140648011715"
## [1] "2014: 0.0250983739834497"
## [1] "2014: (NO Hawaii/Alaska) 0.0293815482091747"
## [1] "2015: 0.00204162860211"
## [1] "2015: (NO Hawaii/Alaska) 0.0285509634034103"
## [1] "2016: 0.0315754568568326"
## [1] "2016: (NO Hawaii/Alaska) 0.0418356152848179"
## [1] "2017: 0.0103733848766385"
## [1] "2017: (NO Hawaii/Alaska) 0.0206704172153174"
```

mortality 2013 r squared

```
## [1] "2011: 0.0285374370715816"
## [1] "2011: (NO Hawaii/Alaska) 0.0229136041040015"
## [1] "2012: 0.00512744385520949"
## [1] "2012: (NO Hawaii/Alaska) 0.00875256024148491"
## [1] "2013: 0.00643192087668385"
## [1] "2013: (NO Hawaii/Alaska) 0.0107743144223076"
## [1] "2014: 0.0967962132363055"
## [1] "2014: (NO Hawaii/Alaska) 0.089162721462024"
## [1] "2015: 0.0540464695899684"
## [1] "2015: (NO Hawaii/Alaska) 0.0546523015095623"
## [1] "2016: 0.0384653767725879"
## [1] "2016: (NO Hawaii/Alaska) 0.0317295820351081"
## [1] "2017: 0.0295998137577852"
## [1] "2017: (NO Hawaii/Alaska) 0.0226044358364418"
```

