RISE | RICE INITIATIVE *for the* STUDY *of* ECONOMICS

RISE Working Paper 14-005

"Local Employment Impact from Competing Energy Sources: Shale Gas versus Wind Generation in Texas"

by

Peter Hartley, Kenneth B. Medlock III, Ted Temzelides, Xinya Zhang

RICE

Department of Economics
Baker Hall, MS22
6100 Main Street, Houston, Texas 77005
https://economics.rice.edu

Local Employment Impact from Competing Energy Sources: Shale Gas versus Wind Generation in Texas*

Peter Hartley[†] Kenneth B. Medlock III[‡] Ted Temzelides[§]

Xinya Zhang[¶]

June 28, 2013

*Preliminary draft. Comments are welcome.

[†]Department of Economics, Rice University

[‡]James A. Baker III Institute for Public Policy, Rice University

[§]Department of Economics, Rice University

[¶]Department of Economics, Rice University

Abstract

The rapid development of both wind power and of shale gas has been receiving significant attention both in the media and among policy makers. Since these are competing sources of electricity generation, it is informative to investigate their relative merits regarding job creation. We use a panel econometric model to estimate the historical job-creating performance of wind versus that of shale oil and gas. The model is estimated using monthly county level data from Texas from 2001 to 2011. Both first-difference and GMM methods show that shale-related activity has brought strong employment to Texas: 77 short-term jobs or 6.4 full-time equivalent (FTE) jobs per well. Given that 5482 new directional/fractured wells were drilled in Texas in 2011, this implies that about 35000 FTE jobs were created in that year alone. We did not, however, find a corresponding impact on wages. Our estimations did not identify a non-negligible impact from the wind industry on either employment or wages.

1 Introduction

There is an ongoing discussion about the relative merits of renewable energy versus those of the shale industry. Proponents of renewable energy emphasize its potential to provide greater energy independence and security, as well as its environmental benefits due to reduced CO_2 emissions. Another argument concerns the potential of renewable energy to act as a driver of economic growth. Since energy produced through renewable sources is still more expensive that that produced through fossil fuels, governments around the world have been providing tens of millions of dollars in subsidies to the renewable industry. More than half of all states in the U.S. have put in place Renewable Portfolio Standards to promote electricity generation from renewable sources.[1] Federal production tax credits and grants also contributed to increases in renewable capacity and generation between 2001 and 2011. Partly as a result, the renewable energy sector has developed rapidly in the past 12 years. In particular, as seen in Figure 1, wind was the fastest growing source among non-hydroelectric renewable sources, as many operators of wind turbines have benefited from tax credit programs. Other sources of non-hydroelectric renewable electricity generation have included biomass, geothermal, and wood, but these have remained relatively stable since 2000.[2] While economists concentrate on objectives related to economic efficiency and growth rather than job creation, policy makers often emphasize the job-creating potential of renewable energy sources. In that regard, renewable energy has the potential to be an important source, and many of these jobs are guaranteed to stay domestically, as they involve construction and installation of physical plants and facilities. Indeed, domestic wind-turbine and component manufacturing capacity has increased recently. Eight of the ten wind turbine manufacturers with the largest

[1] Renewable portfolio standards (RPS), also referred to as renewable electricity standards (RES), require or encourage electricity producers within a given jurisdiction to supply a certain minimum share of electricity from designated renewable resources.

[2] In 2011, in the United States, biomass accounted for about 11% of the total renewable electricity generation, wind accounted for 23%, solar (photovoltaics and concentrating solar power) accounted for 1%, and geothermal for 3%.

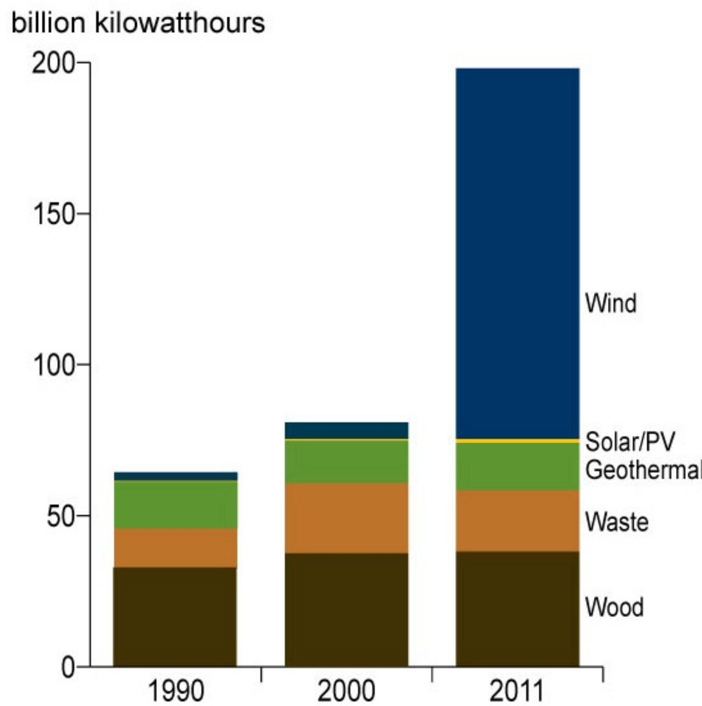


Figure 1: Non hydro-power renewable energy generation, 1990-2011. Data source: EIA

share of the U.S. market in 2011 had at least one manufacturing facility in the United States at the end of 2011. By contrast, in 2004 there was only one active utility-scale wind-turbine manufacturer assembling nacelles in the United States (GE).[3] In addition, a number of new wind turbine and component manufacturing facilities by both foreign and domestic firms either were announced or opened in 2011. The American Wind Energy Association (AWEA) estimates that the entire wind energy sector directly and indirectly employed 75,000 full-time workers in the United States at the end of 2011.

At the same time, another development has been revolutionizing the energy landscape in North America and around the world. This relates to the commercial viability of the vast resources of shale gas and oil. This "shale revolution" is a result of cost-effective technological developments such as horizontal drilling and hydraulic fracturing. The combination of these techniques caused U.S. production of shale oil and gas to boom. The Energy Information

[3]See 2011 Wind Technology Market Report by the U.S. Department of Energy.

Administration's 2012 Annual Energy Outlook (EIA 2012) projects that the share of shale gas as a part of total U.S. natural gas production will increase from 4 percent in 2005 to 34 percent by 2015 and 49 percent by 2025. As shown in Figure 2, shale gas is the largest contributor to natural gas production growth; there is relatively little change in production levels from tight formations, coal-bed methane deposits, and offshore fields. The development of shale

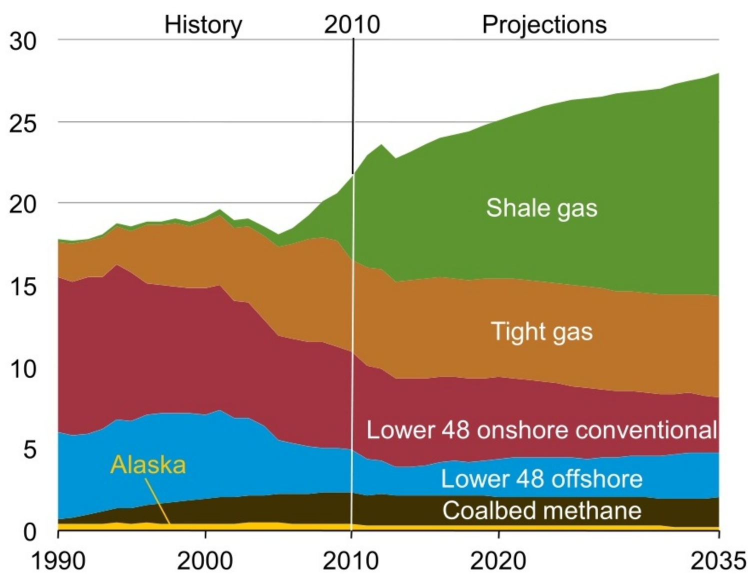
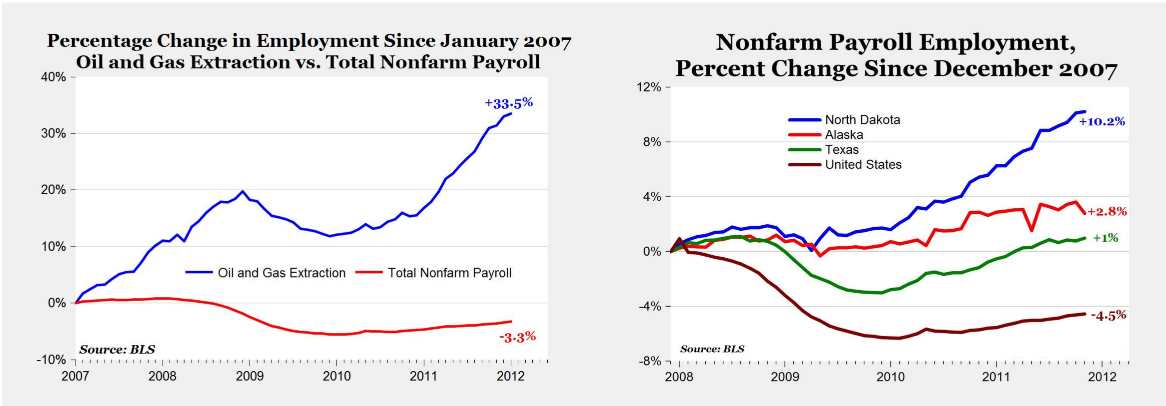


