Employment Impacts of Upstream Oil and Gas Investment in the United States*

Mark Agerton, Peter Hartley, Kenneth B. Medlock III, and Ted Temzelides

James A. Baker III Institute for Public Policy and Department of
Economics, Rice University, 6100 Main St MS-22, Houston, TX 77005

December 1, 2014

Abstract

Technological progress in the exploration and production of oil and gas during the 2000s has led to a boom in upstream investment and has increased the domestic supply of fossil fuels. It is unknown, however, how many jobs this boom has created. We use time-series methods at the national level and dynamic panel methods at the state level to understand how the increase in exploration and production activity has impacted employment. We find robust statistical support for the hypothesis that changes in drilling for oil and gas as captured by rig counts do, in fact, have an economically meaningful and positive impact on employment. The strongest impact is contemporaneous, though months later in the year also experience statistically and economically meaningful growth. Once dynamic effects are accounted for, we estimate that an additional rig count results in the creation of 31 jobs immediately and 315 jobs in the long run, though our robustness checks suggest that these multipliers could be bigger.

1 Introduction

The development of shale has created a new boom in the oil and gas industry.[1] There are also indications that it has led to economic revitalization in places like North Dakota, Texas, Alberta, West Pennsylvania, and Louisiana. This revolutionary increase in the production of

---

[1]See Figure A.1 in the Appendix for an Energy Information Agency (EIA) map of where unconventional oil and gas is located in the continental United States.

oil and natural gas has led to a discussion about its potential effects on employment. Anecdotal evidence about the employment experience of different states over the last recession suggests that such effects may indeed be present. Figure 1 shows the difference in cumulative employment growth of various states since January 2008 from the national average (in approximate percentage terms).[2] The base year roughly coincides with both a severe drop in employment due to recession and the beginning of the boom in shale gas and tight oil production. Casual inspection of the plots suggests that oil and gas states have generally grown much faster than the rest of the nation for the past several years. North Dakota especially has seen blistering employment growth (note the much larger scale of the y-axis). In fact, North Dakota, Texas, Alaska, and Louisiana had the highest cumulative employment growth from 2008 through 2013 out of all states—27.4, 7.86, 5.80, and 3.68 percent, respectively. In 2013, all were top producers of oil and gas.

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 dynamic general equilibrium 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 employment consequences of alternative government policies or other shocks. Input-Output (I/O) analysis, which is commonly used in economic impact studies, is also poorly suited to answer our research question. It cannot incorporate induced price changes or substitution between inputs in production, which means that multipliers are usually overstated. Furthermore, as Kinnaman (2011) points out, because these studies are calibrated with data on a region's existing industries, measuring or predicting the impact of a totally new industry in the region—for example, shale gas in Pennsylvania—is problematic.

Our approach is an empirical one. We first consider the interrelationships between real oil prices, national rig counts, the production of primary energy and employment in the oil and gas extraction industry in a Structural Vector Autoregression (SVAR) at the national

---

[2]Following Blanchard and Katz (1992) cumulative employment growth relative to national is calculated as $\log(Emp_{it}/Emp_{i,Jan2008}) - \log(NatlEmp_t/NatlEmp_{2008})$.
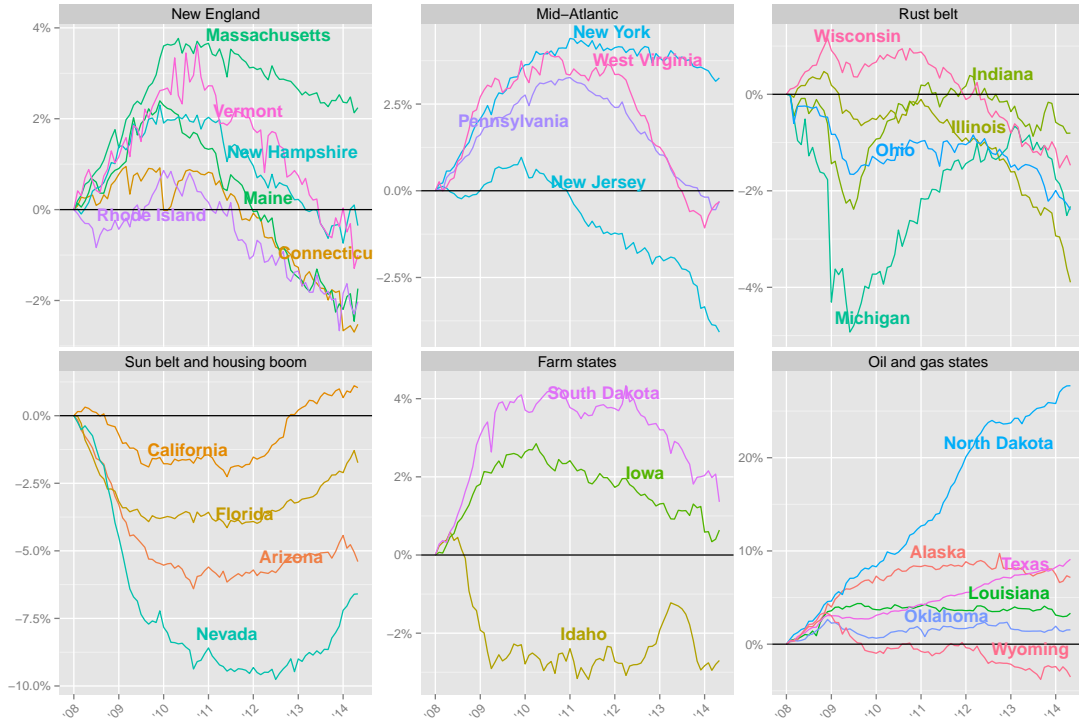
Figure 1: Cumulative Employment Growth (Percent) Relative to National

level. After 24 months, we estimate that a 10% increase in national rig counts results in an approximately 5% increase in employment in the oil and gas industry. Total changes in employment will depend on the extent to which job gains in oil and gas extraction (and related sectors) are offset by job losses in others as workers change industries. Because energy production is a relatively small portion of the economy, a national SVAR framework is not able to speak to a total employment multiplier in a statistically meaningful way.[3]

Our main results are obtained at the state level. The oil and gas boom varies widely over time and states, which we exploit to identify the total employment effect of upstream investment at the state level. Drilling activity varies significantly over time within states. Furthermore, this within-state variation is also quite different between states.[4] For example,

---

[3]For total employment and employment in manufacturing, our point estimates for the cumulative impact of rig-counts after 24 months is approximately 1/10 of the size (half of a percentage point increase in employment for each ten percentage increase in rig counts), and it is not statistically distinguishable from zero.

[4]See Figure A.2 and A.3 in the Appendix for plots of rig counts in per-capita and level terms, respectively.

North Dakota has seen continually increasing investment, while its neighbor, Wyoming, has seen an increase and subsequent decrease in drilling activity during our sample. Based on estimation of a dynamic panel model, we find that an additional rig count per million people, as measured by the working-age civilian non-institutional population, is associated with contemporaneous state-level employment growth of 0.010% and long-run growth of 0.064%. A per-capita specification suggests that a single rig count is associated with 31 new jobs in the same month and 315 jobs in the long run. Finally, we also carry out a range of sensitivity analyses that suggest our results are relatively robust to non-spherical errors, different lag-lengths and measures of population. Furthermore, they are neither driven by one state in particular nor are they driven by the pre- or post-2008 period only.

2   Literature Review

There are several industry related studies of the employment effects of upstream oil and gas development (Considine et al., 2009; Considine, Watson, and Blumsack, 2010; Higginbotham et al., 2010; IHS Global Insight, 2011; Murray and Ooms, 2008; Scott, 2009; Swift, Moore, and Sanchez, 2011). Most of them use the Input-Output (I/O) methodology and predict large increases, not only in employment but also in tax revenues and economic output. Kinnaman (2011) provides a peer-reviewed survey and strong critique of a number of these. He points out that several assume, perhaps counter-factually, that nearly all windfall gains by households are spent locally and immediately, that most inputs are locally produced, and that royalties are accrued locally. By taking an empirical approach, we can infer employment multipliers while remaining agnostic about the parameters underlying consumer and firm behavior. It is important to note that since we examine total employment, our analysis allows for effects of oil and gas development on other industries.

Our analysis gives rise to lower multipliers than these I/O studies. For example, Considine et al. (2009) estimate that upstream investment in the Marcellus shale created 29,284 jobs in

Pennsylvania in 2008, and their follow-up study estimates that 44,098 jobs were created in 2009 (Considine, Watson, and Blumsack, 2010). We assert that the increase in Pennsylvania rig counts in 2008 and 2009 should have led to a long-run increase of 1,812 and 11,930 jobs, respectively, after accounting for dynamic multiplier effects, an ultimate increase of 1,288 and 8,475 jobs (see Figures 2 and 3 for the cumulative impacts of rig-count changes on employment).[5] The IHS Global Insight (2011) study finds that the shale gas industry "supported over 600,000 jobs" in 2010 once direct, indirect, and induced jobs were summed. Using our multipliers, we find that national changes in rig counts in 2010 should lead to 173,794 jobs when dynamic effects are taken into account.[6] Note that 2010 happens to be the largest January to January increase in rig counts, so the estimates of job growth for 2009 were negative and had larger magnitudes.

One empirical approach, which we elect not to take, is the treatment-effect design, which is common in labor economics. The first paper to use this design in relation to the employment impacts of a North American resource boom was Black, McKinnish, and Sanders (2005), which examined the effect of coal booms and busts on county employment in Appalachia. The authors use the presence of large coal reserves as their treatment indicator variable and find modest multiplier effects: for every ten jobs created in the mining sector, 1.74 jobs in the local-goods sector are created, and for a bust, 3.5 local jobs are lost for each ten mining jobs. Weber (2012) uses a similar study design and estimates a triple-difference specification where the treatment is the value of gas production and the instrument used is the percentage of a county covered by shale deposits. He finds that an additional million dollars of natural gas production (resulting from increased price or production) leads to 2.35 additional jobs in the county. In a follow-up study, Weber (2014) examines the possibility that the unconventional gas boom might lead to a resource curse in rural counties. He finds

---

[5]Figures 2 and 3 use the estimated cumulative impulse response function from the per-capita dynamic panel model (2) and rig-count changes to calculate the employment impact over time.

[6]Applying our state-level multipliers to the national level is problematic since job gains in oil and gas producing states may be offset by losses in other non-producing states, but the calculation is illustrative of the different magnitudes obtained in I/O studies versus our empirical model.
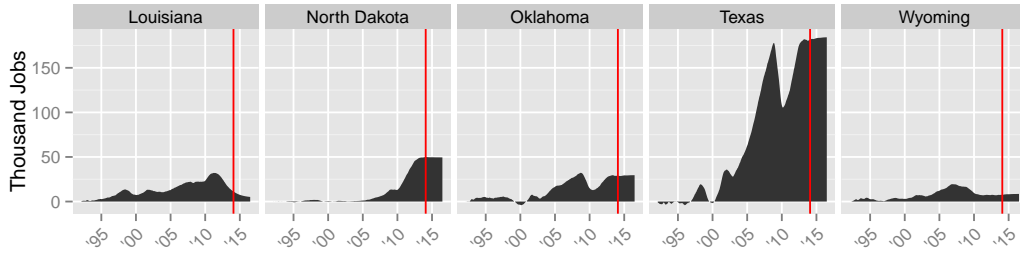
(a) Oil and gas states
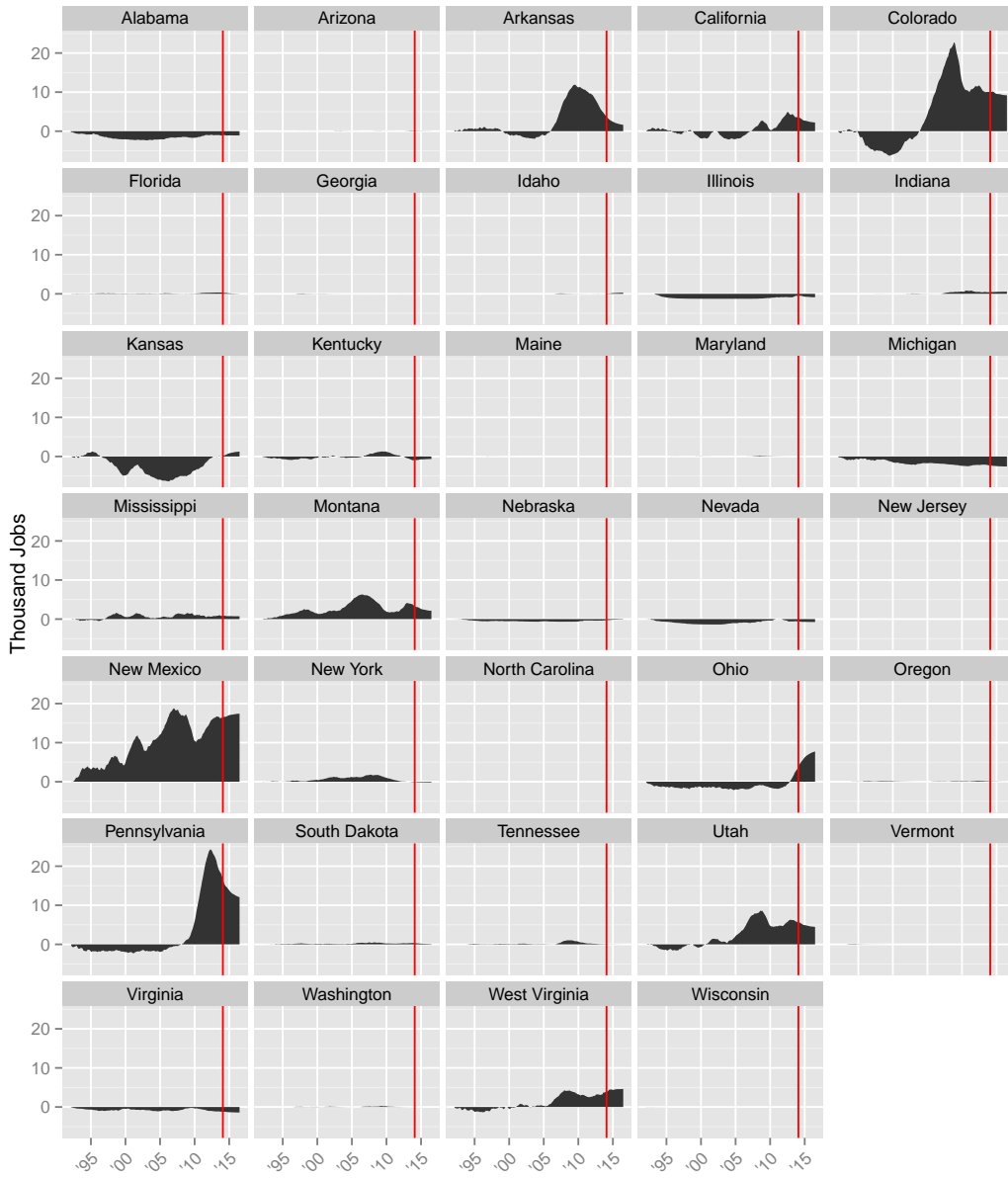
(b) Other states

