AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Resident vs. Nonresident Employment Associated with Marcellus Shale Development

Douglas H. Wrenn, Timothy W. Kelsey, and Edward C. Jaenicke

There is much debate about the employment effect of shale gas development, especially as it relates to extraction counties. Anecdotal evidence suggests that many of the jobs created are filled by nonresidents. We examine the impact shale gas development has on local employment in Pennsylvania using a data set that links workers to their personal residences. We find that activity in the Marcellus shale has had a modest positive impact on job growth. The impact is cut in half, however, when we use data for county residents only. Thus, traditional employment data may overestimate employment impacts from shale development.

Key Words: employment, Marcellus shale, natural gas, natural resources

Advances in horizontal drilling, hydraulic fracturing, and three-dimensional seismic imaging technologies have made extraction of natural gas from deep shale formations economically viable in the United States.[1] Most of the gas produced comes from the Pennsylvania portion of the Marcellus shale formation, which covers a large portion of the northeastern United States (Rahm et al. 2013). Pennsylvania has experienced a substantial amount of shale gas drilling over the past decade, and the industry has had a significant impact on the state. Anecdotal evidence suggests that the shale gas industry has generated increases in employment and economic activity, large royalty payments, and

[1] The rate of U.S. natural gas production from shale has risen rapidly over the last decade, growing from 4.1 percent of total U.S. production in 2005 to 23.1 percent in 2010 (Wang and Krupnick 2013).

Douglas H. Wrenn is an assistant professor of environmental and resource economics, Timothy W. Kelsey is a professor of agricultural economics, and Edward C. Jaenicke is an associate professor of agricultural economics, all in the Department of Agricultural Economics, Sociology, and Education at Pennsylvania State University. Correspondence: *Timothy W. Kelsey ▪ Department of Agricultural Economics, Sociology, and Agricultural Education ▪ 105-A Armsby ▪ Pennsylvania State University ▪ University Park, PA 16802 ▪ Phone 814.865.9542 ▪ Email tkelsey@psu.edu*.

This project was supported by U.S. Department of Agriculture Cooperative State Research, Education, and Extension Service Hatch projects 4503R and 4567. In addition, we appreciate financial assistance for travel and page charges provided by the National Agricultural and Rural Development Policy Center.

This paper is included as part of the special issue of *Agricultural and Resource Economics* related to the workshop "Unconventional Gas and Oil Development: Economic and Environmental Impacts" organized by the Northeastern Agricultural and Resource Economics Association and held in Morgantown, West Virginia, on June 3 and 4, 2014. The authors and editors acknowledge and thank the National Agricultural and Rural Development Policy Center and the U.S. Department of Agriculture National Institute of Food and Agriculture for funding under competitive grant 2012-70002-1938. Additional support was provide by the Davis College of Agriculture, Natural Resources, and Design's Office of Sponsored Programs and the Regional Research Institute, both of West Virginia University. The views expressed are the authors' and do not necessarily represent the policies or views of any sponsoring agencies.

Copyright 2015 Northeastern Agricultural and Resource Economics Association

migration of workers into Marcellus counties as gas development has expanded. While state-level increases in economic activity associated with the Marcellus are viewed as an economic benefit, some policymakers and researchers have questioned the impact and significance of these economic effects for local residents in terms of economic activity and employment. Unconventional gas development is a highly specialized task in which extraction is coordinated across the state, and many of the companies involved are regional, national, or multinational and have little or no formal connection to the counties in which drilling occurs. Specialized employees often commute between counties, never taking up permanent residence where they work. Consequently, much of the economic impact associated with hiring and industry spending can occur outside of the counties where drilling occurs. Employment changes associated with the Marcellus reflect not only the direct impact of industry spending and hiring but also additional indirect employment generated in local firms via downstream industry spending and individuals spending revenue from natural gas leases and royalties. From a policy perspective, it is important to know how much economic benefit is localized since the shale-gas counties bear most of the economic and social costs of increased gas development.[2]

A majority of recent studies of the impact of Marcellus activity on employment found positive impacts but considerable variation in the magnitudes of the employment effect. In an industry-funded study involving input-output analysis, Considine, Watson, and Blumsack (2010, 2011) found that development of Marcellus shale created more than 44,000 jobs in 2009 and more than 139,000 in 2010. Other studies, however, have questioned those results and found much smaller impacts on employment. Kelsey et al. (2011), using a geographic information system (GIS) analysis and individual surveys to fine-tune an input-output analysis, concluded that the Marcellus created or supported about 23,884 jobs in 2009, including 9,372 jobs directly or indirectly related to industry spending. In a study of shale gas development in Pennsylvania, Weinstein and Partridge (2011) found that increased gas drilling from 2004 to 2010 led to between 10,000 and 20,000 direct, indirect, and induced jobs. Other studies have generated similarly conservative estimates (Brundage et al. 2011, Herzenberg 2011).

Some studies of other regions have examined the effect of natural resource development on employment, including specific investigations at a county-level (e.g., Allcott and Keniston 2014, Brown 2014, Weber 2012, 2014, Marchand 2012, Black, McKinnish, and Sanders 2005). In general, those studies showed modest growth in employment in response to coal and gas booms. For example, Weber (2012) found an increase in annual employment of 1.5 percent for natural gas booms in counties in Colorado, Texas, and Wyoming between 1999 and 2007. Black, McKinnish, and Sanders (2005) found that employment grew 2.0 percent faster in coal-producing counties than in non-producing counties in Kentucky, Ohio, Pennsylvania, and West Virginia between 1970 and 1989.

We extend this previous research and address a number of important issues related to measuring local employment effects from shale gas development in the Marcellus region of Pennsylvania. Considering the wide range of estimates of shale-gas-related employment in the literature, we examine critical empirical

[2] For example, in 2013 in the midst of a (failed) re-election campaign, Governor Tom Corbett claimed that 200,000 Pennsylvanians had jobs or were made more prosperous because of the industry (T. Puko, *Pittsburgh Tribune Review*, November 14, 2013).

issues associated with data and model specification when measuring and interpreting such estimates.

