Trump Digs Votes – The Effect of Trump’s Coal Campaign on the Presidential Ballot in 2016

Marina Di Giacomo, Wolfgang Nagl, Philipp Steinbrunner
Trump Digs Votes – The Effect of Trump’s Coal Campaign on the Presidential Ballot in 2016

Abstract

In this paper we investigate the effect of Donald Trump’s campaign for coal in his successful race for the White House in 2016. Using a spatial Durbin model we estimate the effect of coal production on the Republicans vote share in the US Presidential Election of 2016 on the county level. To avoid biased estimates we take spillover effects into account and use spatial clustering. We find a significant positive effect. The effect becomes even more pronounced when we use the vote-share difference between Mitt Romney in 2012 and Donald Trump in 2016 as the dependent variable. The positive effect of coal production on the Republican vote share are retained after allowing for non-linear effects of coal production and using coal production per worker and per working hours as main explanatory variable.

JEL-Codes: D720, P160, P180, R110.

Keywords: US Presidential Election 2016, coal production, Durbin model.

Marina Di Giacomo
Department of Economics, Social Studies, Applied Mathematics and Statistics
School of Management and Economics
University of Turin / Italy
marina.digiacomo@unito.it

Wolfgang Nagl
Deggendorf Institute of Technology
Deggendorf / Germany
wolfgang.nagl@th-deg.de

Philipp Steinbrunner*
Department of Economics, Social Studies, Applied Mathematics and Statistics
School of Management and Economics
University of Turin / Italy
philipp.steinbrunner@unito.it

*corresponding author

We would like to thank the participants at the 2021's CESifo workshop on Political Economy, and the 2022’s EPCS and SES conferences for their comments.
1. Introduction

Donald Trump promised in his 2016 Presidential campaign to end the ’war on coal’ and put U.S. miners back to work. The peak of his coal campaign was probably a speech at Charleston Civic Center on May 5, 2016 in Charleston, West Virginia in the Appalachian coal region. He wore a miners hard hat and promised his crowd “For those miners, get ready because you’re going to be working your asses off”. ¹ This speech happened just a few weeks after his rival, the Democrat nominee Hilary Clinton, stated in a speech in Columbus Ohio, which is one of the U.S. largest coal producers, as part of a longer statement, that her government would “put a lot of coal miners and coal companies out of business”.² Trump capitalised on Clinton’s statement, which was taken out of context, to built his own campaign for coal and therewith to secure a significant number of votes.

In this paper we investigate the effect of this campaign pledge on the Republican’s vote share on the county level. In our analysis we study the relative impact of the electoral promise on areas more or less ‘exposed’ to coal mining, and we use the coal production in a county as the main predictor. Our aim is to answer to the question of whether the electoral consent for Republicans increased more (or less) in counties that were characterized by coal extraction on a larger scale. We model electoral outcomes in a reduced-form, where the share of the votes obtained by the Republican party depends on the economic and institutional characteristics of the counties. To the best of our knowledge this paper is the first to focus on the impact of coal production in a county (measured in short tons) on the political outcome. Furthermore, we are considering spillover effects and apply spatial clustering to avoid biased estimates due to trade, migration and information flows between counties.

In the empirical political economy literature there is extensive debate on the impact of economic conditions on presidential voting (Lewis-Beck & Stegmaier (2000), Besley & Case (2003)). In general, high unemployment and difficult economic conditions benefit Democratic candidates (Rees et al. (1962), Wright (2012), Burden & Wichowsky (2014)). At the same time, economic shocks, such as rising import competition or energy transition are found to explain ideological polarization, expanding support for both far-left and far-right views (Autor et al. (2020)). A large literature empirically investigates the rise of populist parties in many high-income countries, and many authors find that economic insecurity, financial distress and low income are among the driving forces of the increasing support for ‘populist’ policies (Acemoglu et al. (2013), Guiso et al. (2017)).

We contribute to this literature, empirically establishing the role of economic distress on presidential vote, focusing on the coal industry, which represents an excellent case study. The U.S coal production is mainly concentrated in two large regions. In the eastern Appalachian

¹https://www.washingtonpost.com/video/politics/trump-receives-warm-welcome-in-coal-country/2016/05/06/9259c5ea-1327-11e6-a9b5-bf703a5a7191_video.html
²https://www.npr.org/2016/05/03/476486650/fact-check-hillary-clinton-and-coal-jobs?t=1652451309596
region (mainly Alabama, Pennsylvania, Kentucky, Virginia, and West Virginia) mining is underground and labour intensive, while the Western Powder River Basin region (mainly North Dakota, Wyoming, and Montana) is characterized by surface mining which is less labour intensive. The U.S. coal production fell by one third between 2011 and 2016, and the impact on employment was even more dramatic: from 130,000 workers in 2011 to less than 70,000 in 2016 (Houser et al. (2017)). The coal industry collapse had also downstream effects on whole communities where coal companies are located, reducing employment, wealth, tax revenues and finally resulting in service cuts. These communities represent an interesting quasi-natural experiment, as they were differently exposed to the coal industry crisis, depending on the different degree of economic dependence on coal production. Our aim is precisely to exploit these differences in exposure to the coal industry collapse to estimate the effect on presidential vote outcome.

We find a positive effect of coal production in a county on the Republican vote share. If the coal output in a county rises by an additional 1 mill. short tons (approximately corresponding to one standard deviation), the vote share of Donald Trump significantly increases by 0.059-0.095 percentage points.

To estimate the populist effect of Donald Trump we follow Goetz et al. (2019) and substitute in our model the Republican’s vote share in 2016 with the difference of the vote share of Donald Trump in 2016 and Mitt Romney in 2012. We learn that Donald Trump receives disproportionately more votes in the Midwest Counties and the Rust Belt. In this populist model the effect of an additional 1 mill. short tons results in an significant increase of the Republican’s vote share by 0.080-0.123 percentage points.

To test the validity of our estimates we apply several robustness checks. First, we apply the inverse-distance not just for estimation but also for spatial clustering. Second, we substitute coal output in short tons with some binary variables to examines non-linear effects of coal production. Last, we substituted the overall coal output of a county with the output per employed worker and with the output per working hour. Of course, the different specifications lead to slightly different results. The positive relationship between coal production and Trump’s electoral share remains, however.

The paper is organized as follows. Next section describes the U.S coal industry, while section 3 details the main variables introduced in the empirical analysis, and the data sources. In section 4 we present our empirical strategy, and identification issues. In section 5 we present our results, apply robustness checks and discuss our findings in detail. Section 6 concludes.
2. The U.S. coal industry

This section documents a few facts about the U.S. coal industry. Employment in the U.S. coal industry declined since decades with a slight increase in the 2000s, as illustrated in Figure 1. From a peak in June 1985 with 177.8 thousand employed miners to 49.6 thousand just before the Presidential election in October 2016. After taking office in the White House nothing has changed significantly in terms of employment in the coal industry. During the first three years of the Trump Administration the decline in employment stalled, only to continue to decline a bit further at the current edge.

![Number of workers in coal mining industry from 1985](https://fred.stlouisfed.org/series/CES1021210001)

Figure 1: *Seasonally adjusted number of employees in coal mining industry since 1985*

*Data Source: FRED*[^3]

The decline in employment goes along with a decline of coal consumption in the U.S. that was fueled by the rise of fracking. The consumption of coal went up over decades to the peak level in 2005 with a consumption of 22.8 Quadrillion Btu (equivalent to 1.2 billion short tons).[^4] Since then the consumption of coal is declining in the U.S. This is mainly because consumption of coal as source of electric energy has been declining (see Figure 2). Davis *et al.* (2021) point out that the coal based electric generating capacity decreased even further since 2011 because of more severe environmental regulations, increased use of renewable energies and a lower price for natural gas as well as lower driving peak electricity prices.

Together with the coal consumption the coal production went down. In fact, in 2020 coal consumption decreased to 7.7 Quadrillion Btu (equivalent to 0.45 billion short tons).

[^3]: https://fred.stlouisfed.org/series/CES1021210001
[^5]: https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T06.02#/?f=M&start=197301&end=202104&charted=1-5-12-13-14-15
production in the U.S. fell under the level of 1985, as displayed in Figure 3. As mentioned before, the lower prices for natural gas led to higher demand for natural gas and correspondingly to less demand for coal, in the U.S. and internationally. The COVID-19 pandemic did also contribute to the most recent decline in 2020. U.S. coal mines temporarily shut down to prevent further spread of the Coronavirus. Consequently, U.S. coal exports decreased by 26% in 2020 compared to 2019.
The Trump Administration did not stop the decline of the US coal industry but one has to admit it was not because of a lack of trying. Bloomberg reports that the Trump Administration spent over 1 billion US dollar for the coal industry in its legislative period. Environmental rules have also been relaxed and attempts have been made to keep power plants from being closed. The attempt to revive the US coal industry was unsuccessful, but was Donald prominent election pledge in 2016 with the well-known slogan 'Trump digs coal' successful? As mentioned above, in this paper we analyse the effect of this pledge on the ballot outcome of the 2016 Presidential election at the county level.

