Regional Economic Impacts of the Shale Gas and Tight Oil Boom: A Synthetic Control Analysis

Abdul Munasib and Dan S. Rickman

University of Georgia, Oklahoma State University

27. July 2014

Regional Economic Impacts of the Shale Gas and Tight Oil Boom: A Synthetic Control Analysis

by

Abdul Munasib[1]

and

Dan S. Rickman[2]

Abstract: The dramatic increase in oil and gas production from shale formations has led to intense interest in its impact on local area economies. Exploration, drilling and extraction are associated with direct increases in employment and income in the energy industry, but little is known about the impacts on other parts of local economies. Increased energy sector employment and income can have positive spillover effects through increased purchases of intermediate goods and induced local spending. Negative spillover effects can occur through rising local factor and goods prices and adverse effects on the local area quality of life. Therefore, this paper examines the net economic impacts of oil and gas production from shale formations for key shale oil and gas producing areas in Arkansas, North Dakota and Pennsylvania. The synthetic control method (Abadie and Gardeazabal 2003; Abadie et al., 2010) is used to establish a baseline projection for the local economies in the absence of increased energy development, allowing for estimation of the net regional economic effects of increased shale oil and gas production.

JEL Codes: Q32; R11

1. Department of Agricultural and Applied Economics, University of Georgia; 1109 Experiment Street Griffin, GA 30223-1797; Phone: 770-228-7231; Email: munasib@uga.edu.

2. Department of Economics, Business College, Oklahoma State University, Stillwater, OK 74078, USA. Phone: 405-744-1434; Fax: 405-744-5180; Email: dan.rickman@okstate.edu; (corresponding author)

1. Introduction

Following decades of concerns with U.S. dependence on energy imports, a new combination of horizontal drilling and hydraulic fracturing during the previous decade led to dramatic increases in U.S. energy production. The percentage of all wells drilled horizontally increased from around ten percent at the beginning of 2005 to over fifty eight percent by the end of 2011, and to over sixty seven percent by the middle of 2014 (Baker Hughes, 2014). Production of natural gas increased over thirty five percent from 2005 to 2013, while production of oil increased nearly forty four percent from 2005 to 2013.[1] In its 2013 Annual Energy Outlook, the U.S. Energy Information Administration (U.S. Energy Information Administration, 2013) projects a one hundred thirteen percent increase in U.S. shale gas production by 2040, raising its share of total natural gas production from thirty four to fifty percent. EIA projects tight oil production, which includes oil produced from "very low permeability shale, sandstone, and carbonate formations" (U.S. Energy Information Administration, 2013, page 82), to peak in 2021 at nearly triple the 2011 level. The dramatic increase and projected growing importance of unconventional oil and gas extraction has spawned intense interest in both its potential economic benefits and potential adverse impacts on local populations.

A study by IHS (2012) that was supported by the American Petroleum Institute**,** the Institute for 21st Century Energy, the American Chemistry Council, and the Natural Gas Supply Association, estimated the number of U.S. jobs associated with unconventional oil and gas production to be 1.7 million jobs in 2012, projecting them to reach 3.5 million by 2035. The study was based on the use of IMPLAN, a widely used input-output economic impact modeling

[1]Natural gas production data were obtained from the U.S. Energy Information Administration at http://www.eia.gov/dnav/ng/hist/n9050us2M.htm, while oil production data were obtained at http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=MCRFPUS1&f=A, both on June 13, 2014.

system. However, in a review of several studies sponsored by the gas industry that use similar methodology, Kinnaman (2011) finds the studies of economic impacts to be based on questionable assumptions that likely overstate the economic benefits of shale gas extraction: e.g., assumptions of excess supply in the economy that ignore potential crowding-out effects and a lack of economy-wide consistency in attributing exogenous impacts that lead to over counting economic impacts of energy development. Others contend that adverse effects on the local environment and quality of life may negatively affect agriculture and tourism (White, 2012; Lydersen, 2013). Reduced quality of life also may inhibit in-migration of households, reducing population and employment growth. Input-output models and standard econometric models used in the industry-sponsored studies do not account for these potential adverse effects. Kinnaman (2011) notes the paucity or near absence of relevant studies that have gone through the peer review process of an economic journal.[2]

Therefore, this study examines the net economic impacts of oil and gas production from shale formations for key energy producing areas. The areas chosen are located in the states of Arkansas, North Dakota and Pennsylvania; all three states were ranked in the top-ten oil and gas producing states by IHS (2012), but had more limited energy sector employment prior to the shale oil and gas boom, unlike states like Oklahoma and Texas. The synthetic control method (Abadie et al., 2010) (SCM) is used to establish a baseline projection for the local economies in the absence of increased shale-based energy extraction, allowing for estimation of its net regional economic effects. The estimated effects reveal the balance of potential positive economic

[2] A number of studies have examined whether there is a resource curse in the United States (e.g., Black et al., 2005; Papyrakis and Gerlagh, 2007; James and Aadland, 2011; Michaels, 2011; Douglas and Walker, 2012), but they are focused on energy development in regions broadly, where fluctuations in energy prices and other long-term trends related to energy development in the local areas would confound any estimates of the effects of unconventional energy extraction in the areas of study.

impacts versus negative economic impacts, making SCM preferred over input-output-based studies, which by design only capture potential increases in local spending.

An advantage of SCM is transparency in constructing the counterfactual. It is a weighted-average of comparison/control units based on demonstrated affinities. In SCM no single match with all the comparable characteristics to the shale oil and gas areas is required as it is in case studies or some matching approaches. We employ permutations or randomization tests for inference which, given the problem in hand, is an improvement over standard-error-based inference in regression models. Regional economic variables examined include total employment, wage and salary employment, per capita personal income, population and the poverty rate. Wage and salary employment also is examined for the sectors Accommodation and Food Services, Construction and Retail. Because of geographic spillovers aggregates of counties are examined. The impacts also are only estimated for nonmetropolitan counties to avoid potential confounding influences in metropolitan areas given their economic size. The impacts are first estimated for all nonmetropolitan oil and gas counties in each state. Then to capture potential broader geographic spillovers the impacts are estimated for all nonmetropolitan counties in each state.

The next section discusses the potential channels of influence, both positive and negative, of unconventional oil and gas extraction on the regional economy. In so doing, key findings of related studies are presented and critiqued. Section 3 presents the empirical approach, including a description of the use of the synthetic control method, variable selection, and data sources. The results are presented and discussed in Section 4. The results suggest there are significantly positive benefits across nonmetropolitan North Dakota oil and gas counties for a wide range of regional economic measures. There are limited geographic spillovers, however, from the oil and gas counties into other North Dakota nonmetropolitan counties. Significantly positive effects are found in some of the employment measures for only a subset of Arkansas oil and gas producing counties, while no effects are found for Pennsylvania, including for subsets of its oil and gas producing counties. Back of the envelope calculations of likely wage and salary employment multipliers suggest that actual multiplier effects of shale oil and gas extraction likely fall well below estimates produced by input-output models. Section 5 contains summary statements and conclusions.

2. Unconventional Gas and Oil Development and the Local Economy

Exploration, drilling and extraction of unconventional gas and oil are associated with direct increases in employment and income in the industry. Increased energy sector activity can have positive local spillover effects through increased firm purchases of locally-produced goods and services from other sectors (intermediate goods and services) and increased local spending by energy sector workers. Yet, negative spillover effects can occur through rising local factor and goods prices associated with increased local demand and adverse effects on the local area quality of life (White, 2012; Lydersen, 2013).

Industry sponsored studies of the economic impacts of unconventional gas and oil activities focused on the positive spillover effects, using tools that are designed to solely capture the positive spending effects (Kinnaman, 2011). Using the IMPLAN input-output model and their own database on trade flows, IHS (2012) estimates that over 1.2 million jobs were created during 2012 in unconventional gas and oil producing states, ranging from 576 thousand in Texas to 33 thousand in Arkansas. For the two other states of interest in this study, over 100 thousand jobs were estimated to be created in Pennsylvania and over 70 thousand in North Dakota. Over the forecast period, on average across the states only 20 percent of total estimated job gains are direct, implying an employment multiplier of 5.

Using the IMPLAN model and adjusting it with survey data, Considine et al. (2009) estimate that over 29 thousand jobs were created in Pennsylvania during 2008 by unconventional oil and gas activity. Considine et al. (2010) updated the earlier study to estimate that over 44 thousand Pennsylvania jobs were created in 2009 by unconventional oil and gas activity. Kinnaman (2011) questions the assumptions of the study: 1) that all lease and royalty payments are made in Pennsylvania; and 2) that 95 percent of all direct expenses occur in Pennsylvania. It also is not clear whether the payments should be entered in the input-output model as direct payments, as it appears from the descriptions in the studies, because the IMPLAN SAM may already account for them when the energy sector is directly stimulated, which would lead to double counting.

To be sure, Kelsey et al. (2009) reported that 37 percent of workers in shale gas development in Pennsylvania were out-of-state residents, while landowners saved about 55 percent of their royalties/lease payments. Thus, the total job impacts were estimated to be in the range of 23 to 24 thousand in 2009. They report an output multiplier of approximately 1.9.

Similarly, using the IMPLAN model and an industry survey, Center for Business and Economic Research (2008) of the University of Arkansas estimated total job impacts of nearly 10 thousand in Arkansas during 2007 from production in the Fayetteville shale play. The total employment impact projections for 2008-2012 ranged between 11 and 12 thousand. Employment multipliers across the years were in the range of 2.5-2.64.

Other research suggests that input-output models overstate the economic impacts of export-based activity in general. For example, Edmiston (2004) finds that input-output models overstated the multiplier effects of large new manufacturing plants. Computable general equilibrium (CGE) analysis of Harrigan and McGregor (1989) and Rickman (1992) suggest that the general overstatement of multiplier effects by input-output models relates to the absence of prices in the models and implicit assumptions of perfectly elastic supply. In CGE models with less than perfectly elastic supply, increased direct economic activity places upward pressure on prices, making other industries less competitive and reducing demand. This offsets the positive spillover effects from increased intermediate purchases and induced spending captured by input-output models. This is a phenomenon often noted in the resource curse literature (see footnote 2) and is very possible in some areas where unconventional oil and gas extraction is occurring.

Adverse effects on the natural environment and local quality of life also can offset economic gains associated with energy development, both to area resident well-being and to economic growth through negative feedback effects on tourism and migration. A number of potential risks to the local areas have been identified in the literature (Lipscomb et al., 2012; Rahm, 2011, White, 2012; Atkin, 2014): contamination of ground water, accidental chemical spills, reduction in air quality (e.g., dust, diesel fumes), noise, land footprint of drilling activities, earthquake frequency, pipeline placement and safety and the volume of water used in unconventional energy extraction.

Evidence of the adverse economic effects of these can be found in studies of housing values near unconventional energy producing sites. Gopalakrishan and Klaiber (2014) find negative effects on property values in Washington County, Pennsylvania, for proximity to shale gas exploration during 2008 to 2010. Muehlenbachs et al. (2014) find large negative effects for homes dependent on groundwater, though small positive effects are reported for homes with piped-in water. Using contingent valuation surveys in Florida and Texas, Throupe et al. (2013)

find reductions in bid values for homes located near what they refer to as “fracking” sites, with less of an effect in areas that are more familiar with fracking.

Therefore, to capture the net effects, other studies examine aggregate data to estimate the impacts of unconventional oil and gas development on local area economies. Murray and Ooms (2008) conducted case studies of Denton, Texas, Faulkner and White Counties in Arkansas, and counties in northeast Pennsylvania where limited shale development occurred prior to 2006. The authors examined growth in these areas, concluding it was dramatic compared to other areas. However, as Kinnaman (2011) points out, a difference-in differences approach would be much better suited to reach such a conclusion than simple case study analysis.

