

NBER WORKING PAPER SERIES

DUTCH DISEASE OR AGGLOMERATION? THE LOCAL ECONOMIC EFFECTS  
OF NATURAL RESOURCE BOOMS IN MODERN AMERICA

Hunt Allcott  
Daniel Keniston

Working Paper 20508  
<http://www.nber.org/papers/w20508>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2014, Revised June 2017

We thank Yunmi Kong and Wendy Wei for superb research assistance and Julia Garlick for help in data preparation. We are grateful for feedback from Costas Arkolakis, Nathaniel Baum-Snow, Dan Black, Rick Hornbeck, Guido Imbens, Ryan Kellogg, Pat Kline, Erin Mansur, Enrico Moretti, Reed Walker, and seminar participants at the ASSA Annual Meetings, Bates White, Berkeley, Bocconi, the Census Center for Economic Studies, Centre de Recherche en Economie et Statistique, Cornell, Helsinki Center of Economic Research, LSE and UCL, Michigan State, NYU, Oxford, Sussex, Toulouse, Wharton, the World Bank, and Yale. Thanks also to Randy Becker, Allan Collard-Wexler, Jonathan Fischer, Todd Gardner, Cheryl Grim, Javier Miranda, Justin Pierce, and Kirk White for their advice and help in using U.S. Census of Manufactures data, and to Jean Roth for advice on the Current Population Survey. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or the National Bureau of Economic Research. All results have been reviewed to ensure that no confidential information is disclosed.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Hunt Allcott and Daniel Keniston. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms  
in Modern America

Hunt Allcott and Daniel Keniston

NBER Working Paper No. 20508

September 2014, Revised June 2017

JEL No. J2,L6,O4,Q43,R1

**ABSTRACT**

Do natural resources benefit producer economies, or is there a “Natural Resource Curse,” perhaps as the crowd-out of manufacturing productivity spillovers reduces long-term growth? We combine new data on oil and gas endowments with Census of Manufactures microdata to estimate how oil and gas booms affect local economies in the United States. Local wages rise during oil and gas booms, but manufacturing is not crowded out—in fact, the sector grows overall, driven by upstream and locally-traded subsectors. Tradable manufacturing subsectors do contract during resource booms, but their productivity is unaffected, so there is no evidence of foregone local learning-by-doing effects. Over the full 1969-2014 sample, a county with one standard deviation additional oil and gas endowment averaged about one percent higher real wages. Overall, the results provide evidence against a Natural Resource Curse within the United States.

Hunt Allcott

Department of Economics

New York University

19 W. 4th Street, 6th Floor

New York, NY 10012

and NBER

hunt.allcott@nyu.edu

Daniel Keniston

Yale University

P. O. Box 208269

New Haven, CT 06520-8269

and NBER

daniel.keniston@yale.edu

# 1 Introduction

A long literature has debated whether natural resource abundance is good for economic growth (van der Ploeg 2011). If markets are efficient, then standard trade models predict that resource-abundant regions benefit from an increase in resource prices. However, there are many accounts of a potential “Natural Resource Curse,” in which resource extraction interacts with market failures to make producer economies worse off. Indeed, Sachs and Warner (1995, 1999, 2001) find that resource abundance is negatively associated with growth in cross-country data, concluding that the Resource Curse is a “reasonably solid fact.” The empirical literature remains unsettled, however, due to problems of small samples, data quality, and endogeneity.<sup>1</sup> Understanding these questions is crucial for the many countries and local areas whose economies rely on natural resource extraction.

The Resource Curse literature has identified several types of market failures through which natural resource extraction could reduce growth. In countries with strong political institutions that limit corruption and conflict, concern has centered on the loss of local learning-by-doing spillovers from manufacturers that might be crowded out during a natural resource boom. Indeed, a recent literature has shown that manufacturing plants exert positive productivity spillovers on other nearby plants.<sup>2</sup> Thus, if natural resource growth crowds out manufacturing, this also reduces productivity spillovers, which could reduce long-term overall growth. This possible chain of events—the short-run crowd out of manufacturing, leading eventually to decreased manufacturing productivity, causing long-run manufacturing contraction and reduced overall growth and welfare—is often called “Dutch Disease,” and the idea has been well-developed theoretically.<sup>3</sup> Harding and Venables (2013) and Ismail (2010) find that natural resource exports crowd out manufacturing exports in cross-country data, and Sachs and Warner (2001) argue that Dutch Disease is what causes their empirical finding of slower growth in resource-abundant countries.

In this paper, we ask *what have been the local economic impacts of oil and gas booms and busts within the U.S. over the past half century?* In particular, we consider whether oil and gas extraction has benefited people who live in resource-abundant areas, or whether there might be a Natural Resource Curse acting through a Dutch Disease-style chain of events. Aside from being

---

<sup>1</sup>Auty and Mikesell (1998) and Gylfason, Herbertsson, and Zoega (1999) also find that natural resource abundance is negatively correlated with growth, but others have arrived at different results when instrumenting for resource abundance (Brunnschweiler and Bulte 2009), including country fixed effects (Manzano and Rigobon 2001), using different measures of resource intensity (Lederman and Maloney 2007b), or conditioning on the quality of institutions (Collier and Goderis (2009), Mehlum, Moene, and Torvik (2006)).

<sup>2</sup>See Ellison, Glaeser, and Kerr (2010), Greenstone, Hornbeck, and Moretti (2010), Kline and Moretti (2014), Serafinelli (2012) and others for empirical results, and see Glaeser and Gottlieb (2009) and Moretti (2010) for literature overviews on agglomeration and local labor markets.

<sup>3</sup>See Corden and Neary (1982), Krugman (1987), Matsuyama (1992), van Wijnbergen (1984), and others. Some alternative definitions of “Dutch Disease” include only the crowd-out of tradable manufacturing, but this would arise even in an efficient market due to standard Rybczynski effects, and is thus not really a “disease.” Like much of the literature, we therefore think of Dutch Disease as the full possible chain of events including decreased manufacturing productivity and decreased long-run growth.

important for understanding the pathways of local economic growth, this question is very relevant to local policymakers who are deciding whether to encourage or restrain resource booms. Indeed, Maryland, New York State, the Delaware River Basin, Bulgaria, France, Germany Scotland, South Africa, Wales, and many other areas have recently banned hydraulic fracturing for environmental reasons (Rosenbaum 2016), and it is important for policymakers to know the foregone economic gains, if any.

To organize the empirical work, we first lay out a Rosen-Roback spatial equilibrium model in the spirit of Moretti (2010). There are two initially symmetric counties, each with natural resource, local, and tradable goods sectors, and labor is imperfectly mobile across counties. In the local and tradable sectors, we model intertemporal productivity spillovers from both “learning-by-doing” and “agglomeration”—that is, we allow local productivity to be influenced both by past employment in that sector and by the past overall population in the county. We model a resource boom as a temporary increase in resource sector revenue productivity in one of the two counties. The resource boom affects both counties through general equilibrium effects.

The model predicts that during a resource boom, local population and wages increase, and local welfare increases during the boom even as higher wages crowd out the tradable sector. In the absence of productivity spillovers, the counties return to parity after the resource boom ends, there are no market failures, and the resource boom unambiguously increases cumulative welfare, i.e. the sum of welfare over time. If there are productivity spillovers, however, the cumulative welfare effect is ambiguous. If the resource boom crowds out enough tradable sector learning-by-doing spillovers, then tradable sector productivity decreases and the boom can reduce cumulative welfare. This illustrates how an analogue of Dutch Disease might arise *within* a country, where local manufacturing is initially crowded out by higher local wages, not by currency appreciation. On the other hand, if the effects of crowd-out are limited and agglomerative effects are strong, this further augments the boom’s welfare gains. In the extreme, population could concentrate in the resource-abundant county even after the boom ends, just as population continues to be concentrated in historical river portage sites long after these sites lost any productive advantage (Bleakley and Lin 2012).

To empirically estimate the effects of natural resource booms, we combine county-level panel data on employment, earnings, and population with a new dataset of county-level oil and gas production and reserves that we constructed for this paper. Crucially, we also use restricted-access plant-level microdata from the U.S. Census of Manufactures, which allow us both to observe manufacturing revenue productivity and to measure differential effects across local vs. tradable subsectors. Our empirical strategy is a “shift-share” approach inspired by Bartik (1991): the key independent variable interacts time series variation in national oil and gas employment with cross-sectional variation in counties’ initial oil and gas endowment. Extending our analysis back to the 1960s allows us to exploit dramatic time series variation, as hundreds of thousands of oil and gas

jobs are created and destroyed nationwide in the boom of the 1970s and early 1980s, the bust of the late 1980s, and the second boom of the past decade.

Our empirical results trace out the local economic impacts of oil and gas booms and busts, with special attention to the possible chain of events suggested by the within-country version of Dutch Disease. First, local wages, population, and employment in resource-abundant counties are all procyclical with oil and gas: they rise during oil and gas booms and fall during busts. During the most recent boom (2007-2014), wages were 1.6 to 3.0 percent higher in counties with one standard deviation more oil and gas endowment, and wages also rose significantly in resource-abundant counties during the boom of the 1970s and 1980s. This introduces the possibility that some manufacturers could be crowded out.

Second, however, while manufacturing is often thought of as largely producing tradable goods, the sector in aggregate is not crowded out by resource booms. Instead, manufacturing in resource-abundant counties grows during resource booms and shrinks during busts. More disaggregated analyses show that manufacturing's procyclicality is driven by locally-traded subsectors and those with upstream or downstream input-output linkages to oil and gas. By contrast, more highly-tradable manufacturing subsectors contract during booms. Thus, these tradable subsectors could still generate a Dutch Disease effect, if their contraction leads to reduced productivity.

Third, local manufacturing plants' revenue-based total factor productivity (TFP-R) is procyclical with local resource booms. This could either be because local output prices are procyclical with resource booms, or because transport costs are countercyclical, as local buyer and supplier bases grow during booms and shrink during busts. In tradable manufacturing subsectors, despite the temporary contraction during booms, TFP-R never decreases. This rules out the productivity decline that is central to a Dutch Disease mechanism.

Fourth, while county-level population, employment, wages, and revenue productivity are all procyclical, the booms are canceled out by the busts. By the end of the 1990s, we see no significant remaining long-term effects of the boom and bust cycle of the 1970s and 1980s.

The procyclicality of local wages plus the lack of adverse long-term effects foreshadows our final result, on the cumulative effects of natural resource abundance. Over our full 1969-2014 sample, counties with higher oil and gas endowments average 1.0 percent higher real wages relative to otherwise-similar counties with one standard deviation less endowment. Under additional assumptions, we can control for general equilibrium effects (geographic spillovers) using an empirical approach parallel to Miguel and Kremer (2004). Over the years in our sample, counties with one standard deviation more endowment average 0.4 percent higher real wages compared to their own counterfactual.

Real wages are not a measure of local welfare if resource booms also affect local amenities such as environmental quality, and Muehlenbachs, Spiller and Timmins (2015) and others provide evidence of adverse environmental effects from shale gas development. Our model shows how population

changes identify a resource boom’s welfare effects: people “vote with their feet” by moving to counties where they are better off, taking into account amenities that we might not observe. We find that one standard deviation additional oil and gas endowment increases population by an average of 0.8 percent over our sample, implying that oil and gas endowments increase cumulative local welfare even in the presence of changes in unobserved amenities.

In summary, we find evidence against each link in the hypothesized Dutch Disease chain of events. Manufacturing plants in resource-abundant counties benefit a surprising amount from local demand growth, manufacturing productivity does not decrease, and overall, there is no Natural Resource Curse: oil and gas abundance increases cumulative local wages and welfare.

In the remainder of this first section, we discuss related literature. Section 2 provides background on oil and gas booms in the United States since the 1960s. Section 3 presents the model. Section 4 details the data, Section 5 outlines the empirical strategy, and Section 6 presents results. Section 7 concludes.

## 1.1 Related Literature

We build on a growing literature that uses within-country variation to identify the effects of resource booms, both in the United States (Carrington (1996), Glaeser, Kerr, and Kerr (2015), Haggerty *et al.* (2014), Jacobsen and Parker (2014), James and Aadland (2011), James and James (2012), Papyrakis and Gerlagh (2007), and others) and other countries (Aragon and Rud (2011), Asher and Novosad (2014a, 2014b), Caselli and Michaels (2013), Domenech (2008), Dube and Vargas (2013), Monteiro and Ferraz (2012), and others). Recent papers by Bartik, Currie, Greenstone, and Knittel (2017), Feyrer, Mansur, and Sacerdote (2015), Jacobsen (2015), Muehlenbachs, Spiller and Timmins (2015), Munasib and Rickman (2015), Paredes, Komarek, and Loveridge (2015), and Weber (2012) study the impacts of the recent hydraulic fracturing (“fracking”) boom on income, housing values, environmental quality, and other outcomes. Acemoglu, Finkelstein, and Notowidigdo (2013) study the effects of oil price shocks on health spending. Black, McKinnish, and Sanders (2005a) study the Appalachian coal boom and bust, finding positive employment spillovers to construction, retail, and services sectors but no effect on manufacturing employment. An important paper by Michaels (2010) studies the long-term effects of oil abundance in the southern United States, using cross-sectional estimators to compare counties that have major oil fields to other nearby counties that do not. He finds that resource discoveries cause oil-abundant counties to specialize in oil production, but this did not reduce growth in other sectors: higher incomes increased population, which increased the provision of local public goods, which in turn increased employment in agriculture and manufacturing.

Our paper differs from existing work in four main ways. First, the U.S. Census of Manufactures microdata allow a significant advance beyond previous work that used only publicly-available data on sector-level aggregate employment. Unlike Michaels (2010) and Black, McKinnish, and Sanders

(2005a), we find some evidence that resource booms cause some manufacturers to contract, but this finding is only possible because our microdata allow us to look at the subset of manufacturers producing more highly-traded goods. Furthermore, we directly study how resource booms affect manufacturing productivity, which until now was a hypothesized but unmeasured mechanism.

Second, we have gathered very detailed data on resource abundance. Black, McKinnish, and Sanders (2005a) proxy for resource abundance with binary variables based on initial coal sector employment, Michaels (2010) proxies with a binary variable for whether a county lies over a large existing oilfield, and Sachs and Warner (1995, 1999, 2001) and other cross-country studies use rough proxies such as the share of natural resource exports in GDP. By contrast, we have constructed a complete measure of each county’s economically recoverable oil and gas endowment as of 1960, including subsequent production, proven reserves, and undiscovered reserves. Undiscovered reserves are particularly important because production and proven reserves could be endogenous to a county’s potential outcomes. For example, if local economic conditions were worse, New York State might not have banned fracking, and its production would thus be higher.

Third, we pay more attention to the cyclicity of natural resource extraction, as boom and bust cycles are a fundamental feature of oil and gas production. Our 46-year panel allows us to track counties and firms through multiple booms and busts. This is important empirically because it allows us to control for secular time trends that could be correlated with resource endowments. By contrast, analyses focusing only on the recent fracking boom must make stronger assumptions about such confounders, which could be particularly problematic given that the fracking boom was contemporaneous with the Great Recession and with the rapid decline of U.S. manufacturing.

Fourth, we provide a more careful treatment of general equilibrium effects and other forms of geographic spillovers. Our model explicitly captures these effects, and our empirical strategy delivers plausibly unbiased estimates of average treatment effects on the treated (or what we call “absolute effects”) in the presence of geographic spillovers. By contrast, most previous papers on resource booms estimated what we call “relative effects” without fully considering spillovers.<sup>4</sup> One recent exception is Feyrer, Mansur, and Sacerdote (2015), who estimate cross-county geographic spillovers from the fracking boom.

Our work is also connected to the broader empirical literature on agglomeration and productivity spillovers. As Ellison and Glaeser (1997) point out, agglomerative spillovers are difficult to identify because both heterogeneous local natural advantages and spillovers can cause firms to co-locate. To identify spillovers, several recent papers have exploited natural experiments, including the siting of large manufacturing plants (Greenstone, Hornbeck, and Moretti 2010), portage sites (Bleakley and Lin 2012), exogenous worker flows across firms (Serafinelli 2012), and the boundaries of the Tennessee Valley Authority development region (Kline and Moretti 2014). Our analysis identifies

---

<sup>4</sup>The broader literature on local economic shocks has been more precise on this issue; see, for example, Hornbeck and Moretti (2015) and Kline and Moretti (2014).

spillovers using the natural experiment of local resource booms and busts.

More broadly, our work connects to the literature on local economic shocks, including studies of local effects of national-level sectoral trends (Bartik 1991, Blanchard and Katz 1992), military base closures (Hooker and Knetter 2001), place-based economic development policies (Busso, Gregory, and Kline 2013), Chinese import competition (Autor, Dorn, and Hanson 2013), and other factors. Against this literature, our paper is novel in that it studies oil and gas booms, which have helped lead the recovery from the Great Recession (Moretti 2013). Our paper is also novel in that it studies cross-industry spillovers, while much of the local labor market literature studies how exogenous shocks affect county aggregates. Using the tech industry as an analogy, much of this literature essentially studies how the tech boom affects wages, unemployment, and other aggregates in places like Silicon Valley. But while a tech boom clearly benefits local land owners and labor suppliers, it hurts other industries that draw on the same pool of land or labor but are not experiencing positive revenue productivity shocks. As our model and empirical strategy make clear, the policy implications then depend on whether externalities from the booming industry are more positive than from the industries being crowded out.

## 2 Background: Oil and Gas Booms in Modern America

Figure 1 presents oil and natural gas prices in the U.S. from 1960 to 2014, using data from the U.S. Energy Information Administration (EIA). Like all prices in this paper, these are in real 2010 dollars. Real oil and gas prices were steady and slowly declining from the end of the Second World War through the early 1970s. Prices rose suddenly in October 1973 due to the Arab oil embargo and again in 1979-1980 due to the Iranian Revolution and the Iran-Iraq war. Natural gas prices closely follow oil prices over the entire study period, except for the most recent few years, when a marked natural gas supply increase pushed natural gas prices down.

Figure 2 shows that U.S. oil production peaked in 1970 and began to decline. The decline was monotonic for the first few years of the 1970s, until it was arrested by the supply response to the 1973 price shock. This supply response, coupled with the recession of the early 1980s, caused prices to drop steadily from March 1981 to the end of 1985, and then sharply in the first six months of 1986. In the early 2000s, global demand growth spurred a second boom of high prices and increased production.

Oil and gas production requires significant labor input. Figure 2 shows total national employment in the oil and gas sector, as reported in the Regional Economic Information System (REIS). Closely mirroring the price trend, employment rose from under 400,000 people in the early 1970s to over one million in the early 1980s, then dropped sharply in 1986 and continued to decline steadily until 2002.<sup>5</sup> Employment then surged again, more than doubling between 2002 and 2014.

---

<sup>5</sup> We switch from the SIC to the NAICS classification system in 2000 and plot both SIC and NAICS data points in that year.



Of course, these large fluctuations in oil and gas employment are concentrated in the more resource-abundant counties. Prior to the late 1990s, oil and gas were almost exclusively recovered from “conventional” reserves: oil and gas accumulations trapped beneath an impermeable rock layer where the resulting reservoir could be reached with a vertical well. Figure 3 shows each county’s early endowment: that is, the value of oil and gas per square mile that was in the ground in 1960 and was economically recoverable using technologies available during the boom and bust of the 1970s and 1980s. (We detail the construction of this variable in Section 4.1.)

More recently, prospecting and extraction has focused on extracting “unconventional” oil and gas trapped in small pores within impermeable rocks. “Fracking” was pioneered for oil and gas wells in 1947 and was used in some areas during the 1970s boom. There were significant advances in horizontal drilling in the 1980s and 1990s, and commercially-viable shale gas extraction was pioneered in the Barnett Shale in northern Texas in 1997. Since then, large amounts of tight oil and shale gas have become economically recoverable. Figure 4 maps the additional endowment that became economically exploitable only after the end of the 1990s. This illustrates regions where large amounts of shale gas or tight oil are newly economically recoverable, such as the Bakken Shale in western North Dakota, the Niobara shale in eastern Colorado, the Marcellus and Utica Shales in Pennsylvania, Ohio, New York, and West Virginia, the Barnett Shale, Granite Wash, and Eagle Ford in Texas, the Woodford Shale in Texas and Oklahoma, and the Haynesville Shale on the border of Texas and Louisiana.

### 3 Theoretical Framework

#### 3.1 Overview

Our theoretical model nests multi-period productivity spillovers into Moretti’s (2011) version of the Rosen-Roback spatial equilibrium framework, illustrating a within-country version of Dutch Disease. There are two counties,  $c \in \{a, b\}$ , with imperfectly mobile population. There are three production sectors, indexed  $j \in \{m, l, r\}$ , representing tradable goods, local (non-tradable) goods, and natural resources, respectively. Each sector has revenue productivity  $X_{jc}$ . County  $b$  does not have natural resources, so  $X_{rb} = 0$ .

There are three time periods, indexed by  $t$ .  $t = 0$  represents the initial equilibrium, in which the two counties are symmetric, so  $X_{ra} = X_{rb} = 0$ . In  $t = 1$ , county  $a$  experiences a “resource boom,” meaning that  $X_{ra} > 0$ . This could reflect either a technological improvement or a price increase. The boom is temporary, so we again have  $X_{ra} = X_{rb} = 0$  for  $t = 2$ . Below, we provide detail on production and consumption and describe the effects of the resource boom on equilibrium outcomes.<sup>6</sup>

---

<sup>6</sup>We chose to simplify functional forms so that key results can be seen in closed form. Notwithstanding, the basic predictions in Section 3.6 are very general. Our 2014 NBER working paper has a similar model with more general

### 3.2 Production

Each production sector  $j \in \{m, l, r\}$  comprises a composite firm that employs  $N_{jc}$  workers and earns revenue  $R_{jc} = X_{jc}N_{jc}^{1-\gamma}$ , with  $\gamma \in (0, 1)$ .<sup>7</sup> Profit maximization gives labor demand  $N_{jc} = \left[\frac{X_{jc}(1-\gamma)}{W_c}\right]^{\frac{1}{\gamma}}$ , and county total labor demand is  $N_c = N_{mc} + N_{lc} + N_{rc}$ . Revenue productivity  $X_{jc}$  is  $P_{jc}\Omega_{jc}$ , price times physical productivity. Resource revenue productivity  $X_{rc}$  is exogenous, as described above. The tradable good price is exogenous and normalized to  $P_{mc} = 1$ , and the local good price  $P_{lc}$  is endogenously determined.