3. Data

The dependent variable is the Republican’s percentages of votes. Data on the outcomes of the 2016’s ballot is sourced from Harvard Dataverse, providing election data from 2000 collected by the MIT Election Data and Science Lab. We compute the Republican’s percentage of votes by dividing the total number of votes for the Republicans by the total number of votes and multiplying with 100.

The main regressor of interest, is coal output. We obtain data about coal output, coal employment and hours worked of active and inactive, surface and underground mines and preparation plants from the U.S. Energy Information Administration (EIA). County-level coal production is calculated as follows. For each county, coal output is summed across 'active', 'active, men working, not producing', 'permanently abandoned' and 'temporarily closed' plants, that operate at least a mine only or a mine and a preparation plant in 2016. We also include inactive plants ('permanently abandoned' and 'temporarily closed'), for three main reasons. First, they have produced a non-negative output quantity and have employed non-negative input amounts, but have closed in 2016. Second, according to the EIA, some mines have been defined as 'permanently abandoned' by mistake. Third, counties in which relevant plants are located may have believed Trump’s campaign pledge of spurring coal production in the US. Given Trump’s election pledge and the fact that coal regions appear to have an important role in Donald Trump’s victory, a positive effect is expected (Goetz et al. (2019)). The impact of spillovers from neighbours is ambiguous. On the one hand, economic benefits from trade support Donald Trump and the Republican’s. Nevertheless, gains from trade are not that large given the quite simplistic supply chains characterizing the coal industry, since the most important buyer of coal is the electricity sector (see Figure 2). On the other hand, environmental disadvantages (e.g. emission, pollution, inhabitants of a given county might fear that a coal mine or power plant is

---

6https://www.eia.gov/todayinenergy/detail.php?id=48696
8https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/VQCHQ
9https://www.eia.gov/coal/data.php#production
opened in the county they are living) will compensate economic benefits to some degree, particularly in closer regions. Moreover, coal mining counties are often neighbours to each other suggesting that a given coal county does not really care whether its neighbour also hosts a coal mine.

Following the literature (e.g. Steinmayr (2021)), we also control for the Republican’s percentage of votes of the previous ballot in 2012 (Obama vs. Romney) accounting for the county’s general ideological preference, for which we expect a positive impact. The variable is constructed from the same data as the Republican’s percentage of votes in 2016 in the analogous way. It’s spillover effect should be negative, because the Republicans are more successful in the fly-over states located closely to the geographical center of the US. The farther away from the center, the less powerful will be the Republicans.

To control for county-level economic conditions, we involve a set of selected controls.

First, we include the unemployment rate in percentage points (Rodríguez-Pose et al. (2021), Steinmayr (2021), Halla et al. (2017), Madestam et al. (2013)). Given the literature, its effect is ambiguous. On the one hand, voters living in counties with high unemployment rates might prefer the democrats, as they expect them to keep welfare programs in place. On the other hand, the same voters might favor Trump, since they have a greater confidence that he would be able to restore present jobs and create new ones (Goetz et al. (2019)). Furthermore, populist parties benefit from high unemployment rates (Algan et al. (2017)). Spillovers from neighbouring counties might impact the dependent variable in a given county negatively, because higher unemployment rates signals a weak economic performance in an entire region. A widespread poor economic performance, however, may help the Democrats that are more prone to fight such crisis with expansive fiscal policy. Data is provided by the US Department for Agriculture.\footnote{https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/}

As not stated otherwise all the following variables are sourced from the US Census.\footnote{https://www.census.gov/data.html}

Second, the share of workers in manufacturing over the total numbers of workers, in percentage points, is introduced (Steinmayr (2021), Autor et al. (2020), Ochsner & Roesel (2020), Goetz et al. (2019), Halla et al. (2017)).\footnote{While annual data is not available, we use the five-year estimates by the US Census. https://www.census.gov/data.html} Its impact is ambiguous. While this group is still an important target group of the democrats, the phenomenon that blue-collar workers tend to vote right-wing extremist is observed in many European countries (e.g. Austria, Germany, Italy, Sweden) (Rydgren & Tyrberg (2020), Adorf (2018), Stockemer et al. (2018)). Furthermore, instead of representing US workers, Clinton was perceived as a Wall Street representative. These arguments would favor a positive impact of this variable on the dependent variable, while, on the other hand, Goetz et al. (2019) estimate a significantly negative impact of the share of this variable on Trump’s percentage of votes in the 2016’s ballot. Furthermore, Goetz et al. (2019) also conclude that this variable significantly raises the margin between Trump’s outcome in 2016 and Romney’s one in 2012. The effect of its spillover is ambiguous. On the one hand,
a given county may benefit from spread effects due to the positive impacts of manufacturing and jobs in neighbouring counties, while, on the other hand, booming counties attract young, highly educated and possibly the Democrats favouring workers (e.g. engineers) from other counties. Moreover, booming counties also attract firms suggesting negative backwash effects on the dependent variable.

Third, we control for poverty (Goetz et al. (2019)) and introduce the share of households living below the poverty line, in percentage points.\textsuperscript{13} A negative effect on our dependent variable is likely, as Republicans have usually opposed build-ups of the welfare state. Its spillover should impact the dependent variable negatively for the same reasons as for the spillover of unemployment.\textsuperscript{14}

Fourth, we include the share of people benefiting from either public or private social insurance over the total number of inhabitants (Goetz et al. (2019)).\textsuperscript{15} A positive effect is expected for two reasons. First, the higher the income in a given county, the more people can afford social insurance. Richer counties, on the other hand, might have a higher tendency to vote for the Republicans. Second, the larger the share of inhabitants benefitting from either private or public social insurance, the smaller is its complement (the share of people lacking social insurance) that may prefer the democrats to benefit from Medicaid or Medicare, suggesting a positive impact on Trump’s percentage of votes. As booming counties attract high-skilled workers, the spillover is expected to have a negative influence. As a last economic control variable we include the growth rate of import penetration from China on commuting zone level by Autor et al. (2020). As Donald Trump used a lot of anti-China rhetoric during his campaign we expect the effect to be positive. The spillover effect might be negative as a high import penetration from China leads to a weak economic performance.

We additionally include controls for socio-demographic characteristics. First, we include the share of females in the total population (Rodríguez-Pose et al. (2021), Steinmayr (2021), Autor et al. (2020), Goetz et al. (2019), Halla et al. (2017)).\textsuperscript{16} The variable is constructed by dividing the number of women by the total number of inhabitants. Due to sexual harassments allegations against Donald Trump and his pejorative reactions (“Grab’em by the pussy!”), we expect a negative impact. Furthermore, in many countries, women have a stronger tendency to vote left-wing (Goetz et al. (2019)). The spillover effect is ambiguous. As females are more mobile than males, a higher share of females signals better economic performance which would result in a positive spillover effect. But as mentioned, as females vote more anti-Trump the spillover effect could also be negative.

Next, the share of black households over the total number of households is introduced (Rodríguez-Pose et al. (2021), Goetz et al. (2019), Madestam et al. (2013)).\textsuperscript{17} Due to Trump’s several

\textsuperscript{13} Again, we use five-year estimates by the US Census as annual data is not available.\url{https://www.census.gov/data.html}

\textsuperscript{14} Five-year estimates are sourced from the same source.

\textsuperscript{15} Five-year estimates are sourced from the same source.\url{https://www.census.gov/data.html}

\textsuperscript{16} Annual population estimates for every county are provided by the US Census. \url{https://www.census.gov/data.html}

\textsuperscript{17} Five-year estimates are collected from the same database. \url{https://www.census.gov/data.html}
racist comments and the fact that blacks represent an important part of the Democrat’s electorate, the variable should decrease the Republican’s percentages of votes (Goetz et al. (2019)). Its spillover might have a positive effect, as voters might feel estranged, when observing growing shares of African-American in neighbouring counties and, therefore, have a stronger tendency to vote for Trump.

For the same reason, we involve the share of latinos over the total number of inhabitants (Autor et al. (2020), Goetz et al. (2019), Madestam et al. (2013)). We calculate the variable from the same data in the analogous way as the share of females and expect a negative impact (Goetz et al. (2019)). As for the share of blacks, a positive effect of its spillover is expected.