Weinstein and Partridge (2011) use difference-in-differences in comparing employment and per capita income growth between Pennsylvania shale drilling counties and non-shale drilling counties matched on population, urbanization and location in the state, before and after 2005 with 2009 as the end year. They find positive effects for per capita income for shale counties in both the north and south, but only find positive employment effects for those in the southern part of the state. In further analysis, they find no significant employment effects in shale counties when compared to all counties in Pennsylvania, where again there are significant estimated per capita income effects. A problem with using other counties in the same state as control groups though is that there may be geographic spillovers associated with unconventional energy extraction, such that what are considered control groups may in fact be receiving a treatment effect, which would tend to understate the treatment effect in the counties where the drilling is occurring.

Weber (2012) compares economic outcomes in natural gas boom counties in Colorado, Texas and Wyoming using a triple difference-in-difference approach. Counties that share a border with a boom county are omitted from the non-boom control group. He finds effects that suggest that the reported input-output estimates for the Fayetteville and Marcellus shale gas formations may be too large. Yet, it is possible that geographic spillovers reach beyond border counties, such that non-boom counties have a treatment effect. These states also have had a long history of energy development, making it difficult to isolate the effects of the unconventional gas drilling from other influences such as the change in energy prices during the period.

Similarly, Brown (2014) focuses on the effects of natural gas production during 2001 to 2011 on 647 nonmetropolitan counties in a nine state region, mostly comprising the 10th Federal Reserve District. Using regression analysis, he finds faster growth of employment, population, real personal income and wages in counties with increased natural gas production relative to those with declining production and with no production. He estimates an employment multiplier of 1.7, well below those reported in the input-output studies. Employment in construction, transportation and services had positive relative employment growth in counties with increased natural gas production, while there were not any relative employment effects in manufacturing and retail. Yet, the study faces the same limitation of others in ignoring potential geographic spillovers in specifying treatment counties versus control counties. As with Weber (2012), the use of counties with long standing energy development in the control set can confound the separate effects of unconventional oil and gas extraction.

3. Empirical Approach

We examine the nonmetropolitan oil and gas counties in Arkansas, North Dakota, and Pennsylvania, counties that experienced unconventional drilling in the Bakken, Fayetteville and Marcellus shale plays. These counties had limited or stable energy activity prior to the emergence of unconventional drilling, making it easier to identify a treatment effect. They also have been subjects of prior studies using simple impact analysis tools. The counties examined for each state are listed in Table 1.

Only nonmetropolitan counties are examined because it is more likely that the effects of energy extraction can be detected given the overall smaller size of nonmetropolitan local economies and greater relative size of the energy sector (Weinstein and Partridge, 2011). Because of potential geographic spillovers across county lines we examine the aggregate of the oil and gas counties in each state. Potential heterogeneity in effects within a shale play, and difficulties in detecting oil and gas shale impacts in larger economies, lead us to also consider sub-groupings of oil and gas counties in Arkansas and Pennsylvania. Finally, to detect geographic spillovers more broadly we examine all nonmetropolitan counties together as an aggregate in each state.

Regional economic variables examined include total employment, wage and salary employment, per capita personal income, population and the poverty rate. Wage and salary employment also is examined for the sectors Accommodation and Food Services, Construction, and Retail. Employment data for the non-oil and gas sectors are used to assess possible positive or negative spillovers on other sectors. A variety of outcome measures are chosen for the analysis because general well-being in an area may not necessarily be reflected by a single variable (Partridge and Rickman, 2003).

We use the synthetic control method (SCM) (Abadie and Gardeazabal 2003; Abadie et al., 2010) to estimate the counterfactual – a baseline projection for the local economies in the absence of unconventional drilling (the intervention) – for each of the exposed/treatment states Arkansas, North Dakota, and Pennsylvania. The counterfactual in each case is a weighted average of the donor pool (the set of unexposed/control units) consisting of the aggregate nonmetropolitan portions of the states that have not been exposed to unconventional drilling. A comparison of the counterfactual and the actual outcome provides an estimate of the net regional economic impact of unconventional oil and gas activities in Arkansas, North Dakota, and Pennsylvania. We estimate the impact on each outcome for regions of each of the three exposed states separately. Below is a detailed description of the method and its advantages.

3.1. Synthetic Control Method (SCM)

There are a number of advantages to using SCM in this study. First, in program evaluation, researchers often select comparisons on the basis of subjective measures of similarity between the affected and the unaffected regions or states. But, neither the set of all non-shale regions or states nor a single non-shale region or state likely approximates the most relevant characteristics of a treatment region or state.

SCM, in contrast, provides a comparison state (or synthetic) that is a combination of the control states – a data-driven procedure that calculates 'optimal' weights to be assigned to each state in the control group based on *pre-intervention* characteristics (Abadie and Gardeazabal 2003; Abadie et al., 2010) – thus, making explicit the relative contribution of each control unit to the counterfactual of interest. With reduced discretion in the choice of the comparison control units, the researcher is forced to demonstrate the affinities between the affected and unaffected units using observed characteristics.

Secondly, when aggregate data are employed (as the case is in this paper) the uncertainty remains about the ability of the control group to reproduce the counterfactual outcome that the affected unit would have exhibited in the absence of the intervention. This type of uncertainty is not reflected by the standard errors constructed with traditional inferential techniques for comparative case studies. As Buchmueller et al. (2011) explain, in a 'clustering' framework, inference is based on the asymptotic assumption, i.e., the number of units grows large. Naturally, this does not apply in our case as our focus is one state at a time. The comparison of a single state against all other states in the control group collapses the degrees of freedom and results in much larger sample variance compared to the one typically obtained under conventional asymptotic framework and can seriously overstate significance of the intervention (Donald and Lang, 2007, Buchmueller et al., 2011). In other words, it becomes difficult to argue that the observed conditional difference in measured outcome is entirely due to the intervention. Bertrand et al. (2004) also emphasize that regression-based difference-in-difference analyses tend to overstate the significance of the policy intervention. We, therefore, apply the permutations or randomization test (Bertrand et al., 2004, Abadie et al., 2010, Buchmueller et al., 2011, Bohn et al., 2014) that the SCM readily provides.

Finally, because the choice of a synthetic control does not require access to post-intervention outcomes, SCM allows us to decide on a study design without knowing its bearing on its findings (Abadie et al., 2010). The ability to make decisions on research design while remaining blind to how each particular decision affects the conclusions of the study is a safeguard against actions motivated by a ‘desired’ finding (Rubin 2001).

3.1.1. The Synthetic Control

The following exposition is based on Abadie and Gardeazabal (2003) and Abadie et al. (2010). For states $i = 1,..., J+1$ and periods $t = 1,...,T$, suppose state $i = 1$ is exposed to the intervention (unconventional drilling) at $T_0 \in (1,T)$. The observed outcome for any state i at time t is,

(1) $$Y_{it} = Y_{it}^N + \alpha_{it} S_{it},$$

where Y_{it}^N is the outcome for state i at time t in the absence of the intervention, the binary indicator variable S_{it} denotes the intervention taking the value 1 if $i = 1$ and $t > T_0$, and α_{it} is the effect of the intervention for state i at time t. We restrict the donor pool based on an absence of proven reserves and/or a lack of oil and gas employment over the periods $t = 1,...,T$, though we experiment with this in sensitivity analysis. We assume that the intervention had no effect on the outcome in the exposed state before it took place, and that the outcome of the donor pool states were not affected by the intervention in the exposed state.

We want to estimate $(\alpha_{1T_0+1},...,\alpha_{1T})$. From equation (1) we note that $\alpha_{1t} = Y_{1t} - Y_{1t}^N$ for $t \in \{T_0+1,...,T\}$, and while Y_{1t} is observed Y_{1t}^N is unobserved. We, therefore, need to estimate Y_{1t}^N. Suppose Y_{it}^N is given by the model,

(2) $$Y_{it}^N = \delta_t + \boldsymbol{\theta}_t \mathbf{Z}_t + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{it},$$

where, δ_t is an unknown common factor constant across states, $\mathbf{Z}_t$ is a $(r \times 1)$ vector of observed predictors (not affected by the intervention), $\boldsymbol{\theta}_t$ is a $(1 \times r)$ vector of unknown parameters, $\boldsymbol{\lambda}_t$ is a $(1 \times F)$ vector of unobserved time-varying common factors, $\boldsymbol{\mu}_i$ is a $(F \times 1)$ vector of unknown unit specific factors, and ε_{it} are the unobserved transitory shocks at the state level with zero mean.

Consider a $(J \times 1)$ vector of weights $\mathbf{W} = (w_2,...,w_{J+1})'$ such that $\{w_j \geq 0 \mid j = 2,...,J+1\}$ and $\sum_{j=2}^{J+1} w_j = 1$. Each value of the vector $\mathbf{W}$ represents a weighted average of the control states and, hence, a potential synthetic control. Abadie et al. (2010) show that there exist $\mathbf{W}^* = (w_2^*,...,w_{J+1}^*)'$ such that, $Y_{1t}^N = \sum_{j=2}^{J+1} w_j^* Y_{jt}$, $t = 1,...T_0$, and $\mathbf{Z_1} = \sum_{j=2}^{J+1} w_j^* \mathbf{Z}_j$, that is, pre-intervention matching is achieved with respect to the outcome variable as well as the predictors (the procedure to obtain $\mathbf{W}^*$ is explained in the Appendix). Then, under standard conditions, we can use,

$$(3) \qquad \hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}, \qquad t \in \{T_0 + 1, ..., T\},$$

as an estimator for α_{1t}. The term $\sum_{j=2}^{J+1} w_j^* Y_{jt}$ on the right-hand-side of (3) is simply the weighted average of the observed outcome of the control states for $t \in \{T_0 + 1, ..., T\}$ with weights $\mathbf{W}^*$.

3.1.2. Inference

Once an optimal weighting vector $\mathbf{W}^*$ is chosen, the "synthetic" is obtained by calculating the weighted average of the donor pool. The post-intervention values of the synthetic control serve as our counterfactual outcome for the treatment state. The post-intervention gap between the actual outcome and the synthetic outcome, therefore, captures the impact of the intervention.

To begin, we follow Bohn et al. (2014) and calculate a difference-in-difference estimate for the treatment state,

$$(4) \qquad \Delta_{TR} = \left|\bar{Y}_{TR,actual}^{post} - \bar{Y}_{TR,synthetic}^{post}\right| - \left|\bar{Y}_{TR,actual}^{pre} - \bar{Y}_{TR,synthetic}^{pre}\right|,$$

where $\bar{Y}_{TR,actual}^{post}$ is the average of the post-intervention actual outcome of the treatment state, $\bar{Y}_{TR,synthetic}^{post}$ is the average of the post-intervention outcome of the counterfactual. Similarly, $\bar{Y}_{TR,actual}^{pre}$ is the average of the pre-intervention actual outcome of treatment state, and $\bar{Y}_{TR,synthetic}^{pre}$ is the average of the pre-intervention outcome of the counterfactual. If the outcome changed in response to the intervention in time T_0 we would expect $\Delta_{TR} > 0$. Note that taking the absolute values in equation (4) makes sure that the estimate is neutral to the direction of change.

To formally test the significance of this estimate, we apply the permutations or randomization test – as suggested by Bertrand et al. (2004), Buchmueller et al. (2011), Abadie et al. (2010) and Bohn et al. (2014) – on this difference-in-difference estimator. Specifically, for each state in the donor pool, we estimate the difference-in-difference as specified in equation (4) as if these states were exposed to unconventional drilling in time T_0 (i.e., apply a fictitious intervention). The distribution of these "placebo" difference-in-difference estimates then provides the equivalent of a sampling distribution for Δ_{TR}. To be specific, if the cumulative density function of the complete set of Δ estimates is given by $F(\Delta)$, the p-value from a one-tailed test of the hypothesis that $\Delta_{TR} > 0$ is given by $F(\Delta_{TR})$ (Bohn et al. 2014). Note that this answers the question, how often would we obtain an effect of shale mining of a magnitude as large as that of the treatment state if we had chosen a state at random, which is the fundamental question of inference (Bertrandet al., 2004, Buchmueller et al. 2011, Abadie et al. 2010).

