

The (Uneven) Spatial Distribution of The Bakken Oil Boom

Johanna Richter* Alliana Salanguit* Alexander James[†]

January 31, 2018

Abstract

Resulting from a booming shale-energy sector, from 2007 to 2014, real income per capita in North Dakota increased 40%. Does this reflect the experience of a few oil-rich counties, or were the gains more evenly distributed across the region? We find that the shale boom generated significant economic gains for counties above and near the Bakken, but not for those further away. We also document significant state-border effects which are not easily explained. Conditional on distance to the Bakken region, the shale boom generated limited outward migration from South Dakota and, perhaps as a result, the economic gains that accrued there were muted.

Keywords: Resource Boom; Migration; Regional Development

JEL Classification: Q32; Q33

*University of Alaska Anchorage

[†]Corresponding author. Department of Economics and Public Policy, University of Alaska Anchorage, Anchorage, AK 99508. alex.james@uaa.alaska.edu

As some residents [of North Dakota] find themselves cashing oil royalty checks worth tens of thousands of dollars a month or more, many of their neighbors are resigned to receiving almost nothing from the wells that pepper the landscape and even their own land — aside from the headaches that go with living in a boomtown.—A. G. Sulzberger, writing for the New York Times.

1 Introduction

Due largely to advancements in hydraulic fracturing technology, in the mid 2000s there was a rapid increase in unconventional fossil fuel recovery in the United States. The local economic impact of this resource boom was particularly profound in North Dakota, home of the Bakken shale formation. To appreciate the magnitude of the resource boom there, and referencing Figure 1, consider the fact that from 2006 to 2014, oil production in North Dakota increased from 40 million barrels to roughly 400 million barrels (data collected from the Energy Information Administration). As a result, from 2007 to 2014, real income per capita in North Dakota increased by nearly 40% and employment increased by more than 25%. For reference, real per capita income and employment in neighboring oil-poor South Dakota increased by roughly 4% and 6% respectively (authors calculations).¹

There is now a large literature that examines both the short run economic impact of resource booms generally, and the shale-boom specifically. Examining the experience of western U.S. counties, Weber (2012, 2014) finds that the shale boom increased local levels of employment, population, and income. Utilizing the synthetic control methodology, Munasib and Rickman (2015) examine the impact of the shale boom within the states of Arkansas, North Dakota and Pennsylvania. Similar to the conclusions drawn by Jeremy Weber, they conclude that the shale boom generated significant economic gains, especially in North Dakota. In a creative natural field experiment, Komarek (2016) exploits the fact that the state of New York placed a variety of moratorium on hydraulic fracturing whereas neighboring Pennsylva-

nia, West Virginia, and Ohio did not. He similarly documents positive economic effects of the shale boom, namely, increased wages and employment (though using a propensity score matching approach, Paredes, Komarek, and Loveridge (2015) document minimal wage and employment effects of fracking in the Marcellus region).²³ Similar to the present paper, Miljkovic and Ripplinger (2016) examine the economic impact of the North Dakotan shale boom using state-level time series data. They find that state-level wages and employment (in mining, agriculture, and other sectors) all rose in response to enhanced oil recovery. However, their analysis is ill-suited to capture the dispersion of the economic effects as state-level impacts may be driven entirely by a small handful of oil-rich counties. It should be mentioned that, along with significant economic gains, the shale boom generated a variety of negative outcomes as well. The shale boom increased local crime rates (James and Smith, 2015), depressed the value of nearby ground-water-dependent homes (Gopalakrishnan and Klaiber, 2014; Muehlenbachs, Spiller and Timmins, 2015), and limited high school and college attainment (Rickman, Wang and Winters, 2017).

Most similar to the present paper, Feyrer, Mansur, and Sacerdote (2017) evaluate the local economic effects of the shale boom, and draw specific attention to propagation effects. Their baseline analysis regresses the change in either sector employment or wages on oil and gas production at the county level. Aggregating to a 100 mile radii around each county, and further aggregating to the state level significantly increases the estimated coefficient on oil production. The authors consider this to be evidence of economic propagation (or spillovers from the county of production to neighboring counties). To more precisely gauge the extent of propagation, they draw concentric circles around all counties and vary the radii from 20 miles to 200 miles in 20 mile increments. Aggregating up both dependent and independent variables, they observe that the effect of oil production on wages and employment ceases to grow for radii in excess of 100 miles. They conclude from this that significant spillovers occur within 100 miles of the point of oil and gas extraction.

We build upon the work of Feyrer, Mansur and Sacerdote by considering the effect of the shale boom in North and South Dakota counties that are above, near, or far away from the Bakken shale formation. We focus on the experience of North and South Dakota for a few reasons. First, as shown in Munasib and Rickman (2014), the economic impact of the shale boom was especially strong in North Dakota and the state had not previously experienced an oil boom near the magnitude of the recent shale boom.⁴ Second, there were no other major oil booms within hundreds of miles. This allows us to identify the unique effect of the Bakken shale boom and our estimates are not confounded by additional nearby oil drilling. Third, neighboring South Dakota did not experience a shale boom; the Bakken shale formation is largely confined to western North Dakota, eastern Montana, and southern Saskatchewan, Canada.⁵ The differential impacts of the shale boom on these two otherwise very similar states makes for an ideal setting to consider the effects of state borders on the spatial dispersion of economic booms. We further contribute to the existing literature by examining the propagation of a wider array of economic outcomes including income, unemployment, poverty, public finance, and regional migration.

To measure the dispersion of the effects of the oil boom, we place each North and South Dakota county into a specific “bin” based off of each county’s distance to the Bakken shale formation. Using a difference-in-difference estimation strategy we then estimate the effect of the energy boom on counties within each of these bins. We find that oil-producing counties experienced large economic gains (measured as reductions in the unemployment and poverty rate, and increases in income per capita, and population). Nearby counties within 200 kilometers also experienced economic gains, though those gains were matched with falling—not rising—populations. Examining county-to-county migration patterns reveals that there was significant outward migration from neighboring counties (especially neighboring North Dakota counties) to oil-producing ones. There is evidence that the emigrating population was relatively uneducated. To the extent that migrants to the Bakken were previously unemployed or

of a low skill type, this may explain the observed dispersion of the economic gains.⁶ Interestingly, we also document significant state-border effects. Conditional on a county's distance to the Bakken, South Dakota counties experienced minimal outward migration, and perhaps as a result, minimal per capita economic gains. This result is not easily explained by state-level variation in public finances, though we do find that, as a result of the oil boom, all North Dakotan counties received enhanced state-level intergovernmental grants for education. While identifying the mechanism for state-border effects is beyond the scope of the paper, this finding is broadly consistent with an extant literature on intra-country home bias and state border effects (e.g., Wolf, 2000; Hilberry and Hummels, 2003), though, that literature tends to focus on wholesale trade flows rather than migration.

2 Identification & Data

Referencing Figure 2, all North and South Dakota counties are placed into one of $n = 6$ bins according to their Euclidean distance from the nearest oil-producing county.⁷ Bins are defined based off of 100 kilometer band widths. For example, $n = 1$ corresponds to oil-producing counties (in 2012)⁸, $n = 2$ corresponds to counties within 100 kilometers, $n = 3$ corresponds to counties between 100 and 200 kilometers away, $n = 4$ corresponds to counties between 200 and 300 kilometers away, and so on. The last bin, $n = 6$, corresponds to counties that are in excess of 400 kilometers away from oil-producing counties.

To estimate the spatial propagation of the North Dakota shale boom, we employ a difference-in-differences estimation technique that exploits both the geological location of unconventional hydrocarbon deposits and the exogenous timing of the shale-energy boom (both of which are exogenous factors from the perspective of economic agents in North Dakota). Our sample period is 2000-2015. We develop an indicator variable ($Post_t$) that is equal to unity for years in which the Bakken shale formation was booming (2007-2015, see Figure 1). We then inter-

act this indicator variable with another indicator variable that defines which bin a county is designated to. More specifically, we estimate equation (1) below:

$$\ln(y_{i,t}) = \sum_{n=1}^5 \beta_n (\text{Post}_t \times D_{i,n}) + Z_t + C_i + \epsilon_{i,t}, \quad (1)$$

where $D_{i,n}$ is the indicator variable that defines which bin a county is designated to, Z_t are year fixed effects that capture any relevant time-varying factors that are not specific to any given county, C_i are county fixed effects that control for any county-specific factors that are time invariant, and $\epsilon_{i,t}$ is an error term that is clustered at the county level. Note that with the inclusion of year and county fixed effects, the variables Post_t and $D_{i,n}$ are directly controlled for in the model and hence, β_n defines the average treatment effect (the effect of a county being in bin n relative to bin 6) during the oil boom (2007 to 2015) relative to the average treatment effect pre-oil boom (2000 to 2006). Because counties in bin 6 are the reference counties, an identifying assumption is that these South Dakotan counties, that are at least 400 kilometers away from oil-producing counties, were not affected by the shale boom. This assumption seems plausible given the relative isolation of these counties from any oil drilling activity (see Figure A1 in the appendix) which was produced using data generously provided by drillinginfo.com and reflects all new oil and gas wells produced from 2007 to 2013).

A key identifying assumption is that preexisting trends were similar across the treatment and control group of counties. Figure 3 provides a graphical description of the evolution of the various outcome variables for counties in bin 1 (oil-producing counties) and those in the two sets of controls (counties in bin 6 and, for robustness, counties in Nebraska, Iowa, and Minnesota). Starting with panel (a), in 2000, average income per capita in bin 1 is similar to that for bin 6 and that for counties in Minnesota, Iowa, and Nebraska. Prior to 2007, average income per capita in the three groups was also growing at roughly similar rates. However, in 2007, income per capita in bin 1 clearly diverges from both sets of controls, reflecting the

North Dakotan oil boom. Similar patterns exist for unemployment (panel c). The results for poverty are less promising. Prior to the oil boom, variation in the poverty rate in bin 1 was not highly correlated with that for the two sets of controls. The results for poverty should therefore be viewed with some added caution. Finally, the results for population reveal that the two sets of controls are, on average, much more densely populated than those counties in bin 1. However, the evolution of population prior to the oil boom across the three sets of counties appear to be roughly similar, suggesting that the controls may still be valid given that all regression specifications include county and year fixed effects.