Fourth, the share of adults with a bachelor degree or more over the total number of adults controls for the county’s education level. Generally, more educated people vote more strongly for the democrats, suggesting a negative impact as well (Rodríguez-Pose et al. (2021), Steinmayr (2021), Autor et al. (2020), Goetz et al. (2019), Scala & Johnson (2017), Barone et al. (2016), Mendez & Cutillas (2014)). Furthermore, universities attract particular groups of workers that might not only live in the same county the university is located, but also in neighbouring counties, affecting ballot outcomes.18

Since young people represent an important target group of the democrats, we expect a negative effect of this variable controlling for the county’s shape of the age distribution (Rodríguez-Pose et al. (2021), Steinmayr (2021), Autor et al. (2020), Halla et al. (2017), Mendez & Cutillas (2014)). Since young people are generally more mobile than older generations (e.g. study, work) and communicate their ideas and views, spillovers might negatively effect Trump’s percentage of votes. The variable is constructed from the annual population estimates of the US Census by summing the number of people aged up to 30 and dividing it by the total number of inhabitants.

For the same reason, we incorporate the share of persons aged above 60 years in the total number of inhabitants. Similarly to the share of young people, the variable is constructed from the annual population estimates of the US Census. Unlike Goetz et al. (2019) we expect a positive effect on the Republicans’ share of votes, as older generations tend to vote for and back the Republican party more strongly (Center (2018)). Spillover effects are ambiguous, as older generations are usually less mobile than younger generations suggesting insignificant spillovers.

Last, we control for the quality of public infrastructure and urbanization by introducing the popularity of public means of transport. We calculate the share of workers (aged above 16) going to work by public means of transport over the total number of workers (aged above 16) who do not work at home.19 This variable also captures urbanization, suggesting a negative impact, as cities represent an important part of the democrat’s electorate. The effect of its spillover is ambiguous. On the one hand, spillovers might have a positive effect if local public transport networks are connected loosely with each other, implying frustration and envy. On the other hand, inhabitants of neighbouring counties see the benefits of a good public infrastructure when working there. On the whole we expect a positive impact. Big cities characterized by good

18 Five-year estimates are sourced from the same source https://www.census.gov/data.html.

19 Five-year estimates are sourced from the same source https://www.census.gov/data.html.
public infrastructure are mostly located at the East and West coast. The Republicans, however, have received more votes in counties that are located far away from the coasts (e.g. fly-over states), suggesting a positive impact that increases with the distance cutoff.

As in Monnat & Brown (2017), we exclude Alaska due to the lack of election data for those counties. Furthermore, we exclude Campbell/Wyoming (FIPS: 56005), because this county is an extreme outlier in terms of coal output. To include the change in commuting zone-specific import penetrations by China, Hawaii is also dropped, since Autor et al. (2020) do not cover this state. As in other studies (e.g. Rodríguez-Pose et al. (2021), Goetz et al. (2019)), all the covariates are of the year 2016. The one exception is the data from Autor et al. (2020) as the focus on the change in import competition.

Summary statistics are shown in Table 1. The entire county-level data covers 3,107 observations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unit</th>
<th>Mean (SD)</th>
<th>Min-Max (Median)</th>
<th>RQR (CV)</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share Republican 2016</td>
<td>Percentage Points</td>
<td>63.28 (5.67)</td>
<td>0.00 - 66.94</td>
<td>43.64 (20.85)</td>
<td>NA</td>
</tr>
<tr>
<td>Share Republican 2012</td>
<td>Percentage Points</td>
<td>59.6 (4.81)</td>
<td>0.00 - 66.75</td>
<td>39.28 (20.85)</td>
<td>3.65</td>
</tr>
<tr>
<td>Difference Percentages of Votes between 2016</td>
<td>Percentage Points</td>
<td>3.67 (5.21)</td>
<td>47.93 - 347.23</td>
<td>6.08 (1.15)</td>
<td>NA</td>
</tr>
<tr>
<td>Share Manufacturing</td>
<td>Percentage Points</td>
<td>23.21 (7.37)</td>
<td>0.00 - 45.0</td>
<td>30.63 (9.28)</td>
<td>2.43</td>
</tr>
<tr>
<td>Share Unemployment Rate</td>
<td>Percentage Points</td>
<td>5.21 (1.81)</td>
<td>0.00 - 9.02</td>
<td>6.08 (1.15)</td>
<td>2.62</td>
</tr>
<tr>
<td>Share Poverty</td>
<td>Percentage Points</td>
<td>15.01 (6.27)</td>
<td>0.00 - 34.14</td>
<td>37.28 (9.28)</td>
<td>4.89</td>
</tr>
<tr>
<td>Share Insurance</td>
<td>Percentage Points</td>
<td>87.02 (11.11)</td>
<td>31.41 - 98.87</td>
<td>66.6 (6.06)</td>
<td>3.26</td>
</tr>
<tr>
<td>Growth Rate Import Penetration from China</td>
<td>Percentage Points</td>
<td>0.26 (0.68)</td>
<td>0.26 - 0.63</td>
<td>0.84 (0.85)</td>
<td>1.50</td>
</tr>
<tr>
<td>Share Female</td>
<td>Percentage Points</td>
<td>40.01 (22.21)</td>
<td>0.00 - 90.16</td>
<td>71.38 (9.04)</td>
<td>1.50</td>
</tr>
<tr>
<td>Share Black</td>
<td>Percentage Points</td>
<td>9.5 (1.87)</td>
<td>0.00 - 23.47</td>
<td>9.79 (1.45)</td>
<td>4.36</td>
</tr>
<tr>
<td>Share Latino</td>
<td>Percentage Points</td>
<td>39.15 (13.37)</td>
<td>0.00 - 62.64</td>
<td>37.28 (9.28)</td>
<td>3.31</td>
</tr>
<tr>
<td>Share Education</td>
<td>Percentage Points</td>
<td>21.96 (9.41)</td>
<td>0.00 - 62.92</td>
<td>10.02 (9.44)</td>
<td>3.29</td>
</tr>
<tr>
<td>Share Young</td>
<td>Percentage Points</td>
<td>75.7 (5.28)</td>
<td>10.01 - 90.67</td>
<td>7.84 (0.84)</td>
<td>7.08</td>
</tr>
<tr>
<td>Share Old</td>
<td>Percentage Points</td>
<td>21.25 (5.01)</td>
<td>0.00 - 34.05</td>
<td>6.7 (0.23)</td>
<td>6.08</td>
</tr>
<tr>
<td>Share Public Transport</td>
<td>Percentage Points</td>
<td>1.25 (0.25)</td>
<td>0.00 - 0.25</td>
<td>0.83 (0.27)</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Table 1: Descriptive statistics

Figures 4 and 5 show maps of Donald Trump’s and Mitt Romney’s percentages of votes of the 2016’s ballot. The higher these shares, the brighter is the county’s colour. Furthermore, the map at the bottom illustrates the same variable for Hawaiian counties. Coal producing counties are framed green. The Democrats received higher percentages of votes in the coastal regions where bigger cities are located, whereas Trump was successful in the Midwest, Mid-Atlantic, South East, and the eastern parts of the West, North and South West. Particularly, states in the Rustbelt, the coal and industrial region (e.g. Michigan, Ohio, Pennsylvania), were important epicentres of Trump’s victory. Figure 6 displays the difference in the Republican’s percentages of votes between 2016 and 2012. In the relevant counties, Trump was even more successful.

---

20 Election data is only available for districts that are a combination of multiple counties (boroughs).

21 In this county, 257.54 mio shorttons are produced. In comparison, the second highest value is 29.79 shorttons.

When including this observation, the results do not change, as almost the same coefficients are observed.
than Romney in 2012, suggesting that Trump’s campaign pledges to be an effective tool to gain votes in counties suffering from the declines in manufacturing and coal mining. Furthermore, the Republicans won the election in many other coal producing counties located in Wyoming, Illinois and the Appalachians (e.g. Kentucky, Pennsylvania, West Virginia). On the other hand, Romney received higher percentages in Arizona and Utah and Arizona benefiting from the disproportionally higher Mormon votes (Goetz et al. (2019)).