We carry out a second test following Abadie et al. (2010). We calculate what we call DID rank, which is the ranking of the absolute value of the magnitude of the difference-in-difference of the treatment state against all the placebo difference-in-difference magnitudes. For example, if DID rank is 1 then the estimated impact of the intervention in the treatment state is greater than any of the estimated placebo impacts.

3.2. Data and Selection of Predictors

All data for the outcome variables are collected beginning with 2001 to establish a period prior to unconventional oil and gas drilling in the areas of interest, and ending in 2011, the last period available at the time of the study for the outcome variables and energy production data used for identification. Annual data for the outcome variables, total employment, wage and salary employment, per capita income and population are from the U.S. Bureau of Economic

Analysis (U.S. BEA, 2013). Annual county poverty data are from the U.S. Census Bureau Small Area Income & Poverty Estimates (U.S. Bureau of the Census, 2013a).

In addition, because of widespread nondisclosure of industry level data at the county level by the U.S. Bureau of Economic Analysis, proprietary wage and salary employment data for a number of industries were purchased by the authors from Economic Modeling Specialists International (EMSI): Oil and Gas Extraction; Accommodation and Food Services; Retail Trade; Construction.[3] Employment in Oil and Gas extraction was used to identify the counties with oil and gas drilling and extraction. We use wage and salary employment for identification because total employment counts individuals receiving investment income as employees, making it less likely to be correlated with the location of energy extraction. Energy extraction activity used in identification is provided by county-level oil and gas production data for 2000-2011 from Economic Research Services of the U.S. Department of Agriculture (USDA, 2013a).

Several variables were used as predictors to construct the synthetic control for each grouping of oil and gas counties. First, variables were included to reflect industry composition of the counties. Whether the county was heavily dependent on farming, manufacturing, or mining was included, all from the ERS (USDA, 2013b). Industry dependence of the county is based on the shares of earnings during the period of 1998 to 2000, making industry dependence pre-determined to the intervention periods. We also use an industry mix employment growth measure over the period 2002 to 2007. Industry mix employment growth is the growth expected over the period based on an area's initial composition of fast- and slow-growing industries nationally and has a long history of use in regional science. The measure was calculated at the county level using four-digit NAICs data by Dorfman et al. (2011) and aggregated to

[3] EMSI data has been used elsewhere for county level analysis, having been shown to be close to actual estimates at the detailed sectoral level by state employment agencies (Dorfman et al., 2011; Fallah et al., 2014, Rickman and Guettabi, forthcoming).

nonmetropolitan portions of states by Rickman and Guettabi (forthcoming).[4] Because the industry mix measure uses beginning period county employment shares and sector employment growth at the national level, it is predetermined to the intervention periods at the sub-state level.

Because of their importance for growth and income levels (Partridge et al., 2008; Partridge et al., 2009), we also include measures of urbanization and urban proximity. The rural-urban continuum code from ERS is included, which is based on population and contiguity to a metropolitan area (USDA, 2013b). More refined measures of urban proximity are provided by measures of county distances to metropolitan areas in different tiers of the urban hierarchy. Included is the distance from a nonmetropolitan county to the nearest metropolitan areas, while also included are variables representing the incremental distances to reach metropolitan areas with population of at least 250 thousand, 500 thousand and 1.5 million (see Partridge et al., 2008 for details).

The natural amenity attractiveness of the county is included because of its importance for regional employment and population growth (Deller et al., 2001, Rickman and Guettabi, forthcoming). Amenity attractiveness is measured by a ranking from 1 to 7 reported by ERS (USDA, 2013b). The ranking reflects the following characteristics of the county and their relationships to population growth during 1970 to 1996 (McGranahan, 1999): (1) average January temperature; (2) average January days of sun; (3) a measure of temperate summers; (4) average July humidity; (5) topographic variation; and (6) water area as a proportion of total county area. We also include whether a county is a retirement destination, designated so by ERS if the number of residents over 60 years of age increased by more than 15 percent between 1990 and 2000 (USDA, 2013b). Retirement migration may reflect amenity attractiveness not

[4] We are grateful to the authors of Dorfman et al. (2011) for use of these data for the nonmetropolitan aggregates.

reflected in the ERS natural amenity measure, such as the presence of man-made amenities (e.g., health care facilities in the county).

Finally, variables are included to capture the influence of educational attainment of the population and natural population growth. Natural population growth over the period 2002-2007, measured as the excess of births over deaths (U.S. Bureau of the Census, 2013b), is included as it was found to be a primary determinant of population growth differences across nonmetropolitan counties during the period (Rickman and Guettabi, forthcoming). For education, we include the percent of the population 25 years and older with at least a bachelor's degree, only an associates college degree, and only a high school degree from 2000 Census of Housing and Population.

All population-based variables are weighted by beginning-period county population in aggregation, while beginning-period employment is used for employment-based variables not in levels. Only thirty of the lower forty eight states were included in the pool of donor states. Delaware, Rhode Island and New Jersey were omitted because of the absence or paucity of nonmetropolitan areas in the states. Other states were omitted because of proven shale gas reserves or other significant energy extraction in the state: Alabama, Colorado, Kentucky, Louisiana, Michigan, Montana, New Mexico, Ohio, Oklahoma, Texas, West Virginia and Wyoming.

4. Results and Discussion

Two sets of predictor variables are employed. First, we perform the analysis with all the predictor variables discussed above included, plus pre-intervention per capita income to capture convergence effects, and the pre-intervention outcome variable (Abadie et al., 2010). Second, for parsimony given the limitations on the number of donor states we reduce the number of predictors based on their importance and pre-intervention fits. Using the smaller set of predictor variables produces good pre-intervention fits (the absolute pre-intervention prediction error is typically below 1 percent of the pre-intervention mean of the outcome). The parsimonious set of predictor variables includes: percent of college graduates, proportion of farm dependent counties, proportion of manufacturing dependent counties, industry mix growth rate, natural amenity ranking, incremental distances to metropolitan areas of 500 thousand and 1.5 million people, per capita income in 2001, and the pre-intervention outcome variable. Therefore, the parsimonious runs serve as our base results.

Full SCM results for total employment in Arkansas, North Dakota and Pennsylvania appear in Tables 3, 5 and 6. Each table contains the weights that the donor state non-metropolitan portions contribute in the construction of the synthetic control for total employment in the oil and gas county aggregate. Each table also includes the statistical results of the permutations or randomization tests; i.e., the difference in the post- and pre-intervention mean gap between actual and synthetic outcomes of the treatment units are ranked and statistically compared to those from placebo runs for each of the donor non-metropolitan portions. Table 4 contains these statistical tests for all other variables, where for brevity the weights are not shown. Figures 1-3 show the pictures of the impact for total employment, per capita income, population and the poverty rate; each picture shows the fit prior to the intervention and a comparison post-intervention of the actual values and the synthetic control. The figures show that the employment and population variables were measured as ratios to their corresponding 2001 values to avoid issues associated with differences in sizes of county economies. Thus, the own-predictor variable for employment and population are the last pre-intervention values. Per capita income and poverty are measured in the levels for each year, where the own-predictor variables include values for both 2001 and the last pre-intervention year in each case.

4.1. Arkansas

As shown in Table 2, relative to all nonmetropolitan counties in the state, oil and gas counties in Arkansas had a slightly larger share of the adult population with a college degree, were slightly further from the nearest metropolitan area and further down the urban hierarchy, were more likely to be a retirement destination, less likely to be dependent on farming or manufacturing, had slightly lower poverty, and comparable per capita personal income.

We identified 2006 as the year of intervention because oil and gas wage and salary employment dramatically rose afterwards. As shown in the middle column of Panel B, the primary contributors to the synthetic control for total employment in Arkansas oil and gas counties were Mississippi, Virginia and Nebraska, with Mississippi receiving approximately half of the weight. Mississippi is the primary synthetic control contributor to all regional labor market measures for the Arkansas oil and gas counties, with the weights ranging from 0.43 to 0.52 (not shown).

The middle column of Panel B of Table 3 shows that the post-intervention total employment level for Arkansas oil and gas counties falls below the synthetic control estimates, though statistically insignificant based on the placebo analysis. The top half of Figure 1 shows the pre- and post-intervention comparisons of actual to synthetic control total employment estimates. Also shown are the pre- and post-intervention comparisons for personal income, the poverty rate and population. Post-intervention actual per capita income exceeds the synthetic control estimate, while the actual poverty rate lies below. However, Table 4 shows that the differences are not statistically significant. Therefore, shale oil and gas development does not appear to have significantly affected the aggregate economy of Arkansas's oil and gas counties.

To examine whether there were positive spillovers extending statewide beyond the borders of the oil and gas counties we re-ran the synthetic control analysis for the entire nonmetropolitan portion of Arkansas as the treated unit. The first column of Panel A in Table 3 shows that Mississippi became even more the dominant contributor to the synthetic control estimates, with over ninety percent of the weight, with Nebraska the only other state to significantly contribute. As for the Arkansas oil and gas counties, Tables 3 (Panel B) and 4 show that there were not any statistically significant regional labor market effects for nonmetropolitan Arkansas. This suggests that the lack of effect found for the oil and gas counties was not because of neglecting broader positive geographic spillovers.

Because the lack of effect could be attributable to difficulties in detecting the oil and gas activity within the broader aggregate of counties, where oil and gas activity may be too small of an economic component in many counties, we next more narrowly focus on the counties containing the most energy extractive activities. Most shale drilling has occurred in the counties north of Little Rock, running eastward from Conway to White Counties.[5] Therefore, we examine the aggregate of the top four natural gas producing counties in 2011, which were responsible for over eighty percent of natural gas production in Arkansas during the year: Cleburne, Conway, Van Buren and White (USDA, 2013a).[6]

Relative to all Arkansas oil and gas counties, these counties are less remote given their proximity to Little Rock, have lower adult population shares of college graduates, more likely to be retirement destinations, less likely to be manufacturing or farm dependent counties, and

---

[5] http://lingo.cast.uark.edu/LINGOPUBLIC/about/, last accessed July 1, 2014.
[6] The four counties are among the seven Arkansas counties indicated by Center for Business and Economic Research (2008) as engaged in shale gas extraction.

slightly less amenity attractive. The difference in characteristics led to different contributors to the synthetic control estimates. For total employment, the third column of Panel A in Table 3, shows that the non-metropolitan portion of New York, which has almost no nonmetropolitan oil and gas wage and salary employment, served as the primary contributor to the synthetic control for the four county aggregate, followed by Arizona and Florida with smaller weights.

The bottom half of Figure 1 shows that post-intervention actual total employment, per capita income and population are higher than the synthetic control estimates, while the actual poverty rate is lower. Tables 3 (Panel B) and 4 show that only total wage and salary employment, along with employment in Construction and Accommodation and Food Services, are statistically significant at or below the 0.10 level, though per capita income and population were close with the fifth largest DID ranking. In 2011, total wage and salary employment is predicted to be 8.2 percent higher than it otherwise would be without oil and gas development (5.3 percent for total employment but insignificant), where the comparable figures for Construction and Accommodation and Food Services are 11.3 and 15 percent, respectively. Although insignificant, the differences are 7.8 and 3.6 percent higher for per capita income and population.

Because of the significance of the wage and salary employment effect, and because the EMSI employment data for the oil and gas sector are wage and salary, we calculate an estimated wage and salary employment multiplier for energy extraction activity in these four counties. The estimated SCM difference in total wage and salary jobs in 2011, calculated as the difference between the actual and synthetic control estimate is 3,615 (not shown). Using EMSI data for the four counties, the change in the total number of wage and salary jobs in the mining sector attributable to increased oil and gas activity is 2,045.6, which produces a wage and salary multiplier of 1.77.[7] This is considerably lower than the multiplier of 2.5-2.64 produced by Center for Business and Economic Research (2008) for the Fayetteville Shale Play but close to the multiplier of 1.7 estimated by Brown (2014) for several states that had counties involved in natural gas production.

