COAL KEEPS THE LIGHTS ON?
Coal Mine Closures and Local Labor Markets in the US

by

Szilvia Flanek

Submitted to Central European University
Department of Economics and Business

In partial fulfillment of the requirements for the degree of
Master of Arts in Economics

Supervisor: Michael LaBelle

Vienna, 2022

ABSTRACT

Energy transitions are inclined to have winners and losers, which contributes to the polarization of nations. In the US the widespread „Coal Keeps the Lights on” narrative is not only pointing out that coal is necessary to generate electricity, but also infers that the economy is dependent on a thriving coal industry on a deeper level. One concern is that closing coal mines increase local unemployment, which can have spillover effects on other local sectors. In this thesis, I aim to find out if coal mine closures were associated with increasing local unemployment in the Western and Interior coal mine basins of the US during the first two decades of the twenty-first century. To do so I apply the partially pooled synthetic control method in an event study setting with staggered adoption based on the propositions of Ben-Micahel et al (2019) on the county-level unemployment rate and detailed mine level characteristics information. I found no significant effect of mine closures on the local unemployment rate, which can be explained by the flexibility of the US labor markets both by local adjustment and migration. My results suggest that the „Coal Keeps the Lights On” narrative is at least in parts incorrect, however further research on the impact on other outcomes is required.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1. Introduction

Coal mining has a long and volatile history within and outside of the US. In booming cycles, it provides the local economies with employment opportunities, economic growth, and other social spillover effects. However, during declines, when coal is pushed to the background and replaced by other technologies and energy sources, it has the potential to cause serious social unrest.

During the last decade, the US energy sector has experienced a structural change, with the focus diverted to cheaper gas, oil, and the emerging green renewable energy sources. Besides, the regulatory environment has set the agenda to decrease CO_2 emissions and dismantle pollution industries to reduce the environmental impact of energy creation. With the increasing competing industries and the regulatory changes, the coal industry is facing a great decline, that might cause economic and social despair within local and global economies. Policymakers need to consider a wide range of potential impacts when planning and fostering a just energy transition to avoid leaving behind the "losers" and focusing only on the "winners" of the process.

The advocates of the coal sector and many political actors have claimed that the existence of the coal industry is necessary for the economic prosperity of regions with strong coal mining pasts. The slogan "Coal Keeps the Light on!" became a popular catchphrase in coal mining communities, highlighting the real or putative importance of the coal industry within the US economy. This narrative is not only pointing out that coal is necessary to generate electricity, but also infers that the economy is dependent on a thriving coal industry on a deeper level.

The most typical concern is that the decline in coal production and thus the rapidly growing number of closed coal mines increase local unemployment and causes displaced miners to stay permanently out of jobs, which can have spillover effects on other sectors. One of the main messages of Donald Trump's campaign was that "We're going to put miners back to work", (Davenport and Lipton, 2017) which build on the assumption that there is a large magnitude of currently unemployed ex-miners who are waiting to be employed.

But is this the case? Is the decline of coal production increasing the unemployment in the affected communities? In this paper, I will try to find an empirical answer to this question by estimating the effect of coal mine closures on local unemployment using a partially pooled synthetic control method with staggered adoption and detailed mine and county-level data.

I choose the partially pooled synthetic control method since it can be applied using the available data, which includes monthly county-level unemployment rate data for the period between 2000 and 2020 from the Local Area Unemployment Statistics (LAUS) and mine characteristics data available from the Mine Safety and Health Administration reports.

As I will describe in chapter 5, after a few sample restrictions, there are only 63 counties in the dataset with 45 experiencing mine closure during this period. The small number of observations and more importantly, the only 18 potential control counties, would be insufficient for a simple Difference-in-Differences method. With the synthetic control method, I can make sure that the treatment effect is measured compared to an appropriately chosen counterfactual. Using the staggered adoption and partial pooling, all counties that do not experience mine closure up to a certain time can be potential control observations, increasing the pool for the synthetic control.

Using the partially polled synthetic control method with staggered adoption, I found no meaningful short-term effect of mine closure on local unemployment rates in the Western and Interior mine regions of the US between 2000 and 2020, with average treatment effect pooled over all post-treatment periods of -0.064 with a standard error of 0.156. These results are consistent by magnitude across all specifications. The causal interpretation of these results comes with certain limitations. Due to the restricted sample, and great differences between mining regions, the effect might differ for other locations and other periods. Due to the geographic closeness of the observations, both positive and negative spillover effects on the control counties might bias the estimates. Since mine closure might be a foreseeable event, anticipation, and pre-adaption to the labor market shock might also influence the labor market outcomes.

The paper is structured as follows: Chapter 2 presents the characterization and trends of the US coal industry to explain the institutional background and find the underlying characteristics that might affect coal mine closures. Chapter 3 summarizes the current literature on the labor market effect of energy transition and places this paper within that literature. Chapter 4 describes the chosen methodology and argues for this decision, then chapter 5 explains how this methodology is applied to answer the research question. Results are reported in chapter 6 and their implication for policy approaches and limitations are described in chapter 7, finally, chapter 8 concludes.

2. The coal industry in the US

To better understand the nature of the coal industry, this chapter describes the trends and the characteristics of coal production in the US in the last 51 years. All figures, information, and data included here are based on the open-source data of the U.S. Energy Information Administration (EIA), available through API-s provided on their website. (EIA, 2022)

2.1. The process of coal production

The process of coal production is explained in detail by the Macmillan Encyclopedia of Energy. (Macmillan Encyclopedia of Energy, 2022) As they describe it, coal is extracted from underneath the surface level of the earth in coal mines across the globe. It is mainly used as an energy source; the extracted coal is transported to neighboring preparation plants to clean the coal and then to coal-fired power plants to utilize its energy content. The main distinction between coal mines is whether the coal beds or seems it is exploiting lie close to the surface (less than 200 feet) or are buried deep underground.

The depth of the coal seam also defines the mining method. To reach surface level coal, miners remove the top layer (mostly soil or rocks) called overburden and then directly access the underlying coal seam. The overburden might be removed using large machines, and placed over previously mined areas, which is called stripping. In mountainous regions, full mountain tops might be dynamited with explosive devices, called mountaintop removal. The removed excess soil and rocks are placed in near valleys, commonly called "holler fills". The main environmental issue with surface mines lies with the removed soil and surfaces as well as the valleys where the excess material is placed disturbing over- and underground water flows.

To extract coal from seams that are buried several hundred feet underground, coal miners use deep or underground mining. This includes building elevators that run down to the level where the coal lies and tunnels to travel between shafts and using large machines to dig out the coal. Underground mining is more expensive and more dangerous; so, it is less common: in 2020, over 63% of US coal production came from surface-level mines. From an environmental perspective, underground mines have a high risk of methane release and disturbing underground water reserves which also remains an issue long after a mine is closed.

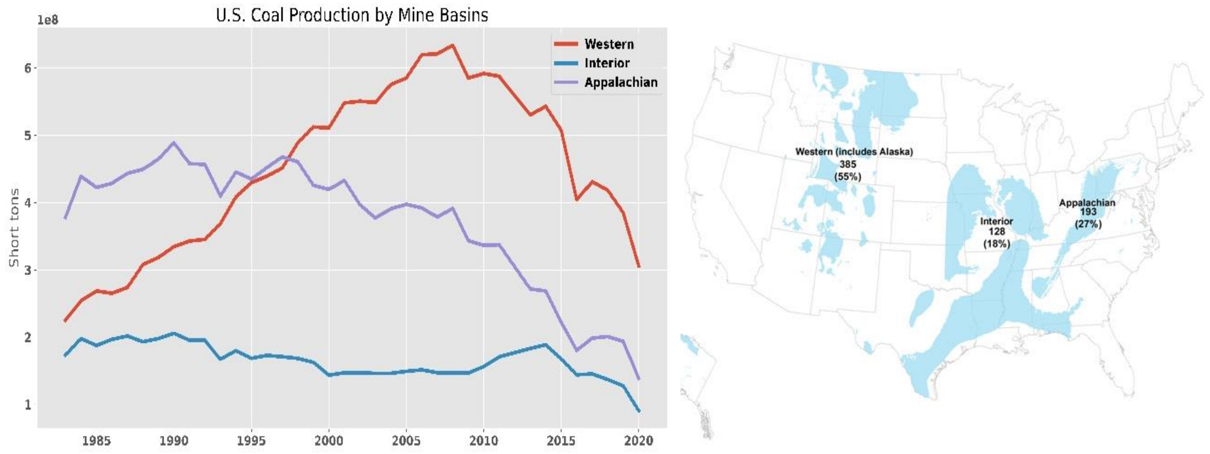
In the US three main type of coal is mined: bituminous, subbituminous, and lignite. Bituminous coal has the highest carbon ratio (45-86%) and thus the highest heating value[1], so it is mainly used in both generating electricity, and making coking coal. It is followed in carbon content by subbituminous coal (35-45%), used almost exclusively by electricity-generating coal-fired power plants. Lignite has the lowest carbon ratio (25-35%) and thus the lowest energy content, also used for electricity generation. Bituminous and subbituminous coal made up about 90% of coal production in the US in 2020, with the remaining 10% being lignite.