4. Empirical strategy

4.1. Model specification

Trump promise to revive the U.S. coal industry was an exogenous event, easily understood by the electorate and heavily publicized, especially during the last months of the election campaign. We use a cross-section county-based research design where counties whose economies are more coal-dependent are compared to those whose economies are unaffected by the coal production. In particular, we exploit variation in the coal industry size across counties and the timing of elections (2012 and 2016). We use a measure of country coal ‘exposure’ as the annual coal production. The variation in coal production will generate differential response of the counties’ electoral choices to the exogenous electoral promise. Our baseline specification takes the following form:

\[ r_{2016} = \alpha \text{Coal} + \gamma r_{2012} + X \beta + \text{state} + \epsilon \]  

where \( r_{2016} \) is the Republican’s percentage of votes in 2016, \( \text{Coal} \) is the coal output in millions of shorttons, and \( r_{2012} \) is the Republican’s percentage of votes in 2012. \( X \) denotes a \( N \times K \) matrix, containing county-level controls, while \( \text{state} \) denotes dummies for the state of the county, and \( \epsilon \) defines the error term. The coefficient of interest is \( \alpha \) that captures the average effect of coal production on county electoral outcomes. The inclusion of state fixed effects controls for unobserved state-specific heterogeneity, i.e. state-specific preferences, policies (e.g. states differ in legislations on early and postal voting), and other characteristics (e.g. swing states).

To establish a link between coal production and the Republican’s percentage of votes at the county-level, we also consider spillover effects and spatial clustering (Scala & Johnson (2017)). Equation (1) implicitly assumes that the stable unit treatment value assumption (SUTVA) is satisfied, but such an approach neglects trade, migration and information flows between
Figure 4: Republican’s percentage of votes in the 2016’s ballot Trump vs. Clinton
Figure 5: Republican’s percentage of votes in the 2012’s ballot Romney vs. Obama
Figure 6: Difference in Republican’s percentage of votes between 2016 and 2012
counties, resulting in biased estimates. Therefore, we also estimate a spatial Durbin model, as shown in equation (2):

\[
\begin{align*}
r_{2016} &= \alpha_{\text{Coal}} + \rho_1 W_{\text{coal}} \text{Coal} + \gamma r_{2012} + \rho_2 W_r r_{2012} + X \beta + \rho_3 W_X X + \text{state} + \epsilon \\
\text{with } \epsilon &= \lambda W_i \epsilon + \eta
\end{align*}
\]

(2)

where \(W_{\text{coal}}, W_X, \text{and } W_i\) define the spatial weight matrix for the coal production, the covariates, and the error term, respectively. \(\epsilon\) defines the error term, consisting of a spatially correlated term and an independent, but heteroskedastically distributed innovation \(\eta\).

To verify robustness, we employ a large set of specifications, that mainly differ in the specification of the weighting matrices. First, we use a binary adjacency matrix. In this matrix, a link between the counties \(i\) and \(j\) equals one, if county \(j\) shares a border with county \(i\) in at least one point, and is zero otherwise (queening). The matrix is constructed from shapefiles obtained from the US Census Bureau.\(^{22}\) The same binary weighting matrix \(W_i\) applies to the errors.

Second, we introduce several inverse-distance matrices characterized by a distance decay (Basile (2009), Dall’erba & Le Gallo (2008), Ertur et al. (2006), Pede et al. (2007)). Circle distances between every pair of counties are computed from the centroids of each polygon. The further away county \(j\) is from \(i\), the smaller is the influence (weight) of county \(j\) on \(i\). To only consider links between counties located closely to each other, we follow the literature (e.g. Basile (2009), Dall’erba & Le Gallo (2008), Ertur et al. (2006), Pede et al. (2007)) and introduce several cutoffs starting from 200 km up to 750 km (first quarter of distances) to check the robustness of the results.\(^{23}\) In other words, circles with a given radius are drawn around every counties’ centroid defining the area for which spillovers across counties are expected. If the distance exceeds this threshold, i.e. county \(j\)’s centroid is located outside this circle around county \(i\)’s centroid, it is assumed that county \(j\) does not influence county \(i\) and, thus, its weight is zero, \(w_{i,j} = 0\). On the other hand, if county \(j\)’s centroid is located inside this circle, county \(j\) is assumed to impact county \(i\) and, hence, gets a weight based on the inverse distance \(w_{i,j} = \frac{1}{d_{i,j}}\).

For the error terms, we impose \(w_{\epsilon,i,j} = 1\), if county \(j\)’s centroid is located within the circle around county \(i\)’s centroid, and zero otherwise. Additionally, we also introduce the inverse distance matrices to check robustness.

All matrices are row-normalized implying that the effect of county \(i\) on the other \(-i\) counties decreases with the number of neighbours. The normalization also facilitates the interpretation, as \(W_{\text{coal}} \text{Coal}, W_X X\) are interpretable as a distance weighted average (Weiss et al. (2015)).

\(^{22}\)https://www2.census.gov/geo/tiger/TIGER2016/COUNTY/?sec_ak_reference=18.e0fd717.1515267074.5aed87d

\(^{23}\)We also estimate the same equations using 50, 100 and 150 km as cutoffs. Nevertheless, there are some islands that are excluded from weighting matrix and the regressions. The results barely change.
The coefficients $\alpha$ and $\beta$ quantify the effect of coal and other covariates in county $i$ on the dependent variable in the same county, while the coefficients $\rho_1$, $\rho_2$ and $\rho_3$ measure the degree to which the given county’s dependent variable is influenced by its neighbouring coal production and other covariates’ value.

For the variable of interest, we expect a positive impact due to the strong economic benefits, while for the spillover the effect is ambiguous. On the one hand, closer neighbours benefit from more employment and economic growth, whereas, on the other hand, they will also suffer from pollution. Furthermore, more liberal Republican voters might generally oppose the support of coal mining. The farer away the neighbours, the weaker are the effects on economic prosperity and pollution. For medium cutoffs, we, however expect positive coefficients, as pollution merely implies regional disadvantages that are exceeded by economic advantages. For larger thresholds, advantages and disadvantages may compensate each other again, as more distant counties neither benefit nor lose substantially from coal mining in a given county.

Other variables such as average and median household income, the relatives sizes of other age cohorts and race groups, and employment shares of other industries are not included due to the strong multicollinearity as suggested by high bivariate correlation coefficients and variance inflation factors. In comparison to Goetz et al. (2019), the voter turnout is excluded, as it can be classified as another outcome of ballot, suggesting it to be a bad control in the sense of Angrist & Pischke (2009).

In a spatial Durbin model, only spatial spillovers of the covariates and/or the errors are introduced but no spatial lag of the dependent variable. There are two main reasons for this modelling. First, strategic interactions and coordination between counties on ballot outcomes is unlikely, suggesting the exclusion of the spatial lag of the dependent variable. Second, the characteristics of one county plausibly influence the Republican’s vote share in other counties via trade and migration flows. For instance, coal production in a given county does not only create jobs in the same county, but also in neighbouring ones.

Following Rodriguez-Pose et al. (2021) and Goetz et al. (2019), we also introduce an alternative specification, where the dependent variable is the difference between the percentages of all votes received by Donald Trump in 2016 and Mitt Romney in 2012:

$$r_{2016} - r_{2012} = \alpha_{Coal} + \rho_1 W_{coal} Coal + X \beta + \rho_2 W_{X} X + state + \epsilon$$

with $\epsilon = \lambda W_{\epsilon} \epsilon + \eta$ (3)

In this specification, the coefficients quantify the extent by how much more or less Donald Trump appeals to voters in a county, when the value of a given covariate rises by one unit, beyond just being the Republican’s candidate. Results are robust when using a different reference election.
4.2. Identification

Concerning identification, there are some issues to mention. First, the ‘reflection problem’ outlined by Manski (1993) and discussed by Pinske & Slade (2010) and Gibbons & Overman (2012) is generally difficult to assess. The question is whether correlation between the Republican’s percentages of votes of neighbouring counties is either caused by direct correlation with the percentages of votes or caused by correlation with the characteristics of neighbouring counties, implying an indirect correlation through percentages of votes, or both (Weiss et al. (2015)). We believe that there is no strategic interaction in voting behaviour among people located in neighbouring counties, but characteristics of a given county influence its neighbours via trade and migration flows, suggesting a spatial Durbin model.

Second, correlation in Republican’s percentages of votes can be caused by spatial correlation between Republican’s percentages of votes and by spatial clustering of the residuals. To disentangle the effects, we follow Weiss et al. (2015) by taking account of spatial clustering which is possible if the model is correctly specified, in particular the spatial weight matrices. Although we are providing arguments favouring our specification and the decision on the design of the spatial weight matrices, it is not possible to test whether the underlying assumptions are correct.

