Effects of Oil Booms on the Local Environment

Jingjing Cai* Dakshina G. De Silva† Aurelie Slechten‡

Abstract

We analyze the impact of oil and gas booms on local environmental quality in school districts in Texas between 2010 and 2014. Using data from the Toxic Release Inventory (TRI) and County Business Patterns, we distinguish economic activity associated with potential and actual polluters. We find that the presence of oil and gas resources in a school district has spillover effects in terms of economic activity by attracting more potentially polluting firms. Oil abundance also generates an actual environmental burden for school districts located in MSAs as the proportion of firms that actually report a release of toxic chemicals to the TRI increases with oil revenue.

Keywords: oil boom, environmental quality, TRI, local pollution

JEL Classification : Q32, R12, H72, H75

*Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: j.cai2@lancaster.ac.uk).

†Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: d.desilva@lancaster.ac.uk).

‡*Corresponding Author*: Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: a.slechten@lancaster.ac.uk).

1 Introduction

New oil and gas drilling brings economic activity to the local communities, but there are substantial concerns about potential impacts on the quality of life of local residents, including pollution, traffic congestion, and crime. In this paper, we contribute to the debate by investigating whether the activity and employment spillovers generated by oil and gas booms are associated with indirect adverse environmental effects on local communities. In most countries, local governments have some degree of autonomy when it comes to the decision to allow new resource extraction activities. With the development of hydraulic fracturing (fracking) technologies and the existence of vast shale deposits around the world, providing a broad picture of the costs associated with these activities is important. It will help local governments design complementary policies that ensure that the local benefits of oil and gas development outweigh any potential cost.

This paper uses school district-level data from the state of Texas for the period 2010-2014. Texas is an ideal setting in which to observe the impacts of oil and gas operations on local environmental quality. This state has experienced an oil and gas boom over the last 10 years due to the development of extracting technology. Annual crude oil production nearly tripled between 2009 and 2015. Texas is the biggest crude oil-producing state and it produces one-third of U.S. crude oil and one-fourth of U.S. natural gas (U.S. EIA, 2015). The Permian Basin in West Texas has become the world's most productive oil field (U.S. EIA, 2019). The Texan economy relies heavily on the oil and gas sector.[1] Further, there are few local environmental restrictions imposed in the state of Texas beyond local zoning laws, and the state itself takes a relatively light hand to regulation in general. Thus, jurisdictions (school districts in our case) in the state are largely subject to a practically identical regulatory environment.

This study makes use of the unique features of the data from the Toxic Release Inventory (TRI) to study the local environmental effects of oil and gas exploitation. The TRI is a mandatory reporting program that requires private and government facilities from a set of industries to report annually how they manage certain toxic chemicals. The chemicals covered by the TRI Program are typically local and are known to have harmful health effects.[2] More importantly, TRI data allow us to distinguish between potential polluters and actual polluters.

Under the TRI Program, only firms in a subset of the North American Industry Classification System (NAICS) that employ at least 10 full-time employees (FTEs) and exceed the Environmental Protection Agency (EPA) threshold limits in terms of their processing or usage of designated hazardous or toxic chemicals are subject to mandatory reporting within the TRI.[3] The firms subject to mandatory reporting are denoted in this paper as TRI reporting firms. TRI reporting firms responsible for toxic chemical releases to the environment that exceed TRI limits are identified and treated in this paper as TRI polluters. A potentially polluting firm (or TRI-type firm) is then defined as any firm, regardless of size or reporting requirements, in a

[1] The value of oil and gas production in Texas represented 13.5% of its GDP in 2014 (*https://businessintexas.com/sites/default/files/txoil.pdf*).

[2] See Currie et al. (2015) or Aizer et al. (2018).

[3] https://www.epa.gov/toxics-release-inventory-tri-program

NAICS code identified by the TRI. So, TRI polluters are a subset of TRI reporting firms and TRI reporting firms are a subset of TRI-type (potentially polluting) firms.

Our analysis proceeds in two steps. We first investigate whether oil and gas revenue influences the location choices of potentially polluting firms from all sectors covered by the TRI program. A larger number of TRI-type firms shouldn't necessarily be seen as a negative effect because more potentially polluting firms implies more economic activity and more job opportunities for local residents. However, if this additional economic activity generates toxic chemical releases, then oil booms result in actual environmental costs for the local community. We examine this possibility in the second step of our analysis by estimating the impact of oil and gas revenue on the number and proportion of TRI polluters.