2.2. Geography of the coal mines

Coal production is a location-bound technology, since transportation costs are high, coal-fired power plants and coal mines are built close to each other. This means that the main coal-producing regions will be the same as the ones that are rich in coal. In the US, there are three main basins, as shown in figure 1., the Western Basins, the Interior Basins, and the Appalachian Basins.

The Western basin includes Alaska, Arizona, Colorado, Montana, New Mexico, North Dakota, Utah, Washington, and Wyoming. In 2020, 57% of the US coal was produced in the western basin, mostly made up of subbituminous coal and lignite, with 91% from surface mines.

[1] Apart from Anthracite coal, which contains 86-96% carbon, but makes up less than 1% of coal production.



1. Figure: U.S. Coal Production in short tons by mine basins between 1973 and 2020. and location of the basins on the US map. Created based on data from the EIA.

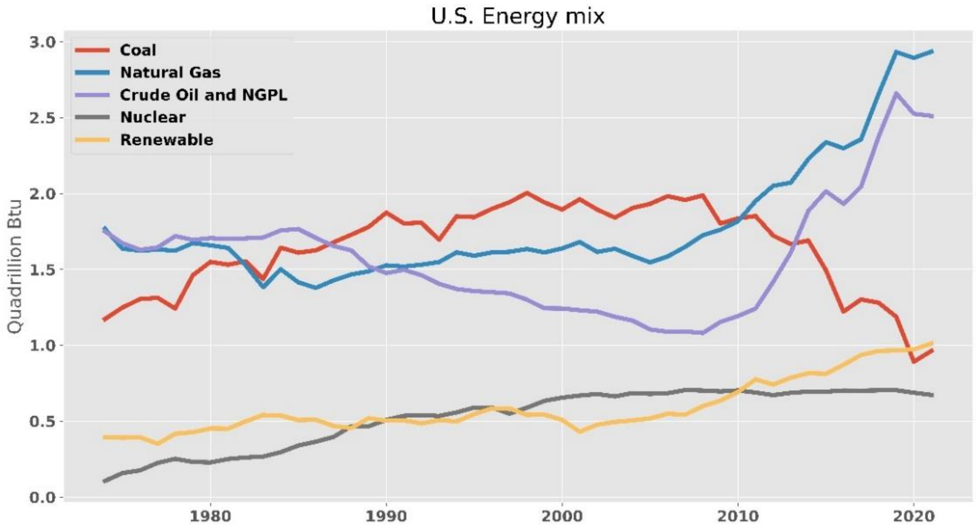
The Interior basin includes Arkansas, Illinois, Indiana, Kansas, Louisiana, Mississippi, Missouri, Oklahoma, Texas, and Western Kentucky. The interior basin produced 17% of the total coal production in 2020 in the US, made up of bituminous coal and lignite, mined in 60% underground and 40% surface-level mines.

The Appalachian coal region includes Alabama, Eastern Kentucky, Maryland, Ohio, Pennsylvania, Tennessee, Virginia, and West Virginia. In Appalachia, the main type of coal is bituminous coal, 81% of which is mined in deep underground mines and the rest is mostly accessed through mountaintop removal. Appalachia gave 26% of US coal production in 2020.

During the expansion of the coal industry in the 1980s-1990s, the Western region took over, by almost tripling its production between 1983 and 2008, while Appalachia slowly declined, and the Interior basin remained about the same. This is due to both regulatory and natural differences between these regions. In the west, the lands are federally owned and rented by the coal-producing companies for low costs.

2.3. The role of coal in the US energy sector

The coal industry has played a prominent role in the energy sector of the US since the second half of the twentieth century. During the 1970s energy crisis, it gain importance as an alternative energy source that was encouraged as a substitute for oil due to the 1973-4 embargo. Between 1975 and its peak in 2008, coal production increased by 70% and became the highest producing energy source, outranking oil, and natural gas for 2 decades.

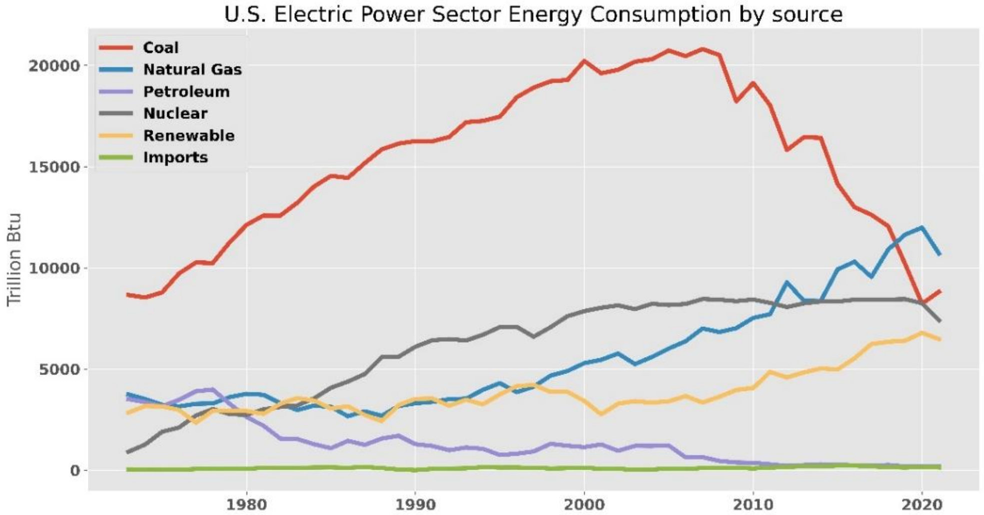


2. Figure: US energy mix: production of energy in quadrillion btu by sources in the US between 1973 and 2021. Edited based on data from EIA.

After 2008, the coal industry experienced a steady decline, decreasing by 52% in the next 13 years, while natural gas, oil, and renewable energy sources gained a stronger position as producers of energy. The declining trend is mostly associated with both increasing production costs due to environmental regulations and the changes in the demand for coal.

In the US the Environmental Protection Agency (EPA) is responsible for combating human interference with the environment, including regulations that affect the coal industry. These laws and regulations focus on water and air pollution both within and across states. The main provisions are the 2009 Celan Water Act, the 2010 Tailoring Rule, the 2011 Mercury and Air Toxics Standards, the 2015 Clean Power Plan, and the 2016 Cross-State Air Pollution Rule. (EPA, 2016a, 2016b, 2016c) These regulations affect both directly the coal mines and the coal-fired power plants, that in turn indirectly affect the demand for coal mines.

On the demand side, coal is mostly consumed by the electric power sector, in 2020, over 90% of coal produced was used to generate electricity. As seen in figure 3, a similar decline in coal production is associated with a decline in coal consumption by the electric power sector. In 2020, 60% less coal was used to generate electricity than in 2008, which was offset by a 76% increase in natural gas, and 87% increase in renewable consumption in the same period.



3. Figure: Consumption of the US electric power sector in trillion btu by source between 1973 and 2021. Edited based on data from EIA.