To assess whether the difference in results for the top-four oil and gas producing counties versus the larger areas in Arkansas is because of differences in donor states in the SCM, where Mississippi has one nonmetropolitan county that experienced increased oil and gas employment during the treatment period, we re-ran the SCM for the larger areas. However, with one exception, there were not any changes in results. Only for the poverty rate in the aggregate of oil and gas counties was there estimated to be a statistically significant economic benefit from shale gas activity. With Mississippi removed from the donor pool, the states contributing the most to the total employment synthetic control estimate in the oil and gas aggregate, in order, were North Carolina, Nebraska, and New York (they were correspondingly Nebraska and Tennessee for the nonmetropolitan area aggregate); these are states with little or no oil and gas employment during the treatment period.

4.2. North Dakota

Compared to the rest of nonmetropolitan North Dakota, its oil and gas counties have a higher share of college graduates, are more remote, had negative natural population growth

[7] The EMSI data estimates we purchased indicate a change of 340.4 oil and gas extraction wage and salary jobs from 2006-2011 in the four counties. Using an estimated change in support jobs in the mining sector of 1,705.2 over the period, the total change in mining sector wage and salary jobs associated with increased energy activity in the four counties equals 2,045.6. Total employment in mining support for the four counties was obtained from Dorfman (2011) and was scaled to reflect the statewide ratio of wage and salary employment to total employment in the sector. The change in mining support jobs is almost exclusively believed to support oil and gas extraction because prior to 2006 mining support jobs was virtually nonexistent, despite non-oil and gas mining activity in the counties (in Nonmetallic Mineral Mining and Quarrying). For a list of support activities for oil and gas operations see http://www.naics.com/censusfiles/ND213112.HTM (accessed July 26, 2014).

(likely reflecting an older population), are more natural amenity-attractive, and are less farm dependent. We identified 2007 as the intervention year as oil and gas wage and salary employment significantly rose in 2008, after having been fairly steady in previous years.

As shown in the second column of Panel A in Table 5, the primary contributor to the synthetic control for North Dakota oil and gas counties is South Dakota, with a weight of approximately 0.65. South Dakota also is the primary contributor for aggregate wage and salary employment, and wage and salary employment in Accommodation and Food Services, while significantly contributing to the synthetic estimates for Construction and Retail employment (not shown). Oil and gas wage and salary employment in South Dakota was miniscule and remained virtually unchanged from 2006-2007 to 2011. In order of contribution, other contributors to the synthetic control for total employment include Maine, Massachusetts, Florida, Washington, and Minnesota.

The second column of Panel B in Table 5 shows that the post-intervention actual employment level rises above the synthetic control estimate. From the bottom half of Figure 2, it can be seen that in the absence of oil and gas activity, total employment in the North Dakota oil and gas counties would have been expected to fall slightly during the Great Recession and rise back to the pre-recession level by 2011, whereas, it actually increased dramatically throughout the post-intervention period. The difference is statistically significant below the 0.001 level, with the DID ranking 1.

Also shown in the bottom half of Figure 2, post-intervention, per capita income and population rise above the synthetic control estimates, while the poverty rate falls below. To be sure, from Table 4, all labor market measures for the North Dakota oil and gas counties statistically differ from the synthetic control estimates with a high level of confidence.[8] Only for population was the DID rank not first, placing second.

From calculations not shown, total employment is estimated to be 32,744 greater in 2011 because of oil and gas activity, representing a 32 percent increase. Correspondingly, wage and salary employment is estimated to be 31,752 greater, a 42 percent increase. Construction wage and salary employment more than doubled, while Retail employment increased 13 percent and Accommodation and Food Services employment increased 30 percent. By comparison, population is estimated to have only increased nearly 10 percent. Also notably, per capita income is estimated to be nearly 40 percent higher, while the poverty rate dropped over 5 percentage points from what it otherwise would have been.

Because of coal mining in western North Dakota, including in the oil and gas counties, not all of the increase in wage and salary employment in the region can be attributed to increased oil and gas extraction. Excluding Nonmetallic Mineral Mining and Quarrying (and Support Activities for Mining), approximately 91.25 percent of total employment jobs in Mining in North Dakota's oil and gas counties were related to oil and gas extraction (Dorfman, 2011), suggesting they were responsible for 28,973 of the SCM estimate of 31,752 wage and salary jobs.[9] Using EMSI wage and salary data for North Dakota oil and gas counties, the total change in the number of wage and salary jobs in Mining sector attributable to increased oil and gas activity is 8,592, which produces an estimated wage and salary multiplier of 3.37.[10] This is considerably higher

[8] Although not shown, the differences were all positive, except for the expected negative difference for the poverty rate.

[9] The scaling of the SCM wage and salary estimate by 0.9125 assumes equal multiplier effects between coal mining and oil and gas extraction.

[10] The change in oil and gas extraction jobs from 2007-2011 is estimated by EMSI to be 763.7. Approximately 91.2 percent of the increase in mining support jobs is attributed to oil and gas extraction, the other 8.8 percent is attributable to increased coal production. The estimated increase in total employment in mining support (Dorfman, 2011) is scaled by the U.S. Bureau of Economic Analysis ratios of mining support wage and salary employment to

than the multiplier estimated above for Arkansas, but less than the national average of five reported by IHS (2012). But the oil and gas counties in North Dakota are much more isolated than those in Arkansas and had considerably fewer other types of economic activity prior to the boom, reducing the likelihood of crowding out. North Dakota primarily extracts oil, while Arkansas primarily extracts natural gas.

To examine whether there were broader positive or negative spillover effects we next examine all North Dakota nonmetropolitan counties as a treated aggregate unit. Compared to the oil and gas counties in the state, all nonmetropolitan counties are more farm dependent, are less amenity attractive, and are less remote (Table 2). The composition of the synthetic control group for the nonmetropolitan aggregate also differs. Possibly reflecting the much greater farm dependence of all of nonmetropolitan North Dakota, the primary contributor is Iowa, followed by Nebraska, and then South Dakota.

The first column of Panel B in Table 5 shows that total employment significantly increased in the entire nonmetropolitan portion of North Dakota. The top half of Figure 2 shows similar patterns as those for the oil and gas counties, except that population did not increase. Total employment increased 17 percent, while wage and salary employment increased nearly 19 percent. This translates into additional total employment of 38,475, and additional wage and salary employment of 29,884 (not shown). The larger total employment estimate suggests additional spillovers to other counties outside of the oil and gas boundaries, while no net additional wage and salary employment is estimated to have occurred. Total employment includes those receiving payments but not actively working in the industry, so there likely are many of those individuals living in non-oil and gas nonmetropolitan North Dakota counties. For

---

comparable total employment for both 2007 and 2011, producing an estimated increase of oil and gas related mining support jobs of 7,823.3.

wage and salary employment, any positive spending spillovers may be offset by drawing labor from other counties, which is suggested by the absence of a population effect for the nonmetropolitan counties as a whole, or crowding out effects through increased factor costs. Among all the labor market measures, only Accommodation and Food Services, and population are not statistically significant at or below the 0.10 level.

4.3. Pennsylvania

Pennsylvania oil and gas counties have lower than the U.S. nonmetropolitan average share of population with a college degree but a greater high school completion rate, much greater manufacturing dependence, and are less remote from urban agglomeration economies (Table 2). Relative to all of nonmetropolitan Pennsylvania, the oil and gas counties are less likely to be a retirement destination, have slightly larger natural population growth and are slightly less likely to be manufacturing dependent.

The intervention year is identified as 2006 because yearly oil and gas wage and salary employment increases become much more significant afterwards. For total employment, the primary donors to the synthetic control in order of importance are Illinois, Oregon, Virginia, and New York (Table 6, middle column of Panel A). The same ordering of contributors is found for aggregate wage and salary employment, where Illinois is the top contributor for all variables, except for poverty where Virginia is the top contributor; all total employment contributors are found to be major contributors to the SCM estimates of the other variables.

Although actual total employment in 2011 lies 1.6 percent above the synthetic control estimate (top half of Figure 3), the differences over the post-intervention period are not statistically significant. Of all the regional labor market measures, only Retail wage and salary employment is significant at or below the 0.10 level (Table 4), though the actual value lies below the synthetic control estimate (not shown). This may reflect the loss of tourism dollars in these counties from oil and gas activity (White, 2012; Lydersen, 2013). Employment in Construction and Accommodation and Food Services appeared positively affected, but the differences are not statistically significant. Not surprisingly then, overall wage and salary employment was statistically unaffected.

The results also suggest that the absence of significant regional labor market effects was not from ignoring broader geographic spillovers. For the total nonmetropolitan area of Pennsylvania, total employment (Table 6, Panel B, first column) and all other regional variables (Table 4) were not statistically affected by oil and gas activity. The composition of the synthetic control for total employment changed somewhat, with New York as the largest contributor, followed by Virginia, Tennessee and Oregon with weights greater than ten percent in order of importance.

As we did for Arkansas, we also examine the most shale energy active counties in the nonmetropolitan portion of Pennsylvania as a group. The five counties of Bradford, Clinton, Potter, Susquehanna and Tioga in the northeast part of the state comprise forty nine percent of natural gas production in 2011, where production in each county increased dramatically in the immediately preceding years (USDA, 2013a). The counties also mostly did not have any oil production during the period.

From Tables 4 and 6, we see there are not any statistically significant effects at the 0.10 level or below. Actual total employment and wage and salary employment in 2011 are higher but the employment effect only becomes noticeably positive after 2009. Per capita income is 2.7 percent higher and population 1.2 percent lower in 2011, though both are statistically insignificant. The pattern may occur because oil and gas extraction wage salary employment only exceeds one hundred in 2010, doubling to over two hundred in 2011; this corresponds to the boom in production that began in 2009. Possibly, if the boom in production and oil and gas employment continues in future years, significant effects could be found.

We also ran the SCM for ten counties (indicated in Table 1) in western Pennsylvania (not reported). The only statistically significant result is for a reduction in Retail wage and salary employment. The rise in oil and gas wage and salary employment may not have been sufficiently dramatic to generate detectable effects. Oil and gas employment was substantial prior to the middle of the decade, with more gradual growth throughout the second half of the decade, in contrast to the dramatic rise in the five northeastern counties that began in 2009.

To examine the robustness of our results, for all Pennsylvania county samples we first changed the intervention year from 2006 to 2005 and secondly, we exclude Illinois and Virginia from the donor pool. Changing the intervention year to 2005 could capture the effects of the modest increases in oil and gas employment during the year and capture anticipatory economic activity. Yet, moving the intervention year to 2005 produced virtually identical results, with the same major donor states in each SCM and an absence of statistically significant effects, except for the negative effect on Retail employment previously found. We removed Illinois and Virginia because of increased oil and gas activity in one or two counties in the state, potentially confounding the synthetic control baseline estimates. Again, the results are robust, where none of the changes in the outcome measures are close to statistical significance; the negative Retail employment result also becomes insignificant. Notably, Tennessee gained in importance as a donor state in the SCMs for total employment, and Maine emerged as a major contributor; across all estimates, both Maine and Tennessee become more prominent contributors, in which both states had little or no oil and gas wage and salary employment during the decade.

4.4. Results for the Full Set of Predictor Variables

We next re-run the SCM for all regional labor market measures and all areas examined above using the complete set of predictor variables. The results are not shown for brevity but are available from the authors upon request. The additional variables are: the percent of the population with only a high school degree; the percent of the adult population with only an associate college degree; ERS rural-urban influence; distance to the nearest metropolitan area; incremental distance to a metropolitan area with over 250 thousand people; Wharton Residential Land Use Regulatory Index, and retirement destination status of counties in the area.

As before, no significant regional labor market effects were found for the Arkansas oil and gas county aggregate or for the Arkansas nonmetropolitan aggregate. For the four county aggregate, not only are the wage and salary employment variables statistically significant as before, but so is the poverty rate, while total employment remains insignificant. For North Dakota, all regional labor market measures continue to be statistically significant for the non-metro and the oil and gas county aggregates. Only Retail wage and salary employment and population are insignificant for the nonmetropolitan county aggregate for North Dakota, where before it was population and Accommodation and Food Services employment that were insignificant. For Pennsylvania, again only Retail wage and salary employment is statistically significant in the oil and gas county aggregate SCM. In contrast to the previous SCM's for the five-county Pennsylvania aggregate, the poverty effect is statistically significant (p-value=0.1), while the other variables remain insignificant.