Beyond economic spillovers to neighboring counties, there may be spillovers that are specific to counties within the state of North Dakota. For example, enhanced state-level revenue resulting from the shale boom may increase local public expenditures on various public goods and this may spur additional economic activity. Equation (1) cannot reveal these effects because a distinction is not made between North and South Dakota counties. We therefore estimate an additional model that interacts each of the interactions from equation (1) with an additional indicator variable equal to unity for counties in South Dakota. We specifically estimate:

$$\ln(y_{i,t}) = \sum_{n=1}^5 \gamma_n(\text{Post}_t \times D_{i,n}) + \sum_{n=1}^5 \lambda_n(\text{Post}_t \times D_{i,n} \times SD_i) + Z_t + C_i + \epsilon_{i,t}, \quad (2)$$

where SD_i is the indicator variable for South Dakota counties. All of the other variables are defined as before. Note from Figure 2 that there is only 1 South Dakota county in bin 1, and only 1 North Dakota county in bin 5. For this reason we focus our North Dakota analysis on bins 1-4 and our South Dakota analysis on bins 2-5.

The outcome variable, $y_{i,t}$ takes on a variety of definitions. As a starting we point, we define it as state personal income per capita (nominal values were converted to real ones using the CPI, 2000 is the base year). But average income levels say little about the distribution of income within a county. To better evaluate the welfare implications of the energy boom we

also define $y_{i,t}$ as the poverty rate and unemployment rate. We also consider broader economic and migration effects by defining $y_{i,t}$ as total employment and population. We also consider the effect of the boom on various state and local government finances including total revenue, expenditures, and education expenditures specifically.

The population, employment, and income data were collected from the Bureau of Economic Analysis (BEA), Regional Database. The unemployment data were collected from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics database and the poverty data were collected from the Census Bureau's Small Area Income and Poverty Estimates database. All public finance data were collected from the Census Bureau's State & Local Government Finance database and reflects revenues and expenditures of county and local governments (including cities and school districts for example). See Table 1 for summary statistics.

3 Results

3.1 Income & Employment

We start by estimating equation (1) defining $y_{i,t}$ as either income per capita, population, the poverty rate, or the unemployment rate. Those results are given in Figure 4. There are large economic gains associated with the energy boom for oil-producing counties, though some spatial spillovers are present. For oil-producing counties, the treatment effects for income, poverty, and unemployment are .321, -.236, and -.365, respectively. These estimates imply that for those oil-producing counties, averaged from 2007 to 2015, the energy boom generated a 37.8% increase in per capita income, a 21.0% decrease in the poverty rate, and a 30.5% decrease in the unemployment rate. For counties within 100 kilometers of oil-producing counties (counties in Bin 2), the treatment effects for income, poverty, and unemployment fall to .025, -.040, and -.134, respectively. While the results for poverty and unemployment remain statistically significant (at least at the 10% confidence level), the estimates are qualitatively

reduced relative to those for oil-producing counties. Similar treatment effects are estimated for bin 3 (counties between 100 and 200 kilometers from oil-producing counties). Beyond 200 kilometers away, the treatment effects corresponding to all three outcomes are again significantly reduced qualitatively, and all become statistically insignificant.

Turning to population, there is some evidence that the energy boom elevated the population of oil-producing counties, though this result is not statistically significant. Interestingly, there is also some evidence that populations were reduced in neighboring counties within 200 kilometers of oil-producing ones. Presumably this reflects the fact that relatively high wages were being offered in oil-producing counties during the boom period and this attracted labor from local communities. In the next section we bore this idea out in more detail by examining county-to-county migration patterns.

Estimating equation (2) reveals the extent to which the spillovers documented above are state specific. The results for the four outcomes, income, population, the poverty rate, and the unemployment rate, are given in Figures 5 and 6. The spatial spillovers are largely specific to the state of North Dakota. North Dakota counties in bins 2 and 3 (non-oil-producing counties within 200 kilometers of oil-producing counties) experienced increases in income per capita and reductions in unemployment and poverty. For unemployment and population, the treatment effects in North Dakota are negative (and mostly significant) for bins 2-5. This suggests that the oil boom pulled people away from even distant North Dakota counties, and perhaps as a result reduced the unemployment rate in those places. These results do not carry over to South Dakota however. While the shale boom increased per capita income, and decreased both poverty and unemployment in the single oil-producing South Dakota county, there is little evidence of any spatial spillovers in that state.

3.2 Local Public Revenue & Expenditures

In this section we similarly examine the effect of the shale boom on a variety of public finance outcomes. Those outcomes (all measured in per capita terms) are: 1) local revenue, 2) local expenditure, 3) state intergovernmental revenue (local revenue granted from the state government), and 4) total education expenditures (which, because we are examining local education, education expenditures refer specifically to K-12 education).

Starting again with the estimation of equation (1), referencing Figure A2 in the appendix, the effect of the oil boom on revenue and expenditures are largely confined to oil-producing counties. While intergovernmental grant spending is positive for all bins, the treatment effect decays somewhat as distance from the Bakken increases. Similar results are documented for total local education spending. Estimating equation (2) again suggests that the oil boom only enhanced public revenue and expenditures in oil-producing counties. As can be seen in Figure A3, while the North Dakota treatment effects for bin 5 are positive, these results should be viewed with caution as there is only one North Dakota county in bin 5. Referencing Figure A4, there is little evidence that any non-oil-producing county in South Dakota experienced an increase in either intergovernmental grants or local education spending. However, all counties in North Dakota experienced an increase in intergovernmental grants, which in turn explains why local education expenditures were enhanced in all North Dakota counties. These are both statistically and quantitatively significant results. From panel (a) of Figure A4, a treatment effect of 0.30 implies that the shale boom generated a 35% increase in intergovernmental grants (recall that the reference observation is bin 6, which consists only of South Dakota counties). Averaged from 2000-2006, the typical North Dakota county received \$852 in state intergovernmental grants, per capita. A 35% increase amounts to roughly an additional \$300 dollars per capita. This is a relatively small amount when considered relative to average county income per capita (1.2%) though significant when compared to total county government revenue per capita (12%).

In conclusion, all North Dakota counties did benefit from the shale boom in at least one way: they all received additional funding for education. Could this explain why the economic gains were relatively muted in South Dakota? Probably not. If this were the case one would expect to see the economic gains in North Dakota more evenly distributed throughout the state. Another potential explanation is that the oil boom attracted labor from neighboring counties. To the extent that people living in South Dakota were, for some reason, relatively less likely to move to the Bakken region, this could explain why the state experienced such meager economic gains. We explore this possibility in the following section.

3.3 County-to-County Migration

Consistent with basic economic theory (e.g., Corden and Neary, 1982), resource booms inflate local wages. To attract labor, other non-resource sectors must also offer high wages and this attracts labor from neighboring communities. Even if all workers were of the same type (in terms of, say, human capital accumulation and employment status) one should expect outward migration to be matched with rising wages. But suppose human capital and employment status were heterogeneously distributed across the working age population. In this case, a resource boom that generates high paying employment opportunities may be especially appealing to nearby residents that are either unemployed or receiving a low wage. In this case, a resource boom is not only likely to pull labor away from neighboring communities, but it is likely to pull away the most unproductive members of those communities. If people living near the Bakken were relatively likely to move in response to the oil boom, this may explain the observed spatial decay of the economic gains.

Here we examine county-to-county migration patterns within and between the states of North and South Dakota. We collected this data from the U.S. Census Bureau; the corresponding survey was administered by the American Community Survey (ACS). The ACS asks respondents (aged 1 year and older) whether they lived in the same residence one year ago,

and if not, where they lived. At the county level, only 5 year moving averages are available. We make use of the 2010-2014 migration flows data set and estimate the following equation:

$$\ln(\text{Migration}_i/\text{Pop}_i) = \alpha + \sum_{n=2}^5 \psi_1(D_{i,n}) + \sum_{n=2}^5 \psi_2(D_{i,n} \times SD_i) + \epsilon_i, \quad (3)$$

where Migration_i is migration from county i to any oil-producing county and Pop_i is the population of that county in 2014. All other variables are defined as before. Note that the first bin has been dropped from the data set as the dependent variable is the number of migrants moving to bin 1 (the bin containing all of the oil-producing counties). As before, the reference observation is bin 6 which consists only of South Dakota counties.

The results from estimating equation (3) are given in Figure A8 in the appendix. Similar to the earlier findings, the treatment effects are strongest for bins near the Bakken region. The treatment effect is statistically significant for all bins in North Dakota (though only at the 10% confidence level for bin 5). The treatment effects are insignificant for all bins in South Dakota.

To help gauge whether these effects are large enough to explain some of the observed income and employment effects, note that averaged from 2010-2014, the average annual number of migrants moving from bin 2 North Dakota counties to the Bakken region was 100. While the migration data is not available over the entire sample period, here we assume that this trend started at the event date (2007) and continued each year until 2015. In 2006, the average bin 2 North Dakota county had a resident population of approximately 10,200, and roughly 1,700 of them were living in poverty. If one third of all migrants to the Bakken region were previously living in poverty (an admittedly arbitrary supposition), each year 100 people would migrate away from the average bin 2 county, and 33 of them would have been living in poverty. Holding all else equal, this implies that in 2007, the average population would fall from 10,200 to 10,100, the number of people living in poverty would fall from 1,700 to 1,667, and the poverty rate would fall from $1,700/10,200 = 16.6\%$ to $1,667/10,100 = 16.5\%$. The next year

in 2008, the average poverty rate would fall to $1,634/10,000 = 16.3\%$. By 2015, the average poverty rate would be just 15%. Averaging over the entire post-event date sample period (2007-2015), the average poverty rate in bin 2 North Dakota counties would be 15.8%. And note that this corresponds to a $(15.8-16.6)/16.6 = 4.8\%$ decrease in the poverty rate. Recall from panel (c) of Figure 5 that within bin 2 counties, the estimated treatment effect for the poverty rate was approximately 5%.