state year bivariate longitude figure a2 appendix

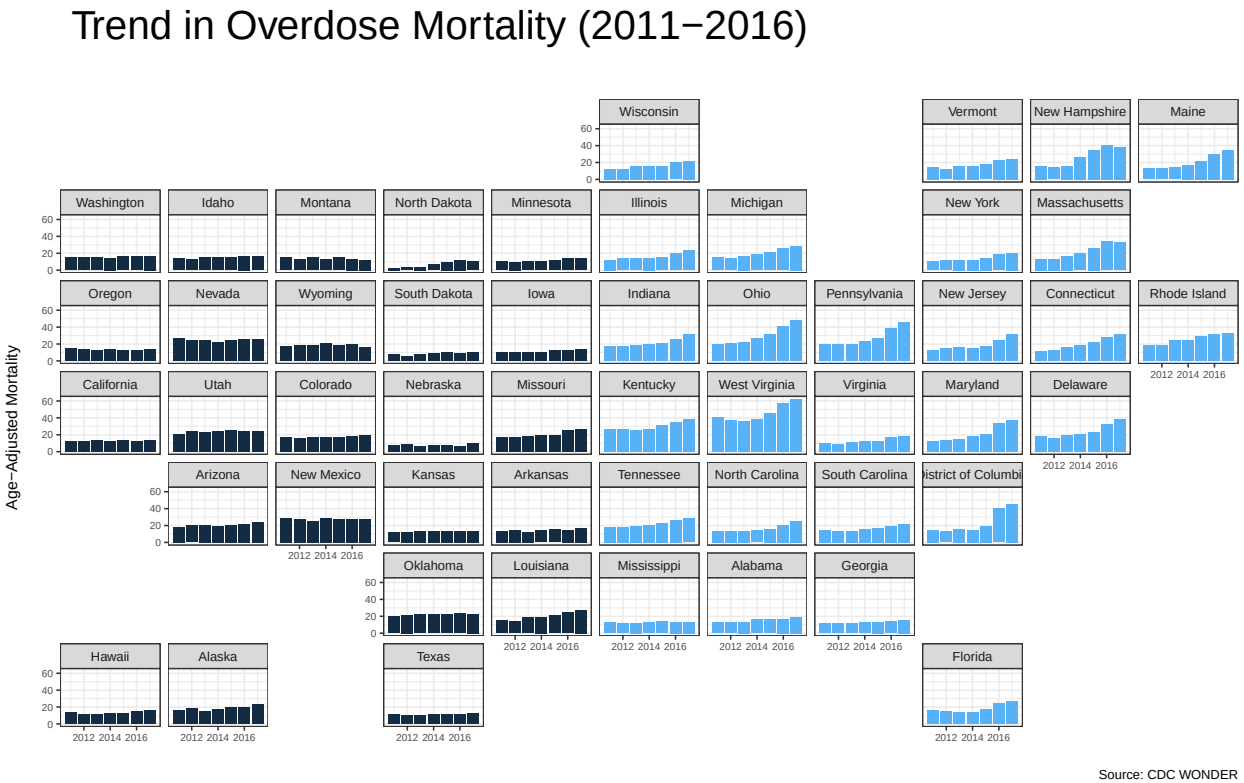
table 1

begin figure 2a and figure 2b paper

Table 3:

	Dependent variable:				
	fent_r				
	(1)	(2)	(3)	(4)	(5)
longitude	0.002	0.019***	0.030***	0.044***	0.053***
	(0.002)	(0.004)	(0.006)	(0.006)	(0.007)
latitude	0.008	0.025*	0.017	0.049**	0.037*
	(0.005)	(0.013)	(0.020)	(0.020)	(0.021)
MORT_2013	−0.003	0.034**	0.035	0.033	0.028
	(0.006)	(0.014)	(0.022)	(0.022)	(0.023)
Constant	0.241	0.985	2.853**	3.498***	5.338***
	(0.290)	(0.679)	(1.070)	(1.084)	(1.130)
Observations	51	51	51	51	51
R^2	0.076	0.406	0.366	0.536	0.590
Adjusted R^2	0.017	0.368	0.325	0.507	0.564
Residual Std. Error (df = 47)	0.233	0.546	0.861	0.872	0.909
F Statistic (df = 3; 47)	1.286	10.706***	9.029***	18.110***	22.591***

Note: *p<0.1; **p<0.05; ***p<0.01



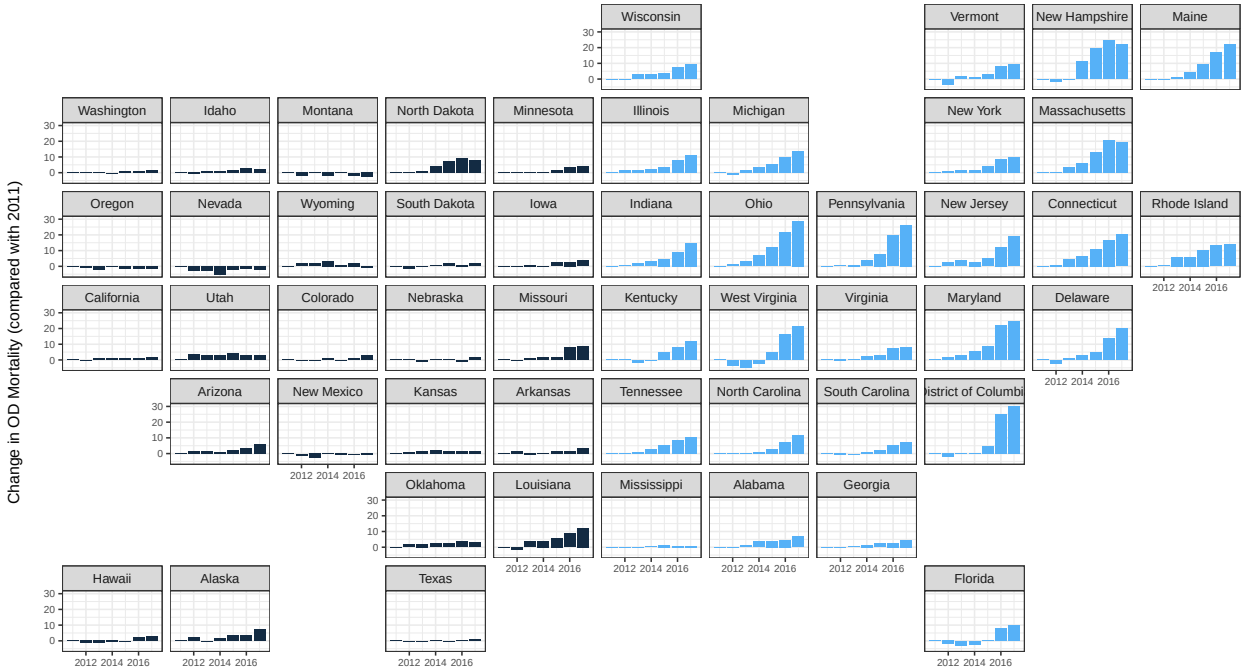


figure 3