To deal with a potential endogeneity between firms' location decisions and oil and gas revenue and accurately estimate the environmental impact of oil abundance, we use an Instrumental Variable (IV) approach. We create an indicator that equals 1 if this school district is in an oil/gas basin county (defined as a county located in any of the Texan oil/gas basins). As the boundaries of Texan counties were defined before the discovery of oil, the location of oil resources does not directly affect our dependent variables.[4] The school district oil and gas revenue is then instrumented by the interaction between our basin dummy and year indicators to allow for temporal variation (as in Feyrer et al., 2017 or Jacobsen, 2019).

To ensure that our results are not driven by the most rural school districts, we estimate our empirical models separately for school districts located in a Metropolitan

[4]This is similar to the approach used by Michaels (2010), who proxies oil abundance with a dummy variable for whether a county lies on a large existing oilfield.

Statistical Area (MSA) and those located outside MSA boundaries. Our findings suggest that the presence of oil and gas resources attracts more potentially polluting firms to both MSA and non-MSA school districts. We also find that oil abundance generates an actual environmental burden for school districts located in MSAs as the proportion of firms that actually report a release of toxic chemicals to the TRI Program is higher in MSA school districts experiencing an oil boom. This is problematic as MSA school districts are more densely populated than rural areas.

These findings provide new insights into the impact of resource abundance on local amenities by identifying indirect environmental effects at the local level. Papers by Bartik et al. (2019), Muehlenbachs et al. (2015), and Jacobsen (2019) study the impacts of the recent fracking booms on various measures of local amenities, including crime, noise, traffic, and housing values. The effects of natural resource abundance on local public goods provision and local finance have also been explored (Caselli and Michaels, 2013; Borges et al., 2015; James, 2015; James, 2017; Marchand and Weber, 2019).

Further, our results contribute to the strand of the literature studying the effects of resource booms on the local economy and local labor markets. Expanded oil and gas exploitation has been shown to create jobs and increase wages (Weber, 2012; Feyrer et al., 2017; Wang, 2018; Allcott and Keniston, 2018; De Silva et al., 2020). This literature has also identified employment spillover effects of oil and gas abundance. However, there is no consensus on the size or on the sectors benefiting from these effects. Some papers document the existence of employment spillover effects in traded goods industries, e.g., manufacturing (Michaels, 2010; Weber, 2014;

Allcott and Keniston, 2018), while other studies show that these effects are found only in local sectors, e.g., retail or construction (Black et al., 2005; Marchand, 2012; Brown, 2014). By and large, the potentially polluting activities considered in our paper (and not related to oil and gas extraction) result from industrial activities whose output is not dependent on the local market, i.e., production of tradeables. Our findings therefore provide some evidence supporting the existence of spillover effects in traded goods sectors.

The remainder of the paper is organized as follows. Section 2 describes the data and variables used in the empirical analysis. In Section 3, we evaluate the environmental costs associated with oil and gas booms. Section 4 concludes the study.