Figure 2: Natural gas production by source, 1990-2035 (TCF). Data source: EIA

gas resources has created an investment boom in the oil and gas industry and led to economic revitalization in places like North Dakota, Alberta, West Pennsylvania, Texas, and Louisiana to name a few. During 2007-2011, employment in the oil and gas extraction sector grew at an annual rate of 7.49 percent, or, 33.5 percent in total. By comparison, during the same period, total employment declined 3.3 percent below its starting value (Figure 3a). Meanwhile, there is anecdotal evidence that rich in shale gas states have experienced increased employment, while the nationwide employment growth rate remains negative (Figure 3b). Furthermore, the relatively low prices resulting from the expanded natural gas supply are stimulating downstream investment in manufacturing,[4] as well as in electricity generation (Figure 4)

[4]This is especially relevant for those sectors that are sensitive to energy costs, such as basic chemicals,

and in transportation.



(a) (b)

Figure 3: Oil and gas extraction employment, 2007-2011

While the aggregate effect on employment from developing different energy sources is an important question, it cannot be readily answered in the context of traditional macroeconomic models. As these models assume market clearing, they cannot account for variations in unemployment rates and, thus, are not well suited to study the consequences of alternative government policies for employment.

Generally speaking, there are two existing approaches focusing on employment impacts of the energy industry. The first uses the input-output (I/O) model. The second approach is based on survey responses from employers, and uses simple descriptive techniques.[5] In this study we collect data on the historical job creation per unit of energy produced by each energy source. We then use this data, together with an econometric model, to estimate the historical job-creating performance of wind versus that of shale gas. Like the analysis based on surveys, ours is a bottom-up approach. At the same time, the econometric techniques used allow us to compare the employment impact in a more systematic and consistent way.

The next section contains the literature review. Section 3 describes the data. The

plastics & rubber, pharmaceuticals, aluminum, pesticides, paints, and fertilizers.

[5]See Section 2 for a detailed discussion.

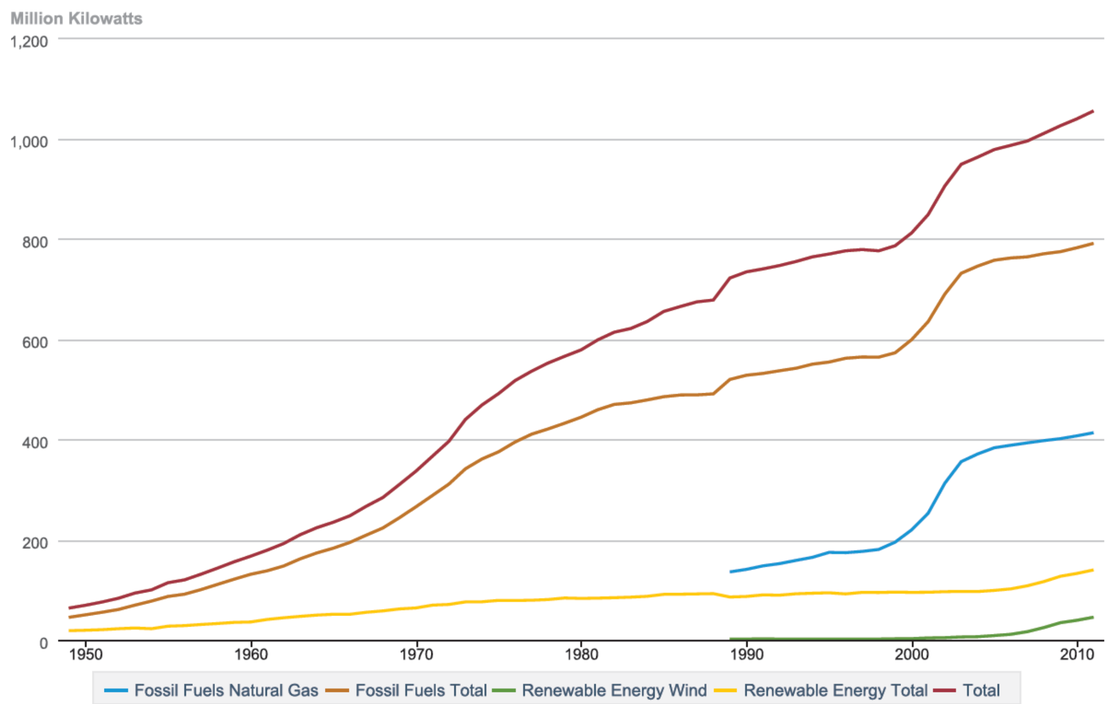


Figure 4: Electricity net summer capacity by source (all sectors), 1949-2011. Data source: EIA

general econometric model is introduced in section 4, while estimation methods and results are reported in section 5. Section 6 offers a short conclusion.

2 Literature Review

In the last few years, a large number of reports have emerged studying employment in shale and in wind. Most of these studies have been completed by non-government organizations and consulting firms, but there have also been a few peer-reviewed journal publications.

Generally speaking, there are two types of studies that focus on the employment impacts in the energy industry. The first uses the input-output(I/O) model. This model is intended to capture the entire economy as an interaction of goods and services between various industrial sectors and consumers. The second approach is based on survey responses from employers. Two widely used I/O studies of the oil & gas industry are the IMPLAN model (see IHS (2012), UTSA (2012), Considine et al. (2009), and American-Chemistry-Council (2011)) and the RIMS II model, used by the U.S. Bureau of Economic Analysis (BEA) (see Scott&Associates (2009)).[6] These studies typically find significant positive effects from the shale oil and gas sectors on jobs, income, and economic growth.

A study of the Eagle Ford Shale (UTSA 2012) estimates the total economic impact on output from shale activity on the local (14-county) region in 2011. It finds that this was just under $20 billion dollars and that it supported 38,000 full-time jobs. If the region is extended to 20 counties, the study finds that 47,097 full-time jobs were supported. A nationwide shale industry report(IHS 2012) has found that the shale gas industry supported 600,000 jobs in 2010. This is projected to grow to nearly 870,000 in 2015, and to over 1.6 million by 2035. Two reports on the Marcellus shale by Pennsylvania State University (Considine et al. 2009) and by West Virginia University (Higginbotham et al. 2009) show that the oil and gas industry in Pennsylvania generates $3.8 billion in value added, and over 48,000 jobs in 2009. In west Virginia, the economic impact of the oil and natural gas industry in 2009 was estimated to be $3.1 billion in total value added, while approximately 24,400 jobs were created.

The Jobs and Economic Development Impact (JEDI) model developed by the National Renewable Energy Laboratory (NREL) is a series of spreadsheet-based I/O models that estimate the economic impacts of constructing and operating power plants, fuel production facilities, and other projects at the local (usually state) level. Slattery et al. (2009) employs the JEDI Wind Energy Model to examine the economic impact of the large-scale wind-farm construction. They then test the model validation using data from NextEra's Capricorn Ridge and Horse Hollow facilities. They find that the JEDI model overestimates local share

[6]The IMPLAN model uses a national input-output dollar flow table called the Social Accounting Matrix (SAM) to model the way a dollar injected into one sector is spent and re-spent in other sectors of the economy. RIMS II provides I/O multipliers that measure the effects on output, employment, and earnings from any changes in a region's industrial activity.

of jobs during the construction phase in smaller, rural counties, and that it underestimates by more than 50% the number of jobs in large, urban counties. This is because the JEDI model sets the same local share value to all counties, and it does not consider urban effects. In addition, the JEDI model assumes 100% local share for operations and maintenance (O&M) jobs, which might be implausible, especially in small rural counties.

I/O models provide the most complete picture of the economy as a whole. They capture employment multiplier effects, as well as the macroeconomic impacts of shifts between sectors. Hence they could account for losses in one sector (e.g., conventional oil industry) created by the growth in another sector (e.g., wind energy). One drawback is that collecting data for an I/O model is highly labor-intensive. As a result, the calibration process for default multiplier parameters may be biased due to lack of information. Bottom up estimates, on the other hand, are based on industry/utility surveys, on outlooks by project developers and equipment manufacturers, and on primary employment data from companies across manufacturing, construction, installation, and O&M. For wind energy, most reports are analytical-based studies, and only calculate direct employment impacts. A case study on the economic effects of the Gulf wind project in Texas reports the estimated creation of 250 - 300 jobs during the peak construction period (9 months), and 15 - 20 permanent jobs.[7]

A report on the wind industry from the Natural Resources Defense Council (NRDC) measures the number of direct jobs that a typical wind farm may create across the entire value chain. They analyze each of the 14 key value-chain activities independently to determine the number of workers involved at each step in the wind farm building. They find that one typical wind farm of 250MW would create 1079 jobs over the lifetime of the project. Similarly, the Renewable Energy Policy Project (REPP) has developed a spreadsheet-based model using data based on a survey of current industry practices. They use it to calculate the number of direct jobs from wind, solar photovoltaic, biomass, and geothermal activities

[7] Gulf Wind: Harnessing the Wind for South Texas

that would result from the enactment of a Renewable Portfolio Standard . They find that every 100 MW of wind power installed creates 475 jobs in total (313 manufacturing jobs, 67 installation jobs, and 95 jobs in O&M).