Figure 2: Estimated cumulative job creation due to rig-count changes
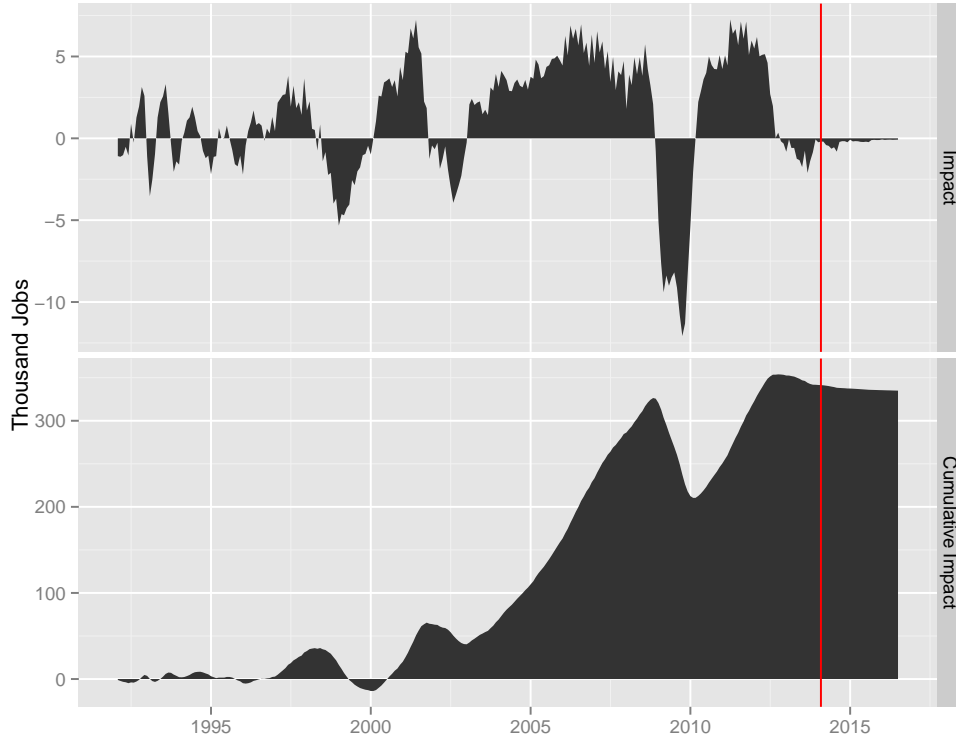
Figure 3: Estimated job creation due to rig-count changes at national level

that an additional 22 billion cubic feet (bcf) of gas production creates 18.5 jobs in each county (7.5 mining jobs and 11.5 non-mining jobs—a multiplier of 1.4 non-mining jobs per mining job), but increased gas production does not lead to crowding out in manufacturing or lower educational levels, which are outcomes associated with a Resource Curse. Other studies which follow a similar design include a manuscript by Fetzer (2014), who concludes that the natural gas boom does not lead to Dutch Disease and crowding-out of traded goods, possibly because of the salutatory effects of lower energy prices, and Marchand (2012), who examines the effect of energy booms on Canadian labor market outcomes and also finds modest multiplier effects in the non-trade sector and no crowding out of manufacturing. Our study is different because we use higher-frequency data and more fully exploit the time-series variation in employment and upstream investment. This allows us to understand the dynamics of job creation, though the use of higher-frequency data may limit our ability to account for lower-frequency, structural shifts in the economy.

7

A recent Baker Institute study (Hartley et al., 2013) considered a similar question to ours using county-level data from Texas. Before the authors allow for cross-county spillovers, they find that each well-count results in 77 short-term jobs, with within-county employment impacts occurring primarily in months zero, one, five, and six. They find some evidence that by allowing for county spillovers using a spatial auto-regressive model, the long-run employment effects are almost three times as large. Our research design is a similar, dynamic panel model, and we use a longer dataset (1990 to 2014) for the entire nation. Additionally, where they focus on unconventional oil and gas production, we focus on all oil and gas activity. We find a long-run multiplier that is approximately twice as big as in the basic specification considered by Hartley et al. (2013), but still below the effects the authors estimate after allowing for cross-county spillovers.

Hooker and Knetter (1997) investigate the employment impact of changes in military spending using a panel of state-level data from 1936–1994. The authors find that exogenous spending shocks (government procurement spending per capita) do lead to changes in employment growth, but the effects of shocks are nonlinear: large, adverse shocks lower employment growth more than smaller ones, while positive shocks also are less effective at increasing employment growth than negative ones at decreasing it. While we use higher-frequency data, we adopt their general model specification, including a set of state-specific intercept terms and time fixed effects as Hooker and Knetter do.

Blanchard and Katz (1992) examine the dynamics of regional labor markets using a vector autoregression (VAR) approach with annual state labor-market data. They argue that labor-demand shocks are the primary drivers of changes in employment, unemployment and participation. Furthermore, they provide evidence that migration of workers—not firms—in response to low unemployment rates—not high wages—is the primary mechanism that returns markets to their long-run equilibria. One of the implications of this for us is that transforming exogenous variables into per-capita variables may be problematic if population movements are endogenous. We run a series of robustness checks related to this issue (see

Section 7.4).

Finally, Arora and Lieskovsky (2014) examine national economic impacts of the natural gas boom using a Structural Vector Autoregression (SVAR) framework. They find evidence that lower real natural gas prices due to increased supply positively impact industrial production. Furthermore, they find that this effect is greater post-2008 and conclude that the shale gas revolution has altered the relationship of natural gas to the macroeconomy. We estimate a SVAR as well. However, while Arora and Lieskovsky look at the supply side of the economy, we focus on job creation and do not distinguish between the positive supply-side effects of lower energy prices and positive demand-side effects of additional upstream investment.

3   Rig Counts

To capture upstream oil and gas investment we use the Baker Hughes rig counts, which are publicly available on the firm's website.[7] Baker Hughes is a major supplier of oil-field services and has been publishing reports on the industry since 1944. Each week, the firm surveys rotary rig operators in North America and publishes a count of the number of rigs that are "actively exploring for or developing oil or natural gas" in each state. Many firms and industry analysts use the reports to gauge investment activity in the sector. Weekly state-level data from Baker Hughes begin in 1990, and we average these figures to a monthly frequency. Unfortunately, the weekly (and monthly) data do not provide detail on whether rigs are engaged in unconventional activity or not, so we must use the total number of land-based rigs in each state. We choose to exclude offshore rigs in the total since the investment required for offshore wells is very different than onshore wells.

In our state-level model, we scale rig-counts by the US Census estimate of state population in the year 2000. First, state population endogenous to state employment as shown by

---

[7]Baker Hughes publishes rig counts at `http://phx.corporate-ir.net/phoenix.zhtml?c=79687&p=irol-reportsother`.
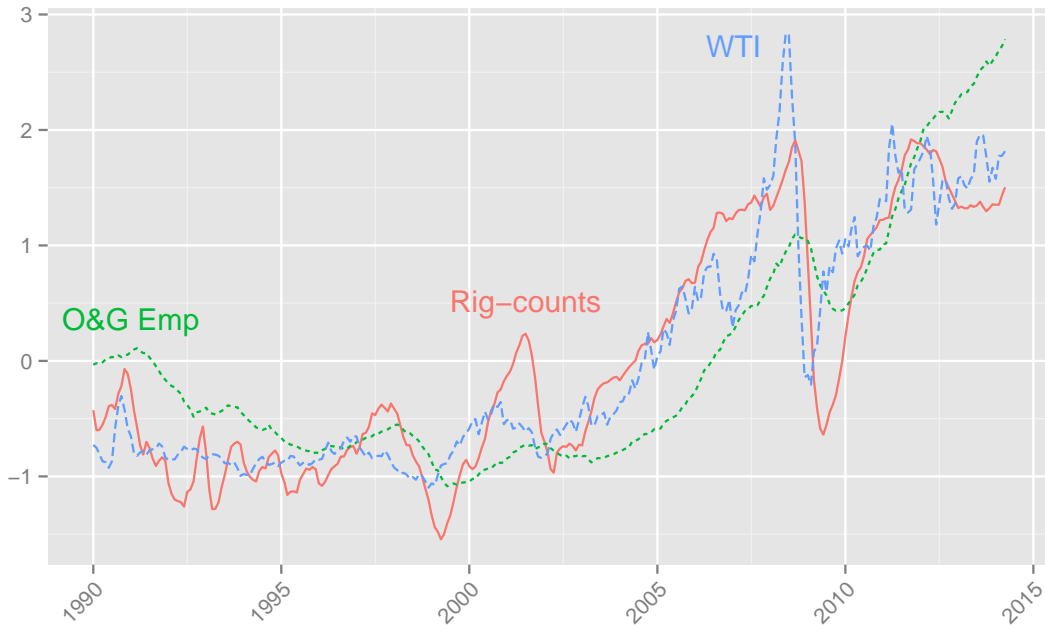
Figure 4: Labor, Capital, and Prices in Oil and Gas (Standardized Variables)

Blanchard and Katz (1992), so we eliminate this variation from our model. Second, since we take differences, if we scale rig-counts by time-varying population then differences will capture both rig-counts as well as migration. Thus, multipliers will be "polluted" by this extra variation. We check this by using a variety of constant and time-varying population measures. See Section 7.4 for details.

Figure 4 plots seasonally adjusted national employment in oil and gas extraction plus support activities along with the national average monthly rig count and the West Texas Intermediate crude benchmark. The three are clearly linked. Plots of the rig count data in per-capita and level terms are displayed in the Appendix (Figures A.2 and A.3, respectively).

Three points are worth making. First, in states with significant exploration and production, there is substantial within-state time series variation in rig counts, which can provide identification of employment impacts. Second, the magnitudes and pattern of variation are different between states. For example, the maximum number of rigs per month in Texas is 946, while many states have no drilling activity (Connecticut, for example). Third, the states

with the highest mean rig count levels (Texas at 488 and Louisiana at 156) are not those with the highest per-capita count (Wyoming and North Dakota are the two highest, with 105 and 68 rig counts per million people respectively). Thus, estimates will be sensitive to whether rig counts are in level or per-capita terms,[8] and they may be overly influenced by outliers like Texas and Wyoming since these have the largest rig-counts in level and per-capita terms, respectively. We address the issue of outliers in our robustness checks.

4 Vector Autoregression Model

State-level variation in rig counts and employment may provide precision in estimating the employment impacts of upstream activity, but these estimates cannot account for inter-state effects. On one hand, an energy boom in one state may induce higher labor demand and job creation in another as the boom state's income rises. On the other, the boom state may pull workers from nearby states, reducing net job creation at the national level. To address this ambiguity, we model the interaction between real oil price shocks, upstream investment, industrial production, and employment. Specifically, we examine three Structural Vector Autoregression (SVAR) models similar to those estimated by Arora and Lieskovsky (2014). Our three models focus on (1) total IP and employment, (2) IP and employment in manufacturing, and (3) IP in Primary Energy and employment in oil and gas extraction.

A vector autoregression (VAR) captures the interrelationships between variables that are jointly determined. Unfortunately, shocks to each variable may be correlated, so causation is unidentified without more structure. By imposing restrictions on the shocks, one can identify causal effects of shocks and their relative importance to the system. We follow notation from

---

[8]We did consider a log-log specification where we used $\log Employment_{i,t}$ and $\log(Rigs + 1)$. This specification means we don't need to include population measures. However, a number of states have very low rig-counts, so a one unit change in rig-counts is a very large percentage change, while a state like Texas (which has hundreds more rigs than most other state) may undergo a large change in levels but not percentages. This means that identifying variation comes from the states *without* much activity—precisely the wrong ones. Multipliers are much smaller in this model, but this is expected.

Lütkepohl (2005) and define our vector of variables as

$$y_t = \left[\log(\text{Real WTI}_t) \quad \log(\text{Rigs}_t) \quad \log(\text{IP}_t) \quad \log(\text{Employment}_t)\right]^\top$$

Since production and employment can both be highly seasonal, and since this seasonality can be deterministic and stochastic in nature, we include monthly dummy variables and specify our VAR in first differences as

$$A(L)\Delta y_t = \Delta y_t + \nu(t) + u_t$$

where $A(L) = I - A_1(L)^1 - \ldots - A_p(L)^p$ is a polynomial in the lag operator, $\nu(t)$ corresponds to a month-specific intercept that captures any deterministic seasonality, and Δ is the first-difference operator needed to make the variables stationary. The covariance of the error term $\Sigma_u = E[u_t u_t^\top$ may not be diagonal, so we cannot make causal inference about which shocks drive the system without further restrictions on the model. We choose to use the standard Cholesky decomposition of Σ_u to do this. This corresponds to the following restriction

$$u_t = P\varepsilon_t$$

where P is a lower-triangular matrix such that $\Sigma_u = PP^\top$ and implies a recursive system. We are primarily interested in the cumulative orthogonalized impulse responses (COIRFs), which show the cumulative impact of a one-standard deviation structural shock over time. We calculate the COIRF of a rig-count shock in period t on variables in period $t + h$ as

$$\sum_{j=0}^{h} \hat{\Phi}_j P^\top$$

where $\hat{\Phi}(L) = \hat{A}(L)^{-1}$ is the VMA representation of the system and $\hat{\Phi}_j$ is the coefficient matrix of the j-th order lag polynomial.

4.1 Data and Identification

The real domestic price of oil is calculated as the West Texas Intermediate (WTI) spot price scaled by the Producer Price Index (PPI) for all commodities. Oil is globally traded, and its price is set in liquid, international markets. Thus, oil price shocks in time t should be exogenous to the other variables in the model. Rig counts capture upstream investment activity. Exploration and production firms make decisions based on the expected future profitability of a well, which is determined by drilling costs and expected future oil prices. Drilling changes in time t should respond to oil prices insomuch as they are an indicator of future oil prices; however, since oil and gas production lags drilling activity, drilling changes in t should not drive oil prices in t. By the same token, drilling decisions at time t should not be affected by other contemporaneous economic shocks to industrial production or employment. Three different measures of Industrial Production (not seasonally adjusted) are taken from the Federal Reserve: Total IP, Manufacturing IP (as defined by SIC codes), and IP in Primary Energy. It is generally accepted that employment is a lagging macroeconomic indicator; thus, we order IP before employment. Employment is taken from the BLS Current Establishment Survey and is not seasonally adjusted. Specifically, we use total private non-farm employment, total employment in manufacturing, and the sum of employment in oil and gas extraction plus oil and gas support activities. For more details on data, see the Appendix.

4.2 Estimation and Results

We estimate the three versions of the above model. Lag selection was done using the minimum of the AIC and checking ACFs of residuals to confirm that they are white noise. The three models (of total, manufacturing and oil & gas activity) have twelve, twelve and three lags, respectively. Since both IP and employment are very seasonal, it makes intuitive sense that 12 lags would be required. Fossil-fuel production, however, is generally not seasonal (though its consumption is), so the fact that only three lags are required is not surprising,

either.

