Renewable Energy Plan Policy Analysis

Gamze Bilsen

Background:

I'm interested in seeing how policy within a given states impacts energy generation in that state ("U.S. State"). To do my analysis, I'll choose among these 5 states from the Midwest: Kansas and Oklahoma have voluntary targets (Voluntary Renewable Energy goals for utilities in each state to meet) set in around the same time, 2009 and 2010 respectively, whereas Arkansas and Nebraska don't have any targets set at all and Missouri set a requirement in 2007 (Renewable Energy Standard) for utilities to meet with regards to percent electricity that's delivered to be sourced from renewable energy sources (Megan).

My hypothesis is that a given directive, voluntary or required mandates such as Renewable Energy Standards, will direct higher renewable energy generation in a given state compared to a state without a requirement, net of other factors.

For this lab, I set out to choose among Kansas, Oklahoma, Arkansas, Nebraska, Missouri to compare, given their more or less similar climates which is assumed to equalize their renewable generation capacity which is an inherently geographic resource. These five states are also Republican states, which is assumed to equalize the state governments visions hence policies and actions such as communications including and beyond the energy sector. Additionally, the policies impact utility energy consumption, however, granular energy consumption data that distinguished between source of energy seems to be unavailable. Given this, I will use state energy generation data from EIA: all five states either import/export less than +-10% of their electricity or export 10-30% of their energy generation - Oklahoma and Arkansas (California), thus I assume this generation data is a good substitute for the utility consumption hence a good indicator for policy impact.

I will be using percent of electricity generation to total generation per state as the target variable, rather than total generation to control for factors that might be impacting overall energy demand such as long-term population change or energy efficiency changes.

```
setwd("/Users/gamzebilsen/Documents/GitHub/state-climate-policy/data")
total_gen <- read.csv('total_utility_gen_per_month.csv')
ren_gen <- read.csv('total_utility_renewable_gen.csv')
```

Below, I'm creating a dataframe for total generation and specifically renewable generation per each state to then identify percent of generation per state coming from renewables.

```
#Preparing the dataframe
library(tidyr)
library(dplyr)

target <- c('Oklahoma : all sectors','Kansas : all sectors','Nebraska : all sectors','Arkansas : all se
total_gen2 <- filter(total_gen, description %in% target)
total_gen2<-total_gen2[,c(4:length(colnames(total_gen2)))]
total_gen2$State <- c('Kansas','Missouri','Nebraska','Arkansas','Oklahoma')
total_gen2.T <- t(total_gen2[,1:ncol(total_gen2)-1])
colnames(total_gen2.T) <- total_gen2[,'State']
```

```r
gcolnames <- data.frame(colnames(total_gen2))
date<- data.frame(gcolnames[1:nrow(gcolnames)-1,])
total_gen3 <- merge(date,total_gen2.T)
colnames(total_gen3) <- c('Date','Kansas','Missouri','Nebraska','Arkansas','Oklahoma')
total_gen2.T <- data.frame(total_gen2.T)

target <- c('Oklahoma : all sectors','Kansas : all sectors','Nebraska : all sectors','Arkansas : all se
ren_gen2 <- filter(ren_gen, description %in% target)
ren_gen2<-ren_gen2[,c(4:length(colnames(ren_gen2)))]
ren_gen2$State <- c('Kansas','Missouri','Nebraska','Arkansas','Oklahoma')
ren_gen2.T <- t(ren_gen2[,1:ncol(ren_gen2)-1])
colnames(ren_gen2.T) <- ren_gen2[,'State']
ren_gen2.T <- data.frame(ren_gen2.T)
head(ren_gen2.T)
```

```
##        Kansas Missouri Nebraska Arkansas Oklahoma
## Jan.01     --        1        2      137       20
## Feb.01     --        1        2      126       19
## Mar.01     --        1        2      119       17
## Apr.01     --        1        2      136       15
## May.01     --        1        2       69       20
## Jun.01     --        1        1       88       12
```

Writing the files to a csv to find percentages per month per state

```r
write.csv(total_gen2.T,"/Users/gamzebilsen/Documents/GitHub/state-climate-policy/data/total_gen2.csv")
write.csv(ren_gen2.T,"/Users/gamzebilsen/Documents/GitHub/state-climate-policy/data/ren_gen2.csv")

setwd("/Users/gamzebilsen/Documents/GitHub/state-climate-policy/data")
percent <- read.csv('percent_renewable_gen.csv')
head(percent)
```

```
##     Date Kansas    Missouri    Nebraska   Arkansas    Oklahoma
## 1 1/1/01      0 0.000140449 0.000689893 0.03510120 0.004453351
## 2 2/1/01      0 0.000160875 0.000824742 0.03736655 0.004711133
## 3 3/1/01      0 0.000169750 0.000888889 0.03808000 0.004464286
## 4 4/1/01      0 0.000196194 0.000959693 0.04059702 0.004152824
## 5 5/1/01      0 0.000182749 0.000800961 0.01869412 0.004601933
## 6 6/1/01      0 0.000147016 0.000374672 0.02129204 0.002381898
```

Treated group vs control group

Among the five, I dug into average temperature and total population across the state and in their largest cities to determine my main comparison group. Oklahoma is 4M in total population with top two cities being ~1M population in total. Missouri's top two cities have a total ~800K population with the state population at 6M. Nebraska is 2M total with top two cities being ~700K population in total. Kansas is 3M in total with top two cities being ~600K population in total. Arkansas is 3M in total with top two cities being ~300K population in total ("List of Largest").