Third, the coefficients are not likely to suffer from inconsistency stemming from simultaneities or omitted variable biases. At a first glance, the coefficient of interest could be prone to simultaneities. Nevertheless, simultaneities are unlikely, since winning votes via promoting coal mining and supporting mine operators to relocate back is a difficult political endeavour due to the strict environmental legislation, opposition from residents, high wage costs and the long construction times required to relocate and buildup capacities and production. Besides, more feasible alternatives are available to win votes in the relevant counties. Hence, in the short-run, coal output can reasonably classified as an exogenous variable. Furthermore, the previous government of Barack Obama had little interest in coal counties, being Republican strongholds that are located in regions important for the Republican party (e.g. Midwest, South-West, Central-East and South-East). Additionally, omitted variable biases are overcome by including a large set of controls and state dummies.

Fourth, the sample is incomplete due to the lack of election data for Alaskan counties. Kelejian & Prucha (2010), however, show that the GLS-2SLS-GMM estimator stays asymptotically normal and consistent, if the number of missing observations in the dependent variable is not too large. As there is only one coal producing county in Alaska, this issue can be neglected.
5. Results and discussion

5.1. Results of the core models

5.1.1. Explaining the Republican’s percentages of votes

Table 2 provides the results of the spatial Durbin regressions of equations (1) and (2). In every regression, the dependent variable is the county-level Republican’s percentage of votes. Column (1) displays the estimates of the simple OLS model with standard errors clustered at the state-level excluding spatial spillovers. Column (2) shows the estimates of the regression using the adjacency matrix obtained by queening for spatial spillovers of the covariates and the residuals, while columns (3)-(14) provide the analogous of the regressions using the inverse-distance matrices and their binary pendants for the spatial spillovers of the covariates and residuals.

Quantifying the effect of coal output, as measured in mill. shorttons, on the Republican’s percentage of votes, the coefficients are interpreted as follows. If coal output in county \( i \) rises by 1 mill. shorttons, the Republican’s percentage of votes significantly increases by \( 0.064-0.101 \) percentage points. Row-normalization of the spatial weighting matrices allows to interpret the spillovers’ coefficients as the effects of changes in weighted averages, i.e.: when the distance weighted average of coal production of county \( i \)'s neighbours located within the circle around its centroid is raised by 1 mill. shorttons, the dependent variable in county \( i \) changes by \( -1.091-2.049 \) percentage points. Spillovers, however, only significantly affect the dependent variable in three out of 14 models, suggesting that benefits and disadvantages from pollution compensate each other (the relatively small gains from the trade given the simplistic supply chains do not sufficiently exceed environmental disadvantages) within given regions (up to 350 km, from 550 to 750 km), but then benefits exceed the latter (from 400 to 500 km). These spillover pattern are reasonable due to two further reasons. First, the supply chain of the coal industry is quite simplistic one. The coal is mined and than transported to a coal-fired power plant. As shown before the electricity industry is by far the most important purchaser of coal. Given the simplistic supply chains, economic benefits from trade might not be that large. Second, neighbouring coal counties constitute a coal region (Appalachian). If neighbour \( j \) produces coal, this fact might be irrelevant to county \( i \), if it also produces coal because there will be no trade and the pollution in county \( i \) comes also and foremost from the county’s mines themselves. That will make spillovers within coal reagions weaker. But overall, coal producing counties indeed show a stronger tendency to vote for Trump than other counties, confirming the hypothesis stating that Trump has been more successful in this counties due to his election pledges.

As expected, the Republican’s percentage of votes of the 2012’s ballot significantly increases the same party’s outcome of 2016, suggesting some degree of persistence of preferences. In comparison, the spillover is mostly significantly negative, implying that a lower share of voters
in county $i$ votes for Trump, when the weighted average of its neighbours rises by one percentage point. As can be seen from Figures 4 and 5, the Republicans have been successful in the fly-over states and the Democrats in the coastal states and border regions. Hence, the farther away from the country’s geographical center, the lower is the Republican’s share of votes resulting in a negative effect. The negative impact will be stronger for the Western part of US in which counties tend to be larger.

Next, the unemployment rate does not significantly affect Trump’s percentage of votes, though showing the assumed sign. The same is true for its spillover. Plausibly, a more dominant manufacturing industry significantly positively impacts the dependent variable, as Trump’s agenda, compared to the one of Clinton, has been more pro-business (e.g. the pledge of reducing corporate taxes). On the other hand, the spillover’s impact is significantly negative suggesting that neighbouring counties’ booming manufacturing sectors cause losses of income, unemployment and poverty supporting the Democrats.

On the other hand, poverty significantly drops the Republican’s share of votes. Its spillover is significantly negative as well, because high poverty rates in neighbouring counties might raise worries and anxiety in county $i$ spurring its inhabitants to vote for the democrats due to the Republican’s restrictive social policy.

In comparison, the share of insured people significantly raises Trump’s share of voters, as a larger share coincides with a robust economics development allowing employees and employers to purchase more social insurance. The effect of its spillover only significantly differs from zero in three regressions.

For the growth rate of the import penetration we find a positive effect on Donald Trump’s vote share in almost every model, except for the simple OLS model. The spillover effect is not uniform but never significant.

As expected, females, blacks, latinos, better-educated and younger people are characterized by a weaker tendency to vote for Trump, which is confirmed by the findings. While the spillover of females does not significantly affect the dependent variable, the one of the share of black households significantly raises it in five models, suggesting that voters of a given county might be afraid of a rising share of blacks at the cost of the share of whites. On the other hand, the spillover of the share of latinos significantly decreases the dependent variable only in models employing a larger cutoff. In comparison, the share of people with a bachelor degree or more and the share of young people are always significantly negative, because well-educated and young people are more mobile and keen on social media.

Confirming the generation gap in American politics, older generations back the Republicans more strongly than younger generations. As expected spillover effects do not significantly deviate from zero.

Besides, the share of workers aged above 16 travelling to work by public means of transport significantly decreases Trump’s share of votes in four models, as quality of living tends to be higher in relevant counties implying a smaller pool of frustrated voters prone to Trump. Besides, the spillover is significantly positive and rises with the cutoff, suggesting that counties
that are located farer away from the coasts where the big cities with good public transport systems are vote more intensively for the Republicans.

Last, the spillovers of the residuals are significantly positive. Thus, a positive shock to the dependent variable is likely to affect the outcomes of neighbouring counties in the similar way, because the same or similar shocks might also affect them. Second, the coefficient’s magnitude decreases with the size of the cutoff.

<table>
<thead>
<tr>
<th>15K</th>
<th>20K</th>
<th>25K</th>
<th>30K</th>
<th>35K</th>
<th>40K</th>
<th>45K</th>
<th>50K</th>
<th>55K</th>
<th>60K</th>
<th>65K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
<td>1.6</td>
<td>1.7</td>
<td>1.8</td>
<td>1.9</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Table 5.1.2: Examination of populism effects

<table>
<thead>
<tr>
<th>Table 5.1.2: Examination of populism effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal output (300,000)</td>
</tr>
<tr>
<td>Share Education</td>
</tr>
<tr>
<td>Impact Penetration</td>
</tr>
<tr>
<td>Share Manufacturing</td>
</tr>
<tr>
<td>Share Old Town</td>
</tr>
<tr>
<td>Share Female</td>
</tr>
<tr>
<td>Share Filer</td>
</tr>
<tr>
<td>Coal Output (300,000)</td>
</tr>
<tr>
<td>Notes:</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>-</td>
</tr>
</tbody>
</table>

5.1.2. Examination of populism effects

When estimating the populist equation, the results, as shown in Table 3, are generally robust, though the interpretation changes. Concerning the coefficient of interest, Donald Trump has been more successful in 2016 than Romney in 2012 in a given county, the more coal is produced. When coal output rises by 1 mill. shorttons, the difference in percentages significantly increases.

Table 2: Spatial Durbin regressions of Republican’s percentage of votes

<table>
<thead>
<tr>
<th>Table 2: Spatial Durbin regressions of Republican’s percentage of votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>203</td>
</tr>
</tbody>
</table>

Note: The data used is from the Pennsylvania’s previous vote in all jurisdictions. All other factors are, by and large, taken from the Pennsylvania’s election and have been for the past election. This is due to the use of this data set being the Pennsylvania’s election results from the Pennsylvania’s election.
0.080-0.124 percentage points. In other words, Trump’s campaign pledge indeed played a role, as inhabitants may have believed that Trump, as a business man, might be able to keep his promise. Spillovers, however, turn insignificant. Hence, the output of neighbouring counties has no significant impact on the outcome in a given county. Trump has also focused more intensively on the coal counties themselves and not that much on their neighbours. Besides, the simplistic supply chains and the fact that coal counties are clustered also decrease spillover effects.