```
## `geom_smooth()` using formula 'y ~ x'
```

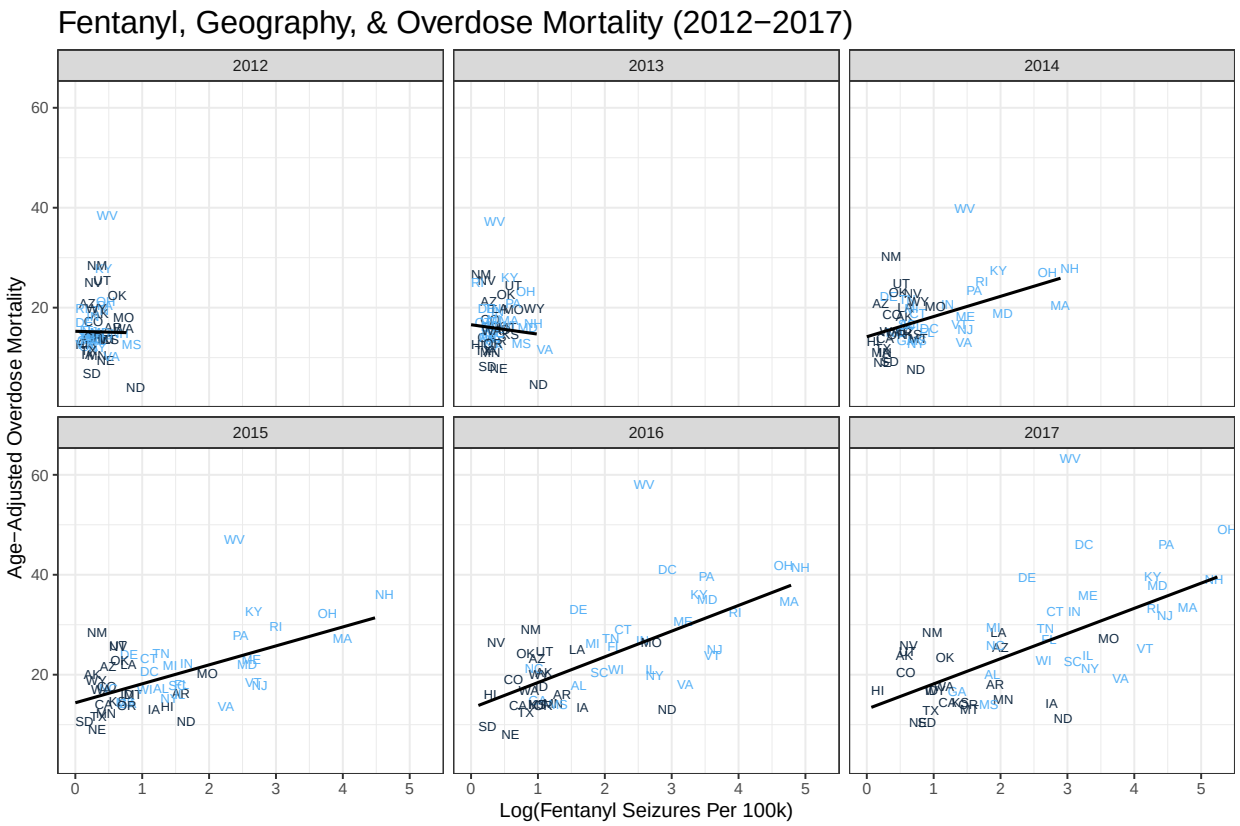
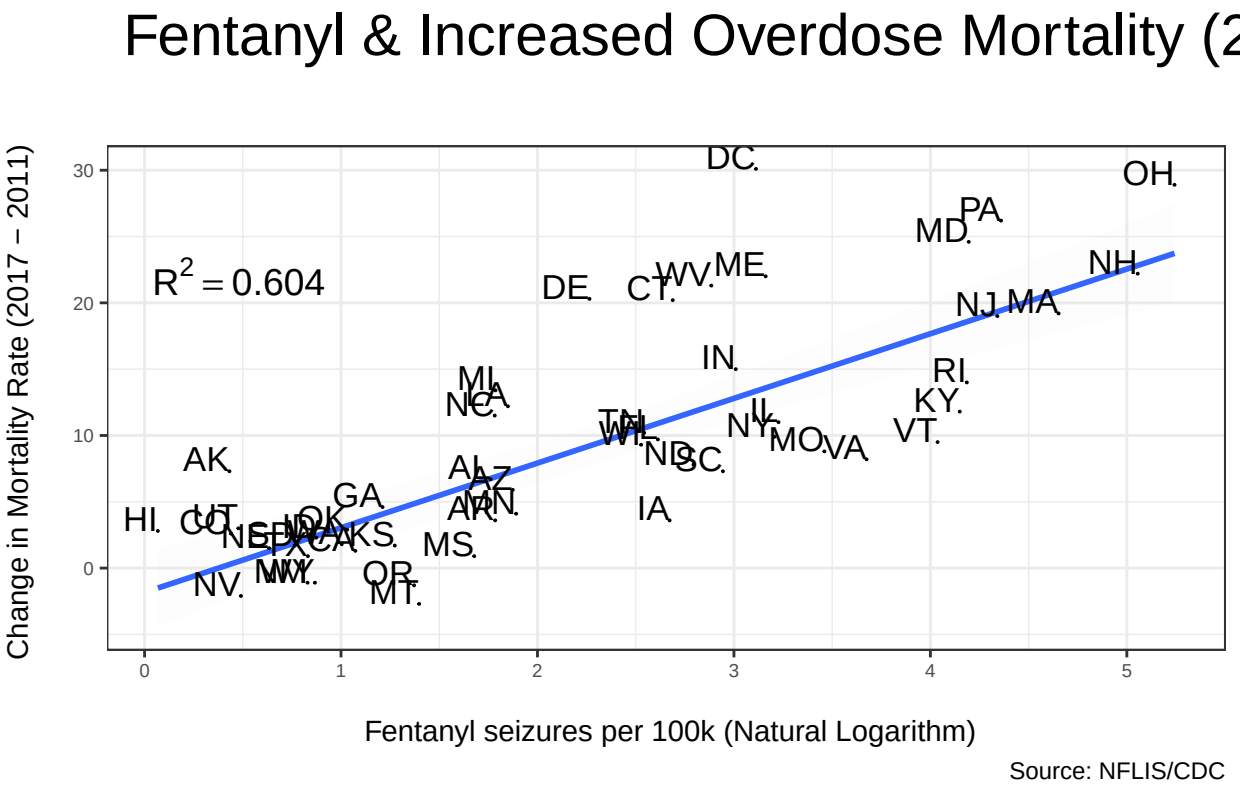


figure 4