Tradable and local sector physical productivities  $\Omega_{jc}$  evolve over time subject to two types of possible productivity spillovers. First, “learning-by-doing spillovers”  $\theta_j \geq 0$  would imply that a sector’s current productivity increases in that sector’s past employment. Second, “agglomeration spillovers”  $\Lambda \geq 0$  would imply that current productivity increases with past local population, echoing the findings of Bleakley and Lin (2012) and others. Specifically, using lower case letters for natural logs (so  $\omega_{jc} = \ln(\Omega_{jc})$ ,  $n_{jc} = \ln(N_{jc})$ , and  $n_c = \ln(N_c)$ ), we assume that physical productivity in sector  $j \in \{m, l\}$  is

$$\omega_{jc,t+1} = \rho_j \omega_{jct} + \theta_j n_{jct} + \Lambda n_{ct} + \zeta_j. \quad (1)$$

We conceptualize the possible learning-by-doing spillovers as between-firm, local, and intertemporal. Although the exact functional forms vary, this conceptualization is consistent with Arrow’s (1962) original formulation of learning-by-doing, with Kline and Moretti (2014), and with the previous theoretical literature on Dutch Disease, including Krugman (1987), Matsuyama (1992), and van Wijnbergen (1984). Modeling productivity spillovers as between-firm and local matches the empirical results of Ellison, Glaeser, and Kerr (2010), Greenstone, Hornbeck, and Moretti (2010), Kline and Moretti (2014), and Serafinelli (2012), who document productivity spillovers to other firms in the same area. While within-firm learning may also occur (see Benkard (2000) and others), we focus on between-firm spillovers because within-firm spillovers are not externalities meriting policy intervention. We model intertemporal learning-by-doing instead of only contemporaneous spillovers because this admits the possibility that a natural resource boom could affect (and possibly reduce) local productivity and welfare in the long term, after the boom is over.

One could imagine any number of other market failures, including contemporaneous productivity spillovers and environmental externalities. The social welfare function we specify and empirically calibrate below captures these and any other factor that would manifest through real wages or migration.

Following Moretti (2010), we also model a housing sector that does not require local labor, with

---

functional forms, and our 2015 NBER working paper has a model that generates similar results in a Melitz (2003) model of firm heterogeneity.

<sup>7</sup>The model’s basic predictions generalize to the case with different  $\gamma_m$ ,  $\gamma_l$ , and  $\gamma_r$ , but we choose a homogeneous  $\gamma$  to highlight geographic rather than intersectoral differences.

constant elasticity supply function  $p_{hc} = kn_c + k_0$ .

### 3.3 Consumers and Workers

Each individual decides what to consume and where to live, supplying one unit of labor. Individual  $i$  consumes  $C_{il}$ ,  $C_{ih}$ , and  $C_{im}$  units of local goods, housing, and tradable goods, respectively. Individuals also receive utility  $A_c \epsilon_{ic}$  from living in county  $c$ , where  $A_c$  is county  $c$ 's "amenity" and  $\epsilon_{ic}$  is individual  $i$ 's idiosyncratic taste for county  $c$ . Amenities could represent factors such as local public goods, congestion, and environmental quality, some of which could change during a resource boom. In this theoretical framework, we consider a resource boom that does not affect  $A_c$ , but we also introduce some results that allow the empirical welfare analysis to accommodate possible effects on  $A_c$ . Individuals have Cobb-Douglas preferences, maximizing utility

$$U_{ic} = C_{il}^\alpha C_{ih}^\beta C_{im}^{1-\alpha-\beta} A_c \epsilon_{ic}, \quad (2)$$

subject to budget constraint  $W_c \geq P_{lc} C_{il} + P_{hc} C_{ih} + C_{im}$ . We assume  $\alpha, \beta, (1 - \alpha - \beta) \in (0, 1)$ .

Again using lower case letters for natural logs, indirect utility is  $u_{ic} = w_c - \alpha p_{lc} - \beta p_{hc} + a_c + \ln \epsilon_{ic}$ , plus a constant. Individuals choose to live in county  $a$  instead of county  $b$  if  $u_{ia} > u_{ib}$ . We assume that  $\ln \epsilon_{ic}$  is distributed type I extreme value with scale parameter  $s^2$ , where  $s \in (0, \infty)$ . This gives a logit model, and the inverse labor supply difference between the two counties is

$$w_a - w_b = \alpha(p_{la} - p_{lb}) + \beta(p_{ha} - p_{hb}) + s(n_a - n_b) - (a_a - a_b). \quad (3)$$

In equilibrium, consumers and firms optimize, and the labor, local goods, and housing markets clear.

### 3.4 Social Welfare

Of course, the first-best response to unpriced productivity spillovers would be for a policymaker to measure and price these externalities. The general problem of the social planner with learning-by-doing or agglomeration spillovers is well-explored; see, for example, Arrow (1962) and Glaeser and Gottlieb (2008, 2009). We therefore do not study this issue further. Instead, we focus on a different welfare question: *does the resource boom increase local social welfare in the producer county?* This question is important because it helps local policymakers to evaluate policies such as fracking bans that restrain or prevent resource booms.

As a measure of social welfare, we consider the indirect utility of people who live in county  $c$  in all periods. Up to a constant, this is

$$u_c = \sum_t u_{ct} = \sum_t \delta^t [w_{ct} - \alpha p_{lct} - \beta p_{hct} + a_{ct}], \quad (4)$$

where  $\delta$  is the discount factor.<sup>8</sup>

One might also be interested in national or global-level welfare. We focus on local welfare effects for two reasons. First, many of the significant policy decisions to discourage oil and gas drilling, such as fracking bans, are made by state or local authorities. Second, significant additional assumptions are required to empirically identify national social welfare impacts, due to general equilibrium effects.

### 3.5 Relative and Absolute Effects

Our welfare question can be asked in two different ways, depending on what county  $a$  is compared to. First, we can consider the *relative effects* of a natural resource boom: does the resource boom increase cumulative social welfare in county  $a$  relative to county  $b$ ?

$$u_a - u_b = \sum_t \delta^t [(w_{at} - w_{bt}) - \alpha (p_{lat} - p_{lbt}) - \beta (p_{hat} - p_{hbt}) + (a_{at} - a_{bt})] \quad (5)$$

Substituting labor supply from Equation (3), we see that this can also be written as a function of only relative population:

$$u_a - u_b = \sum_t \delta^t s (n_{at} - n_{bt}). \quad (6)$$

Intuitively, people vote with their feet by migrating to the county with higher welfare. This equation will be useful empirically, as it will allow us to sign the relative welfare effect even without a direct estimate of how the resource boom affects local prices and amenities.

Second, we can consider *absolute effects*: does the resource boom increase cumulative social welfare in county  $a$  relative to counterfactual? Specifically, we denote with (0) a counterfactual outcome if  $X_{ra} = X_{rb} = 0$  always.

---

<sup>8</sup>We assume that the social planner focuses on permanent residents both because this may be more realistic and because it eliminates the need to keep track of migrants' idiosyncratic taste shocks  $\epsilon_{ic}$ . We ignore producer surplus because many firms are owned by shareholders outside of the county. In the empirical section, however, we will estimate effects on manufacturing revenues and housing rents. We note that the social planner maximizes indirect utility in units of log dollars instead of in units of dollars. Under the usual log approximation to percent differences, however, both the absolute and relative effects comparisons described below give approximate percent differences in welfare in units of dollars.

$$u_a - u_a(0) = \sum_t \delta^t [(w_{at} - w_{at}(0)) - \alpha (p_{lat} - p_{lat}(0)) - \beta (p_{hat} - p_{hat}(0)) + (a_{at} - a_{at}(0))] \quad (7)$$

Of course, relative and absolute effects are different because  $u_b \neq u_a(0)$  due to general equilibrium effects. Relative effects are important from a positive perspective: they describe how resource producers vs. non-producers fare during a resource boom. Absolute effects are important from a normative perspective: they answer the question of whether a local policymaker should restrain a resource boom. In the remainder of the paper, we distinguish between relative and absolute effects for welfare and for other outcomes.

### 3.6 Effects of Resource Booms, and the Definition of Dutch Disease

We now consider the predicted effects of a natural resource boom – that is, an increase in county  $a$ 's resource sector revenue productivity  $X_{ra}$  in period  $t = 1$ . Appendix A solves for the equilibrium and derives these predictions in absolute as well as relative terms.

#### 3.6.1 Contemporaneous Effects

Predictions 1 and 2 consider the contemporaneous ( $t = 1$ ) effects of a resource boom.

**Prediction 1:** A resource boom increases population and wages in  $t = 1$ .

**Prediction 2:** A resource boom increases local sector employment and prices, and decreases tradable sector employment, in  $t = 1$ .

These two predictions would arise in any standard model, and in our model they arise either with or without productivity spillovers. Prediction 1 is important because the entire Dutch Disease mechanism begins with an initial wage increase during the resource boom. If migration moderates the wage increase, then any subsequent Dutch Disease effects will also be moderated. Although manufactures are often thought of as goods that are traded across counties, Prediction 2 highlights that the effect of a resource boom on overall manufacturing employment depends on the share of manufactures that are traded vs. sold locally. If manufacturing is largely traded outside the county, then it will be crowded out during a resource boom. If more of manufacturing is sold locally, then it will be “crowded in.” Of course, the decrease in tradable sector employment does not imply any economic inefficiency.

#### 3.6.2 Long-Term Effects with and without Productivity Spillovers

Predictions 3 and 4 demonstrate the long-term ( $t = 2$ ) effects.

**Prediction 3:** If and only if there are agglomeration spillovers or local sector learning-by-doing spillovers (i.e.  $\Lambda > 0$  or  $\theta_l > 0$ ), a resource boom in  $t = 1$  increases local sector productivity in  $t = 2$ . Furthermore, the sign of relative tradable sector productivity in  $t = 2$  is the same as the sign of

$$\Lambda - \theta_m \frac{\alpha\gamma + \beta k + s}{\gamma(1 - \alpha)}. \quad (8)$$

For the local sector, this prediction follows from applying Predictions 1 and 2 to the productivity evolution formula in Equation (1). Since both population and local sector employment increase in county  $a$  in  $t = 1$ , either learning-by-doing ( $\theta_l > 0$ ) or agglomeration ( $\Lambda > 0$ ) would increase local sector productivity in  $t = 2$ . For the tradable sector, Equation (8) says that relative productivity  $\omega_{ma,t=2} - \omega_{mb,t=2}$  decreases if the foregone learning-by-doing  $\theta_m \frac{\alpha\gamma + \beta k + s}{\gamma(1 - \alpha)}$  outweighs the agglomeration effect  $\Lambda$ . Alternatively, tradable sector relative productivity could increase if  $\Lambda$  is sufficiently large, or stay the same if learning-by-doing and agglomeration are zero or exactly offset.

Prediction 4 then illustrates the possibility that foregone learning-by-doing spillovers could generate a perverse result in which the initial increase in resource sector revenue productivity actually reduces cumulative welfare.

**Prediction 4: A resource boom increases cumulative relative welfare  $u_a - u_b$  if and only if**

$$\frac{\gamma + \beta k + s}{\delta} + \Lambda + \alpha\theta_l > \theta_m \frac{\alpha\gamma + \beta k + s}{\gamma}. \quad (9)$$

This prediction shows that relative welfare is more likely to increase under several conditions. First, if the social planner is less patient ( $\delta$  is small), then growth during the boom in  $t = 1$  can outweigh any losses in  $t = 2$ . Second, if agglomeration spillovers  $\Lambda$  are large, then county  $a$  benefits more from the population increase during the boom. Third, if local sector learning-by-doing  $\theta_l$  is strong, and/or the local good expenditure share  $\alpha$  is large, then county  $a$  benefits more from an increase in local sector learning-by-doing in  $t = 2$ . Fourth, if tradable sector learning-by-doing  $\theta_m$  is weak, the sector's contraction in  $t = 1$  has smaller negative consequences in  $t = 2$ . Fifth, if housing supply is more elastic (i.e.  $k$  is lower), the housing expenditure share  $\beta$  is lower, locational preferences are weaker (i.e.  $s$  is smaller), or production is less labor-intensive (i.e.  $\gamma$  is larger), then the tradable sector contracts less in  $t = 1$ .

Predictions 3 and 4 allow us to contrast the cases with vs. without productivity spillovers. If there are no productivity spillovers (i.e.  $\Lambda = \theta_m = \theta_l = 0$ ), then the two counties have equal productivity, population, and wages after  $t = 1$ , and a resource boom unambiguously increases relative welfare.<sup>9</sup> If there are productivity spillovers, then local sector relative productivity will

<sup>9</sup>As we show in Appendix A, this is also true for absolute welfare, although the formulas differ because the absolute welfare effect equals the relative effect plus an adjustment to account for the fact that the resource boom also increases

increase, and the signs of both tradable sector relative productivity and relative welfare will depend on the relative strengths of the learning-by-doing vs. agglomeration spillovers.

Of course, one could easily imagine extending our stylized model beyond three periods, and our empirical work spans many years. In that case, any productivity differences in  $t = 2$  could grow over time if the spillovers  $\theta_j$  and  $\Lambda$  and/or the persistence of past productivity differences  $\rho$  are sufficiently large.<sup>10</sup> In the theoretical extreme, this could cause all population to concentrate in one of the two counties. On the other hand, if spillovers and/or productivity persistence are small, then the differences between counties in  $t = 2$  will disappear as the counties converge back to parity in long-run steady state.

We can formalize what we mean by a “Resource Curse” and “Dutch Disease” in the context of these four predictions. A “Resource Curse” is when “natural resource abundance reduces cumulative welfare.” We think of the within-country analogue of Dutch Disease as the possible chain of events through which an increase in resource sector revenue productivity increases local wages (Prediction 1), leading to decreased tradable sector employment (Prediction 2), leading to decreased tradable sector productivity (i.e. the sign of Equation (8) is negative in Prediction 3), which could be large enough to cause a decrease in cumulative welfare (Prediction 4).

The model generates a set of five specific empirical tests that both quantify the local economic effects of natural resource booms and test the chain of events that would arise under a within-country analogue of Dutch Disease. First, do local wages rise substantially during a boom, or is this dampened by migration? Higher local wages increase local welfare during the boom but also could crowd out the tradable sector. Second, even if wages rise, is manufacturing actually crowded out during a boom, or might the sector instead grow as local demand increases? Third, how does manufacturing productivity evolve over time? Some productivity decline, either transitory or permanent, is a necessary condition for Dutch Disease. Fourth, what are the long-term effects after the boom and bust are over? In other words, are any economic effects permanent or transitory? Fifth, what are the cumulative welfare effects over the boom and bust cycle? Depending on the discount rate, a boom and bust could still generate positive cumulative welfare gains even if the effects in  $t = 2$  are negative. Before answering these questions, we detail the data and estimating equations.

## 4 Data

### 4.1 Resource Data

In the model, there are two counties, one with natural resources and one without. In reality,

---

welfare in county  $b$ .

<sup>10</sup>The condition for county productivity in sector  $j$  to diverge from period to period is just  $\omega_{ja,t+1} - \omega_{jb,t+1} > \omega_{jat} - \omega_{jbt}$ . Substituting this into Equation (1) gives  $(\rho_j - 1)(\omega_{jat} - \omega_{jbt}) + \theta_j(n_{jat} - n_{jbt}) + \Lambda(n_{at} - n_{bt}) > 0$ .

there is continuous variation in resource endowments across many counties. To match the model, we would like a measure of each county’s oil and gas revenue productivity, or perhaps more generally, the oil and gas supply curve given the technology available in each year. We proxy for this with the density of “economically recoverable” oil and gas that was in the ground before our analysis begins. We want this endowment measure to be determined only by geological factors that are exogenous to economic outcomes. There are no nationally-comprehensive 1960s-era estimates of economically recoverable oil and gas reserves, although such data are available more recently. We thus construct 1960 endowments by adding production from year 1960 to year  $T$  plus remaining reserves at time  $T$ . There are two types of reserves, “proven” and “undiscovered,” so our measure of county  $c$ ’s endowment density is

$$R_c = \frac{\sum_{t=1960}^T Production_{ct} + Proven\ Reserves_{cT} + Undiscovered\ Reserves_{cT}}{Area\ (Square\ Miles)_c} \quad (10)$$

We divide by county area (including both land and water area) to account for variation in county size.

Because “economically recoverable” depends on the extraction technology, we construct different endowment variables for pre- vs. post-fracking periods. The variable  $R_c^{early}$  measures endowment in the first boom, with  $T = 1995$ , using reserves estimates from the late 1990s.  $R_c^{total}$  measures total endowments including unconventional resources requiring fracking, with  $T = 2011$ , using the latest estimates of proven and undiscovered reserves. For our primary empirical specifications, we define an endowment variable  $R_{ct}$ , which takes value  $R_c^{early}$  until 2000, and  $R_c^{total}$  beginning in 2001. The precise year in which  $R_c^{early}$  ends does not make much difference, and we show that the conclusions are not sensitive to using either  $R_c^{early}$  or  $R_c^{total}$  for the entire period.

We now detail the data sources for production, proven reserves, and undiscovered reserves. Oil and gas production data are from a new county-by-year panel dataset from 1960 to 2011, the first such comprehensive county-level dataset. Much of this data is drawn from a well-level database of oil and gas production provided by DrillingInfo, a market research company, which we collapse to the county-by-year level. The DrillingInfo data do not include several states and are incomplete in early years for some other states, so we have also gathered county-level oil and gas production data from local authorities in 14 additional states; see Appendix Table A1 for details. In several states with relatively low production, county-level production data do not exist for some early years, so we impute production by multiplying state-level production by the county’s share of state production in the earliest year when it is observed.

Proven reserves data are from the EIA’s survey 23L, which collects proven reserves and production by each firm in each oil field (U.S. EIA 2014). The EIA granted us confidential access to county-by-year totals.<sup>11</sup> This is an improvement over publicly-available proven reserves data used

---

<sup>11</sup>We were required to use these EIA data inside a Census Research Data Center, which is cumbersome for analyses



in previous work, which include only the largest fields.

Undiscovered reserves are estimated by the U.S. Geological Survey (USGS) on the basis of the expected oil, gas, and natural gas liquid yield using current technology, including estimated future discoveries over the next 30 years. We map these data to counties by intersecting the most detailed available USGS geological maps with county outlines. The USGS defines undiscovered reserves as “undiscovered petroleum is that which is postulated from geologic knowledge and theory to exist *outside of known accumulations*” [emphasis added].<sup>12</sup> These resources are unaffected by proven reserves or past production except insofar as greater prospecting in an area might transfer resources from the “undiscovered” to the “proven” category. Thus, while both production and proven reserves could be endogenous to economic activity, our measure of  $R_c$  including undiscovered reserves is in principle determined only by geological factors. This is an improvement over previous papers that measure endowments from production, exports, or proven reserves.

We transform physical units of oil and gas to dollar values using their average prices (as always, in real 2010 dollars) over 1960-2011: \$34.92 per barrel of oil and \$3.20 per mmBtu of gas. To make regression results more easily interpretable, we normalize  $R_c^{early}$  and  $R_c^{total}$  into standard deviation units. Across U.S. counties,  $R_c^{early}$  and  $R_c^{total}$  have standard deviations of 6.3 and 9.1 million dollars per square mile, respectively.

Table 1 presents descriptive statistics for the resource data, focusing on total endowment  $R_c^{total}$ . There are several basic facts to highlight. First, there are 715 more counties with undiscovered oil than counties that produce oil, and 1165 more counties with undiscovered gas than counties that produce gas. This highlights the exogeneity of our measure of resource endowments:  $R_c^{total}$  does not just include counties that produce in equilibrium. Second, multiplying total endowments by average prices shows that oil and gas are roughly equal constituents in total endowments, representing 42 and 58 percent of  $R_c^{total}$ , respectively. Third, there is substantial variation in  $R_c^{total}$  across counties: the standard deviation across all counties (including those with zero endowments) is about three times the mean. This variation is also highlighted in Figures 3 and 4, which show  $R_c^{total}$  and  $R_c^{total} - R_c^{early}$ , respectively.

Fourth, the primary constituent of oil endowment is the actual production over 1960-2011 – proven and undiscovered reserves are relatively small. For natural gas, by contrast, reserves are a much larger share of endowment. Combining oil and gas, 58 percent of  $R_c^{total}$  is from production, while the remaining 42 percent is from reserves. For  $R_c^{early}$ , which is not reported in the table, reserves are smaller, so production is 82 percent of endowment. This difference in composition between  $R_c^{early}$  and  $R_c^{total}$  highlights that one effect of the recent unconventional oil and gas boom has been to dramatically increase economically recoverable reserves, which have not yet been exploited.

---

that do not involve the Census of Manufactures data. For estimates that do not use Census of Manufactures data, we approximate early period proven reserves with data from the 1999 Oil and Gas Journal Data Book and current proven reserves using the EIA list of top 100 oil fields in 2009 (U.S. EIA 2009).

<sup>12</sup>See U.S. Geological Survey (2013) for more details on assessment methodology.

## 4.2 County-Level Economic Outcomes

Our primary source of data on employment, earnings, and population is the Regional Economic Information System (REIS), which at the time of this writing has county-level data available from 1969 to 2014. We use the REIS for national-level coal employment and oil and gas employment, as well as county-level population, total employment, total earnings, manufacturing employment, and manufacturing earnings. We construct earnings per worker by dividing total earnings by total employment. Due to REIS confidentiality restrictions, manufacturing earnings and manufacturing employment are not disclosed for some smaller counties, so the sample sizes in those regressions will be smaller. See Appendix B.2 for additional details on the REIS. Because earnings per worker is an imperfect proxy for labor input costs, we study hourly wage data from the Current Population Survey in Appendix A5.

Nationwide historical housing rent data are not available on an annual basis. We use Decennial Census data on mean rents until the year 2000 and add the American Community Survey five-year estimates for 2009 and 2014. We use rents instead of housing prices because this better reflects the current cost of living in our model, whereas house prices presumably capitalize expectations of future market conditions, including the future duration of a natural resource boom or bust.

Table 2 describes the data on county-level economic outcomes. A few counties merge or split during our sample period. In these cases, we define counties at the most disaggregated level at which data are observed for a consistent geographic area. This gives a population of 3075 counties in the lower 48 states.

## 4.3 Manufacturing Census Microdata

The Census of Manufactures (CM) includes microdata for all manufacturing plants in the United States. The data include the county where the plant is located, its four-digit SIC code, as well as number of employees, total wage bill, value of materials inputs, and total revenues.<sup>13</sup> The CM microdata are available for 1963 and every five years beginning in 1967, i.e. 1972, 1977, ..., 2007. We convert all industry codes to four-digit year-1987 SIC codes.

We use plant-level revenue productivity (TFP-R) estimates made available by Foster, Grim, and Haltiwanger (2014) for years 1972-2007. These are standard Cobb-Douglas log-TFP-Rs estimated in OLS, with separate production function coefficients for each industry.

For about 6000 relatively-homogeneous products defined at the 7-digit SIC level, the CM records both physical production quantities and sales revenues.<sup>14</sup> We divide revenues by physical output to arrive at a plant-by-product-by-year dataset of manufacturing output prices. We drop imputed

---

<sup>13</sup>Employment and earnings data for non-responders and small plants with fewer than five employees is imputed and/or marked as an “administrative record.” For these plants, we use the employment and earnings data drawn from tax records, but we do not use any imputed variables.

<sup>14</sup> These are the data used by Foster, Haltiwanger, and Syverson (2008) to study the distinction between physical productivity and revenue productivity.

data, as well as any reported prices that differ from the 7-digit median by a factor of more than five.

We also construct a “county-level Manufacturing Census dataset” by collapsing across plants to the county-by-year level.

### 4.3.1 Industry Classifications

When using the CM, we also examine subsets of manufacturers that may be differentially affected by resource booms. We distinguish subsectors along two dimensions: tradability and linkage to oil and gas.