3 Data

We use data from the state of Texas. Texas has rich shale gas and oil resources and, at the same time, it is the national leader in wind installations and a manufacturing hub for the wind energy industry. According to EIA, Texas accounted for 40 percent of U.S. marketed dry shale gas production in 2011, making it the leading unconventional gas producer in the U.S. Meanwhile, Texas leads the nation in wind-powered generation capacity and it is the first state to have reached 10,000 megawatts of wind capacity. Next, we discuss the variables that will be included in our empirical analysis, our data sources, and some summary statistics.

In Texas, there are 254 counties.[8] For each county $i = 1, \ldots, 254$, we have collected observations for $T = 132$ months, or a total of 11 years (2001 - 2011), making the panel balanced. We will use total employment in all industries as a dependent variable. We did not use data from specific industries since, besides direct job creation, we want to measure total employment effects, including indirect job-creation. This includes jobs created in upstream and infrastructure supplying industries, as well as induced jobs, such as jobs added in sectors supplying consumer items (food, auto, housing, etc.) and services. Another dependent variable of interest is the average weekly wage, since, according to economic theory, it should also be impacted by an increase in the demand for workers. We use monthly employment data and quarterly wage data from the Quarterly Census of Employment and Wages (QCEW) Database of the Bureau of Labor Statistics (BLS).[9] The latter has been adjusted to a real

[8] Out of these, 77 are urban counties.

[9] The QCEW employment and wage data is derived from micro data summaries of 9.1 million employer reports of employment and wages submitted by states to the BLS in 2011. These reports are based on place of employment rather than place of residence. Average weekly wage values are calculated by dividing

wage using the implicit price GDP deflator (IPD) from BEA.[10]

In order to evaluate the impact from shale and wind development on employment and on the local economy, we need to first devise a method for measuring the activity in the shale and wind industries. The key explanatory variables we will use are the number of unconventional wells completed and the newly installed wind capacity in each county and each month, respectively.

Of course, other variables could also be used to reflect other aspects of activity in the shale industry. These include the number of permits issued, rig counts, the number of wells spudded, and total shale gas production. We chose the number of wells completed because the well completion date indicates the end of the construction period for each well. We suspect that more direct and on-site jobs are created during that period. To fully describe the impact of shale on employment, especially the multiplier effects on job creation in the local economy, we allow well drilling activities to affect employment with a lag, and we study both pre-completion and post-construction effects.

In the shale industry, the entire process from spudding to producing marketed output can take up to 3-4 months. Horizontal drilling itself currently takes approximately 18-25 days from start to finish. Then wells are fractured to release the gas before the well is completed. The well is then connected to a pipeline, which transports the gas to the market. Among these activities, hydraulic fracturing and the last step before completion are the most labor intensive. Hence, we expect drilling activities to have a peak impact on employment in the pre-completion period and, in particular in the month of well completion.

We use the Drilling Info Database for well information. We concentrate our study on wells that are both directional/horizontally drilled and hydraulically fractured.[11] Thus, we

quarterly total wages by the average of the three monthly employment levels and then dividing the result by 13, as there are 13 weeks in a quarter.

[10]The implicit price GDP deflator is the ratio of the current-dollar value of GDP to its corresponding chained-dollar value multiplied by 100.

[11]This filter option is only available for Texas data.

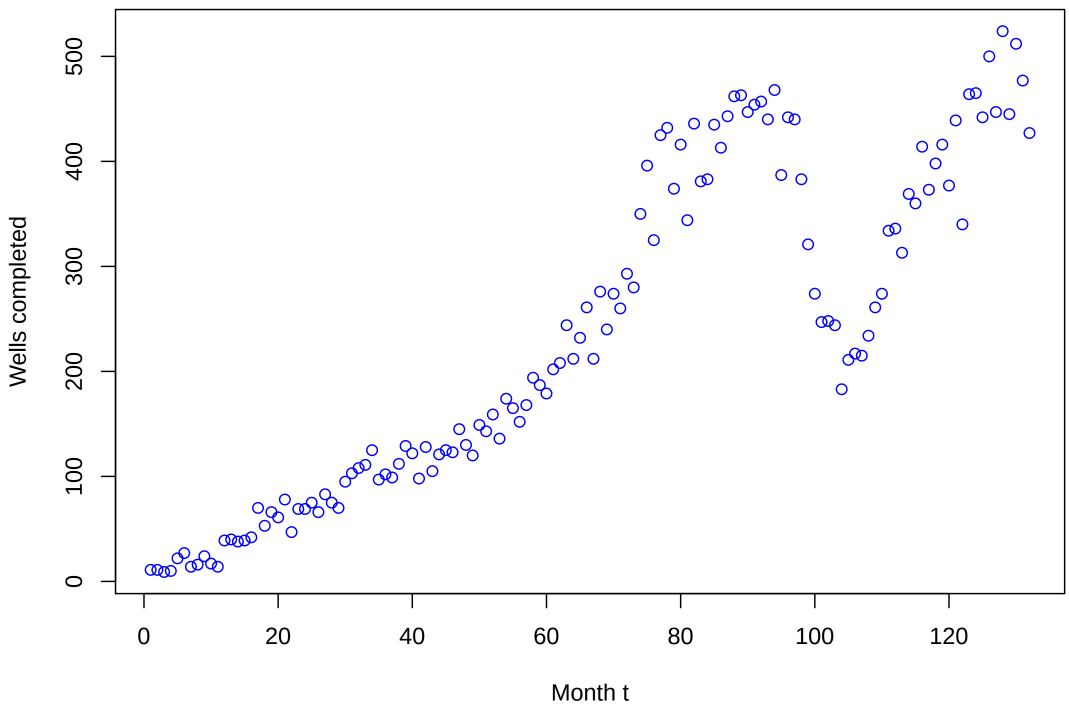


Figure 5: Number of completed new wells per county per year, Jan, 2001 - Dec, 2011

exclude conventional oil/gas wells from our data set. There were 31050 directional/horizontal and fractured wells completed in 174 Texas counties during 2001 - 2011, including 25467 gas wells, 4963 oil wells, and 620 wells classified as "other types." Figure 6 indicates that shale gas developed very quickly in the past 12 years, from 1 well per month in Jan, 2001 to around 500 in 2011. The completion date and location of each well are used to count the number of wells completed in each county each month. We then compute the cumulative number of wells, assuming that this number is 0 at the starting point of Jan, 2001.

To measure wind activity in each county we will use the installed nameplate capacity online per month. We are not using power generation as more jobs are created during the construction period than during the O&M period. The installed capacity and online year for wind projects in Texas through the 2007-2011 is taken from the American Wind Energy

Association (AWEA). For wind projects before 2007, we used EIA electricity data on plant level output and a wind industry progress report by Wind Today. To find the online month and county location for each wind project, we needed to refer to additional sources, such as specific project information from individual projects' websites, as well as local news stories for the projects' online years. For those wind farms that cover several neighboring counties, we divided installed capacity equally between each of the counties involved. Until 2011, we found that 125 wind projects had been constructed in 40 counties, with a total installed capacity of 10006MW (compared to 6 counties and 920MW in 2001).