```
## `geom_smooth()` using formula 'y ~ x'
```



Discussion

All of the figures and tables appear to be nearly the same, with a few exceptions. Firstly, the tables presented in the paper are of a different format than those produced by stargazer; I suspect the author might have just copied the information into a different table format in Excel for example for clarity (the values are all about the same). Secondly, Table 2 in the paper does not include the residual standard errors when presented in the paper, but I think maybe it should be? Finally, and perhaps most importantly, I see some differences between the small table after table one and this one provided in the paper:

B: Total Estimated Deaths Attributable to Fentanyl by Model.		
	Model 1 Deaths	Model 2 Deaths
2011	2,295	2,705
2012	2,365	2,788
2013	3,312	3,904
2014	8,870	10,458
2015	15,446	18,211
2016	23,188	27,339
2017	30,398	35,841

I will have to look into this further to see whether this is perhaps due to chance or the way I organized the code. The differences are not insignificant as some estimates are off by nearly 6,000 deaths.

Extension

For my extension I am going to use an alternative regression method in place of model 2. Instead of two-stage least squares I am going to use _____ to analyze the instrumental variables.

my extension attmept

```
## stan_glm
##  family:       gaussian [identity]
##  formula:      fent_r ~ longitude + latitude + MORT_2013
##  observations: 51
##  predictors:   4
## ------
##             Median MAD_SD
## (Intercept)  0.233  0.287
## longitude    0.002  0.002
## latitude     0.008  0.005
## MORT_2013   -0.003  0.006
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.236  0.024
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
```