5. Summary and Conclusions

In this study, we examined the regional economic effects of unconventional oil and gas production in the Bakken, Fayetteville and Marcellus shale plays. Because of potential spillovers across local area economies we examined the effects for aggregates of counties, including the entire metropolitan portions of Arkansas, North Dakota and Pennsylvania, aggregates of the oil and gas producing counties in the states, and subsets of the states' oil and gas counties. To broadly gauge the effects on economic well-being (Partridge and Rickman, 2003), we examined a wide range of local economic indicators: total employment, wage and salary employment, per capita income, the poverty rate, and population. Wage and salary employment in Accommodation and Food Services, Construction and Retail also were examined to identify possible positive or negative spillover effects of unconventional energy production on other local industries.

We use the synthetic control method (SCM) (Abadie and Gardeazabal 2003; Abadie et al., 2010) to predict economic activity that would occur in the absence of increased unconventional energy development, which can then be compared to actual outcomes, where the differences are the estimated effects of increased shale oil and gas production. We find large and statistically significant positive effects for the oil and gas counties in North Dakota across the entire wide range of regional labor market measures, as well as for the entire nonmetropolitan portion of the state (except for population). The only statistically significant positive effects found for Arkansas are for the four counties most intensive in shale gas production, suggesting mostly localized positive economic benefits of unconventional oil and gas production. No statistically significant positive effects were found for any aggregation of counties in Pennsylvania.

In addition, back of the envelope calculations with the results where, significantly positive wage and salary employment effects were found, suggest that actual multiplier effects on local economies in shale plays are smaller than commonly reported by the use of input-output models. The absence of reported effects for Pennsylvania could at least partly be because of the more recent, or slow pace of, development in key nonmetropolitan portions of the state compared to Arkansas and North Dakota. But the lack of effects for Pennsylvania contrasts with the large predicted effects by Considine et al. (2009; 2010) that should show up in our analysis and are more in line with the findings of Weinstein and Partridge (2011).[11]

The less sanguine findings for the regional economic benefits of unconventional oil and gas development caution that it may not be the panacea for what ails many local area economies. Areas that contain significant levels of other types of economic activity such as agriculture, retiree migration, tourism, may more likely experience offsetting adverse economic effects. The isolation of counties in the Bakken area mitigates some of these but energy activity in and near more populated areas may come with greater economic costs. To be sure, the analysis in this paper is short run, more likely emphasizing the positive effects, such as those related to exploration and initial construction (White, 2012), but may not cover a time span of sufficient length to capture most of the long-term adverse effects that may arise from, for instance, contamination of groundwater or increased frequency of earthquakes in areas with fault lines. State and local area residents and policy makers must weigh the balance of both the potential benefits and costs in the permitting and taxation of unconventional energy extraction to ensure it enhances regional overall economic well-being.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* no. 105:493-505.

[11] Because of our focus on nonmetropolitan areas, we did not examine Washington County, as did Weinstein and Partridge (2011), because it is part of the Pittsburgh metropolitan area.

Abadie, Alberto, and Javier Gardeazabal, 2003. "The Economic Costs of Conflict: A Case-Control Study for the Basque Country." *American Economic Review* no. 93 (1):113-132.

Atkin, Emily, 2014. "2,500 Oklahoma Earthquakes Linked to Fracking, Scientists Say," *Climateprogress*, July 7. Accessed on July 20, 2014 at http://thinkprogress.org/climate/2014/07/07/3456931/oklahoma-frackquakes/

Baker Hughes, Inc., 2014. Accessed at http://phx.corporate-ir.net/phoenix.zhtml?c=79687&p=irol-reportsother, June 13, 2014.

Bertrand, M., E Duflo and S Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics* 119 (1), 249-275.

Black**,** Dan, McKinnish, Terra, Sanders, Seth, 2005. "The Economic Impact of the Coal Boom and Bust," *Economic Journal* 115(April), 449-76.

Bohn, Sarah, Magnus Lofstrom and Steven Raphael, 2014. "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?" *Review of Economics and Statistics*, Vol. 96, No. 2, Pages 258-269.

Brown, Jason P., 2014. "Production of Natural Gas from Shale in Local Economies: A Resource Curse or Blessing? *Federal Reserve Bank of Kansas City Economic Review* 99(1), 5-33.

Buchmueller, Thomas C., John DiNardo, and Robert G. Valletta. 2011. "The Effect of an Employer Health Insurance Mandate on Health Insurance Coverage and the Demand for Labor: Evidence from Hawaii." *American Economic Journal: Economic Policy*, 3(4): 25-51.

Center for Business and Economic Research, 2008. "Projecting the Economic Impact of the Fayetteville Shale Play for 2008-2012," Sam M. Walton School of Business, University of Arkansas.

Considine, T.J., R. Watson, J. Sparks, R. Entler, 2009. "An Emerging Giant: Prospects and Economic Impacts of Developing the Marcellus Natural Gas Play," The Pennsylvania State University, Department of Energy and Mineral Engineering.

Considine, T.J., R. Watson and S. Blumsack, 2010. "The Economic Impacts of the Pennsylvania Marcellus Shale Natural Gas Play: An Update," The Pennsylvania State University, Department of Energy and Mineral Engineering.

Deller, Steven, Tsung-Hsiu Tsai, David Marcouiller and Donald English. 2001. "The Role of Amenities and Quality of Life in Rural Economic Growth," *American Journal of Agricultural Economics*, 83(2), 352-365.

Donald, Stephen G. and Kevin Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data," *The Review of Economics and Statistics*, May 2007, Vol. 89, No. 2, Pages 221-233.

Dorfman, Jeffrey, Mark D. Partridge, and Hamilton Galloway, 2011. "Are High-tech Employment and Natural Amenities Linked: Answers from a Smoothed Bayesian Spatial Model." *Spatial Economic Analysis* 6: 397-422.

Douglas**,** Stratford M.; Walker, Anne W., 2012. "Sample Selection in Appalachian Research," *Review of Regional Studies* 42(2), 143-59.

Edmiston, Kelly D., 2004. "The Net Effects of Large Plant Locations and Expansions on County Employment," *Journal of Regional Science* 44(2), 289-320.

Fallah, Belal, Mark Partridge, Dan Rickman, 2014, "Geography and High-Tech Employment Growth in U.S. Counties", *Journal of Economic Geography*, 14 (4), 683-720.

Gopalakrishnan, S and Klaiber, HA, 2014. "Is the Shale Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania," *American Journal of Agricultural Economics* 96 (1), 43-66.

Harrigan**,** Frank; McGregor**,** Peter G., 1989. "Neoclassical and Keynesian Perspectives on the Regional Macro-economy: A Computable General Equilibrium Approach," *Journal of Regional Science*, 29(4), 555-73.

IHS, 2012. *America's New Energy Future: The Unconventional Oil and Gas Revolution and the US Economy*. Volume 2, State Economic Contributions, December.

James, Alex, Aadland, David, 2011. "The Curse of Natural Resources: An Empirical Investigation of U.S. Counties," *Resource and Energy Economics* 33(2), 440-53.

Kelsey, Timothy J., Martin Shields, James Ladlee, Melissa Ward, 2009. "Economic Impacts of Marcellus Shale in Pennsylvania: Employment and Income in 2009," Penn State Extension and Penn College.

Kinnaman, Thomas C., 2011. "The Economic Impact of Shale Gas Extraction: A Review of Existing Studies," *Ecological Economics* 70, 1243-1249.

Lipscomb, Clifford A., Wang, Yongsheng, Kilpatrick, Sarah J, 2012. "Unconventional Shale Gas Development and Real Estate Valuation Issues," *Review of Regional Studies* 42(2), 161-75.

Lydersen, Kari, 2013. "U.S. Chamber's Fracking Boom: Behind the Numbers," *Midwest Energy News* (January 10), last accessed at http://www.midwestenergynews.com/2013/01/10/u-s-chambers-fracking-job-boom-behind-the-numbers/ June 14, 2014.

McGranahan, David A. 1999. "Natural Amenities Drive Rural Population Change," AER 781. Washington D.C.: Economic Research Service, U.S. Department of Agriculture.

Michaels**,** Guy, 2011. "The Long-Term Consequences of Resource-Based Specialization," *Economic Journal* 121(March), 31-57.

Muehlenbachs, Lucija; Spiller, Elisheba; Timmins, Christopher; 2014. "The Housing Market Impacts of Shale Gas Development," National Bureau of Economic Research, Inc., NBER Working Papers: 19796.

Murray, Sherry and Teri Ooms, 2008. "The Impact of the Marcellus Shale in Northeastern Pennsylvania," Joint Urban Studies Center. Last accessed On June 15, 2014 at http://www.institutepa.org/PDF/Marcellus/mtwhitepaper.pdf.

Papyrakis, Elissaios, Gerlagh, Reyer, 2007. "Resource Abundance and Economic Growth in the United States," *European Economic Review* 51(4), 1011-39.

Partridge, Mark D. and Dan S. Rickman, 2003. "Do We Know Economic Development When We See It**?"** *Review of Regional Studies* 33(1), 17-39.

Partridge, Mark D., Rickman, Dan S., Ali, Kamar, Olfert, M. Rose, 2008. "Lost in Space: Population Growth in the American Hinterlands and Small Cities," *Journal of Economic Geography* 8(6), 727-57.

_____, 2009. "Agglomeration Spillovers and Wage and Housing Cost Gradients across the Urban Hierarchy," *Journal of International Economics* 78(1), 126-40.

Rahm, Dianne, 2011. "Regulating Hydraulic Fracturing in Shale Gas Plays: The Case of Texas," *Energy Policy* 39(5), 2974-81.

Rickman, Dan S., 1992, "Estimating the Impacts of Regional Business Assistance Programs: Alternative Closures in a Computable General Equilibrium Model", *Papers in Regional Science*, 71(4), 421-435.

Rickman, Dan S. and Mouhcine Guettabi, forthcoming. "The Great Recession and Nonmetropolitan America," *Journal of Regional Science*.

Rubin, D. B. 2001. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation," *Health Services and Outcomes Research Methodology*, 1, 169–188.

Throupe, Ron, Simons, Robert A., Mao, Xue, 2013. "A Review of Hydro Fracking and its Potential Effects on Real Estate," *Journal of Real Estate Literature* 21(2), 205-32.

U.S. Bureau of the Census, 2013a. Data accessed December 28, 2013 at http://www.census.gov/did/www/saipe/.

U.S. Bureau of the Census, 2013b. Data last accessed February 21, 2013 at http://www.census.gov/popest/.

U.S. Bureau of Economic Analysis, 2013. Data accessed July 30, 2013 at http://www.bea.gov/regional/index.htm.

U.S. Department of Agriculture, Economic Research Service, 2013a. Data accessed April 27, 2014 at http://www.ers.usda.gov/data-products/county-level-oil-and-gas-production-in-the-us.aspx#.U5xhbLHb7ZU.

U.S. Department of Agriculture, Economic Research Service, 2013b. Data last accessed June 14, 2014 at http://www.ers.usda.gov/data-products.aspx#.U5xkxbHb7ZV.

U.S. Energy Information Administration, 2013. Annual Energy Outlook 2013 with Projections to 2040, April.

Weber, Jeremy G., 2012. "The Effects of a Natural Gas Boom on Employment and Income in Colorado, Texas, and Wyoming," *Energy Economics* 34(5), 1580-88.

Weinstein, Amanda L. and Mark D. Partridge, 2011. "The Economic Value of Shale Natural Gas in Ohio," The Swank Program in Rural-Urban Policy Summary and Report.

White, Nancy E., 2012. "A Tale of Two Shale Plays," *The Review of Regional Studies* 42(2), 107-119.