Using the Commodity Flow Survey, Holmes and Stevens (2014) calculate a measure of transportation costs for each four-digit SIC industry that is closely correlated with average product shipment distance. Ready-mixed concrete, ice, and newspapers have the highest  $\eta$ , while watches, x-ray equipment, space propulsion units, and aircraft parts have the lowest. We define a four-digit SIC industry as “tradable” if the Holmes and Stevens  $\eta^{loglog}$  is less than 0.8, which corresponds to an average shipment distance of approximately 500 miles. By this definition, 69 percent of four-digit manufacturing industries are tradable.

We classify four-digit SIC industries as upstream or downstream of the oil and gas sector using the Bureau of Economic Analysis Input-Output tables for 1987 (U.S. BEA 2014). For each industry, we calculate the direct oil and gas output share (the share of output purchased by the oil and gas sector) and the indirect oil and gas output share (the share of output purchased by the oil and gas sector through an intermediate industry), and we define the “upstream linkage share” as the sum of these two quantities. We define an industry as “upstream” of oil and gas if this upstream linkage share is larger than 0.1 percent. An industry is “downstream” if the oil and gas input cost share is larger than 0.1 percent.<sup>15</sup> We refer to an industry as “non-linked” if it is neither upstream nor downstream. Using such small cutoff values in defining upstream and downstream is conservative in the sense that “non-linked” industries have very limited linkage to oil and gas and thus should not be directly affected by that sector. 27 percent of industries are upstream, 2.5 percent are downstream, 73 percent are non-linked, and 2.1 percent (largely chemical plants) are both upstream and downstream. Appendix Table A2 presents the most-linked upstream industries (such as oil and gas field machinery and equipment, cement, lubricants, chemicals, and pipes) and downstream industries (such as petroleum refining, fertilizers, chemicals, and plastics).

---

<sup>15</sup> We do not add the analogous “indirect input share” because this is primarily a measure of electricity intensity, given that the electricity generation sector uses substantial amounts of natural gas.

## 5 Empirical Strategy

### 5.1 County-Level Data

As suggested by the model, our initial regressions use county-level panel data, measuring differences in economic outcomes for higher- vs. lower-endowment counties through the boom and bust cycles between 1969 and 2014. We use separate strategies to estimate relative effects vs. absolute effects, which we denote by  $\tau^r$  and  $\tau^a$ , respectively. In Appendix A.3, we use the Rubin Causal Model to show formally how  $\tau^r$  and  $\tau^a$  translate from the two-county model to the many-county empirical setting. Intuitively,  $\tau^r$  is the effect of a resource boom on the average difference in outcomes between higher- vs. lower-endowment counties. General equilibrium effects and other types of geographic spillovers are (by definition) part of the relative effect. By contrast,  $\tau^a$  is the average treatment effect on the treated: the average across counties of each county's difference in potential outcomes if it had higher vs. lower endowments. We must control for geographic spillovers when estimating the absolute effect.

Define  $Y_{ct}$  as an outcome in county  $c$  in year  $t$ , and define  $\mathbf{Y}_{0c}$  as a vector of two baseline values of the outcome from two different years at the beginning and end of the 1960s, with separate coefficients  $\mu_t$  for each year of the sample.<sup>16</sup> Also define  $\phi_{dt}$  as a vector of Census division-by-year indicator variables. To illustrate the relative effects of resource booms, we begin by graphing the estimates of  $\tau_t^r$  from the following equation, where the subscript  $t$  on  $\tau_t^r$  indicates that separate coefficients are estimated for each year of the sample:

$$\ln Y_{ct} = \tau_t^r R_{ct} + \lambda R_{ct} + \mu_t \ln \mathbf{Y}_{0c} + \phi_{dt} + \varepsilon_{ct}, \quad (11)$$

with 1969 as the omitted year in  $\tau_t^r$ . The estimated  $\tau_t^r$  show how outcomes such as population and wages covary with oil and gas endowments  $R_{ct}$  in each year. For example, with population or wages as the outcome, Prediction 1 is that  $\hat{\tau}_t^r$  should be relatively large during the boom of the late 1970s, as wages rise in counties with higher  $R_{ct}$  and population migrates to those counties. More generally, the  $\hat{\tau}_t^r$  for population and wages should covary with national oil and gas employment, both rising during booms and falling during busts. We omit the  $\tau_t^r$  for 1969, so  $\lambda$  measures the association between endowments and the outcome in 1969, while the  $\tau_t^r$  measure the difference between the associations in year  $t$  and in 1969.

---

<sup>16</sup>When the outcome  $Y_{ct}$  is population,  $\mathbf{Y}_{0c}$  is county  $c$ 's population in 1966 and 1960, from the U.S. Census Bureau. When the outcome is employment or earnings per worker,  $\mathbf{Y}_{0c}$  is county  $c$ 's values of that variable for 1968 and 1964, from County Data Books. There are no housing price or rent data available for 1950 or 1960, so when the outcome is housing rents,  $\mathbf{Y}_{0c}$  is 1940 mean rents and county  $c$ 's earnings per worker in 1968 and 1964. In regressions using the county-level Manufacturing Census dataset,  $\mathbf{Y}_{0c}$  is the 1967 and 1963 values of  $Y_c$ . Including two logged baseline values and allowing their coefficients to vary by year means that we control for how baseline levels and trends in  $\ln Y$  are associated with  $\ln Y$  in any later year. We add one to each variable before logging so that zero values do not become missing. The decision of how to address zeros makes little difference because there are very few zeros in the data.

This intuition suggests a “shift-share” regression in the spirit of Bartik (1991), in which our key independent variable is the interaction of county-level oil and gas endowment  $R_{ct}$  with time-series variation in national oil and gas employment, denoted  $E_t$ . For our primary estimates, we use a differencing approach instead of fixed effects, because errors  $\varepsilon_{ct}$  are sufficiently serially correlated that differences are more efficient. Our specification to estimate relative effects is:

$$\Delta \ln Y_{ct} = \tau^r R_{ct} \Delta \ln E_t + \lambda R_{ct} + \mu_t \ln \mathbf{Y}_{0c} + \phi_{dt} + \varepsilon_{ct}. \quad (12)$$

This and all other regressions use robust standard errors, clustered by state. Because  $Y$  and  $E$  are logged,  $\tau$  is the differential elasticity of outcomes with respect to  $E_t$  for counties with one standard deviation additional endowment per square mile.

Equation (12) provides an unbiased estimate of the relative effect  $\tau^r$  if  $R_{ct} \Delta \ln E_t \perp \varepsilon_{ct} \mid \{R_{ct}, \ln \mathbf{Y}_{0c}, \phi_{dt}\}$ , i.e. as long as resource booms are uncorrelated with unobserved economic trends, conditional on baseline outcomes and other controls. The fact that our sample covers multiple booms and a bust means that  $\Delta \ln E_t$  increases, then decreases, then increases again in particular years. It is difficult to imagine a confounder that has this same particular time trend and differentially affects resource-abundant counties.

The absolute effect  $\tau^a$  can be identified only under some additional assumption about the structure of spillovers; see Manski (2013). In our context, most geographic spillovers are local: for example, people are more likely to move across the state for a job than across the country. We thus assume that geographic spillovers from other counties scale with the distance from county  $c$ , and that counties outside a maximum radius of 400 miles are unaffected. Specifically, we calculate total oil and gas endowment (in units of dollars) for all other counties with centroids within eight different doughnuts around county  $c$ : 0-50, 50-100, ... , and 350-400 miles. We denote this total endowment in doughnut  $d$  around county  $c$  as  $R_{ctd}$ , where the  $t$  subscript again reflects the fact that we use early endowment until the year 2000 and total endowment thereafter. The following specification estimates the absolute effect:

$$\Delta \ln Y_{ct} = \tau^a R_{ct} \Delta \ln E_t + \sum_{d=1}^8 [\nu_d R_{ctd} \Delta \ln E_t + \vartheta_d R_{ctd}] + \lambda R_{ct} + \mu_t \ln \mathbf{Y}_{0c} + \phi_{dt} + \varepsilon_{ct}. \quad (13)$$

In this equation,  $\nu_d$  represents the effect of additional endowment in doughnut  $d$  on a county’s economic outcomes. This specification is analogous to the Miguel and Kremer (2004) approach to estimating geographic spillovers. Clarke (2015) shows that this provides an unbiased estimate of  $\tau^a$  if there are no spillover effects outside the maximum radius. This assumption would be violated if, for example, cities such as Denver and Houston grow during the boom as they provide management and financing for business operations in far-away counties. We consider this particular concern in

Section 6.7, and in general we present the absolute effects estimates with additional caution.

## 5.2 Census of Manufactures Microdata

We also use plant-level Census of Manufactures microdata to estimate how changes in resource booms affect continuing plants. We focus on relative effects so as to not need to impose assumptions about the nature of geographic spillovers. In the estimating equation below,  $f$  indexes plants,  $c$  indexes counties,  $\lambda_{nt}$  denotes the full interactions of four-digit SIC codes and years, and the difference operator  $\Delta$  now reflects a 5-year difference between rounds of the Census of Manufactures. The regression is:

$$\Delta \ln Y_{fct} = \tau^r R_{ct} \Delta \ln E_t + \beta R_{ct} + \phi_{dt} + \lambda_{nt} + \varepsilon_{fct}. \quad (14)$$

## 5.3 Long Differences

Finally, we also use a long difference specification to test for pre-1973 trends and measure long-term effects of the boom and bust of the 1970s and 1980s. To do this, we construct a county-level (not county-by-year) dataset of long differences  $\Delta \ln Y_c$  and regress this on endowment  $R_c^{early}$ , controlling for two baseline outcomes  $\ln \mathbf{Y}_{0c}$  and for census division fixed effects  $\phi_d$ :

$$\Delta \ln Y_c = \tau^r R_c^{early} + \mu \ln \mathbf{Y}_{0c} + \phi_d + \varepsilon_c. \quad (15)$$

# 6 Results

In this section, we present the results of the five empirical tests suggested at the end of Section 3, illustrating the possible Dutch Disease chain of events. We first present relative effects in Sections 6.1 through 6.6, including initial conditions and pre-1973 trends, effects on county-level aggregates, effects on the manufacturing sector, and long-term effects. In Section 6.7, we estimate geographic spillovers and absolute effects, and in Section 6.8 we estimate the welfare effects of resource booms.

## 6.1 Initial Conditions and Pre-1973 Trends

Table 3 describes initial conditions before the 1970s oil boom. Of course, these are “initial conditions” only in the sense that our main sample begins at the end of the 1960s. Most counties and states that produced oil and gas during the 1970s boom had already been producing for many years, so their local economies may have already been shaped by oil and gas abundance. The first two rows present population and total employment data from the 1969 REIS, while the remaining five rows show manufacturing employment from the 1967 Census of Manufactures. Column 1 shows the mean across all 3075 counties in the data. Manufacturing was about one-fifth of employment

at the time. Within manufacturing, 73 percent is “non-linked” by our conservative definition. Of non-linked employment, 73 percent is tradable by our definition, while the remainder is local.

Column 2 shows coefficients and standard errors from regressions of natural logs of each variable on oil and gas endowment per square mile  $R_c^{early}$ , controlling for census division fixed effects. Endowments are positively and significantly correlated with population and employment, and positively but not significantly correlated with manufacturing employment. Consistent with the Michaels (2010) results for several southern states, this implies that resource-abundant counties had grown faster over the decades since they began producing oil and gas. The confidential Census data show that the manufacturing sector correlation depends heavily on the type of manufacturing. A one standard deviation increase in oil and gas endowment is associated with an 8.9 percent increase in linked manufacturing employment. By contrast, the relationships between endowments and employment in other manufacturing subsectors are insignificant, and point estimates are negative.<sup>17</sup>

As Figures 1 and 2 in Section 2 showed, oil and gas prices and employment were relatively steady between 1969, when our main sample begins, and the 1973 oil price shock. The 1969-1972 period is thus an opportunity for “placebo tests”: given that national oil and gas activity was relatively stable, changes in economic outcomes should be uncorrelated with oil and gas endowments, conditional on our standard control variables. Table 4 provides such tests: each column presents a regression of the 1972-1969 difference in a logged outcome from the REIS on early endowment  $R_c^{early}$ , controlling for census division indicators and controls for baseline levels  $\ln \mathbf{Y}_{0c}$ , as formalized in Equation (15). The absence of pre-1973 trends supports our causal interpretation of the association of resource abundance with changes in outcomes during resource booms.<sup>18</sup> Of course, even if there were monotonic trends associated with resource abundance, our identification would still be credible because it exploits non-monotonic changes in the oil and gas sector: a boom, a bust, and a boom over a 46-year period.

## 6.2 Relative Effects on County-Level Aggregates

Prediction 1 of the model is that natural resource booms increase local population and wages, and this prediction is important because the wage increase is the first step in the potential Dutch Disease mechanism. How do oil and gas booms differentially affect these county-level aggregate economic outcomes for counties with higher vs. lower endowments? Figure 5 presents the estimates

---

<sup>17</sup>Including division fixed effects does affect the magnitudes and even the signs of these correlations. This is partially because as the map in Figure 3 shows, a large amount of oil and gas is in sparsely-populated Rocky mountain states. The sensitivity of the cross sectional correlations to the inclusion of such controls emphasizes the importance of exploiting time series variation in oil and gas booms.

<sup>18</sup>The table includes five of the six county-level outcomes that we consider in Tables 5, 8, and 10. The sixth is housing rents, which we cannot test in this framework because it is not observed annually (and furthermore is not available for 1950 or 1960).

of  $\tau_t^r$  from Equation (11) for employment, population, and earnings per worker using the REIS data. Our key time-series variable  $E_t$  (national oil and gas employment) is plotted in gray against the right axis. Each of the outcomes is highly procyclical with the resource boom. The figure also illustrates the dynamic adjustment to a local economic shock highlighted by Blanchard and Katz (1992). As the resource sector expands, total employment and wages rise immediately. Population adjusts more slowly, however, meaning that the short-run effects of a resource boom are to increase wages and decrease unemployment. Within one to two years, people migrate in search of higher wages, and this migration moderates wage changes. To account for this delayed population adjustment, we use the one-year lag of  $\Delta \ln E_t$  in all regressions with population as the outcome. Appendix Figure A3 plots each outcome separately and includes confidence intervals.

Table 5 presents formal estimates of  $\tau^r$ , the relative effects of oil and gas booms, using Equation (12).<sup>19</sup> The positive coefficients mean that outcomes in higher-endowment counties are statistically significantly more procyclical with national oil and gas employment than in lower-endowment counties. In other words, a boom significantly increases relative growth, and a bust significantly decreases relative growth. A boom that increases national oil and gas employment by 100 log points increases relative population, employment, housing rents, and earnings per worker by approximately 1.19, 2.82, 2.70 and 1.79 percent, respectively, for counties with one standard deviation additional endowment.

There are several reasons why an increase in county average earnings per worker during a resource boom might not translate to an increase in manufacturers' labor input costs. First, average labor quality could change during a resource boom. If higher-skill workers enter the labor force during a resource boom, either by transitioning from unemployment or by migrating from other counties, this could cause average earnings per worker to increase even as wages are unchanged for workers of the same skill type. Second, labor input per worker could change: if people work more hours during resource booms, earnings per worker would increase even if unit labor costs did not.

In Appendix A5, we use Current Population Survey microdata to address these concerns. The CPS records hourly wages and worker demographics. Using these data, we find estimates comparable to those in Table 5. Controlling for observable and unobservable worker characteristics does not statistically significantly change the estimates, and we find no effect of resource booms on average hours worked.

A third reason why county average wages or earnings per worker might not affect manufacturers is that manufacturing workers may be imperfectly substitutable with other workers. In the extreme, manufacturing wages could even be completely unaffected, even as county average wages rise. Using the REIS data, Column 5 of Table 5 shows that relative manufacturing earnings per worker are

<sup>19</sup>The population regression in column 1 has 3075 fewer observations than the employment regression in column 2 because we lose one year of data when we use the one-year lag of  $\Delta \ln E_t$ . The housing rent sample size is smaller because we observe this variable approximately every ten years instead of every year. Manufacturing earnings per worker and manufacturing employment have smaller sample sizes due to the REIS non-disclosure protocols.



also procyclical with resource booms. The coefficient is smaller, but not statistically different, than the coefficient in column 4. In Appendix A5, the point estimates of  $\tau^r$  for manufacturing workers are actually larger than for the average worker, although this again is not statistically different.

Appendix D presents further robustness checks. Appendix Table A6 shows that the results are similar when using  $R_c^{early}$  or  $R_c^{total}$  in all years in place of  $R_{ct}$ , measuring resource boom intensity by national-level oil and gas wage and salary employment or oil and gas prices instead of total oil and gas employment, or using an analogous fixed effects estimator instead of a difference estimator.<sup>20</sup> In Appendix Table A7, we show that for most of the six main outcomes, the overall procyclicality is driven both by increases during oil and gas booms and decreases during busts, although coefficients tend to be smaller for the post-2001 period. We have also constructed nationwide county-level coal endowment data analogous to our oil and gas data, which we use to study the effects of coal booms and busts in Appendix Table A8. Results in our nationwide data are broadly consistent with the Black, McKinnish, and Sanders (2005a, 2005b) estimates for several eastern states. As shown in Appendix Figure A4, however, the 1970s coal boom was about an order of magnitude smaller than the 1970s oil and gas boom in terms of national employment growth. Thus, although the coal boom occurred in a similar period and in some of the same counties as the oil and gas boom, controlling for its effects has very little impact on our estimates of  $\tau^r$  from oil and gas booms.

Per the model's Prediction 2, the magnitude of a resource boom's effects on tradable goods producers depends on the magnitude of the wage increase. The coefficients for Figure 5 show that during the second oil and gas boom in our sample (2007-2014), earnings per worker average 2.3 percent higher than their 1969 or 2002-2003 levels in counties with one standard deviation additional endowment. Should we think of this relative increase as large or small? We can benchmark these magnitudes in three ways.

First, given that the average labor input revenue share in manufacturing is on the order of 0.25, total costs increased by 0.6 percent of revenues over the 2007-2014 period. While this is small, it represents a much larger share of variable profits, and it may be exacerbated by increases in costs of any locally sourced materials inputs. Second, for 2011, the coefficient of variation across counties in earnings per worker is 21.7 percent. Thus, an increase in earnings per worker of 2.3 percent represents 0.11 of a standard deviation.

Third, we can compare the cross-county wage increase to estimates from the cross-*country* Dutch Disease literature. Sachs and Warner (2001, page 834) show that natural resource exports are strongly negatively correlated with GDP growth and argue that Dutch Disease (through currency appreciation) is the primary mechanism. They estimate that near the peak of the boom in 1979, a 10 percentage point increase in a country's share of natural resource exports in GDP was associated with a 6.9 percent increase in price level. For context, Nigeria's natural resource exports equaled

---

<sup>20</sup>When using oil and gas prices to measure booms, coefficients have comparable t-statistics but are mechanically smaller, because oil and gas booms historically have entailed larger proportional increases in oil and gas prices than in national oil and gas employment. This can be seen by comparing Figure 1 to Figure 2.

10 percent of GDP, and one standard deviation across countries is about eight percent. Thus, a one standard deviation increase in natural resources leads to about a 5.5 percent increase in exporters' local costs. This is about 2.4 times larger than the 2.3 percent cross-county increase we estimate for 2007-2014.

### 6.3 Relative Effects on County-Level Manufacturing Employment

After wages rise, that the next step of the Dutch Disease mechanism is that the tradable sector is crowded out, per the model's Prediction 2. Is that what happens in U.S. counties?

Figure 6 presents estimates of Equation (11) for manufacturing employment. There is no evidence of crowding out. To the contrary, relative manufacturing employment in resource-abundant counties is clearly procyclical with oil and gas booms. Manufacturing employment grows during the boom of the 1970s, drops off during the bust in the late 1980s, and grows again during the boom of the 2000s. In counties with one standard deviation additional endowment, manufacturing employment was two to three percent larger at the peaks of the two booms in our sample, compared to the 1969-1972 pre-boom period.

Returning to Table 5, Column 6 presents presents formal estimates of  $\tau^r$  for manufacturing employment. A boom that increases national oil and gas employment by 100 log points increases relative manufacturing employment by approximately 2.93 percent in counties with one standard deviation larger endowment. Remarkably, this point estimate is slightly larger than the estimate for total employment across all sectors presented in column 2.

### 6.4 County-Level Manufacturing Census Dataset: Subsectors and Alternative Outcomes

In the context of Prediction 2, manufacturing's procyclicality with resource booms suggests that a significant share manufacturing is locally-traded, and that growth in locally-traded manufacturing subsectors outweighs any contraction in traded subsectors. Do the data bear this out? Table 6 presents estimates of Equation (12) using the county-level Manufacturing Census dataset, which we constructed by collapsing the CM microdata across plants to the county-by-year level. Each panel presents a different outcome, while each column presents estimates with outcomes collapsed from different manufacturing sub-sectors.

Column 1 presents estimates for all manufacturing plants. The first panel examines employment, confirming the result from Table 5 that the sector is procyclical with oil and gas. The point estimates are slightly larger in the CM data but not statistically different. The second and third panels consider revenues ("total value of shipments") and investment. Revenues are even more procyclical than employment, and investment is even more procyclical than revenues.

Column 2 presents analogous results for plants that are upstream or downstream of oil and gas. Column 3 considers non-linked plants, and columns 4 and 5 divide the non-linked plants into subsamples that produce more vs. less tradable goods. Results are generally consistent with the

model’s Prediction 2. Non-linked and local industries are clearly procyclical with resource booms. In contrast, column 5 suggests that tradable subsectors are countercyclical. Point estimates for all variables other than investment have similar magnitude but opposite signs as the point estimates for local non-tradables, although revenue results are more noisy. Crucially, these results arise for well-defined subsets of manufacturing plants that would not be possible to isolate in aggregate data.

Do firms adjust to resource booms on the “intensive margin,” by hiring more workers within the same plant, or on the “extensive margin,” by opening and keeping open more plants? Because opening physical plants may involve larger sunk costs than hiring workers into existing plants, this could affect the persistence of a resource boom’s effects. In the bottom two panels, we see that different sub-sectors adjust differently on the intensive vs. extensive margins. In column 2, the linked sector adjusts largely on the extensive margin, with a much larger coefficient on number of plants than on employees per plant. In contrast, the non-linked sectors adjust more on the intensive margin, with the non-linked local sector significantly increasing the number of employees per plant, and the non-linked tradable sector significantly contracting on that margin.

## 6.5 Effects of Resource Booms on Manufacturing Revenue Productivity

In this section, we first study the contemporaneous effects of resource booms on manufacturing productivity. These effects could be an important part of the mechanism making aggregate manufacturing wages and employment procyclical with resource booms. We then study the evolution of productivity in the highly tradable manufacturing subsectors.

Table 7 presents estimates of the contemporaneous effects of oil and gas booms on manufacturing revenue productivity, using Equation (14) with plant-level data from the Census of Manufactures. The dependent variable is first differences of the natural log of outcomes within surviving plants over the five year periods between each CM. As in the previous table, each panel examines a different outcome, and the columns include the same subsets of plants. For each outcome, we present two sets of estimates: sub-table (a) has only census division-by-year fixed effects, while sub-table (b) also includes four-digit industry-by-year effects.