Figure 5a displays COIRFs from rig counts to employment for the three models. Since all variables are in logs, the vertical axis represents the approximate percentage change in the response variable with respect to a one standard-deviation shock in the impulse variable. COIRFs from the Total and Manufacturing models do not appear to be statistically significant at almost any lag. Given that these two models leave the majority of the economy unmodeled (in our sample, employment in oil and gas extraction is, on average, just 0.28% of total employment), this finding is not surprising. Our Oil and Gas model, on the other hand, shows very positive and significant cumulative employment impacts of additional rig counts. After re-scaling the COIRF(24) by standard deviation of $u_{Rigs,t}$ (which is the square-root of the second element of the diagonal of $\hat{\Sigma}_u$ and represents the magnitude of a one-standard deviation rig count shock), we calculate that a 1.0% increase in rig counts leads to approximately 0.5% increase in employment in oil and gas extraction.[9] Figure 5b, which displays the Forecast Error Variance Decompositions for the impulse-response of rig counts on employment, shows that rig count changes do not lead to major changes in employment at the aggregate but do have significant impacts (both statistically and economically) on employment in the oil and gas sector.

The set of COIRFs for the Oil and Gas VAR is in Figure 6 and the FEVD, in Figure 7. A few causal relationships appear to be important. First, higher oil prices increase drilling activity and employment in extraction, as expected, but, as shown by the FEVD, oil shocks are less important than rig counts in generating oil and gas jobs. Oil prices do not appear to have a strong impact on IP in Primary Energy, which suggests that energy production at monthly frequencies is fairly price inelastic—a result theory predicts. Rig counts also have an impact on employment and IP in Primary Energy and are associated with decreases in oil prices, again, as expected. Employment shocks in oil and gas extraction do not seem

---

[9]A similar calculation for the total and manufacturing employment models suggests elasticities that are about 1/10 the size. This could represent a substantial number of jobs, but the estimates are not at all statistically meaningful.
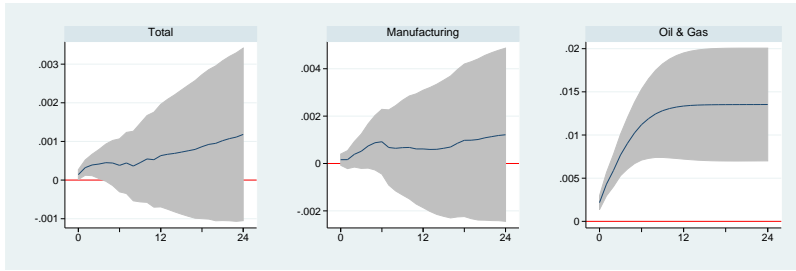
(a) COIRF of rig-count shocks
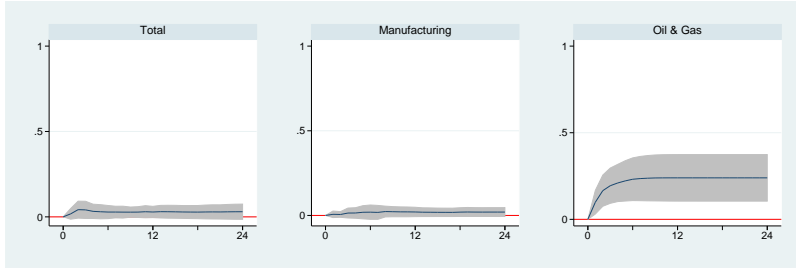
(b) FEVD of rig-count shocks

Figure 5: Employment responses in SVARs (95% CI) to rig-counts
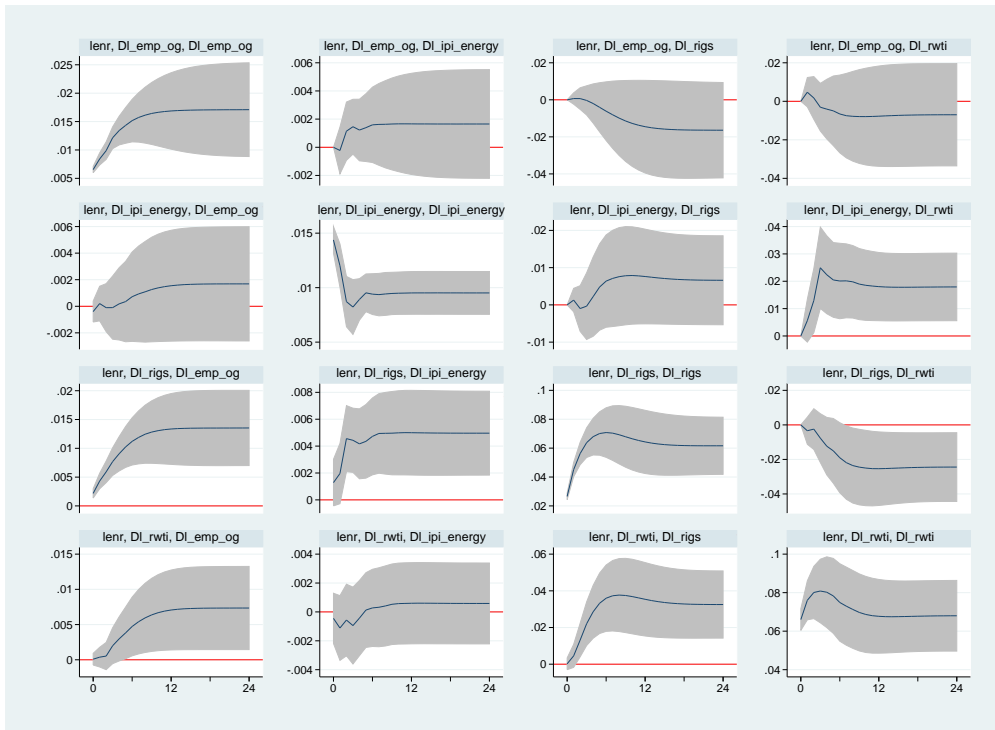
Figure 6: COIRFs for One-Standard Deviation Shocks in Oil and Gas SVAR (95% CI)
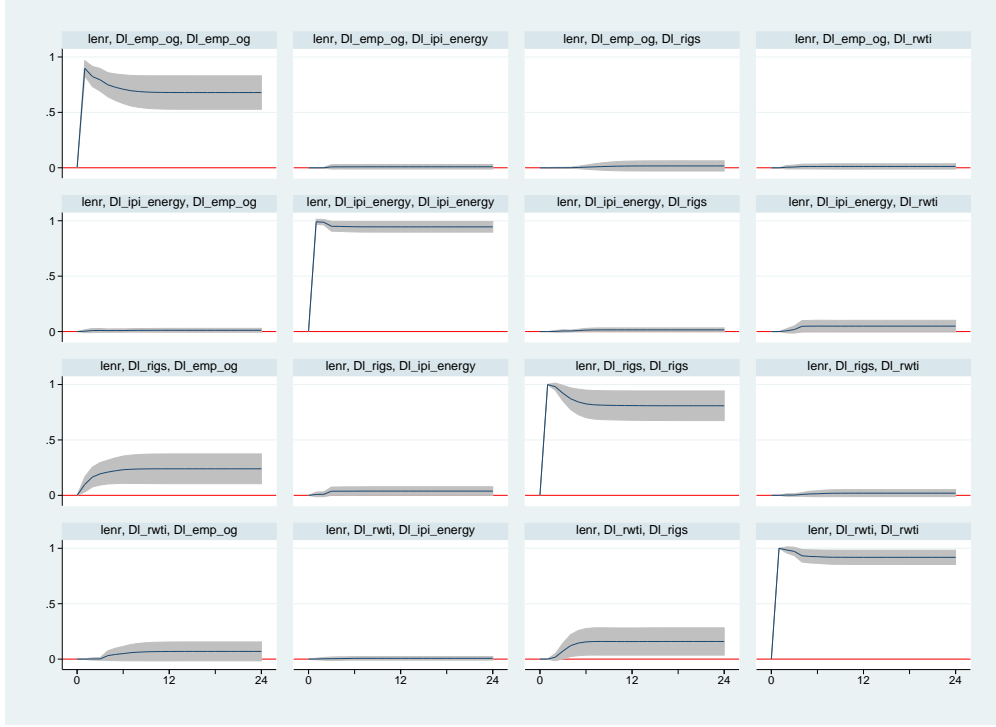
Figure 7: FEVD for Oil and Gas SVAR

to cause changes in other variables. Since labor demand in extraction is driven by drilling decisions, not the converse, this is to be expected. Finally a shock to IP in Primary Energy seems to prompt a positive, statistically significant response in oil prices but not in the other variables.

5 State Dynamic Panel Model

At the state level, our goal is to understand the response of private non-farm employment growth to changes in upstream drilling activity. Our primary model—which is least affected by outliers or choice of population variable—is

$$\Delta \log(Emp_{it}) = \alpha(L)\Delta \log(Emp_{it}) + \beta(L)\Delta(Rigs/Pop)_{it} + \nu_i(t) + \nu_t + u_{it}, \qquad (1)$$

though we also estimate a secondary model in per-capita terms

$$\Delta(Emp/Pop)_{it} = \alpha(L)\Delta(Emp/Pop)_{it} + \beta(L)\Delta(Rigs/Pop)_{it} + \nu_i(t) + \nu_t + u_{it} \qquad (2)$$

since estimates from the latter can be interpreted as "jobs per rig count."[10] The difference operator is Δ, and $\alpha(L)$ and $\beta(L)$ are polynomials in the lag operator. We allow states to have different equilibrium growth rates and individual, deterministic seasonality, which is captured by state-month fixed effects, $\nu_i(t)$. We do not try to model the impact of national macroeconomic shocks on state employment; rather we simply theorize that national shocks impact state employment growth rates uniformly across states and capture this with a time fixed effect, ν_t. It is very important that the time fixed effect controls for oil prices, which are common across states.[11] Employment does not adjust immediately back to equilibrium after a shock, and ACFs of state employment still show autocorrelation and seasonality after state-month intercepts are removed. Therefore, we include lags one through twelve of state employment.[12] Similarly, employment may take time to adjustment to additional upstream investment. Additional processing and transportation infrastructure may be needed and constructed after initial production, and spillovers into housing, entertainment, and retailing may take time. Finally, labor supply may be slow to respond, as appears to have been the case with North Dakota and its labor shortages. Thus, we include the contemporaneous change in rig counts per-capita plus ten lags.[13] Omitted variable bias is a valid concern for

[10]See the Robustness Checks section for a further discussion on the precise interpretation of multipliers.

[11]We also estimated a model which included lags 0–12 of industrial production and lags 1–13 of national employment in lieu of time fixed effects. The results were largely the same, though the savings on degrees of freedom did not provide substantially improved precision. These results are in Tables 9 and 10, along with three lag-length robustness checks similar to those in Tables 3 and 4.

[12]The BIC has a minimum with the inclusion of 24 lags and a local minimum at 13 lags. (See Figure A.4 in the Appendix.) However, Han, Phillips, and Sul (2013) show that the BIC is an inconsistent estimator in dynamic panels with fixed effects and overestimates lag-length. Since the majority of the improvement in both the BIC and residual ACFs is achieved with 12 lags, we use 12 lags of employment. Furthermore, intuition suggests that 12 lags is an appropriate number of lags for a monthly series with substantial seasonality. Robustness checks for different lag-lengths are discussed later.

[13]We pared the model back to 12 lags of both employment and rig counts using the BIC, inspection of ACFs, and intuition. Since lags 11 and 12 of rig counts were not jointly significant at the 10% level even with OLS standard errors, we eliminated them.

our model since we do not include time-varying, state-specific controls. For instance, we do not capture the housing bust, which did not affect oil states like Texas and North Dakota as severely as non-oil states like Florida and Nevada. If housing markets are systematically related to the oil and gas sector, this would bias our results. We are comfortable taking this risk.

Rig counts are scaled by the population in the 2000 Census[14] to account for the idea that an additional rig count in Texas (with a 2000 population of almost 20 million) is a proportionally much smaller shock to the economy than in North Dakota (with a population of 0.64 million in 2000). Log employment and rig counts per capita both have unit roots, and for some states with a very large share of employment in the upstream sector, the two might be cointegrated. However, some states have zero upstream activity, so the two cannot be cointegrated for all states. Since we cannot capture each state's underlying labor-market drivers (which would be cointegrated with total employment), we work in differences.[15]

The primary estimate of interest is the long-run multiplier (LRM), which is the employment impact of upstream activity once state employment returns to equilibrium. Long-run multipliers from a VARX process are $LRM = \beta(1)/[1 - \alpha(1)]$, and they are guaranteed to be positive if $\beta(1) > 0$.[16] The impulse response function (IRF) and cumulative impulse response function (CIRF) trace out the dynamic impact of drilling activity over time and are also of interest.[17] In model (1), we interpret $\beta_0 \times 100$ and $LRM \times 100$ to be the approximate percentage change in employment (contemporaneously and long-run, respectively) to an increase of one rig count per million people. In model (2), we interpret β_0 and LRM as jobs created per rig count.

---

[14]See Section 7.4 for robustness checks about the choice of population variables.

[15]One limitation of this is that we do not allow the possibility that upstream investment can cause permanently higher growth rates. However, monthly data may not be well suited to assessing questions of endogenous growth.

[16]In time-series notation, replacing the lag operator L by 1 gives $\beta(1) = \beta_0 + \beta_1 + \ldots + \beta_s$.

[17]Formulas for calculating the IRF, CIRF, and confidence intervals for both in the multivariate VARX context are available in Lütkepohl (2005, Chapter 10). Note that what we call (Cumulative) Impulse Response Functions, Lütkepohl calls (Cumulative) Dynamic Multipliers. Our single equation context is just a special case.

While our baseline estimation method is ordinary least squares (OLS), the error term may suffer from heteroskedasticity and autocorrelation, as well as correlation between state employment shocks.[18] Therefore, in addition to reporting OLS point estimates and t-statistics, we report t-statistics and confidence intervals for IRFs and CIRFS using multi-way clustered standard errors and Driscoll and Kraay (1998) serial correlation and cross-sectional dependence-consistent (SCC) standard errors as well as OLS standard errors. The former, detailed in Cameron, Gelbach, and Miller (2011), allow for errors to be serially correlated within a state as well as contemporaneously across states.[19] Driscoll-Kraay standard errors are essentially a Newey-West estimator applied to cross-sectional averages of the model's moments (e.g., the cross-product of $\bar{x}_t \bar{u}_t$, not $x_{it} u_{it}$, where $\bar{x}_t$ denotes $\bar{x}_t = n^{-1} \sum_{i=1}^{n} x_{it}$).

6   State Panel Results

Coefficients and the three standard errors for model (1) are displayed in Table 1, and results for model (2) are in Table 2. The OLS standard errors are tightest, and Driscoll-Kraay, the largest. Even with the larger standard errors, however, t-tests that $\beta(1)$ and the LRM are less than zero are strongly rejected, and the contemporaneous rig count is highly significant. Thus, we conclude that upstream investment does, in fact, create a statistically significant, positive number of jobs. We estimate that an additional rig per million people increases a state's employment by 0.010% in the same month and 0.064% in the long run, though standard error bounds imply some uncertainty about this impact. If the model is estimated in per-capita terms, the immediate impact multiplier $\hat{\beta}_0$ indicates that one additional rig count is associated with 31 additional jobs in the same month, and the LRM, 315 jobs in the long run.