Table 1: Shale Oil and Gas Counties Examined

Arkansas	North Dakota	Pennsylvania
Bradley	Billings	Bedford
Chicot	Bottineau	Bradford[b]
Cleburne[a]	Bowman	Clarion[c]
Columbia	Burke	Clearfield[c]
Conway[a]	Divide	Clinton[b]
Lafayette	Dunn	Crawford[c]
Logan	Golden Valley	Elk[c]
Nevada	McKenzie	Forest[c]
Ouachita	McLean	Greene
Pope	Mercer	Indiana[c]
Union	Mountrail	Jefferson[c]
Van Buren[a]	Renville	Lawrence
White[a]	Slope	McKean[c]
Woodruff	Stark	Northumberland
	Steele	Potter[b]
	Williams	Somerset
		Susquehanna[b]
		Tioga[b]
		Venango[c]
		Warren[c]

[a]denotes part of Arkansas sub-group of counties
[b]denotes part of northeastern Pennsylvania sub-group of five counties
[c] denotes part of western Pennsylvania sub-group of ten counties

Table 2: Summary Statistics

	Donor pool, N=319		Non-metro (Mean), N=11			Oil and gas counties (Means), N=11				
	Mean	SD	AR	ND	PA	AR	AR top 4	ND	PA	PA top 5
Outcomes										
Total employment	507000	302000	554000	234000	963000	176000	64725	112000	518000	95264
PC personal income, 2005 US GDP	29774	6522	24757	33396	27183	24332	25303	34821	26716	25751
Population	988000	616000	1150000	346000	2000000	361000	136000	157000	1090000	204000
Poverty rate (%)	14.19	4.37	20.01	12.33	12.37	19.03	16.71	10.79	13.74	13.37
Total wage and salary employment	382000	234000	421000	168000	746000	135000	45741	84706	402000	68987
Number of jobs in accommodation	31645	16826	27566	11856	57168	8974	3566	6646	28191	5200
Number of jobs in construction	25562	15297	25634	10170	44690	10003	4159	5637	22746	4487
Number of jobs in retail	48330	27928	49404	17719	95870	16397	6043	9712	51208	8435
Predictors										
Percent with at least Bachelor's degree	11.49	3.71	8.19	13.43	8.77	9.64	9.12	14.73	8.39	8.44
Percent high school graduate	34.23	4.25	36.50	30.45	46.96	36.02	0.37	30.03	48.56	46.82
Percent with associate degree	6.48	1.36	3.53	8.96	5.21	3.59	0.04	9.27	5.34	6.31
Status as a retirement county: 1=yes	23.77	25.00	20.20	0.00	10.09	25.43	71.40	0.00	0.45	0.00
Percent manufacturing counties	31.40	27.72	54.23	0.00	59.53	44.37	18.82	0.00	55.00	70.19
Percent farm counties	6.54	10.22	12.56	46.30	0.00	11.53	0.00	19.16	0.00	0.00
USDA natural amenities scale	3.81	1.10	3.49	2.54	3.49	3.37	2.95	2.96	3.51	3.62
Natural population growth 2002-2007	1.39	1.28	0.65	0.60	-0.13	0.63	0.76	-1.33	0.31	-1.38
Industry mix emp. growth 2000-2007	6.76	3.01	4.58	5.14	5.95	6.53	8.09	6.78	6.92	5.18
Rural-urban index from ERS	5.71	0.74	6.36	7.33	4.86	6.98	5.60	8.32	5.12	5.97
Wharton regulatory index	-0.18	0.74	-0.98	-0.53	-0.16	-0.98	-0.98	-0.53	-0.16	-0.16
Incremental distance to metro over 1.5m	86.44	99.42	242.30	0.00	55.53	283.23	154.07	0.00	27.22	32.14
Incremental distance to metro over 0.5m	45.45	43.64	20.72	24.64	6.97	21.25	0.00	22.98	46.41	82.48
Incremental distance to metro over 0.25m	42.48	41.65	33.41	445.47	21.92	18.04	0.00	554.99	8.22	14.61
Distance to nearest metro area	80.44	29.26	76.14	145.70	49.53	81.31	35.90	175.98	55.40	52.43

Notes: (a) Donor pool consists of the non-metro counties of 29 states. (b) The time period is 2001-2011. (c) Estimates for employment (jobs) and population are totals (not means) for each aggregate area. (d) 'AR top 4' refers to the top 4 oil and gas counties in Arkansas, 'PA top 5' refers to top 5 oil and gas counties in Pennsylvania.

Table 3: SCM Estimation of the Impact of Shale Gas and Tight Oil Boom on Total Employment in Arkansas Counties

Panel A: W-weights

State	Non-metro counties	All oil-gas counties	Top 4 oil-gas counties
Arizona	0.000000	0.000000	0.095085
California	0.000000	0.000000	0.000000
Connecticut	0.000000	0.000000	0.000000
Florida	0.000092	0.000104	0.024444
Georgia	0.000000	0.000000	0.000000
Idaho	0.000000	0.000000	0.000000
Illinois	0.000000	0.000000	0.000002
Indiana	0.000000	0.000000	0.000000
Iowa	0.000000	0.000000	0.000000
Maine	0.000000	0.000000	0.000000
Maryland	0.000000	0.000000	0.000000
Massachusetts	0.000000	0.000000	0.000000
Minnesota	0.000000	0.000000	0.000000
Mississippi	0.907790	0.502556	0.000000
Missouri	0.000000	0.000000	0.000000
Nebraska	0.092116	0.164440	0.000000
Nevada	0.000000	0.000000	0.000000
New Hampshire	0.000000	0.000000	0.000000
New York	0.000000	0.000000	0.880469
North Carolina	0.000000	0.001684	0.000000
Oregon	0.000002	0.000000	0.000000
South Carolina	0.000000	0.000000	0.000000
South Dakota	0.000000	0.000000	0.000000
Tennessee	0.000000	0.000000	0.000000
Utah	0.000000	0.000000	0.000000
Vermont	0.000000	0.000000	0.000000
Virginia	0.000000	0.331216	0.000000
Washington	0.000000	0.000000	0.000000
Wisconsin	0.000000	0.000000	0.000000

Panel B: Estimation Statistics

	Non-metro counties	All oil-gas counties	Top 4 oil-gas counties
SCM: Pre-intervention Fit			
Absolute prediction error to mean ratio	0.0113	0.0039	0.0014
SCM Inference: Permutations Test			
Pre-intervention difference (D1)	-0.0112	-0.0022	-0.0006
Post-intervention difference (D2)	-0.0250	-0.0084	0.0327
DID = |D2|-|D1|	0.0138	0.0062	0.0321
P-value: DID	0.4000	0.6667	0.2000
DID rank	13	21	7

Notes:

(a) List of Predictors

Percent college graduates

Proportion of manufacturing counties

Proportion of farm counties

Industry mix employment growth rate

Natural amenities scale

Incremental distance to 1.5m metro

Incremental distance to 0.5m metro

Natural population growth

Per capital personal income

Total employment of 2005

(b) Intervention is in 2006.

Table 4: SCM Estimation of the Impact of Shale Gas and Tight Oil Boom

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Arkansas: Non-metro Counties							
Pre-intervention fit	0.0454	0.0089	0.0097	0.0057	0.0071	0.0199	0.0022
DID = |D2|-|D1|	-0.0268	0.0016	0.0548	-0.0043	0.9072	0.5927	0.0087
P-value: DID	0.9667	0.9333	0.1667	0.9333	0.4333	0.3667	0.5333
DID rank	30	29	6	29	14	12	17
Arkansas: Oil-gas Counties							
Pre-intervention fit	0.0474	0.0029	0.0228	0.0019	0.0048	0.0217	0.0013
DID = |D2|-|D1|	0.0012	0.0028	-0.0021	0.0111	0.5925	0.4998	0.0009
P-value: DID	0.9000	0.7333	0.9667	0.4667	0.5333	0.3667	0.9000
DID rank	28	23	30	15	17	12	28
Arkansas: Top 4 Oil-gas Counties							
Pre-intervention fit	0.0170	0.0022	0.0164	0.0020	0.0044	0.0330	0.0014
DID = |D2|-|D1|	0.1738	0.0058	0.1279	0.0441	1.7985	0.2613	0.0253
P-value: DID	0.1000	0.7000	0.0000	0.0667	0.1333	0.5667	0.1333
DID rank	4	22	1	3	5	18	5
North Dakota: Non-metro							
Pre-intervention fit	0.0137	0.0017	0.0021	0.0036	0.0185	0.0244	0.0092
DID = |D2|-|D1|	0.5464	0.0500	0.0734	0.0816	6.7491	2.6214	-0.0081
P-value: DID	0.0000	0.1000	0.1667	0.0000	0.0000	0.0333	0.9667
DID rank	1	4	6	1	1	2	30
North Dakota: Oil-gas Counties							
Pre-intervention fit	0.0217	0.0045	0.0078	0.0021	0.0203	0.0372	0.0045
DID = |D2|-|D1|	0.7930	0.1209	0.1902	0.2071	8.8230	3.3302	0.0473
P-value: DID	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0333
DID rank	1	1	1	1	1	1	2
Pennsylvania: Non-metro Counties							
Pre-intervention fit	0.0027	0.0042	0.0088	0.0027	0.0054	0.0257	0.0028
DID = |D2|-|D1|	0.0333	0.0159	0.0346	0.0178	0.0687	0.0913	0.0088
P-value: DID	0.6333	0.5667	0.3333	0.3333	0.8000	0.7000	0.6000
DID rank	20	18	11	11	25	22	19
Pennsylvania: Oil-gas Counties							
Pre-intervention fit	0.0045	0.0144	0.0078	0.0039	0.0025	0.0287	0.0062
DID = |D2|-|D1|	0.0714	0.0483	0.0064	-0.0003	0.4290	-0.3513	0.0159
P-value: DID	0.3000	0.1000	0.9333	0.9000	0.4333	0.9667	0.3333
DID rank	10	4	29	28	14	30	11
Pennsylvania: Top 5 Oil-gas Counties							
Pre-intervention fit	0.0218	0.0122	0.0187	0.0033	0.0096	0.0224	0.0030
DID = |D2|-|D1|	0.1381	-0.0005	0.0263	0.0087	0.2813	0.6732	0.0126
P-value: DID	0.1667	0.9333	0.4667	0.6333	0.5667	0.3000	0.3667
DID rank	6	29	15	20	18	10	12

Notes: (a) Columns: 1 = Construction employment, 2 = Retail employment, 3 = Accommodation and food service employment, 4 = Wage-salary employment, 5 = PC personal income, 6 = Poverty rate, 7 = Population. (b) PC income is in '000 of 2005 dollars, employment and population are in ratio of current to 2001 ratio, and poverty is in percentage. (c) Pre-intervention fit = Absolute prediction error to mean ratio, D1 = Pre-intervention difference, D2 = Post-intervention difference. (d) Except for poverty rate, whenever the estimated impact on the outcome is significant, it indicates an increase in the outcome. In case of poverty, the impact is a decline.