The first and second panels analyze two different measures of revenue productivity: value added per worker and revenue total factor productivity. Both tell qualitatively similar stories. For the manufacturing sector overall, revenue productivity is positively associated with oil and gas booms. These effects are smaller for TFP-R than value added per worker, although the value added effects are substantially attenuated by including industry-year controls. Effects appear to be stronger for local industries and those linked to the oil and gas sector. By contrast, in non-linked and tradable industries, neither measure of revenue productivity covaries with resource booms, with point estimates close to zero.

One concern might be that even though these are within-plant estimates, the apparent productivity increase for linked and local sectors could be driven by selection. Specifically, if resource

booms allow lower-productivity firms to operate during booms, a boom would allow more firms to remain in operation even if they receive a negative productivity shock. However, this potential selection effect would bias against our results, which show that the average continuing plant in linked and non-linked local industries experiences a productivity increase during a boom.

As Foster, Haltiwanger, and Syverson (2008) highlight, the revenue productivity effects in Table 7 could reflect either higher markups or higher physical productivity. Indeed, our model predicts higher prices for locally traded goods during resource booms. Given that the CM does not have physical output data for most plants, this cannot be tested conclusively. In Appendix Table A9, we show that resource booms do not statistically significantly affect prices in the subset of plants that report physical output. Because prices can only be calculated for relatively homogeneous goods, however, this subset of plants is both relatively unusual and relatively small. Thus, the results only tentatively suggest that the TFP-R effects may reflect physical productivity.

If the results in Table 7 do reflect physical productivity, what might be the mechanism? A natural starting point is reductions in Marshall’s (1890) three types of transport costs: goods, people, and ideas. “Goods” linkages could increase productivity if local economic growth allows goods to be procured from and delivered to more nearby suppliers and buyers, either in the resource sector or in other procyclical local sectors. This is consistent with Table 7 results, which show TFP-R growth in linked and local sectors, but not in tradables.

Worker flows seem to be less important than in Greenstone, Hornbeck, and Moretti (2010) or Serafinelli (2012). For the entire CPS-MORG panel, which includes 281,301 manufacturing workers, there are only 220 that transition from oil and gas to manufacturing. Of these, 130 go to refining, and in every other two-digit industry, only around 0.001 percent of incoming workers come from oil and gas. The “ideas” channel seems even less likely: much of the oil and gas sector’s innovation does not occur at the drilling site, and much of it seems unlikely to translate to non-resource industries.

Beyond the three classical transport costs, several other channels might contribute to the productivity effects. For example, plant-level economies of scale could be relevant, perhaps due to higher capacity utilization, although the CM microdata do not suggest strong scale economies. Furthermore, local government actions, from tax cuts to the infrastructure improvements studied by Michaels (2010), might improve measured productivity. However, this channel does not seem particularly important given the near-zero effects on tradable sector plants that would presumably also benefit. Tax cuts and infrastructure take time to implement, so they might be more relevant in a long-term cross-sectional analysis like Michaels (2010) than in these relatively short-term difference estimates. In summary, the contemporaneous TFP-R gains from natural resource booms are likely driven by some combination of higher output prices and lower transport costs for inputs and outputs.

Table 6 showed that only the highly tradable manufacturing subsectors appear to contract during natural resource booms. Thus, if we are to see any evidence of Dutch Disease, it will be in

these subsectors. Per the model’s Prediction 3, the next step in the Dutch Disease chain of events would be that the tradable sector contraction leads to reduced productivity. Is this what we see in the data, or is productivity unaffected?

Figure 7 shows the evolution of employment and TFP-R for the tradable manufacturing subsectors, presenting the  $\tau_t^r$  coefficients from estimates of Equation (11) in county-by-year data. (We calculate county-level TFP-R by taking the mean across all plants in the county-year cell.) Employment is generally countercyclical, as suggested by the negative coefficient in Table 6: relative employment in resource-abundant counties contracted during the boom of the 1970s and 1980s before recovering through the 1990s. Thus, if there were foregone learning-by-doing effects, relative productivity would also eventually become negative. In reality, however, we see that relative TFP-R is never negative, implying no significant loss of learning-by-doing spillovers. With 90 percent confidence, we can rule out the TFP-R dropped more than 0.39 percent in any census year in counties with one standard deviation additional endowment. Connecting to Prediction 3, this shows that any effects of resource booms on tradable sector productivity are economically and statistically insignificant. Thus, this key link in the possible Dutch Disease chain of events is broken.

## 6.6 Long-Term Effects

Predictions 3 and 4 highlight the importance of measuring the long-term effects of resource booms. Prediction 3 states that if either local sector learning-by-doing  $\theta_l$  or agglomeration  $\Lambda$  is positive, local sector productivity will be higher in higher-endowment counties after the boom is over. For the tradable sector, post-boom relative productivity effects depend on whether foregone learning-by-doing outweighs the agglomeration effect; Figure 7 showed that the overall effects are insignificant. Of course, in a more general context outside of our model, a resource boom could have long-term effects through other mechanisms, such as adjustment costs, labor market search frictions, or other frictions that can deepen recessions following booms, as well as persistent changes in environmental quality or development of local infrastructure and public goods. These issues motivate empirical tests of the long-term impacts of oil and gas booms on productivity and other outcomes.

The most recent boom and bust cycle is not complete as of the end of our sample in 2014. However, we can measure the relative effects of the 1970s and 1980s boom and bust, as of the end of the 1990s—safely before the more recent boom began. To do this, we use the long differences regression in Equation (15), regressing the change in log outcomes between 1972 and 1997 on  $R_c^{early}$ , controlling for baseline outcomes  $\ln \mathbf{Y}_{0c}$  and census division fixed effects  $\phi_d$ .

Table 8 examines the long-term effects of the boom and bust on log TFP-R. For the top panel, we collapse plant-level log TFP-R to county means and difference between 1972 and 1997 to generate the long difference, then estimate Equation (15). The bottom panel of Table 8 instead uses plant-level data for 1972 and 1997, where we run a difference-in-differences estimator, controlling for county fixed effects and 4-digit SIC-by-year effects and reporting the interaction of the 1997

year indicator with resource endowment  $R_c^{early}$ . Both specifications find no significant differences in productivity in any sector. Thus, the TFP-R effects measured in Table 7, plus any potential longer-term effects, cancel out over the boom and bust of the 1970s and 1980s.

Table 9 presents estimates of Equation (15) for each of the six main county-level outcomes from Table 5. To match the CM years, we also use 1997-1972 long differences, except for housing rents, where we use 2000-1970. The results here are foreshadowed by Figure 5, as well as Appendix Figure A3: these figures show that the coefficients in the late 1990s are very close to zero, implying that the bust has unwound the effects of the boom. Indeed, for all six outcomes, the long difference results confirm zero long-term effects of the 1970s and 1980s boom and bust.

In the context of our model, these results also provide insight as to whether resource booms have lasting impacts on amenities such as environmental quality. If there are no productivity differences, then the model predicts that population and wages differ if and only if there are amenity differences.<sup>21</sup> The fact that there are no statistically significant long-term effects on productivity, population, or wages therefore suggests that people perceive equal amenities in resource-abundant counties. Thus, the oil and gas boom and bust of the 1970s and 1980s had no detectable negative long-term effects on average amenities. This counters concerns about potential harm from oil and gas extraction using conventional technologies. However, we emphasize that our estimates from the conventional boom of the 1970s and 1980s do not generalize to some current questions about the environmental effects of fracking.

## 6.7 Geographic Spillovers and Absolute Effects

So far, we have studied the relative effects  $\tau^r$  of natural resource booms on higher- vs. lower-endowment counties. We now estimate geographic spillovers and absolute effects  $\tau^a$ , using Equation (13). While the model focuses on general equilibrium effects induced by migration, geographic spillovers could also occur through other channels. For example, producer states may redistribute tax revenues to their non-producer counties, and there may be business demand spillovers as firms in non-producer counties expand to serve higher demand in nearby producer counties.

The expected sign of the spillovers depends on the outcome and the channel. Migration and business demand spillovers both suggest positive wage spillovers to nearby counties. By contrast, direct migration predicts negative population spillovers, but business demand spillovers increase labor demand in nearby counties, which would draw migrants from other more distant counties.

Even if the sign of the geographic spillover is known, the sign of the difference between  $\tau^r$  and  $\tau^a$  is ambiguous, depending on the geographic arrangement of treatment units in the sample, i.e. the geographic clustering of oil and gas endowments across the United States. Imagine momentarily that the average treatment effect on the treated  $\tau^a$  is positive. In one extreme, if treatment counties

---

<sup>21</sup>See Equations (29) and (30) in Appendix A.

are uniformly geographically dispersed,  $\tau^r < \tau^a$  with positive geographic spillovers, and  $\tau^r > \tau^a$  with negative geographic spillovers. Intuitively, positive spillovers increase outcomes for many control units, attenuating the difference between treatment and control, and negative spillovers reduce outcomes for many control units, increasing the difference between treatment and control.

In another extreme, however, treatment counties could be clustered together such that all geographic spillovers impact other treatment counties, with no impact on control. In this case, we have the opposite result:  $\tau^r > \tau^a$  with positive spillovers, and  $\tau^r < \tau^a$  with negative spillovers. Intuitively, positive spillovers increase outcomes for treatment units only, magnifying the difference between treatment and control, while negative spillovers reduce outcomes for treatment units only, reducing the difference between treatment and control. Reality is somewhere between these two extremes: figures 3 and 4 show that oil and gas endowments are moderately geographically clustered. Thus,  $\tau^a$  could be larger or smaller than  $\tau^r$ .

Table 10 presents estimates of  $\tau^a$  from Equation (13) for the six main county-level outcome variables. In each case,  $\hat{\tau}^a$  is smaller than the corresponding  $\hat{\tau}^r$  from Table 5. Figure 8 presents the estimates of  $\nu_r$  for each of the eight different distance doughnuts. Spillovers from nearby counties are statistically significantly positive for all outcomes.

Population is close to a zero-sum game, as resource booms will have limited effect on total nationwide population. This explains why population spillovers in the top left panel of Figure 8 tend negative from counties with centroids between 300 and 400 miles from county  $c$ : the migration into producer counties and other nearby counties has to come from more distant counties. Consistent with reduced population (and thus reduced housing demand), housing rents also tend negative in the 300-400 mile doughnuts. For earnings per worker and manufacturing earnings per worker, spillovers are statistically zero after about 150 miles. This is roughly consistent with Feyrer, Mansur, and Sacerdote (2015), who find that drilling during the recent fracking boom only impacted wages and employment in other counties within about a 100-mile radius.

As discussed in Section 5, in order for the  $\tau^a$  from Equation (13) to be an unbiased estimate of the absolute effect, there must be no spillover effects not captured in the empirical model – i.e. no spillovers from counties more than 400 miles away. This is particularly important for earnings per worker and housing rents, given that these variables enter our welfare analysis in the next section. The statistically zero effects for these outcomes from endowments more than 150 miles from county  $c$  supports the identifying assumption. One additional factor that could violate this assumption is if large cities such as Denver or Houston benefit from resource booms by providing management, financial services, or inputs to firms in distant counties. To test this, Appendix Table repeats the estimates in Table 10, after excluding counties whose populations were larger than 250,000 people in 1960. The estimates are almost exactly the same.

## 6.8 Social Welfare Effects

In our final set of empirical results, we present the social welfare effects of oil and gas booms over our sample. In our model, the zero long-term impacts on productivity documented in Section 6.6 already imply that resource booms increase welfare in both relative and absolute terms. Two questions remain. First, could there be mechanisms not considered in the model that cause resource booms to reduce welfare? Second, quantitatively how large are the welfare effects?

We first quantify effects on “real earnings,” i.e. earnings net of the increase in local prices. These equal welfare effects under the assumption that resource booms do not affect amenities. To quantify real earnings effects, we combine the formal estimates of  $\tau^r$  and  $\tau^a$  from Equations (12) and (13) into an empirical analogue of the welfare formulas in Equations (5) and (7).<sup>22</sup> Specifically, beginning with Equations (5) and (7), we substitute empirical estimates  $\hat{\tau}^r$  for  $(w_{at} - w_{bt})$  and other relative effects, and  $\hat{\tau}^a$  for  $(w_{at} - w_{at}(0))$  and other absolute effects. We must also incorporate the actual length and magnitude of booms and busts over our 1969-2014 sample. To do this, we multiply the treatment effects by the sum of annual changes in national oil and gas employment  $\Delta E_t$ , where each annual change is weighted by  $(T - t + 1)$ , the number of remaining years in the sample. Dividing by 45 then gives the annual average effect. Thus, from the perspective of the 1969 policymaker, the average real earnings effect of natural resource booms and busts over 1969-2014 is

$$\hat{W} = \frac{1}{45} \sum_{t=1970}^{T=2014} [\delta^{t-1969} \Delta \ln E_t \cdot (T - t + 1) \cdot (\hat{\tau}_{earnings} - \alpha \hat{\tau}_{rent} - \beta \hat{\tau}_{local})], \quad (16)$$

where  $\hat{\tau}$  can be either  $\hat{\tau}^r$  or  $\hat{\tau}^a$ . Because  $\hat{\tau}$  is in log units, this formula approximates the percent change in real earnings. We use  $\delta = 1$ . Following Hornbeck and Moretti (2015), we assume that the prices of all locally-traded goods move proportionally with housing prices, and we assume that the expenditure share on local goods plus housing is  $\alpha + \beta = 0.3$ .

Table 11 presents results, with standard errors in parentheses. Column 1 presents relative effects using  $\hat{\tau}^r$ , while column 2 presents absolute effects using  $\hat{\tau}^a$ . The first two rows present regression coefficients for earnings per worker and housing rents from Equations (12) and (13). For housing rents, these are the same as the estimates in Tables 5 and 10. For consistency, we use the same estimation sample (with data from 1970, 1980, 1990, 2000, 2009, and 2014) to estimate both  $\hat{\tau}_{earnings}$  and  $\hat{\tau}_{rent}$ , so the coefficients for earnings per worker are slightly different than in the previous tables. Row 3 shows that the relative and absolute real earnings gains from a boom causing a 100 log point increase in national oil and gas employment are 1.6 and 0.7 percent, respectively.

<sup>22</sup>To see the intuition, notice that the relative earnings gains over our sample are the integral under the 1969-2014 earnings effects illustrated in Figure 5. To calculate the real relative earnings gains, we would then subtract off the local good and housing expenditure shares times the integral under the equivalent figures for those goods. This would then reflect the relative real earnings effect of one additional standard deviation oil and gas endowment over the 1969-2014 period.



Row 4 presents calibrations of Equation (16). The annual average real earnings gain from one standard deviation additional oil and gas endowment is approximately 1.0 percent relative to other counties during the boom, and approximately 0.4 percent relative to a county’s own counterfactual.

These real earnings gains equal social welfare gains only if resource booms do not affect amenities. As Muehlenbachs, Spiller and Timmins (2015) and others have shown, however, oil and gas production may affect environmental quality. Equation (6) from the model shows that the sign of the relative welfare effect is the same as the average population change. Intuitively, people migrate to producer counties if and only if their utility is higher there. Row 5 of Table 11 presents estimates of  $\tau^r$  and  $\tau^a$  for population using the same sample of years as rows 1 and 2, and row 6 presents calibrations of Equation (16) after substituting  $\hat{\tau}_{population}$  for  $\hat{\tau}_{earnings} - \alpha\hat{\tau}_{rent} - \beta\hat{\tau}_{local}$ . That final row shows that population averages approximately 0.8 percent higher in counties with one standard deviation additional endowment over the full 1969-2014 period. Thus by this metric as well, oil and gas endowments have increased relative welfare.

## 7 Conclusion

The dramatic oil and gas price changes and improvements in drilling technology over the past 15 years raise the important question of how resource booms and busts shape the economic geography of the United States. The agglomeration literature suggests that booms might bring both short-term benefits and, due to persistent productivity spillovers, cause long-term economic transformations in resource rich areas. Alternatively, there might instead be a “Natural Resource Curse,” perhaps due to a version of the Dutch Disease mechanism in which wage increases crowd out externality-generating manufacturing establishments. To test between these alternative mechanisms, we combine a new panel dataset of oil and gas production and reserves with restricted-access microdata from the Census of Manufactures to estimate how oil and gas booms have differentially affected U.S. counties since the 1960s.

Our empirical results document that oil and gas booms significantly increase wages, even as people migrate to producer counties. Notwithstanding, overall manufacturing employment, output, and revenue productivity are actually procyclical with resource booms. Although resource booms do seem to crowd out more highly-tradable manufacturing, there are no temporary or permanent reductions in tradable subsectors’ revenue productivity. This implies that a key link in the Dutch Disease chain of events is not active. Across outcomes, we see no “long-run” effects: by the end of the 1990s, the boom and bust of the 1970s and 1980s had zero significant effects on TFP-R and other key outcomes. Over the full 1969-2014 sample, a county with one standard deviation additional oil and gas endowment averaged 1.0 percent higher wages when compared to less resource-abundant counties, and 0.4 percent higher real wages compared to its own counterfactual.

What might be the relationship between these results and the cross-country estimates? Might

a comparable strategy in country-level data “generalize down” to the U.S., and might our results “generalize up” to the cross-country setting? As discussed in Section 6, the currency appreciation caused by country-level natural resource exports appears to be larger than the wage appreciation caused by county-level oil and gas production. Furthermore, our estimates do not capture any learning-by-doing spillovers that happen at the country level but not at the county level. Thus, it is possible that the mechanism of crowding out plus foregone positive productivity spillovers is more relevant for countries than for counties. On the other hand, while manufacturers are often thought of as producing nationally- or internationally-traded goods, this paper echoes Holmes and Stevens (2014) in highlighting how a meaningful share of manufacturers benefit from highly localized demand growth. Since we find that natural resource booms exert positive spillovers to manufacturers in the same *county*, they surely do so for manufacturers in the same *country*.

In summary, while post-1960 resource booms did not produce strong agglomerative forces, we also find no evidence for a Dutch Disease mechanism in the United States. On net, our estimates suggest that oil and gas abundance has increased welfare in producer areas in modern American history.

## References

- [1] Acemoglu, Daron, Amy Finkelstein, and Matthew Notowidigdo (2013). "Income and Health Spending: Evidence from Oil Price Shocks." *Review of Economics and Statistics*, Vol. 95, No. 4 (October), pages 1079-1095.
- [2] Aragon, Fernando, and Juan Pablo Rud (2011). "Natural Resources and Local Communities: Evidence from a Peruvian Gold Mine." Working Paper, Simon Fraser University (April).
- [3] Asher, Sam, and Paul Novosad (2014a). "Dirty Politics: Natural Resource Wealth and Politics in India." Working Paper, Oxford University (March).
- [4] Asher, Sam, and Paul Novosad (2014b). "Digging for Development: Mining Booms and Local Economic Development in India." Working Paper, Oxford University (April).
- [5] Autor, David, David Dorn, and Gordon Hanson (2013). "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, Vol. 103, No. 6 (October), pages 2121-2168.
- [6] Auty, Richard, and Raymond Mikesell (1998). Sustainable Development in Mineral Economies. Oxford, UK: Oxford University Press.
- [7] Bartik, Alexander, Janet Currie, Michael Greenstone, and Chris Knittel (2016). "The Local Economic and Welfare Consequences of Hydraulic Fracturing." Work in progress. Abstract available from <http://economics.mit.edu/grad/abartik/research>.
- [8] Bartik, Timothy (1991). Who Benefits from State and Local Economic Development Policies? Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.
- [9] Benkard, Lanier (2000). "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review*, Vol. 90, No. 4 (September), pages 1034-1054.
- [10] Black, Dan, Terra McKinnish, and Seth Sanders (2005a). "The Economic Impact of the Coal Boom and Bust." *The Economic Journal*, Vol. 115, No. 503 (April), pages 449-476.
- [11] Black, Dan, Terra McKinnish, and Seth Sanders (2005b). "Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust." *Industrial and Labor Relations Review*, Vol. 59, No. 1 (October), pages 3-16.
- [12] Blanchard, Olivier, and Lawrence Katz (1992). "Regional Evolutions." *Brookings Papers on Economic Activity*, Vol. 1992, No. 1, pages 1-75.
- [13] Bleakley, Hoyt, and Jeffrey Lin (2012). "Portage and Path Dependence." *Quarterly Journal of Economics*, Vol. 127, No. 2, pages 587-644.
- [14] Brown, Chip (2013). "North Dakota Went Boom." *The New York Times Magazine*, January 31, 2013.
- [15] Brown, Eliot (2015). "Oil Boom Swells North Dakota Town; What Now?" *The Wall Street Journal* March 17, 2015.
- [16] Brunnschweiler, Christa, and Erwin Bulte (2008). "The Resource Curse Revisited and Revised: A Tale of Paradoxes and Red Herrings." *Journal of Environmental Economics and Management*, Vol. 55, No. 3, pages 248-264.
- [17] Busso, Matias, Jesse Gregory, and Patrick Kline (2013). "Assessing the Incidence and Efficiency of a Prominent Place-Based Policy." *American Economic Review*, Vol. 103, No.2 (April), pages 897-947.
- [18] Carrington, William (1996). "The Alaskan Labor Market during the Pipeline Era." *Journal of Political Economy*, Vol. 104, No. 1 (February), pages 186-218.

- [19] Caselli, Francesco, and Guy Michaels (2013). “Do Oil Windfalls Improve Living Standards? Evidence from Brazil.” *American Economic Journal: Applied Economics*, Vol. 5, No. 1 (January), pages 208-238.
- [20] Clarke, Damian (2015). “Estimating Difference-in-Differences in the Presence of Spillovers: Theory and Application to Contraceptive Reforms in Latin America.” Working Paper, Catholic University of Chile (August). Available from <http://economia.uc.cl/wp-content/uploads/2015/07/Clarke-Spillovers.pdf>
- [21] Collier, Paul, and Benedikt Goderis (2009). “Commodity Prices, Growth, and the Natural Resource Curse: Reconciling a Conundrum.” Working Paper, Oxford University (December).
- [22] Corden, Max, and Peter Neary (1982). “Booming Sector and De-Industrialisation in a Small Open Economy.” *The Economic Journal*, Vol. 92, No. 368 (December), pages 825-848.
- [23] Domenech, Jordi (2008). “Mineral Resource Abundance and Regional Growth in Spain, 1860-2000.” *Journal of International Development*, Vol. 20, pages 1122-1135.
- [24] Dube, Oeindrila, and Juan Vargas (2013). “Commodity Price Shocks and Civil Conflict: Evidence from Colombia.” *Review of Economic Studies*, Vol. 80, pages 1384-1421.
- [25] Ellison, Glenn, and Edward Glaeser (1997). “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach.” *Journal of Political Economy*, Vol. 105, No 5, pages 889-927.
- [26] Ellison, Glenn, Edward Glaeser, and William Kerr (2010). “What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns.” *American Economic Review*, Vol. 100, No. 3 (June), pages 1195-1213.
- [27] Feyrer, James, Erin T. Mansur, and Bruce Sacerdote (2015). “Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution.” National Bureau of Economic Research Working Paper No. 21624 (October).
- [28] Foster, Lucia, Cheryl Grim, and John Haltiwanger (2014). “Reallocation in the Great Recession: Cleansing or Not?” NBER Working Paper No. 20427 (August).
- [29] Foster, Lucia, John Haltiwanger, and Chad Syverson (2008). “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, Vol. 98, No. 1 (March), pages 394-425.
- [30] Glaeser, Edward, and Joshua Gottlieb (2008). “The Economics of Place-Making Policies.” *Brookings Papers on Economic Activity*, Vol. 1, pages 155-239.
- [31] Glaeser, Edward, and Joshua Gottlieb (2009). “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature*, Vol. 47, No. 4, pages 983-1028.
- [32] Glaeser, Edward, Sari Pekkala Kerr, and William Kerr (2015). “Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines.” *Review of Economics and Statistics*, Vol. 97, No. 2 (May), pages 498-520.
- [33] Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010). “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings.” *Journal of Political Economy*, Vol. 118, No. 3, pages 536-598.
- [34] Gylfason, Thorvaldur, Tryggvi Herbertsson, and Gylfi Zoega (1999). “A Mixed Blessing: Natural Resources and Economic Growth.” *Macroeconomic Dynamics*, Vol. 3, pages 204-225.
- [35] Haggerty, Julia, Patricia Gude, Mark Delorey, and Ray Rasker (2014). “Long-Term Effects of Income Specialization in Oil and Gas Extraction: The U.S. West, 1980-2011.” *Energy Economics*, Vol. 45, No. C, pages 186-195.