```
##  family:       gaussian [identity]
##  formula:      fent_r ~ longitude + latitude + MORT_2013
##  observations: 51
##  predictors:   4
## ------
##             Median MAD_SD
## (Intercept) 0.975  0.685
## longitude   0.019  0.004
## latitude    0.025  0.013
## MORT_2013   0.034  0.014
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.552  0.058
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      fent_r ~ longitude + latitude + MORT_2013
##  observations: 51
##  predictors:   4
## ------
##             Median MAD_SD
## (Intercept) 2.857  1.097
## longitude   0.030  0.007
## latitude    0.017  0.019
## MORT_2013   0.035  0.022
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.869  0.085
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg

## stan_glm
##  family:       gaussian [identity]
##  formula:      fent_r ~ longitude + latitude + MORT_2013
##  observations: 51
##  predictors:   4
## ------
##             Median MAD_SD
## (Intercept) 3.485  1.074
## longitude   0.044  0.007
## latitude    0.049  0.021
## MORT_2013   0.033  0.023
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.881  0.091
```

```
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg


## stan_glm
##  family:       gaussian [identity]
##  formula:      fent_r ~ longitude + latitude + MORT_2013
##  observations: 51
##  predictors:   4
## ------
##             Median MAD_SD
## (Intercept) 5.325  1.128
## longitude   0.053  0.007
## latitude    0.037  0.021
## MORT_2013   0.028  0.023
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.917  0.094
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```

References

Abby Goodnough, Margot Sanger-Katz, Josh Katz. 2019. *Drug Overdose Deaths Drop in U.s. For First Time Since 1990*. The New York Times. https://www.nytimes.com/interactive/2019/07/17/upshot/drug-overdose-deaths-fall.html.

Drug Overdose Deaths. n.d. Centers for Disease Control; Prevention. https://www.cdc.gov/drugoverdose/data/statedeaths.html.

Joshua A. Barocas, Jianing Wang, Jake R. Morgan. 2020. *Outcomes Associated with Medications for Opiod Use Disorder Among Persons Hospitalized for Infective Endocarditis. Infectious Diseases Society of America*. https://watermark-silverchair-com.ezp-prod1.hul.harvard.edu/ciaa062.pdf?token=
AQECAHi208BE49Ooan9kkhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAn0wggJ5BgkqhkiG9w0BBwagggJqMIICZgIBADCCA
EQx-LXUgTgLIiPAgEQgIICMC-7gICW1kwD-HanETepgOj5x6deCq5Fr3n9WXriHm-9kI1nEaZorUoyzw9fFR28xBPraP8ad7
x9ke4fdLGmXZ6KXr9G06vH0T45mov4AIT93pZ0u8a1WCqluFe92iBuye2CAUOl0ZOOHmI8fTSPxhwTFSZdu6Uncelk9wrR
IfiALtxGG5zMzG9jlEnpSaWEmsBeYMmr7IHFRmuKvFkseVcPMtAgtO3N9Wjo7qzWYW48WLME18g2uFRaidSMtzVMCf
DKiIiFGvq7LbF3z0MNRaohlFQ2WpZagYsnscG41hy5OUQJME4K7VpCChcM0V6H159z1wZf4ooVo65ipAjiWs1tgbeXssyjuk
4elwIf0YBWCouCB-vWRbtfJRZnWm4etPfARlgtTKLD4DL4kxW6fes-609BTXekgqTMc0qTqoo4StQUMxGL55SsTayp1EF
rxb-ASlH8-TZlAKAiAW6RneBSD3ClYw2twDMGAZyEsxNON6veX7tgL3_6WcvJ_zgCUPhBkEkJS22-
03MXt6XY0bOflFasTqjYgRBTAj4pbvHAYpRdRckKWft2iuSPaY06a7zF1P1yah5vWnoPbDvZjlzYC22.

Zoorob, Michael. 2019. *Fentanyl Shock: The Changing Geography of Overdose in the United States. International Journal of Drug Policy*. https://www.sciencedirect.com/science/article/abs/pii/S0955395919301136.