Top two city population to total state population: Oklahoma: 25%, Missouri: 13%, Nebraska: 35%, Kansas: 20%, Arkansas: 10%

Looking at the average temperature: Arkansas (15.7C), Kansas (12.4), Missouri (12.5), Nebraska (9.3) and Oklahoma (15.3). Average temperature is used as a proxy for generation to find best groupings ("Average Annual").

All five states currently have Republican governments.

With this information in addition to the goal of comparing between policy promotion versus not, I will be mainly comparing Kansas to Nebraska given more or less similar levels of average temperature and city population to compare volunteerary renewable integration policy compared to none (Kansas and Nebraska respectively).

I will also compare Missouri to Nebraska to understand the impact of mandatory renewable integration policy versus no policy. This is due to more or less similar population distributions and similar temperatures across each grouping.

I will not be comparing voluntary to mandatory renewable policies due to different time frames enacted, leading to analysis issues.

To analyze the impact, I will be using difference-in-difference regression by assuming that absent a policy change, the compared states would have seen a parallel trend in the integration of renewable into the generation mix in the chosen time-frame.

Difference-in-difference regressions

To recap, Kansas introduced a voluntary renewable energy goal of utilities to meet 20% of energy demand with renewables by 2020 in May 2009. Missouri introduced a requirement of utilities to meet the renewable energy standard of meeting 15% of energy demand with renewables by 2020 in November-2007. Nebraska has no plan or standards as of yet ("U.S. State") in April 2022.

To give further background, all three states likely took advantage of federal energy policies such as the federal government provided $1billion in tax credits for renewable energy generation development in March 2002 and energy policy act that developed these tax credits to promote renewable energy and energy efficiency into law in August 2005 ("Eere Timeline"). So, I will use 2005 as the start year assuming the acceleration of renewable integration will be more evident after that date.

For the first difference in difference comparing Kansas and Nebraska, to make it exactly 4 year difference, I'll have the dates start 08-2005, the policy date as 05-2009 and end date at 02-2013.

For the second difference in difference (Missouri vs Nebraska), I'll have the dates start 08-2005, the policy date as 11-2007 and end date at 01-2010. The time differences are shorter than the previous regression, however to be able to have an equal before-and-after comparison and also be able to take into the August 2005 federal policy change, I have chosen these dates.

```
library(ggplot2)
nebraska <- data.frame(percent$Nebraska)
nebraska$Group <- 'Nebraska'
nebraska$Treatment <- FALSE
date <- data.frame(percent$Date)
nebraska$ID <- seq.int(nrow(nebraska))
date$ID <- seq.int(nrow(date))
nebraska<-merge(nebraska,date)
colnames(nebraska)<-c('ID','Value','Group','Treatment','Date')
nebraska$Date <- as.Date(nebraska$Date, "%m/%d/%y")
head(nebraska)
```

```
##   ID       Value    Group Treatment       Date
## 1  1 0.000689893 Nebraska     FALSE 2001-01-01
```

```
## 2  2 0.000824742 Nebraska     FALSE 2001-02-01
## 3  3 0.000888889 Nebraska     FALSE 2001-03-01
## 4  4 0.000959693 Nebraska     FALSE 2001-04-01
## 5  5 0.000800961 Nebraska     FALSE 2001-05-01
## 6  6 0.000374672 Nebraska     FALSE 2001-06-01
```

```
missouri <- data.frame(percent$Missouri)
missouri$Group <- 'Missouri'
missouri$Treatment <- TRUE
date <- data.frame(percent$Date)
missouri$ID <- seq.int(nrow(missouri))
date$ID <- seq.int(nrow(date))
missouri<-merge(missouri,date)
colnames(missouri)<-c('ID','Value','Group','Treatment','Date')
missouri$Date <- as.Date(missouri$Date, "%m/%d/%y")

kansas <- data.frame(percent$Kansas)
kansas$Group <- 'Kansas'
kansas$Treatment <- TRUE
date <- data.frame(percent$Date)
kansas$ID <- seq.int(nrow(kansas))
date$ID <- seq.int(nrow(date))
kansas<-merge(kansas,date)
colnames(kansas)<-c('ID','Value','Group','Treatment','Date')
kansas$Date <- as.Date(kansas$Date, "%m/%d/%y")
head(kansas)
```