---

[18]We ignore the issue of lagged endogenous regressors since estimates are still consistent as $T \longrightarrow \infty$ and we have 265 observations per state. Judson and Owen (1999) perform a Monte Carlo study on the relative performance of LSDV versus a corrected estimator and GMM-based estimators like the Arellano-Bond approach, and they find that the LSDV estimator performs well.

[19]Point estimates and multi-way clustered standard errors were calculated using the **felm** function from **R** package **lfe**. With multi-way clustered standard errors the variance matrix is not guaranteed to be positive semi-definite. We adopt the standard time-series approach of simply setting negative eigenvalues to zero.

Figure 8a shows the IRF and CIRF for model (1), along with 95% confidence intervals for OLS, Multi-Clustered, and Driscoll-Kraay serial and cross-sectional correlation consistent (SCC) errors. Figure 8b shows the same for model (2). While the individual impulse responses themselves may not all be statistically significant at the 5% level, their sum is. The biggest employment increase is coincident with an increase in rig counts, but months $t+1$ and $t+8$ through $t+10$ also see sizeable impacts. There are many possible explanations for the delayed impact: subsequent infrastructure build-out; delayed labor supply responses; and delay in increased consumption by owners of labor, capital, and minerals are all reasonable possibilities.

Figures 2 and 3 plot the cumulative job-creation implied by the cumulative impulse response function from model (2). As mentioned earlier, these estimates are much smaller than those obtained via I/O models. Furthermore, as documented in the new EIA Drilling Productivity Report, shale gas and tight oil production per new well has increased dramatically since 2007.[20] Thus, seemingly lower rig counts do not correspond to lower production. Without more structure on our model or detail in our rig counts, we are unable to account for these changes in productivity. Thus, it seems reasonable to view years such as 2009—which corresponds to a drop in national rig counts and a resulting decrease of 182,934 jobs[21] in the long run—with some caution.

7 Four Robustness Checks

7.1 Lag Lengths

The first robustness check is to verify that multipliers do not vary greatly with lag-lengths. In addition to our base model, we estimate our baseline model with three different sets of lag-lengths. In the first set, we only include the contemporaneous and first lag of rig counts

---

[20]The EIA Drilling Productivity Report is available at `http://www.eia.gov/petroleum/drilling/`.
[21]Calculated as $\sum_{months} \sum_{states} \Delta rigs_{it} \times \widehat{LRM}$.
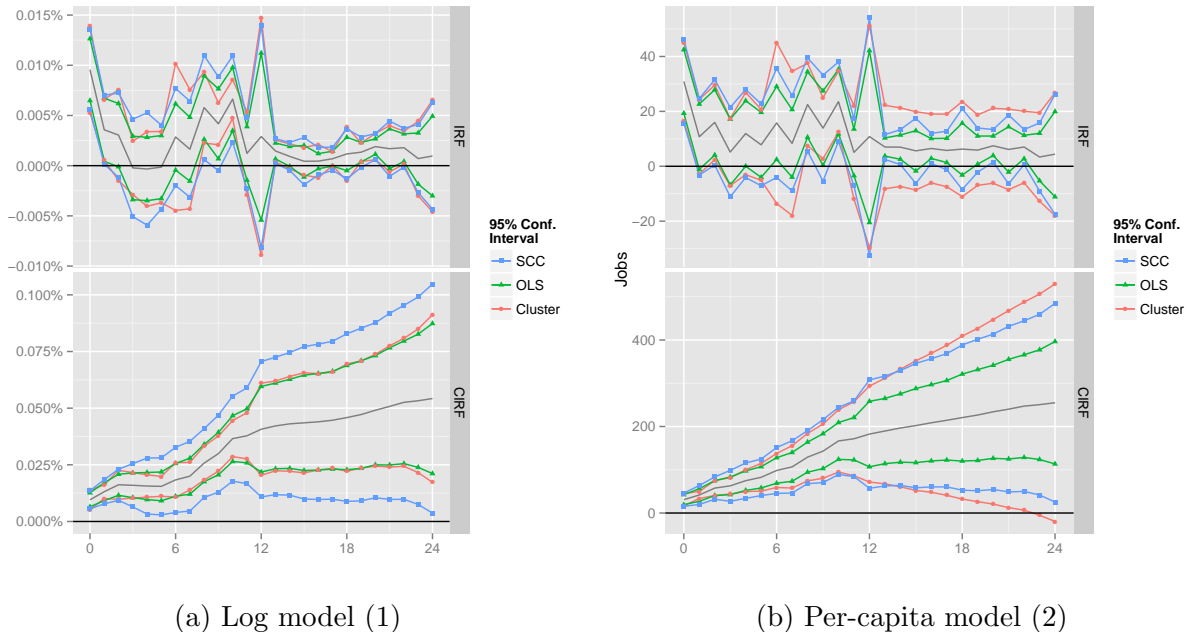
(a) Log model (1)

(b) Per-capita model (2)

Figure 8: IRFs and CIRFs for base models

and 12 lags of employment. In the second and third we include through lags 13 and 24, respectively, of both employment and rig counts since these lag lengths corresponded to local minima in the BIC (see Figure A.4 in the Appendix). Plots of the IRFs and CIRFs for model (1) are shown in Figure 9. When lags of rig-counts are truncated, the employment response is more muted, which suggests that much of the employment activity happens with a lag. As lags 11 to 13 and lags 11 to 24 are added, the shapes of the IRF and CIRF do not change much, but the size of the effect does grow. A substantial portion of the increase in the multiplier seems to be due to the inclusion of more positive rig count coefficients since the ratio of $\hat{\beta}(1)$ to $\widehat{LRM}$ is relatively stable across specifications (this implies that $1 - \alpha(1)$ varies little with lag lengths greater than 12). It is not unreasonable to expect a boom in upstream activity to have effects beyond one year, but our ability to accurately estimate such long-run effects in a monthly regression is limited.
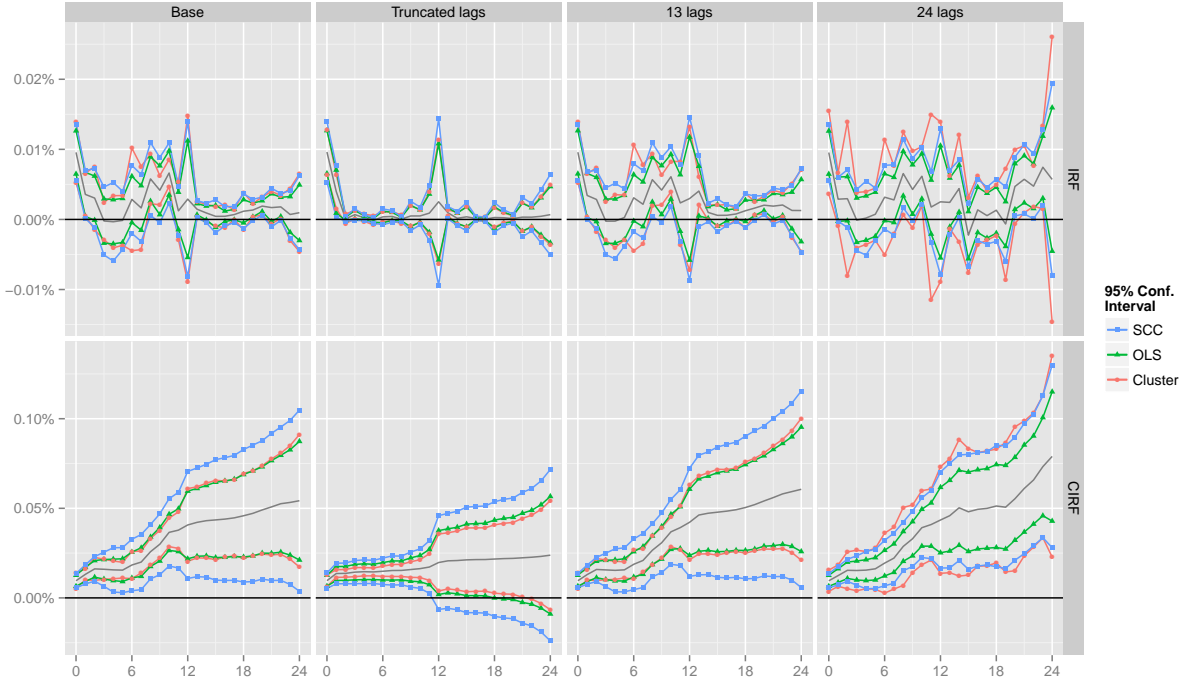
Figure 9: IRF and CIRF for log model (1) with varying lags

7.2 Outliers

In addition to checking whether various lag-lengths and different controls changed estimates, we estimated both models while dropping each of the 48 states. As shown in Tables 5 and 6, the estimates for both $\beta(1)$ and LRM from models (1) and (2) are very stable with respect to the subset of states included. In fact, the difference between the minimum and maximum of each set of multipliers is only slightly larger than the smallest standard deviation reported for each statistic. Interestingly, when North Dakota and Wyoming—the two states with per-capita rig counts that are by far the largest—are dropped, the sum of the rig count coefficients rises instead of falling in the log model. Given that the difference is not statistically significant, however, not much can be made of this.

7.3 Structural Breaks

One other concern is that the recession and its aftermath, not the unprecedented boom in unconventional oil and gas resources from 2008–2013, drives our results. If there is a structural break, it is unclear whether we would expect the pre-2008 long-run employment multiplier to be smaller or larger than the post-2008 multiplier. On the one hand, the post-2008 period corresponds to relatively high-cost drilling in areas that had historically seen less production and, therefore, would require more infrastructure investment to support drilling. On the other, firms drilling in these new areas might prefer to purchase inputs from areas like Texas and Louisiana with a historical oil and gas presence if these are cheaper or of higher quality, which would attenuate the state-level relationship between drilling and jobs.

To examine this issue, we estimate the following models

$$\Delta \log(Emp_{it}) = \alpha(L)\Delta \log(Emp_{it}) + \beta_{pre}(L)\Delta(Rigs/Pop_{it})\mathbb{1}[t < Jan2008]$$

$$+ \beta_{post}(L)\Delta(Rigs/Pop_{it})\mathbb{1}[t >= Jan2008] + \nu_i(t) + \nu_t + u_{it} \qquad (3)$$

$$\Delta(Emp/Pop)_{it} = \alpha(L)\Delta(Emp/Pop)_{it} + \beta_{pre}(L)\Delta(Rigs/Pop_{it})\mathbb{1}[t < Jan2008]$$

$$+ \beta_{post}(L)\Delta(Rigs/Pop_{it})\mathbb{1}[t >= Jan2008] + \nu_i(t) + \nu_t + u_{it} \qquad (4)$$

and test both that $\beta_{pre} = \beta_{post}$ as well as $\beta_{pre}(1) = \beta_{post}(1)$ using OLS, multi-way clustered, and Driscoll-Kraay estimates for the variance matrix. Under all three variance estimates, we reject both null hypotheses in favor of structural change at the 0.1% level (with the exception of the second hypothesis for model (4) under Driscoll-Kraay estimates, where the hypothesis was rejected at 1.5%). Post-2008 multipliers are more than twice as big as pre-2008 multipliers: 0.038% versus 0.078%. In the per-capita model, this translates to 201 versus 375 jobs per rig count.[22]

Figures 10a and 10b show IRFs and CIRFs corresponding to the two periods for mod-

---

[22]This outcome would also be consistent with the larger employment effects found in Hartley et al. (2013), where the wells drilled were all aimed at unconventional resources.

els (3) and (4), respectively. It is clear that employment impacts are larger in the second period.

7.4   Population Scaling

The choice of a population scaling variable does have implications for the magnitudes of the employment-multiplier. There are two issues that may arise from the choice of population variable. First, as pointed out by Blanchard and Katz (1992), population movements are generally in response to changes in labor demand. This means that exogenous variation—rig-counts—is polluted by an endogenous variable—population. Second, population growth—not just the variables of interest—affects results if using time-varying population measures (which are endogenous to employment) are used. To see this, first note the following equivalence:

$$\Delta(x/z)_t = z_t^{-1}\left[\Delta x_t - x_{t-1}(\Delta z_t/z_{t-1})\right].$$

In the case of model (2), the increase in employment per capita due to an increase in rig counts per capita is simply $\widehat{LRM} \times \Delta(Rigs/Pop)_t$. The the resulting change in jobs can be decomposed as

$$\frac{\partial \Delta Emp_{t+\infty}}{\partial \Delta Rigs_t} = \widehat{LRM} \times \Delta Rigs_t + (\Delta Pop_t/Pop_{t-1})(Emp_{t-1} - \widehat{LRM} \times Rigs_{t-1}).$$

While $\partial \Delta Emp_{t+\infty}/\partial \Delta Rigs_t = LRM$, if long-run growth in jobs due to rigs is computed using rigs per capita, the estimate will be polluted by population growth. In the log employment case, the problem is also present since the change in employment is computed as

$$\frac{\partial \Delta Emp_{t+\infty}}{\partial \Delta Rigs_t} = Emp_{t-1} \times \exp\left\{\left[\widehat{LRM} \times \Delta Rigs_t/Pop_t\right] - \left[\widehat{LRM} \times Rigs_{t-1}(\Delta Pop_t/Pop_{t-1})\right]\right\}.$$