One of the main difficulties in identifying the true impact of gas development on local labor markets is obtaining data that adequately capture its impact on local employment, which we define as employment of residents of the counties in which gas drilling is taking place. The most popular sources of data for measuring changes in county-level employment are the U.S. Bureau of Labor Statistics (BLS) and U.S. Bureau of Economic Analysis (BEA). Two characteristics of the data collected by those agencies can result in inflated estimates of local employment impacts. First, both data sets are based on employer reports and thus do not take employees' place of residence into account, a limitation that is potentially problematic in industries such as shale gas development as many workers in that industry do not reside in the counties in which they work. In addition, full-time and part-time jobs are treated as equivalent so one cannot know whether changes in employment are primarily part-time or full-time jobs. Both data characteristics suggest that many of the jobs that are attributed to shale gas could actually be filled by out-of-county residents or could be part-time, and these types of jobs may not contribute as strongly to the long-run health of the county. Thus, while overall gains in employment may be positive, the impact of those gains on local residents may be more modest.

Furthermore, overall state and national changes in employment from unconventional gas development may have little bearing on the welfare of the counties that most directly bear the costs of increased well activity, making local impacts an important issue for researchers and policymakers (e.g., White 2012). A potentially better source of data for examining local employment effects would account for employee place of residence rather than employer location and identify whether a job is full-time or part-time.

In terms of modeling, accurate estimates of local employment effects require econometric models that can (i) account for spatial and temporal differences in observable and unobservable factors between well and non-well counties and (ii) model the heterogeneous levels of treatment associated with different stages of well activity. While the locations of shale gas deposits are exogenous, the impacts they have on county-level economic activity likely are endogenous and are a product of complex interactions between *ex ante* economic factors and policies, increases in economic activities associated with drilling, and the amount of drilling in each county.[3] To identify the true impact of shale gas activity on local employment, researchers need models that can isolate the treatment effects associated with various degrees of drilling—models that incorporate exogenous variation across space and time in well activity and empirical measures of local employment that are conditioned on those exogenous shifts.

We address each of these concerns to provide a more robust and consistent estimate of the impact of Marcellus shale development on employment of local residents. Our results confirm that the data sets commonly used in measures of employment can overestimate employment impacts on local residents, especially in counties in which there is a significant amount of well activity. We combine a unique set of data on local employment based on the county of residence of each employee with a panel-data difference-in-difference-in-differences (DDD) model that uses well activity at the county level as

[3] We define the level of well activity as the number of wells drilled in a given time period.

a multi-value economic treatment. The employment data come from the Pennsylvania Department of Revenue, and we use the number of resident tax returns from each county that reported gross compensation for 2002 through 2011 as a representation of resident employment in a given year. These data allow us to account for variations in employment between counties and across time within a county.

We combine the local employment data with conventional BEA and BLS measures of county-level employment and estimate a multivariate specification of the standard DDD regression model, which allows us to control for covariance in the error terms across our three specifications for employment (local, BEA, and BLS). It also allows for joint hypothesis tests of the coefficients on our treatment variables, which we use to make statistical comparisons of the impact of well activity on our employment measures. In our DDD model, we measure treatment several ways using different distributions of well activity in the state. We specify both a binary and a multi-value treatment to control for changes in the degree of drilling activity. Because we estimate a panel data model, we can control for time-invariant county-level heterogeneity that may bias our parameter estimates.

Several recent studies have employed methods similar to ours and modeled county-level employment as a function of a treatment variable defined by variation in the resource base along with other covariates as controls. The definition of the treatment counties in these studies has often been based on production data (e.g., Allcott and Keniston 2014, Weber 2012) or the percentage of total earnings from the energy-extraction sector (e.g., Marchand 2012, Black, McKinnish, and Sanders 2005). We use the number of wells drilled in a given year. Many of the studies used only BEA employment data (e.g., Allcott and Keniston 2014, Brown 2014, Weber 2012, 2014, Black, McKinnish, and Sanders 2005), and some investigated only nonmetropolitan counties (e.g., Brown 2014, Weber 2012, 2014).[4]

Our models produce a number of interesting results. First, while Marcellus activity has a positive effect on employment, the effect is statistically significant only for counties in which 90 or more wells were drilled in a given year. This result holds for BEA, BLS, and tax data. We also find that, at low levels of well activity, the coefficients on our treatment-effect variables for all three data sets are statistically equivalent and insignificant. However, at high levels of well activity, the coefficients for the BLS and BEA data sets are positive and statistically different from the coefficients for the tax-based employment measure. Finally, for counties in which there is significant well activity, the employment impacts estimated using the BEA and BLS data are more than twice as large as the impact estimated from the tax data. For example, for treatment counties in which 90 or more wells were drilled in a year, the increase in employment is 1.53 percent when using the tax data, 3.10 percent when using BEA data, and 3.86 percent when using BLS data. While the BEA and BLS estimates are modest, their divergence from the tax-based estimate provides evidence that employment of local residents is not well captured by traditional measures.

[4] In the results section, we provide an explicit connection between our results and this earlier research, especially for Weber (2012) and Weinstein and Partridge (2011), which are closest to ours in terms of data and model specification.

From a policy perspective, if we translate the aforementioned percentages into estimates of the overall employment impact of Marcellus development in Pennsylvania for 2005 through 2011, we find that the BEA and BLS data predict an increase of 18,000 to 20,000 additional jobs while our tax data predict a resident increase of only 7,346 to 9,602 jobs. Our estimates based on household tax data should be taken as lower bounds on the actual aggregate impact of development of the Marcellus on employment, but they suggest that actual employment effects for local residents are far more modest than previously thought.

Data

As noted, we use three sources of data in this study: BLS and BEA employment measures and Pennsylvania state tax revenue filings. The BLS data set comes from the bureau's quarterly Census of Employment and Wages and provides state- and county-level employment measures collected from unemployment and workers compensation reports of employers. Employment is reported by the location of the employer irrespective of where the employees reside. The BEA data rely on the same base data but adjust for employment and wages not covered by these programs. These two data sources provide a good understanding of gross changes in employment among businesses located in Pennsylvania but cannot be used to understand the extent to which county residents are being affected.

An important contribution of this study is the ability to estimate local employment effects more accurately thanks to our third data set, state-level tax information from the Pennsylvania Department of Revenue that reports the gross number of resident tax returns by county. This data set reflects changes observed in the employment of each county's residents regardless of where they work. Employment, as implied by the tax data, is a reflection of the number of tax returns that report positive gross compensation from wages and salaries. In a novel approach, we use the job numbers from this tax data to account for both employment and household location by county. The BLS and BEA data are also based on positive wage and salary values but are calculated by the location of the employer. Our measure uses positive compensation listed in tax returns as a measure of employment but categorizes the information by the location of the employee—the tax return filer—instead of the employer. Finally, because this employment measure is based on the number of tax returns, a person who has two jobs (potentially part-time jobs) is counted as a single employment opportunity.