This “back of the envelope” exercise arbitrarily assumes that one third of the migrants moving from bin 2 counties were previously living in poverty. While this may be unlikely, this exercise nonetheless provides some context for the estimates, and highlights the potential for migration to explain some of the earlier findings.

There is little evidence that the shale boom generated outward migration from South Dakota (see panel (b) of Figure A8 in the appendix). One, somewhat speculative interpretation is that South Dakota did not benefit from the oil boom because the oil boom failed to attract South Dakota labor. The economic impact of outward migration might be especially pronounced if the migrating labor is low skill because that labor is more likely to be unemployed, or have a low level of income. We provide evidence of this by estimating the effect of the Bakken shale boom on education attainment in neighboring counties. We collected data on the percent of the county population with 1) at most a high school degree and 2) at least a bachelor’s degree in the years 2000 and 2013.⁹ We compute the change in the percent of the county population with at most a high school diploma or at least a college degree from 2000 - 2013, and regress this value on the various North and South Dakota bin indicators. The results are given in Figure A5. There is evidence that, from 2000 - 2013, North Dakotan counties near the Bakken shale play became more educated; the percent of the population with at most a high school degree fell by 2 points (significant only at the 10% confidence level for bin 2), while the percent of the population with at least a college degree increased by more than 2 points (significant at the 5% confidence for bins 2 and 3). Furthermore, this

effect is unique to North Dakota, and are consistent with those of Wilson (2016) who finds that migrants to energy boom towns tend to be male, unmarried, young, and less educated.

Why were people in South Dakota more hesitant to move than people living in North Dakota? One possibility is that, prior to the boom, people living in North Dakota were relatively poor, and hence more likely to move once the boom occurred. But a year before the shale boom, in 2006, income per capita in South Dakota was \$26,037, whereas that in North Dakota was \$25,444 and the unemployment rate in both states was roughly 3.7%. The poverty rate in South Dakota was also slightly *larger* than that in North Dakota (16% relative to 13%).

There is surprisingly little existing literature that identifies or explains the existence of sub-national migratory border effects. Poncet (2006) finds significant inter-provincial border effects in China and Kone et al. (2016) documents similar results for India. Song (2017) documents significant U.S. state-border migration effects and suggests that “What drives the border effect must also be consistent with my earlier finding that it is the more educated, high income-earning, young professionals that cease to move across states over time”. According to Bishop (2009), migrants are most likely to move between two similar regions and Song (2017) additionally finds that differences in both race and urbanization contribute to state border migration effects. However, this is not likely to explain our results; prior to the shale boom, western counties within North and South Dakota were similar both in terms of racial makeup and urbanization. One potentially meaningful difference between the two states is that, unlike North Dakota, South Dakota does not have an individual state income tax. In 2000, well before the shale boom, the bottom marginal income tax rate in North Dakota was 2.67% and the top marginal rate was 12% (for people earning in excess of \$50,000).¹⁰ Another difference is weather: South Dakota is warmer and dryer than North Dakota. There is a large literature that links regional development and population density to environmental amenities such as weather (see for example Blomquist, Berger, and Hoehn, 1988). Taken together, the

net benefit of moving west from North Dakota to the Bakken region may be high relative to the net benefit of moving north from South Dakota. But until additional research is carried out, this explanation remains speculative.

4 Robustness

The validity of the results outlined above depend on whether the comparison units used are appropriate. Recall that counties in bin 6 are used as controls. Two key identifying assumptions are that, (1) prior to the boom, the control group was evolving similarly to other counties in North and South Dakota and (2) the evolution of the control group was unaffected by the boom. In this section we explore the validity of these assumptions.

We start by considering the possibility that the counties in bin 6 were actually affected by the shale boom. In this case, our estimates would be biased towards failing to reject the null hypothesis of no effect. We address this concern by replacing the baseline control group with Nebraskan, Iowan, and Minnesotan counties. These three states all border North and South Dakota, and are not endowed with either oil or natural gas. It is possible that the counties within these states may have also been affected by the boom. However, if this is the case, it is then reasonable to think that all U.S. counties may have been affected by the boom and hence no good control group exists. Rather than re-estimating all of the previous effects, we focus on the baseline results for income, population, unemployment and poverty. The results are given in Figure A6 and are largely consistent with the earlier findings, with two exceptions. The treatment effects for income and poverty are significant for all bins; suggesting the economic gains associated with the Bakken oil boom were far reaching. While using counties in surrounding states as controls offers the advantage of exogeneity, these counties may not be suitable controls. For example, the city of Minneapolis is in Hennepin county, Minnesota. This is arguably not a great control for North and South Dakotan counties which

are much more sparsely populated. Given this, equation (1) is re-estimated for the baseline socio-economic outcomes after restricting the data set to low population counties. The results are surprisingly similar to the previous ones and so are not reported.¹¹

Casual observation suggests that, with the exception of poverty, prior to the boom bin 1 counties were evolving similarly to bin 6 counties (and Minnesota, Iowa, and Nebraskan counties). We formally test whether there was meaningful preexisting variation in treatment and controls prior to the shale boom by estimating a variation of our main estimation equation. Rather than interacting bin indicators with the *post* dummy, we interact them with year dummies (for the years 2001-2015). With the inclusion of county and year fixed effects, this model reveals whether treatment effects rose in tandem with the timing of the energy boom, or whether our baseline results simply reflect pre-existing trend. For income, poverty, and unemployment, preexisting trend does not appear to be a significant problem (see Figure A7). Prior to 2007, variation in the treatment effects for all bins is relatively small, suggesting that prior to 2007, counties in bins 1-6 were evolving similarly to each other. This, however, was not the case with population. As can be seen, all bins (with the exception of those in Bin 5) were shrinking in population relative to bin 6, and this was especially true for counties in bins 1, 2, and 3.

5 Discussion

Our results suggest there may be wide spread gains associated with hydraulic fracturing that reach well beyond the point of extraction. In fact, they suggest that the Bakken shale boom reduced unemployment and poverty rates in counties that were in excess of one hundred miles away. Informing both voters and policy makers of the regional economic effects of resource extraction is of clear importance. This is especially true given that oil and gas-related regulations are typically negotiated at the state level by representatives that may not

believe their oil-poor districts have much to gain (or lose) from the legislation.

These results inform previous and future related research as well. To account for spatial spillovers, previous literature has dropped counties that are immediately adjacent to oil-producing counties (Michaels, 2011; Jacobsen and Parker, 2016). The idea is that counties adjacent to treatments may also be affected by a resource boom; using them as controls would cause estimates to be biased toward failing to reject the null hypothesis of no effect. Our results suggest this methodology is warranted, but may not sufficiently eliminate spillover bias as the propagation extends beyond contiguous counties.

To illustrate this point, we estimate the effect of the shale boom in bin 1 counties using bin 3 counties as the reference set, and then compare these estimates to those obtained using bin 6 counties as the reference set. Continuing to assume that bin 6 counties serve as suitable counterfactuals, this should reveal the extent of spillover bias that remains even after omitting contiguous counties. Starting with income, and using bin 3 as the reference set, the estimated treatment effect is .266 (and is significant at the 1% confidence level). But this may be a lower bound as we have previously shown that significant spillovers accrue even to bin 3 counties (which are at least 100 miles from the Bakken). Using bin 6 counties as the reference set increases the estimate to .321 (a $.321 - .266 = 5.5\%$ point increase). This same procedure suggests that the effects for unemployment and poverty are underestimated by roughly 8.8% points, and 5.3% points, respectively. Because the shale boom attracted labor from nearby counties, using nearby (bin 3) counties as reference causes the estimated treatment effect for population to be biased upward by 7% points. In the future, researchers should consider using more distant counties as controls to more fully avoid spillover bias, though doing so also raises concerns of unobserved heterogeneity bias.

Following a methodology similar to Feyrer et al. (2017), we further explore propagation effects within North Dakota by computing the percent of the income gains accruing to each bin. Referencing panel (a) of Figure 5, the treatment effects for bins 1-5 are 0.33, 0.06, 0.14,

0.03, and 0.03, respectively. Multiplying each treatment effect by the average income level for each respective bin reveals an estimate of the total income gains accruing to each bin. We estimate that 45% of the total income gains accrued to bin 1 counties (approximately \$80 million), 9% accrued to bin 2 counties (approximately \$16 million), and 14% accrued to bin 3 counties (roughly \$25 million). Qualitatively there were gains to bins 4 and 5, but the treatment effects were not statistically significant. These results are in contrast to those of Feyrer et al. who estimate that just 15% of the wage gains that accrue within 100 miles of extraction remain within the county of extraction.¹²

6 Conclusion

Due to advancements in hydraulic fracturing and horizontal drilling technology, in addition to high energy prices, energy production in the United States dramatically surged in the mid 2000s. The regional impact of this energy boom was perhaps most profound in North Dakota, home to the Bakken Shale formation.

We document large economic gains associated with the production of Bakken shale oil in North Dakota (measured as reductions in poverty and unemployment, and increases in income), and these gains spilled over to counties within 200 miles. Further, all North Dakota counties benefited from enhanced public spending on education. We also find that the boom generated significant outward migration from neighboring counties, and there is evidence that the migrants were relatively uneducated.

While the economic gains propagated within the state of North Dakota, they did not spill over into neighboring South Dakota, nor were residents of South Dakota as likely to move to the Bakken region. This result does not appear to be driven by across-state variation in demographics or economic opportunities as the western parts of North and South Dakota were, prior to the boom, very similar. However, the two states did differ in terms of both

tax policy (South Dakota does not have a state income tax), and climate (South Dakota is warmer and dryer). Additional research should ultimately be completed to better understand the effects of state borders on migration and the dispersion of economic shocks.