Again, the second bracketed term involves ΔPop_t, which distorts the multiplier.
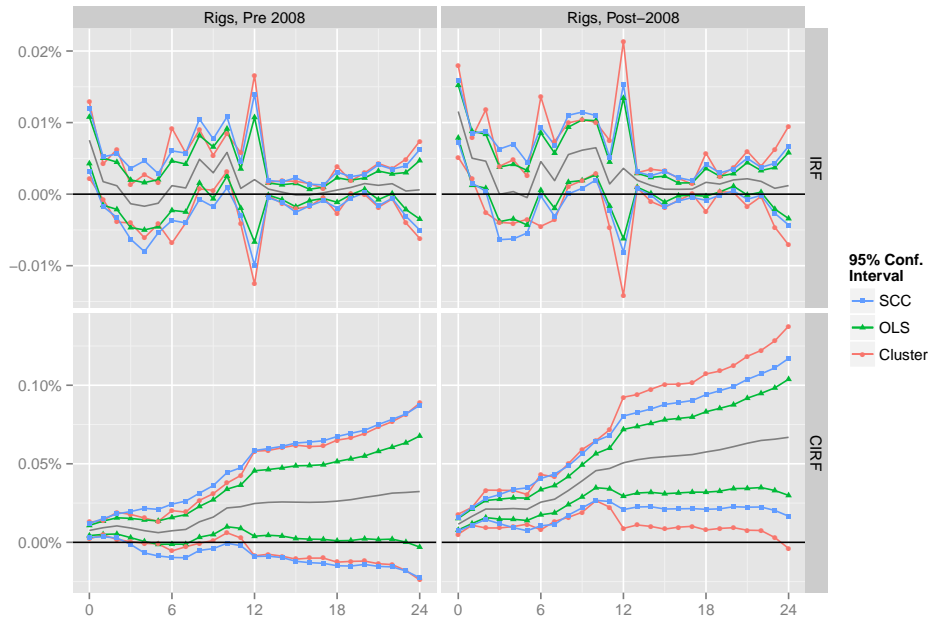
(a) Log model (3)
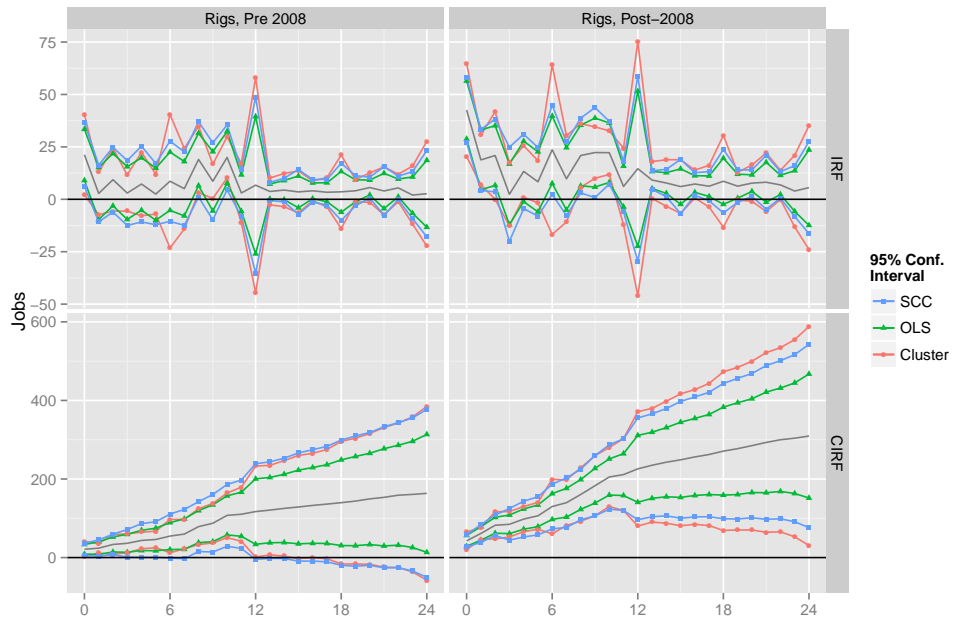
(b) Per-capita model (4)

Figure 10: IRF and CIRF for Rig Counts, Pre- and Post-2008

To address this concern, we estimate equations (1) and (2) using a variety of different population measures:

1. Population in 1990 as given by the US Census,

2. Population in 2000 as given by the US Census (our base model),

3. Population in 2010 as given by the US Census,

4. Total population from the Current Population Survey, where months between June of each year are linearly interpolated,

5. Working-age population as given in the Local Area Unemployment statistics from BLS, and

6. Working-age population lagged by one month, which should be pre-determined with respect to rig counts, which might drive population movements in time t.

The advantage of the first three population variables is that, with no population change, the "polluting" population growth term drops out of the estimation. When we use a fixed population parameter, model (2) is almost equivalent to weighted least squares (WLS) with constant, state-specific weights. To see this, suppose that the variance of state's idiosyncratic shock u_{it} is proportional to the square of its population. Then in OLS using first-differences of level variables, large states receive many times more weight than smaller states. If variables are all scaled by population, however, each state receives approximately the same weight in the model. Scaling by a constant population is not exactly WLS since the common time-shock ν_t is also implicitly divided by state population. Doing this, however, makes sense since a common shock should induce a larger absolute employment change in larger states and smaller absolute changes in smaller states.

Estimation results of these robustness checks are available in Tables 7 and 8 for the log and per-capita models, respectively. The general result—that upstream drilling activity causes statistically significant increases in employment—holds true in all of the models,

as does the pattern of dynamic response to rig counts. Nevertheless, point estimates are affected. In the log model (1), both $\hat{\beta}(1)$ and $\widehat{LRM}$ decrease by around a quarter when working-age population from the Local Area Unemployment statistics is used, though using a time-varying CPS population measure has minimal impact. In the per-capita model (2), the use of a time-varying population variable seems to drive a decrease of around a third, not the change in dataset.

8   Conclusion

To date, most estimates of the impact of unconventional oil and gas activity have been done through industry-related studies and using an Input-Output analysis. Analysis using a national-level VAR framework suggests that a 10% increase in rig counts raises employment in oil and gas extraction by approximately 5% after 24 months, but general employment impacts and employment impacts in manufacturing are not statistically distinguishable from zero. Using a dynamic panel model we confirm the hypothesis that increased drilling for oil and gas creates jobs, but we find the employment impacts to be much smaller than figures put forward by Input-Output studies. In particular, we estimate that the long-run employment impact of an additional rig count per million people is an increase of 0.064% in total employment. A per-capita formulation of the model yields a LRM of 315 jobs per rig count. Our estimates appear robust to autocorrelation, heteroskedasticity, and cross-sectional dependency in the error terms. They also appear to be consistent across sets of states included in our sample and different lag-lengths. We find some evidence that upstream activity may have had a larger impact post-2008, though we cannot say as to why this would be. Additionally, using other measures of population decreases our job-creation estimates by up to a third, though these measures of population do not seem as appropriate. The fact that different modeling approaches result in substantially different estimates of the employment impacts suggests the need for further research in order to better understand the role of alternative modeling

assumptions on these results.

	Coefficient (%)	OLS	Cluster	SCC
$\Delta \log Employment_{i,t-1}$	-2.1305	$(0.8906)^*$	$(3.7387)^*$	(2.4636)
$\Delta \log Employment_{i,t-2}$	1.4896	(0.8882)	(2.2458)	(1.7444)
$\Delta \log Employment_{i,t-3}$	6.3648	$(0.8870)^{***}$	$(1.8624)^{***}$	$(1.4389)^{***}$
$\Delta \log Employment_{i,t-4}$	0.1043	(0.8873)	(1.6728)	(1.2686)
$\Delta \log Employment_{i,t-5}$	-1.3672	(0.8879)	(2.3544)	(1.1426)
$\Delta \log Employment_{i,t-6}$	3.8131	$(0.8863)^{***}$	$(2.5049)^{***}$	$(1.2945)^{**}$
$\Delta \log Employment_{i,t-7}$	2.6306	$(0.8828)^{**}$	$(2.3930)^{**}$	$(1.1599)^*$
$\Delta \log Employment_{i,t-8}$	-0.8629	(0.8832)	(1.7269)	(1.0862)
$\Delta \log Employment_{i,t-9}$	4.3981	$(0.8811)^{***}$	$(1.5566)^{***}$	$(1.2407)^{***}$
$\Delta \log Employment_{i,t-10}$	2.3412	$(0.8756)^{**}$	$(1.7832)^{**}$	$(1.1793)^*$
$\Delta \log Employment_{i,t-11}$	8.0504	$(0.8744)^{***}$	$(2.9623)^{***}$	$(1.8155)^{***}$
$\Delta \log Employment_{i,t-12}$	22.1630	$(0.8733)^{***}$	$(3.8013)^{***}$	$(2.7968)^{***}$
$\Delta Rigs/Pop2000_{i,t-0}$	0.0096	$(0.0016)^{***}$	$(0.0022)^{***}$	$(0.0020)^{***}$
$\Delta Rigs/Pop2000_{i,t-1}$	0.0038	$(0.0016)^*$	$(0.0017)^*$	$(0.0018)^*$
$\Delta Rigs/Pop2000_{i,t-2}$	0.0030	(0.0016)	(0.0021)	(0.0022)
$\Delta Rigs/Pop2000_{i,t-3}$	-0.0008	(0.0016)	(0.0017)	(0.0025)
$\Delta Rigs/Pop2000_{i,t-4}$	-0.0006	(0.0016)	(0.0018)	(0.0028)
$\Delta Rigs/Pop2000_{i,t-5}$	-0.0002	(0.0016)	(0.0018)	(0.0022)
$\Delta Rigs/Pop2000_{i,t-6}$	0.0026	(0.0016)	(0.0035)	(0.0023)
$\Delta Rigs/Pop2000_{i,t-7}$	0.0014	(0.0016)	(0.0030)	(0.0024)
$\Delta Rigs/Pop2000_{i,t-8}$	0.0057	$(0.0016)^{***}$	$(0.0019)^{***}$	$(0.0026)^*$
$\Delta Rigs/Pop2000_{i,t-9}$	0.0036	$(0.0016)^*$	$(0.0015)^*$	(0.0021)
$\Delta Rigs/Pop2000_{i,t-10}$	0.0062	$(0.0016)^{***}$	$(0.0011)^{***}$	$(0.0022)^{**}$
$\hat{\beta}(1)$	0.0341	$(0.0042)^{***}$	$(0.0038)^{***}$	$(0.0079)^{***}$
$\widehat{LRM}$	0.0643	$(0.0082)^{***}$	$(0.0102)^{***}$	$(0.0174)^{***}$

NOTE: Estimates and SEs are scaled by 100 to represent percentages.
$^{***}p < 0.001$, $^{**}p < 0.01$, $^*p < 0.05$. Standard errors in parenthesis. 48 states, 265 months, 12,720 observations. State-month and Time FE included, as well as lags 1–12 of state employment.

Table 1: Base model (1) in logs with different standard errors

	Coefficient	OLS	Cluster	SCC
$\Delta Emp/Pop00_{i,t-1}$	-0.039	$(0.009)^{***}$	$(0.036)^{***}$	(0.024)
$\Delta Emp/Pop00_{i,t-2}$	0.017	(0.009)	(0.076)	(0.017)
$\Delta Emp/Pop00_{i,t-3}$	0.071	$(0.009)^{***}$	$(0.068)^{***}$	$(0.015)^{***}$
$\Delta Emp/Pop00_{i,t-4}$	0.014	(0.009)	(0.059)	(0.015)
$\Delta Emp/Pop00_{i,t-5}$	0.012	(0.009)	(0.060)	(0.012)
$\Delta Emp/Pop00_{i,t-6}$	0.069	$(0.009)^{***}$	$(0.029)^{***}$	$(0.013)^{***}$
$\Delta Emp/Pop00_{i,t-7}$	0.045	$(0.009)^{***}$	$(0.035)^{***}$	$(0.012)^{***}$
$\Delta Emp/Pop00_{i,t-8}$	-0.001	(0.009)	(0.016)	(0.012)
$\Delta Emp/Pop00_{i,t-9}$	0.046	$(0.009)^{***}$	$(0.022)^{***}$	$(0.012)^{***}$
$\Delta Emp/Pop00_{i,t-10}$	0.011	(0.009)	(0.015)	(0.011)
$\Delta Emp/Pop00_{i,t-11}$	0.062	$(0.009)^{***}$	$(0.063)^{***}$	$(0.016)^{***}$
$\Delta Emp/Pop00_{i,t-12}$	0.217	$(0.009)^{***}$	$(0.052)^{***}$	$(0.025)^{***}$
$\Delta Rigs/Pop00_{i,t-0}$	30.905	$(5.912)^{***}$	$(7.240)^{***}$	$(7.801)^{***}$
$\Delta Rigs/Pop00_{i,t-1}$	11.962	$(6.025)^{*}$	$(6.874)^{*}$	(7.073)
$\Delta Rigs/Pop00_{i,t-2}$	15.885	$(6.019)^{**}$	$(6.132)^{**}$	$(8.058)^{*}$
$\Delta Rigs/Pop00_{i,t-3}$	3.416	(6.028)	(6.502)	(7.644)
$\Delta Rigs/Pop00_{i,t-4}$	10.661	(6.053)	(6.844)	(8.055)
$\Delta Rigs/Pop00_{i,t-5}$	6.524	(6.005)	(5.920)	(7.699)
$\Delta Rigs/Pop00_{i,t-6}$	13.011	$(6.023)^{*}$	$(12.159)^{*}$	(8.036)
$\Delta Rigs/Pop00_{i,t-7}$	5.576	(6.026)	(13.022)	(8.194)
$\Delta Rigs/Pop00_{i,t-8}$	20.196	$(6.010)^{***}$	$(6.255)^{***}$	$(8.486)^{*}$
$\Delta Rigs/Pop00_{i,t-9}$	10.759	(5.998)	(6.812)	(7.733)
$\Delta Rigs/Pop00_{i,t-10}$	20.904	$(5.885)^{***}$	$(2.846)^{***}$	$(7.450)^{**}$
$\hat{\beta}(1)$	149.800	$(15.726)^{***}$	$(17.782)^{***}$	$(27.943)^{***}$
$\widehat{LRM}$	315.187	$(35.448)^{***}$	(275.434)	$(77.946)^{***}$

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Standard errors in parenthesis. 48 states, 265 months, 12,720 observations. State-month and Time FE included.

Table 2: Base model (2) in per-capita terms with different standard errors

	Base	Truncated lags	13 lags	24 lags
$\Delta Rigs/Pop2000_{i,t-0}$	0.0096 (0.0020)***	0.0096 (0.0022)***	0.0096 (0.0020)***	0.0095 (0.0020)***
$\Delta Rigs/Pop2000_{i,t-1}$	0.0038 (0.0018)*	0.0042 (0.0019)*	0.0037 (0.0017)*	0.0033 (0.0016)*
$\Delta Rigs/Pop2000_{i,t-2}$	0.0030 (0.0022)		0.0029 (0.0021)	0.0030 (0.0021)
$\Delta Rigs/Pop2000_{i,t-3}$	−0.0008 (0.0025)		−0.0008 (0.0024)	−0.0006 (0.0023)
$\Delta Rigs/Pop2000_{i,t-4}$	−0.0006 (0.0028)		−0.0005 (0.0027)	0.0000 (0.0027)
$\Delta Rigs/Pop2000_{i,t-5}$	−0.0002 (0.0022)		0.0002 (0.0021)	0.0006 (0.0020)
$\Delta Rigs/Pop2000_{i,t-6}$	0.0026 (0.0023)		0.0028 (0.0023)	0.0028 (0.0021)
$\Delta Rigs/Pop2000_{i,t-7}$	0.0014 (0.0024)		0.0020 (0.0024)	0.0025 (0.0025)
$\Delta Rigs/Pop2000_{i,t-8}$	0.0057 (0.0026)*		0.0056 (0.0026)*	0.0064 (0.0025)*
$\Delta Rigs/Pop2000_{i,t-9}$	0.0036 (0.0021)		0.0037 (0.0020)	0.0038 (0.0020)
$\Delta Rigs/Pop2000_{i,t-10}$	0.0062 (0.0022)**		0.0058 (0.0022)**	0.0058 (0.0021)**
$\Delta Rigs/Pop2000_{i,t-11}$			0.0013 (0.0021)	0.0007 (0.0020)
$\Delta Rigs/Pop2000_{i,t-12}$			0.0002 (0.0020)	0.0000 (0.0019)
$\Delta Rigs/Pop2000_{i,t-13}$			0.0022 (0.0022)	0.0007 (0.0021)
$\Delta Rigs/Pop2000_{i,t-14}$				0.0031 (0.0021)
$\Delta Rigs/Pop2000_{i,t-15}$				−0.0029 (0.0020)
$\Delta Rigs/Pop2000_{i,t-16}$				0.0006 (0.0022)
$\Delta Rigs/Pop2000_{i,t-17}$				−0.0005 (0.0021)
$\Delta Rigs/Pop2000_{i,t-18}$				0.0003 (0.0019)
$\Delta Rigs/Pop2000_{i,t-19}$				−0.0021 (0.0026)
$\Delta Rigs/Pop2000_{i,t-20}$				0.0027 (0.0020)
$\Delta Rigs/Pop2000_{i,t-21}$				0.0043 (0.0023)
$\Delta Rigs/Pop2000_{i,t-22}$				0.0035 (0.0023)
$\Delta Rigs/Pop2000_{i,t-23}$				0.0062 (0.0021)**
$\Delta Rigs/Pop2000_{i,t-24}$				0.0042 (0.0023)
$\hat{\beta}(1)$	0.0341 (0.0079)***	0.0138 (0.0029)***	0.0387 (0.0095)***	0.0578 (0.0112)***
$\widehat{LRM}$	0.0643 (0.0174)***	0.0269 (0.0070)***	0.0759 (0.0210)***	0.1112 (0.0220)***
BIC	−106146.4584	−106172.7398	−106147.9075	−106149.0476

NOTE: Estimates and SEs are scaled by 100 to represent percentages. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. State-month and time fixed effects included. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.001 level.