We use 10 years of data (2002–2011) for Pennsylvania's 67 counties to generate year-over-year county-level measures of percent changes in employment. Following conversion of the raw data to annual percentage changes, we obtain a panel data set of 603 observations for nine years (2003–2011). These data serve as dependent variables in the three equations composing our multivariate regression model.

We also use spatial data on the location of all Marcellus well activity in Pennsylvania for 2005 through 2011 (there was no activity before 2005). These data come from the Spud reports (mandatory reports by oil and gas companies of initiation of drilling) made to the Pennsylvania Department of Environmental Protection. The data include total number of wells initiated per county per year. We assign each county to one of four categories of drilling

activity that we identified from natural breaks observed in the annual number of Marcellus wells drilled: high-activity counties exceeded the 90th percentile for number of wells (90 or more), moderate-activity counties had between the 50th and 90th percentiles (10–89 wells), low-activity counties had 1–9 wells, and no-activity counties had none. We use these breaks to generate the treatment variables. Table 1 shows the total number of Marcellus shale wells in Pennsylvania by year.

Additional data from the U.S. Census Bureau provides GIS and population information. Descriptive data for the variables in our econometric model are shown in Table 2.

The study period best reflects the changes that occurred in the Marcellus counties as gas development progressed. Development of the Marcellus shale began slowly but increased rapidly in 2007 and 2008. Thus, beginning with 2003 data allows us to establish pre-existing employment trends in well and non-well counties. We use these pre-existing trends and the treatment effects from different levels of well activity to identify the impact of Marcellus development on our three measures of employment.

Table 1. Pennsylvania Marcellus Shale Wells Drilled by Year

Year	No. of Wells
2002	0
2003	0
2004	0
2005	8
2006	37
2007	116
2008	332
2009	817
2010	1,602
2011	1,961
Total	4,873

Notes: Well data are from the Pennsylvania Department of Environmental Protection's Office of Oil and Gas Management.

Table 2. Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Dependent Variables				
Pennsylvania Dept. of Revenue data	0.00	0.02	–0.14	0.07
Bureau of Economic Analysis data	0.01	0.02	–0.11	0.14
Bureau of Labor Statistics data	0.00	0.03	–0.17	0.20
Treatment Variables				
Model 1 treatment 1-1: any wells	0.24	0.43	0.00	1.00
Model 2 treatment 2-1: 10 or more wells	0.10	0.30	0.00	1.00
Model 3 treatment 3-1: 1–9 wells	0.14	0.34	0.00	1.00
Model 3 treatment 3-2: 10–89 wells	0.08	0.27	0.00	1.00
Model 3 treatment 3-3: 90 or more wells	0.03	0.16	0.00	1.00
County Controls				
Population (in thousands)	186.77	263.20	5.09	1,526.01
Well county trend (in thousands)	2,224.45	2,004.68	0.00	4,044.12

Notes: The descriptive statistics for all of the fixed effects have been suppressed.

Econometric Model

We employ a simple reduced-form econometric model in which the percentage change in the employment in each county i from one year to the next is a function of a county-specific fixed effect represented by α_i, an average-year effect represented by γ_t, a set of lagged county-level variables represented by X_{it-1}, an indicator variable I_{it} for whether the county received a positive treatment in the form of Marcellus well activity during the year, and a normally distributed idiosyncratic error term ε_{it}:

$$PCT\Delta EMP^{d}_{it,t-1} = \alpha_i + \gamma_t + X'_{it-1}\beta + I_{it}\delta + \varepsilon_{it}. \quad (1)$$

The superscript d demonstrates that the dependent variable in each equation varies by the source of employment data: $d \in \{tax, bea, bls\}$. Since each dependent variable measures the same outcome, we stack the equations for the data sources and estimate this equation as a multivariate regression that accounts for unobservable correlation in the error terms from the three data sources.

In this empirical specification, the coefficient of interest, δ, captures the impact of Marcellus well activity on the change in employment. At this point, we specify the variable for the treatment effect only broadly to capture the basic intuition. We later describe in detail how this indicator varies based on the number of wells drilled each year in a county.

A primary challenge in identifying the causal impact of Marcellus well activity on employment is that the researcher can rarely construct a clear experimental design. For example, Marcellus well activity in Pennsylvania began in late 2005 and 2006 but only in a few counties. Growth in these early years was slow, and significant drilling activity did not take off until 2007 and 2008. While the location of the Marcellus gas deposit is exogenously determined, the slow ramping up of Marcellus activity, which is shown in Table 1, makes empirical identification of a binary treatment effect of well activity difficult, especially given the panel nature of our employment and well data.[5] To account for the lack of a clear binary treatment effect and the fact that the level of well activity in a county is likely to be closely tied to county employment, we estimate the model using both a binary and a multi-value treatment and allow the treatments to vary depending on the model specification and amount of well activity in the county (Imbens and Wooldridge 2009). To be consistent with the intuition of the literature on treatment effects, we assign each county to the treatment or control group based on whether there was any shale gas development in the county during the study period. Then, we specify a series of treatment levels according to the amount of drilling activity, which is the number of wells drilled per year. Instead of treating the variables with a binary, homogeneous shale-gas policy that stays constant over time, we assume that the level of shale-gas activity influences the level of treatment for counties in the treatment group, and we estimate the impact of those levels of treatment on our employment measures.