7 References

- Alcott, Hunt and Daniel Keniston. (2017). Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America, *The Review of Economic Studies*, <https://doi.org/10.1093/restud/rdx042>.
- Bishop, B. The big sort: Why the clustering of like-minded America is tearing us apart, Houghton Mifflin Harcourt, 2009.
- Brown, Jason P. (2014). Production of natural gas from shale in local economies: a resource blessing or curse? *Economic Review-Federal Reserve Bank of Kansas City*, 5.
- Blomquist, G.C., Berger, M.C, and Hoehn, J.P. (1988). New estimates of quality of life in urban areas. *The American Economic Review*, 89-107.
- Bureau of Labor Statistics, Local Area Unemployment Statistics. Available at: <https://www.bls.gov/lau/> (accessed June 2017).
- Corden, W. Max and Peter Neary. (1982). Booming sector and de-industrialization in a small open economy. *The Economic Journal*, 825-848.
- Fetzer, T. (2014). Fracking Growth. CEP Discussion Paper. Rep. 1278, Center for Economic Performance, London School of Economics and Political Science, London.
- Feyrer, J., Mansur, E. T., and Sacerdote, B. (2017). Geographic dispersion of economic shocks: Evidence from the fracking revolution. *The American Economic Review*, 107(4), 1313-1334.
- Gopalakrishnan, S. and Klaiber, H.A. (2014). Is the shale energy boom a bust for nearby residents? Evidence from housing values in Pennsylvania. *American Journal of Agricultural Economics*, 96(1), 43-66.

- Hilberry, R. and Hummels, D. (2003). Intranational home bias: Some explanation. *Review of Economics and Statistics*, 85(4), 1089-1092.
- James, Alexander, and Brock Smith. (2017). There will be blood: crime rates in shale-rich U.S. counties. *Journal of Environmental Economics and Management*, 84, 125-152.
- Jacobsen, Grant D., and Dominic Parker. (2016). The economic aftermath of resource booms: evidence from boomtowns in the American West. *The Economic Journal*, 126(593), 1092-1128.
- Jacobsen, Grant. (2016). Who wins in an energy boom? Evidence from wage rates and housing. Working paper. University of Oregon.
- Komarek, T. (2016). Labor market dynamics and the unconventional natural gas boom: Evidence from the marcellus region. *Resource and Energy Economics*, 45, 1-17.
- Kone, Z. L., Maggie, Y.L., Asditya, M., Caglar, O., Siddharth, S. (2016). Internal borders and migration in India. *Working Paper*.
- Lee, J. (2015). The regional economic impact of oil and gas extraction in Texas. *Energy Policy*, 87, 60-71.
- Michaels, Guy. (2011). The long term consequences of resource based specialization. *The Economic Journal*, 121(551), 31-57.
- Miljkovic, D. and Ripplinger, D. (2016). Labor market impacts of US tight oil development: The case of the Bakken. *Energy Economics*, 60, 306-312.
- Maniloff, P., and Mastromonaco, R. (2017). The local employment impacts of fracking: A national study. *Resource and Energy Economics*, 49, 62-85.
- Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *The American Economic Review*, 3633-3659.

- Munasib, A. and Rickman, D.S. (2015). Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis. *Regional Science and Urban Economics*, 50, 1-17.
- Paredes, D., Komarek, T. and Loveridge, S. (2015). Income and employment effects of shale gas extraction windfalls: Evidence from the Marcellus region. *Energy Economics*, 47, 112-120.
- Poncet, S. (2006). Provincial migration dynamics in China: Borders, costs, and economic motivations. *Regional Science and Urban Economics*, 36(3), 385-398.
- Rickman, D.S., Wang, H., and Winters, J.V. (2017). Is shale development drilling holes in the human capital pipeline? *Energy Economics*, 62, 283-290.
- Song, Y. (2017). Internal migration and crossing the state borders: 1960-2000. *Working Paper*.
- Smith, Brock. (2015). The resource curse exorcised: Evidence from a panel of countries. *Journal of Development Economics*, 116, 57-73.
- Tax Foundation, State Individual Income Tax Rates, 2000-2014. Available at: <https://taxfoundation.org/state-individual-income-tax-rates/> (accessed November 2017).
- U.S. Bureau of Economic Analysis, Regional Economic Accounts. Available at: <https://www.bea.gov/regional/index.htm> (accessed June 2017).
- U.S. Census Bureau, Small Area Income and Poverty Estimates (SAIPE) Program, available at: <https://www.census.gov/programs-surveys/saipe.html> (accessed June 2017).
- U.S. Census Bureau, State & Local Government Finance. Available at: <https://www.census.gov/govs/local/> (accessed June 2017).

- U.S. Department of Agriculture, County-Level Data Sets Download Data, available at: <https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/> (accessed October 2017).
- Weber, Jeremy. (2012). The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. *Energy Economics*, 34(5), 5, 1580-1588.
- Weber, Jeremy. (2014). A decade of natural gas development: The makings of a resource curse? *Resource and Energy Economics*, 37, 168-183.
- Wilson, R. (2016). Moving to economic opportunity: The migration response to the fracking boom. Availalbe at SSRN: <https://ssrn.com/abstract=2814147>
- Wolf, H.C. (2000). Intranational home bias in trade. *Review of Economics and Statistics*, 82(4), 555-563.

Notes

¹Calculations were based on data collected from the Bureau of Economic Analysis (BEA), Regional Database

²Fetzer (2014), Jacobsen (2016), Maniloff and Mastromonaco (2017), Alcott and Keniston (2017), Lee (2015), and Brown, (2014) also document significant short run localized economic gains stemming from the shale boom.

³Examinations of the recent shale boom are unable to say very much about the potential long run implications of short-lived resource booms. In a seminal paper, Jacobsen and Parker (2016) find that U.S. counties that experienced a resource boom in the late 1970s suffer today from higher levels of unemployment and lower levels of income. This raises the important question of whether modern shale-boom towns will prosper in the long run. Michaels (2011) considers the long run effect of oil discoveries made in the southern United States in the early 1900s. He documents modest economic gains associated with resource-based specialization.

⁴The shale boom may have had unique effects in places like Texas, Oklahoma, Louisiana, and Kansas because the pre-existing labor force and infrastructure in these states might reflect earlier oil and gas booms.

⁵Harding county, residing in the very northwest region of South Dakota does produce some Bakken oil, but a relatively small amount. In 2014, Harding county produced just 1.1 million barrels per year, compared to North Dakota which produced nearly 400 million barrels. While oil production in Harding did surge from 2011 to 2014 by 300 thousand barrels, as noted earlier, from 2006 to 2014, oil Production in North Dakota increased by 360 million barrels.

⁶James and Smith (2016) find that the oil boom in North Dakota attracted a disproportionate number of registered sex offenders, a group of people that have notoriously difficult time finding employment.

⁷More specifically, for each county we compute the Euclidean distance from a county's geographic center to the border of the nearest oil-producing county.

⁸We chose to base our treatment definition off of 2012 production levels because this year is arguably the height of the North Dakota shale boom. The set of oil-producing counties is stable over time; from 2009 to 2013, the set of oil-producing counties is constant. Mercer county in North Dakota, which is a treatment county in our analysis did not produce oil prior to 2009, and stopped producing any oil in 2014.

⁹This data was collected from the U.S. Department of Agriculture (USDA). The USDA collected the year 2000 data from the Decennial Censuses of Population. The 2013 data is actually a five year average (2011-2015) and is from the American Community Survey. This data can viewed at: <https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/>.

¹⁰Data collected from the Tax Foundation and is available at: <https://taxfoundation.org/state-individual-income-tax-rates/>

¹¹Low population counties are defined as having an average population (from 2000 to 2015) less than 30,000. The estimations obtained using low-population counties are available from the authors upon request.

¹²Feyrer et al. specifically find that \$1 million of oil and gas extraction raises BLS wages by \$33,957 at the county level, and by \$221,153 within a one hundred mile radius. Therefore, just $\$33,957/\$221,153 = 15.3\%$ of the wage increase remains in the producing county.

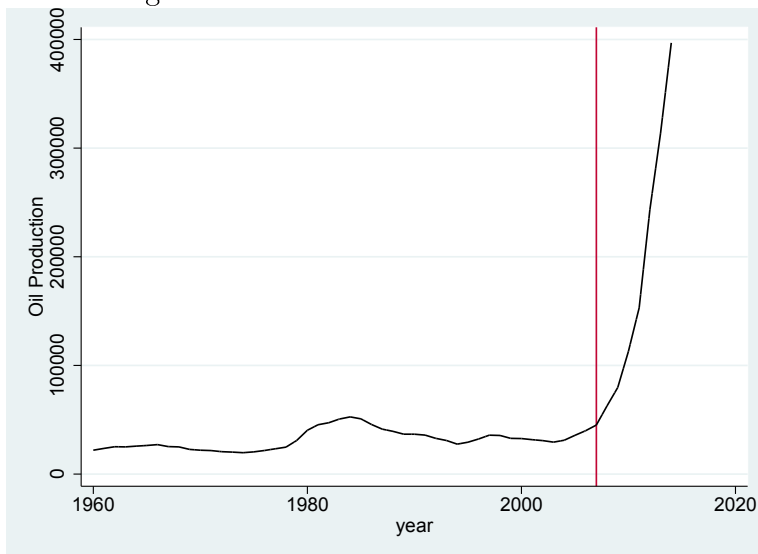
Tables and Figures

Table 1: Summary Statistics

	Mean	Std. Dev.	Min.	Max
Income Per Capita	31,166	10,013	10,035	93,530
Population	12,370	23,885	711	185,197
Unemployment	3.97	1.78	1.1	16.2
Poverty	14.04	7.67	3.7	62
%HS Deg.	33.84	4.63	26.05	47
%Coll. Deg.	31.38	3.98	25.35	40.85
Migrants/Pop	.007	.013	0	.082
Expenditures	2,655	1,430	.189	33.68
Edu. Exp.	1.407	.419	.343	6.357
Revenue	2.708	1.457	.208	33.924
Intergov. Revenue	.882	.520	.043	7.913