Table 3: Log model (1) with differing lag-lengths

	Base	Truncated lags	13 lags	24 lags
$\Delta Rigs/Pop2000_{i,t-0}$	30.91 (7.80)***	31.21 (8.09)***	31.17 (7.92)***	31.86 (7.57)***
$\Delta Rigs/Pop2000_{i,t-1}$	11.96 (7.07)	15.26 (7.60)*	12.19 (7.15)	11.25 (6.53)
$\Delta Rigs/Pop2000_{i,t-2}$	15.89 (8.06)*		15.81 (8.08)	17.43 (8.21)*
$\Delta Rigs/Pop2000_{i,t-3}$	3.42 (7.64)		3.56 (7.62)	4.63 (7.31)
$\Delta Rigs/Pop2000_{i,t-4}$	10.66 (8.06)		10.99 (7.97)	12.10 (8.09)
$\Delta Rigs/Pop2000_{i,t-5}$	6.52 (7.70)		7.08 (7.77)	7.75 (7.51)
$\Delta Rigs/Pop2000_{i,t-6}$	13.01 (8.04)		13.25 (8.13)	12.84 (7.62)
$\Delta Rigs/Pop2000_{i,t-7}$	5.58 (8.19)		6.74 (8.59)	8.12 (8.76)
$\Delta Rigs/Pop2000_{i,t-8}$	20.20 (8.49)*		20.34 (8.44)*	22.86 (8.15)**
$\Delta Rigs/Pop2000_{i,t-9}$	10.76 (7.73)		11.37 (7.66)	11.78 (7.25)
$\Delta Rigs/Pop2000_{i,t-10}$	20.90 (7.45)**		20.96 (7.61)**	21.84 (7.84)**
$\Delta Rigs/Pop2000_{i,t-11}$			−0.21 (7.30)	−1.32 (7.02)
$\Delta Rigs/Pop2000_{i,t-12}$			−3.46 (7.88)	−3.39 (8.16)
$\Delta Rigs/Pop2000_{i,t-13}$			3.54 (8.86)	−2.07 (8.74)
$\Delta Rigs/Pop2000_{i,t-14}$				12.04 (7.83)
$\Delta Rigs/Pop2000_{i,t-15}$				−8.13 (8.03)
$\Delta Rigs/Pop2000_{i,t-16}$				12.29 (7.15)
$\Delta Rigs/Pop2000_{i,t-17}$				4.45 (7.90)
$\Delta Rigs/Pop2000_{i,t-18}$				4.84 (7.49)
$\Delta Rigs/Pop2000_{i,t-19}$				−5.77 (9.00)
$\Delta Rigs/Pop2000_{i,t-20}$				14.20 (7.46)
$\Delta Rigs/Pop2000_{i,t-21}$				18.20 (9.40)
$\Delta Rigs/Pop2000_{i,t-22}$				13.74 (8.11)
$\Delta Rigs/Pop2000_{i,t-23}$				23.96 (8.03)**
$\Delta Rigs/Pop2000_{i,t-24}$				14.16 (8.61)
Employment lags included	12	12	13	24
$\hat{\beta}(1)$	149.80 (27.94)***	46.47 (11.94)***	153.34 (34.96)***	259.67 (47.16)***
$\widehat{LRM}$	315.19 (77.95)***	102.29 (34.89)***	331.77 (90.43)***	524.39 (97.64)***
BIC	44762.48	44743.25	44786.77	44807.14

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. State-month and time fixed effects included. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.001 level.

Table 4: Per-capita model (2) with differing lag-lengths

	Alabama	Arizona	Arkansas	California	Colorado	Connecticut	Delaware	Florida
$\hat{\beta}(1)$	0.0339	0.0344	0.0340	0.0344	0.0340	0.0343	0.0339	0.0343
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0041)	(0.0042)
$\widehat{LRM}$	0.0643	0.0621	0.0644	0.0635	0.0636	0.0646	0.0660	0.0634
	(0.0083)	(0.0079)	(0.0083)	(0.0081)	(0.0082)	(0.0082)	(0.0084)	(0.0081)

	Georgia	Idaho	Illinois	Indiana	Iowa	Kansas	Kentucky	Louisiana
$\hat{\beta}(1)$	0.0339	0.0343	0.0342	0.0339	0.0340	0.0344	0.0338	0.0348
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0040)
$\widehat{LRM}$	0.0642	0.0644	0.0648	0.0641	0.0644	0.0652	0.0644	0.0710
	(0.0082)	(0.0082)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0083)	(0.0087)

	Maine	Maryland	Massachusetts	Michigan	Minnesota	Mississippi	Missouri	Montana
$\hat{\beta}(1)$	0.0335	0.0340	0.0338	0.0338	0.0340	0.0343	0.0339	0.0343
	(0.0041)	(0.0042)	(0.0042)	(0.0041)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
$\widehat{LRM}$	0.0639	0.0644	0.0635	0.0637	0.0646	0.0649	0.0644	0.0656
	(0.0083)	(0.0083)	(0.0082)	(0.0081)	(0.0083)	(0.0082)	(0.0083)	(0.0084)

	Nebraska	Nevada	New Hampshire	New Jersey	New Mexico	New York	North Carolina	North Dakota
$\hat{\beta}(1)$	0.0342	0.0351	0.0340	0.0339	0.0341	0.0341	0.0341	0.0349
	(0.0042)	(0.0041)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0055)
$\widehat{LRM}$	0.0651	0.0612	0.0645	0.0641	0.0637	0.0634	0.0645	0.0652
	(0.0083)	(0.0075)	(0.0083)	(0.0083)	(0.0082)	(0.0081)	(0.0083)	(0.0107)

	Ohio	Oklahoma	Oregon	Pennsylvania	Rhode Island	South Carolina	South Dakota	Tennessee
$\hat{\beta}(1)$	0.0341	0.0331	0.0341	0.0340	0.0336	0.0339	0.0339	0.0340
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0041)	(0.0042)	(0.0042)	(0.0042)
$\widehat{LRM}$	0.0644	0.0628	0.0640	0.0637	0.0634	0.0644	0.0645	0.0641
	(0.0083)	(0.0084)	(0.0082)	(0.0082)	(0.0082)	(0.0083)	(0.0083)	(0.0083)

	Texas	Utah	Vermont	Virginia	Washington	West Virginia	Wisconsin	Wyoming
$\hat{\beta}(1)$	0.0333	0.0339	0.0333	0.0340	0.0342	0.0346	0.0340	0.0364
	(0.0042)	(0.0042)	(0.0041)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0053)
$\widehat{LRM}$	0.0626	0.0627	0.0641	0.0643	0.0645	0.0655	0.0645	0.0685
	(0.0083)	(0.0081)	(0.0083)	(0.0083)	(0.0082)	(0.0083)	(0.0083)	(0.0102)

NOTE: Estimates and SEs are scaled by 100 to represent percentages. Each regression omits the named state. OLS standard errors in parenthesis. $***p < 0.001$, $**p < 0.01$, $*p < 0.05$. 47 states, 265 months, 12,455 observations. State-month and Time FE included.

Table 5: Base model (1) in logs omitting one state at a time

	Alabama	Arizona	Arkansas	California	Colorado	Connecticut	Delaware	Florida
$\hat{\beta}(1)$	149.16	151.03	149.15	150.35	149.88	150.55	148.60	150.86
	(15.83)	(15.78)	(15.81)	(15.83)	(15.83)	(15.78)	(15.43)	(15.82)
$\widehat{LRM}$	315.09	307.06	314.79	313.41	312.59	318.44	324.97	310.98
	(35.85)	(34.30)	(35.76)	(35.35)	(35.38)	(35.85)	(36.37)	(34.90)

	Georgia	Idaho	Illinois	Indiana	Iowa	Kansas	Kentucky	Louisiana
$\hat{\beta}(1)$	148.77	150.14	150.21	149.12	149.69	151.45	148.65	152.28
	(15.70)	(15.74)	(15.85)	(15.79)	(15.82)	(15.79)	(15.79)	(15.26)
$\widehat{LRM}$	315.23	316.06	317.34	314.56	316.19	320.08	315.05	341.18
	(35.66)	(35.53)	(35.91)	(35.69)	(35.84)	(35.87)	(35.86)	(37.26)

	Maine	Maryland	Massachusetts	Michigan	Minnesota	Mississippi	Missouri	Montana
$\hat{\beta}(1)$	147.49	149.64	148.82	148.00	149.37	149.89	149.10	151.35
	(15.66)	(15.77)	(15.80)	(15.63)	(15.78)	(15.75)	(15.80)	(15.83)
$\widehat{LRM}$	313.66	316.24	312.38	312.45	317.54	316.51	315.72	321.49
	(35.70)	(35.76)	(35.49)	(35.34)	(35.97)	(35.69)	(35.88)	(36.14)

	Nebraska	Nevada	New Hampshire	New Jersey	New Mexico	New York	North Carolina	North Dakota
$\hat{\beta}(1)$	150.22	157.88	149.20	149.24	151.94	149.60	149.90	139.29
	(15.72)	(15.54)	(15.76)	(15.79)	(16.05)	(15.84)	(15.81)	20.62
$\widehat{LRM}$	320.67	290.99	316.50	315.47	317.12	312.01	316.14	285.30
	(36.09)	(30.56)	(35.85)	(35.81)	(35.75)	(35.37)	(35.75)	44.52

	Ohio	Oklahoma	Oregon	Pennsylvania	Rhode Island	South Carolina	South Dakota	Tennessee
$\hat{\beta}(1)$	148.43	149.49	149.69	149.53	147.76	148.81	149.05	149.45
	(16.08)	(15.83)	(15.80)	(15.85)	(15.65)	(15.73)	(15.72)	(15.83)
$\widehat{LRM}$	313.24	314.99	316.50	315.47	317.12	316.23	315.83	314.47
	(36.15)	(35.75)	(35.85)	(35.81)	(35.75)	(35.83)	(35.72)	(35.67)

	Texas	Utah	Vermont	Virginia	Washington	West Virginia	Wisconsin	Wyoming
$\hat{\beta}(1)$	146.83	148.61	146.21	149.72	150.09	151.98	149.58	166.11
	(15.93)	(15.79)	(15.54)	(15.82)	(15.77)	(15.80)	(15.83)	(20.03)
$\widehat{LRM}$	307.26	307.89	313.57	315.80	312.70	318.77	316.34	346.41
	(35.55)	(35.03)	(35.69)	(35.76)	(35.70)	(35.58)	(35.89)	(43.88)

Each regression omits the named state. OLS standard errors in parenthesis. $p < 0.001$, $^{**}p < 0.01$, $^*p < 0.05$. 47 states, 265 months, 12,455 observations. State-month and Time FE included.

Table 6: Base model(2) in per-capita terms omitting one state at a time

	CPS 1990	CPS 2000 (Base)	CPS 2010	CPS Interpolated	LAU	Lagged LAU
$\Delta Rigs/Pop_{i,t-0}$	$0.0089\ (0.0019)^{***}$	$0.0096\ (0.0020)^{***}$	$0.0104\ (0.0022)^{***}$	$0.0099\ (0.0020)^{***}$	$0.0072\ (0.0015)^{***}$	$0.0072\ (0.0015)^{***}$
$\Delta Rigs/Pop_{i,t-1}$	$0.0036\ (0.0016)^{*}$	$0.0038\ (0.0018)^{*}$	$0.0041\ (0.0019)^{*}$	$0.0037\ (0.0018)^{*}$	$0.0029\ (0.0014)^{*}$	$0.0029\ (0.0014)^{*}$
$\Delta Rigs/Pop_{i,t-2}$	$0.0028\ (0.0020)$	$0.0030\ (0.0022)$	$0.0031\ (0.0024)$	$0.0025\ (0.0022)$	$0.0019\ (0.0017)$	$0.0019\ (0.0017)$
$\Delta Rigs/Pop_{i,t-3}$	$-0.0008\ (0.0023)$	$-0.0008\ (0.0025)$	$-0.0009\ (0.0027)$	$-0.0010\ (0.0024)$	$-0.0007\ (0.0018)$	$-0.0007\ (0.0018)$
$\Delta Rigs/Pop_{i,t-4}$	$-0.0004\ (0.0026)$	$-0.0006\ (0.0028)$	$-0.0007\ (0.0031)$	$-0.0011\ (0.0029)$	$-0.0009\ (0.0022)$	$-0.0009\ (0.0022)$
$\Delta Rigs/Pop_{i,t-5}$	$-0.0003\ (0.0020)$	$-0.0002\ (0.0022)$	$-0.0001\ (0.0024)$	$-0.0003\ (0.0022)$	$-0.0002\ (0.0017)$	$-0.0002\ (0.0017)$
$\Delta Rigs/Pop_{i,t-6}$	$0.0025\ (0.0021)$	$0.0026\ (0.0023)$	$0.0025\ (0.0025)$	$0.0025\ (0.0023)$	$0.0019\ (0.0017)$	$0.0019\ (0.0017)$
$\Delta Rigs/Pop_{i,t-7}$	$0.0013\ (0.0022)$	$0.0014\ (0.0024)$	$0.0015\ (0.0027)$	$0.0015\ (0.0024)$	$0.0012\ (0.0018)$	$0.0012\ (0.0018)$
$\Delta Rigs/Pop_{i,t-8}$	$0.0054\ (0.0024)^{*}$	$0.0057\ (0.0026)^{*}$	$0.0062\ (0.0029)^{*}$	$0.0060\ (0.0026)^{*}$	$0.0045\ (0.0019)^{*}$	$0.0045\ (0.0019)^{*}$
$\Delta Rigs/Pop_{i,t-9}$	$0.0033\ (0.0019)$	$0.0036\ (0.0021)$	$0.0041\ (0.0023)$	$0.0038\ (0.0021)$	$0.0029\ (0.0016)$	$0.0029\ (0.0016)$
$\Delta Rigs/Pop_{i,t-10}$	$0.0057\ (0.0020)^{**}$	$0.0062\ (0.0022)^{**}$	$0.0068\ (0.0024)^{**}$	$0.0065\ (0.0022)^{**}$	$0.0049\ (0.0017)^{**}$	$0.0049\ (0.0017)^{**}$
$\hat{\beta}(1)$	$0.0322\ (0.0076)^{***}$	$0.0341\ (0.0079)^{***}$	$0.0372\ (0.0085)^{***}$	$0.0340\ (0.0083)^{***}$	$0.0256\ (0.0064)^{***}$	$0.0256\ (0.0064)^{***}$
$\widehat{LRM}$	$0.0608\ (0.0166)^{***}$	$0.0643\ (0.0174)^{***}$	$0.0701\ (0.0186)^{***}$	$0.0645\ (0.0181)^{***}$	$0.0485\ (0.0139)^{***}$	$0.0485\ (0.0139)^{***}$

NOTE: Estimates and SEs are scaled by 100 to represent percentages. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. 12 employment lags plus state-month and time fixed effects included. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.001 level.