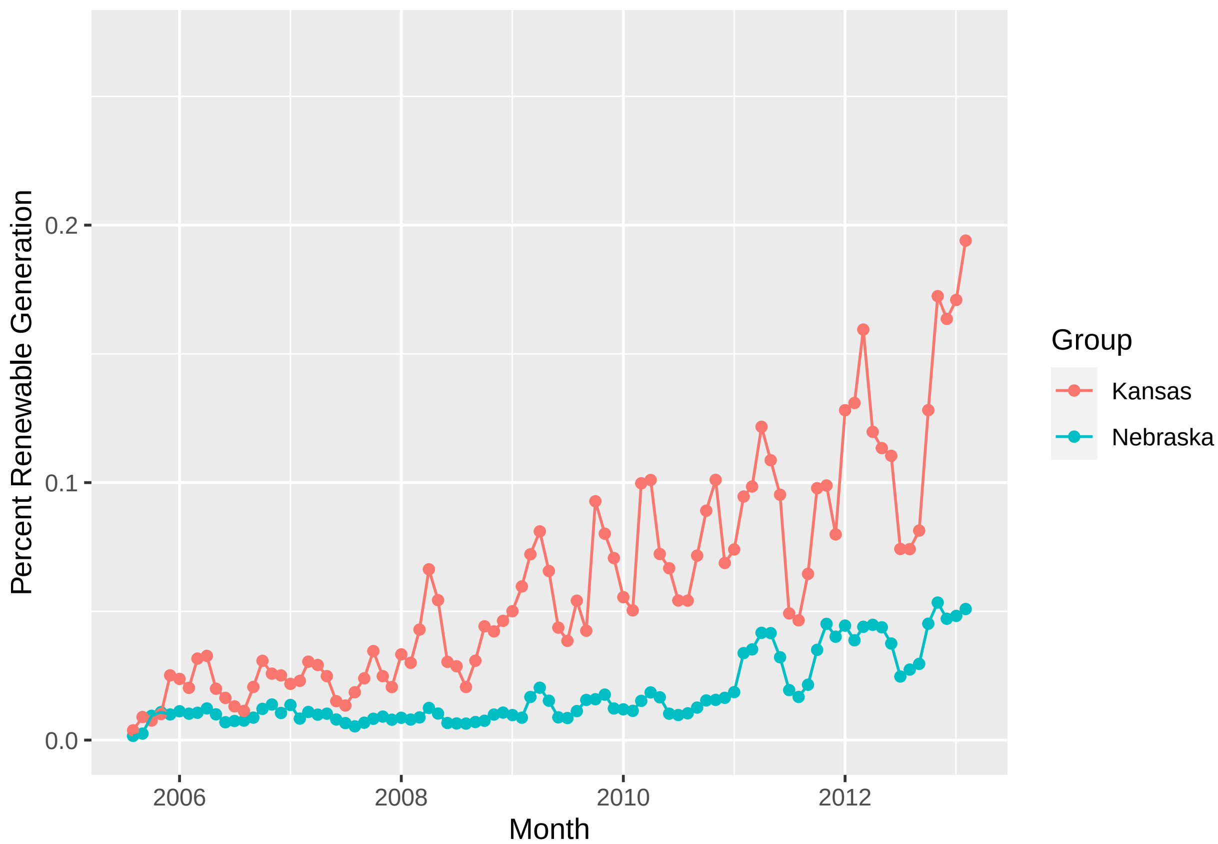
```
##   ID Value  Group Treatment       Date
## 1  1     0 Kansas      TRUE 2001-01-01
## 2  2     0 Kansas      TRUE 2001-02-01
## 3  3     0 Kansas      TRUE 2001-03-01
## 4  4     0 Kansas      TRUE 2001-04-01
## 5  5     0 Kansas      TRUE 2001-05-01
## 6  6     0 Kansas      TRUE 2001-06-01
```

```
diff1<-rbind(nebraska,kansas)
head(diff1)
```