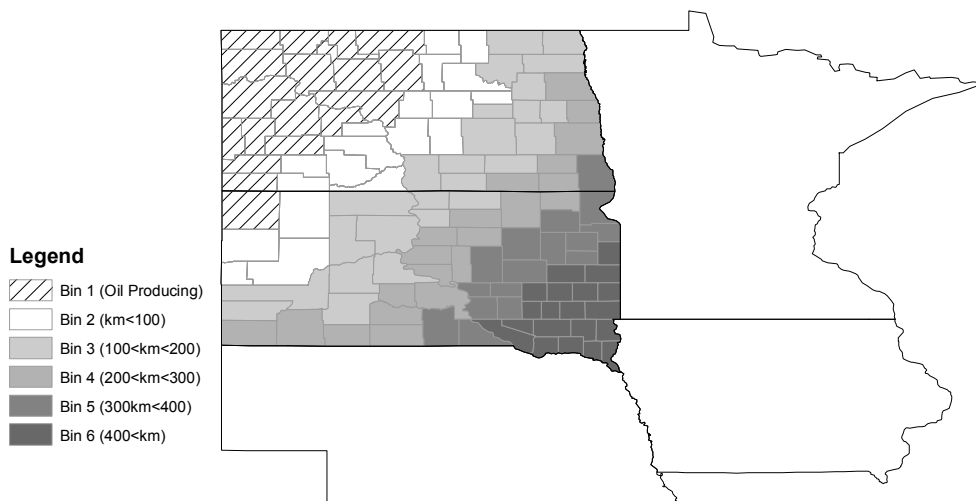
Note. Income per capita is real, and the base year is 2000. %HS Deg. refers to the percent of the county population with a maximum education level of a high school degree (averaged across 2000 and 2011-2015 periods). %Coll Deg. refers to the percent of the county population with a bachelor's degree or higher. Migrants refers to the number of people moving from a neighboring county to a Bakken county, relative to the host county's total population. All public expenditure and revenue variables are measured relative to the county population and are expressed in thousands of dollars.

Figure 1: North Dakota Oil Production



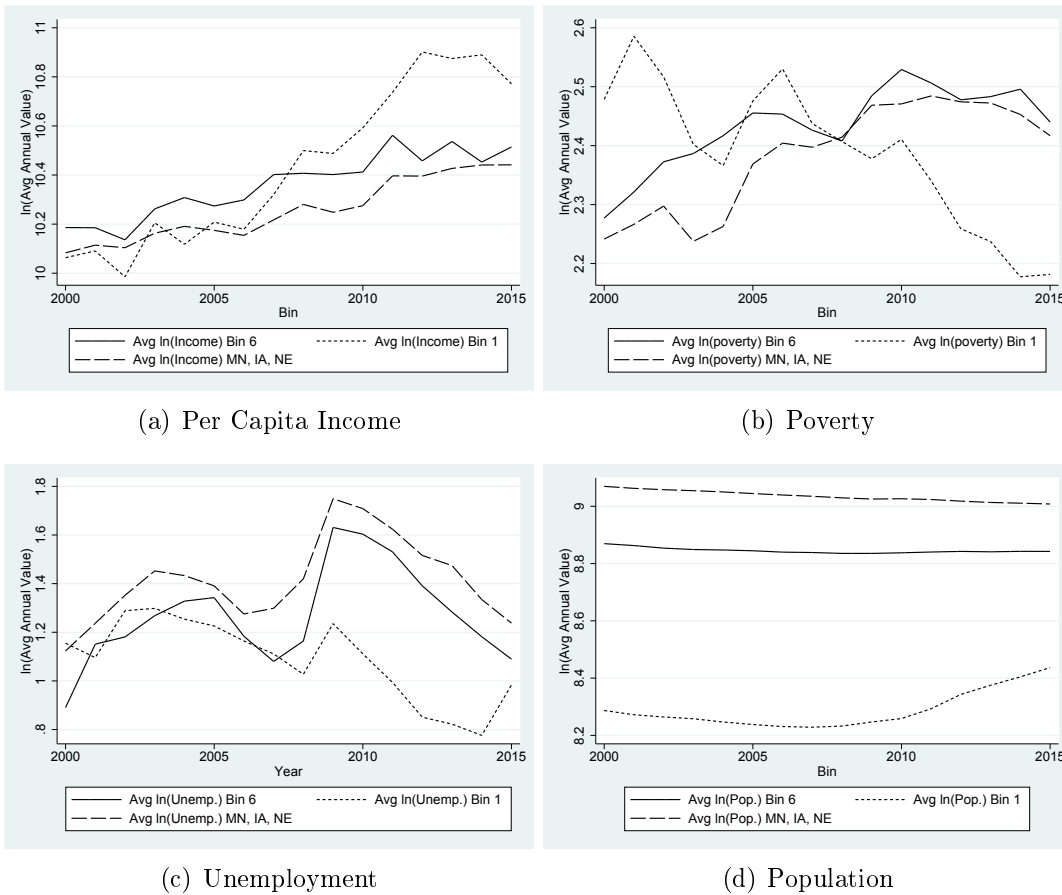
Note: Oil production is measured in thousands of barrels. A vertical line is drawn at the event date, 2007. Data collected from the Energy Information Administration and is available at: www.eia.gov/state/

Figure 2: North and South Dakota Counties



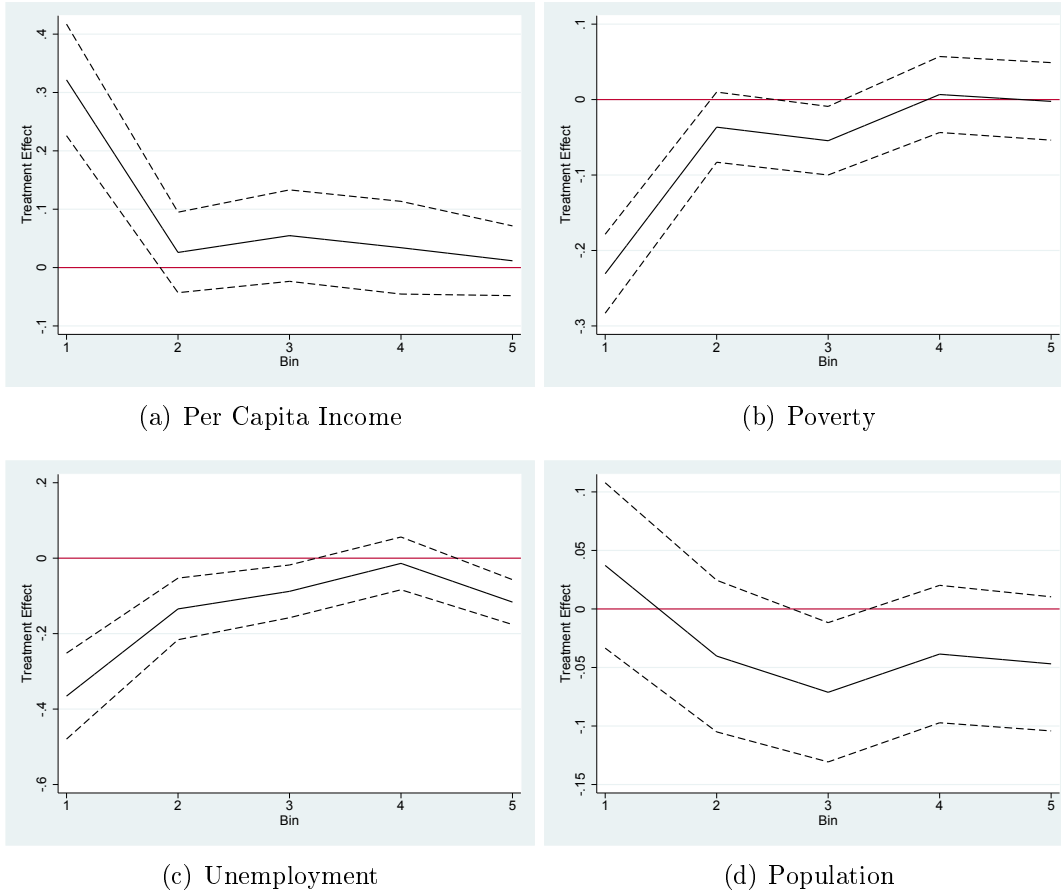
Note: State borders are defined in bold. North Dakota is the top left state, South Dakota is the bottom left state. The other three states are Minnesota, Iowa, and Nebraska.

Figure 3: Treatment and Controls Over Time



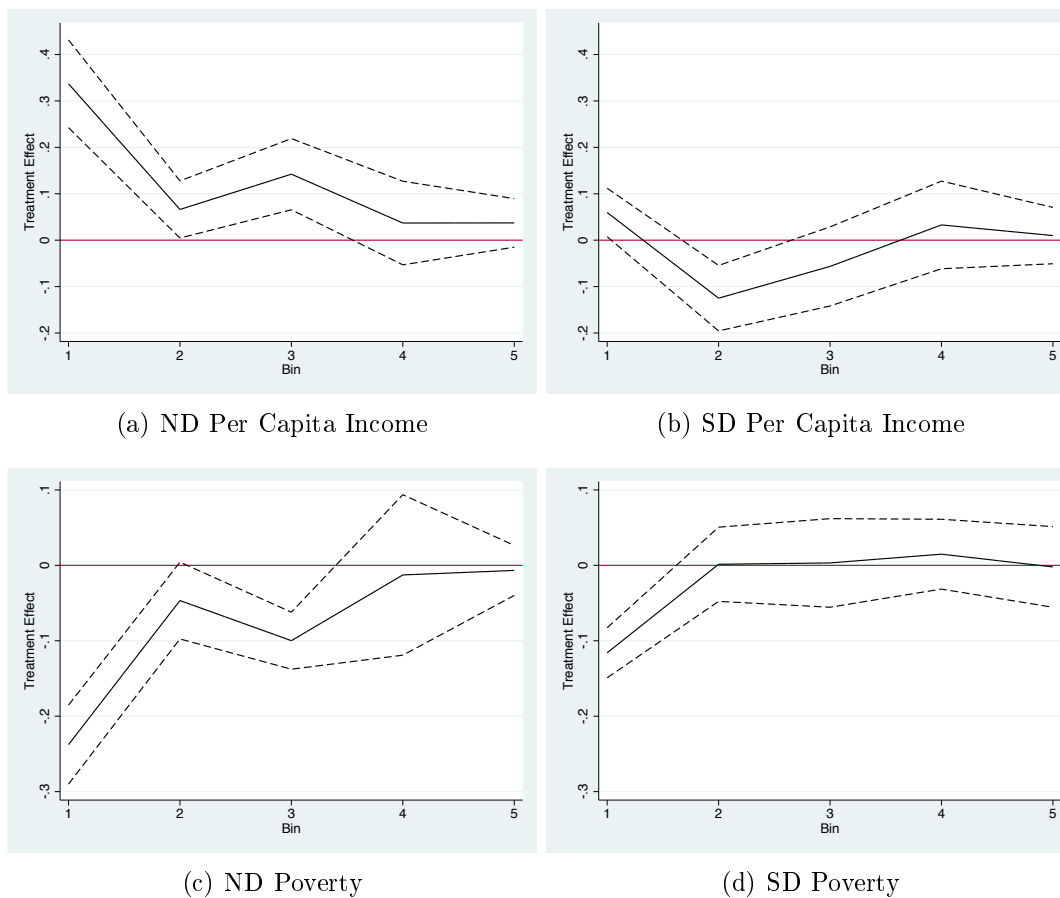
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate.