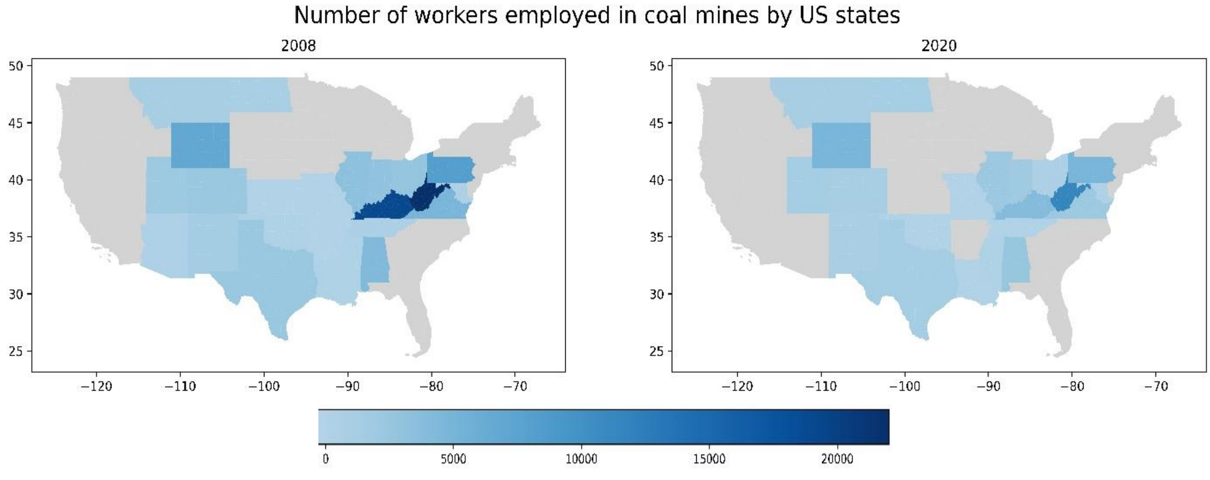
As one of the main substitutions, the natural gas market puts high demand pressure on the coal industry. The price of natural gas has always been a very volatile value, however, it experienced an overall decrease of 78% between 2007 and 2020, serving as a cost-efficient alternative to coal. Together with the increasing presence and popularity of renewable energy sources, they create a strong demand-side pressure, which can also result in the decline of coal production.

As a result of supply and demand-side changes, the number of coal mines decreased substantially during this declining period. Jordan et al (2018) use detailed mine level data to estimate the effect of both these supply and demand-side changes on mine closure. They found that mine closure is affected both by supply and demand-side shocks, with cost increases in Appalachia having the largest effect.

3. The employment effect of energy transitions

Due to the decline of the coal industry, a large share of mines were closed in the US. Both the smaller number of mines and less human labor-intensive technology resulted in fewer people being employed at coal mines on an aggregate level. Between 2008 and 2020, the number of people employed in US coal mines decreased by 51%, which came from all mine basins but in different intensities.

Figure 4. shows the comparison of the number of coal miners by state in 2008 and 2020 based on the Quarterly Employment and Production Data available through the Mine Safety and Health Administration. The Appalachia mining region has suffered the most employment loss, while the decline is similar in the Interior and Western regions. In most of the affected states, coal mines are important employers, providing a large portion of the jobs. It is quite common that the surrounding local economy is built around the mine and its revenue-generating activity, which makes these regions' economies dependent on the mine's existence.



4. Figure: Number of employed people in coal mines by US states in 2008 and in 2020. The color of the states corresponds to the number of people employed in coal mines. Edited based on data from the EIA.

The response of the labor market to a local demand-side shock (like a mine closure) depends on the mobility of the workers and the elasticity of the labor supply. If the workers have perfect information, and mobility has no barriers, the local demand shock should not translate to any change in local unemployment, the displaced workers will just move to other locations with better employment opportunities. Marston (1985) argues that in the US, moving costs and willingness to move across states are high enough that local changes in the labor demand do not have any persistent effect on labor market outcomes.

However, if mobility is not instantaneous or if workers do not have perfect information, a mine closure should be associated with increasing unemployment, and the magnitude of the change depends on the elasticity of the labor supply. Based on Bartik's (2014) argument, both positive and negative local labor market shocks have a persistent impact on labor supply if migration is limited.

While there is extensive research on both the labor market effect of aggregate demand shocks and the labor and poverty-related impact of oil and gas extraction cycles (Marchand 2012, Weber 2012, Brown 2014, Michaels 2011) and some limited research on the economic impact of the development of the coal industry after the 1970s boom (Black et al 2005, Douglas and Walker 2012), the literature on the local labor market effect of the recent decline in the US coal sector is very sparse.

Nevertheless, due to demand changes and environmental regulations, the future of the US coal industry has been on the agenda of many recent political debates both within and outside of the US, thus a thorough examination is necessary. The main question is whether, through local labor market adjustment or migration, the labor markets are flexible enough to respond to coal mine closures and displaced miners can find new employment, or they will remain unemployed after the mine is closed and might even cause negative spillover effects due to decreasing local economic activity.

Black et al (2005) examine the coal boom and bust cycle of the second half of the 1900s and find that non-mining sectors, especially local trading sectors had a positive spillover from the boom of the coal sector. The results show that an additional miner employed during the boom period creates a modest, but significant 0.174 local jobs, and one miner job lost during the bust period eliminates an additional 0.349 local jobs. However, the general economic circumstances were quite different during the 1990-80s, then in the 2010s, thus it is worth analyzing the labor market effects of the recent decline of the coal market.

Burke et al (2019) focus on Australia and use an event study design with regional and time fixed effects to estimate the local labor market effect of coal-fired power stations. Including state-specific month dummies help with both depersonalizing the data and controlling for any labor shocks that are correlated across states for a given month. They find around 0.7 percent of persistent unemployment increase associated with a power-plant closure. As opposed to Burke et al (2019), I will look at responses to coal mines, in the US and will use the synthetic control method to have a better counterfactual realization.

Betz et al (2015) estimate the net local economic impact of the energy transition, specifically the decline of the coal sector and the shift of importance between US basins. They focus on the periods between 1990 and 2010 and use an instrumental variable to control for the endogeneity of the location of coal mining, they find a small effect of coal mining that differs across the basins. The research shows that energy transitions and environmental policies have both winners and losers on a local and wider geographic level.

Severnini (2014) looks at the agglomeration spillover effects of energy transitions, by comparing US counties with and without hydroelectric dams. He finds that the electric dams, that were built before 1950 had positive local labor market effects and argues for the presence of long-term spillover effects. Although this paper focuses on a different energy sector, it is important, since it uses a similar empirical method, a Pooled Synthetic Control Method. To build on his results, I will use a similar setting, but adjust it with a partially pooled model and implement it for the coal mining sector.

4. Methodology

To unveil the association between mine closure and the changes in the local labor market, I attempt to measure the local labor market effect of mine closure using the Synthetic Control Method (SCM) combined with an event study setting. To describe the methodology, I will first introduce the Synthetic Control Method, described by Abadie (2021), and then go on to explain how I can implement this for the staggered mine closure based on the Partially Pooled Synthetic Control Method defined by Ben-Michael et al (2019).

4.1. The Synthetic Control Method

Estimating the effect of some policy intervention or large-scale event with comparative case studies has been the interest of many scientific papers. These methods build on the Potential Outcome Framework (Neyman, 1923; Rubin, 1974) and define the effect of an intervention as the difference between the outcome of interest in the unit (city, state, etc) in which the intervention or treatment is conducted and some other *control unit* that did not receive the treatment that represents the potential outcome of the treated unit without treatment. Thus, for a reasonable identification, we need to define the treatment (policy intervention), the unit that is being treated, and find a control unit that is similar to the treated one in every aspect but the presence of the treatment.

In this paper, I define the observed units as the US counties, the treatment as a mine being closed in the specific county, and my outcome of interest is the unemployment rate in the given county. The potential control units are any other counties that have an active coal mine that is not closing at this time.