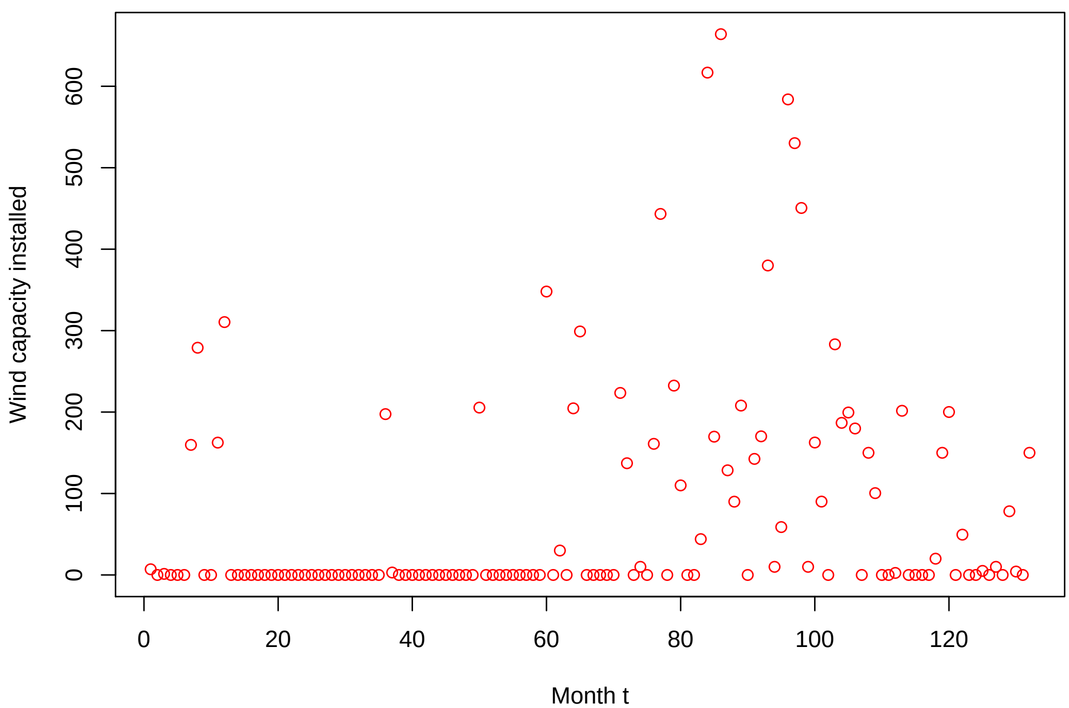


Figure 6: New wind capacity installed during Jan, 2001 - Dec, 2011

Since regression test statistics do not have the usual asymptotic distributions when variables are non-stationary, we first need to investigate the stationarity properties of our variables before we use them in regression analysis. To test for non-stationarity in our panel data setting, we consider the following model written in difference form:

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{L=1}^{p_i} \delta_i \Delta y_{i,t-L} + \alpha_0 + \alpha_1 t + u_{it},\ t = 1, 2, \ldots \tag{1}$$

We then test the hypothesis that $\rho = 0$. Note that the term $\alpha_0 + \alpha_1 t$ allows for a constant and deterministic time trend. When $\rho = 1$, the series y_{it} has a unit root and it is a random walk. The random walk process is simply a sum of all past random shocks, which also implies that the effect of any one shock lasts forever. When $\rho < 1$, y_t is stationary and the correlation between variables declines as they get farther apart in time.[12]

We use the Im, Pesaran, and Shin (IPS) test (Im et al. 2003) to test for $\rho = 0$. This test is based on the estimation of the above augmented Dickey-Fuller (ADF) regressions for each time series. A statistic is then computed using the t-statistic associated with the lagged variable. Note that this procedure does not require ρ to be the same for all counties. The null hypothesis is that all series have a unit root, and the alternative is that some have a unit root while others have different values of $\rho_i < 0$. To run the test, we must first determine the optimal number of lags, p_i, for each time series in the panel. If we assume too few lags, u_{it} will be serially correlated and the test statistics will not have the assumed distribution. If we assume too many lags, the power of the test statistic declines. Since we are working with monthly data, we set the maximum p_i at 14, which is slightly larger than an annual cycle. We then use both the Swartz information criteria (SIC) and the Akaike information criteria (AIC) to determine the optimal value of p_i. We also used a test based on Hadri (2000) as a complement. The Hadri statistic does not rely on the ADF regression. It constitutes the cross-sectional average of the individual KPSS statistics (Kwiatkowski et al. 1992), standardized by their asymptotic mean and standard deviation. It tests the null hypothesis that all panels are stationary, while the alternative is that some panels contain

[12]Specifically, y_t is covariance stationary, which means that the correlation between y_t and y_{t+h} only depends on h.

unit roots. For both the employment and the wage series, we found that the p-value of the IPS test is close to zero. Hence, H_0 is rejected and we conclude that some counties may have no unit roots. On the other hand, the Hadri test rejects H_0 as well, implying that at least one county has a unit root. Hence, we conclude that the employment and wage series have unit roots in some counties, while others are stationary. We then applied the Dickey-Fuller Generalized Least Squares test (DF-GLS) (at 5% level) and the KPSS test (at 10% level) to each county. We found that in 156 counties, the DF-GLS test cannot reject a unit root, while the KPSS test shows the presence of a unit root.[13]

4 Econometric Modeling

We estimate a regression relationship via a panel data approach using a data set of 254 counties in Texas covering the years 2001-2011. Working with panel data allows us to study dynamic relationships, which we cannot do using a single cross section. It also allows us to test for the presence of specific effects in counties with shale and wind activities versus those without. One potential problem with time series analysis is that many exogenous factors might change at the same time, making it difficult to attribute an outcome to any one particular change. The panel enables us to interpret differences between counties since other factors are presumed to affect all counties symmetrically. Comparing to time series and cross sections using over-time data from the same counties is useful as it allows us to look at dynamic relationships which we cannot do with a single cross section. In addition, a panel data set also allows us to control for unobserved cross section heterogeneity.

[13]Among the 156 counties, 46 are urban. That is, employment in 60% of the urban counties and 62% of the rural counties may have unit roots. On the other hand, based on results from both tests, only 34 counties have stationary employment series.

4.1 Assumptions

We start with a static linear unobserved effects model.

$$y_{it} = \mathbf{x}_{it}\beta + \theta_t + c_i + u_{it}, t = 1, 2, \ldots, T, \quad (2)$$

where y_{it} is a scalar, $\mathbf{x}_{it}$ is a $1 \times K$ vector for $t = 1, 2, \ldots, T$, and β is a $K \times 1$ vector. Here, c_i indicates a time-invariant unobservable effect, and θ_t represents a series of time fixed effects.

To make the model more realistic, we allow for arbitrary dependence between the unobserved effects, c_i, and the observed explanatory variables, $\mathbf{x_{it}}$. For example, underground geology characteristics would be included in c_i, and these would very well be correlated with the number of wells drilled in county i. Also, wind capacity highly depends on the climate, and especially the wind resource of the county, which is also part of the variable c_i.

We also assume that the explanatory variables are strictly exogenous conditional on the unobserved effect, c_i. This terminology, introduced by Chamberlain (1982), requires that

$$E(u_{it}|\mathbf{x}_i, c_i) = 0, t = 1, 2, \ldots, T. \quad (3)$$

That is, once $\mathbf{x_{it}}$ and c_i are accounted for, $\mathbf{x_{is}}$ has no partial effect on y_{it}, for $s \neq t$. In addition, u_{it} has zero mean conditional on all explanatory variables in all time periods. This is a stronger assumption than contemporaneous exogeneity, which requires that $E(u_{it}|\mathbf{x_{it}}, c_i) = 0$. This says nothing about the relationship between $\mathbf{x_s}$ and u_t for $s \neq t$. Sequential exogeneity, which requires that $E(u_{it}|\mathbf{x_{it}}, \mathbf{x_{i,t-1}}, \ldots, \mathbf{x_{i1}}, c_i) = 0$, for $t = 1, 2, \ldots, T$, is stronger than contemporaneous exogeneity. It implies that $\mathbf{x_s}$ is uncorrelated with u_t for all $s \leq t$, but imposes no constraints on the correlation between $\mathbf{x_s}$ and u_t for $s > t$.

It is standard to assume zero contemporaneous correlation; i.e., that u_{it} is uncorrelated with the number of wells drilled, or the wind capacity installed at t. But what about the correlation between u_{it} and, say, $\mathbf{x_{i,t+1}}$? Does future well drilling activity or wind-farm construction depend on past shocks to the county's employment? We do not believe that such feedback is important for our study, since total employment in a county is not the main goal of energy companies. Therefore, it seems reasonable to assume that past employment has a negligible effect on energy companies' future plans.

Another issue is that the explanatory variables could have lasting effects, so that a correlation might exist between u_{it} and past $\mathbf{x_{i,t-1}}, \ldots \mathbf{x_{i,1}}$, leading to a failure of sequential exogeneity. This be the case if well-drilling activity and wind-activity have lasting effects on local employment. One way to deal with this kind of correlation is to include lags of the explanatory variables into the model. Strict exogeneity would then hold, if enough lags are included.[14]

A test of strict exogeneity is based on Wooldridge (2002), 10.7.1.

$$\Delta y_{it} = \mathbf{\Delta x_{it}}\beta + \mathbf{w_{it}}\gamma + \Delta u_{it}, t = 2, \ldots, T, \tag{4}$$

In the above equation, $w_{i,t}$ is a subset of $x_{i,t}$. Under strict exogeneity, none of the $\mathbf{x_{it}}$s should be significant explanatory variables in the first difference (FD) equation. That is, we should find support of the hypothesis H_0: $\gamma = 0$. Carrying out this test, the F statistic on γ is 0.32, with $p - value = 0.5695$. Thus, we could not reject H_0. The strict exogeneity assumption fails in models with unobserved effects and lagged dependent variables. The reason is that y_{it} is correlated with u_{it} and would show up as part of explanatory variables at $t+1$, implying that $E(u_{it}|\mathbf{x_{i,t+1}}) \neq 0$. Additional care is required when we include lagged dependent variables as explanatory variables on the right hand side.