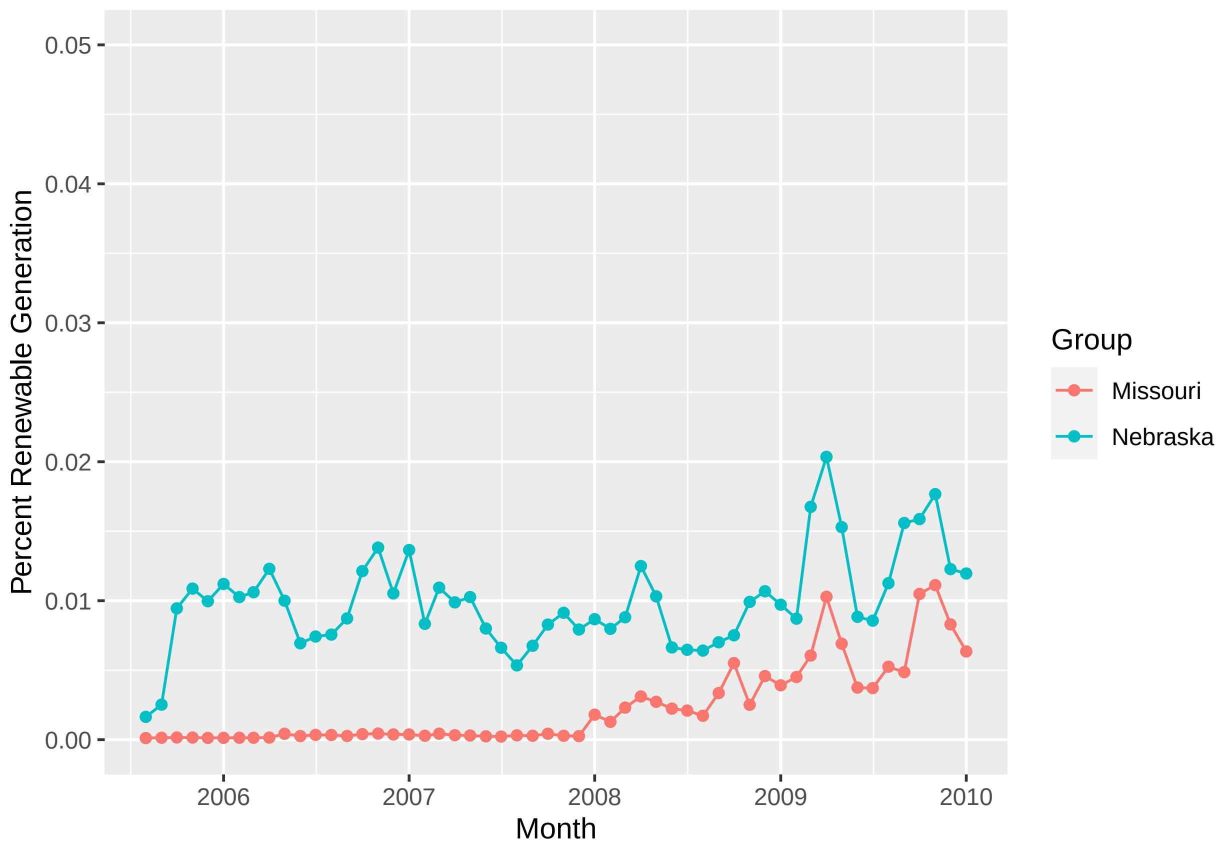
```
##   ID       Value    Group Treatment       Date
## 1  1 0.000689893 Nebraska     FALSE 2001-01-01
## 2  2 0.000824742 Nebraska     FALSE 2001-02-01
## 3  3 0.000888889 Nebraska     FALSE 2001-03-01
## 4  4 0.000959693 Nebraska     FALSE 2001-04-01
## 5  5 0.000800961 Nebraska     FALSE 2001-05-01
## 6  6 0.000374672 Nebraska     FALSE 2001-06-01
```

```
diff2<-rbind(nebraska,missouri)
```

```
qplot(Date, Value, data=diff1, geom=c("point","line"), colour=Group,
        xlab="Month", ylab="Percent Renewable Generation",
      xlim=as.Date(c('2005-08-01','2013-02-01')),ylim=(c(0,0.27)))
```




```
qplot(Date, Value, data=diff2, geom=c("point","line"), colour=Group,
        xlab="Month", ylab="Percent Renewable Generation",
      xlim=as.Date(c('2005-08-01','2010-01-01')),ylim=(c(0,0.05)))
```



For the first graph, looking at Kansas, we can see an overall increase throughout the recent history, however at 2007 there seems to be a significant acceleration in renewable energy generation distribution compared to total generation whereas Nebraska saw a much lower increase throughout increase, however seems to have have a largest acceleration/slope after 2011/2012.

For the second graph, we can see that Missouri and Nebraska both have very similar levels of renewable generation integration throughout history at low levels: looking at Missouri, we can see an overall increase throughout the recent history, however at 2008 there seems to be a shift from levels around 0 to more positive levels, signalling 11-2007 policy impacts. This is compared to Nebraska which seems to have a much higher baseline integration of renewables into the grid already, and who didn't see much of a pronounced increase throughout this timeframe.

Kansas vs Nebraska ~ Voluntary Target vs None

```
diff1 <- diff1 %>% filter(Date>=as.Date('2005-08-01')) %>% filter(Date<=as.Date('2013-02-01'))
diff1$post2009 <- ifelse(diff1$Date>=as.Date('2009-05-01'),1,0)
diff1$Kansas <- ifelse(diff1$Group=='Kansas',1,0)

npre = colMeans(subset(diff1, post2009 == 0 & Kansas == 0, select=Value))
kpre = colMeans(subset(diff1, post2009 == 0 & Kansas == 1, select=Value))
naft = colMeans(subset(diff1, post2009 == 1 & Kansas == 0, select=Value))
kaft = colMeans(subset(diff1, post2009 == 1 & Kansas == 1, select=Value))
(kaft-naft)-(kpre-npre)
```

```
##      Value
```

```
## 0.04458757
```

```
library(MASS)
reg1 = lm(Value ~ post2009 + Kansas + post2009*Kansas, data = diff1)
summary(reg1)
```

```
##
## Call:
## lm(formula = Value ~ post2009 + Kansas + post2009 * Kansas, data = diff1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.053262 -0.011151 -0.001013  0.007065  0.102181
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.009321   0.003284   2.838 0.005063 **
## post2009        0.017429   0.004619   3.773 0.000219 ***
## Kansas          0.020474   0.004644   4.409 1.80e-05 ***
## post2009:Kansas 0.044588   0.006532   6.826 1.32e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02203 on 178 degrees of freedom
## Multiple R-squared:  0.674,  Adjusted R-squared:  0.6685
## F-statistic: 122.7 on 3 and 178 DF,  p-value: < 2.2e-16
```

For the first regression that compares Kansas to Nebraska - voluntary renewable policy vs none - we can see that the difference in difference value of volunteer based renewable energy policy to no policy at all led to a 4.4pp difference in the integration of renewables. Thus, although both states saw an increase in percent of generation sourced from renewables to overall generation, Kansas saw a higher increase after the voluntary policy that promoted higher sourcing of renewables from utilities net of other factors, and at a 99% confidence level.