The most challenging part of this analysis is choosing the appropriate control unit(s). In some cases, researchers used a single, similar-sized unit and argued that the parallel trends in the outcome before the treatment assures that they are similar enough. (e.g., Card 1990, Card and Kruger 1997). However, most of the time this choice of control is very arbitrary and relies on the informal affinity between treated and control units, besides, a single perfect comparison unit might not even exist in many cases (such as this paper). The synthetic control method, first introduced by Abadie and Gardeazabal (2003) formalizes the choice of the control unit by creating a combination of all potential control units and aggregating them into a *synthetic control* using observable data.

According to Susan Athey and Guido W. Imbens "the synthetic control approach developed by Abadie, Diamond, and Hainmueller (2010, 2014) and Abadie and Gardeazabal (2003) is arguably the most important innovation in the policy evaluation literature in the last 15 years." (Athey and Imbens 2017 page 9.). Indeed, in the past two decades, synthetic controls gained immense popularity for estimating the effect of policy interventions and treatments among economists and social scientists (Acemoglu et al 2016, Peri and Yasenov 2019, Allegretto et al 2019) as well as natural scientists (Proppe et at 2018, Cole, Elliott and Liu 2020).

The main advantage of synthetic controls over a simple difference-in-differences method is that it eliminates the arbitrary nature of the choice of control by implementing a data-driven procedure to reduce the discretion of choice of control units. They are best fitted if there are a finite (mostly small) number of observed units, e.g. states or counties, one (or few) of which are treated. Synthetic control design also provides a setup that does not allow for specification searching such as "p-hacking", since the aggregating weights are calculated by a pre-defined algorithm.

To estimate the local labor market effect of a single mine closure based on the methodology described by Abadie (2021), I will compare the change in the unemployment rate of a treated county after the mine is closed in that unit to the change in a synthetic control county which is a weighted average of all potential control units where no mine has been closed.

Or formally: Let the number of observed time periods be T, where the mine is closed at time $t = T_0$, and the observed counties are $c = 1, 2, \dots, C$, with $c = 1$ being the treated unit. Let Y_{ct} be the labor market outcome (unemployment rate) of county c at time t. For all the post-mine-closure periods $t(t \geq T_0)$, the synthetic control estimator of the effect of mine closure is given by the comparison between the outcome for the treated county and the outcome for the synthetic control county at that period:

$$Y_{1t} - \sum_{c=2}^{C} w_c^* Y_{ct}\,, \tag{1}$$

where $W^* = (w_2^*, \dots, w_C^*)$ is chosen as the weights that minimize the mean squared error of the synthetic control estimator. Let X_1 be the vector containing the values of pre-closure characteristics of the treated county and the closing coal mine and X_0 for all potential counties that have an active coal mine. The pre-closure characteristics also include the lagged pre-treatment values of the outcome variable.

Then choose W^* as the value that minimizes

$$\left|\left|X_1 - X_0 W\right|\right|_V = \sqrt{(X_1 - X_0 W)' V\, (X_1 - X_0 W)}\,, \tag{2}$$

where an optimal choice of V assigns weights to a linear combination of the variables in X_1 and X_0 to minimize the MSE of the synthetic control estimator.

To address inference a typical technique is to conduct a permutation test, where each potential control county is used as a placebo-treated unit and then test whether the estimated effect under the true intervention is significantly large compared to the estimated effects using placebo-treated counties.

4.2. Synthetic Control Method with Staggered Adoption

Since multiple counties experience mine closure at different times, I can use a panel data setting combined with the synthetic control method to estimate the average impact of coal mine closures in an event study design with staggered adoption.

To formalize this method, building on Ben-Michel et al (2019) let T_c be the time when county c experiences a mine closure (this would be T_0 in the previous setting where there was only one treated unit) and order counties by the time they are treated. There are M number of counties that are never treated, they have $T_c = \infty$, and there are $N = C - M$ number of treated counties with $T_1 \leq T_2 \leq \cdots \leq T_N$, they are indexed as $n = 1, 2, \ldots, N$. To implement the event study design, the event time $k = t - T_N$ is introduced representing the time relative to treatment time T_N. Now the county-level treatment effect for treated county n at k is:

$$\delta_{nk} = Y_{nk} - \left(\sum_{c=n}^{N} w_c^* Y_{ck} + \sum_{c=N+1}^{C} w_c^* Y_{ck} \right), \qquad (3)$$

since for a given treated county, the "donor pool" for the synthetic control includes both counties that are never treated and counties that have not been treated yet.

In the event study design, we can average across all treated counties to get the Average Treatment Effect on the Treated for each post-treatment period k (Abraham and Sun, 2018):

$$ATT_k = \frac{1}{N} \sum_{n=1}^{N} \delta_{nk} \,. \qquad (4)$$

One step further, we can also calculate the average post-treatment effect ATT by averaging across all post-treatment periods k. The issue remains of how to calculate the weight and thus which distance to minimize.

Dube and Zipperer (2015) and Donohue et al (2019) fit a synthetic control and calculate the weight for each treated unit one by one, separately minimizing the distance between treated and control before treatment. However, this might lead to a biased average estimate due to poor fit on the average outcome if the separate county-specific estimates are biased, meaning that the outcome of the synthetic control county is a biased estimate of the potential counterfactual outcome of the specific treated county.

One other possibility is to implement pooled synthetic control method for the event study setting based on Ben-Michael et al (2019), where all treated units and all control units all pooled and the weights are calculated to minimize the distance between the average across all treated and control units. This will result in a good fit for the average outcome but give poor fits on the separate county levels.

Based on the advantages and disadvantages of the two approaches, Ben-Michael et al (2019) propose the Partially Pooled Synthetic Control Method. This method calculates the weights by minimizing a convex combination of the county level distance (individual imbalance, Δ^{ind}) and the pooled average distance (global imbalance, Δ^{global}) with hyperparameter ν:

$$\min_{W} \quad \frac{\nu}{2}\Delta^{global}(W) + \frac{(1-\nu)}{2}\Delta^{ind}(W) \tag{4}$$

Choosing the appropriate ν, the solution to the partially controlled synthetic control methods will be the same weight implied by the pooled synthetic control method (with $\nu = 1$) or the separate synthetic control method (with $\nu = 0$).

5. Data and Research design

To implement the described methodology, I acquired mine-level data on the US coal mines from the Mine Safety and Health Administration (MSHA). The Quarterly Employment and Production Data set includes data reported by mine operators quarterly to the MSHA about their production and employment. The Mines Dataset includes information on the characteristics (such as commodity codes, physical attributes, location, operators, etc.) of all coal and metal/non-metal mines under MSHA's control from the year 1970. The datasets can be combined using the unique identifier of the mines.

For the outcome variable, I used the monthly county-level unemployment rate reported in the Local Area Unemployment Statistics (LAUS) by the U.S. Bureau of Labor Statistics. I am interested in estimating the effect of coal mine closure during the recent decline of the coal sector in the US, so I used data between 2000 and 2020. The panel starts before the decline in 2008 so I have a long enough time series to balance the synthetic controls on and goes only until February of 2020 to avoid any contamination from the pandemic.

I call a county treated in a month if a coal mine from my dataset located in the county is closed in that month. I have a variable that shows the date the mine obtained the current status from the Mine Information Form (MIF). The month the date of closure falls into, and all following months, the county will be considered treated. In some counties, there are multiple mines, but as soon as one mine is closed, I will consider the county treated.

The included covariates are chosen based on the characteristics of coal mines that might affect the probability of closure based on what is described in chapter 2 of this paper. These are, besides all the lagged values of the outcome of interest for 3 years before treatment, dummies for the coal classification type (bituminous, subbituminous, or lignite), the mine type (surface or underground), and the coal basin. This minimizes the distance between treated and control before treatment not just on the outcome variable but all the other included covariates, which helps to achieve a sample, where control counties have a similar possibility of facing mine closure as the treated county without actually being treated.