In comparison, the effect of the unemployment rate turns significantly positive, suggesting that Trump has won larger percentages than Romney in counties with a higher unemployment rate. Plausibly, voters had more confidence in Trump, because he might have been able to create new jobs due to his vocational background.

The analogous holds for poverty shares. While a higher poverty non-significantly reduced Donald Trump’s percentage, he has received more votes than Romney, suggesting that the electorate might have believed that Trump would have been more able to reduce poverty via job creation than a president Mitt Romney. Its spillover is still significantly negative, suggesting that Trump’s margin to Romney declined, when more neighbouring counties suffer from poverty.

In other words, if poverty is a geographically widespread problem in a given region, then voters preferred the Democrats and Mitt Romney.

While Trump has been more successful than Romney in counties with a strong manufacturing sector as suggested by five models, the spillover is now significantly positive. In other words, Trump has not only received larger percentages of votes in the relevant counties, he has also received more votes than Romney. Although, the percentages of votes decreases, when neighbouring counties also benefit from a strong manufacturing industry, Donald Trump has still been more successful in the given county than Mitt Romney. The reason is that Trump’s campaign also covered the revitalization of the Rust Belt. Given the complex supply chains (long, geographically widespread) characterizing the manufacturing sectors, positive shocks spread out geographically more broadly benefiting larger regions.

In comparison, the share of insured people becomes insignificant, suggesting that a higher share generally favours the Republican party, as Trump has not benefited significantly more from it than Romney in 2012.

Plausibly, the growth rate of import penetration spurs Trump’s margin, as he has intensified the trade conflicts with China. Thus, the campaign against China also turned out to be a successful tool. Spillover effects are significantly negative, as trade restrictions damage neighbouring counties and their industries given the complex system of supply chains.

Similar to the share of insured people, the share of women loses significance. The Republicans generally suffer from a larger share independent of the candidate, implying that scandals about sexual harassment have not ruined Trump’s probability of winning.

As can be seen from the previous table, higher shares of blacks, latinos, highly-educated and young people generally decrease the Republican’s chances to win. However, Donald Trump has even been more unpopular in this groups than Mitt Romney due his response to the Black-Lives-
Matter movement and racist comments. Conversely, the spillover of the share of latinos becomes significantly positive. Hence, Trump's strategy of stoking fears of latinos ('wall to Mexico') has proven to be successful, as Trump received more votes in counties whose neighbors are the home of larger shares of latinos.

On the other hand, older generations backed the Republican party, but not Trump in particular as suggested by the insignificant impact. Trump has not been more successful in urban counties as well.

As in Goetz et al. (2019), the results are robust, when using the Republican's share of votes of the previous ballots (e.g. 2008, 2004) instead of the one of 2012 as the reference values.

<table>
<thead>
<tr>
<th></th>
<th>16-18</th>
<th>18-34</th>
<th>35-54</th>
<th>55-75</th>
<th>Total</th>
<th>12-16</th>
<th>17-21</th>
<th>22-35</th>
<th>36-54</th>
<th>55-75</th>
<th>75+</th>
<th>Total</th>
<th>12-16</th>
<th>17-21</th>
<th>22-35</th>
<th>36-54</th>
<th>55-75</th>
<th>75+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>1.149</td>
<td>1.035</td>
<td>0.848</td>
<td>0.859</td>
<td>1.051</td>
<td>3.133</td>
<td>1.458</td>
<td>2.945</td>
<td>6.834</td>
<td>4.313</td>
<td>1.782</td>
<td>3.133</td>
<td>1.458</td>
<td>2.945</td>
<td>6.834</td>
<td>4.313</td>
<td>1.782</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matter</td>
<td>0.599</td>
<td>0.543</td>
<td>0.510</td>
<td>0.604</td>
<td>0.562</td>
<td>0.724</td>
<td>0.689</td>
<td>0.716</td>
<td>0.698</td>
<td>0.684</td>
<td>0.679</td>
<td>0.689</td>
<td>0.716</td>
<td>0.698</td>
<td>0.684</td>
<td>0.684</td>
<td>0.679</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Racist</td>
<td>0.526</td>
<td>0.469</td>
<td>0.434</td>
<td>0.465</td>
<td>0.457</td>
<td>0.546</td>
<td>0.509</td>
<td>0.528</td>
<td>0.514</td>
<td>0.499</td>
<td>0.495</td>
<td>0.509</td>
<td>0.528</td>
<td>0.514</td>
<td>0.499</td>
<td>0.499</td>
<td>0.495</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotion</td>
<td>0.455</td>
<td>0.404</td>
<td>0.371</td>
<td>0.398</td>
<td>0.395</td>
<td>0.483</td>
<td>0.446</td>
<td>0.466</td>
<td>0.448</td>
<td>0.435</td>
<td>0.431</td>
<td>0.446</td>
<td>0.466</td>
<td>0.448</td>
<td>0.435</td>
<td>0.431</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Spatial Durbin regressions of difference between Republican's percentage of votes in 2016 and 2012

Notes:
- All models include fixed effects for county.
- All models include lagged dependent variable as independent variable in the first order.
- All models include lagged dependent variable as independent variable in the second order.
- All models include lagged dependent variable as independent variable in the third order.
- All models include lagged dependent variable as independent variable in the fourth order.
- All models include lagged dependent variable as independent variable in the fifth order.
5.2. Robustness checks

5.2.1. Republican's percentage of votes

The assumption that all counties located within the great circle around a given county's centroid respond the same way to a shock in the same county might be quite restrictive. When using the inverse-distance matrix for the modelling of spatial clustering as well, the results are robust, though observing significance declines slightly.

Table 4: Spatial Durbin regressions of Republican's percentage of votes using inverse-distance matrices for spatial clustering of residuals.

Second, we substitute coal output with some binary variables for particular size groups to
examine non-linear effects of coal production. The baseline are counties that do not mine for coal. The first dummy $I\{\text{Coal Output} \in [0,1)\}$ equals one, if a county produces $>0$ shorttons, but $\leq$ one mill. shorttons, and zero otherwise. The second dummy is one, when output lies between one and three mill. shorttons, and is zero otherwise. The third and fourth dummies equal one, if production varies between three and five, and five and nine mill. shorttons, and are zero otherwise. Last, the fifth dummy equals one, if more than nine mill. shorttons are produced, and is zero otherwise. This specification allows to examine whether the percentages of votes differ across the various size categories.

The results are displayed in Table 5. When only small amounts of coal are produced, Trump’s percentage of votes does not significantly differ from the ones in non-coal producing counties. On the other hand, the dependent variable is higher for counties that produce larger amounts. Plausibly, for the relevant counties, coal mining is more important source of income affecting local economic conditions and, therefore, favour the Republican party. In comparison, the interpretation of spillovers is not straightforward. Some spillovers show very large coefficients. For instance, if, in a hypothetical sense, all neighbours of a given county produce an output lying between three and five mill. shorttons, the dependent variable in the same county is by $2.29\text{-}115.86$ percentage points lower compared with a situation in which all its neighbours do not produce coal. To give an example, suppose the following two situations. First, a given county has no neighbour hosting coal mines. Second, the same county now has a neighbour producing an amount between three and five mill. shorttons. It’s neighbour gets the average weight of 0.0003. In column (15), the weighted average rises from zero to 0.0003, suggesting an increase of the dependent variable by 0.38 percentage points.

Moreover, this specification allows to distinguish the spillover effects resulting from different class sizes. Spillover effects from neighbouring counties with a lower coal production significantly increase a given county’s dependent variable in some models, as for smaller outputs economic benefits will exceed concerns regarding pollution and emissions. Similarly, the spillovers from the third and fourth dummies show a significant impact. Concerning the former, its spillovers are significantly negative. Thus, people in a given county might fear that coal mines of a similar class category might be opened in their county in the future due to possible liberalizations and the resulting pollution. Nonetheless, the economic advantages may not be large enough to compensate these downsides. Conversely, the spillover of the fourth dummy is significantly positive for larger cutoffs, suggesting that for counties farer away benefits stemming from economic spillovers exceed concerns regarding emission and pollution, while for closer counties environmental disadvantages might be too large to be compensated by economic gains from trade.