We have not ruled out serial correlation in the idiosyncratic error u_{it}, that is, $Corr(u_{it}, u_{is}) \neq 0$, $t \neq s$. Specifically, here we will only consider serial correlation across time, excluding any cross-sectional correlations a priori. If one allows for the u_{it}s to be serially correlated over

[14] Another remedy is to use instrumental variables. This method is somewhat involved, as it relies on finding suitable instruments.

time, the usual pooled ordinary least squares (OLS) and fixed effects (FE) standard errors are not valid, even asymptotically. To test for the existence of serial correlation in the u_{it}s, we use the Breusch-Godfrey/Wooldridge's LM test and the Wooldridge first difference test (Wooldridge 2002). Rather than interpreting serial correlation as a technical violation of the OLS assumption, we will think of time series data in the context of economic dynamics. This leads us to the inclusion of lagged variables.

4.2 The Finite Distributed Lag (FDL) Model

Since we expect that drilling and wind activity can have lasting effects on local employment, we include lags of explanatory variables into the model. A finite distributed lag model is appropriate if the impact of the explanatory variables lasts over a finite number of periods, q, and then stops. The FDL unobserved effects model expands equation (2) to the following form:

$$E_{it} = \sum_{k=0}^{q} \beta_k wells_{i,t-k} + \sum_{k=0}^{q} \delta_k wcap_{i,t-k} + c_i + \theta_t + u_{it} \tag{5}$$

where E_{it} denotes total employment, $wells_{it}$ denotes the number of directional/fractured wells drilled, and $wcap_{it}$ indicates installed wind capacity for $i = 1, 2, \ldots, 254$ and $t = 1, 2, \ldots, T$. Our interest lies in the pattern of coefficients $\{\beta_k, \delta_k\}_{k=0}^{q}$. The values of β_0 and δ_0 capture the immediate change in E_i due to the one-unit increase in $wells_i$ and $wcap_i$, respectively, at time t. Similarly, β_k and δ_k capture the changes in E_i, k periods after the temporary change. At time $t + q$, E_i has reverted back to its initial level, $E_{i,t+q} = E_{i,t-1}$.

We are also interested in the change in E_i due to a permanent increase in any of the explanatory variables. For example, following a permanent increase in $wells_{it}$, $E_{i,t+1}$ increases by $\beta_0 + \beta_1$ after one period, and $E_{i,t+k}$ increases by $\beta_0 + \cdots + \beta_k$ after k periods. After q periods, there are no further changes in E_i. Thus, the sum of the coefficients on current and lagged $wells_i$ is the long-run change in E_i, which is also referred to as the long-run propensity (LRP). The same analysis applies for the impact of variable $wcap_i$ on E_i.[15]

4.3 The Autoregressive Distributed Lag (ADL) Model

One way of bypassing the multicollinearity problem is by including a lagged dependent variable with fewer lags than the explanatory variables. The model becomes an autoregressive distributed lag (ADL) model. The ADL model is similar to the FDL model, except that the impact of explanatory variables persists over time at a geometrically declining rate. Denoting the number of lagged dependent variables by p, an ADL(p, q) model with unobserved effects has the form:

$$E_{it} = \sum_{j=1}^{p} \lambda_j E_{i,t-j} + \sum_{k=0}^{q} \beta_k wells_{i,t-k} + \sum_{k=0}^{q} \delta_k wcap_{i,t-k} + c_i + \theta_t + u_{it} \tag{6}$$

where $\{\lambda_j\}_{j=1}^{p}$ are the autoregressive coefficients. If there is a temporary change in $wells$, E_{it} will initially go up by β_0 in period 1, then by $\beta_1 + \lambda_1\beta_0$ in period 2, by $\beta_2 + \lambda_1(\beta_1 + \lambda_1\beta_0) + \lambda_2\beta_0$ in period 3... etc. If we have a unit level change and provided that the process is stationary, the ADL model eventually reaches a new equilibrium that is

$$\frac{\sum_{k=0}^{q} \beta_k}{1 - \sum_{j=0}^{p} \lambda_j} \tag{7}$$

higher than the original equilibrium.

Another advantage of the ADL model is that the inclusion of a lagged dependent variable can eliminate serial correlation, particularly if additional lags of the dependent variable are included. Lags of the independent variables may also assist with eliminating serial correlation in the error term. Hence, once we introduce lagged values of y_{it}, dynamic completeness implies sequential exogeneity. However, the strict exogeneity assumption is false, as we

[15] Of course, the right lag length is rarely known in advance, or pinned down by theory. In addition, time series data is often limited and the inclusion of lagged variables may reduce degrees of freedom. The fact that the explanatory variables are likely to be correlated might lead to severe multicollinearity.

discussed before. In this case, both the fixed effects (FE) estimator and the first difference (FD) estimator are inconsistent.[16]

4.4 The Spatial Panel Model

We now discuss cross-sectional dependence (XSD) in panels. This can arise, for example, if spatial diffusion processes are present, relating panel members (in our case, counties) in a way that depends on a measure of distance. The CD and $CD(p)$ tests (Pesaran 2004) are used to detect XSD. These tests are based on the averages over the time dimension of pairwise correlation coefficients for each pair of cross-sectional units. The $CD(p)$ test also takes into account an appropriate subset of neighboring cross-sectional units in order to check the null of no XSD against dependence between neighbors only. To do so, a spatial weights matrix, W, is needed for the $CD(p)$ test.

In our data set, both tests show the presence of XSD at 0.000 level. This is not surprising, since it seems likely that employment might be correlated across counties. We use a spatial panel model to study this spatial interaction effect across counties and try to capture the indirect effect of a county's energy sector development on employment within other counties. Spatial interaction effects could be due to competition or complementarity between counties, spillovers, externalities, regional correlations in industry structures, as well as many other factors.

Interactions between spatial units are typically modeled in terms of some measure of distance between them, which is described by a spatial weights matrix W. Here, W is a 254×254 non-negative matrix, in which the element w_{ij} expresses the degree of spatial proximity between the pair of objects i and j. Following Kapoor et al. (2007), the diagonal

[16] Deciding which model to use and how many lags to include is complicated by the fact that we are unlikely to have a theory to distinguish between the different models. As a result, Boef & Keele (2008) and others have advocated starting with a general model like the ADL and testing down to a more specific model, including the optimal values for p and q.

elements w_{ii} are all set to zero, to exclude self-neighbors. Furthermore, only neighborhood effects are considered in this paper, that is, W is a contiguity matrix:[17]

$$w_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are neighbors} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Then the contiguity matrix is transformed into row- standardized form, which assumes the impact on each unit by all other neighboring units are equal. Given a spatial weights matrix W, a family of related spatial econometric models can be expanded from equation (2):

$$E_{it} = \rho \sum_{j=1}^{N} w_{ij} E_{jt} + \beta_1 wells_{it} + \beta_2 wcap_{it} + u_{it}, \tag{9}$$

where ρ is the spatial autoregressive coefficient. The composite error u_{it} can be specified in two ways. First we can have $u_{it} = c_i + \epsilon_{it}$, while ϵ_{it} is a vector that follows a spatial autoregressive process of the form

$$\epsilon_{it} = \lambda \sum_{j=1}^{N} w_{ij} \epsilon_{jt} + \nu_{it} \tag{10}$$

with λ being the spatial autocorrelation parameter. The second specification we consider for the error u_{it} is the one in Kapoor et al. (2007). They assume that spatial correlation applies to both unobserved individual effects and the remainder error components. In this case, u_{it} follows a first order spatial autoregressive process of the form:

$$u_{it} = \lambda \sum_{j=1}^{N} w_{ij} u_{jt} + \epsilon_{it} \tag{11}$$

[17] W is also called as adjacency matrix.

and ϵ follows an error component structure

$$\epsilon_{it} = c_i + \nu_{it} \tag{12}$$

to further allow ϵ_{it} to be correlated over time. Although the two data generating processes share some similarities, they imply different spatial spillover mechanisms governed by a different structure of the implied variance covariance matrix. In what follows, we will consider the implementation of the second error term specification, which leads to a simpler variance matrix that is also easier to invert.

5 Estimation Methods and Results

Let us first turn to the general unobserved effects model (2). The pooled OLS estimator can be used to obtain a consistent estimator for β only if the explanatory variables satisfy contemporaneous exogeneity and zero correlation with the unobserved individual effects. Earlier, we assumed that contemporaneous exogeneity holds. However, as we discussed before, explanatory variables are necessarily correlated with the unobserved individual effects, c_i. In addition, F-tests of poolability show that the pooled OLS estimation is inconsistent. Since random effects analysis also requires orthogonality between the c_is and the observed explanatory variables, it is also inconsistent. Finally, the result of the Hausman test, namely $\chi^2 = 601.67$ with a p-value close to zero, indicates that the random effects approach is inconsistent.