Table 5: SCM Estimation of the Impact of Shale Gas and Tight Oil Boom on Total Employment in North Dakota Counties

Panel A: W-weights

State	Non-metro counties	All oil-gas counties
Arizona	0.000000	0.000000
California	0.000005	0.000000
Connecticut	0.000000	0.000000
Florida	0.000019	0.057381
Georgia	0.000000	0.000000
Idaho	0.000000	0.000000
Illinois	0.000000	0.000000
Indiana	0.000000	0.000000
Iowa	0.454313	0.000000
Maine	0.000826	0.130650
Maryland	0.000000	0.000000
Massachusetts	0.000001	0.071381
Minnesota	0.000000	0.036731
Mississippi	0.000000	0.000000
Missouri	0.000000	0.000000
Nebraska	0.301941	0.000000
Nevada	0.000000	0.000000
New Hampshire	0.000382	0.000004
New York	0.000000	0.000000
North Carolina	0.000000	0.000000
Oregon	0.000000	0.000000
South Carolina	0.000000	0.000000
South Dakota	0.242393	0.649064
Tennessee	0.000000	0.000000
Utah	0.000000	0.000000
Vermont	0.000119	0.000000
Virginia	0.000000	0.000000
Washington	0.000000	0.054783
Wisconsin	0.000000	0.000000

Panel B: Estimation Statistics

	Non-metro counties	All oil-gas counties
SCM: Pre-intervention Fit		
Absolute prediction error to mean ratio	0.0013	0.0022
SCM Inference: Permutations Test		
Pre-intervention difference (D1)	-0.0012	-0.0004
Post-intervention difference (D2)	0.0901	0.1614
DID = |D2|-|D1|	0.0888	0.1610
P-value: DID	0.0333	0.0000
DID rank	2	1

Notes:

(a) List of Predictors

Percent college graduates

Proportion of manufacturing counties

Proportion of farm counties

Industry mix employment growth rate

Natural amenities scale

Incremental distance to 1.5m metro

Incremental distance to 0.5m metro

Natural population growth

Per capital personal income

Total employment of 2006

(b) Intervention is In 2007.

Table 6: SCM Estimation of the Impact of Shale Gas and Tight Oil Boom on Total Employment in Pennsylvania Counties

Panel A: W-weights

State	Non-metro	All oil-gas	Top 5 oil-gas
Arizona	0.000000	0.000000	0.000000
California	0.000000	0.000000	0.000000
Connecticut	0.083671	0.000000	0.000000
Florida	0.000000	0.000000	0.000000
Georgia	0.000000	0.000000	0.000000
Idaho	0.000000	0.000000	0.000000
Illinois	0.036588	0.430079	0.492928
Indiana	0.000000	0.000000	0.049645
Iowa	0.000000	0.000000	0.000000
Maine	0.000000	0.000000	0.000000
Maryland	0.019451	0.000000	0.000000
Massachusetts	0.000000	0.000000	0.000000
Minnesota	0.000000	0.000000	0.000000
Mississippi	0.000000	0.000000	0.000000
Missouri	0.000001	0.000000	0.000000
Nebraska	0.000000	0.000000	0.000000
Nevada	0.000000	0.000000	0.000000
New Hampshire	0.000000	0.000000	0.000000
New York	0.310691	0.115292	0.000000
North Carolina	0.000000	0.000000	0.000000
Oregon	0.104396	0.247529	0.347401
South Carolina	0.000000	0.000000	0.000000
South Dakota	0.000000	0.000000	0.000000
Tennessee	0.193714	0.002551	0.110026
Utah	0.000000	0.000000	0.000000
Vermont	0.000000	0.000000	0.000000
Virginia	0.251486	0.204548	0.000000
Washington	0.000000	0.000000	0.000000
Wisconsin	0.000000	0.000000	0.000000

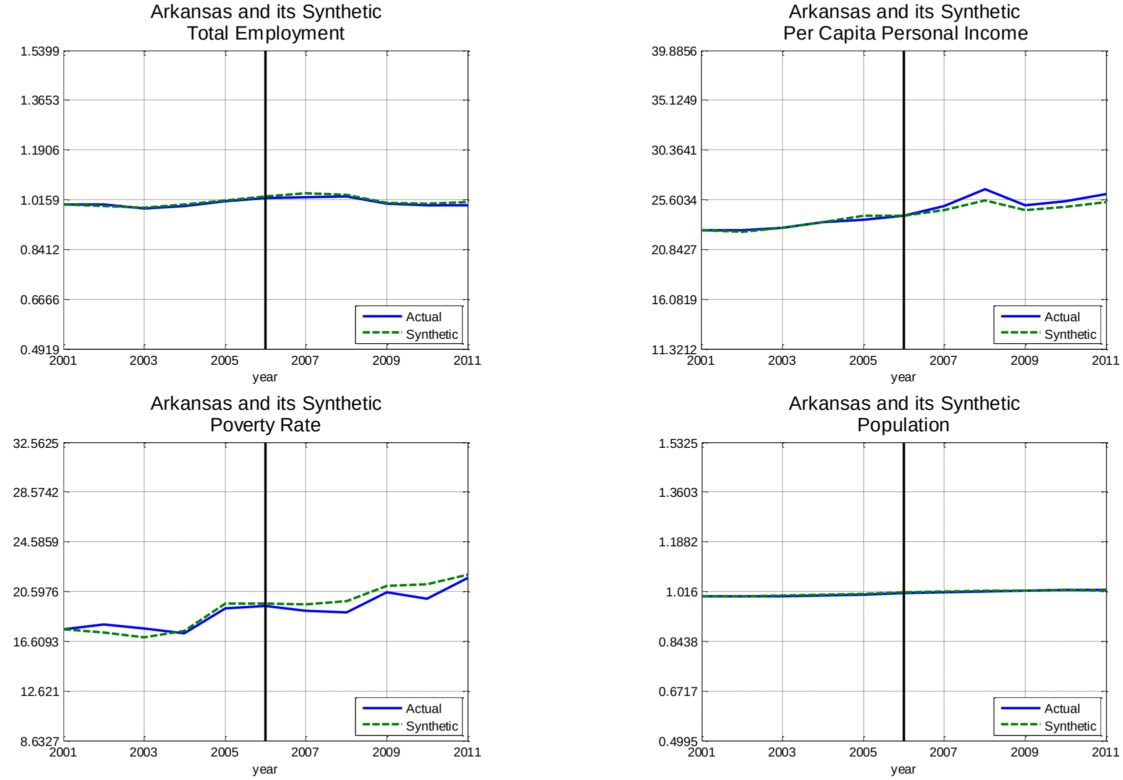
Panel B: Estimation Statistics

	Non-metro	All oil-gas	Top 5 oil-gas
SCM: Pre-intervention Fit			
Absolute prediction error to mean ratio	0.0020	0.0042	0.0035
SCM Inference: Permutations Test			
Pre-intervention difference (D1)	-0.0001	-0.0024	-0.0025
Post-intervention difference (D2)	0.0122	-0.0066	0.0057
DID = |D2|-|D1|	0.0121	0.0042	0.0032
P-value: DID	0.4667	0.7333	0.7000
DID rank	15	23	22

Notes:
(a) List of Predictors
Percent college graduates
Proportion of manufacturing counties
Proportion of farm counties
Industry mix employment growth rate
Natural amenities scale
Incremental distance to 1.5m metro
Incremental distance to 0.5m metro
Natural population growth
Per capital personal income
Total employment of 2006
(b) Intervention is In 2006.

Figure 1: Impact of Shale Gas and Tight Oil Boom in Arkansas – Actual vs Synthetic

Arkansas Oil-gas Counties



Arkansas Top 4 Oil-gas Counties

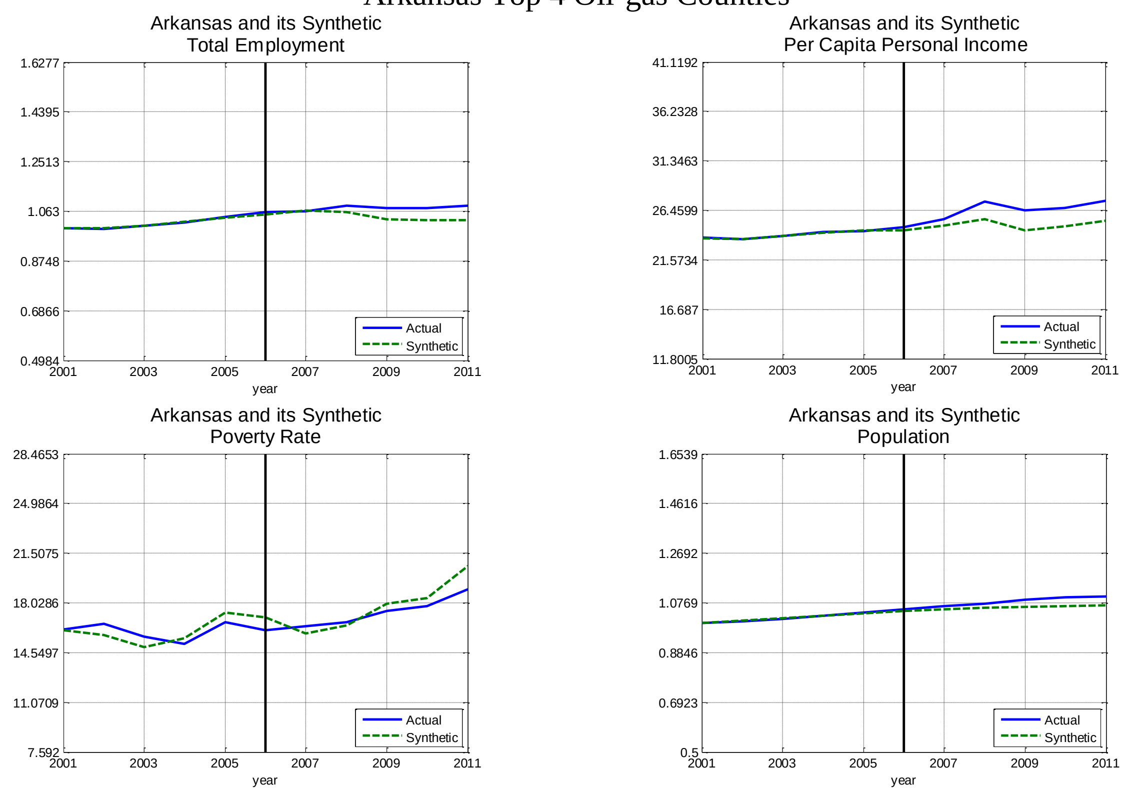
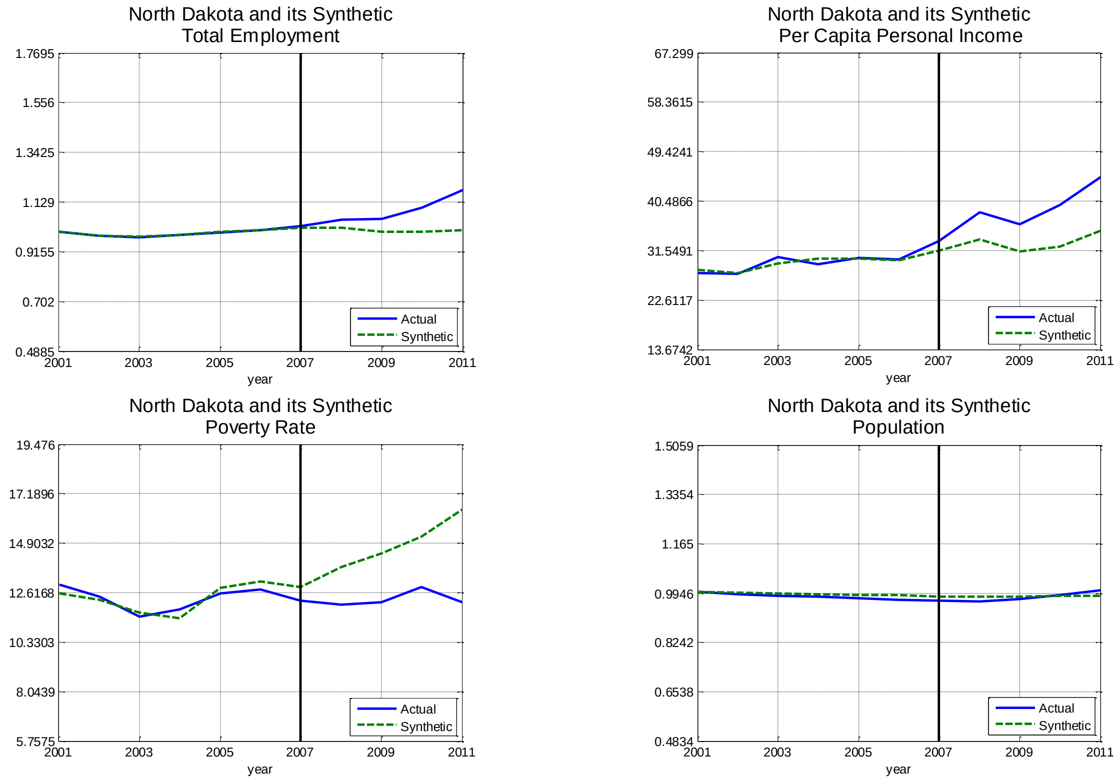