In the Tables 6 and 7, we substitute coal output by output per employed worker and working hour. Both variables are average products of labour measuring productive efficiency. To ease the interpretation without loss of generality, the output per worker is measured in thsnd. tons per employee, while output per working hour is measured in shorttons per working hour. The
<table>
<thead>
<tr>
<th>State</th>
<th>Share Old Female</th>
<th>Share Old Male</th>
<th>Share Young Male</th>
<th>Share Young Female</th>
<th>Young Farm</th>
<th>Young Urban</th>
<th>Share School Male</th>
<th>Share School Female</th>
<th>Share Falls</th>
<th>Falls Male</th>
<th>Falls Female</th>
<th>Share Mid</th>
<th>Share High</th>
<th>Share Low</th>
<th>Share</th>
<th>Share Total</th>
<th>Share Rural</th>
<th>Share Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>N.H.</td>
<td>0.003</td>
<td>0.009</td>
<td>0.008</td>
<td>0.002</td>
<td>0.002</td>
<td>0.011</td>
<td>0.009</td>
<td>0.006</td>
<td>0.009</td>
<td>0.003</td>
<td>0.004</td>
<td>0.010</td>
<td>0.009</td>
<td>0.009</td>
<td>0.010</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>Mass.</td>
<td>0.004</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td>0.003</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
<td>0.004</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>R.I.</td>
<td>0.004</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Vt.</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**Notes:**
- Data are from the U.S. Bureau of the Census, 1970 Census of Population and Housing, 1975
- The data are for the years 1970 and 1975
- The table shows the percentage of Republican votes for each category of coal output amounts.
coefficients of interest decrease, while their p-values rise, though still being significant at the 0.10-significance level. Thus, the results are robust concerning the employed measure.

| 161 | 0.038 | 0.028 | 0.023 | 0.021 | 0.020 | 0.020 | 0.019 | 0.019 | 0.019 | 0.019 |
|     |      | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 |

The dependent variable is the Republican’s percentage of votes in all specifications. All standard errors, in parentheses, take account of heteroskedasticity. Alaskan and Hawaiian counties, and Campbell/Wyoming are excluded.

3107 3107 3107 3107 3107 3107 3107 3107 3107 3107

One famous example is Donald Trump’s campaign pledge of relocating coal production

Table 6: Spatial Durbin regressions of Republican’s percentage of votes using coal output in thousands short tons per average number of employees hired by coal mines

5.2.2. Populism effects

For the population equation, we perform the same robustness checks and observe robust results, as shown in the Tables 8, 9, 10 and 11.

5.3. Discussion

Campaign pledges serve as important tools used by politicians to persuade voters to vote for them. One famous example is Donald Trump’s campaign pledge of relocating coal production to America.
### Table 7: Spatial Durbin regressions of Republican’s percentage of votes using coal output in short tons per working hour

| 1200 | Over 1200 | 1000 km | 800 km | 600 km | 400 km | 200 km | 100 km | 10 km | 5 km | W | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P |
|-----|-----------|---------|--------|--------|--------|--------|--------|-------|------|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.588 | 0.615 | 0.678 a | 0.545 | 0.572 | 0.475 | 0.457 a | 0.374 a | 0.460 a | 0.299 | 0.320 | 0.275 | 0.253 | 0.220 | 0.198 | 0.154 | 0.105 | 0.060 | 0.018 | 0.006 |
| 0.018 | 0.027 | 0.018 | 0.024 | 0.022 | 0.022 | 0.027 | 0.037 | 0.044 | 0.052 | 0.059 | 0.066 | 0.073 | 0.080 | 0.087 | 0.094 | 0.101 | 0.108 | 0.115 |
| 0.733 | 0.663 | 0.603 a | 0.563 a | 0.531 a | 0.495 a | 0.465 a | 0.434 | 0.414 | 0.389 | 0.364 | 0.340 | 0.315 | 0.291 | 0.266 | 0.241 | 0.216 | 0.191 | 0.166 | 0.141 |
| 0.042 | 0.039 | 0.044 | 0.049 | 0.055 | 0.061 | 0.066 | 0.071 | 0.076 | 0.082 | 0.087 | 0.092 | 0.097 | 0.102 | 0.107 | 0.112 | 0.117 | 0.122 | 0.127 |
| 0.758 | 0.692 | 0.635 a | 0.597 a | 0.567 a | 0.535 a | 0.505 a | 0.475 a | 0.444 a | 0.413 a | 0.382 a | 0.351 a | 0.320 a | 0.290 a | 0.260 a | 0.230 a | 0.200 a | 0.170 a | 0.140 a |
| 0.054 | 0.051 | 0.056 | 0.061 | 0.066 | 0.071 | 0.076 | 0.081 | 0.086 | 0.091 | 0.096 | 0.101 | 0.106 | 0.111 | 0.116 | 0.121 | 0.126 | 0.131 | 0.136 |
| 0.077 | 0.074 | 0.079 | 0.084 | 0.089 | 0.094 | 0.099 | 0.104 | 0.109 | 0.114 | 0.119 | 0.124 | 0.129 | 0.134 | 0.139 | 0.144 | 0.149 | 0.154 | 0.159 |
| 0.083 | 0.080 | 0.085 | 0.090 | 0.095 | 0.100 | 0.105 | 0.110 | 0.115 | 0.120 | 0.125 | 0.130 | 0.135 | 0.140 | 0.145 | 0.150 | 0.155 | 0.160 | 0.165 |
| 0.089 | 0.086 | 0.091 | 0.096 | 0.101 | 0.106 | 0.111 | 0.116 | 0.121 | 0.126 | 0.131 | 0.136 | 0.141 | 0.146 | 0.151 | 0.156 | 0.161 | 0.166 | 0.171 |

Note: The Durbin’s correlation in the Republic’s percentage of votes is calculated by the estimated coefficient of the spatial lag in the regression of Republican’s percentage of votes using coal output in short tons per working hour. The coefficient is estimated independently for each lag distance using a Tobit model. The coefficient for the spatial lag is calculated using the following equation: $\hat{\rho} = \frac{\hat{\beta}_1}{1 - \hat{\beta}_1 - \hat{\beta}_2}$, where $\hat{\beta}_1$ is the coefficient of the spatial lag and $\hat{\beta}_2$ is the coefficient of the spatial error term. The standard error of the coefficient is calculated using the Newey-West estimator. The coefficient is considered significant at the 0.10 level if it is greater than 0.10. The coefficient is considered significant at the 0.05 level if it is greater than 0.05. The coefficient is considered significant at the 0.01 level if it is greater than 0.01. The coefficient is considered significant at the 0.001 level if it is greater than 0.001.
In all specifications, 1 is a reference category for the Republican's percentage of votes in 2016 and 2012. All standard errors, unless presented otherwise, are heteroskedasticity-robust. Alabama and Hawaii have one observation, and in parentheses are presented with "150 km" as the breakpoint of the spatial weight. 

\( p \leq 0.05 \quad \star \quad p \leq 0.01 \quad \star\star \quad p \leq 0.001 \)

<table>
<thead>
<tr>
<th>b</th>
<th>100 km</th>
<th>200 km</th>
<th>300 km</th>
<th>400 km</th>
<th>500 km</th>
<th>600 km</th>
<th>700 km</th>
<th>800 km</th>
<th>900 km</th>
<th>1000 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>0.632</td>
<td>0.532</td>
<td>0.432</td>
<td>0.332</td>
<td>0.232</td>
<td>0.132</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
</tr>
<tr>
<td>2</td>
<td>0.532</td>
<td>0.432</td>
<td>0.332</td>
<td>0.232</td>
<td>0.132</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
<td>0.032</td>
</tr>
</tbody>
</table>

**Table 8:** Spatial Durbin regressions of difference between Republican's percentage of votes in 2016 and 2012 using inverse-distance matrices for spatial clustering of residuals.
Table 9: Spatial Durbin regressions of difference between Republican’s percentage of votes in 2016 and 2012 using dummies for categories of coal output amounts

<table>
<thead>
<tr>
<th>Year</th>
<th>Size of Output (1-3)</th>
<th>Size of Output (4-6)</th>
<th>Size of Output (7-12)</th>
<th>Size of Output (13-)</th>
<th>Size of Output (13-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>2017</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>2018</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Notes: For all regressions, $s$ is a dummy variable for each logit regression. The dependent variable is the difference in Republican’s percentage of votes in 2016 and 2012, adjusted for age, sex, race, and urban residence. The fixed effects are state and time. The fixed effects are state and time.
Table 10: Spatial Durbin regressions of difference between Republican’s percentage of votes in 2016 and 2012 using coal output in ths. short tons per average number of employees hired by coal mines
Table 11: Spatial Darwin regressions of difference between Republican’s percentage of votes in 2016 and 2012 using coal output in shorttons per working hour
back to US counties. In this study we establish a link between coal mining and the Republican’s percentages of votes in the 2016’s ballot Trumps vs. Clinton. Besides, we examine populist effects using the shares received by Mitt Romney in the 2012’s ballot as the non-populist baseline. Being the first study examining the effects of coal mining on election outcomes that considers spillover effects, the main results of this research highlight that the Republicans have not only received larger percentages of votes in coal counties, but Trump has been also more successful than Romney in the same counties due to his campaign pledge. Moreover, the robustness checks confirm this conclusion.