With a fixed effects (FE) or first difference (FD) approach, the explanatory variables are allowed to be arbitrarily correlated with c_i, but strict exogeneity conditional on the c_is is still required. The idea behind the fixed effects approach is to transform the equations by removing the intertemporal mean, thereby eliminating the unobserved effects. One can then apply pooled OLS to get FE estimators. Similarly, the FD approach transforms the equations by lagging the model one period and subtracting, then applying pooled OLS to get FD estimators. As we mentioned earlier, we found that more than half of the counties have highly persistent employment series. Using time series with a unit root process in a regression equation could cause a spurious regression problem. In that case, first differencing should be used to remove any potential unit roots.

5.1 The FDL model

As noted earlier, including lagged dependent variables in the model violates strict exogeneity, implying that the resulting autoregressive FD estimator may suffer from asymptotic bias. Therefore, in this section we drop all lagged dependent variables and use the FDL approach. We verified that the strict exogeneity assumption holds as long as enough lags of the explanatory variables are included.

To obtain a First-Difference (FD) estimator, we lagged the model in (5) by one period and subtracted to obtain:

$$\Delta E_{it} = \sum_{k=0}^{q} \beta_k \Delta wells_{i,t-k} + \sum_{k=0}^{q} \delta_k \Delta wcap_{i,t-k} + \theta_0 + \theta_t + \Delta u_{it},\ t = 2, 3, \ldots, T \tag{13}$$

Note that rather than dropping an overall intercept and including the differenced time dummies $\Delta\theta_t$, we estimated an intercept and then included the time dummies θ_t, for $T-2$ of the remaining periods. Because the set of the regressors involving the time dummies are non-singular linear transformations of each other, the estimated coefficients on the other variables do not change.

If we find no serial correlation, the FD estimator is the pooled OLS estimator from the regression model (13). The only remaining issue involves deciding how to choose the number of lags, q. To test for the presence of serial correlation in Δu_{it}, we use the tests by Breusch-Godfrey and Wooldridge. Both tests reject H_0 and show serial correlation remains in the

Variable	Coefficient	(Std. Err.)	(Robust SE.)
$wells_{it}$	16.31	(5.396)**	[6.06]**
$wells_{i,t-1}$	13.17	(6.666)*	[7.081]$^{\cdot}$
$wells_{i,t-2}$	0.932	(7.025)	[2.929]
$wells_{i,t-3}$	-5.519	(7.127)	[6.006]
$wells_{i,t-4}$	12.23	(7.119)$^{\cdot}$	[8.705]
$wells_{i,t-5}$	17.89	(6.875)**	[11.13]
$wells_{i,t-6}$	22.46	(5.686)***	[12.91]$^{\cdot}$
$wcap_{it}$	-0.756	(1.235)	[0.923]
$wcap_{i,t-1}$	-0.755	(1.653)	[0.594]
$wcap_{i,t-2}$	-0.739	(1.864)	[0.332]*
$wcap_{i,t-3}$	-0.212	(1.923)	[0.323]
$wcap_{i,t-4}$	0.111	(1.865)	[0.374]
$wcap_{i,t-5}$	0.250	(1.654)	[0.432]
$wcap_{i,t-6}$	-0.178	(1.236)	[0.266]

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 1: FD Estimation Results, $q = 6$

idiosyncratic errors. We then increased p up to $p = 36$. The results showed that serial correlation remained no matter how many lags of the explanatory variables were included. This serial correlation may imply that the model does not fully capture the actual dynamic adjustment process.

We proceed by computing a robust variance matrix for the FD estimator, which accommodates a fully general structure with respect to heteroskedasticity and serial correlation in Δu_{it}. Following Arellano (1987), this robust variance matrix is consistent. To determine the appropriate lag length, q, we posited a maintained value that should be larger than the optimal q. Here we use $q = 24$. We then performed sequential F tests on the last $24 > p$ coefficients, stopping when the test rejects the H_0 that the coefficients are jointly zero at a 5% level. Using a robust variance matrix to calculate the F statistics, we drop 18 lagged explanatory variables and assign $q = 6$.

The estimation results are reported in Table 1, with both robust standard errors and the usual FD standard errors. Using robust standard errors, we find that all coefficients of wind installed capacity are close to zero and not statistically significant.[18] A joint F-test on $H_0 : \delta_k = 0$ for $k = 0, 1, \ldots, 6$ gives $F(7, 31734) = 0.78$ with p-value $= 0.6001$. Thus, we cannot reject the hypothesis hat the impact of wind activity on employment is not statistically significantly different from zero.

(a) (b)

Figure 7: FD estimation results with $q = 6$: (a) wells (b) wind capacity

Next, we test for the presence of serial correlation in Δu_{it} using the Breusch-Godfrey and Wooldridge tests for serial correlation in panels. Both tests accept the H_0 (p-values equal 0.3656 and 0.3096, respectively). Thus, we cannot reject that no serial correlation remains in the idiosyncratic errors. We then perform OLS and obtain the FD estimator. The estimation results indicate that six out of seven coefficients of the wind installed capacity are negative and all are insignificant. A joint F test on $H_0 : \delta_k = 0$ for $k = 0, 1, \ldots, 6$ gives $F(7, 31734) = 0.78$, with p-value $= 0.6001$. Hence, we cannot reject H_0, and the impact of wind activity on employment is not significant from zero. However, since there is substantial correlation in *wells* at different lags (multicollinearity), it can be difficult to obtain precise estimates for the individual βs. The estimated long run multiplier is 4.4. On the other hand, we found $wells_t$, $wells_{t-1}$,... and $wells_{t-6}$ to be jointly significant: the F statistic has a p-

[18]The order 2 lag is negative and significant at a 5% level.

value equal to 0.0007. Adding the estimated coefficients of the current and lagged variables, we obtain long-term multipliers $LRP_{wells} = 77.46$. Assuming that all the jobs created are short-term (they only last for 1 month), we divide LRP_{wells} by 12 to obtain the number of full-time equivalent (FTE) jobs: 6.42.[19] Given that 5482 new directional/fractured wells were drilled in Texas in 2011, about 35000 FTE jobs would have been created.

We graph the estimated short-run impact of $wells_k$ and $wcap_k$ as a function of k in Figure 7. The lag distribution summarizes the dynamic effects from a temporary increase in the explanatory variables on the dependent variable. Figure 7a shows a mainly declining trend in the impact of wells on the first three months. This is expected, as workers generally leave after the well completion. Employment then increases starting with month 4. We interpret this as the result of the emergence of new business opportunities in the area, resulting from the well-drilling activity. We find that the largest effect is with the first and the last lag. Figure 7b shows the impact from the added new wind capacity. The employment effect is negative at first, and then increases. It peaks about four months after the wind farm construction, and then declines again. After six months, the impact from both well-drilling and wind activities fades and employment falls back to its original level.

Assuming that all jobs created by well-drilling last for one month, we sum up the 6 coefficients of wells and divide by 12 to get the full-time equivalent growth rate: 0.01465%. Given that 3889 new directional/fractured wells were drilled in Texas in 2011, the results imply that about 54446 jobs would have had been created.[20]

[19] This allows us to avoid the double-counting problem.

[20] The total employment in Texas in 2010 was 10,182,150. Note that $R^2 = 0.00084$. Since oil and gas-related employment is only 2.6% of the total employment in Texas, a low explanatory power of the regression model is perhaps to be expected.

5.2 The ADL model

Since the ADL model involves lagged dependent variables, the strict exogeneity assumption is violated and both the FE and the FD estimators are inconsistent. To overcome this we employ the generalized method of moments (GMM).[21] We again need to assign appropriate p and q to the model before we estimate it. As before, we start with p and q that large enough to be guaranteed to be larger than their optimal value: $p = q = 24$. When we include one lagged dependent variable, $E_{i,t-1}$, Wooldridge's test for serial correlation gives $\chi^2 = 30.189$, with p-value $= 3.919e{-8}$. The strong serial correlation implies that the dynamic data generation processes has not been fully captured. When we include one more lagged dependent variable, $E_{i,t-2}$, the test result changes to $\chi^2 = 0.0081$, with p-value $= 0.9285$. We conclude that the error term, u_{it}, is now serially uncorrelated. Henceforth, we set $p = 2$.

As in the previous section we proceed by setting $q = 6$. We then proceed to estimate the two-way Arellano-Bond GMM regression. The full results are shown in Table 2 in Appendix A. Both the Wald and the joint F tests cannot reject that the coefficients for wind capacity $\delta_0 = \ldots = \delta_6 = 0$, which implies no impact of wind activity on local employment. Thus, in the following discussion we will focus solely on the *wells* variable.