Figure 2: Impact of Shale Gas and Tight Oil Boom in North Dakota – Actual vs Synthetic

North Dakota Non-metro Counties



North Dakota Oil-gas Counties

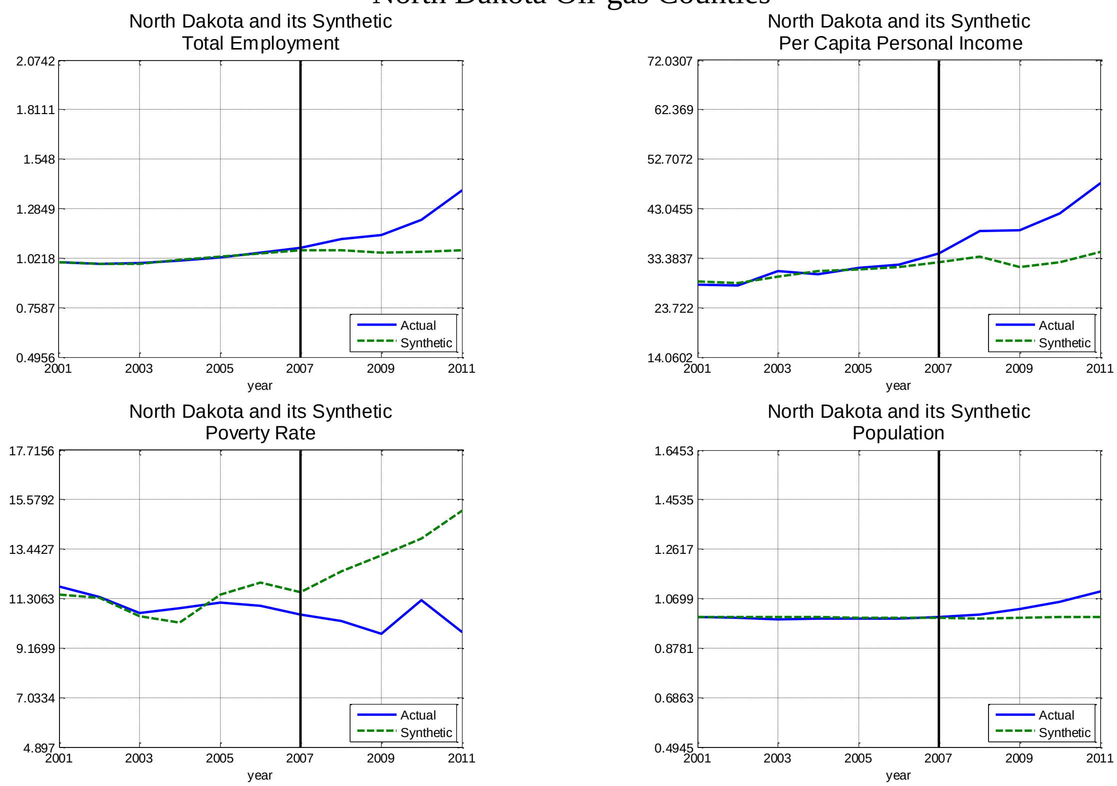
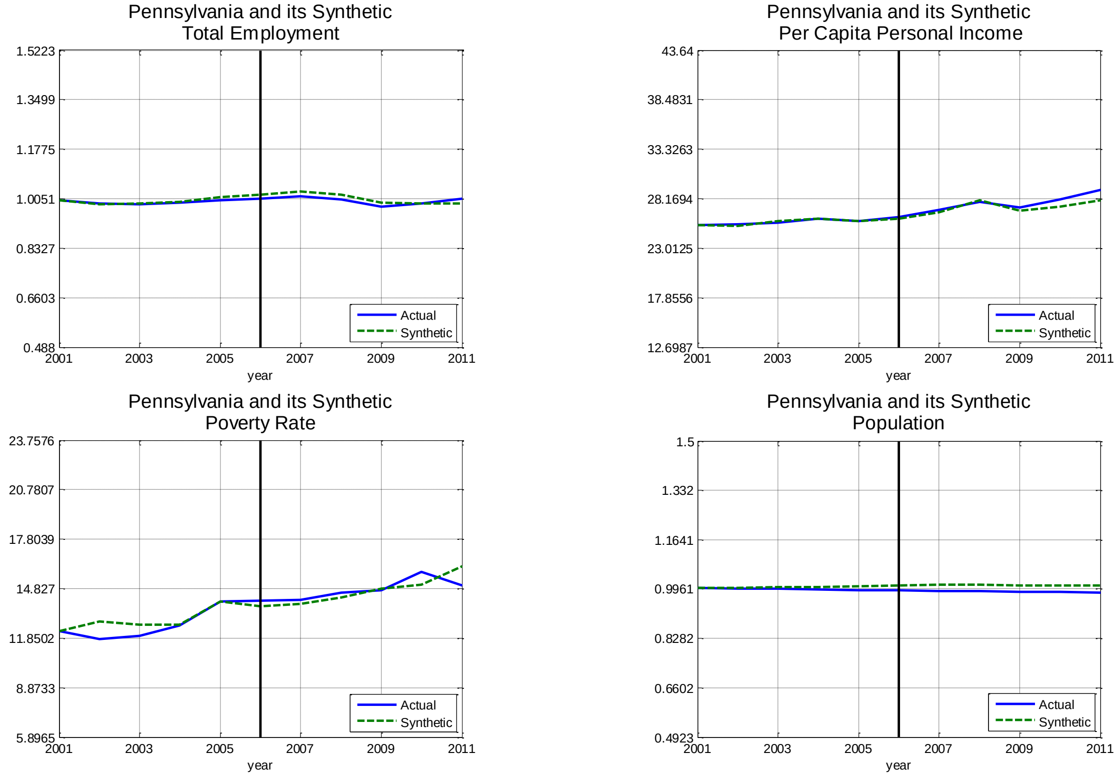
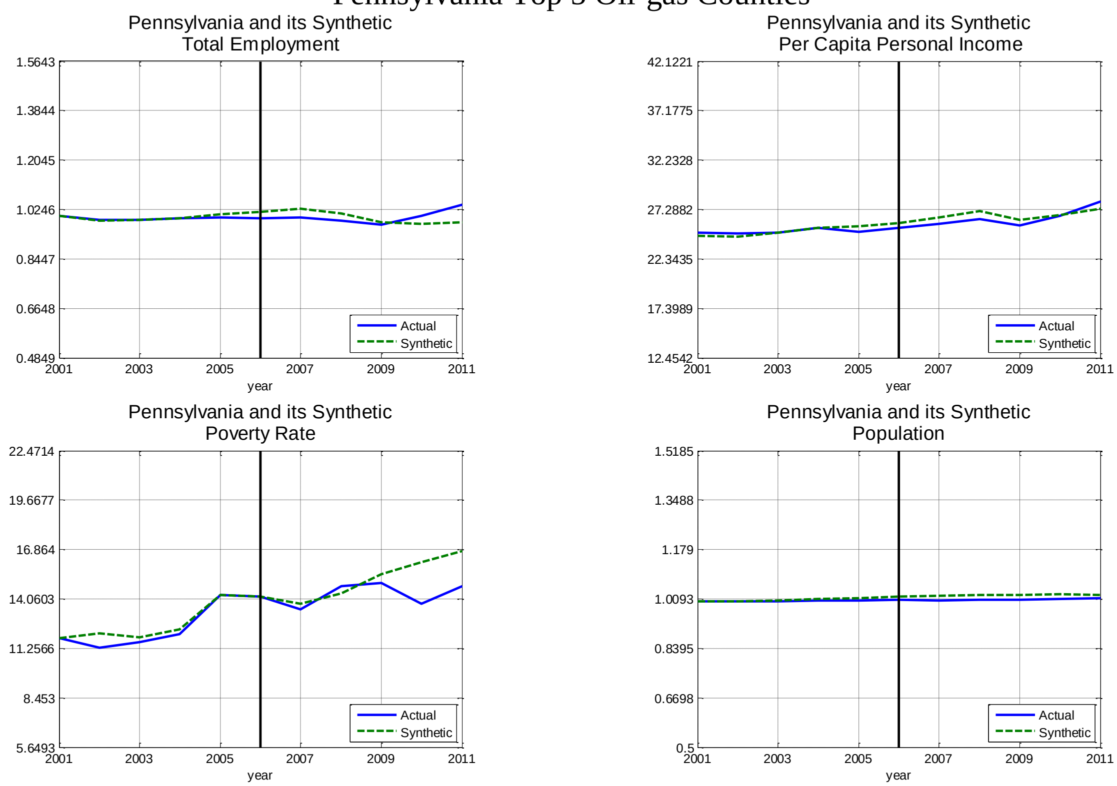


Figure 3: Impact of Shale Gas and Tight Oil Boom in Pennsylvania – Actual vs Synthetic

Pennsylvania Oil-gas Counties



Pennsylvania Top 5 Oil-gas Counties



Appendix: Procedure to Obtain $\mathbf{W}^*$

Let $(T_0 \times 1)$ vector $\mathbf{K} = (k_1, ..., k_{T_0})'$ define a linear combination of pre-intervention outcomes $\tilde{Y}_i^{\mathbf{K}} = \sum_{s=0}^{T_0} k_s Y_{is}$. Define $\mathbf{X}_1 = (\mathbf{Z}_1', \tilde{Y}_1^{\mathbf{K}_1}, ..., \tilde{Y}_1^{\mathbf{K}_M})'$ as a $(k \times 1)$ vector of pre-intervention characteristics for the exposed state where $k = r + M$.[12] Similarly, define a $(k \times J)$ matrix $\mathbf{X}_0$ that contains the same variables for the unexposed states. The $j-\text{th}$ column of $\mathbf{X}_0$, thus, is $(\mathbf{Z}_j', \tilde{Y}_j^{\mathbf{K}_1}, ..., \tilde{Y}_j^{\mathbf{K}_M})'$. Let $\mathbf{V}$ be a $(k \times k)$ symmetric positive semidefinite matrix. Then,

(A1) $$\mathbf{W}^* = \underset{\mathbf{W}}{\text{argmin}}\ (\mathbf{X}_1 - \mathbf{X}_0\mathbf{W})'\mathbf{V}(\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}) \text{ subject to } \{w_j \geq 0 \mid j = 2, ..., J+1\} \text{ and } \sum_{j=2}^{J+1} w_j = 1.$$

Following Abadie and Gardeazabal (2003) and Abadie et al. (2010), we choose $\mathbf{V}$ among positive definite and diagonal matrices such that the mean squared prediction error (MSPE) of the outcome variable is minimized for the pre-intervention periods.

As Abadie et al. (2010) argue, it is important to note that equation (2) is a generalization and the traditional regression-based difference-in-difference model can be obtained if we impose that $\boldsymbol{\lambda}_t$ be constant for all t. Thus, unlike the traditional regression-based difference-in-difference model that restricts the effects of the unobservable confounders to be time-invariant so that they can be eliminated by taking time differences, this model allows the effects of such unobservables to vary with time. In particular, Abadie et al. (2010) show that a synthetic control can fit $\mathbf{Z}_1$ and a long set of pre-intervention outcomes, $Y_{11}, \ldots, Y_{1T_0}$, only as long as it fits $\mathbf{Z}_1$ and μ_1 (unknown factors of the exposed unit).

---

[12] For example, if $M = 2$, $\mathbf{K}_1 = (1,0,...,0)'$, and $\mathbf{K}_2 = (0,0,...,1)'$ then $\mathbf{X}_1 = (\mathbf{Z}_1', Y_1, Y_{T_0})'$, which would mean that the outcome values of the treatment state for the first observed year and the year immediately before the intervention are included in $\mathbf{X}_1$.