2 Data

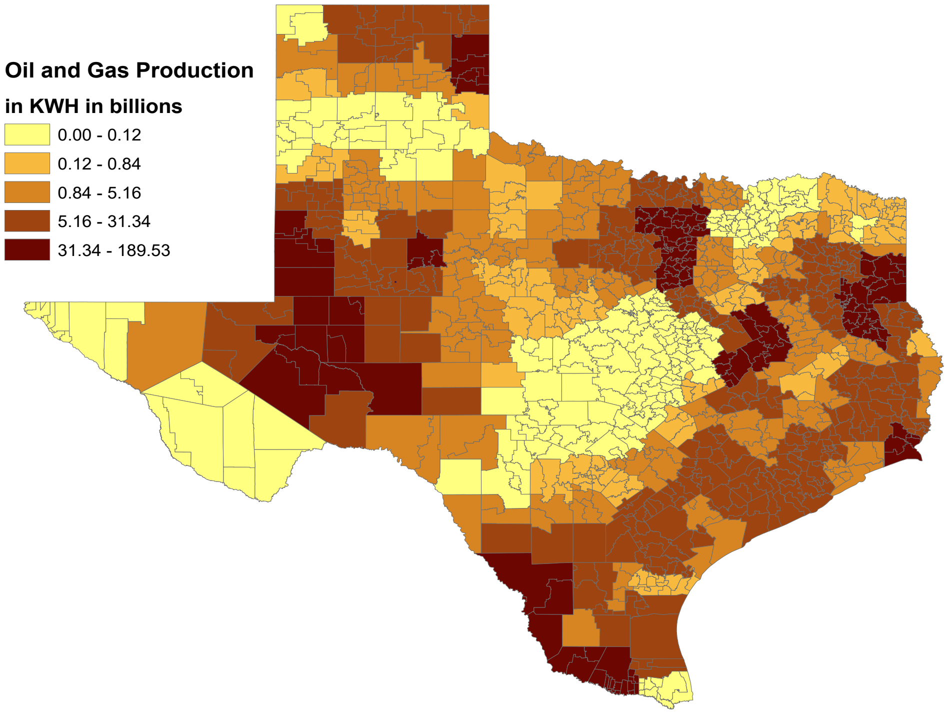
In this section, we describe the data sources, explain the construction of our variables, and provide summary statistics. We use data at the school district level from Texas over a five-year period (2010 to 2014). School districts constitute a good institutional framework to study local environmental impacts of oil booms. They are relatively small areas and closely represent the population that would bear the immediate environmental impact of increased economy activity due to oil abundance. Moreover, a school district (as opposed to a census tract) is an independent government with some fiscal autonomy for the purpose of operating public schools that are situated within that area. In particular, every school district is authorized to set its own property tax and oil and gas companies pay a property tax based on the value of their production. A school district is therefore a small area that might benefit from oil and gas extraction but may also bear the potential environmental costs. During our sample period, there were 1024 school districts in Texas.

2.1 Oil and gas production and revenue

To measure oil abundance at the school district level, we use two alternative explanatory variables: oil and gas production and oil and gas revenue. The data on oil and gas production comes from the Railroad Commission of Texas (RRC). It includes county-level crude oil production in thousands of barrels, condensate oil production in thousands of barrels, gas-well gas production in thousands of cubic feet, and casing-head gas production in thousands of cubic feet. To derive the total level of oil and gas production at the county level, we convert all four types of oil and gas production into kilowatt hours (kWh) and add them up. As our analysis is at the school district level and all school districts are contained within a single county, we level down the county-level production using the proportion of the school district area contained in the county area. Based on the average yearly price of oil in dollars per barrel and gas in dollars per thousand cubic feet (data from U.S. Energy Information Administration), we calculate the revenue generated by oil and gas extraction.

Figure 1 shows school district-level oil and gas production in 2010. The distribution of oil and gas activity in Texas is consistent with the location of the main oil and gas resources. There are four principal zones: the Permian Basin in West Texas,

Figure 1: **School district-level oil and natural gas production in Texas in 2010**



the Eagle Ford shale formation located in the Gulf Coast Basin in South Texas, the Barnett shale formation in North Texas, and the Haynesville/Bossier shale formation in East Texas. For example, the majority of school districts in the Gulf Coast Basin had a production level higher than five billion kWh in 2010 compared to less than 0.12 billion kWh in Central Texas.

2.2 Sample generation

The objective of this paper is to compare school districts that have witnessed an oil boom over the sample period with those that have no specialization in oil and gas production. It is, therefore, important to narrow the analysis to school districts that have some degree of similarity. To this end, we restrict our sample in two different ways.

First, we identify the areas that are specialized in oil and gas production. Because the original data of oil and gas production is at the county level, we identify oil (and gas) counties in Texas. If oil and gas revenue at any time is greater than ten percent of a county's total revenue, that county is treated as an "oil county"; otherwise, it is a "non-oil county". Our first restriction on the sample of school districts is based on population and median income in school districts located in oil counties; the former signals the size of a school district and the latter indicates local living standards. Our analysis excludes school districts with a population less than 69 or larger than 164,642 (the smallest and the largest school districts in terms of population in the oil counties), or a median income less than $15,917 or greater than $92,917 (the lowest and the highest school district median incomes in the oil counties). This restriction reduces the number of school districts under study from 1024 to 980.