We apply three empirical treatment-effect specifications based on the empirical distribution of well activity in the state for counties with shale gas

[5] Most empirical models that employ a treatment-effect strategy to test the impact of particular policies are based on policy environments in which there has been a clear binary change in a policy that has affected a well-defined, exogenously determined subset of the population and thus a clearly distinguishable before and after.

activity. In model 1, we specify a simple binary variable (treatment 1-1) that is positive for counties that have any well activity in a given year and zero otherwise. This specification mimics the intuition of the binary treatment model and assumes that the critical threshold for well activity is any activity after 2005. In model 2, we again specify a binary cut-off variable (treatment 2-1), but this variable is positive for counties and years in which 10 or more wells were drilled where 10 wells is the median level of well activity during the study period. In model 3, we assign levels of well activity to three discrete bins: treatment 3-1 is defined as counties with fewer than 10 wells per year ($wells < 10$), treatment 3-2 is counties with 10–89 wells per year ($10 \leq wells < 90$), and treatment 3-3 is counties with 90 or more wells per year ($wells \geq 90$). The bins are based on the 50th and 90th percentiles from the distribution of well activity and are intended to account for structural breaks in the level of treatment observed for various levels of well activity.[6,7]

This empirical specification has been applied elsewhere to analyze the long-run impact of spatially and temporally varying policy instruments (Banzhaf and Lavery 2010, Banzhaf and Walsh 2008). Our specification, which uses multi-value treatments, differs from those previous works but the basic model is the same. By including county-level fixed effects (α_i) in a model that has a differenced dependent variable, we empirically specify a DDD model—more specifically, a difference-in-differences model in growth rates instead of levels.[8] In our model, the impact of a policy change over time is identified by comparing historical trends to a baseline control outcome. The treatment effect (δ) of a positive increase in Marcellus activity in county i in period t is estimated from the percent change in the dependent variable from the previous year relative to county i's employment trend in the pretreatment period and relative to those changes in other counties.[9] We use the three data sets to identify the effect of Marcellus activity on employment during periods in those counties following a positive Marcellus treatment. The baseline (pre-existing trend in employment)

[6] We also used the number of wells as a continuous treatment variable with its square to allow for nonlinear impacts but chose not to present the results because there is discontinuity in the data. All counties in the pretreatment period have zero wells; in the post-treatment period, more than 70 percent of the counties have zero wells and the counties that have wells mostly have either a very small number or a very large number. Despite the discontinuity, the results from this model are generally consistent with the results presented.

[7] As a robustness check, we estimated the multi-value treatment-effect model with other thresholds. The results were nearly identical to the results for the 50th and 90th percentiles.

[8] We thank an anonymous reviewer for pointing out that our DDD terminology is the same as saying that we are estimating a difference-in-differences model in growth rates instead of levels. Given the close empirical connection between our models and those of Banzhaf and Lavery (2010), we continue to use DDD terminology but this key distinction should be kept in mind.

[9] Treatment-effect models rely on the stable-unit treatment-value assumption (SUTVA), which requires that a treatment cannot affect the nontreated (Imbens and Wooldridge 2009). An anonymous reviewer noted the valid concern that employment gains in shale-well counties could come at the expense of nearby non-well counties. In the Pennsylvania Marcellus, two cases are pertinent: out-of-county workers who do not move to the shale-well county and out-of-county workers who do move. The first case is not a significant problem because the Pennsylvania tax data will identify any positive employment treatment effects in untreated counties as a weaker effect on treated counties. Hence, our estimates would understate the true effect in treated counties. The second case, in which workers move to Marcellus counties, could cause our models to overstate the employment impact of Marcellus activity. However, we have searched documented evidence of migrations of workers from non-Marcellus to Marcellus counties in Pennsylvania and could find none. Typically, studies have found that workers come from other states. For example, Brasier et al. (2011) reported that most jobs directly associated with the industry were filled by out-of-state crews. We conclude, therefore, that the potential for violation of the SUTVA is small.

for each county is established by pretreatment-period outcomes and county-level fixed effects that control for unobservables that affect the rate of change in employment.[10] They also control for any historical differences between counties generated by past development of conventional gas reservoirs.

While the location of Marcellus gas is exogenous, pre-existing state, regional, and county trends could have effects on changes in employment that are unrelated to increased shale gas activity. To control for these potential impacts, we include additional controls in our model. To control for any macro-level changes in economic conditions, such as the global financial crisis of 2008–2009, we add a full set of year fixed effects that control for any intertemporal shocks that affect all Pennsylvania counties symmetrically. Counties also could have experienced increases in shale gas activity that were randomly correlated in space and time with other factors that impact county-level employment. For example, many of the counties that experienced the largest increases in shale activity were among the poorest and most rural in the state, and other regional policies may have been enacted at the time to positively impact employment for reasons not explained by county-level fixed effects. To account for such simultaneous effects, we interact each year fixed effect with the county's latitude, longitude, and interaction between its latitude and longitude. We thus create an entire spatial surface for each year that represents annual deviations for each county from its long-run trend. Identification of the impact of Marcellus activity is based on statistically significant deviations for treated counties away from the additional year-specific spatial deviations. Long-run historical trends for Marcellus and non-Marcellus counties could be different. Some of the Marcellus counties, for example, have experienced conventional (vertical) gas well development. To account for these potential trends, we add a quadratic Marcellus-specific time trend (*WellCntyTrend*) to the model to capture serially correlated time trends that are specific to Marcellus counties. Finally, to control for the size of each county in terms of population, we include lagged and logged yearly population measures (*LogPopulation*) that are based on 2000 and 2010 census projections and change over time.

Results

Before presenting the results from the DDD model, we address empirical issues associated with implementing a treatment-effect panel data model with a first-differenced dependent variable. First, multi-period treatment-effect models introduce the potential for serial correlation in the dependent variable that is not an issue with two-period, two-group binary treatment models. Bertrand, Duflo, and Mullainathan (2004) demonstrated that, as the number of time periods increases before and after treatment, the likelihood that serial correlation in the dependent variable will affect the parameter estimates for the treatment increases.[11] This is particularly true for economic variables such

[10] Using a method similar to ours, Weinstein and Partridge (2011) applied a difference-in-differences model using BLS labor data and county-level well activity and found no difference in employment changes for well and non-well counties. However, because they estimated a difference-in-differences model, they could not account for pre-existing trends. Weber (2012) also used a similar method and specified a difference-in-differences model that was based on changes in the level of employment at a county level.

[11] We thank an anonymous referee for pointing out that reintroducing panel fixed effects in a first-differenced model could cause serial correlation between the error terms within panels.

as taxes and wages, which are usually positively correlated in time. Bertrand, Duflo, and Mullainathan (2004) showed that positive serial correlation could lead to understated standard errors and inflated t-statistics.