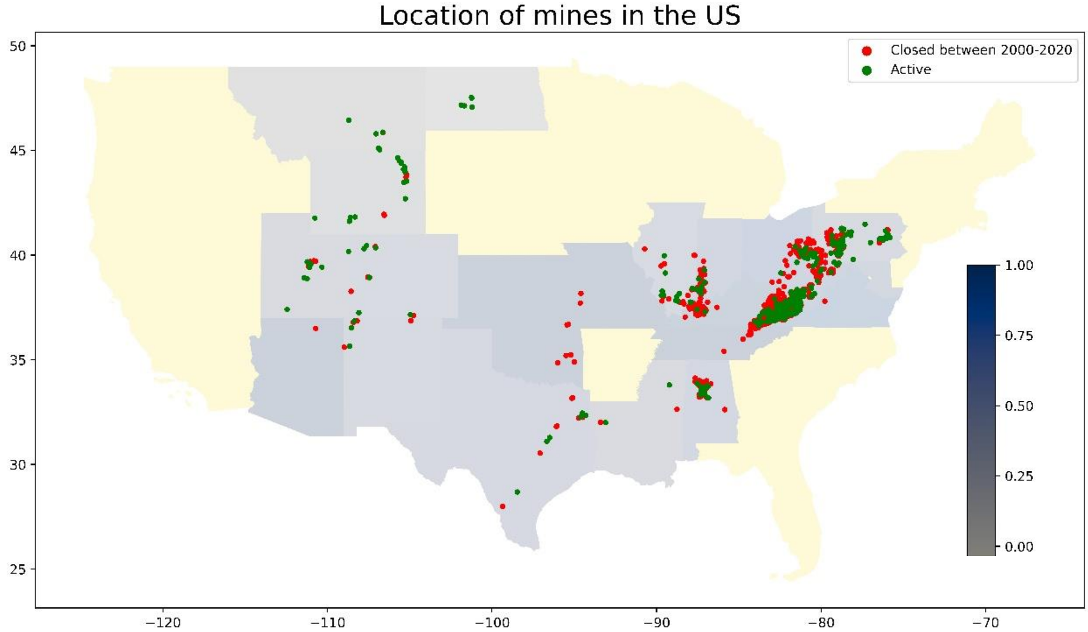
To control for differences between larger and smaller labor markets that might react differently to demand shock, I include the labor force of the county between the balancing covariates. To account for the variation in the number of mines in a county, which might also have an impact on the response, I include the number of mines in the county at the time of treatment as well.

The LAUS does not report deseasonalized unemployment statistics at the county level, so I first deseasonalized the monthly unemployment rate by dividing it with its seasonal index. To calculate the seasonal index for a given month in each county, I first divided each monthly observation by the yearly average and averaged this value for each month within a county. When calculating the treatment effect, I also included unit fixed effects to control for any other time-invariant differences between counties.

I made some sample restrictions through the data cleaning process to get more clear results. The MSHE had data on intermittent, non-producing, and temporarily idled mines, that I did not include. I used only active, abandoned, and abandoned and sealed coal mines. To remove small mines that are not relevant as a significant labor market shock, I removed mines that had less than 30 employees in each quarter during the observed period. The dataset originally also included preparation plants and other coal mining-related facilities that I dropped.

Figure 5. shows all mines that are in this restricted sample, with red for the ones that close during the observed period and green for the ones that stay active. The color of the states corresponds to the share of closed mines in the given state represented by the color bar on the right side of the map. Since the mines and the mining industry in the Appalachian region are very different from the rest of the US and the mines are also a lot closer to each other than

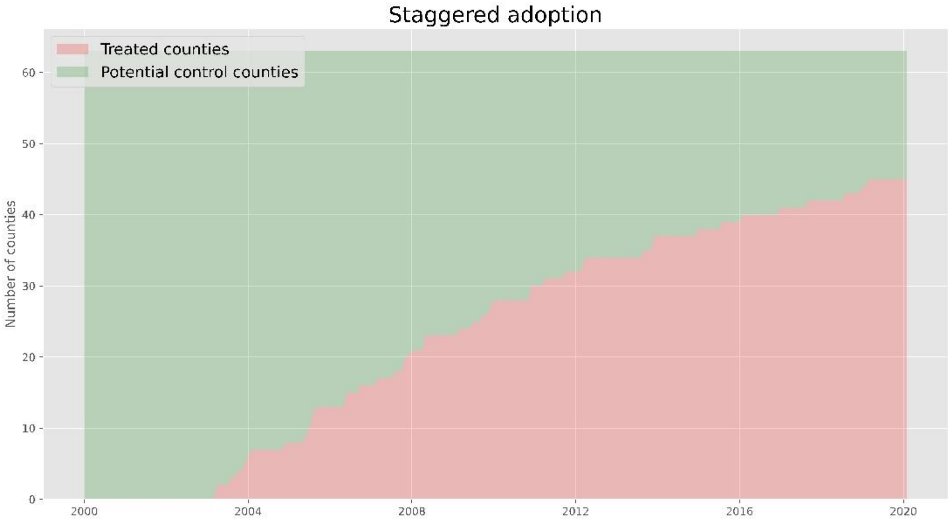
anywhere else, which increases the possible bias from spillover effects, I will not include Appalachia in my model.



5. Figure: Closed and Active coal mines in the US. The coloring of the states represents the share of closed mines in each state, yellow states do not have coal mining activity. Edited based on data from MSHA.

I will focus on the Western and Interior region, where the regulatory circumstance and characteristics of available coal seams are similar, resulting in mines with comparable productivity and potential profit level. Nevertheless, I do balance the included covariates here too, but this will be a less difficult task than if I tried to do that for the whole US.

After the geographic and characteristic-based selection process, there are 166 mines in 63 counties. Within the 63 counties, 18 are never treated, meaning that during the observed period they never experience mine closure. However, since the mines are closed at different times, for most of the counties, there could be more than 18 possible control counties, since all counties that have not been treated in that period will be considered for the donor pool of the synthetic control. Figure 6 shows the staggering adoption as the number of treated and not yet treated, thus possible control counties for each period. The panel includes 16,632 observations, for the period between 2000 January and 2020 February, with monthly data on each county.



6. Figure: Staggered adoption: number of treated and potential control counties in each period. Edited based on data from the MSHA.

If after the treatment starts in a county (a mine is closed) another mine is closed in the same county, the estimated effect not only shows the impact of the first abandoned mine but also of the second. To avoid this, I will only estimate the effect of closure for 4 months past treatment, since there are no counties in the dataset, where the first mine closure is followed by a second one in less than 4 months. I will use 3 years (36 months) of the lagged values of the outcome variable for the balancing of the synthetic controls since there are 3 years of observations available in the data before the first mine closure. The synthetic control counties thus will be by creating a weighted average that is the most similar to the treated counties' outcome for three years before treatment.

6. Results

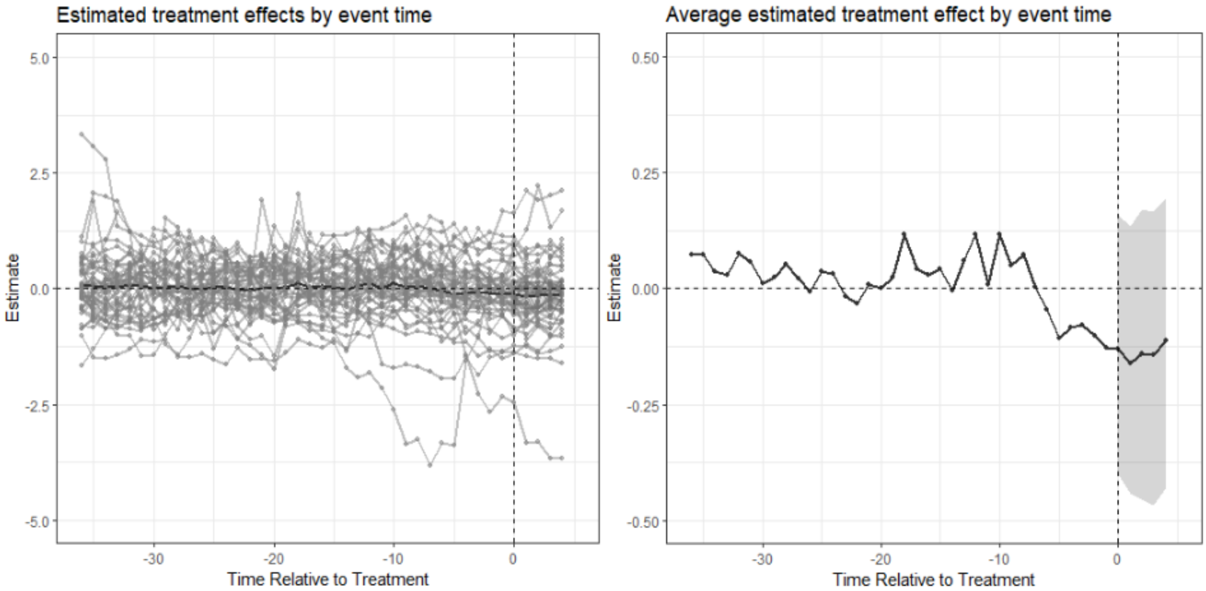
6.1. Separate SCM for each treated county