Figure 4: Estimation of Equation 1: Economic & Demographic Outcomes



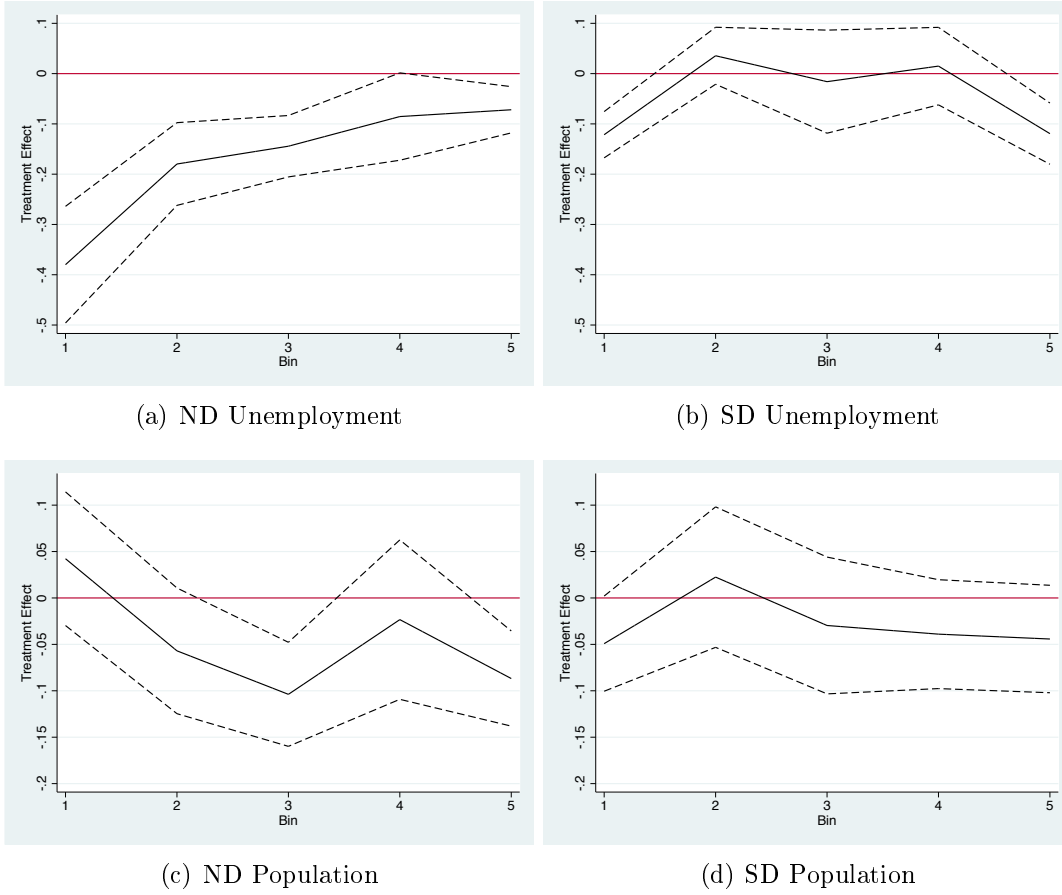
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. “Treatment Effect” is the estimate of β_n from equation 1. 95% confidence intervals are given. For panel (a) $N = 1,904$ and $R^2 = .351$, for panel (b) $N = 1,874$ and $R^2 = .039$, for panel (c) $N = 1,904$ and $R^2 = .144$, and for panel (d) $N = 1,904$ and $R^2 = .002$.

Figure 5: Estimation of Equation 2: Income & Poverty



Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. For North Dakota, “Treatment Effect” is the estimate of γ_n and for South Dakota is the estimate of $(\gamma_n + \lambda_n)$ from equation 2. 95% confidence intervals are given. The corresponding sample size for panels (a) and (b) is $N = 1,904$ and $R^2 = .385$. For panels (c) and (d) the sample size is $N = 1,874$ and $R^2 = .064$.

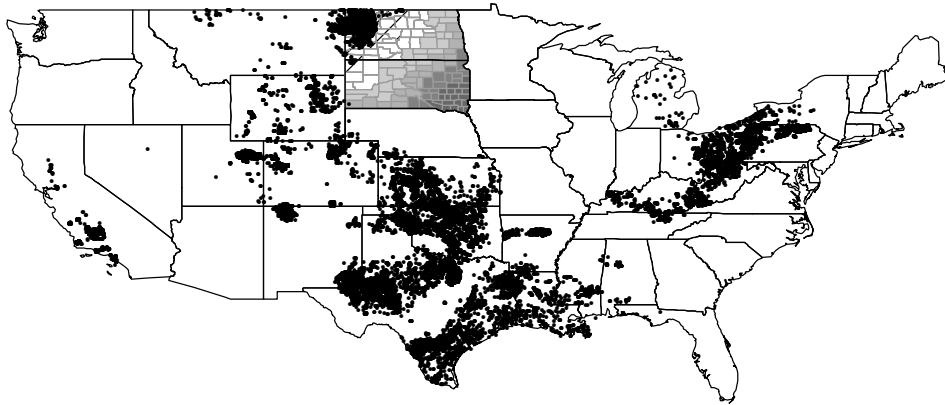
Figure 6: Estimation of Equation 2: Unemployment & Population



Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. For North Dakota, “Treatment Effect” is the estimate of γ_n and for South Dakota is the estimate of $(\gamma_n + \lambda_n)$ from equation 2. 95% confidence intervals are given. The corresponding sample size for panels (a) and (b) is $N = 1,904$ and $R^2 = .154$. For panels (c) and (d) the sample size is $N = 1,904$ and $R^2 = .007$.