- [36] Harding, Torfinn, and Anthony Venables (2013). “The Implications of Natural Resource Exports for Non-Resource Trade.” OxCarre Research Paper 103 (January).
- [37] Holmes, Thomas, and John Stevens (2014). “An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade.” *Journal of Political Economy*, forthcoming.
- [38] Hooker, Mark, and Michael Knetter (2001). “Measuring the Economic Effects of Military Base Closures.” *Economic Inquiry*, Vol. 39, No. 4 (October) pages 583-598.
- [39] Hornbeck, Richard, and Enrico Moretti (2015). “Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality.” Working Paper, University of Chicago (May).
- [40] Ismail, Kareem (2010). “The Structural Manifestation of the ‘Dutch Disease’: The Case of Oil Exporting Countries.” IMF Working Paper 10/103 (April).
- [41] Jacobsen, Grant (2015). “Who Wins in an Energy Boom? Evidence from Wage Rates and Housing.” Working Paper, University of Oregon (October).
- [42] Jacobsen, Grant, and Dominic Parker (2014). “The Economic Aftermath of Resource Booms: Evidence from Boomtowns in the American West.” *Economic Journal*, forthcoming.
- [43] James, Alexander, and David Aadland (2011). “The Curse of Natural Resources: An Empirical Investigation of U.S. Counties.” *Resource and Energy Economics*, Vol. 33, pages 440-453.
- [44] James, Alexander, and Robert James (2011). “Do Resource Dependent Regions Grow Slower than they Should?” *Economics Letters*, Vol. 111, pages 194-196.
- [45] Kline, Patrick, and Enrico Moretti (2014). “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority.” *Quarterly Journal of Economics*, Vol. 129, No. 1 (February), pages 275-331.
- [46] Krugman, Paul (1987). “The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher.” *Journal of Development Economics*, Vol. 27, pages 41-55.
- [47] Lederman, Daniel, and William Maloney, Eds. (2007a). Natural Resources: Neither Curse Nor Destiny. Stanford, CA: Stanford University Press.
- [48] Lederman, Daniel, and William Maloney (2007b). “Trade Structure and Growth.” In Lederman, Daniel, and William Maloney, Eds., Natural Resources: Neither Curse Nor Destiny. Stanford, CA: Stanford University Press.
- [49] Madrian, Brigitte, and Lars Lefgren (1999). “A Note on Longitudinally Matching Current Population Survey (CPS) Respondents.” NBER Technical Working Paper 247 (November).
- [50] Manski, Charles (2013). “Identification of Treatment Response with Social Interactions.” *Econometrics Journal*, Vol. 16, pages S1-S23.
- [51] Manzano, Osmel, and Roberto Rigobon (2001). “Resource Curse or Debt Overhang?” NBER Working Paper 8390 (July).
- [52] Marshall, Alfred (1890). Principles of Economics. New York, NY: MacMillan.
- [53] Matsuyama, Kiminori (1992). “Agricultural Productivity, Comparative Advantage, and Economic Growth.” *Journal of Economic Theory*, Vol. 58, pages 317-334.
- [54] Mehlum, Halvor, Karl Moene, and Ragnar Torvik (2006). “Cursed by Resources or Institutions?” *The World Economy*, pages 1117-1131.
- [55] Melitz, Marc (2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, Vol. 71, No. 6, pages 1695-1725.

- [56] Michaels, Guy (2010). "The Long Term Consequences of Resource-Based Specialisation." *The Economic Journal*, Vol. 121 (March), pages 31-57.
- [57] Miguel, Ted, and Michael Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, Vol. 72, No 1 (January), pages 159-217.
- [58] Monteiro, Joana, and Claudio Ferraz (2012). "Does Oil Make Leaders Unaccountable? Evidence from Brazil's Offshore Oil Boom." Working Paper, Pontificia Universidade Catolica do Rio de Janeiro (May).
- [59] Moretti, Enrico (2010). "Local Multipliers." *American Economic Review*, Vol 100, No. 2 (May), pages 1-7.
- [60] Moretti, Enrico (2010). "Local Labor Markets." *Handbook of Labor Economics*, Vol. 4b, pages 1237-1313.
- [61] Moretti, Enrico (2013). "Where the Good Jobs Are - and Why." *Wall Street Journal*, September 17th. <http://eml.berkeley.edu/~moretti/wsj.pdf>
- [62] Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins (2015). "The Housing Market Impacts of Shale Gas Development." *American Economic Review*, Vol. 105, No. 12 (December), pages 3633-3359.
- [63] Munasib, Abdul, and Dan Rickman (2015). "Regional Economic Impacts of the Shale Gas and Tight Oil Boom: A Synthetic Control Analysis." *Regional Science and Urban Economics*, Vol. 50 (January), pages 1-17.
- [64] Notowidigdo, Matthew (2013). "The Incidence of Local Labor Demand Shocks." Working Paper, Northwestern University (March).
- [65] Papyrakis, Elissaios, and Reyer Gerlagh (2007). "Resource Abundance and Economic Growth in the United States." *European Economic Review*, Vol. 51, pages 1011-1039.
- [66] Paredes, Dusan, Timothy Komarek, and Scott Loveridge (2015). "Assessing the Income and Employment Effects of Shale Gas Extraction Windfalls: Evidence from the Marcellus Region." *Energy Economics*, Vol. 47 (January), pages 112-120.
- [67] Rosenbaum, Liz (2016). "List of Bans Worldwide." <https://keptapwatersafe.org/global-bans-on-fracking/>.
- [68] Sachs, Jeffrey, and Andrew Warner (1995). "Natural Resource Abundance and Economic Growth." NBER Working Paper 5398 (December).
- [69] Sachs, Jeffrey, and Andrew Warner (1999). "The Big Push, Natural Resource Booms, and Growth." *Journal of Development Economics*, Vol. 59, pages 43-76.
- [70] Sachs, Jeffrey, and Andrew Warner (2001). "The Curse of Natural Resources." *European Economic Review*, Vol. 45, pages 827-838.
- [71] Serafinelli, Michel (2012). "'Good' Firms, Worker Flows, and Local Productivity." Working Paper, UC Berkeley (November).
- [72] Smith, Adam (1776 [1976]). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Oxford: Clarendon Press.
- [73] U.S. BEA (Bureau of Economic Analysis) (2014). "Gross-Domestic-Product-(GDP)-by-Industry Data." <http://www.bea.gov/industry/gdpbyind.data.htm>
- [74] U.S. EIA (Energy Information Administration) (2009). "Top 100 Oil and Gas Fields of 2009." [www.eia.gov/naturalgas/crudeoilreserves/archive/2009/pdf/top100fields.pdf](http://www.eia.gov/naturalgas/crudeoilreserves/archive/2009/pdf/top100fields.pdf)
- [75] U.S. EIA (Energy Information Administration) (2014). "EIA-23L Reserves Information Gathering System." <http://www.eia.gov/survey/form/eia.23l/rigs.cfm>

- [76] U.S. Geological Survey (2013). “National Oil and Gas Assessment.” <http://energy.usgs.gov/OilGas/AssessmentsData/NationalOilGasAssessment/Methodology.aspx>.
- [77] van der Ploeg, Frederick (2011). “Natural Resources: Curse or Blessing?” *Journal of Economic Literature*, Vol. 49, No. 2, pages 366-420.
- [78] van Wijnbergen, Sweder (1984). “The ‘Dutch Disease’: A Disease After All?” *The Economic Journal*, Vol. 94, No. 373 (March), pages 41-55.
- [79] Weber, Jeremy (2012). “The Effects of a Natural Gas Boom on Employment and Income in Colorado, Texas, and Wyoming.” *Energy Economics*, Vol. 34, No. 5 (September), pages 1580-1588.
- [80] Wright, Gavin, and Jesse Czelusta (2007). “Resource-Based Growth Past and Present.” In Lederman, Daniel, and William Maloney, Eds., Natural Resources: Neither Curse Nor Destiny. Stanford, CA: Stanford University Press.

## Tables

Table 1: Oil and Gas Endowments: Descriptive Statistics

	Mean	SD	Max	N>0
<b>Oil (million barrels)</b>				
Production 1960-2011	33.4	229	8993	1141
Current proven reserves	2.58	43.5	2152	109
Current undiscovered reserves	7.11	41.0	878	1856
1960-2011 average price (\$/barrel)	34.92			
<b>Natural Gas (billion cubic feet)</b>				
Production 1960-2011	273	1238	23183	1117
Current proven reserves	50.2	457	19676	239
Current undiscovered reserves	314	1471	51014	2282
1960-2011 average price (\$/mcf)	3.20			
1960 oil and gas endowment (\$million/square mile)	3.1	9.1	209	2292

Notes: This table presents the components of total 1960 oil and gas endowment for the 3075 counties in our sample. The 1960 oil and gas endowment is the product of physical quantities of oil and gas production and reserves with their average prices over 1960-2011, divided by county land area. Prices are in real 2010 dollars.

Table 2: County-Level Outcomes: Descriptive Statistics

	N	Mean	SD	Min	Max
Population (000s)	141,424	83.1	274	0.06	10,117
Employment (000s)	141,424	44.7	164	0.06	6088
Housing rent (\$/month)	18,446	576	180	0	1795
Earnings per worker (\$000s)	141,424	30.7	6.9	11.4	105
Manufacturing earnings per worker (\$000s)	114,192	46.6	16.6	2.47	600
Manufacturing employment (000s)	114,238	7.0	24.7	0	950

Notes: Housing rent is from the 1970, 1980, 1990, and 2000 Decennial Censuses and the 2009 and 2014 American Community Survey five-year estimates. All other variables are from the Regional Economic Information System, for years 1969-2014. Prices are in real 2010 dollars.



Table 3: **Baseline Employment Characteristics and Correlation with Oil and Gas Endowments**

	(1)	(2)
	Mean (000s)	Association with endowment $R_c^{early}$
Population	63.1	0.073** (0.023)
Total employment	29.2	0.080** (0.022)
Manufacturing employment	5.99	0.028 (0.046)
Up/downstream manufacturing employment	1.65	0.089* (0.052)
Non-linked manufacturing employment	4.35	-0.020 (0.049)
Non-linked local manufacturing employment	1.16	-0.017 (0.045)
Non-linked tradable manufacturing employment	3.19	-0.034 (0.062)

Notes: Population and total employment are from the 1969 REIS. Manufacturing employment is from the 1967 Census of Manufactures. Column 2 presents the coefficient of a regression of natural log of the listed variable on early period oil and gas endowment  $R_c^{early}$ , controlling for Census division fixed effects. Sample size is 3075 counties. Robust standard errors in parentheses, clustered by state.

Table 4: **Economic Trends Before the 1970's Boom**

	(1)	(2)	(3)	(4)	(5)
Outcome:	Population	Employment	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
Endowment $R_c^{early}$	0.000449 (0.00110)	-0.00223 (0.00162)	0.000340 (0.000815)	-0.00189 (0.00219)	-0.00614 (0.00790)
Observations	3,075	3,075	3,075	2,483	2,484

Notes: This table presents estimates of Equation (15). The dependent variable is the 1972-1969 change in the natural log of the listed outcome. All regressions include census division indicators and controls for the natural log of the outcome variable in two baseline years. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 5: **Effects of Resource Booms on County-Level Outcomes**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.0119***	0.0282***	0.0270***	0.0179***	0.0108***	0.0293***
$\times \text{endowment}$	(0.00201)	(0.00607)	(0.00754)	(0.00510)	(0.00376)	(0.00724)
Observations	135,274	138,349	15,371	138,349	111,709	111,754

Notes: This table presents estimates of Equation (12). The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 6: **Effects on Manufacturing Subsectors and Alternative Outcomes**

	(1)	(2)	(3)	(4)	(5)
	All	Upstream / downstream	Non-linked	Non-linked and local	Non-linked and tradable
Employment					
$\Delta \ln(\text{National oil\&gas employment})$	0.0351**	0.0463**	.0181	0.0433***	-0.0327*
$\times \text{endowment}$	(0.0133)	(0.0176)	(0.0115)	(0.00584)	(0.0172)
Revenue					
$\Delta \ln(\text{National oil\&gas employment})$	0.0676***	0.0489*	0.0456***	0.0670*	-0.0621
$\times \text{endowment}$	(0.0178)	(0.0257)	(0.0133)	(0.0361)	(0.0509)
Investment					
$\Delta \ln(\text{National oil\&gas employment})$	0.118***	0.109**	0.0587*	0.0984***	-.00477
$\times \text{endowment}$	(0.0305)	(0.0438)	(0.0315)	(0.0241)	(0.0547)
Number of plants					
$\Delta \ln(\text{National oil\&gas employment})$	0.0212***	0.0307***	0.00741**	0.0123*	-0.00955**
$\times \text{endowment}$	(0.00233)	(0.00897)	(0.00300)	(0.00645)	(0.00371)
Employees per plant					
$\Delta \ln(\text{National oil\&gas employment})$	0.00906	0.0106	0.00580	0.0297***	-.0316**
$\times \text{endowment}$	(0.0110)	(0.00971)	(0.0135)	(0.00740)	(0.0154)
Observations	24,600	24,600	24,600	24,600	24,600

Notes: This table presents estimates of Equation (12) for different outcomes (in rows) and manufacturing subsectors (in columns), using county-by-Census year data. The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 7: **Effects of Resource Booms on Continuing Manufacturing Plants**

	(1)	(2)	(3)	(4)	(5)
	All	Upstream / downstream	Non-linked	Non-linked and local	Non-linked and tradable
Value added per worker					
$\Delta \ln(\text{National oil\&gas employment})$	0.0131***	0.00996**	0.0134***	0.0201***	0.00156
$\times \text{endowment}$	(0.00373)	(0.00429)	(0.00334)	(0.00479)	(0.00428)
Observations	1,140,000	388,000	752,000	379,000	372,000
Revenue productivity					
$\Delta \ln(\text{National oil\&gas employment})$	0.00240**	0.00226	0.00286***	0.00393**	-0.000565
$\times \text{endowment}$	(0.000951)	(0.00149)	(0.000967)	(0.00163)	(0.00129)
Observations	756,000	280,000	476,000	251,000	225,000
<b>(a) No Industry Controls</b>					
Value added per worker					
$\Delta \ln(\text{National oil\&gas employment})$	0.00544***	0.00512*	0.00489***	0.00832***	0.000680
$\times \text{endowment}$	(0.00182)	(0.00297)	(0.00152)	(0.00166)	(0.00350)
Observations	1,140,000	388,000	752,000	379,000	372,000
Revenue productivity					
$\Delta \ln(\text{National oil\&gas employment})$	0.00253**	0.00261**	0.00184	0.00377*	-0.000448
$\times \text{endowment}$	(0.00115)	(0.00121)	(0.00181)	(0.00197)	(0.00182)
Observations	756,000	280,000	476,000	251,000	225,000

**(b) With 4-digit Industry-by-Year Controls**

Notes: This table presents estimates of Equation (14), using plant-by-Census year data. The dependent variable is the change in the natural log of the outcome variable. All regressions include census division-by-year indicators. Sample sizes are rounded to the nearest 1,000 due to Census disclosure rules. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 8: Long-Term Revenue Productivity Effects of the Boom and Bust of the 1970s and 1980s

	(1)	(2)	(3)	(4)	(5)
	All	Upstream / downstream	Non-linked	Non-linked and local	Non-linked and tradable
County-level average log revenue productivity					
Endowment $R_c^{early}$	0.00145 (0.00681)	0.00667 (0.0103)	-0.0101 (0.00774)	-0.00470 (0.00597)	-0.00771 (0.00673)
Observations	2,769	2,033	2,710	2,525	2,260
Plant-level revenue productivity with 4-digit SIC-by-year controls					
1(year=1997)×endowment $R_c^{early}$	0.00132 (0.00249)	-0.00108 (0.00141)	0.00319 (0.00270)	0.000743 (0.00190)	0.00230 (0.00274)
Observations	338,000	116,000	222,000	110,000	112,000

Notes: The top panel presents estimates of Equation (15), using 1997-1972 changes in within-county mean of log revenue productivity as the dependent variable. The bottom panel presents regressions using 1997 and 1972 data of plant-level log TFP-R on an interaction of a 1997 indicator variable with endowment  $R_c^{early}$ , controlling for county fixed effects and 4-digit industry-by-year interactions. All regressions include census division-by-year indicators. In the plant-level regressions, sample sizes are rounded to the nearest 1,000 due to Census disclosure rules. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 9: Long-Term County-Level Effects of the Boom and Bust of the 1970s and 1980s

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
Endowment $R_c^{early}$	-0.0000928 (0.00475)	-0.00117 (0.00572)	-0.00666 (0.00407)	0.000545 (0.00178)	-0.00980 (0.00660)	-0.00271 (0.0151)
Observations	3,075	3,075	3,075	3,075	2,483	2,484

Notes: This table presents estimates of Equation (15), using 1997-1972 changes in log outcomes as the dependent variable. (Because housing rents are not available in 1997 or 1972, the housing rent regression uses the 2000-1970 change.) All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table 10: Absolute Effects of Resource Booms on County-Level Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.00692***	0.0162***	0.0120*	0.0102***	0.00176	0.0166**
$\times \text{endowment}$	(0.00155)	(0.00579)	(0.00702)	(0.00304)	(0.00212)	(0.00626)
Observations	135,274	138,349	15,371	138,349	111,709	111,754

Notes: This table presents estimates of Equation (13). The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

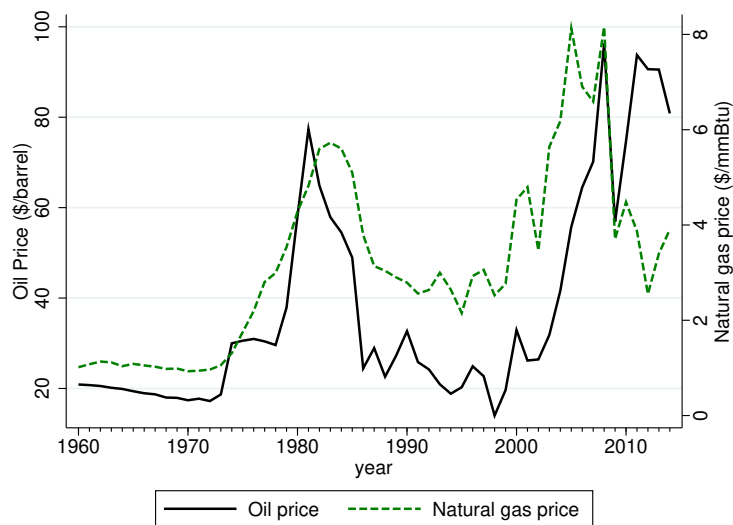
Table 11: 1969-2014 Welfare Gains from Resource Booms

Row	Parameter	(1) Relative effects	(2) Absolute effects
1	$\hat{\tau}_{\text{earnings}}$	0.024 (0.007)	0.011 (0.003)
2	$\hat{\tau}_{\text{rent}}$	0.027 (0.008)	0.012 (0.007)
3	$\hat{\tau}_{\text{earnings}} - 0.3 \cdot \hat{\tau}_{\text{rent}}$	0.016 (0.005)	0.007 (0.001)
4	Annual average real earnings gain (log points)	0.010 (0.003)	0.004 (0.001)
5	$\hat{\tau}_{\text{population}}$	0.012 (0.002)	0.008 (0.002)
6	Annual average population gain (log points)	0.008 (0.001)	0.005 (0.001)

Notes: Rows 1, 2, and 5 present estimates of Equation (12) and (13) in columns 1 and 2, respectively, using data from 1970, 1980, 1990, 2000, 2009, and 2014. Robust standard errors in parentheses, clustered by state. Rows 3, 4, and 6 present additional calculations, with standard errors from the Delta method.

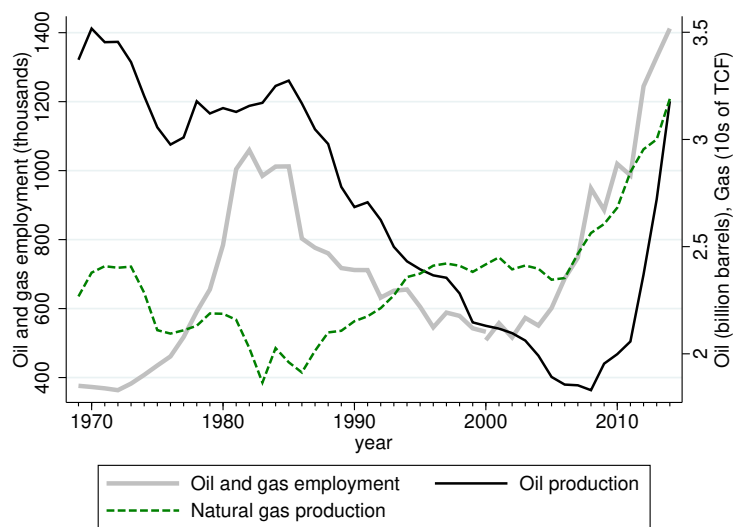
# Figures

Figure 1: Real Oil and Gas Prices



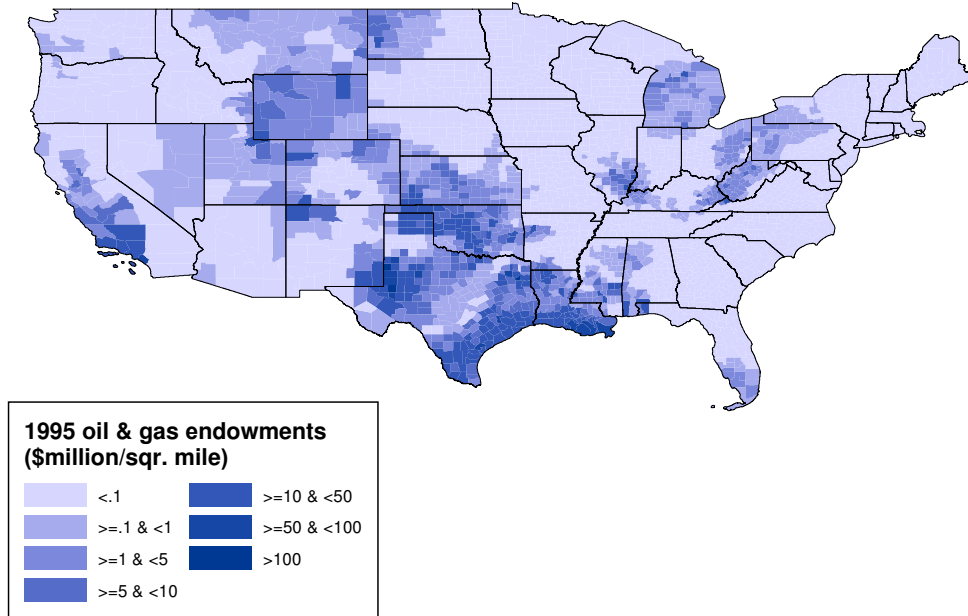
Notes: Oil and gas price data are from the Energy Information Administration. Prices are in real 2010 dollars.