Second, it would not be appropriate to compare the impact of an oil boom between rural and urban school districts as they widely differ in terms of population growth, employment, etc. We, therefore, split our sample into two subsets: school districts located in an MSA and school districts outside MSA boundaries. There are 25 MSAs in Texas, corresponding to 455 school districts in our sample.

Table 1 displays the summary statistics. We provide three categories of data. "Sample SD" refers to our restricted sample of 980 school districts in Texas. "MSA SD" and "non-MSA SD" refer to MSA school districts and non-MSA school districts. For each category, we compare school districts located in oil counties and non-oil counties. The definitions of all the variables are in Table A.1 in the Appendix.

2.3 TRI data

Production and transport of coal, natural gas, and oil provide many opportunities for the release of air pollutants (e.g. carbon dioxide, methane...) which may be hazardous to the health of local residents and speed up climate change. Local opposition to fracking has also emerged due to the potential damage from methane leakages or water contamination. However, in this paper, we are interested in the broader environmental impact of oil and gas production. Oil booms bring in more economic activity from other industries which can potentially result in adverse environmental effects on local communities.

To measure this indirect environmental impact of oil booms, we use data from the TRI. The TRI is a U.S. database established by law which requires private and government facilities to report annually how much of certain chemicals is released to the environment or managed through recycling, energy recovery and treatment. It covers a specific subset of NAICS codes and around 600 different toxic chemicals. We believe the data from the TRI Program constitute a good proxy for local environmental quality. First, most chemicals included in the TRI Program have very localized impacts. Using individual level data, Currie et al. (2015) show that the openings or closings of toxic plants (i.e., plants reporting a release to the TRI Program) have an impact on housing prices and birth outcomes within a 1-mile radius of the plant location.

Second, some of these chemicals have been shown to pose a threat to human health and the environment. For example, Currie et al. (2015) show that a reporting plant's operation is associated with a roughly 3 percent increase in the probability of low birth weight within a mile. Working at the county level, Currie and Schmieder (2009) find strong evidence that fetal exposure to most reported TRI-chemicals has a negative effect on health at birth and subsequent infant mortality. Aizer et al. (2018) show that there might also be long-term health effects.[5]

We define three subsets of TRI-related facilities: TRI reporters, TRI polluters and TRI-type firms. The EPCRA (Emergency Planning and Community Right to Know Act) Section 313 requires TRI reports to be filed by owners and operators of facilities that meet all of the following criteria:

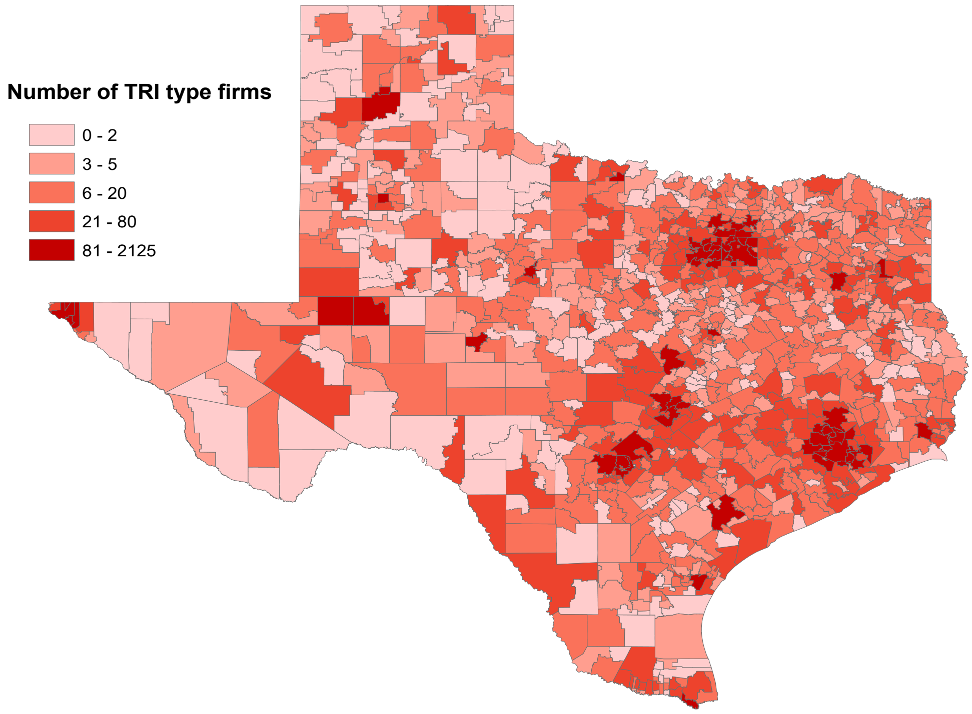
- The facility has 10 or more full-time employee equivalents (FTE);
- The facility is included in a given subset of the NAICS. These NAICS codes are at the 6-digit level. This is the most detailed classification one can get; and
- The facility manufactures (defined to include importing), processes, or otherwise uses any EPCRA Section 313 chemical in quantities greater than the established threshold in the course of a calendar year.[6]