To test for serial correlation, we apply the Harris-Tzavalis panel-data unit-root test (Harris and Tzavalis 1999), which is designed for panels that have a large N and small T; it has been shown to be robust for values of N greater than 25. To control for cross-sectional correlation, we perform the unit root tests on our data by removing cross-sectional means. Application of this test to our data reveals no indication of serial correlation, likely because our dependent variable is already in first-differenced form, which removes any time trend from the data. Bertrand, Duflo, and Mullainathan (2004) made several recommendations for controlling serial correlation in the dependent variable in a multi-period treatment-effect model. However, their recommendations were based on the assumption that the dependent variable had not been adjusted. Because we specify the dependent variable in percentage terms, we are effectively first-differencing the variable, which likely removes any serial correlation. To confirm this, we also tested for serial correlation in the original variables for level of employment and could not reject the null hypothesis of serial correlation, which indicates that differencing the original variables successfully removed serial correlation from our employment measures.[12]

Next, we address the issue of unequal pre-existing trends for the treatment and control groups. The validity of the result of any treatment-effect model is conditional on the assumption that the underlying trend in the outcome variable is the same for the treatment and the control in the pretreatment period. In a two-period difference-in-differences model, this assumption is not testable; with only two observations, the researcher can never determine a pre-existing trend. However, with more than two observations, one can plot the pre-existing trends to compare the pretreatment trajectories in the dependent variables for the treatment and control groups. Figure 1 plots the pretreatment-period trajectories and their respective trend lines and slopes for the county tax data for counties with (treatment) and without (control) well activity.[13] While we have only a two-year pretreatment period, the results demonstrate that the pretreatment trends for well and non-well counties are similar based on the slope of the lines.[14] Thus, we can use these data to compare the treatment and control groups in the post-treatment period.

We report the results of the models in Table 3. The coefficients represent the relative percent change in employment for a change in the covariate. For the treatment variables (δ), the coefficients represent percent changes for the effects relative to the baseline employment trend in the county and to the baseline for control counties. All of the models control for county, year, and spatial effects.

Overall, we find that Marcellus activity has had a positive but modest impact on employment; all but one of the treatment variables have positive signs.

---

[12] We also applied an Im-Pesaran-Shin (Im, Pesaran, and Shin 2003) unit root test, which allows for different serial correlation parameters across panels, and obtained similar results. We do not include the full set of results from these tests to conserve space, but they are available upon request.

[13] We plot results only for tax data to converse space; the BLS and BEA results have similar trends.

[14] Note that it is not necessary for the pretreatment means of the treatment and control groups to be equal. All that is required is that their respective trends follow a similar path.

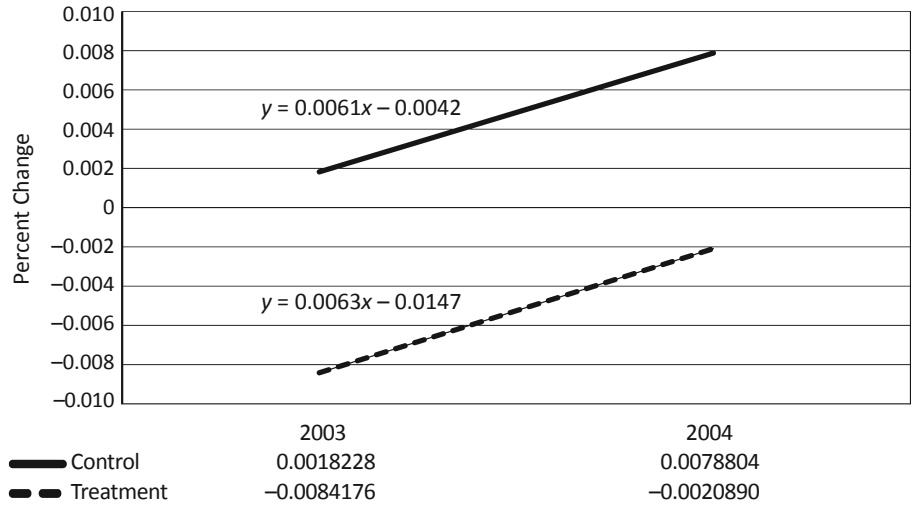

Figure 1. Pretreatment Trends for the County Tax Data

However, the positive effect is statistically significant only when the treatment involves a substantial amount of well activity in a county. For example, treatment 1-1 in model 1, which is based on whether a county has any well activity, has no statistically significant impact on employment relative to long-run trends in the treatment and control counties. The result is the same for each data source.

For treatment 2-1 in model 2, however, in which the treatment is 10 wells or more, we find that the impact of well activity is positive and significant for all data sets. The impact of well activity ranges from a 0.59 percent employment increase for the Pennsylvania tax data to a 1.78 percent increase for the BLS data. We also observe that the coefficients are significantly higher for the BLS and BEA data than for the tax data—the BLS values are almost three times greater. To the extent, then, that the BLS and BEA data misrepresent local employment levels, the tax data set provides more consistent and conservative estimates of the impact of Marcellus activity on employment of residents in each county.

Model 3 estimates the employment effect with the multi-value treatment variable. Once again, we find that a small amount of well activity (treatment 3-1) does not have a significant impact on employment while a large amount of well activity does impact employment. However, while use of the BEA data generates significant employment effects under treatments 3-2 and 3-3, the tax and BLS effects are significant only for high Marcellus well activity (90 or more wells). Moreover, the coefficient of 1.53 percent under treatment 3-3 for the tax data is less than half of the coefficients for the BLS (3.86 percent) and BEA (3.10 percent) data. The extent to which the tax data provides a better representation of employment trends for local residents has significant implications for policymakers who have used BLS and BEA data to evaluate the local employment impacts of Marcellus activity. Caution is advised in assigning causality in such evaluations when using these conventional data sources.