I first set the balancing hyperparameter ν in the minimization problem (4) to be zero so the model fits a separate synthetic control for each treated unit. Table 1. presents the average estimated treatment effect (ATT_k) for each post-treatment period. The average ATT estimate for all $k \geq 0$ is -0.137 with 0.161 standard error. All estimated effects are small, negative, and insignificant since the standard errors are quite large. Based on this result, the closure of a coal mine has no significant effect on the local unemployment rate on average for all post-treatment periods separately and combined compared to the synthetic control counties.

$\boldsymbol{k}$	$\boldsymbol{ATT_k}$	**Standard error**	**Lower bound**	**Upper bound**
0	-0,130	0,151	-0,422	0,164
1	-0,161	0,156	-0,475	0,144
2	-0,139	0,159	-0,457	0,172
3	-0,142	0,160	-0,439	0,178
4	-0,112	0,156	-0,401	0,197

1. Table: Average treatment effect estimates for each post-treatment period in the separate SCM. The outcome of interest is the unemployment rate, the level of observation is county and month.

Figure 7. shows the pre-treatment balance for the pre-treatment periods and the estimated treatment effect for each post-treatment period. The grey lines are representing the estimated effect in each county and the stronger black line on both the left and the right graph shows the average treatment effect for each k event-time (time relative to treatment), which are the estimates for ATT_k. Before treatment, the estimated effect should be around zero, which shows a well-balanced pre-treatment period and a good fit of the synthetic controls. From this graph, it may already be seen that since the county-specific estimates are most likely biased, the average outcome will have a poor fit resulting in biased average estimates. This suggests, that fitting separate SCM for each treated county is not the optimal solution.



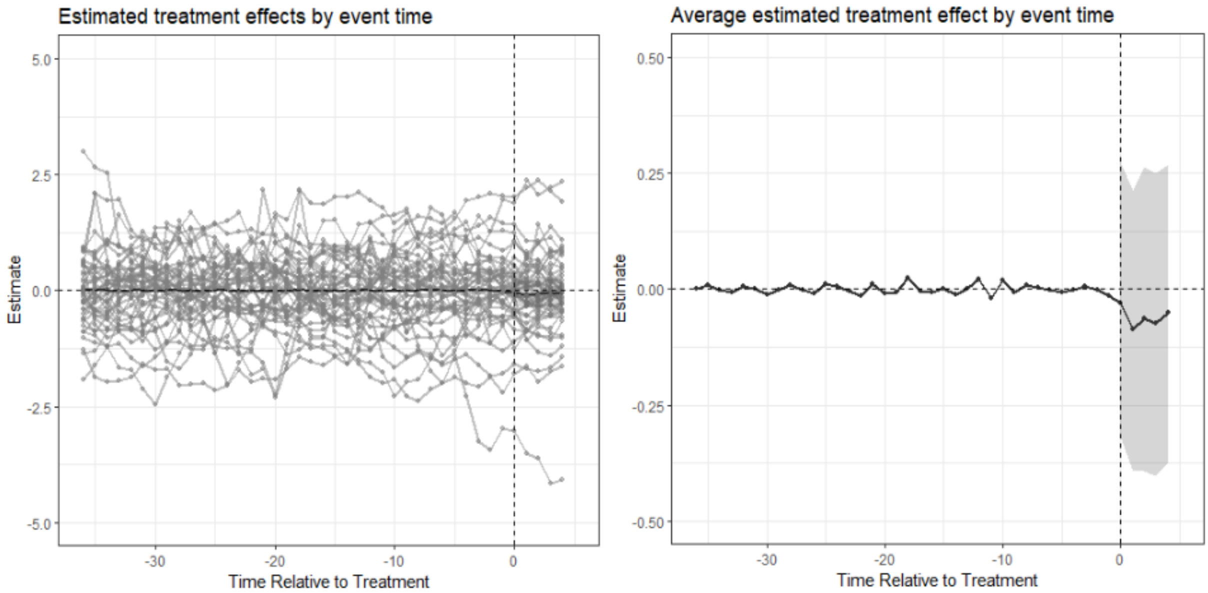
7. Figure: Individual and average balance before treatment and treatment effects with separate synthetic controls. The outcome of interest is the unemployment rate, the level of observation is county and month.

6.2. Pooled SCM

I proceed with estimating the treatment effect with pooled SCM by setting the ν in the minimization problem (4) to be 1. Similar to the previous case, table 2. reports the average treatment effect for each post-treatment period and figure 8. shows the average and individual estimations for each period. The average ATT estimate for all $k \geq 0$ is -0.061 with a 0.151 standard error.

$\boldsymbol{k}$	$\boldsymbol{ATT_k}$	**Standard error**	**Lower bound**	**Upper bound**
0	-0,030	0,146	-0,315	0,259
1	-0,087	0,153	-0,395	0,201
2	-0,064	0,158	-0,376	0,238
3	-0,073	0,162	-0,382	0,246
4	-0,051	0,158	-0,360	0,273

2. Table: Average treatment effect estimates for each post-treatment period in the pooled SCM. The outcome of interest is the unemployment rate, the level of observation is county and month.



8. Figure: Individual and average balance before treatment and treatment effects with pooled synthetic controls. The outcome of interest is the unemployment rate, the level of observation is county and month.

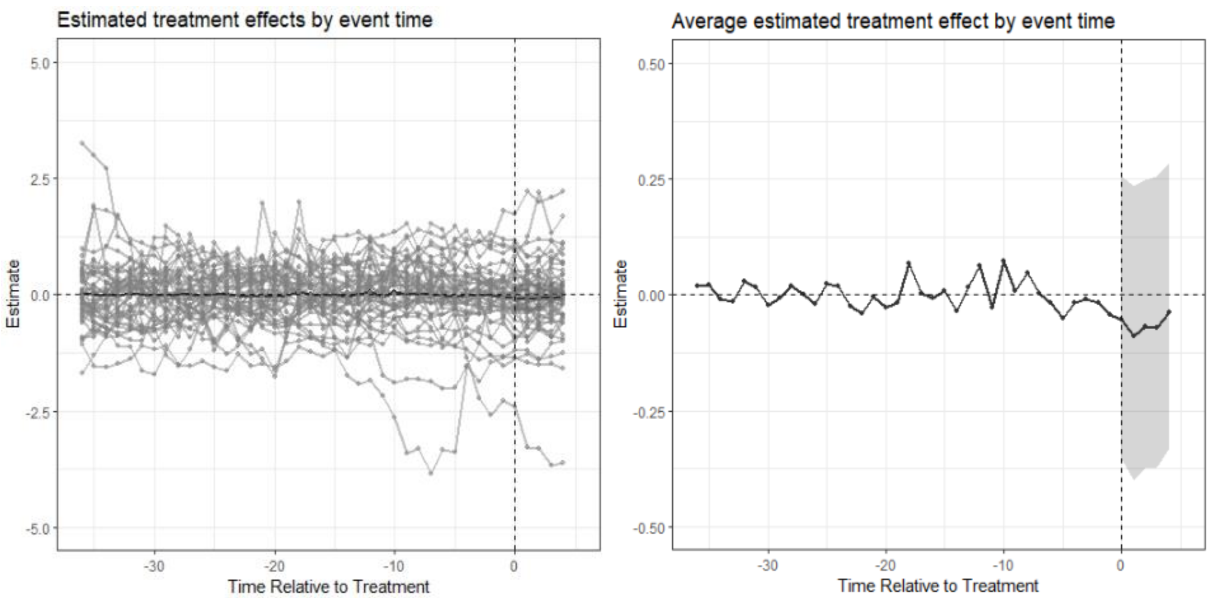
The pooled estimate gives even smaller treatment effects in absolute value, with standard errors of similar magnitude to the separate case. Thus, the effect of mine closure on local unemployment is still insignificant if the synthetic control weights balance the average across all treated units. Compared to the separate case, the average fit is better, since the pre-treatment balance of the average estimator is close to zero. However, the pooled estimator gives poorer county-level fits, since it is choosing the optimal weights without considering unit-level balance.