So, a variety of factors Kansas and Nebraska changed before and after May 2009 to increase percent of energy generation to come from renewables. For this lab, we only know one factor that changed which is the volunteer based policy signalling by Kansas state to have utilities to deliver/consume 20% of their electricity from renewables by a certain date. For reasons that we don't necessarily know, renewables generated at a higher percentage to overall generation for both states before and after May 2009, but for Kansas this increased at a higher rate, because the utilities likely got the signal that the state either now or in the future will impose more stringent policies.

Missouri vs Nebraska ~ Mandatory Target vs None

```
diff2 <- diff2 %>% filter(Date>=as.Date('2005-08-01')) %>% filter(Date<=as.Date('2010-01-01'))
diff2$post2007 <- ifelse(diff2$Date>=as.Date('2007-11-01'),1,0)
diff2$Missouri <- ifelse(diff2$Group=='Missouri',1,0)

npre2 = colMeans(subset(diff2, post2007 == 0 & Missouri == 0, select=Value))
mpre = colMeans(subset(diff2, post2007 == 0 & Missouri == 1, select=Value))
naft2 = colMeans(subset(diff2, post2007 == 1 & Missouri == 0, select=Value))
maft = colMeans(subset(diff2, post2007 == 1 & Missouri == 1, select=Value))
(maft-naft2)-(mpre-npre2)
```

```
##       Value
## 0.002338552
```

```
reg2 = lm(Value ~ post2007 + Missouri + post2007*Missouri, data = diff2)
summary(reg2)
```

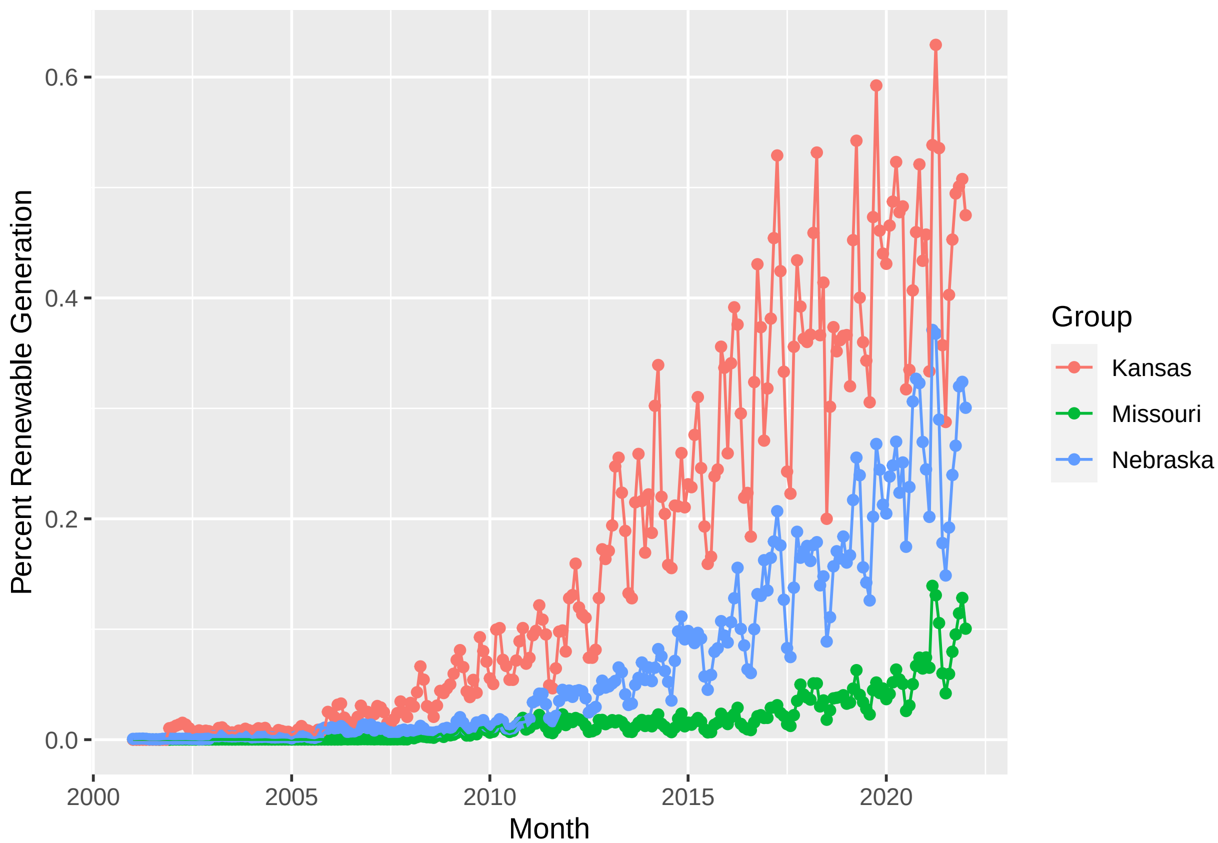
```
##
## Call:
## lm(formula = Value ~ post2007 + Missouri + post2007 * Missouri,
##     data = diff2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0073966 -0.0017636 -0.0000372  0.0010996  0.0095064
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.0090360  0.0005412  16.696   <2e-16 ***
## post2007           0.0018076  0.0007654   2.362    0.020 *
## Missouri          -0.0087671  0.0007654 -11.455   <2e-16 ***
## post2007:Missouri  0.0023386  0.0010824   2.161    0.033 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002812 on 104 degrees of freedom
## Multiple R-squared:  0.6905, Adjusted R-squared:  0.6816
## F-statistic: 77.34 on 3 and 104 DF,  p-value: < 2.2e-16
```