It is important to subject the results to joint hypothesis tests that can account for covariance in the error terms across the data sources and equations. Because

Table 3. Results for Difference-in-difference-in-differences Models

	Data Source for Dependent Variable					
	Income Tax Data		BEA Data		BLS Data	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Model 1						
Treatment 1-1 (any wells)	0.0005	0.0025	0.0005	0.0028	–0.0046	0.00428
LogPopulation	–0.0981 **	0.0272	–0.2034 **	0.0306	–0.2700 **	0.046159
WellCntyTrend	1.6E-07	1.8E-07	3.2E-07	0.0000	4.9E-07	3.0E-07
R-square	0.64		0.61		0.48	
Model 2						
Treatment 2-1 (10 or more wells)	0.0059 **	0.0029	0.0098 **	0.0033	0.0178 **	0.0050
LogPopulation	–0.0947 **	0.0271	–0.1974 **	0.0303	–0.2559 **	0.0457
WellCntyTrend	1.1E-07	1.7E-07	2.3E-07	1.9E-07	1.7E-07	2.8E-07
R-square	0.65		0.61		0.64	
Model 3						
Treatment 3-1 (1–9 wells)	0.0006	0.0025	0.0010	0.0028	–0.0044	0.0042
Treatment 3-2 (10–89 wells)	0.0050	0.0038	0.0073 *	0.0042	0.0097	0.0063
Treatment 3-3 (90 or more wells)	0.0153 **	0.0057	0.0310 **	0.0063	0.0386 **	0.0095
LogPopulation	–0.0942 **	0.0270	–0.1967 **	0.0298	–0.2581**	0.0452
WellCntyTrend	1.0E-07	1.8E-07	2.3E-07	2.0E-07	3.2E-07	3.0E-07
R-square	0.65		0.63		0.50	

Notes: All models include county and time fixed effects. * Significant at a 10 percent level. ** Significant at a 5 percent level.

we estimate the DDD model as a multivariate regression, we can perform joint hypothesis tests (F-tests) on the results for coefficient equivalence across the data sets. Specifically, we want to test the hypothesis that the coefficients on the treatment-effect variables in the model estimated using income-tax data are statistically different from the coefficients in the models that use BLS and BEA data. We present the results of these tests in Table 4, which provides F-statistics and p-values. As expected, the low-activity treatment generates statistically equivalent coefficients. However, the results from our joint hypothesis tests of the high-activity treatment demonstrate that the differences in the coefficients are significantly different from zero, and the coefficients for the BLS and BEA

data are larger than the coefficient estimated for the income-tax data. These results provide further evidence that using traditional data sources to gauge employment in high-activity counties may overestimate the impact of shale gas development on local employment. Moreover, the results indicate that the deficit in those data sources may be important not just for Marcellus development but for other regions and states in which there is significant development of unconventional energy sources. Our results suggest that, as the drilling activity expands, traditional data sources will be increasingly unreliable for estimating local employment precisely.

Robustness Checks and Relationship to Prior Studies

Our results demonstrate that shale gas development in the Marcellus has had a modest positive impact on employment in Pennsylvania only in high-activity counties and that a significant amount of any increase in employment from its development likely comes primarily from nonresidents and/or workers temporarily relocating to those counties. These results are interesting from both academic and policy perspectives, but some concern remains that they may be driven by specification of our treatment variables and/or by the counties used in the models. Therefore, we subject the estimates to several robustness checks. One of these models allows us to draw connections between our study and prior studies of the impact of shale gas development on employment.

Given the more comprehensive nature of the multi-value treatment, we present the results of the robustness checks for that model only (see Table 5). In the first specification, we alter the well-count cut-offs used to create the multi-value treatment to the 33rd (5–124 wells) and 93rd (125 or more wells) percentiles to determine whether the cut-offs we used in the models (50 and 90 or more wells) influenced the results. This change effectively expands the middle-activity group and contracts the low-activity and high-activity groups.

Table 4. Hypothesis Tests of Coefficient Equality

	(Tax – BEA) = 0		(Tax – BLS) = 0	
	F-statistic	**p-value**	**F-statistic**	**p-value**
Model 1				
Treatment 1-1 (any wells)	0.01	0.9952	1.17	0.2792
Model 2				
Treatment 2-1 (10 or more wells)	4.23 *	0.0403	4.70 **	0.0307
Model 3				
Treatment 3-1 (1–9 wells)	0.01	0.9248	1.13	0.2878
Treatment 3-2 (10–89 wells)	0.20	0.6583	0.43	0.5115
Treatment 3-3 (90 or more wells)	4.02 **	0.0454	4.81 **	0.0288

Notes: This table presents results from hypothesis tests of the equivalence of the coefficients from the tax model with the coefficients from the models using BEA and BLS data. These tests account for the covariance between the error terms across the data sets. * Significant at a 10 percent level. ** Significant at a 5 percent level.

Table 5. Results of the Robustness Checks

	Data Source for Dependent Variable					
	Income Tax Data		BEA Data		BLS Data	
Specification	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**	**Coef.**	**Std. Err.**
Change in Alternative Treatment Cut-off						
Treatment 3-1 (1–4 wells)	–0.0001	0.0026	0.0000	0.0029	–0.0065	0.0044
Treatment 3-2 (5–124 wells)	0.0033	0.0033	0.0052	0.0036	0.0051	0.0055
Treatment 3-3 (125 or more wells)	0.0133 **	0.0064	0.0342 **	0.0070	0.0425 **	0.0106
F-Statistic			5.80 **		6.13 **	
p-value			(0.0163)		(0.0136)	
Dry Gas Counties Only						
Treatment 3-1 (1–9 wells)	–0.0019	0.0027	0.0021	0.0030	–0.0031	0.0046
Treatment 3-2 (10–89 wells)	0.0047	0.0040	0.0084 *	0.0045	0.0112	0.0068
Treatment 3-3 (90 or more wells)	0.0130 **	0.0058	0.0328 **	0.0066	0.0407 **	0.0100
F-Statistic			6.10 **		6.35 **	
p-value			(0.0139)		(0.0121)	
Small Counties Only – 90th Percentile or Less						
Treatment 3-1 (1–9 wells)	0.0009	0.0027	0.0005	0.0030	–0.0050	0.0045
Treatment 3-2 (10–89 wells)	0.0062	0.0040	0.0073 *	0.0044	0.0101	0.0067
Treatment 3-3 (90 or more wells)	0.0178 **	0.0061	0.0315 **	0.0068	0.0405 **	0.0120
F-Statistic			2.81 *		3.90 **	
p-value			(0.0963)		(0.0490)	
Absolute Change in Employment – Change in Level of Employment as Dependent Variable for Small Counties Only						
Treatment 3-1 (1–9 wells)	142	89	–36	123	–116	115
Treatment 3-2 (10–89 wells)	131	135	49	186	99	174
Treatment 3-3 (90 or more wells)	494 **	213	1,196 **	293	1,042 **	274
F-Statistic			5.57 **		3.54 *	
p-value			(0.0187)		(0.0605)	

Notes: The F-statistics and p-values in the BEA and BLS columns are for joint hypothesis tests of coefficient equality between the tax data estimates and the BEA and BLS data estimates for treatment 3-3.
* Significant at a 10 percent level. ** Significant at a 5 percent level.