6.3. Partially Pooled SCM

As explained in chapter 4, implementing a partially pooled estimate is a potential solution to optimize the shortcomings of the two "extremes". Therefore, finally, I implement the partially pooled synthetic control method, letting the model choose the value of ν based on how well separate synthetic controls balance the overall average. The optimal ν in this setting is 0.1108, which is reported during the estimation. Again, table 3. reports the average treatment effect for each post-treatment period and figure 9. shows the average and individual estimations for each period. The average ATT estimate for all $k \geq 0$ is -0.064 with 0.156 standard error. This means that the unemployment rate does not increase on average significantly during mine closure and in the following 4 months compared to its synthetic control.

k	ATT_k	Standard error	Lower bound	Upper bound
0	-0,054	0,156	-0,343	0,255
1	-0,088	0,159	-0,380	0,222
2	-0,069	0,162	-0,367	0,251
3	-0,071	0,165	-0,366	0,260
4	**-0,039**	**0,162**	**-0,323**	**0,286**

3. Table: Average treatment effect estimates for each post-treatment period in the partially pooled SCM model. The outcome of interest is the unemployment rate, the level of observation is county and month.



9. Figure: Individual and average balance before treatment and treatment effects with partially pooled synthetic controls. The outcome of interest is the unemployment rate, the level of observation is county and month.

The final estimate of the effect of coal mine closure on local unemployment is also small in absolute value and negative but insignificant. Since the standard errors are quite large compared to the estimate, the wide confidence intervals include zero for all post-treatment periods. The pre-treatment balance stays close to zero, which shows that the weights of the synthetic controls are well balanced.

Therefore, using separate, pooled, and partially pooled synthetic controls all results in no meaningful effect of mine closure on local unemployment, suggesting that if the model specifications are right, the local labor markets either absorb the extra labor supply or migration has low costs, and the mobility of workers is quite instantaneous.

To compare the three methods, the individual and global imbalance values are reported in table 4. The separate SCM is only minimizing the individual imbalance, thus it is the lowest in this model, the pooled version is minimizing the global imbalance, resulting in a well-fit average balance, but a very poor individual fit. The partially pooled model is finally optimizing the focus by setting the hyperparameter to 0.1108, which puts higher weights on individual balance since the global balance was already quite low for the separate case.

	Separate SCM	**Pooled SCM**	**Partially Pooled SCM**
Individual imbalance	0.648	0.783	0.657
Global imbalance	0.025	0.004	0.012

4. Table: Individual and global imbalance for the separate, pooled, and partially pooled synthetic control methods.

7. Discussion

7.1. Applications for policy

As shown in the previous chapters, I found no meaningful effect of mine closure on local unemployment rates in the Western and Appalachian regions using the synthetic control method with separate, pooled, partially pooled weight balancing in a staggered adoption setting between 2000 and 2020.

These results imply that the US labor markets are flexible enough and mobility has few enough limitations that the local labor supply can adjust to the negative demand shocks caused by mine closure. Therefore, to offset the unemployment effect of the decline of the coal sector, policymakers must focus on maintaining and improving conditions for skill transfer and mobility of workers to maintain the level of employment in these regions.

However, in this paper, I focused on unemployment and did not analyze the effect of mine closures on the quality of work and most importantly wages. It is possible that although displaced miners find new employment, their quality of life and the economic and social well-being of the county still decline.

Mine closure does not only affect miners but the whole local economy since these jobs might have a spillover effect on other local sectors. In addition, due to the lack of mining jobs, workers migrate in great numbers to other locations and leave behind the affected regions, which can also cause economic downturns. Thus, policymakers should pay attention to the impact of energy transitions on sectors and workers outside of the energy industry who are also dependent on it or affected by its changes.

Local spillover effects and migration might give some explanation to why these regions have been deprived compared to other non-mining states of the US, with higher-than-average poverty rates, mortality, and lower health outcomes. (Singh at al., 2017) The well-being of these regions has a strong political and social impact on the US on a federal level: they responded to the campaign of Donald Trump and mostly voted for his presidency, in hope of the revival of the coal industry. (Kemp, 2019)

The contentment of a local region cannot be fully measured and described by one economic metric (like unemployment rate), thus changes in other fundamental human factors need to be addressed when evaluating the effect of the decline of the coal industry on the complex well-being and future of these communities. To mitigate the impact on this wide range of outcomes and to support a "fair transition", a coordinated response is required from several societal actors.

Strambo et al (2019) summarizes all the stakeholders and their responses that have been analyzed in the literature. International institutions have a role in funding research that can map out the impact of mine closures and then advise on the appropriate response programs that help the economic diversification. Governments on both national and local levels need to be involved in these response measures, the former by funding and coordinating and the latter by individualizing and implementing them. On the demand and supply side of the market the private sector organizations and the workers both need to cooperate with the applied measures and support the research with data and other inputs. Finally, civil society has a responsibility to adequately represent the interest of the affected actors who would be left behind otherwise.

7.2. Limitations of the methodology and research design

The result of the models should be taken with precaution since there are some limitations to the causal interpretation and the generalization of the estimates. These effects are mostly valid for the region that they are estimated for. I excluded Appalachia since it is quite different in basic characteristics and the intensity and closeness of mines would have given a biased estimate with large spillover effects. Therefore, the impact of mine closure in Appalachia might be completely different from the estimated impact for the Western and Interior basin.

Even with this restricted sample, spillover effects might be present, and a mine closure can be offset by nearby mines within the county or from neighboring counties. This means that the estimated zero effect might only be a result of the fact that there are still other active mines in the county, and the result would differ greatly if these were the last mines that are closing. There also might be negative spillover effects, on the control counties. A mine closure might affect the economy and thus the labor market of not just the county the mine was in but also other neighboring counties.

To avoid contamination of the effect by other mine closures, I estimated the treatment effect for only 4 months post-treatment, thus all my results show only the short-term impact. The spillover effect of declining mining activity might not be measurable in this timeframe, but other negative effects might show up in the long term.

Mine closure can be an anticipated event, with many signs leading up to it, including reduced production and employment. The legislation does not happen over time, so most regulatory impacts are well known before they come into force. This would mean that workers are looking for jobs or planning migration well before the actual shock (mine closure) happens, which mitigates the instant short-term effects of the treatment.

8. Conclusion

These 7 chapters described the US coal industry and its trend in the past decades, the literature on the economic impact of the energy transition, how the methodology proposed by Ben-Michael et al (2019) can be applied to this scenario, and finally reported the estimated results with their limitations. The paper contributes to the literature on energy transitions by estimating the local labor market effect of coal mine closures using the partially pooled synthetic control method with staggered adoption.

The results show no meaningful treatment effect using partially pooled synthetic controls for the counties in the Western and Interior coal basins of the US between 2000 and 2020. The estimated impacts although are negative, are small and compared to the standard errors insignificant, implying that in contrast to the “Coal Keeps the Lights On!” narrative, a decline in coal mining activity (both in production and in the number of coal mines) is not associated with higher levels of local unemployment. Nevertheless, these results have limitations, especially external validity due to the restricted estimated sample as described in chapter 7.