Facilities that meet all these requirements are classified as TRI reporters. When these firms exceed the toxic release limits set by the EPA (25,000 toxic pounds), they are considered TRI polluters for the year a release is reported. A facility located in a NAICS code covered by the TRI Program, regardless of whether it meets the

[5]In particular, they show that one unit decrease in average blood-lead levels (a TRI-listed chemical) reduces the probability of being substantially below proficient in reading.

[6]See https://www.epa.gov/toxics-release-inventory-tri-program for details on on NAICS codes, listed chemicals, and chemical thresholds required for reporting.

Figure 2: **TRI type firms in Texas in 2010**



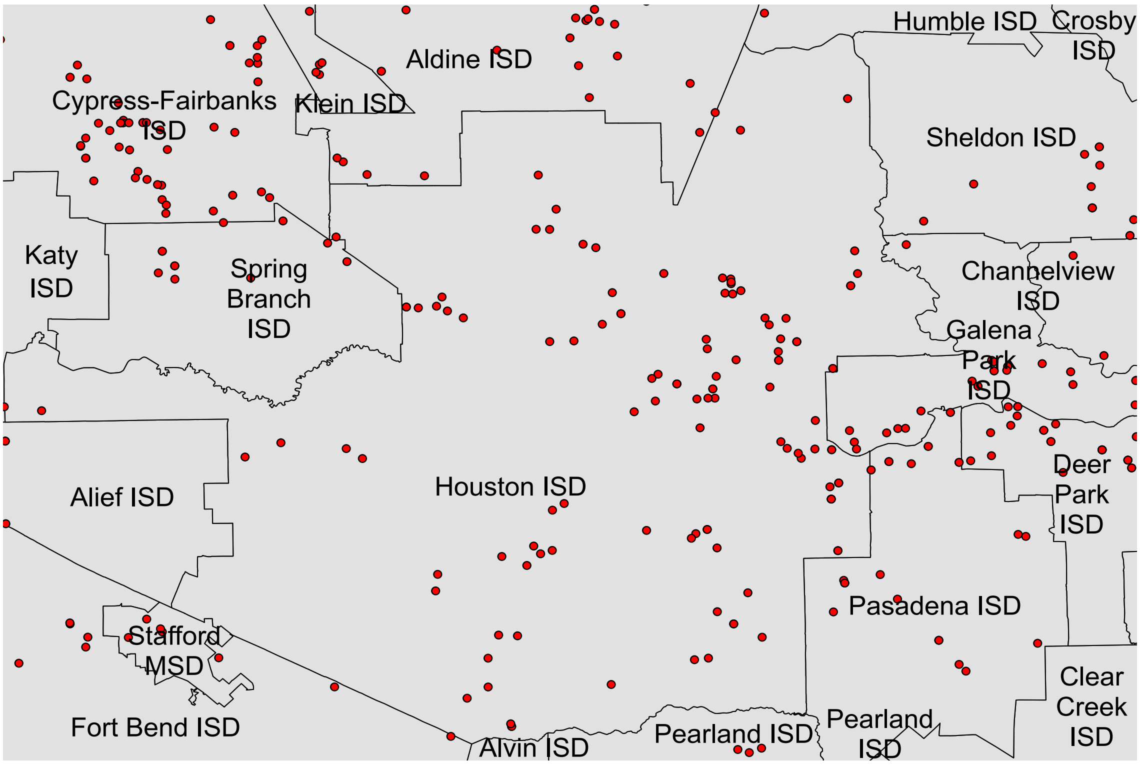
other two requirements for mandatory reporting, is denoted as a TRI-type firm. To count the number of TRI-type firms at the school district level, we use data from the County Business Patterns (CBP). CBP data contain the number of establishments in each NAICS code at the county level. We select the establishments located in a NAICS code subject to TRI reporting in each county. This gives us the number of TRI-type firms at the county level, which we level down to the school district using the population distribution. Figure 2 shows the distribution of TRI-type firms in Texas in 2010. A higher number of TRI-type firms can be observed in the school districts near Dallas and Houston.