Figure 2: National Oil and Gas Employment and Production



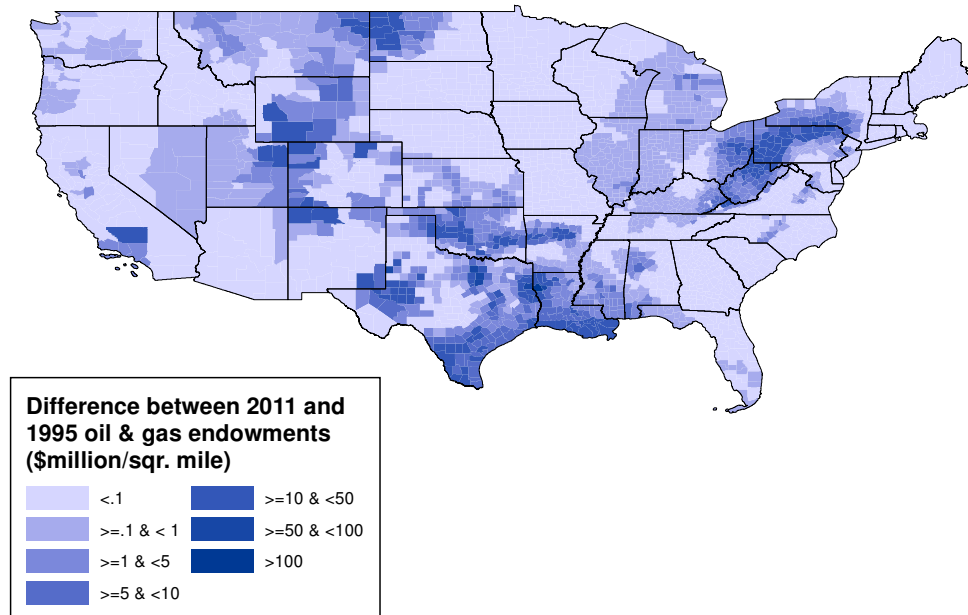
Notes: Oil and gas production data are from the Energy Information Administration. Oil and gas employment data are from the Regional Economic Information System. We switch from the SIC to the NAICS classification system in 2000 and plot both data points in that year.

Figure 3: Early Endowment per Square Mile



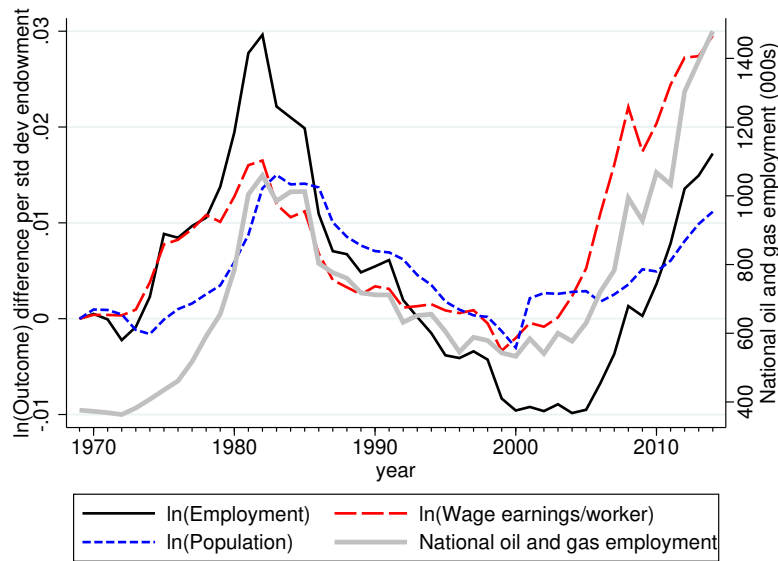
Notes: This figure maps each county's oil and gas endowment as of 1960 that is economically recoverable during the oil and gas boom of the 1970s and 1980s. See Section 4.1 for details of variable construction.

Figure 4: Change in Endowment After Early Period



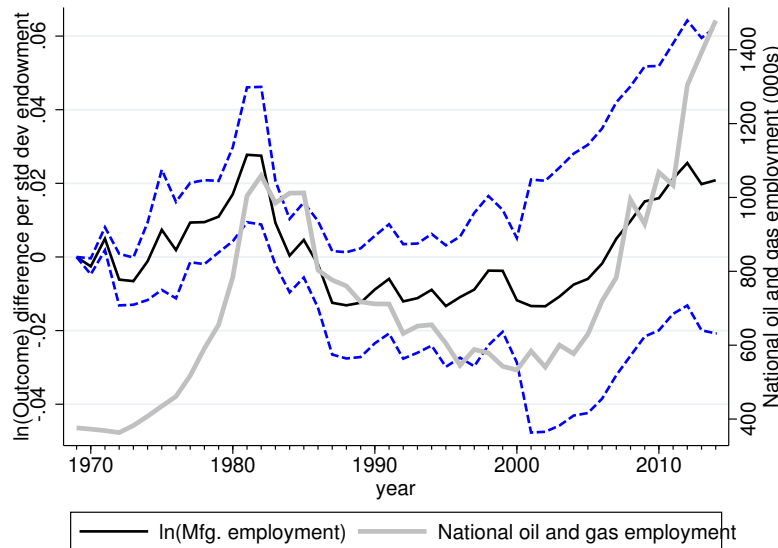
Notes: This figure maps each county's difference between the early period oil and gas endowment graphed in figure 3 and the total endowment that is economically recoverable using current technologies. It largely reflects "unconventional" endowments that are economically recoverable only with hydraulic fracturing. See Section 4.1 for details of variable construction.

Figure 5: County-Level Aggregates Over Time in Resource-Abundant Counties



Notes: This figure presents the  $\tau_t^r$  regression coefficients from Equation (11) with natural log of county employment, population, and earnings per worker as dependent variables. 1969 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 1969 associations between resources and outcomes.

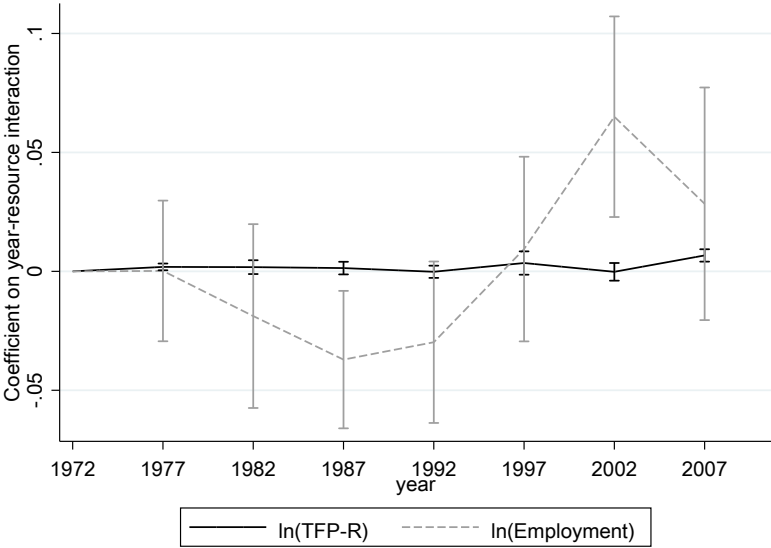
Figure 6: Manufacturing Employment Over Time in Resource-Abundant Counties



Notes: This figure presents the  $\tau_t^r$  coefficients and 90 percent confidence intervals from Equation (11) with natural log of county manufacturing employment as the dependent variable. 1969 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 1969 association between resources and manufacturing employment.

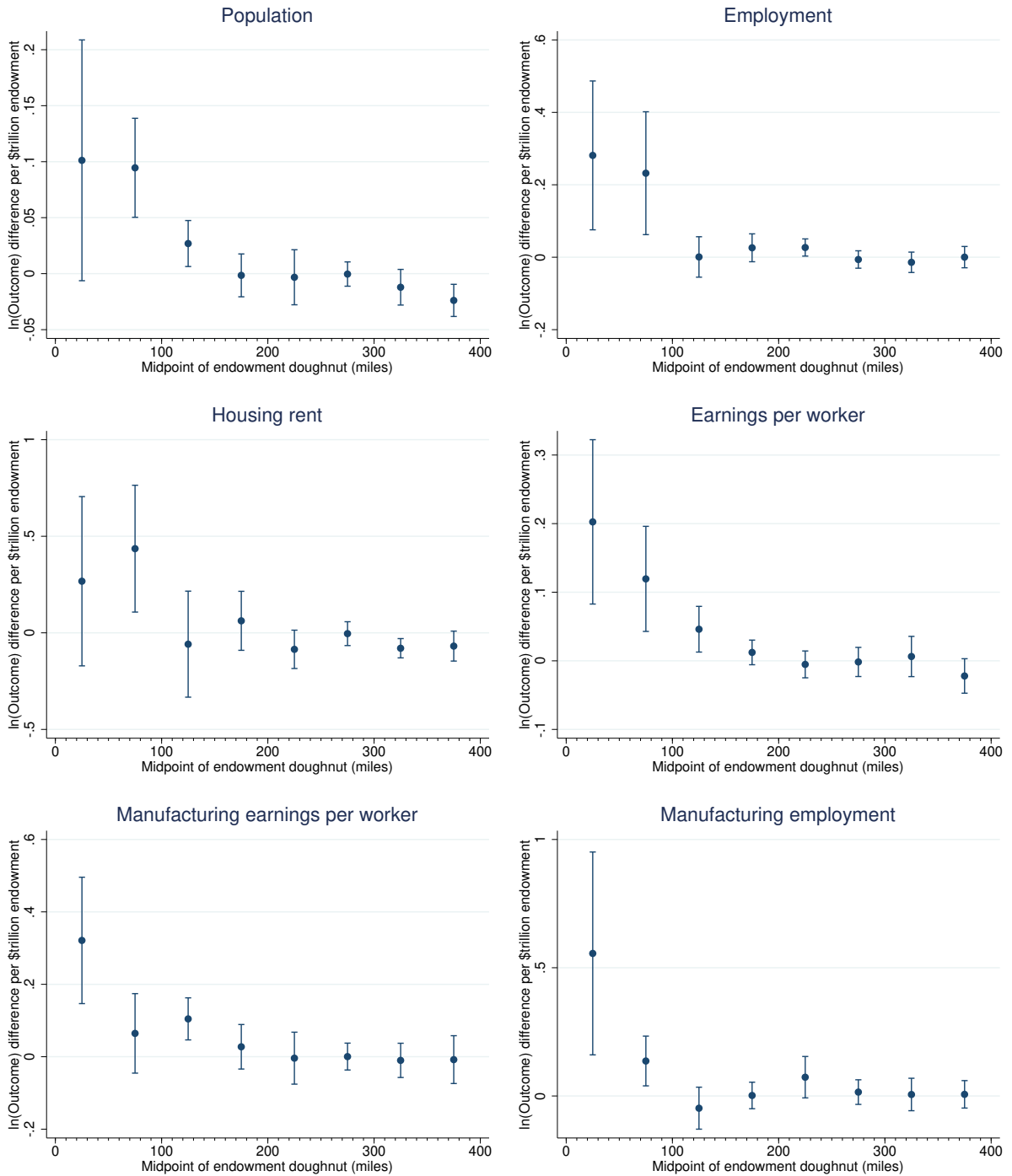


Figure 7: Employment and Productivity Over Time for Tradable Manufacturing Subsectors



Notes: This figure presents the  $\tau_t^r$  coefficients and 90 percent confidence intervals from Equation (11) with the natural log of county mean  $\ln(\text{TFP-R})$  and county manufacturing employment, using data from only the highly tradable manufacturing subsectors, as the dependent variables. 1972 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 1972 associations between resources and outcomes.

Figure 8: Geographic Spillovers



Notes: This figure presents the  $\nu_d$  coefficients and 90 percent confidence intervals from estimating Equation (13) with various dependent variables.

**Appendix: For Online Publication**

*Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America*

Hunt Allcott and Daniel Keniston

## A Model Appendix

### A.1 Solving for Equilibrium

In equilibrium, firms and consumers optimize, and supply equals demand for labor, housing, and the local good.

#### A.1.1 Equilibrium in Housing and Local Good Markets

Subtracting the housing supply function in county  $b$  from county  $a$  gives the relative price of housing:

$$p_{ha} - p_{hb} = k(n_a - n_b). \quad (17)$$

The equilibrium price of the local good is determined by two conditions. First, local good supply equals demand:

$$\Omega_{lc} N_{lc}^{1-\gamma} = N_c \frac{\alpha W_c}{P_{lc}}, \quad (18)$$

where demand is from maximizing the Cobb-Douglas utility function.

Second, the composite firm for any sector  $j$  maximizes profits, giving first-order condition

$$(1 - \gamma) P_{jc} \Omega_{jc} N_{jc}^{-\gamma} = W_c. \quad (19)$$

Combining these two equations gives the local good price:

$$P_{lc} = \frac{\alpha^\gamma W_c N_c^\gamma (1 - \gamma)^{\gamma-1}}{\Omega_{lc}}.$$

Taking logs gives

$$p_{lc} = w_c + \gamma n_c + \gamma \ln \alpha + (\gamma - 1) \ln(1 - \gamma) - \omega_{lc}. \quad (20)$$

Taking the difference between counties  $a$  and  $b$  gives the relative price of the local good:

$$p_{la} - p_{lb} = (w_a - w_b) + \gamma(n_a - n_b) - (\omega_{la} - \omega_{lb}). \quad (21)$$

#### A.1.2 Labor Market Equilibrium

County aggregate labor demand is the sum across sectors  $j \in \{m, l, r\}$ :

$$N_c = \sum_j N_{jc} = \left[ \frac{X_{jc}(1-\gamma)}{W_c} \right]^{\frac{1}{\gamma}} \quad (22)$$

$$= \frac{(1-\gamma)^{1/\gamma}}{W_c^{1/\gamma}} \left[ X_{mc}^{1/\gamma} + X_{rc}^{1/\gamma} + \Omega_{lc}^{1/\gamma} \left( \frac{\alpha^\gamma W_c N_c^\gamma (1-\gamma)^{\gamma-1}}{\Omega_{lc}} \right)^{1/\gamma} \right] \quad (23)$$

$$= \frac{(1-\gamma)^{1/\gamma}}{1-\alpha(1-\gamma)} W_c^{\frac{-1}{\gamma}} \left[ X_{mc}^{1/\gamma} + X_{rc}^{1/\gamma} \right]. \quad (24)$$

Taking logs and re-arranging gives inverse labor demand in county  $c$ :

$$w_c = \gamma \tilde{x}_c + \lambda \gamma - \gamma n_c, \quad (25)$$

where  $\lambda \equiv \ln \left[ \frac{(1-\gamma)^{1/\gamma}}{1-\alpha(1-\gamma)} \right]$ . Subtracting  $w_a - w_b$  gives relative inverse labor demand:

$$w_a - w_b = \gamma (\tilde{x}_a - \tilde{x}_b) - \gamma (n_a - n_b), \quad (26)$$

where  $\tilde{x}_c \equiv \ln \left( X_{mc}^{1/\gamma} + X_{rc}^{1/\gamma} \right)$  measures traded and resource sector productivity. This equation shows how higher productivity raises relative wages, but this is dampened by migration.

To solve for the population difference, substitute housing and local good prices into labor supply, giving

$$w_a - w_b = \alpha ((w_a - w_b) + \gamma (n_a - n_b) - (\omega_{la} - \omega_{lb})) + \beta k (n_a - n_b) + s(n_a - n_b) - (a_a - a_b) \quad (27)$$

$$= \frac{\alpha \gamma + \beta k + s}{1-\alpha} (n_a - n_b) - \frac{\alpha}{1-\alpha} (\omega_{la} - \omega_{lb}) - \frac{\alpha}{1-\alpha} (a_a - a_b) \quad (28)$$

Equating this to relative inverse labor demand from Equation (26) and simplifying gives relative population: shown in Equation (29):

$$n_a - n_b = \frac{1}{\gamma + \beta k + s} [\gamma(1-\alpha) (\tilde{x}_a - \tilde{x}_b) + \alpha (\omega_{la} - \omega_{lb}) + (a_a - a_b)], \quad (29)$$

where  $\omega_{lc}$  is natural log local sector physical productivity. This shows that people will migrate to the county with higher productivity and amenities. There will be more migration if production is more labor intensive (i.e.  $\gamma$  is smaller), if housing supply is more elastic (i.e.  $k$  is lower), if the housing expenditure share  $\beta$  is lower, or if locational preferences are weaker (i.e.  $s$  is smaller).

The equilibrium wage difference is derived by substituting the population difference into the inverse labor demand difference:

$$w_a - w_b = \frac{\gamma}{\gamma + \beta k + s} [(\beta k + s + \gamma \alpha) (\tilde{x}_a - \tilde{x}_b) - \alpha (\omega_{la} - \omega_{lb}) - (a_a - a_b)]. \quad (30)$$

This shows that higher resource or tradable sector productivity increases nominal wages. Higher amenities encourage people to accept lower wages, as does higher local sector productivity  $\omega_{la}$ , as this reduces the local good price.

### A.1.3 Absolute Population and Wages

We can also close the model to solve for absolute population and wages as a function of productivity and primitives. To do this, fix total population in the two counties at  $N_a + N_b = N$ . Rearranging and taking logs gives  $n_b = \ln(N - N_a)$ . Substituting this into the population difference from Equation (29) and simplifying gives population:

$$n_a = \xi - \ln(1 + e^\xi) + n, \quad (31)$$

where  $\Gamma_c \equiv \gamma(1 - \alpha)\tilde{x}_c + \alpha\omega_{lc} + a_c$  and  $\xi \equiv \frac{\Gamma_a - \Gamma_b}{\gamma + \beta k + s}$ . Substituting this into Equation (25) gives wages:

$$w_a = \gamma \cdot \left[ \tilde{x}_a + \lambda - \xi + \ln(1 + e^\xi) - n \right]. \quad (32)$$

Equations (31) and (32) are symmetric for county  $b$ .

## A.2 Predictions: Derivations and Additional Details

### A.2.1 Contemporaneous Effects on County Aggregates

Prediction 1's relative population and wage increases are immediate from Equations (29) and (30). In period  $t = 1$ ,  $\omega_{la} = \omega_{lb}$ ,  $\omega_{ma} = \omega_{mb}$ , amenities are unchanged by assumption, and  $(\tilde{x}_a - \tilde{x}_b) > 0$  because  $\omega_{ra} - \omega_{rb} > 0$ . Thus,  $n_a - n_b = \frac{1}{\gamma + \beta k + s} [\gamma(1 - \alpha)(\tilde{x}_a - \tilde{x}_b)] > 0$  and  $w_a - w_b = \frac{\gamma}{\gamma + \beta k + s} [(\beta k + s + \gamma\alpha)(\tilde{x}_a - \tilde{x}_b)] > 0$  in  $t = 1$ , both of which are positive because  $(\tilde{x}_a - \tilde{x}_b)$ ,  $\alpha$ ,  $(1 - \alpha)$ ,  $\beta$ ,  $\gamma$ , and  $s$  are all positive. Migration is an outward shift in labor supply, so it reduces equilibrium wages.

Prediction 1 also holds in absolute terms. The absolute effect on population can be derived from Equation (31):

$$n_a - n_a(0) = \xi_t - \ln(1 + e^{\xi_t}) + \ln 2 > 0.$$

It is also true that  $n_a - n_a(0) < \tilde{x}_a - \tilde{x}_a(0)$ . Using this and (25), we have  $w_a - w_a(0) > 0$ : the absolute wage effect is also positive.

### A.2.2 Contemporaneous Effects on Manufacturing

For any sector  $j$ , differencing the first-order conditions gives relative sectoral labor demand:

$$n_{ja} - n_{jb} = \frac{1}{\gamma} (\omega_{ja} - \omega_{jb}) + \frac{1}{\gamma} (p_{ja} - p_{jb}) - \frac{1}{\gamma} (w_a - w_b). \quad (33)$$

Substituting local sector relative price from Equation (21) into Equation (33) and simplifying gives

$$n_{la} - n_{lb} = n_a - n_b. \quad (34)$$

Using Prediction 1, this is positive in  $t = 1$ . Of course, this equality is a feature of our Cobb-Douglas functional forms, but the result that non-tradable sector employment grows with the local economy is more general.

For the tradable sector, substituting the wage difference from Equation (30) into Equation (33) and assuming that  $\omega_{la} = \omega_{lb}$  and  $a_a = a_b$  gives relative employment:

$$n_{ma} - n_{mb} = -\frac{\beta k + s + \alpha \gamma}{\beta k + s + \gamma} (\tilde{x}_a - \tilde{x}_b). \quad (35)$$

Because  $(\tilde{x}_a - \tilde{x}_b) > 0$  in  $t = 1$ ,  $(n_{ma} - n_{mb}) < 0$  in  $t = 1$ . This proves Prediction 2 in relative terms.

Prediction 2 also holds in absolute terms. Logging the first-order condition from Equation (19), we have

$$n_{jc} = \frac{1}{\gamma} [\omega_{jc} + p_{jc} + \ln(1 - \gamma) - w_c] > 0. \quad (36)$$

For the local sector, the absolute price difference in  $t = 1$  is  $p_{la} - p_{la}(0) = (w_a - w_a(0)) + \gamma(n_a - n_a(0))$ . This means that the local good price rises more than the wage. Substituting this into Equation (36) and taking absolute effects, we have  $n_{la} - n_{la}(0) = n_a - n_a(0) > 0$ . For the tradable sector, substituting  $w_a - w_a(0) > 0$  into Equation (36) immediately gives  $n_{ma} - n_{ma}(0) < 0$  in  $t = 1$ .