Figure 8 graphs the resulting dynamic response of employment to a unit increase in $wells_{it}$ and $wcap_{it}$ under six lags. The employment variable, E_{it}, initially increases by β_0 in period 1, then by $\beta_1 + \lambda_1\beta_0$ in period 2, by $\beta_2 + \lambda_1(\beta_1 + \lambda_1\beta_0) + \lambda_2\beta_0$ in period 3,... etc. The corresponding graphs in Figure 8 and Figure 7 exhibit similar trends. In addition, the estimation results in Table 2 are in line with that in Table 1.[22]

Note that the sum of the two estimated coefficients for the lagged dependent variables

[21] If we maintain the contemporaneous exogeneity assumption, the FE estimator's inconsistency shrinks to zero at the rate $1/T$, while the inconsistency of the FD estimator is essentially independent of T (Wooldridge, 2002).

[22] For comparison, we have also included the pattern estimated in the previous Section. The long-run effect, referred to as the long-run propensity (LRP) or long-run multiplier, is given by $\frac{\sum 1_{k=0}^{19}\beta_k}{1-\lambda_1-\lambda_2} = 121.13$, which is very close to the one estimated from the FDL model using FGLS. Full results appear in the Appendix.

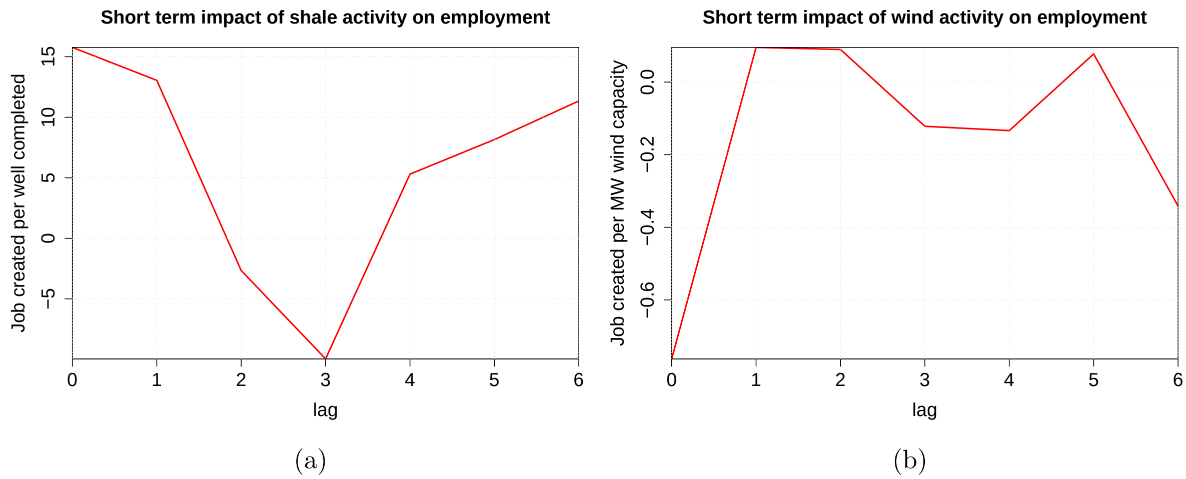


Figure 8: GMM estimation results with $p = 2$, $q = 6$: (a) wells (b) wind capacity

is 0.98. Although the test $\lambda_1 + \lambda_2 = 1$ is rejected at the 1% level, employment might still follow a unit root process. To address this issue, we ran the same estimation using data from only the 98 counties with stationary employment series. We obtained similar estimates: $\hat{\lambda}_1 + \hat{\lambda}_2 = 0.98$. We believe that the large persistence in employment is probably due to the small explanatory power of well drilling, which is reasonable since employment in the shale gas sector is a rather small component of total employment.

5.3 Spatial Panel Models

Earlier we discussed the theory behind the spatial autoregression (SAR) and the spatial error model (SEM). In the SAR model, the inclusion of the dependent variable introduces simultaneity bias and the OLS estimator is no longer unbiased and consistent, while in the SEM, the OLS estimator is unbiased, but inefficient. Therefore, maximum likelihood estimation is used to estimate the parameters of both models.

Both the SAR and SEM models are estimated allowing for two-way fixed effects. The results are reported in Table 3 in Appendix A. We find that the spatial interaction coefficients of both models are significant and very similar: $\rho = 0.1730$, $\lambda = 0.1734$. Also, both models show large and significant coefficients for $wells$, and insignificant coefficients for $wcap$.

Following LeSage & Pace (2009), the expectation of the SAR model $y = \rho Wy + X\beta + \epsilon$ is

$$E(y) = (I_N - \rho W)^{-1} X\beta \tag{14}$$

Employment in county i depends on developments in neighboring counties, as workers in bordering counties migrate to take advantage of new job opportunities in shale and wind. This provides a motivation for the spatial lag variable Wy.

The own- and cross-partial derivatives in the SAR model take the form of an $N \times N$ matrix that can be expressed as:

$$\partial y / \partial x'_r = (I_N - \rho W)^{-1} I_N \beta_r \tag{15}$$

These partial derivatives measure how drilling/wind activities in county j influence employment in county i. For the rth explanatory variable, the average of the main diagonal elements of this matrix is labeled as the "direct effect," while the average of the cumulative off-diagonal elements over all observations corresponds to the "indirect effect." The average total effect will be the sum of the two.

This model implies that direct effect of well-drilling activity on employment is 225, and it is significant at the 0.000 level. The direct effect measures how wells drilled in a particular county affect employment in that same county. The result shows that about 225 jobs would be created per well drilled in the same county. The indirect effect estimate of well drilling activity is 46 which makes the total effect grow to 271. Thus, if we only consider the direct effect, the results would be underestimated by 17%. The result is significantly higher compared to the FD estimation results, $LRP = 77$, indicating that spatial correlation effects are important.

The direct and indirect effects of wind activity are 0.05 and 0.01, respectively, and they are not statistically significant. Hence, wind farm installation and construction was not found to have an impact on employment.

5.4 Wage Effects

Here we report on the impact from shale gas and wind developments on average weekly wages. We again employ the FD approach. Sequential F-tests determine that $q = 12$. The results appear in Table 5 in Appendix A. The results show that the coefficients of the 4^{th} and 9^{th} lagged wells are significant at the 0.05 level. For wind capacity, the coefficients of lag 1 and lag 10 are significant at the 0.05 level, while lags 11 and 12 are significant at the 0.01 level. Figure 9 graphs the resulting dynamic response of wages to a unit increase in $wells_{it}$ and $wcap_{it}$ (12 lags). The impact from wells drilled rises and falls with a 6 month cycle. The peaked value is about 0.3. The impact from wind capacity installation shows a quite different trend: it increases over time from near zero to 0.13.

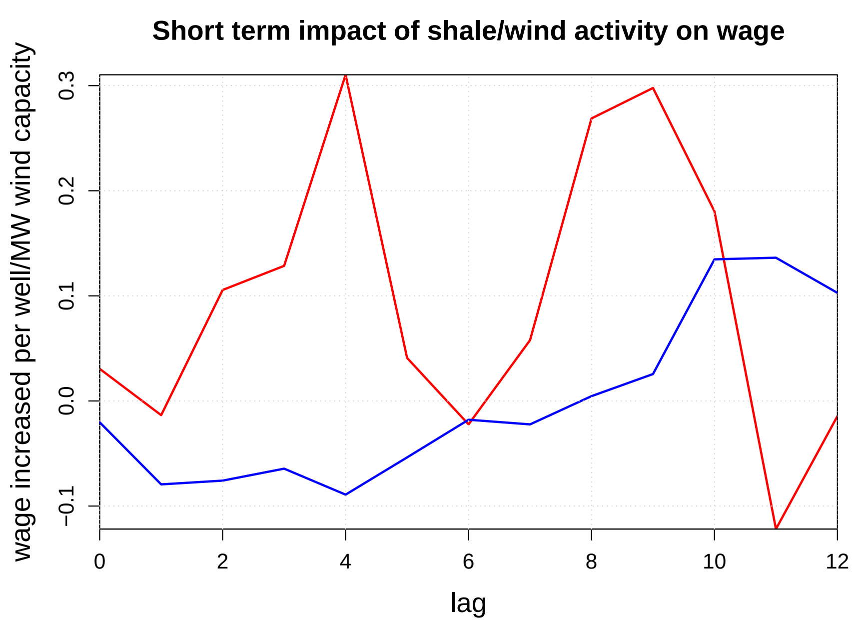


Figure 9: Short run impact of shale/wind activity on wage

The spatial panel regression results for wages are shown in Table 4 in Appendix A. The results are in line with those without including spatial interaction effects. According to the SAR model, the estimate of β_1 is 0.18 and significant at the 0.05 level, while the estimate of β_2 is 0.06 and not significant. Additionally, the results show a strong spatial correlation: $\rho = \lambda = 0.26$. Interestingly, the spatial correlation effects in wages are even larger than in employment: 26% of the increase in average wages is due to indirect effects, while such effects are responsible for 17% of the change in employment.

The direct and indirect effects of well drilling activity on wages are 0.18 and 0.06, respectively. Alternatively, the total effect on wages is 23 cents per well drilled, of which 18 cents are due to drilling activity in the same county, while 6 cents are attributed to drilling activity in neighboring counties. The total effect from wind activity is 8 cents per MW, of which about 6 cents are due to the direct effect, while 2 cents are due to the indirect effect.