The estimates from this model follow a pattern that is similar to the pattern for our primary models and pass both of the joint hypothesis tests for treatment 3.[15]

In the second specification, we drop what are known as "wet gas" counties and use only "dry gas" counties. Dry gas is mostly methane when it comes out of the well; wet gas contains, in addition to methane, other marketable gases such as ethane, butane, and propane that can be separated and sold. For counties in which a significant portion of the gas produced is dry, the price of natural gas has a significant impact on extraction: the profitability of new natural gas wells falls with the price. This impact tends to be muted in wet gas counties since the other gas products have value. Thus, we expect the price of natural gas to affect employment in counties that mainly produce dry gas more than counties that mainly produce wet gas. We identify each county as wet or dry according to the median level of each type of gas it produces. Once again, the results are consistent with our primary model results (Table 3); the coefficients from the tax data model are less than half of the coefficients from the BEA and BLS data models. These results provide further evidence that interactions between the price of natural gas (a macro variable) and the type of gas produced in a county (a local effect) do not drive our overall results.

In the third robustness check, we address concerns about use of large versus small counties in terms of population and how our treatments affect our outcome variable, which is specified in percentage terms. For example, in terms of the actual number of jobs created, the effect of 90 or more wells drilled in a year is likely similar regardless of the population of the counties in which they were drilled but could have a very different impact in terms of the percent change in employment in a county. Moreover, as noted in Weber (2012), counties characterized by urban centers and relatively large populations could exert excessive influence on the regression results. We follow Weber (2012) and drop all of the counties with populations that exceed the 90th percentile. This effectively removes all of the counties in and around large urban centers and assures that our results are not driven by a few highly populated counties. It also greatly reduces the number of observations in the data set.

In this case, the basic pattern of our primary model holds and the results of the joint hypothesis tests of coefficients in the third treatment remain statistically different from zero. We find, however, that the coefficient in treatment 3 when using tax data increases from 1.5 percent to 2.0 percent, which suggests that some of the results from the tax-data model may be driven by high-populaton counties.

In the fourth specification, we again drop the highly populated counties and then replace the original dependent variable, a percentage term, with the absolute change in employment at a county level. This model not only serves as a robustness check for specification of our dependent variable but also allows us to draw direct comparisons between our study and prior studies related to employment and shale gas development.

---

[15] We also estimated a model that used number of wells per capita as the variable of interest. We thank an anonymous reviewer for pointing out that 90 wells drilled in a county of 1,000 people could have a much larger impact than the same number of wells drilled in a county of one million people. The results from this model, while varying in magnitude from the coefficients from the treatment-effect models, have similar patterns. The estimates using BLS and BEA data are larger and statistically different from estimates using the income-tax data. These results are available upon request.

The results from this model once again demonstrate that estimates based on BEA and BLS data are two times larger than estimates based on tax data. In addition, the BEA data estimate (1,196) is similar to the estimate of 1,780 found by Weber (2012) in a "boom" county treatment (in the top 20 percent in terms of change in actual gas production). Our treatment is based on number of wells drilled rather than gas production, but if we assume that our third treatment (90 or more wells drilled) is a close approximation for counties in the top 20 percent in terms of production, our BEA data estimate is close to Weber's. We also estimate this fourth specification using the alternative well cut-offs of 5–124 and 125 or more. The estimate from this variation of the model using BEA data is 1,543, which is closer still to Weber's (2012) estimate, providing further evidence of the consistency of the results of our model with results from prior studies.

Analysis of Actual Employment

Because our dependent variable is specified as a percentage, it is difficult to make a direct comparison of our results and results from previous studies. We therefore use the coefficients from the treatment variables in model 3 to generate predictions from the original raw employment data.

To convert our percentage results into numbers of jobs, we take the mean values for employment in the pretreatment period (2002–2004) for each county (and for each of the data sets) as a baseline and the coefficient values from treatment 3 in model 3 (90 or more wells), the only treatment that is statistically different from zero and significant for all of the data sets. We multiply each coefficient by the baseline average employment value for the high-activity counties. We then sum the resulting estimated job numbers for all of the years and counties in treatment group 3 and use this final value as our estimate of statewide employment gains associated with Marcellus development during the study period. As with the DDD model, the pretreatment period establishes the trend, and a positive significant treatment in the post-treatment period generates the prediction.

The results from this process using our Pennsylvania tax data show that from 2005 to 2011 Pennsylvania counties experienced, on average, an increase of 7,346 jobs above the baseline trend that would have existed in the absence of Marcellus shale development. We also generate this same employment figure using the coefficient (2 percent) from the third section of Table 5, which uses only small counties; the resulting estimate was 9,602 jobs for 2005–2011. Thus, based on the tax data, Marcellus shale gas development increased the aggregate employment of local residents by 7,346–9,602 jobs. The estimates based on BLS and BEA data are much larger—18,761 and 20,385 jobs on average, respectively. Our BLS predictions are in line with Weinstein and Partridge (2011), which found that employment increased by about 20,000 jobs between 2004 and 2010. Once again, using data that identifies local residents, we find significantly less job growth associated with the Marcellus than prior estimates.

Conclusions and Implications

Expectations of local economic growth from development of the Marcellus shale offset many concerns associated with it. The question, then, is whether the communities that bear the costs of gas extraction are also receiving the benefits,

which include increased employment—are jobs created by development of the Marcellus going to local residents and, if so, how large is the overall impact of those jobs? We take up both questions. We first estimate an empirical model that allows us to evaluate statistically the impact of different levels of Marcellus activity based on two conventional and one new, locally focused measure of employment. Then, using the coefficients from the analysis, we translate the results into estimates of actual employment gains from Marcellus development. We are particularly interested in whether the conventional sources of information on employment may, in empirical models, overestimate gains in employment for local residents, an issue of interest to community leaders and policymakers responsible for regulating shale gas development.