### A.2.3 Long-Term Effects

Substituting employment differences from Equations (34) and (35) into the productivity evolution from Equation (1) gives closed-form solutions for local and tradable sector productivity in period  $t = 2$ . Assuming no amenity change, these are

$$\omega_{la} - \omega_{lb} = (\theta_l + \Lambda) \frac{\gamma(1 - \alpha)}{\beta k + s + \gamma} (\tilde{x}_{a,t=1} - \tilde{x}_{b,t=1}) \quad (37)$$

and

$$\omega_{ma} - \omega_{mb} = \frac{\Lambda\gamma(1 - \alpha) - \theta_m(\beta k + s + \alpha\gamma)}{\beta k + s + \gamma} (\tilde{x}_{a,t=1} - \tilde{x}_{b,t=1}). \quad (38)$$

Since  $(\tilde{x}_{a,t=1} - \tilde{x}_{b,t=1}) > 0$  and  $\{\theta_l, \Lambda\} \geq 0$  by assumption, Equation (37) shows that  $\omega_{la} - \omega_{lb} > 0$ . Equation (37) simplifies directly to Equation (8). This proves Prediction 3.

With the exception of Equation (8), Prediction 3 also holds in absolute terms. Substituting the signs of the absolute effects on total population from Prediction 1 and local sector employment from Prediction 2 into the productivity evolution from Equation (1) immediately gives  $\omega_{la} - \omega_{la}(0) > 0$ . For the tradable sector, we have

$$\omega_{ma,t=2} - \omega_{ma,t=1}(0) = \theta_m(n_{ma,t=1} - n_{ma,t=1}(0)) + \Lambda(n_{a,t=1} - n_{a,t=1}(0)). \quad (39)$$

Since  $n_{ma,t=1} - n_{ma,t=1}(0) < 0$  and  $n_{a,t=1} - n_{a,t=1}(0) > 0$  from Predictions 2 and 1, respectively, we again have that the  $t = 2$  productivity impact trades off agglomeration spillovers with foregone learning-by-doing.

#### A.2.4 Overall Welfare Effects

Substituting labor supply from Equation (3) into the relative welfare effect from Equation (5) shows that relative welfare can be measured by migration:<sup>23</sup>

$$u_a - u_b = \sum_t \delta^t s(n_{at} - n_{bt}). \quad (40)$$

Substituting the population difference from Equation (29), assuming  $X_{rc} = 0$  always except  $X_{ra} > 0$  in  $t = 1$ , and simplifying gives

$$u_a - u_b = \frac{s\delta(1 - \alpha)}{\beta k + s + \gamma} (\tilde{x}_{a,t=1} - \tilde{x}_{b,t=1}) \left[ \gamma + \frac{\delta}{\beta k + s + \gamma} (\gamma\Lambda + \alpha\gamma\theta_l - \theta_m(\beta k + s + \alpha\gamma)) \right]. \quad (41)$$

This is positive if and only if the final term (in brackets) is positive, which is true if Equation (9) holds.

The absolute welfare effects depend on the absolute wage and price levels. Substituting the price levels into the absolute welfare effect from Equation (7) gives

---

<sup>23</sup>Note that this equation does not have an analogue for absolute effects, i.e. it is not true that  $u_a - u_a(0) = \sum_t \delta^t s(n_{at} - n_{at}(0))$ . Intuitively, the reason is that county  $a$  can have high population, and thus high population relative to counterfactual, even if social welfare is worse than counterfactual, as long as social welfare is even worse in county  $b$ .



$$u_a - u_a(0) = \sum_t [(1 - \alpha)\gamma(\tilde{x}_{at} - \tilde{x}_{at}(0)) + \alpha(\omega_{lat} - \omega_{lat}(0)) - (\gamma + \beta k)(n_{at} - n_{at}(0))]. \quad (42)$$

This equation states that absolute social welfare effects depend on two factors: the county productivity increases, which determine wages and local prices, net of the population inflows, which reduce wages by  $\gamma$  and increase the utility impact of housing costs by  $\beta k$ . Substituting Equation (31) into Equation (42) gives

$$u_a - u_a(0) = \sum_t \delta^t \left[ (\Gamma_{at} - \Gamma_{at}(0)) - (\gamma + \beta k) \left( \xi_t - \ln(1 + e^{\xi_t}) + \ln 2 \right) \right]. \quad (43)$$

### A.2.5 Special Case: Absolute Welfare with No Productivity Spillovers

We can also consider the special case of absolute welfare with no productivity spillovers, i.e. Equation (43) with  $\Lambda = \theta_l = \theta_m = 0$ . In that case, absolute and relative welfare effects are both zero in period  $t = 2$ , as county  $a$  productivity has returned to equal county  $b$  productivity. Furthermore,  $\Gamma_{bt} = \Gamma_{at}(0)$  for  $t = 1$ . Substituting this into Equation (43) gives

$$u_a - u_a(0) = \delta \left[ s(n_{a,t=1} - n_{b,t=1}) + (\gamma + \beta k) \ln \left( \frac{1 + \exp(n_{a,t=1} - n_{b,t=1})}{2} \right) \right]. \quad (44)$$

This equation is useful because it shows that the absolute welfare effect is unambiguously positive: we know from Prediction 1 that  $(n_{a,t=1} - n_{b,t=1}) > 0$ , and substituting this into Equation (44) implies  $u_a - u_a(0) > 0$ . Intuitively, this just says that if there are no market failures, an increase in county  $a$ 's productivity increases welfare.

For comparison, the relative welfare effect when  $\Lambda = \theta_l = \theta_m = 0$  is

$$u_a - u_b = \delta s(n_{a,t=1} - n_{b,t=1}). \quad (45)$$

Comparing Equations (44) and (45), we see that the absolute welfare effect equals the relative effect plus adjustment factor  $(\gamma + \beta k) \ln \left( \frac{1 + \exp(n_{a,t=1} - n_{b,t=1})}{2} \right)$ . This factor accounts for the fact that the resource boom also increases wages in county  $b$  in  $t = 1$ , so the relative welfare gain is smaller than the absolute welfare gain.

## A.3 Relative and Absolute Effects with Many Counties

While the theoretical model has two counties, the empirical analysis has more than 3000. A brief discussion using the Rubin Causal Model can help clarify the distinction between relative and absolute effects in this more general context. Let  $D_c \in \{1, 0\}$  denote county  $c$ 's treatment

assignment (i.e. whether or not it is endowed with natural resources), and let  $\mathbf{D}_{-c}$  denote the vector of treatment assignments for all counties other than  $c$ . Let  $Y_c$  denote an outcome for county  $c$ , such as wages or employment in some time period, and let  $Y_c(d_c, \mathbf{d}_{-c})$  denote county  $c$ 's potential outcome as a function of any possible treatment assignments  $d_c$  and  $\mathbf{d}_{-c}$ . As in Manski (2013), potential outcomes depend on treatment status both in county  $c$  and in all other counties, due to geographic spillovers.

In the two-county model, the relative effect is the difference between effect of the boom on county  $a$  and the effect of the boom on county  $b$ . In the many-county world, the relative effect  $\tau^r$  is the expectation over all counties of the effects of the boom on treated counties and the effects on untreated counties:

$$\tau^r \equiv \mathbb{E}[Y_c(1, \mathbf{D}_{-c}) - Y_c(0, \mathbf{0}) \mid D_c = 1] - \mathbb{E}[Y_c(0, \mathbf{D}_{-c}) - Y_c(0, \mathbf{0}) \mid D_c = 0]. \quad (46)$$

If the treatment is randomly assigned, then  $\mathbb{E}[Y_c(0, \mathbf{0}) \mid D_c = 1] = \mathbb{E}[Y_c(0, \mathbf{0}) \mid D_c = 0]$ , so this simplifies to

$$\tau^r \equiv \mathbb{E}[Y_c(1, \mathbf{D}_{-c}) \mid D_c = 1] - \mathbb{E}[Y_c(0, \mathbf{D}_{-c}) \mid D_c = 0]. \quad (47)$$

The relative effects estimating equation (Equation (12)) estimates a continuous-treatment analogue to  $\tau^r$  by comparing outcomes in more- vs. less-treated counties.

In the two-county model, the absolute effect is the difference between county  $a$ 's potential outcomes if it does vs. does not have natural resources. In the many-county world, the absolute effect  $\tau^a$  is the expectation over all counties of the difference in potential outcomes with vs. without resources:

$$\tau^a \equiv \mathbb{E}[Y_c(1, \mathbf{D}_{-c}) - Y_c(0, \mathbf{D}_{-c}) \mid D_c = 1]. \quad (48)$$

In other words, this is the average treatment effect on the treated. The absolute effects estimating equation (Equation (13)) estimates a continuous-treatment analogue to  $\tau^a$  by comparing outcomes in more- vs. less-treated counties while controlling for spillovers from other counties.

Both  $\tau^r$  and  $\tau^a$  are causal parameters in the sense that they represent effects of natural resource endowments; both equal zero if natural resources do not affect outcomes. Furthermore, both  $\tau^r$  and  $\tau^a$  compare observed outcomes of treated counties  $\mathbb{E}[Y_c(1, \mathbf{D}_{-c}) \mid D_c = 1]$  to some other quantity. The difference is that  $\tau^r$  compares this to actual outcomes in untreated counties  $\mathbb{E}[Y_c(0, \mathbf{D}_{-c}) \mid D_c = 0]$ , whereas  $\tau^a$  compares this to counterfactual outcomes in treated counties  $\mathbb{E}[Y_c(0, \mathbf{D}_{-c}) \mid D_c = 1]$ .

Both parameters also are conditional on the actual treatment status of other counties. One might also be interested in treatment effects conditional on some  $\mathbf{d}_{-c}$  other than actual treatments  $\mathbf{D}_{-c}$  – for example, if fracking were banned throughout the eastern U.S. Our empirical analysis does

not consider those cases because they would require significant additional empirical assumptions.

## **B Data Appendix**

### **B.1 Oil and Gas Endowments**

Table A1: State-Level Sources of Oil and Gas Production Data

State	Resource	Title	Source	Years
CA	Oil, Gas	Summary of Operations: California Oil Fields	<a href="ftp://ftp.consrv.ca.gov/pub/oil/Summary_of_Operations/">ftp://ftp.consrv.ca.gov/pub/oil/Summary_of_Operations/</a>	1960-1977
IL	Gas	Natural Gas Production in Illinois	Bryan Huff, Illinois State Geological Survey	1973-1992
IL	Oil	Historic County Production in Illinois	Bryan Huff, Illinois State Geological Survey	1932-2011
IN	Gas	Petroleum Data Management System	<a href="http://igs.indiana.edu/PDMS/WellSearch.cfm">http://igs.indiana.edu/PDMS/WellSearch.cfm</a>	1863-2011
IN	Oil	Petroleum Data Management System	<a href="http://igs.indiana.edu/PDMS/Fields.cfm">http://igs.indiana.edu/PDMS/Fields.cfm</a>	1965-2011
KS	Oil, Gas	County Production	<a href="http://www.kgs.ku.edu/PRS/petro/interactive.html">http://www.kgs.ku.edu/PRS/petro/interactive.html</a>	1960-2011
KY	Oil	Oil and Gas Production	<a href="http://kgs.uky.edu/kgsmap/OGProdPlot/OGProduction.asp">http://kgs.uky.edu/kgsmap/OGProdPlot/OGProduction.asp</a>	1883-2011
KY	Gas	Oil and Gas Production	<a href="http://kgs.uky.edu/kgsmap/OGProdPlot/OGProduction.asp">http://kgs.uky.edu/kgsmap/OGProdPlot/OGProduction.asp</a>	1986-2011
LA	Oil, Gas	Crude and Natural Gas Production by Parish	Sharron Allement, Louisiana Office of Conservation	1965-1977
MI	Oil, Gas	Michigan's Oil and Gas Fields, 1965-1982	<a href="http://www.michigan.gov/deq">http://www.michigan.gov/deq</a>	1965-1982
MT	Oil, Gas	Annual Reviews for the Years 1965-1985	<a href="http://bogc.dnrc.mt.gov/annualreview/">http://bogc.dnrc.mt.gov/annualreview/</a>	1965-1985
NM	Oil, Gas	Annual Report	New Mexico Oil & Gas Engineering Committee	1960-1985
NV	Oil, Gas	Historical Production	Lowell Taylor, Nevada Division of Minerals	1954-2011
NY	Oil, Gas	New York Natural Gas and Oil Production	<a href="http://www.dec.ny.gov/energy/1601.html">http://www.dec.ny.gov/energy/1601.html</a>	1967-2011
OK	Oil	Report on Oil and Natural Gas Activity	Jason Lawter, Oklahoma Corporation Commission	1963-2011
PA	Oil	Oil and Gas Developments in Pennsylvania	<a href="http://www.libraries.psu.edu/">http://www.libraries.psu.edu/</a>	1960-1991
UT	Oil, Gas	Pre-1984 Production Download File	<a href="http://oilgas.ogm.utah.gov/">http://oilgas.ogm.utah.gov/</a>	1965-1983

Notes: This details additional state-level sources of oil and gas production data that are used to augment the DrillingInfo database.

## B.2 Regional Economic Information System

The REIS gathers two measures of employment and earnings. Series 7 and 27 measure wage and salary earnings and employment, based primarily on unemployment insurance payment records. This corresponds closely to data from the Quarterly Census of Employment and Wages. Series 5 and 25 measures total earnings and employment, adding sole proprietors (who file a Schedule C on their tax returns) and general partners in partnerships (who file a Form 1065). Neither series includes limited partners, who are likely to be only passive investors in oil and gas businesses. Series 5 and 25 are intended to be a more comprehensive measure of total earnings and employment, so we use these series in all specifications. Substituting Series 7 and 27 does not meaningfully change the estimates.

In general, these measures appear to correctly allocate economic activity to counties. Unemployment insurance payments are assigned to the county where the employing *establishment* is located, not the firm headquarters. Sole proprietor and partnership earnings and employment are assigned to the tax-filing address of the recipient, which is typically the person’s residence. This will misallocate employment when the filer is not working in his or her county of residence.

For population, the REIS uses annual mid-year estimates constructed by the U.S. Census Bureau. In recent years, the Census Bureau estimates county  $c$ ’s population as

$$population_{ct} = population_{t-1} + births_{ct} - deaths_{ct} + net\ migration_{ct},$$

where the population base ( $population_{t-1}$ ) comes initially from the decennial census, and births and deaths are measured from vital statistics data provided by the National Center for Health Statistics. Net domestic migration is primarily measured using Internal Revenue Service income tax data: individuals are matched across years using their social security numbers, and they are counted as migrants when their tax address changes. For the elderly, tax data are augmented with Medicare enrollment data. Net international migration is measured using the American Community Survey. While the ACS is only a sample survey, international migration is a relatively small share of a county’s total migration.<sup>24</sup>

Earlier in our sample, the Census Bureau used county-level data on school enrollment and number of occupied housing units, along with county-level population estimates from state agencies, to arrive at each year’s total population count. Thus, throughout the sample, the Census Bureau used actual data to count migration, and did not simply interpolate between decennial census counts.<sup>25</sup>

County-level land area are from [http://quickfacts.census.gov/qfd/download\\_data.html](http://quickfacts.census.gov/qfd/download_data.html).

There are 3075 counties in the sample through 2001. In 2002, Broomfield County, Colorado,

---

<sup>24</sup>See <http://www.census.gov/popest/methodology/2015-natstcopr-meth.pdf> for full methodology for recent mid-year population estimates.

<sup>25</sup>See <http://hdl.handle.net/2027/coo.31924000774731> for full methodology for the 1966 population estimates.

was created by breaking off pieces of other counties, primarily Adams and Boulder Counties. We never include Broomfield County in the sample, and we include Adams and Boulder Counties in Colorado only in 2001 and before. Thus, in 2002 and later, there are 3073 counties in the sample.

### B.3 Industry Linkages

Table A2: **Linked Manufacturing Industries**

<b>Top Ten Upstream Industries</b>		
SIC Codes	Industry	Upstream Linkage Share
3533	Oil and gas field machinery and equipment	0.23
324	Hydraulic cement	0.12
3295	Ground or treated minerals	0.086
2899	Chemicals and chemical preparations, n.e.c.	0.066
3491, 3492, 3494, 3498	Pipe, valves, and pipe fittings	0.037
3441	Fabricated structural metal	0.034
3312	Blast furnaces and steel mills	0.033
2892	Explosives	0.031
2992	Lubricating oils and greases	0.031
3313	Electrometallurgical products, except steel	0.028
<b>All Downstream Industries</b>		
SIC Codes	Industry	Input Cost Share
291	Petroleum refining	0.69
2999	Products of petroleum and coal, n.e.c.	0.31
2873, 2874	Nitrogenous and phosphatic fertilizers	0.081
2895	Carbon black	0.062
281, 2865, 2869	Industrial inorganic and organic chemicals	0.021
308	Miscellaneous plastics products, n.e.c.	0.001
285	Paints and allied products	0.001

Notes: Linkages are calculated using data from the 1987 Bureau of Economic Analysis input-output tables (U.S. BEA 2014). Upstream linkage share is the sum of oil and gas output share and the share of output purchased by the oil and gas sector through an intermediate industry. “n.e.c.” stands for “not elsewhere classified.”

## C Current Population Survey Data and Regressions

### C.1 Current Population Survey Data

In order to directly measure wages instead of earnings per worker, we construct two datasets from the Current Population Survey (CPS). The first is a repeated cross section formed by combining all observations from the May CPS for 1977 and 1978 with all observations from the Merged Outgoing Rotation Group (MORG) files beginning in 1979. Table A3 describes these data for hourly wage and hours worked. The second dataset is a panel based on the MORG, which includes each individual’s change in hourly earnings in the 12 months between two “outgoing rotation” interviews.<sup>26</sup> We include only workers employed by the private sector or government and exclude self-employed and unemployed.

Except for large counties in recent years, the CPS does not include county identifiers.<sup>27</sup> In regressions with CPS data, we thus use state-level oil and gas endowment  $R_{st}$ , which is constructed analogously to the county-level measure described earlier and normalized to a standard deviation of one across states. For context, one standard deviation of state-level total endowments is \$4.5 million per square mile.

Table A3: Current Population Survey: Descriptive Statistics

	N	Mean	SD	Min	Max
Hourly wage (\$/hour)	5,970,555	18.6	14.0	0.0001	2826
Hours worked per week	6,001,940	38.3	10.8	0	99

Notes: Data are at the individual-by-interview level. Wages are in real 2010 dollars.

### C.2 CPS Empirical Strategy

We estimate relative effects using two parallel specifications in the two different datasets, using  $i$  to index individuals. For the CPS repeated cross section, we regress natural log of an outcome variable on the resource boom variable, using controls  $X_i$  for individual  $i$ ’s age, education, gender, race, and industry. We index states by  $s$ , months by  $m$ , and years by  $t$ , and include vectors of state

<sup>26</sup>Since the CPS sampling frame is the household, not the individual, this panel includes only individuals who do not change residence between the two interviews. The CPS does not include unique individual identifying codes, so we use the approach of Madrian and Lefgren (1999) to match individuals between interviews.

<sup>27</sup>It is not possible to construct a panel over our entire study period using any consistently-defined geographical area less aggregated than the state. Before 1977, the May CPS data do not include a complete set of state identifiers. From 1977 to June 1985, there are geographic identifiers for state and approximately 45 Standard Metropolitan Statistical Areas (MSAs). From July-December 1985, only state identifiers are included. In 1986 and again in 1993, the CPS changes to different and more disaggregated MSA definitions, and beginning in late 2004 there are identifiers for large counties and for Core Based Statistical Areas (CBSAs). While these geographical areas comprise precisely-defined sets of counties in any particular year, counties are often moved between areas over time. To avoid potentially confounding effects of changes in geographic definitions, we use only state identifiers.



indicators  $v_s$ , month indicators  $\mu_m$ , and year indicators  $\nu_t$ . With  $Y_{ismt}$  as an outcome variable (either wages or hours worked per week), the specification is

$$\ln Y_{ist} = \tau^r R_{st} \ln E_t + \lambda R_{st} + \rho X_i + v_s + \mu_m + \nu_t + \varepsilon_{ismt}. \quad (49)$$

To illustrate effects, we also plot estimates from another regression in which a separate  $\tau_t^r$  is estimated for each year of the sample:

$$\ln Y_{ist} = \tau_t^r R_{st} + \lambda R_{st} + \rho X_i + v_s + \mu_m + \nu_t + \varepsilon_{ismt}, \quad (50)$$

with 2001 as the omitted year in  $\tau_t^r$ .

For the MORG panel, we regress individual  $i$ 's change in wages in the 12 months between the interview on the change in the resource boom:

$$\Delta \ln Y_{ismt} = \tau^r R_{st} \Delta \ln E_t + \lambda R_{st} + v_s + \mu_m + \nu_t + \varepsilon_{ismt}. \quad (51)$$

### C.3 CPS Results

Appendix Table A4 presents the results of the Current Population Survey wage regressions. The top and bottom panels present estimates for all workers and manufacturing workers, respectively. Column 1 presents estimates of Equation (49) with natural log of wages as the dependent variable. The coefficient of 0.00962 implies that an oil and gas boom that increases national oil and gas employment by 100 log points increases relative wages by 0.962 percent in states with one additional standard deviation oil and gas endowment.

Appendix Figure A1 illustrates these results by plotting the  $\tau_t^r$  coefficients from Equation (50). The figure shows that states with one standard deviation additional endowment per square mile saw relative wages rise 1.7 to 3.2 percent higher than their year-2001 equilibrium for the entire eleven-year period between 1977 and 1987. Figure A2 presents analogous results, limiting to manufacturing workers only. While the estimates for manufacturing workers are noisier, the basic trend is very similar to the estimates for all workers.

Column 2 of Table A4 presents estimates of Equation (49), except with natural log of hours worked as the dependent variable. There are no statistically-significant effects on hours worked, and the standard errors rule out that a boom that increases national oil and gas employment by 100 log points increases hours worked by more than about one-third of one percent in states with one standard deviation additional endowment.

Column 3 presents estimates of Equation (49), but excluding controls  $X_i$  for age, education, race, and industry. The point estimates are statistically indistinguishable from those in column 1.

Because worker-level demographic controls do not affect the estimates and there are no significant effects on hours worked, this suggests that measures of earnings per worker available in the REIS provide reasonable approximations to the effects of oil and gas booms on “wages,” i.e. quality-adjusted per-unit labor input costs.

Column 4 presents estimates of Equation (51) using the MORG panel data, with the log of each individual’s 12-month wage change as the dependent variable. The sample size is much smaller because the MORG panel does not begin until 1979, because it includes only individuals who can be matched between their first and second outgoing rotations, and because each person is counted as one observation when calculating differences, while estimates of Equation (49) include both of an individual person’s observations. The results are almost exactly identical.

When limiting the sample to manufacturing workers only, the standard errors widen, but estimates are statistically indistinguishable from the effects on all workers, and the point estimates are actually larger. Columns 1-4 of Appendix Table A5 confirm that these results are robust to the same set of robustness checks that we carry out for the county-level data in Appendix Table A6; we omit the fixed effects robustness check from column 5 of Appendix Table A6 because the primary CPS specification already includes state indicators. Instead, column 5 of Appendix Table A5 shows that the estimates are robust to dropping outlying wages.