Table 7: Log model(1) with different population normalizations

	CPS 1990	CPS 2000 (Base)	CPS 2010	CPS Interpolated	LAU	Lagged LAU
$\Delta Rigs/Pop_{i,t-0}$	29.83 (7.76)***	30.91 (7.80)***	30.14 (7.69)***	33.40 (7.67)***	36.28 (7.72)***	36.28 (7.72)***
$\Delta Rigs/Pop_{i,t-1}$	11.64 (6.87)	11.96 (7.07)	12.25 (7.16)	13.92 (7.09)*	11.05 (8.01)	11.05 (8.01)
$\Delta Rigs/Pop_{i,t-2}$	13.56 (8.03)	15.89 (8.06)*	15.69 (8.06)	14.29 (7.71)	8.59 (10.24)	8.59 (10.24)
$\Delta Rigs/Pop_{i,t-3}$	0.24 (7.36)	3.42 (7.64)	4.28 (7.76)	1.91 (7.89)	1.50 (9.06)	1.50 (9.06)
$\Delta Rigs/Pop_{i,t-4}$	9.42 (7.92)	10.66 (8.06)	11.74 (8.13)	5.88 (8.76)	1.02 (9.62)	1.02 (9.62)
$\Delta Rigs/Pop_{i,t-5}$	4.34 (7.79)	6.52 (7.70)	8.24 (7.93)	2.70 (7.35)	−1.23 (7.54)	−1.23 (7.54)
$\Delta Rigs/Pop_{i,t-6}$	12.64 (7.82)	13.01 (8.04)	11.98 (8.14)	11.44 (8.18)	10.12 (9.34)	10.12 (9.34)
$\Delta Rigs/Pop_{i,t-7}$	4.89 (8.03)	5.58 (8.19)	6.98 (8.52)	5.34 (7.87)	8.05 (7.67)	8.05 (7.67)
$\Delta Rigs/Pop_{i,t-8}$	20.27 (8.27)*	20.20 (8.49)*	20.25 (8.59)*	21.40 (8.82)*	21.70 (8.79)*	21.70 (8.79)*
$\Delta Rigs/Pop_{i,t-9}$	8.92 (7.61)	10.76 (7.73)	12.54 (7.88)	12.55 (7.18)	13.84 (7.57)	13.84 (7.57)
$\Delta Rigs/Pop_{i,t-10}$	19.77 (7.42)**	20.90 (7.45)**	21.62 (7.54)**	23.22 (7.46)**	28.31 (8.91)**	28.31 (8.91)**
$\hat{\beta}(1)$	135.51 (26.44)***	149.80 (27.94)***	155.70 (28.86)***	146.05 (30.24)***	139.23 (39.14)***	139.23 (39.14)***
$\widehat{LRM}$	366.53 (120.84)***	315.19 (77.95)***	290.76 (65.01)***	214.50 (49.63)***	213.43 (64.94)***	213.43 (64.94)***

Table 8: Per-capita model(2) with different population normalizations

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. 12 employment lags plus state-month and time fixed effects included. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.025 level.

	Base	Truncated lags	13 lags	24 lags
$\Delta Rigs/Pop2000_{i,t-0}$	0.0115 (0.0023)***	0.0116 (0.0024)***	0.0116 (0.0023)***	0.0114 (0.0021)***
$\Delta Rigs/Pop2000_{i,t-1}$	0.0035 (0.0019)	0.0038 (0.0020)	0.0030 (0.0019)	0.0031 (0.0017)
$\Delta Rigs/Pop2000_{i,t-2}$	0.0023 (0.0025)		0.0022 (0.0025)	0.0030 (0.0023)
$\Delta Rigs/Pop2000_{i,t-3}$	−0.0004 (0.0028)		−0.0001 (0.0028)	−0.0002 (0.0026)
$\Delta Rigs/Pop2000_{i,t-4}$	−0.0011 (0.0034)		−0.0011 (0.0033)	−0.0003 (0.0031)
$\Delta Rigs/Pop2000_{i,t-5}$	−0.0015 (0.0024)		−0.0012 (0.0024)	−0.0011 (0.0021)
$\Delta Rigs/Pop2000_{i,t-6}$	0.0020 (0.0026)		0.0020 (0.0026)	0.0016 (0.0023)
$\Delta Rigs/Pop2000_{i,t-7}$	0.0012 (0.0027)		0.0021 (0.0027)	0.0016 (0.0027)
$\Delta Rigs/Pop2000_{i,t-8}$	0.0054 (0.0026)*		0.0054 (0.0027)*	0.0057 (0.0024)*
$\Delta Rigs/Pop2000_{i,t-9}$	0.0048 (0.0022)*		0.0049 (0.0021)*	0.0045 (0.0020)*
$\Delta Rigs/Pop2000_{i,t-10}$	0.0059 (0.0025)*		0.0050 (0.0025)*	0.0048 (0.0024)*
$\Delta Rigs/Pop2000_{i,t-11}$			0.0025 (0.0022)	0.0019 (0.0021)
$\Delta Rigs/Pop2000_{i,t-12}$			−0.0003 (0.0022)	−0.0007 (0.0020)
$\Delta Rigs/Pop2000_{i,t-13}$			0.0020 (0.0023)	0.0004 (0.0022)
$\Delta Rigs/Pop2000_{i,t-14}$				0.0033 (0.0024)
$\Delta Rigs/Pop2000_{i,t-15}$				−0.0024 (0.0021)
$\Delta Rigs/Pop2000_{i,t-16}$				0.0006 (0.0025)
$\Delta Rigs/Pop2000_{i,t-17}$				−0.0002 (0.0022)
$\Delta Rigs/Pop2000_{i,t-18}$				0.0007 (0.0020)
$\Delta Rigs/Pop2000_{i,t-19}$				−0.0037 (0.0027)
$\Delta Rigs/Pop2000_{i,t-20}$				0.0028 (0.0025)
$\Delta Rigs/Pop2000_{i,t-21}$				0.0030 (0.0025)
$\Delta Rigs/Pop2000_{i,t-22}$				0.0036 (0.0023)
$\Delta Rigs/Pop2000_{i,t-23}$				0.0054 (0.0022)*
$\Delta Rigs/Pop2000_{i,t-24}$				0.0029 (0.0025)
$\hat{\beta}(1)$	0.0336 (0.0085)***	0.0154 (0.0032)***	0.0380 (0.0101)***	0.0518 (0.0122)***
$\widehat{LRM}$	0.0627 (0.0174)***	0.0296 (0.0079)***	0.0736 (0.0214)***	0.1043 (0.0249)***
BIC	−105020.3853	−105058.1508	−104962.3525	−105071.1135

NOTE: Estimates and SEs are scaled by 100 to represent percentages. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. State-month effects ONLY included. Control variables include national log $Employment_{t-1}$ and log IPI_{t-0} with 12, 12, 13 and 24 lags. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.001 level.

Table 9: Log model with differing lag-lengths and controls

	Base	Truncated lags	13 lags	24 lags
$\Delta Rigs/Pop2000_{i,t-0}$	39.22 (9.00)***	39.73 (9.45)***	39.39 (9.00)***	40.91 (8.25)***
$\Delta Rigs/Pop2000_{i,t-1}$	11.31 (7.80)	15.42 (7.94)	10.40 (7.86)	10.25 (6.93)
$\Delta Rigs/Pop2000_{i,t-2}$	17.64 (8.75)*		17.96 (8.58)*	18.47 (8.56)*
$\Delta Rigs/Pop2000_{i,t-3}$	6.83 (8.70)		6.90 (8.73)	6.66 (8.40)
$\Delta Rigs/Pop2000_{i,t-4}$	11.00 (9.89)		11.92 (9.74)	13.53 (9.55)
$\Delta Rigs/Pop2000_{i,t-5}$	−1.71 (8.04)		−0.40 (8.15)	−1.06 (7.82)
$\Delta Rigs/Pop2000_{i,t-6}$	10.89 (9.02)		10.50 (9.11)	9.13 (8.25)
$\Delta Rigs/Pop2000_{i,t-7}$	3.45 (9.05)		4.58 (9.48)	3.46 (9.52)
$\Delta Rigs/Pop2000_{i,t-8}$	19.73 (8.26)*		19.78 (8.40)*	20.41 (7.91)**
$\Delta Rigs/Pop2000_{i,t-9}$	14.45 (8.07)		14.97 (7.99)	14.64 (7.69)
$\Delta Rigs/Pop2000_{i,t-10}$	21.58 (8.35)**		20.47 (8.37)*	20.30 (8.49)*
$\Delta Rigs/Pop2000_{i,t-11}$			4.05 (7.61)	2.33 (7.17)
$\Delta Rigs/Pop2000_{i,t-12}$			−4.87 (8.47)	−5.61 (8.37)
$\Delta Rigs/Pop2000_{i,t-13}$			5.34 (9.18)	−2.93 (8.97)
$\Delta Rigs/Pop2000_{i,t-14}$				15.28 (8.70)
$\Delta Rigs/Pop2000_{i,t-15}$				−5.10 (8.64)
$\Delta Rigs/Pop2000_{i,t-16}$				13.55 (7.84)
$\Delta Rigs/Pop2000_{i,t-17}$				1.89 (8.39)
$\Delta Rigs/Pop2000_{i,t-18}$				6.63 (7.73)
$\Delta Rigs/Pop2000_{i,t-19}$				−13.01 (9.08)
$\Delta Rigs/Pop2000_{i,t-20}$				14.61 (8.63)
$\Delta Rigs/Pop2000_{i,t-21}$				11.96 (9.83)
$\Delta Rigs/Pop2000_{i,t-22}$				15.17 (7.97)
$\Delta Rigs/Pop2000_{i,t-23}$				21.29 (8.44)*
$\Delta Rigs/Pop2000_{i,t-24}$				10.58 (9.14)
Lagged controls included	12	12	13	24
$\hat{\beta}(1)$	154.38 (29.35)***	55.15 (13.58)***	160.99 (36.13)***	243.36 (49.13)***
$\widehat{LRM}$	311.09 (76.59)***	115.78 (38.37)***	334.47 (89.43)***	493.97 (106.68)***
BIC	45917.38	45885.50	45905.45	45925.13

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Driscoll-Kraay standard errors in parenthesis. 48 states, 265 months (Feb 1992–Feb 2014), 12,720 observations. Only state-month fixed effects included. Control variables include national log $Employment_{t-1}$ and log IPI_{t-0} with 12, 12, 13 and 24 lags. Tests that $\beta(1) < 0$ and that $LRM < 0$ both rejected at 0.001 level.

Table 10: Per-capita model with differing lag-lengths and controls

	Obs.	N	T	Mean	SD	Min	Median	Max
Land-based rigs	13,920	48	290	22.36	80.23	0.00	0.40	935.50
CES Employment (thou), NSA	13,920	48	290	2,173.02	2,230.86	131.60	1,473.60	13,062.50
CPS Population (millions)	1,104	48	23	5.86	6.27	0.45	4.20	38.04
LAU Population (millions)	13,920	48	290	4.48	4.77	0.33	3.16	29.88
Total rigs	290	1	290	1,170.43	437.39	496.40	1,008.75	2,017.25
CES Employment (thou), NSA	290	1	290	105,677.40	8,452.42	88,010	108,855	116,938
CES Employment in Mfg. thou), NSA	290	1	290	15,170.36	2,188.94	11,340	15,444.5	17,870
CES Employment in Pri. Energy (thou), NSA	290	1	290	304.43	75.58	221.70	272.95	509.50
Total Industrial Production (IP)	290	1	290	84.73	12.78	59.93	88.97	102.31
IP in Mfg - SIC	290	1	290	82.53	13.53	55.96	87.49	103.04
IP in O&G Extraction	290	1	290	103.44	5.89	87.90	102.02	126.38
PPI All commodities (PPIACO)	290	1	290	148.34	30.25	114.10	133.75	205.70
WTI	290	1	290	45.13	30.90	11.35	29.41	133.88
LAU Population (million)	290	1	290	216.56	17.93	188.41	216.58	247.09
CPS Population	23	1	23	283.65	19.83	249.62	284.97	313.91

Time period is Jan 1990 through Feb 2014. Excluded states are Alaska and Hawaii.

Table 11: Summary statistics

References

Arora, Vipin and Jozef Lieskovsky (2014). "Natural Gas and U.S. Economic Activity". In: *The Energy Journal* 35.3.

Black, Dan, Terra McKinnish, and Seth Sanders (2005). "The Economic Impact Of The Coal Boom And Bust". In: *The Economic Journal* 115.503, pp. 449–476.

Blanchard, Olivier Jean and Lawrence F. Katz (1992). "Regional evolutions". In: *Brookings papers on economic activity* 1.1992, pp. 1–61.

Cameron, A. Colin, J.B. Gelbach, and D.L. Miller (2011). "Robust inference with multiway clustering". In: *Journal of Business and Economic Statistics* 29.2, pp. 238–249.

Considine, Timothy et al. (2009). *An emerging giant: Prospects and economic impacts of developing the Marcellus shale natural gas play.*

Considine, Timothy J., Robert Watson, and Seth Blumsack (2010). *The economic impacts of the Pennsylvania Marcellus shale natural gas play: An update.* Funded by Marcellus Shale Gas Coalition.

Driscoll, John C. and Aart C. Kraay (1998). "Consistent covariance matrix estimation with spatially dependent panel data". In: *Review of economics and statistics* 80.4, pp. 549–560.

Fetzer, Thiemo (2014). "Fracking Growth". Working paper.

Han, Chirok, Peter CB Phillips, and Donggyu Sul (2013). "Lag length selection in panel autoregression". In: *Yale University Department of Economics Working Pa-pers.*

Hartley, Peter et al. (2013). *Local Employment Impact from Competing Energy Sources: Shale Gas versus Wind Generation in Texas.* Baker Institute for Public Policy, Rice University.

Higginbotham, Amy et al. (2010). *The economic impact of the natural gas industry and the Marcellus shale development in West Virginia in 2009.* Prepared for West Virginia Oil and Natural Gas Association.