The policy implications of the results call for an extension of the paper examining the effect of mine closure on the quality of jobs and migration. To estimate the impact of mine closure on wages, a similar methodology could be implemented, with the local average wages as the outcome variable. Looking at a specific percentile of the wage distribution or restricting the sample to workers with a specific level of education might give detailed results on how different socio-economic groups are affected by the decline of the coal sector.

Future research on the energy transitions also needs to address migration since if mobility has low barriers, displaced workers can easily relocate to other regions with better employment possibilities. This helps the workers but does not solve the issue of declining economic activity in the affected region. Again, a similar methodology, but with percentage changes in the local population and controlling for changes in fertility could find some answers to whether coal mine closures cause people to migrate from these counties. However, on top of this, the impact of a declining population on economic outcomes also needs to be analyzed.

The energy transition requires a detailed analysis with a multi-level perspective, considering the role of policymakers and the socio-technical landscape at the same time. While economic impacts might not be significant or hard to measure given the available data, fundamental human factors and the overall well-being of local communities might still suffer. Therefore, research on the impact of coal mine closures on health outcomes might also be of interest.

References

Abadie, A. (2021). *Using synthetic controls: Feasibility, data requirements, and methodological aspects.* Journal of Economic Literature, 59(2), 391-425.

Abadie, A., & Diamond, A. (2014). Hainmueller, J. (2010). *Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program.* Journal of the American Statistical Association, 105(490), 493-505.

Abadie, A., & Gardeazabal, J. (2003). *The economic costs of conflict: A case study of the Basque Country*. American economic review, 93(1), 113-132.

Abraham, S., & Sun, L. (2018). *Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.* Available at SSRN, 3158747.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., & Mitton, T. (2016*). The value of connections in turbulent times: Evidence from the United States*. Journal of Financial Economics, 121(2), 368-391.

Allegretto, S., & Mishel, L. (2020). *Teacher Pay Penalty Dips but Persists in 2019: Public School Teachers Earn about 20% Less in Weekly Wages than Nonteacher College Graduates.* Economic Policy Institute.

Athey, S., & Imbens, G. W. (2017). *The state of applied econometrics: Causality and policy evaluation.* Journal of Economic Perspectives, 31(2), 3-32.

Bartik, T. (2014). *How effects of local labor demand shocks vary with local labor market conditions.* Upjohn Institute Working Paper 14-202. (January 15, 2014).

Ben-Michael, E., Feller, A., & Rothstein, J. (2019). *Synthetic controls and weighted event studies with staggered adoption*. UC Berkely, (December 6, 2019). arXiv:1912.03290.

Betz, M. R., Partridge, M. D., Farren, M., & Lobao, L. (2015). *Coal mining, economic development, and the natural resources curse*. Energy Economics, 50, 105-116.

Black, D., McKinnish, T., & Sanders, S. (2005). *The economic impact of the coal boom and bust*. The Economic Journal, 115(503), 449-476.

Brown, J. P. (2014). *Production of natural gas from shale in local economies: a resource blessing or curse.* Economic Review, 99(1), 119-147.

Burke, P. J., Best, R., & Jotzo, F. (2019*). Closures of coal-fired power stations in Australia: local unemployment effects.* Australian Journal of Agricultural and Resource Economics, 63(1), 142-165.

Card, D. (1990). *The impact of the Mariel boatlift on the Miami labor market.* ILR Review, 43(2), 245-257.

Card, D. and Kruger, A.B. 1997. *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton, NJ: Princeton University Press.

Cole, M. A., Elliott, R. J., & Liu, B. (2020). *The impact of the Wuhan Covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach.* Environmental and Resource Economics, 76(4), 553-580.

Davenport, C. & Lipton E. (2017). *How G.O.P. Leaders Came to View Climate Change as Fake Science.* New York Times, June 3, 2017. Available at: https://www.nytimes.com/2017/06/03/us/politics/republican-leaders-climate-change.html

Donohue, J. J., Aneja, A., & Weber, K. D. (2019). *Right-to-carry laws and violent crime: A comprehensive assessment using panel data and a state-level synthetic control analysis.* Journal of Empirical Legal Studies, 16(2), 198-247.

Douglas, S. M., & Walker, A. W. (2012*). Sample selection in Appalachian research*. Review of Regional Studies, 42(2), 143-159.

Dube, A., & Zipperer, B. (2015). *Pooling multiple case studies using synthetic controls: An application to minimum wage policies*. IZA Discussion Paper No. 8944. (April 6, 2015). Available at SSRN 2589786.

EPA. (Jun 27, 2016). *Clean Power Plan for Existing Power Plants.* Clean Power Plan. Environmental Protection Agency.

EPA. (Nov 3, 2016). *Cross-State Air Pollution Rule (CSAPR*). Environmental Protection Agency.

EPA. (Oct 5, 2016). *Clean Air Act Permitting for Greenhouse Gasses.* New Source Review Permitting. Environmental Protection Agency.

EPA. (Oct 6, 2016). *Regulating Under the Clean Water Act*. Environmental Protection Agency.

Jordan, B., Lange, I., Linn, J. (2018) *Coal Demand, Market Forces, and US Coal Mine Closures* (April 24, 2018). CESifo Working Paper Series No. 6988

Macmillan Encyclopedia of Energy (2022). *Coal, Production of* . Encyclopedia.com. 23 May 2022.

Marchand, W. R. (2012). *Mindfulness-based stress reduction, mindfulness-based cognitive therapy, and Zen meditation for depression, anxiety, pain, and psychological distress*. Journal of Psychiatric Practice, 18(4), 233-252.

Marston, S. T. (1985). *Two views of the geographic distribution of unemployment.* The Quarterly Journal of Economics, 100(1), 57-79.

Michaels, G. (2011). *The long term consequences of resource-based specialisation*. The Economic Journal, 121(551), 31-57.

Neyman, J. S. (1923). *On the application of probability theory to agricultural experiments. essay on principles.* Annals of Agricultural Sciences, (1990), 5, 465-480, 10, 1-51.

Peri, G., & Yasenov, V. (2019). *The labor market effects of a refugee wave synthetic control method meets the mariel boatlift.* Journal of Human Resources 54, no. 2 (2019): 267-309.

Quintero-Bermudez, R., Gold-Parker, A., Proppe, A. H., Munir, R., Yang, Z., Kelley, S. O., ... & Sargent, E. H. (2018). *Compositional and orientational control in metal halide perovskites of reduced dimensionality.* Nature materials, 17(10), 900-907.

Rubin, D. B. (1974). *Estimating causal effects of treatments in randomized and nonrandomized studies*. Journal of educational Psychology, 66(5), 688.

Severnini, E. (2014). *The power of hydroelectric dams: Agglomeration spillovers*. IZA Discussion Paper No. 8082.

Singh, G. K., Daus, G. P., Allender, M., Ramey, C. T., Martin, E. K., Perry, C., Reyes, A., & Vedamuthu, I. P. (2017). *Social Determinants of Health in the United States: Addressing Major Health Inequality Trends for the Nation,* 1935-2016. International journal of MCH and AIDS, 6(2), 139–164.

Strambo, C., Aung, M. T., & Atteridge, A. (2019). *Navigating coal mining closure and societal change: learning from past cases of mining decline*. Stockholm Environment Institute..

Weber, J. G. (2012). *The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming*. Energy Economics, 34(5), 1580-1588.

DATABASES

Local Area Unemployment Statistics, LAUS (2022). *Monthly and annual employment, unemployment, and labor force data for Census regions and divisions, States, counties, metropolitan areas, and many cities, by place of residence reported by the U.S. Bureau of Labor Statistics.* Avaliable at: https://www.bls.gov/lau/

The Mine Safety and Health Administration, MSHA (2022). *Quarterly Employment and Production Data set.* Available through the website of the MSHA at: https://www.msha.gov/mine-data-retrieval-system

The Mine Safety and Health Administration, MSHA (2022). *Mines Dataset.* Available through the website of the MSHA at: https://www.msha.gov/mine-data-retrieval-system

The U.S. Energy Information Administration, EIA (2022). *Open data source on the coal industry*. Available through APIs at: https://www.eia.gov/coal/