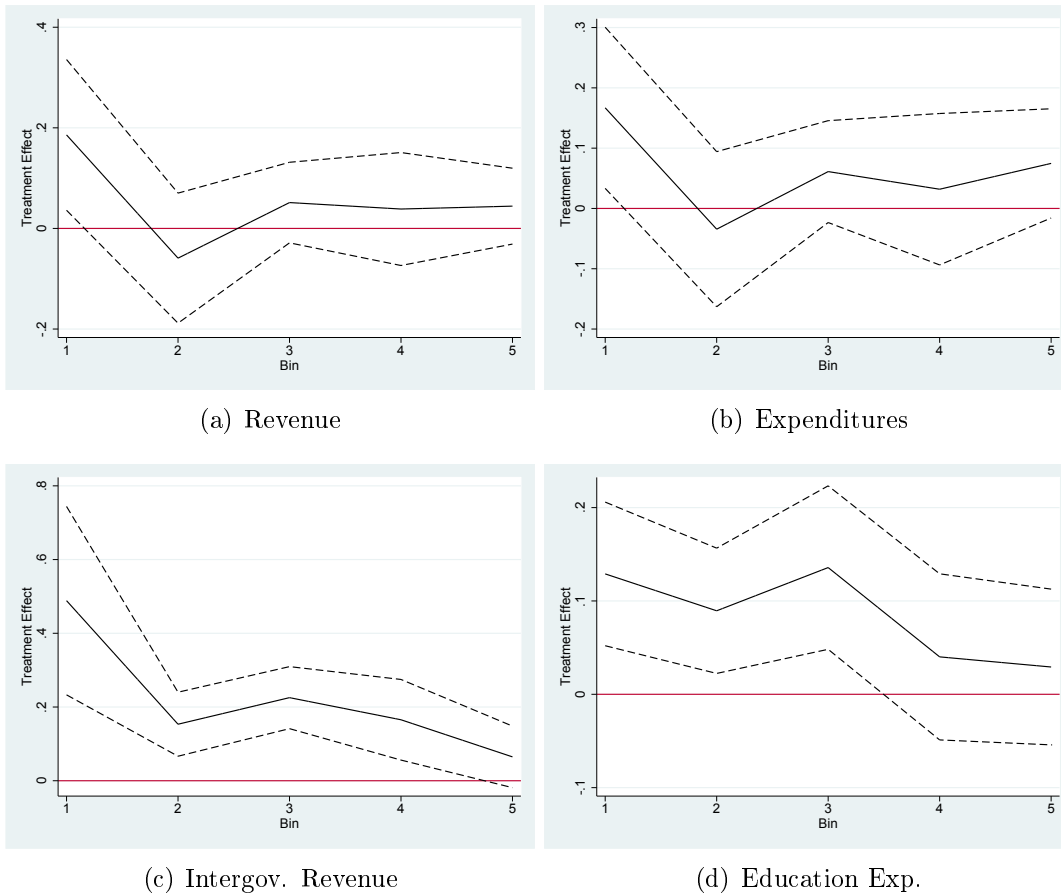
A Appendix

Figure A.1: New Oil Wells 2012



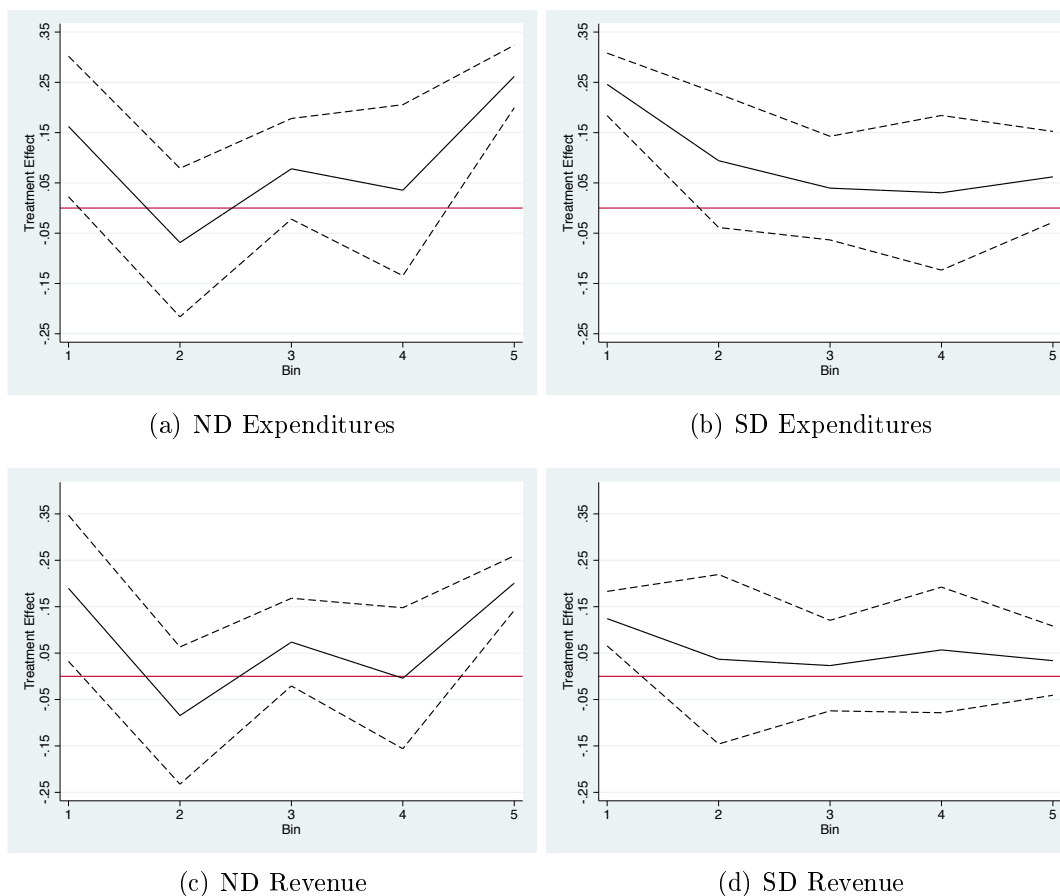
Note: Each point represents a new oil or gas well drilled in 2012. Drilling data was generously provided by drillinginfo.com.

Figure A.2: Estimation of Equation 1: Public Finance Outcomes



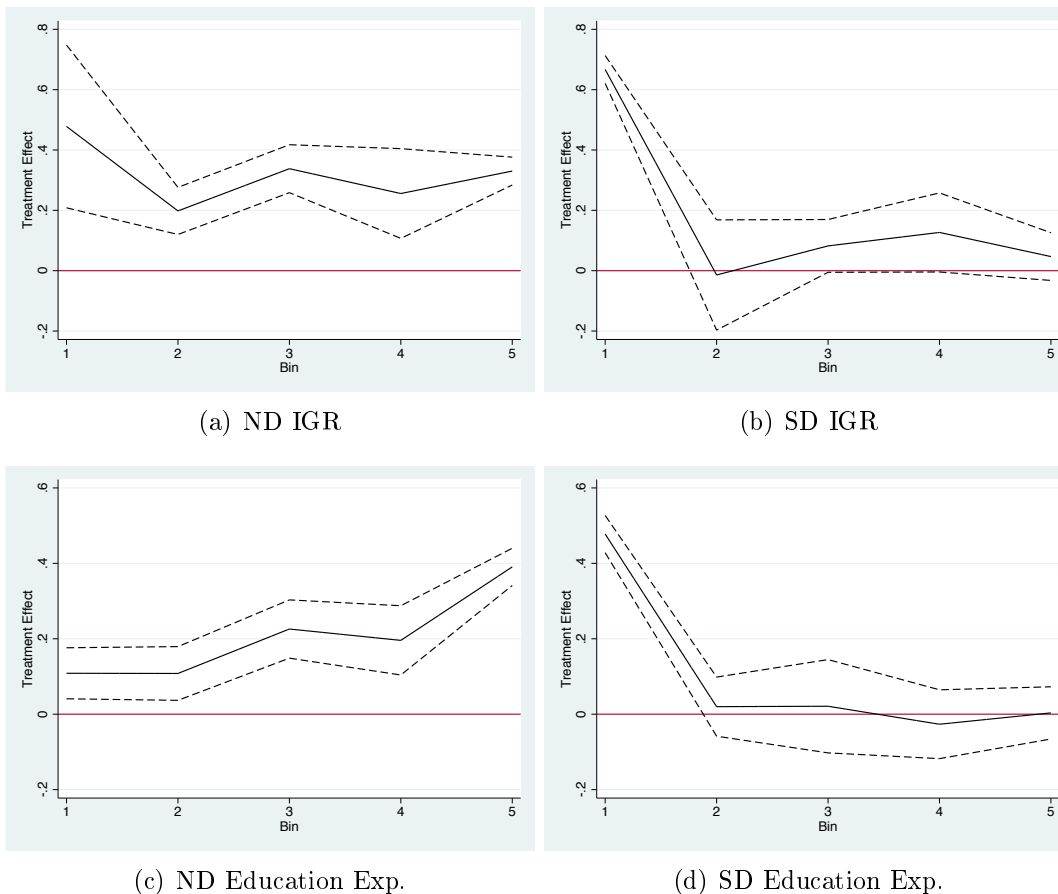
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Intervog. Revenue” refers to intergovernmental revenue to local governments from the state government. “Education Exp.” refers to total (k-12 and higher education) local expenditures on education. “Treatment Effect” is the estimate of β_n from equation 1. 95% confidence intervals are given. The corresponding sample size for panels (a), (b), and (c) is $N = 1,531$. For panel (d) $N = 1,521$. The overall R^2 for panel (a) is .215, for panel (b) is .178, for panel (c) is .207, and for panel (d) is .066.

Figure A.3: Estimation of Equation 2: Public Finance, Expenditures & Revenue



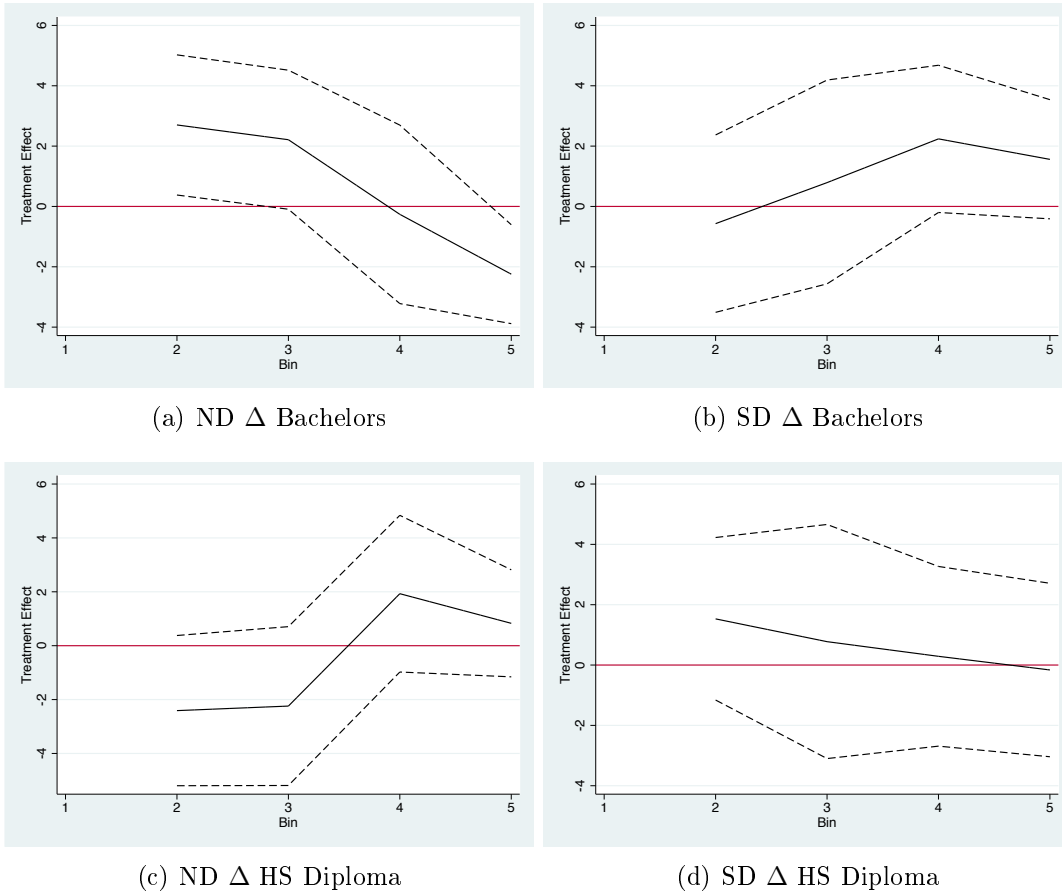
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. For North Dakota, “Treatment Effect” is the estimate of γ_n and for South Dakota is the estimate of $(\gamma_n + \lambda_n)$ from equation 2. 95% confidence intervals are given. For panels (a) and (b) $R^2 = .181$ and $N = 1,531$. For panels (c) and (d) $R^2 = .216$ and $N = 1,531$.