Hooker, Mark A. and Michael M. Knetter (1997). "The effects of military spending on economic activity: evidence from state procurement spending." In: *Journal of Money, Credit & Banking (Ohio State University Press)* 29.3, pp. 400–421.

IHS Global Insight (2011). *The economic and employment contributions of shale gas in the United States*. Prepared for America's Natural Gas Alliance.

Judson, Ruth A. and Ann L. Owen (1999). "Estimating dynamic panel data models: a guide for macroeconomists". In: *Economics letters* 65.1, pp. 9–15.

Kinnaman, Thomas C. (2011). "The economic impact of shale gas extraction: A review of existing studies". In: *Ecological Economics* 70.7, pp. 1243–1249.

Lütkepohl, Helmut (2005). *New Introduction to Multiple Time Series Analysis*. New York: Springer. 792 pp.

Marchand, Joseph (2012). "Local labor market impacts of energy boom-bust-boom in Western Canada". In: *Journal of Urban Economics* 71.1, pp. 165–174.

Murray, Sherry and Teri Ooms (2008). *The Economic Impact of Marcellus Shale in Northeastern Pennsylvania*. Joint Urban Studies Center.

Scott, Loren C. (2009). *The economic impact of the Haynesville shale on the Louisiana economy in 2008*. Loren C. Scott & Associates.

Swift, T. K., M. G. Moore, and E. Sanchez (2011). *Shale Gas and New Petrochemicals Investment: Benefits for the Economy, Jobs, and US Manufacturing*. American Chemistry Council.

Weber, Jeremy G. (2012). "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming". In: *Energy Economics* 34.5, pp. 1580–1588.

— (2014). "A decade of natural gas development: The makings of a resource curse?" In: *Resource and Energy Economics* 37, pp. 168–183.

A   Appendix

A.1   National Data

Summary statistics for national data are in Table 11. All data begin in January 1990 and end in February 2014.

We measure oil prices by using the West Texas Intermediate (WTI) benchmark, which we take from the Energy Information Agency (EIA) website. WTI is the primary North American crude oil benchmark, and it is traded in a very deep and liquid market. We then deflate this price by the Producer Price Index for All Commodities (PPIACO), which we take from the FRED data service of the Federal Reserve Bank.

National rig counts are taken from two spreadsheets on the Baker Hughes website[23]: historical data beginning in 1990 ("North America Rotary Rig Counts through 2013") and 2014 data ("Rigs by State - Current and Historical"). We include both land and offshore rigs.

Employment data are taken from the Bureau of Labor Statistics' (BLS) Current Establishment Survey. We use three measures, which are all in units of 1,000 employees. Data are not seasonally adjusted.

1. Private, Non-Farm Employment (CEU0500000001)

2. Employment in manufacturing (CEU3000000001)

3. Employment in Oil and Gas Extraction plus Oil and Gas Support Activities (CEU1021311201 plus CEU1021100001, which correspond to NAICS codes 211000 and 213112)

Industrial production is taken from the Federal Reserve website.[24] We use three measures, which correspond to the three employment measures. None of the variables are seasonally adjusted.

---

[23]http://phx.corporate-ir.net/phoenix.zhtml?c=79687&p=irol-reportsother.
[24]http://www.federalreserve.gov/releases/g17/ipdisk/ip_nsa.txt.

1. Total Industrial Production (B50001)

2. Industrial Production in Manufacturing as defined by SIC (B00004)

3. Industrial Production in Primary Energy (B53310), which includes natural gas, Louisiana and other crude oil, nuclear electric power generation, coal mining, support activities for oil and gas operations, Texas crude oil, Alaska and California crude oil, and hydro-electric power generation.

A.2   State Data

Summary statistics for state-level data are in Table 11. All data begin in January 1990 and end in February 2014. Land-based rig counts are taken from the Baker Hughes website. Employment data are taken from the BLS Current Establishment Survey State and Metro dataset. As in the national case, we use Private, Non-Farm Employment, which is not seasonally adjusted and measured in thousands.

We work with a variety of population measures. Our primary measure of population is the Census Bureau's midyear population estimate, and we download it from the "Annual State Personal Income and Employment" tables on the Bureau of Economic Analysis website.[25] In particular, we use the population in July 2000 as our base-line population estimate for the sample, though in robustness checks we also use the measures from 1990 and 2010, as well as a linear interpolation for intervening months. Our other measure of population is the BLS Local Area Unemployment (LAU) Statistics series "Civilian noninstitutional population." As defined by the BLS, this panel captures the population of 16 years of age or older who are not institutionalized or in the armed services.[26]

---

[25]http://www.bea.gov/iTable/iTable.cfm?ReqID=9&step=1#reqid=9&step=1&isuri=1.

[26]See http://www.bls.gov/lau/rdscnp16.htm for the exact definition.
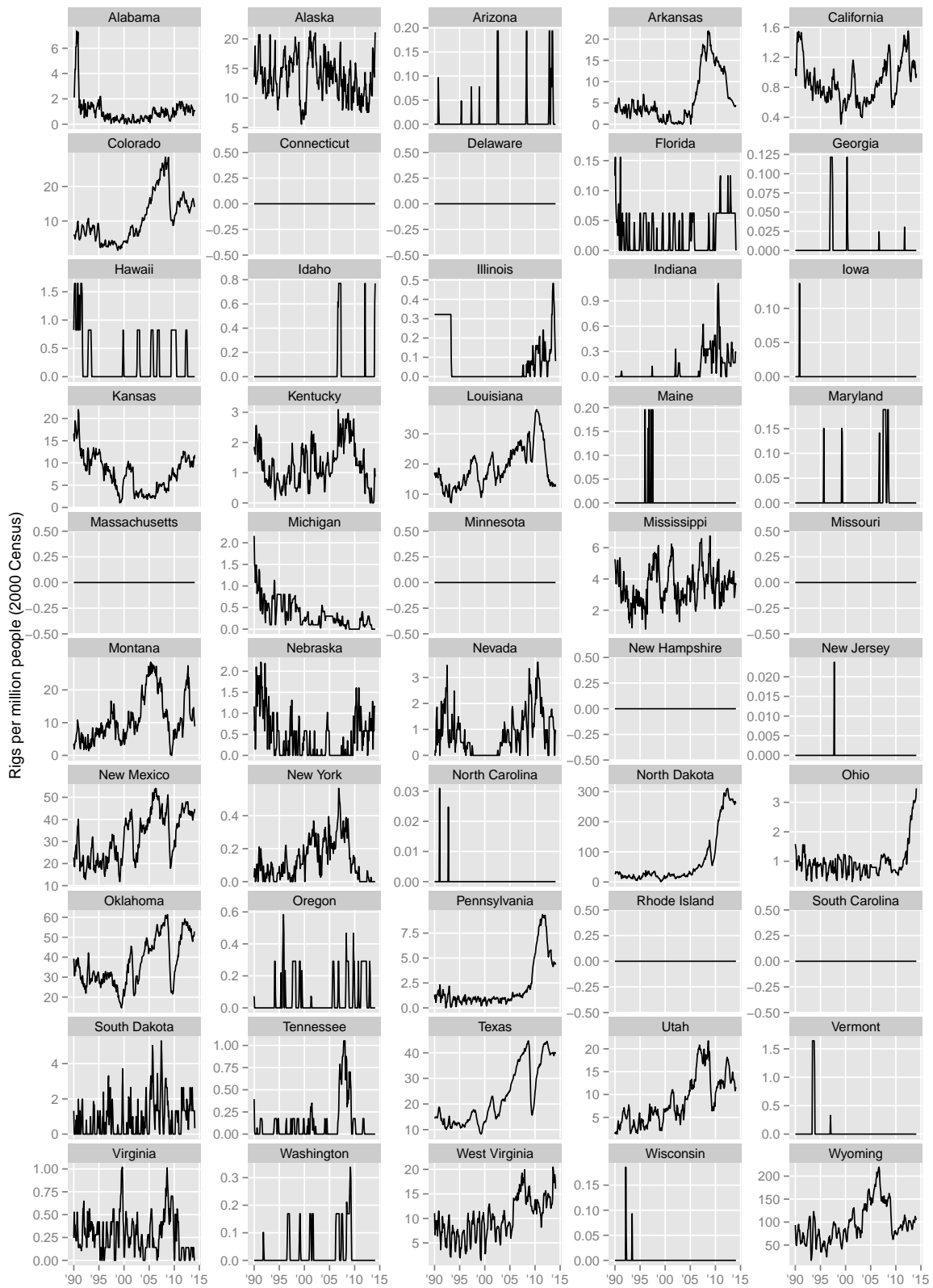
Figure A.1: Shale basins

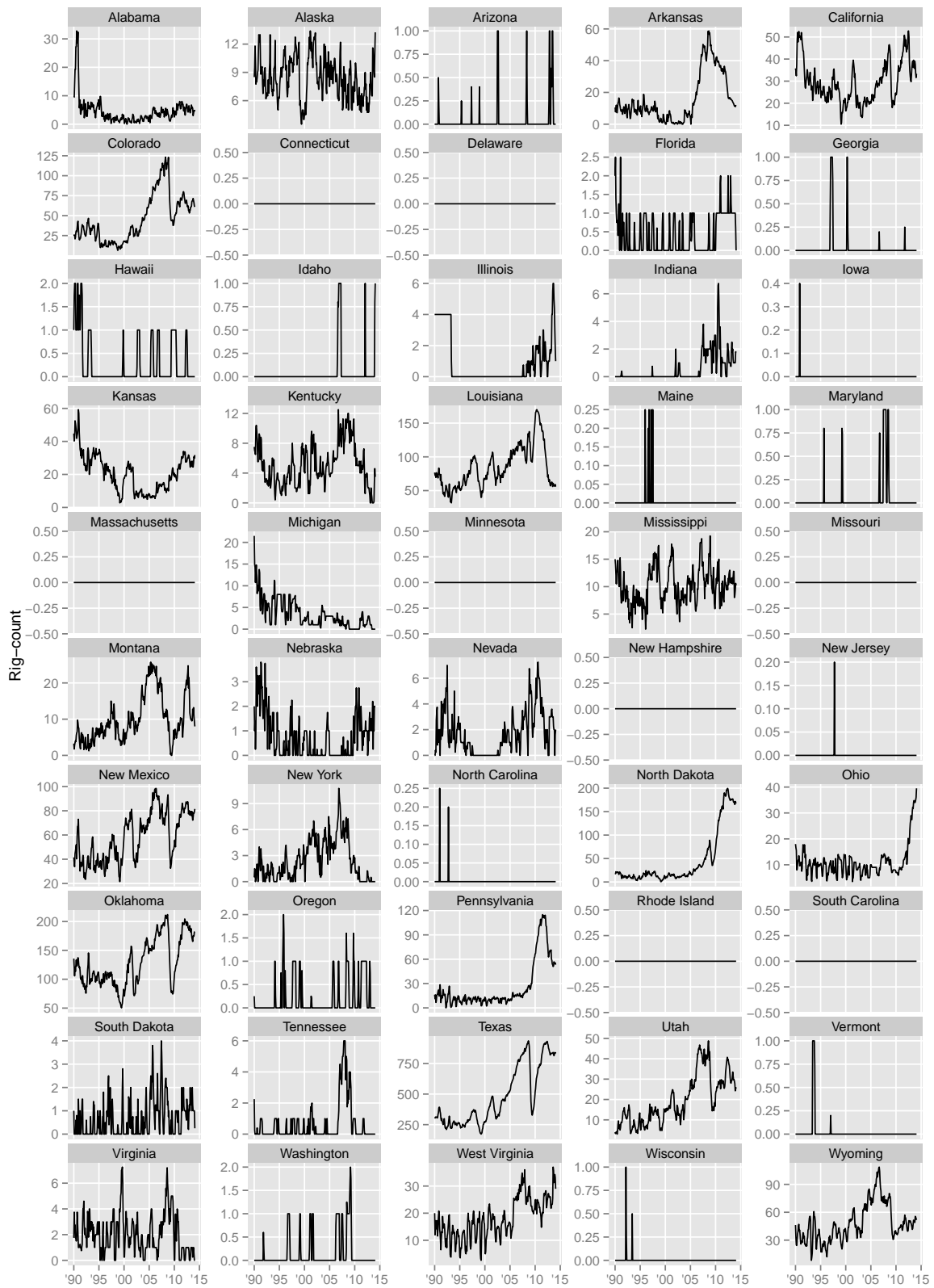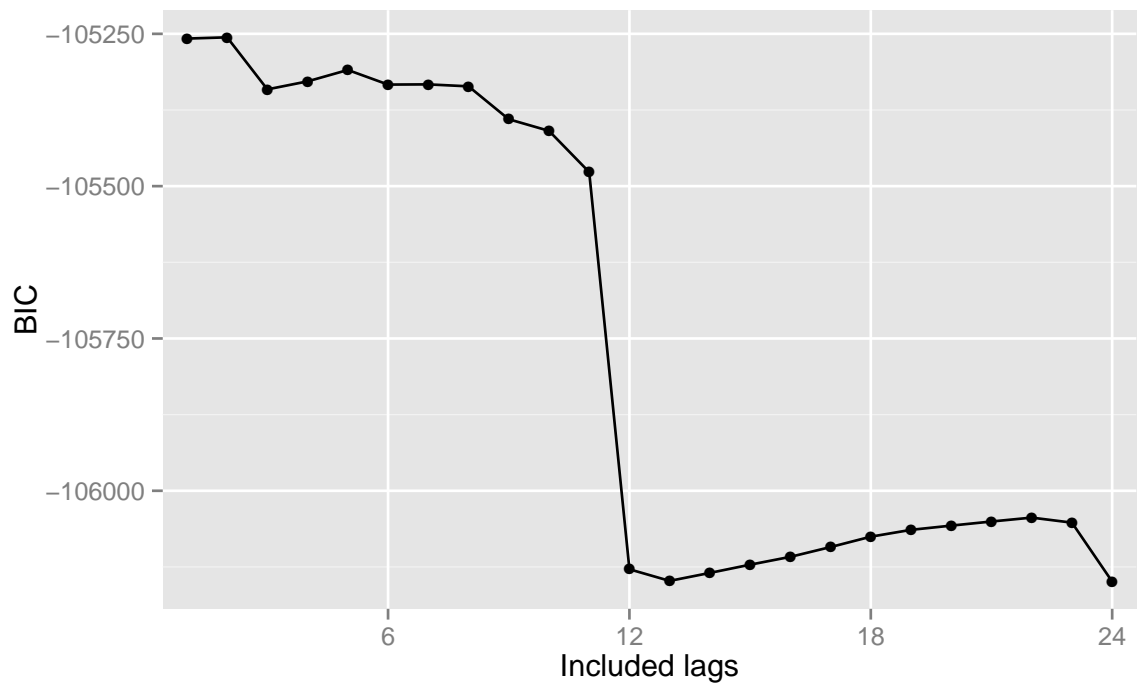Figure A.2: Rig-count per million people

45

Figure A.3: Rig-counts (level)

Figure A.4: BIC for lag-length selection of model (1)