For the second regression that compares Missouri to Nebraska - mandatory renewable policy vs none - we can see that the difference in difference value of volunteer based renewable energy policy to no policy at all led to a 0.23pp difference in the integration of renewables. Thus although both states saw an increase in percent of generation sourced from renewables to overall generation, Missouri saw a slightly higher increase after the mandator policy that promoted higher sourcing of renewables from utilities net of other factors, at a statistically insignificant level.

So, a variety of factors Missouri and Nebraska changed before and after November 2007 to increase percent of energy generation to come from renewables. For this lab, we only know one factor that changed which is the mandatory policy in Missouri whose state government mandated utilities to deliver/consume 15% of their electricity from renewables by a certain date. For reasons that we don't necessarily know, renewables generated at a higher percentage to overall generation for both states before and after August 2008, but for Missouri this increased at a slightly higher rate, because the utilities likely got the signal that the state either now or in the future will impose more stringent policies.

Unlike my prior assumption, volunteer based target led to a higher difference in difference. Consequently, I want to look at the overall change of all three states throughout recent history which I didn't do before to not impact the dates or states I chose:

```
total_graph <- rbind(missouri,kansas,nebraska)
qplot(Date, Value, data=total_graph, geom=c("point","line"), colour=Group,
        xlab="Month", ylab="Percent Renewable Generation")
```



We can see that Missouri overall has a significantly lower integration: Kansas started accelerating the integration much earlier, and the integration understandably stalled off after a few decades likely due to increased marginal cost. Nebraska seems so have accelerated throughout recent history, especially after late mid-2010s: thus despite not having any renewable energy standards, the state accelerated renewable generation integration significantly although not as fast as Kansas likely due to later start. Among the three surprisingly, Missouri has had the least integration of renewables: they were also not able to meet the 15% requirement in 2020 though they finally did in 2022 under the assumption that generation in a state is equal to consumption of electricity by utilities in the said state. This lower increase in renewables in Missouri despite policy is interesting: there is either a characteristic of Nebraska/Missouri that I didn't take account into, leading to a breach of the parallel lines assumption from possible factors such as lower renewable generation capacity across the state, and/or there is likely something about mandatory policy that doesn't impact renewable integration as much as voluntary targets or private industry action for reasons other than following government directives such as monetary gains/high return on investment.

I've also learned that Nebraska released a renewable energy plan in March of 2011, which explains the overall positive trend beginning in early 2010s. Thus despite not having an official voluntary or mandatory directive for renewable integration, we can see the overall positive trend after 2011 for Nebraska as well. However, I don't believe this discredits renewable energy plans as we can clearly see that Missouri saw a significant increase in integration of renewables after the launch of the policy in 2007 whereas Nebraska always had a higher baseline of renewable integration, that throughout history naturally increased in integration, likely due to lower upfront costs of renewewables.