Figure A.4: Estimation of Equation 2: Public Finance, IGR & Education



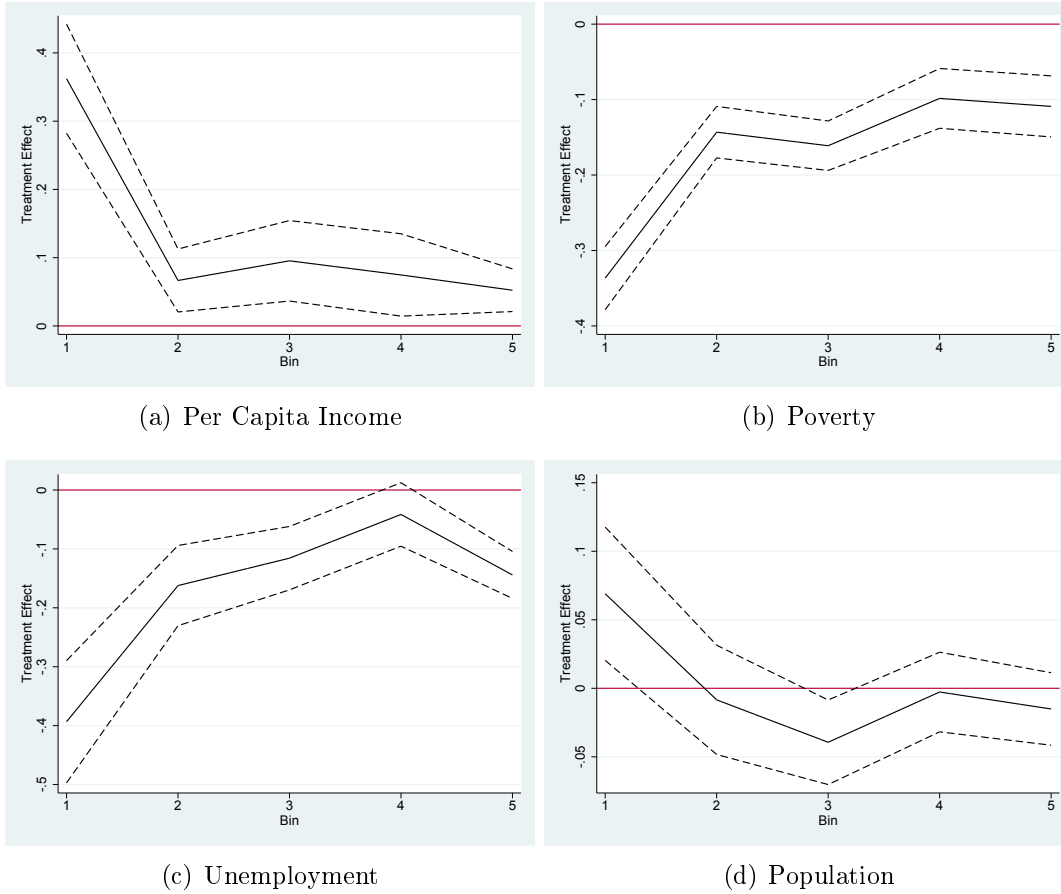
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. For North Dakota, “Treatment Effect” is the estimate of γ_n and for South Dakota is the estimate of $(\gamma_n + \lambda_n)$ from equation 2. 95% confidence intervals are given. For each outcome variable $N = 1,531$. For panels (a) and (b) $R^2 = .250$ and $N = 1,531$. For panels (c) and (d) $R^2 = .086$ and $N = 1,521$.

Figure A.5: Change in Education Attainment



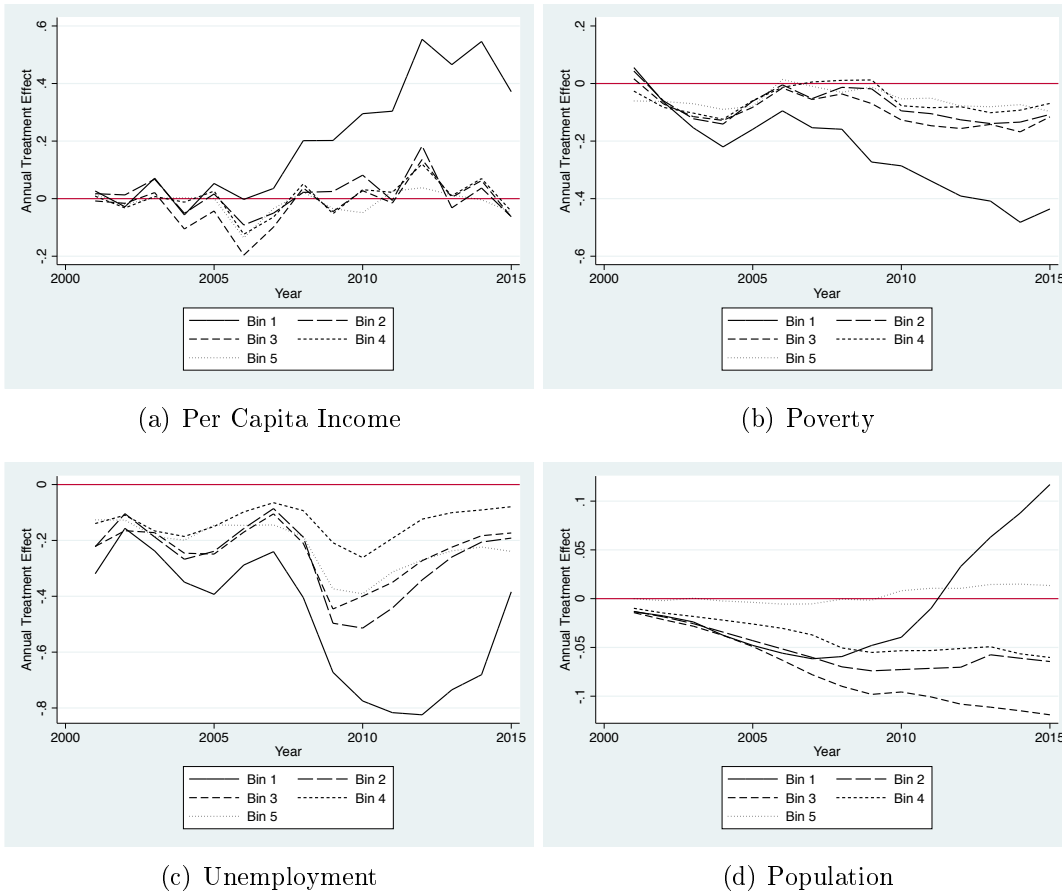
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. Panels (a) and (b) report the results from a cross-county regression in which the dependent variable is the absolute change in the percent of county population with at least a college diploma. For panels (c) and (d) the dependent variable is the percent of the population with at most a high school diploma. 95% confidence intervals are given. Independent variables include the bin indicator variables, and their interaction with the indicator variable for South Dakota. The corresponding sample size is $N = 101$. For panels (a) and (b) $R^2 = .112$, and for panel (c) and (d) $R^2 = .101$

Figure A.6: Economic & Demographic Outcomes: Alternative Controls



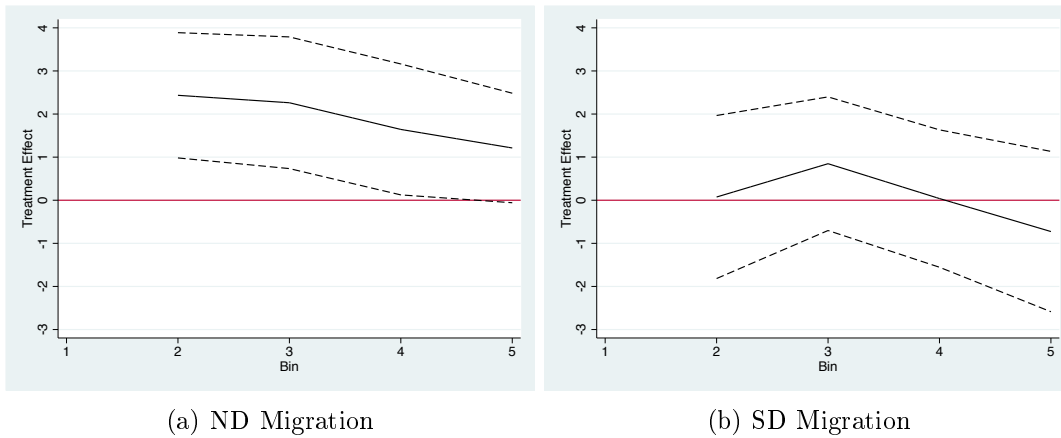
Note: Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on. “Poverty” refers to the poverty rate. “Unemployment” refers to the unemployment rate. “Treatment Effect” is the estimate of β_n from equation 1. 95% confidence intervals are given. For panel (a) $N = 6,048$ and $R^2 = .357$, for panel (b) $N = 5,961$ and $R^2 = .023$, for panel (c) $N = 6,048$ and $R^2 = .251$, and for panel (d) $N = 6,048$ and $R^2 = .0001$.

Figure A.7: Heterogeneous Treatment Effects



Note: These are the results from re-estimating the baseline specification in which each bin indicator is interacted with year indicators. Year and county fixed effects are included in each regression. Confidence intervals are available upon request. Each line represents an annual treatment effect; it is the effect of being in a given bin, in a given year, relative to being in that bin in the year 2000.

Figure A.8: County-to-County Migration



(a) ND Migration

(b) SD Migration

Note: The solid line gives the treatment effects estimated using equation 3. Bin 1 refers to oil-producing counties, Bin 2 refers to those within 100 kilometers of oil-producing counties, Bin 3 refers to those within 200 but in excess of 100 kilometers from oil-producing counties, and so on (see Figure 2). 95% confidence intervals are given. The corresponding $R^2 = .462$ and $N = 55$.