Table A4: **Effects of Resource Booms on Wages in Current Population Survey**

	(1)	(2)	(3)	(4)
Outcome Variable:	ln(wage)	ln(hours)	ln(wage)	$\Delta$ ln(wage)
All workers				
ln(National oil&gas employment) ×endowment	0.00962*** (0.00328)	-0.000411 (0.00218)	0.00832** (0.00402)	0.00969*** (0.00231)
Observations	5,696,671	5,723,794	5,696,671	1,587,855
Manufacturing workers				
ln(National oil&gas employment) ×endowment	0.0162** (0.00630)	-0.00194 (0.00397)	0.0102 (0.00690)	0.0162 (0.00982)
Observations	980,388	985,243	980,388	288,868
Age, education, gender, race, industry	Yes	Yes	No	No

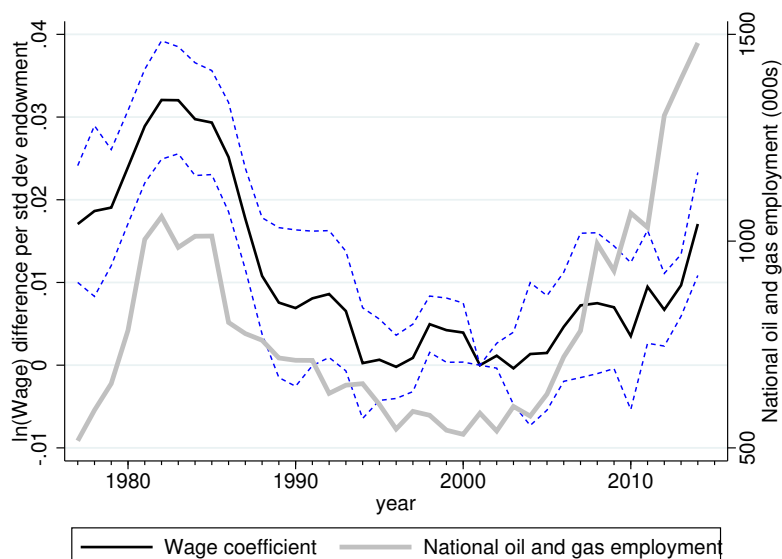
Notes: Columns 1-3 present estimates of variants of Equation (49), while column 4 presents estimates of Equation (51). All regressions include year, month, and state indicator variables. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent confidence, respectively. Robust standard errors in parentheses, clustered by state.

Table A5: Current Population Survey: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)
	Use early endow- ment $R_c^{early}$	Use total endowment $R_c^{total}$	Measure intensity by wage emp. only	Measure intensity by oil and gas price	Drop outlying wages
All workers					
ln(National intensity) ×endowment	0.00712*** (0.00232)	0.00834** (0.00315)	0.0289*** (0.00322)	0.00464** (0.00208)	0.00936*** (0.00300)
Observations	5,696,671	5,696,671	5,696,671	5,696,671	5,678,234
Manufacturing workers					
ln(National intensity) ×endowment	0.0148** (0.00554)	0.0145** (0.00651)	0.0289*** (0.00595)	0.00970*** (0.00357)	0.0165*** (0.00599)
Observations	980,388	980,388	980,388	980,388	978,752

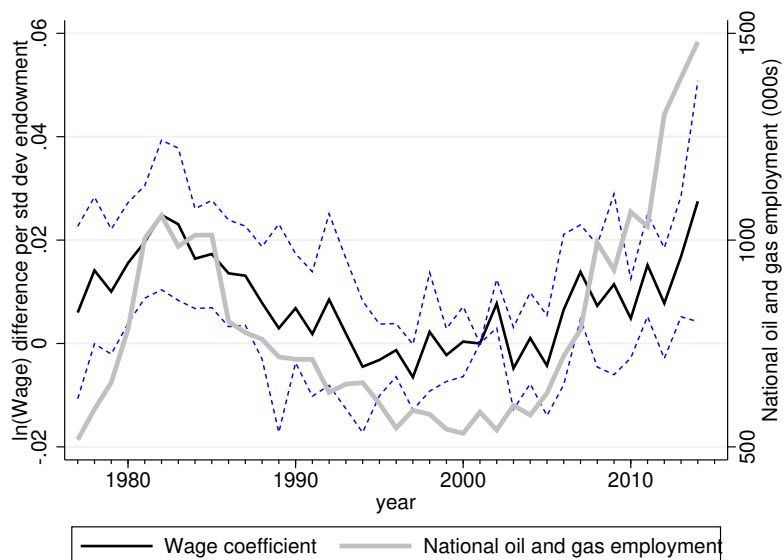
Notes: This table presents alternative estimates of Equation (49). Columns 1, 2, and 5 measure “National intensity” with “National oil&gas employment,” as in the main estimates. All regressions include year, month, and state indicator variables, plus age, education, gender, race, and industry controls. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent confidence, respectively. Standard errors are robust and clustered by state.

Figure A1: CPS Wages Over Time in Resource Abundant Counties



Notes: This figure presents the  $\tau_t^r$  coefficients and 90 percent confidence intervals from Equation (50) with natural log of hourly wage as the dependent variable. 2001 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 2001 association between endowments and wages.

Figure A2: CPS Manufacturing Wages Over Time in Resource Abundant Counties



Notes: This figure presents the  $\tau_t^r$  coefficients and 90 percent confidence intervals from Equation (50) with natural log of hourly wage as the dependent variable, using only the sample of manufacturing workers in the CPS. 2001 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 2001 association between endowments and wages.

## D Additional Empirical Results

## D.1 Additional Tables and Figures

Table A6: Effects on County-Level Outcomes: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)
	Use early endow- ment $R_c^{early}$	Use total endowment $R_c^{total}$	Measure intensity by wage emp. only	Measure intensity by oil and gas price	Fixed effects
Population					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0114*** (0.00194)	0.0130*** (0.00272)	0.0129*** (0.00205)	0.00324*** (0.000521)	0.00869 (0.00539)
Employment					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0273*** (0.00622)	0.0308*** (0.00684)	0.0316*** (0.00710)	0.00648*** (0.00138)	0.0174** (0.00703)
Housing rent					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0257*** (0.00723)	0.0351*** (0.0108)	0.0252*** (0.00761)	0.0108** (0.00418)	0.0121 (0.00784)
Earnings per worker					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0167*** (0.00468)	0.0206*** (0.00636)	0.0220*** (0.00633)	0.00609*** (0.00140)	0.0177*** (0.00640)
Manufacturing earnings per worker					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0114*** (0.00405)	0.0122*** (0.00423)	0.0164*** (0.00534)	0.00280 (0.00177)	0.00975 (0.00593)
Manufacturing employment					
$\Delta \ln(\text{National intensity})$ $\times \text{endowment}$	0.0309*** (0.00726)	0.0249*** (0.00737)	0.0357*** (0.00991)	0.00305* (0.00167)	0.0267** (0.0103)

Notes: This table presents alternative estimates of Equation (12). Columns 1, 2, and 5 measure “National intensity” with “National oil&gas employment,” as in the main estimates. The dependent variable is the change in the natural log of the listed outcome. Sample sizes are 135,274 (138,349) for the population estimates, with column 5 fixed effects sample size in parenthesis, 138,349 (141,424) for employment and earnings per worker, 15,371 (18,446) for housing rents, 111,709 (114,192) for manufacturing earnings per worker, and 111,754 (114,238) for manufacturing employment. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table A7: **Effects of Resource Booms on County-Level Outcomes: Separate Effects by Time Period**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/worker	Mfg. earnings/worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.0151***	0.0368***	0.0503**	0.0117**	-0.00331	0.0437***
$\times \text{endowment} \times 1(\text{boom})$	(0.00267)	(0.0106)	(0.0232)	(0.00532)	(0.00482)	(0.00876)
$\Delta \ln(\text{National oil\&gas emp})$	0.0145***	0.0372***	-0.0197	0.0215***	0.0309***	0.0327**
$\times \text{endowment} \times 1(\text{bust})$	(0.00312)	(0.00838)	(0.0233)	(0.00459)	(0.00792)	(0.0129)
$\Delta \ln(\text{National oil\&gas emp})$	-0.0103***	-0.0172**	0.0242	0.00285	-0.000779	-0.0365***
$\times \text{endowment} \times 1(\text{post-2001})$	(0.00263)	(0.00686)	(0.0192)	(0.00285)	(0.00754)	(0.00862)
Observations	135,274	138,349	15,371	138,349	111,709	111,754

Notes: This table presents estimates of Equation (12), allowing  $\tau^r$  to differ in three types of years: during “booms” (when  $\Delta E_t > 0$ ) and “busts” (when  $\Delta E_t \leq 0$ ), as well as for 2001 and after. The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table A8: **Effects on County-Level Outcomes: Controlling for Coal Boom**

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/worker	Mfg. earnings/worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.0116***	0.0278***	0.0265***	0.0175***	0.0106***	0.0294***
$\times \text{endowment}$	(0.00193)	(0.00585)	(0.00711)	(0.00478)	(0.00362)	(0.00730)
$\Delta \ln(\text{National coal emp})$	0.00554**	0.0142***	0.0184***	0.0189***	0.0127**	-0.00644
$\times \text{coal endowment} \times 1(\text{East})$	(0.00215)	(0.00477)	(0.00624)	(0.00450)	(0.00622)	(0.00605)
$\Delta \ln(\text{National coal emp})$	0.0300***	0.0469***	0.0319	0.0456***	0.0104	0.0259***
$\times \text{coal endowment} \times 1(\text{West})$	(0.00700)	(0.0136)	(0.0329)	(0.0130)	(0.0213)	(0.00721)
Observations	135,274	138,349	15,371	138,349	111,709	111,754

Notes: This table presents estimates of Equation (12), including effects of coal endowments. The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$  and  $\Delta \ln(\text{National coal emp})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Table A9: Effects of Resource Booms on Prices for Plants with Physical Output Data

	(1)	(2)	(3)	(4)	(5)
	All	Upstream / downstream	Non-linked	Non-linked and local	Non-linked and tradable
$\Delta \ln(\text{National oil\&gas employment})$	-0.00310	-0.00739	-0.000847	-0.00173	0.00280
$\times \text{endowment}$	(0.00307)	(0.00904)	(0.00151)	(0.00150)	(0.00316)
Observations	420,000	87,000	333,000	248,000	86,000

Notes: This table presents estimates of Equation (14), with the sample limited to plants that report physical output. The dependent variable is the natural log of price, at the product-by-plant-by-year level. All specifications use differenced outcomes; the time between each Census is five years. All regressions include Census division-by-year fixed effects. \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by county.

Table A10: Absolute Effects of Resource Booms on County-Level Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.00692***	0.0162***	0.0120*	0.0102***	0.00176	0.0166**
$\times \text{endowment}$	(0.00155)	(0.00579)	(0.00702)	(0.00304)	(0.00212)	(0.00626)
$\Delta \ln(\text{National oil\&gas emp})$	0.101	0.281**	0.267	0.203***	0.321***	0.556**
$\times \text{endowment from 0-50 miles}$	(0.0652)	(0.125)	(0.266)	(0.0726)	(0.106)	(0.240)
$\Delta \ln(\text{National oil\&gas emp})$	0.0945***	0.232**	0.436**	0.119**	0.0645	0.136**
$\times \text{endowment from 50-100 miles}$	(0.0268)	(0.103)	(0.199)	(0.0464)	(0.0665)	(0.0587)
$\Delta \ln(\text{National oil\&gas emp})$	0.0269**	0.000801	-0.0587	0.0460**	0.105***	-0.0475
$\times \text{endowment from 100-150 miles}$	(0.0124)	(0.0338)	(0.166)	(0.0202)	(0.0352)	(0.0496)
$\Delta \ln(\text{National oil\&gas emp})$	-0.00152	0.0260	0.0622	0.0122	0.0275	0.00203
$\times \text{endowment from 150-200 miles}$	(0.0116)	(0.0234)	(0.0925)	(0.0109)	(0.0372)	(0.0313)
$\Delta \ln(\text{National oil\&gas emp})$	-0.00322	0.0268*	-0.0854	-0.00521	-0.00392	0.0732
$\times \text{endowment from 200-250 miles}$	(0.0149)	(0.0144)	(0.0602)	(0.0119)	(0.0435)	(0.0489)
$\Delta \ln(\text{National oil\&gas emp})$	-0.000353	-0.00626	-0.00416	-0.00167	0.000436	0.0152
$\times \text{endowment from 250-300 miles}$	(0.00658)	(0.0146)	(0.0375)	(0.0129)	(0.0225)	(0.0291)
$\Delta \ln(\text{National oil\&gas emp})$	-0.0121	-0.0140	-0.0796**	0.00631	-0.00997	0.00589
$\times \text{endowment from 300-350 miles}$	(0.00965)	(0.0170)	(0.0303)	(0.0178)	(0.0287)	(0.0383)
$\Delta \ln(\text{National oil\&gas emp})$	-0.0239***	0.000292	-0.0685	-0.0221	-0.00793	0.00645
$\times \text{endowment from 350-400 miles}$	(0.00870)	(0.0178)	(0.0470)	(0.0152)	(0.0400)	(0.0324)
Observations	135,274	138,349	15,371	138,349	111,709	111,754

Notes: This table presents estimates of Equation (13). The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

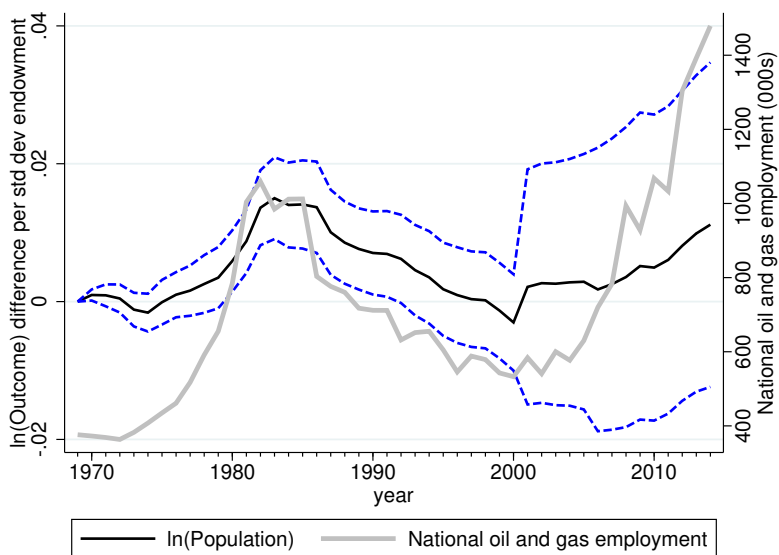


Table A11: **Absolute Effects of Resource Booms on County-Level Outcomes: Excluding Large Counties**

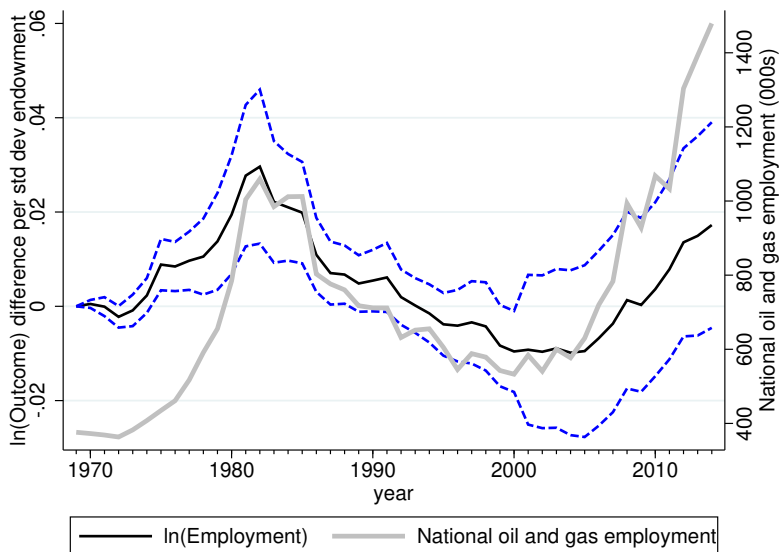
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome:	Population	Employment	Housing rent	Earnings/ worker	Mfg. earnings/ worker	Mfg. employment
$\Delta \ln(\text{National oil\&gas emp})$	0.00736***	0.0171***	0.0112	0.0104***	0.00195	0.0166**
$\times \text{endowment}$	(0.00176)	(0.00627)	(0.00718)	(0.00343)	(0.00226)	(0.00664)
Observations	129,818	132,769	14,751	132,769	106,264	106,309

Notes: This table presents estimates of Equation (13). It parallels Table 10, except excluding counties with population larger than 250,000 people in 1960. The dependent variable is the change in the natural log of the listed outcome. All regressions include census division-by-year indicators and controls for year interacted with natural log of the outcome variable in two baseline years. Population effects in column (1) use the one year lag of  $\Delta \ln(\text{National oil\&gas employment})$ . \*, \*\*, \*\*\*: Statistically different from zero with 90, 95, and 99 percent certainty, respectively. Robust standard errors in parentheses, clustered by state.

Figure A3: County Aggregates Over Time in Resource Abundant Counties



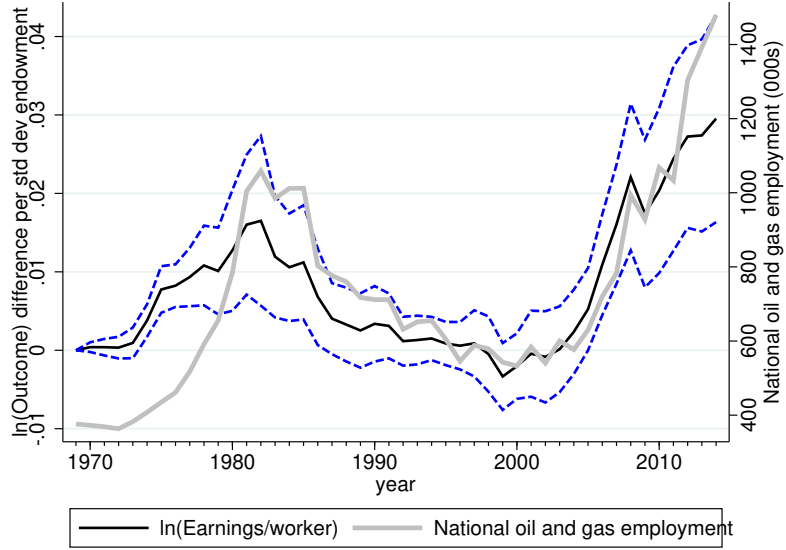
(a) Population



(b) Employment

Figure A3

(c) Earnings per Worker



(d) Manufacturing Earnings per Worker

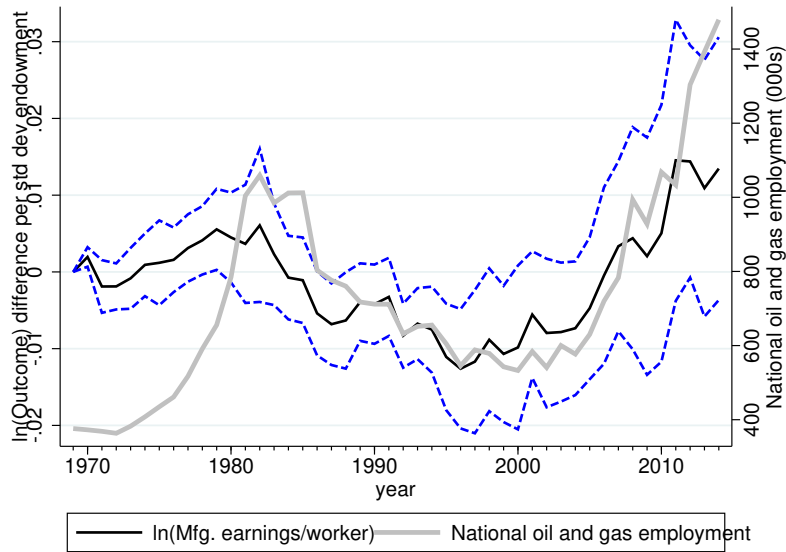
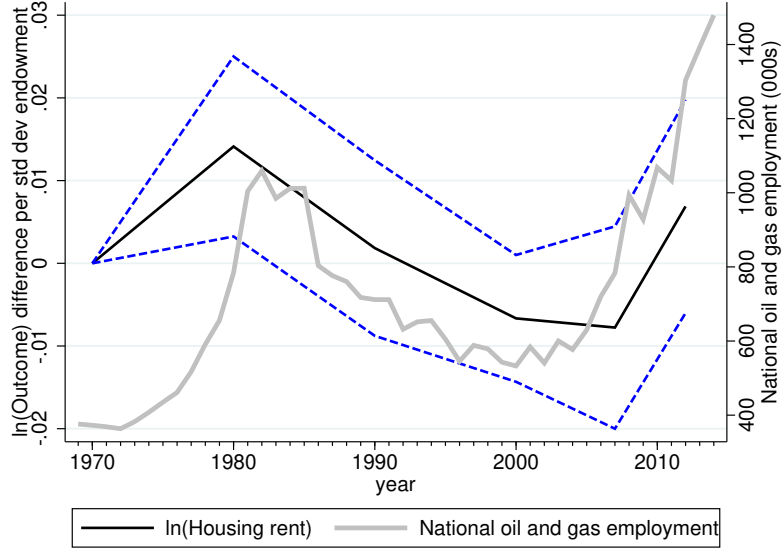


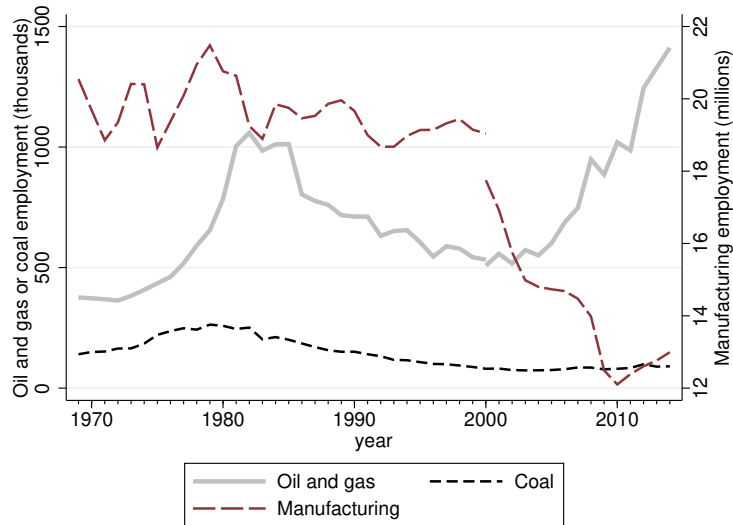
Figure A3

(e) Housing Rents



Notes: These figures present the  $\tau_t^r$  coefficients and 90 percent confidence intervals from Equation (11) with various dependent variables. 1969 is the omitted year in  $\tau_t^r$ , so the regression coefficients are relative to the 1969 associations between resources and outcomes. The point estimates for the population, employment, and earnings per worker graphs are the same as in Figure 5.

Figure A4: National Manufacturing, Oil and Gas, and Coal Employment



Notes: Employment data are from the Regional Economic Information System. We switch from the SIC to the NAICS classification system in 2000 and plot both data points in that year.