Temperature as a control variable

I created a average temperature dataframe for the selected dates of the three states selected. Average temperature per month per state will be used as a conditional variable to reduce overall variance and account for other changes between two periods. Temperature overall is a good proxy for renewable generation capacity with higher average temperatures signalling higher generation capacity from solar. I would've liked to use more variables such as precipitation and wind speed to control to signal higher generation from other renewables such as hydro and wind in the future.

```r
setwd("/Users/gamzebilsen/Documents/GitHub/state-climate-policy/data")
temp <- read.csv('final_temp.csv')
temp$Date <- as.Date(temp$Date, "%m/%d/%y")
head(temp)
```

```
##   X Kansas_Temperature Kansas_Anomaly Missouri_Temperature Missouri_Anomaly
## 1 0               30.4            1.8                 29.8              0.4
## 2 1               31.8           -1.8                 33.1             -0.7
## 3 2               49.2            6.8                 48.3              4.7
## 4 3               54.9            1.4                 56.2              1.7
## 5 4               66.9            3.5                 67.0              2.8
## 6 5               71.2           -2.2                 70.8             -2.2
##         Date Nebraska_Temperature Nebraska_Anomaly
## 1 2004-01-01                 24.0              1.9
## 2 2004-02-01                 26.3             -0.4
## 3 2004-03-01                 43.1              7.3
## 4 2004-04-01                 50.5              2.8
## 5 2004-05-01                 60.4              2.0
## 6 2004-06-01                 65.2             -3.1
```

```r
nebraska <- data.frame(percent$Nebraska)
nebraska$Group <- 'Nebraska'
nebraska$Treatment <- FALSE
date <- data.frame(percent$Date)
nebraska$ID <- seq.int(nrow(nebraska))
date$ID <- seq.int(nrow(date))
nebraska<-merge(nebraska,date)
colnames(nebraska)<-c('ID','Value','Group','Treatment','Date')
nebraska$Date <- as.Date(nebraska$Date, "%m/%d/%y")
nebtemp <- temp[,c('Nebraska_Temperature','Date')]
nebraska_2 <- drop_na(merge(x=nebtemp,y=nebraska,by="Date",all.x=TRUE))
colnames(nebraska_2)<-c('Date','Temperature','ID','Value','Group','Treatment')

missouri <- data.frame(percent$Missouri)
missouri$Group <- 'Missouri'
missouri$Treatment <- TRUE
date <- data.frame(percent$Date)
missouri$ID <- seq.int(nrow(missouri))
date$ID <- seq.int(nrow(date))
missouri<-merge(missouri,date)
colnames(missouri)<-c('ID','Value','Group','Treatment','Date')
missouri$Date <- as.Date(missouri$Date, "%m/%d/%y")
misemp <- temp[,c('Missouri_Temperature','Date')]
missouri_2 <- drop_na(merge(x=misemp,y=missouri,by="Date",all.x=TRUE))
colnames(missouri_2)<-c('Date','Temperature','ID','Value','Group','Treatment')

kansas <- data.frame(percent$Kansas)
kansas$Group <- 'Kansas'
```

```r
kansas$Treatment <- TRUE
date <- data.frame(percent$Date)
kansas$ID <- seq.int(nrow(kansas))
date$ID <- seq.int(nrow(date))
kansas<-merge(kansas,date)
colnames(kansas)<-c('ID','Value','Group','Treatment','Date')
kansas$Date <- as.Date(kansas$Date, "%m/%d/%y")
kaemp <- temp[,c('Kansas_Temperature','Date')]
kansas_2 <- drop_na(merge(x=kaemp,y=kansas,by="Date",all.x=TRUE))
colnames(kansas_2)<-c('Date','Temperature','ID','Value','Group','Treatment')
```

```r
diff1_2<-rbind(nebraska_2,kansas_2)
diff2_2<-rbind(nebraska_2,missouri_2)
diff1_2 <- diff1_2 %>% filter(Date>=as.Date('2005-08-01')) %>% filter(Date<=as.Date('2013-02-01'))
diff2_2 <- diff2_2 %>% filter(Date>=as.Date('2005-08-01')) %>% filter(Date<=as.Date('2010-01-01'))

diff1_2$post2009 <- ifelse(diff1_2$Date>=as.Date('2009-05-01'),1,0)
diff1_2$Kansas <- ifelse(diff1_2$Group=='Kansas',1,0)
diff2_2$post2007 <- ifelse(diff2_2$Date>=as.Date('2007-11-01'),1,0)
diff2_2$Missouri <- ifelse(diff2_2$Group=='Missouri',1,0)

head(diff1_2)
```

```
##         Date Temperature ID       Value    Group Treatment post2009 Kansas
## 1 2005-08-01        72.6 56 0.001639344 Nebraska     FALSE        0      0
## 2 2005-09-01        67.5 57 0.002519798 Nebraska     FALSE        0      0
## 3 2005-10-01        51.9 58 0.009447166 Nebraska     FALSE        0      0
## 4 2005-11-01        41.1 59 0.010869565 Nebraska     FALSE        0      0
## 5 2005-12-01        25.0 60 0.009960159 Nebraska     FALSE        0      0
## 6 2006-01-01        37.1 61 0.011204482 Nebraska     FALSE        0      0
```

```r
head(diff2_2)
```

```
##         Date Temperature ID       Value    Group Treatment post2007 Missouri
## 1 2005-08-01        72.6 56 0.001639344 Nebraska     FALSE        0        0
## 2 2005-09-01        67.5 57 0.002519798 Nebraska     FALSE        0        0
## 3 2005-10-01        51.9 58 0.009447166 Nebraska     FALSE        0        0
## 4 2005-11-01        41.1 59 0.010869565 Nebraska     FALSE        0        0
## 5 2005-12-01        25.0 60 0.009960159 Nebraska     FALSE        0        0
## 6 2006-01-01        37.1 61 0.011204482 Nebraska     FALSE        0        0
```

```r
reg3 = lm(Value ~ post2009 + Kansas + Temperature + post2009*Kansas , data = diff1_2)
summary(reg3)
```

```
##
## Call:
## lm(formula = Value ~ post2009 + Kansas + Temperature + post2009 *
##     Kansas, data = diff1_2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
```

```
## -0.051285 -0.009831 -0.001160  0.007983  0.093739
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       2.779e-02  5.253e-03   5.291 3.57e-07 ***
## post2009          1.794e-02  4.401e-03   4.077 6.89e-05 ***
## Kansas            2.262e-02  4.451e-03   5.083 9.39e-07 ***
## Temperature      -3.806e-04  8.695e-05  -4.377 2.05e-05 ***
## post2009:Kansas   4.468e-02  6.222e-03   7.181 1.85e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02098 on 177 degrees of freedom
## Multiple R-squared:  0.7058, Adjusted R-squared:  0.6992
## F-statistic: 106.2 on 4 and 177 DF,  p-value: < 2.2e-16
```