We find that development of the Marcellus has had only a modest impact on overall employment and that the local employment impact is likely to be even smaller. Based on the mean levels of employment in the pretreatment period in our raw data, a 3.86 percent increase in employment (based on BLS data) from 2005 to 2011 translates to a statewide employment increase of 18,761.[16]

When we account for the likelihood that a significant portion of that employment is attributable to individuals who do not live in the counties in which they work, the increase in employment of local residents based on the 1.53 percent coefficient from our analysis using tax data is only 7,346. The difference likely reflects the number of nonresident workers, on average, who commute to counties where drilling is occurring. Note, however, that our analysis does not allow us to determine how many of those commuters reside in other states or are in Pennsylvania counties in which there is no drilling activity. Thus, it is best to consider these results as bounds on the total employment impact of the Marcellus with the tax data estimates being an approximate lower bound on total employment gained solely by Pennsylvania residents and the BLS and BEA data estimates being an approximate upper bound that includes residents of Pennsylvania and other states.

Our employment estimates based on the BLS and BEA data are nearly identical to estimates by Weinstein and Partridge (2011) and are similar in scale to the results of other studies (Brundage et al. 2011, Herzenberg 2011, Kelsey et al. 2011). Moreover, when we analyze absolute changes in the number of jobs created, our estimates are similar to other studies of the impact of shale gas development on employment (Weber 2012). By combining our results with those of prior studies, we establish a consistent baseline by which to compare the results from our analysis of locally oriented tax data. Assuming that the tax data provide a measure of employment that is comparable to traditional sources of employment data, our results demonstrate that local impacts of gas shale development on employment are minimal. This result is consistent with anecdotal reports from residents and policymakers that many shale gas jobs are filled by workers from other counties and states.

Our results provide broad insight into the implications of using conventional data sources when assessing local employment impacts of shale development. When a significant number of workers commute to jobs within an industry, conventional data sources based on employer records are likely to lead to overestimation of employment of local residents. In most cases, the geographic component of employer versus employee location is minimal. However, in the case of the Marcellus, many of the jobs created are regional, transitory, and

[16] The same calculation using the BEA data produced an estimate of a 20,385 increase in jobs.

filled by nonresident workers, making the potential for such overestimation significant with important political and policy implications. The state of Pennsylvania currently neither promotes nor subsidizes gas industries, but politicians and policymakers have argued against taxing their revenue because of presumed employment benefits for the communities impacted by the industry.

This analysis shows that there is a gap between the total employment benefit experienced in Pennsylvania counties where the drilling takes place and changes in employment of local residents of those counties. The analysis confirms what many local residents have been saying: much of the increase in employment from Marcellus shale development in Pennsylvania has largely benefitted out-of-county and out-of-state residents.

References

Allcott, H., and D. Keniston. 2014. "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in North America." Working paper, New York University.

Banzhaf, H.S., and N. Lavery, 2010. "Can Land Tax Help Curb Urban Sprawl? Evidence from Growth Patterns in Pennsylvania." *Journal of Urban Economics* 67(2): 169–179.

Banzhaf, H.S., and R. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review* 98(3): 843–863.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How Much Should We Trust Differences-in-differences Estimates?" *The Quarterly Journal of Economics* 119(1): 249–275.

Black, D., T. McKinnish, and S. Sanders. 2005. "The Economic Impact of the Coal Boom and Bust." *The Economic Journal* 115(503): 449–476.

Brasier, K.J., M.R. Filteau, D.K. McLaughlin, J. Jacquet, R.C. Stedman, T.W. Kelsey, and S.J. Goetz. 2011. "Residents' Perceptions of Community and Environmental Impacts from Development of Natural Gas in the Marcellus Share: A Comparison of Pennsylvania and New York Cases." *Journal of Rural Studies* 26(1): 32–61.

Brown, J.P. 2014. "Production of Natural Gas from Shale in Local Economies: A Resource Blessing or Curse?" *Economic Review, Federal Reserve Bank of Kansas City* 99(1): 119–147.

Brundage, T.L., J. Jacquet, T.W. Kelsey, J.R. Ladlee, J. Lobdell, J.F. Lorson, L.L. Michael, and T.B. Murphy. 2011. "Pennsylvania Statewide Marcellus Shale Workforce Needs." Marcellus Shale Education and Training Center, Williamsport, PA.

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

———. 2011. "The Pennsylvania Marcellus Natural Gas Industry: Status, Economic Impacts, and Future Potential." Department of Energy, Environmental, and Mineral Economics, Pennsylvania State University, University Park.

Harris, R.D.F., and E. Tzavalis. 1999. "Inference for Unit Roots in Dynamic Panels Where the Time Dimension Is Fixed." *Journal of Econometrics* 91(2): 201–226.

Herzenberg, S. 2011. "Drilling Deeper into Job Claims: The Actual Contribution of Marcellus Shale to Pennsylvania Job Growth." Keystone Research Center, Harrisburg, PA.

Im, K.S., M.H. Pesaran, and Y. Shin. 2003. "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics* 115(1): 53–74.

Imbens, G.W., and J.M. Wooldridge. 2009. "Recent Development in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47(1): 5–86.

Kelsey, T.W., M. Shields, J.R. Ladlee, and M. Ward. 2011. "Economic Impacts of Marcellus Shale in Pennsylvania: Employment and Income in 2009." Marcellus Shale Education and Training Center, Williamsport, PA.

Marchand, J. 2012. "Local Labor Market Impacts of Energy Boom-Bust-Boom in Western Canada." *Journal of Urban Economics* 71(1): 165–174.

Rahm, B.G., J.T. Bates, L.R. Bertoria, A.E. Galford, D.A. Yoxtheimer, and S.J. Riha. 2013. "Wastewater Management and Marcellus Shale Gas Development: Trends, Drivers, and Planning Implications." *Journal of Environmental Management* 120(1): 105–113.

Wang, Z., and A. Krupnick. 2013. "A Retrospective View of Shale Gas Development in the United States." Discussion paper 13-12, Resources for the Future, Washington, DC.

Weber, J.G. 2012. "The Effects of a Natural Gas Boom on Employment and Income in Colorado, Texas, and Wyoming." *Energy Economics* 34(5): 1580–1588.

_____. 2014. "A Decade of Natural Gas Development: The Makings of a Resource Curse?" *Resource and Energy Economics* 37(1): 168–183.

Weinstein, A.L., and M.D. Partridge. 2011. "The Economic Value of Shale Natural Gas in Ohio." College of Food, Agricultural, and Environmental Sciences, Ohio State University, Columbus.

White, J. 2012. "Why Aren't There More Marcellus Shale Jobs for Pennsylvania Workers?" Message posted to http://patch.com/pennsylvania/canon-mcmillan/why-aren-t-there-more-marcellus-shale-jobs-for-pa-workers (accessed May 24, 2014).