6 Conclusion

We followed an econometric approach to compare job creation in wind power versus that in the shale gas sector. We have discussed the advantages and disadvantages of a number of different models. We then estimated them using county level data in Texas from 2001 to 2011. The results were quite consistent. Both first-difference and GMM methods show that shale development and well-drilling activity have brought strong employment to Texas: 77 short-term jobs or 6.4 FTE jobs per well. Given that 5482 new directional/fractured wells were drilled in Texas in 2011, an estimated 35000 FTE jobs were created. In contrast, we did not find a corresponding effect on wages. The effect on wages corresponds to a 30-cent increase in month 4 and month 9 after each well completion.

All our estimations show that the impact from wind industry development on employment is not significantly different from zero. Its impact on wages increases gradually after construction and peaks about one year later. We found that 13 cents are added to wages in months 10 to 12 after construction.

While there are undoubtedly several advantages to expanding wind-enerdy installation, the effects on employment appear to be insignificant. In contrast, unconventional gas production appears to be a significant force behind employment growth in Texas. Further research is needed to confirm these effects in other states.

A Tables of Estimation Results

Variable	Coefficient	(Std. Err.)
$E_{i,t-1}$	0.88241	0.0056054***
$E_{i,t-2}$	0.1005518	0.0056067***
$wells_{it}$	15.86961	5.586163**
$wells_{i,t-1}$	-.8539942	5.714926
$wells_{i,t-2}$	-15.90933	5.853195**
$wells_{i,t-3}$	-8.917589	5.964613
$wells_{i,t-4}$	14.38985	5.922299 *
$wells_{i,t-5}$	4.476252	5.851388
$wells_{i,t-6}$	3.622863	5.794333
$wcap_{it}$	-9.00e-06	0.0000248
$wcap_{i,t-1}$	-0.0000184	0.0000243
$wcap_{i,t-2}$	0.0000239	0.0000237
$wcap_{i,t-3}$	0.0000107	0.0000236
$wcap_{i,t-4}$	0.0000242	0.0000237
$wcap_{i,t-5}$	-0.0000129	0.0000243
$wcap_{i,t-6}$	-9.63e-06	0.0000249
Signif. Code: *** 0, ** 0.01, * 0.05, · 0.1, 1.		

Table 2: GMM Estimation Results, $p = 2$, $q = 6$

SAR Coefficients:				
	Estimate	Std. Error	t-value	Pr($>\|t\|$)
ρ	0.1730	0.0081	21.43	$< 2e-16^{***}$
wells	224.72	12.99	17.29	$< 2e-16^{***}$
newcap	0.05	6.366	0.0079	0.9937
SEM Coefficients:				
λ	0.1734	0.0081	21.42	$< 2e-16^{***}$
wells	235.81	13.63	17.30	$< 2e-16^{***}$
newcap	0.47	6.374	0.704	0.4814
Significant Code: *** 0, ** 0.01, * 0.05, $\cdot$ 0.1, 1.				

Table 3: Spatial interaction effects on employment

SAR Coefficients:				
	Estimate	Std. Error	t-value	Pr($>\|t\|$)
ρ	0.26	0.01	33.88	$< 2e-16^{***}$
wells	0.18	0.09	2.04	0.0412^{*}
newcap	0.06	0.04	1.51	0.1319
SEM Coefficients:				
λ	0.26	0.01	33.87	$< 2e-16^{***}$
wells	0.12	0.09	1.27	0.20
newcap	0.07	0.04	1.60	0.11
Significant Code: *** 0, ** 0.01, * 0.05, $\cdot$ 0.1, 1.				

Table 4: Spatial interaction effects on wage

Variable	Coefficient	Robust SE.	t-value	$Pr(>\|t\|)$
$wells_t$	0.030579	0.112459	0.2719	0.785688
$wells_{t-1}$	-0.013512	0.149363	-0.0905	0.927917
$wells_{t-2}$	0.105605	0.160727	0.6570	0.511159
$wells_{t-3}$	0.128526	0.159170	0.8075	0.419399
$wells_{t-4}$	0.310342	0.132325	2.3453	0.019018 *
$wells_{t-5}$	0.040933	0.119525	0.3425	0.732003
$wells_{t-6}$	-0.022280	0.128208	-0.1738	0.862040
$wells_{t-7}$	0.057849	0.121926	0.4745	0.635176
$wells_{t-8}$	0.268792	0.137164	1.9596	0.050047 .
$wells_{t-9}$	0.297772	0.133959	2.2229	0.026232 *
$wells_{t-10}$	0.180401	0.108571	1.6616	0.096603 .
$wells_{t-11}$	-0.121860	0.155113	-0.7856	0.432095
$wells_{t-12}$	-0.014441	0.144573	-0.0999	0.920434
$wcap_t$	-0.020312	0.022022	-0.9224	0.356349
$wcap_{t-1}$	-0.079381	0.039836	-1.9927	0.046304 *
$wcap_{t-2}$	-0.075806	0.057576	-1.3166	0.187978
$wcap_{t-3}$	-0.064414	0.053879	-1.1955	0.231889
$wcap_{t-4}$	-0.089130	0.058455	-1.5248	0.127325
$wcap_{t-5}$	-0.053716	0.051733	-1.0383	0.299123
$wcap_{t-6}$	-0.017880	0.041650	-0.4293	0.667710
$wcap_{t-7}$	-0.022297	0.065829	-0.3387	0.734828
$wcap_{t-8}$	0.004571	0.056724	0.0806	0.935773
$wcap_{t-9}$	0.025598	0.058720	0.4359	0.662890
$wcap_{t-10}$	0.134719	0.058776	2.2921	0.021907 *
$wcap_{t-11}$	0.136272	0.042045	3.2411	0.001192 **
$wcap_{t-12}$	0.102829	0.035770	2.8748	0.004046 **
Significant Code: *** 0, ** 0.01, * 0.05, $\cdot$ 0.1, 1.				

Table 5: FD estimation results with robust se. on wage, $q = 12$

References

American-Chemistry-Council (2011), Shale gas and new petrochemicals investment: Benefits for the economy, jobs, and US manufacturing, Technical report, Economics & Statistics American Chemistry Council.

Arellano, M. (1987), 'Computing robust standard errors for within-groups estimators', *Oxford Bulletin of Economics and Statistics* **49**(4), 431–34.

Boef, S. D. & Keele, L. (2008), 'Taking time seriously', *American Journal of Political Science* **52**(1), pp. 184–200.

Chamberlain, G. (1982), 'Multivariate regression models for panel data', *Journal of Econometrics* **18**(1), 5–46.

Considine, T., Watson, R., Entler, R. & Sparks, J. (2009), An emerging giant: Prospects and economic impacts of developing the Marcellus shale natural gas play, Technical report, The Pennsylvania State University College of Earth & Mineral Sciences Department of Energy and Mineral Engineering.

EIA (2012), Annual energy outlook 2012 with projections to 2035, Technical report, U.S. Energy Information Administration (EIA).

Hadri, K. (2000), 'Testing for stationarity in heterogeneous panel data', *Econometrics Journal* **3**(2), 148–161.

Higginbotham, A., Pellillo, A., Gurley-Calvez, T. & Witt, T. S. (2009), The economic impact of the natural gas industry and the Marcellus shale development in West Virginia in 2009, Technical report, Bureau of Business and Economic Research College of Business and Economics West Virginia University.

IHS (2012), The economic and employment contributions of shale gas in the United States, Technical report, America's Natural Gas Alliance.

Im, K. S., Pesaran, M. & Shin, Y. (2003), 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics* **115**(1), 53 – 74.
URL: *http://www.sciencedirect.com/science/article/pii/S0304407603000927*

Kapoor, M., Kelejian, H. H. & Prucha, I. R. (2007), 'Panel data models with spatially correlated error components', *Journal of Econometrics* **140**(1), 97 – 130.
URL: *http://www.sciencedirect.com/science/article/pii/S0304407606002259*

Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* **54**(13), 159 – 178.
URL: *http://www.sciencedirect.com/science/article/pii/030440769290104Y*

LeSage, J. & Pace, R. K. (2009), *Introduction to Spatial Econometrics*, Taylor & Francis/CRC Press.

Pesaran, M. (2004), General diagnostic tests for cross section dependence in panels, Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

Scott&Associates (2009), The economic impact of the Haynesville shale on the Louisiana economy in 2009, Technical report, Loren C. Scott & Associates.

Slattery, M., Richards, B., Schwaller, E., Swofford, J. & Thompson, L. (2009), Job and economic development impact model (JEDI) validation Capricorn Ridge and Horse Hollow, Texas, Technical report, The Institute for Environmental Studies at Texas Christian University and NextEra Energy.

UTSA (2012), Economic impact of the Eagle Ford shale, Technical report, The University of Texas at San Antonio Institute for Economic Developments Center for Community and Business Research.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, number 0262232197 *in* 'MIT Press Books', The MIT Press.