When temperature as the controlling factor is used, we can see a statistically significant results for all variables: now the difference in difference between Kansas and Nebraska is much lower, meaning a large chunk of the change in increase of renewables across both states is explained by changes in average temperature across the same time frame rather than the state policy differences.

```r
reg4 = lm(Value ~ post2007 + Missouri + Temperature + post2007*Missouri , data = diff2_2)
summary(reg4)
```

```
##
## Call:
## lm(formula = Value ~ post2007 + Missouri + Temperature + post2007 *
##     Missouri, data = diff2_2)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0066064 -0.0013468 -0.0001194  0.0009595  0.0095316
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.106e-02  9.623e-04  11.491   <2e-16 ***
## post2007            1.546e-03  7.537e-04   2.051   0.0428 *
## Missouri           -8.549e-03  7.516e-04 -11.375   <2e-16 ***
## Temperature        -3.874e-05  1.541e-05  -2.513   0.0135 *
## post2007:Missouri   2.358e-03  1.056e-03   2.233   0.0277 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.002743 on 103 degrees of freedom
## Multiple R-squared:  0.7084, Adjusted R-squared:  0.697
## F-statistic: 62.55 on 4 and 103 DF,  p-value: < 2.2e-16
```

When temperature as the controlling factor is used, we can see a statistically insignificant results except the Missouri variable: however, the results are similar to the difference between Kansas and Nebraska as now the difference in difference between Missouri and Nebraska is much lower, meaning a large chunk of the change in increase of renewables across both states is explained by changes in average temperature across the same time frame rather than the state policy differences.

Interpretation

The initial results, before the control variables, fit my expectations as there is a positive difference from states that had a volunteer or mandatory renewable energy standard to Nebraska who has no renewable energy standard, in the percent generation coming from renewable sources in a given states energy generation net of other factors. The part that I didn't expect was that this difference wasn't statistically significant for the difference between Missouri and Nebraska, and after looking at the changes between all three states throughout history/the graph of three states across the available timeframe, I didn't expect Missouri to have a much lower integration of renewables into their grid compared to Nebraska, nor Nebraska to have such high integration, almost catching up to Kansas who significantly increased renewable integration after the voluntary policy launch. Thus, there is likely something that I didn't account for that led Missouri to overall see a much less integration into the grid. However, I don't believe this means policy doesn't impact generation actions, such as investing in more renewables, much given that Missouri went from almost 0% energy generation from renewables to a positive levels in 2008 which is when the mandatory policy was enacted. This however does show although mandatory directive policy does impact renewable investment, likely agents in a given state has/needs other actions/characteristics of a region that cause additional integration of renewables into the grid such as possible infrastructure needs such as access to transmission lines or tax rebates that the data used for this project doesn't touch upon, especially looking at Nebraska's very promising increase in renewables into their generation portfolio after a vague renewable energy plan in 2011.

Looking at the results of the regressions with the control variables, the difference in differences are still positive however much closer to 0, meaning even one variable besides policy such as temperature explained the variation well. Hence, we can possibly conclude that policy is a positive push towards higher integration into the grid, even without being mandatory standards for utilities - Kansas voluntary policy saw a great acceleration in the integration of renewables into the grid and Nebraska after the renewable energy plan yet no utility based policy in 2011 also saw an increase - however other factors such as renewable generation capacity, represented by temperature for this project, explain this integration well as well, and likely other omitted variables such as the state of the grid/ability to transmit the produced energy likely impact the results.

Works Cited

"Average Annual Temperature for Each US State." Average Annual Temperatures by USA State - Current Results, https://www.currentresults.com/Weather/US/average-annual-state-temperatures.php.

"California Was the Largest Net Electricity Importer of Any State in 2019." Homepage - U.S. Energy Information Administration (EIA), https://www.eia.gov/todayinenergy/detail.php?id=46156#:~:text=In%202019%2C%20California

"Eere Timeline." Energy.gov, https://www.energy.gov/eere/timeline/eere-timeline.

"List of Largest Cities of U.S. States and Territories by Population." Wikipedia, Wikimedia Foundation, 17 Apr. 2022, https://en.wikipedia.org/wiki/List_of_largest_cities_of_U.S._states_and_territories_by_population.

Megan Cleveland, Laura Shields. "State Renewable Portfolio Standards and Goals." State Renewable Portfolio Standards and Goals, https://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx.

"U.S. State Electricity Portfolio Standards." Center for Climate and Energy Solutions, 16 Feb. 2022, https://www.c2es.org/document/renewable-and-alternate-energy-portfolio-standards/.