Our results are generally in line with the available literature. Concerning the variable of interest, Goetz et al. (2019) also find a significantly positive impact of the share of employment in coal industry in both models, the regression using the Republican’s percentage of votes as dependent variable and the populist equation. Despite the different measure, the magnitudes of their coefficients and ours are comparable.

Second, Steinmayr (2021) also observes a significantly positive impact of the right-wing party’s percentage of votes in the previous ballot on the share of the same party in the ballot of interest. Third, Goetz et al. (2019) get an insignificant negative effect of the share of manufacturing in both models. In comparison, we obtain significantly positive ones. The reason is the inclusion of spatial spillovers. For instance, when performing OLS regressions of our models excluding the spillovers, as shown in the first columns of the tables, the impact of the variable as non-significantly negative as well. The result suggests that controlling for spatial spillovers is necessary to avoid omitted variable biases. Fourth, in Goetz et al. (2019), the unemployment rate significantly decreases the Republican’s percentages of votes, while, in the populist equation, it significantly raises the difference in the percentages of votes between Trump and Romney. The analogous holds for the populist equation in Rodríguez-Pose et al. (2021). In our first model, however, the effect of the same variable is insignificantly positive, while, in the populist equation, it significantly raises the difference in votes as in Goetz et al. (2019). When estimating the simple OLS model excluding the spillovers, the effect is still insignificantly positive as in Rodríguez-Pose et al. (2021). It, however, becomes significantly negatively as in Goetz et al. (2019), if excluding the Republican’s percentage of votes in the 2012’s ballot. Thus, excluding the outcome of the previous ballot results in an omitted variable bias. Fifth, Goetz et al. (2019) estimate an insignificantly negative impact of the Gini index. In comparison, we find that higher poverty shares significantly drop the Republicans’ share of votes. Besides, the result contradicts Rodríguez-Pose et al. (2021) who conclude that higher Gini coefficients significantly decrease the difference in the share of votes between 2016 and 2012, as we find significantly positive impacts. Sixth, contradicting Goetz et al. (2019), the share of insured people significantly raises the Republican’s percentage of votes. Goetz et al. (2019) observe a significant positive effect of the share of uninsured people on the same dependent variable. In the populist equation, however, its effect is insignificant. Limiting the number of covariates, as suggested by the variance inflation factors, might be the reason for the different conclusion, as excluding the spillovers does not change the result. Additionally excluding the
outcome of the 2012’s ballot does not change the conclusion as well. Seventh, the positive effect of the growth of import penetration is to some degree consistent with Autor et al. (2020). Particularly, they find that the same variable raised the growth of the Republicans’ probability of winning, while it insignificantly decreases the growth of the party’s share of votes. There are two reasons. First, dependent variables are differently defined (levels vs. first differences). Second, in the OLS regressions excluding the spillovers, the its impact is also insignificant. Hence, the inclusion of the spillovers may solve omitted variable biases. Next, the share of females significantly decreases the dependent variable in the regression of Trump’s percentage of votes as in Goetz et al. (2019), but turns insignificant in the populist equation. In the simple OLS equation, the impact, however, stays insignificant, implying that the smaller number of covariates causes this conclusion. Like Rodríguez-Pose et al. (2021), Goetz et al. (2019) and Scala & Johnson (2017), we obtain also get a significant negative impact of black households, a significantly negative effect of the share of hispanics, and significantly negative impacts of the shares of people aged up to 30 years, that is comparable to their share of millenials and highly educated people. Last, confirming the generation gap, older generations back the Republican party contracting Goetz et al. (2019) who find a negative effect on Trump’s share of votes. In comparison, older generations do not support Trump in particular consistent with the same study.

6. Conclusions

Our analysis shows that Donald Trump’s campaign for coal was a success at the ballot boxes in his 2016 Presidential campaign. In general, Donald Trump’s campaign was particularly successful in coal regions and their surroundings.

Using a spatial Durbin model we find a very robust positive effect of coal production in a county on the vote share of Donald Trump in the respective county. In our baseline model the effect This positive effect amounts to 0.059-0.095 percentage points per additional mill. short tons of coal production. This effect is even more pronounced in our populist model in which we estimate the vote difference between Mitt Romney 2012 and Donald Trump 2016 with the county’s coal production and further control variables. An additional coal production of 1 mill. short tons results in an significant increase of the Republican’s vote share by 0.080-0.123 percentage points.
7. References


Appendices

A. Estimation method

Both, Maximum Likelihood (ML) and GLS-2SLS-GMM techniques, produce consistent estimates. According to Lee (2004), ML is asymptotically efficient and consistent, if some regularity conditions are satisfied. However, this only holds true under homoskedastic errors, while GLS-2SLS-GMM, however, also generates efficient and consistent estimates under heteroskedasticity (Arraiz et al. (2010), Badinger & Egger (2011), Drukker et al. (2013b), Kelejian & Prucha (1998), Kelejian & Prucha (1999), Kelejian & Prucha (2010)). Furthermore, Gibbons & Overman (2012) criticize MLE for assuming prior knowledge of the data-generating processes which is not usual in empirical studies. Given these issues, we choose the GLS-2SLS-GMM taking account of heteroskedasticity.

Empirical studies employing instrumental variables techniques follow Kelejian & Prucha (1999) who use spatial lags of all covariates as instruments. Kelejian & Prucha (1999) suggest a procedure consisting of three steps. In the first step, consistent estimators for all coefficients are obtained by a Two-Staged Least Square (2SLS) estimation. The coefficients are combined to the single vector $\delta = (\alpha, \gamma, \beta', \rho')'$, for which an estimator is obtained by $\hat{\delta} = (Z'Z)^{-1}Z'y$, with $Z$ being the matrix containing the covariates and the spatially lagged covariates, $\bar{Z} = P_{H_1}Z$, $P_{H_1} = H_1(H_1'H_1)^{-1}H_1'$ and $H_1 = X_f$ whereby $X_f$ is a matrix containing the covariates and their spatial lags. Spatial clustering in the residuals is ignored, as only asymptotically efficient and consistent estimators of the listed coefficients are required. In the second step, coefficient $\lambda$ is obtained with GMM by solving the sample equivalent of the population moment conditions by using the residuals obtained from the first step (Badinger & Egger (2011), Drukker et al. (2013a), Kelejian & Prucha (1998), Kelejian & Prucha (1999), Kelejian & Prucha (2004), Kelejian & Prucha (2010))

\[
\frac{1}{N} E[\eta' W_\epsilon \eta] = 0 \\
\frac{1}{N} E[\eta' B_\epsilon \eta] = 0 \\
\text{with } B_\epsilon = W_\epsilon' W_\epsilon - \text{diag}(W_\epsilon' W_\epsilon) \tag{4}
\]

In the third step, the estimator $\hat{\lambda}$ is employed to perform a Cochrane-Orcutt transformation,
as shown in equation (5).

\[
y_{nt} = Z_s(\lambda)\delta + \eta
\]

with \( y_{nt} = (I_n - \sum_{s=1}^{S} \lambda W_s)y \)

and \( Z_s(\lambda) = (I_n - \sum_{s=1}^{S} \lambda W_s)Z \)

\( I_n \) and \( S \) denoting an \( n \times n \) identity matrix and the order of spatial lags of the error term (in our case, \( S = 1 \)).

By using the instrument matrix \( H_2 \) and substituting \( \lambda \) with the estimator \( \hat{\lambda} \), the GS2SLS estimator of \( \delta \) is

\[
\hat{\delta} = \left( \hat{Z}_s(\hat{\lambda})' \hat{Z}_s(\hat{\lambda}) \right)^{-1} \hat{Z}_s(\hat{\lambda})' y_s(\hat{\lambda})
\]

with \( y_s(\hat{\lambda}) = (I_n - \sum_{s=1}^{S} \hat{\lambda} W_s)y \)

and \( Z_s(\hat{\lambda}) = (I_n - \sum_{s=1}^{S} \hat{\lambda} W_s)Z \)

and \( \hat{Z}_s(\hat{\lambda}) = P_{H_2} Z_s(\hat{\lambda}) \)

and \( P_{H_2} = H_2(H_2' H_2)^{-1} H_2' \)

whereby \( H_2 \) contains the linearly independent columns in \( H_2 = [H_1, W_1 H_1] \).