The number of TRI-type firms is an indicator of economic activity associated with potential polluters. We are working at sufficient industry detail (six-digit NAICS codes) that a reasonable level of homogeneity in activity can be assumed. If establishments in a given industry have been identified as having experienced a release via TRI reporting, we assume that other establishments in that same industry have largely similar activities and could potentially experience a similar release. In that respect, the number of TRI-type firms is also a proxy for potential environmental risks (as it affects the likelihood of toxic releases in the school district). A larger number of TRI-type firms in a school district can be seen as a positive effect, as it brings in more economic activity. At the same time, a larger number of potentially polluting firms also implies that the likelihood of toxic releases is higher.

Some oil and gas facilities are included in the TRI Program because they deal with around 25 different TRI-listed chemicals, including hydrogen sulfide, benzene, toluene, ethylbenzene, and xylene.[7] Table A.2 in the Appendix provides a list of oil-based TRI NAICS codes, i.e., NAICS codes covered by the TRI Program and related to oil and gas exploitation. As the focus of our paper is the indirect environmental costs associated with oil abundance, we divide TRI-type firms into two categories. The first category, "oil-based TRI-type firms", refers to TRI-type firms that extract oil and gas or produce oil- and gas-related products (NAICS codes listed in Table A.2). The second category, "non-oil-based TRI-type firms," covers the remaining TRI-type firms that do not relate to the oil and gas industry. This classification allows us to analyze whether the presence of oil resources attracts firms from other potentially polluting industries. The average number of TRI-type firms in non-oil counties is higher than in oil counties (see Table 1). The difference is the largest in MSA school districts: an average of 62 TRI-type firms in non-oil counties compared

[7] https://www.reginfo.gov/public/do/eAgendaViewRule?pubId=201610& RIN=2070-AK16.

Figure 3: **TRI polluters in Houston Independent School District in 2010**



to 20 in oil counties.

The number of TRI-type firms is related to potential pollution. To obtain a measure of the actual environmental cost, we use the number of TRI polluters and the proportion of TRI polluters to TRI-type firms. As TRI data provide the address of reporting facilities, we can easily compute the number of TRI polluters at the school district level. Over our sample period, MSA school districts have, on average, more TRI polluters than non-MSA school districts. Figure 3 shows the distribution of the 114 TRI polluters in the Houston Independent School District in 2010. This is the largest number of TRI polluters at the school district level in our sample.

2.4 Other controls

The literature on firm location decisions postulates that, in a profit maximization framework, a firm considering the location of a new plant will choose the area with attributes that allow this plant to operate at the lowest cost (Keller and Levinson, 2002; De Silva et al., 2016). Therefore, in our analysis of the environmental cost induced by an oil boom, we include a set of input factors at the school district level that may affect firm location decisions: unemployment rates and population to capture labor availability, median income to account for living conditions, house rental ratio to explain house-occupying status, transportation costs (such as the number of roads and railways), and the size of the school district to measure land availability. To incorporate insights from the environmental justice literature, we also include the non-white ratio, defined as the proportion of non-white residents in the school district.

School district-level data comes from the American Community Survey (ACS), except for the information regarding number of roads and railways that are computed using the U.S. Census Bureau's Census Feature Class Codes (CFCC) and ESRI Data & Maps (2000). The ACS is a series of survey databases including detailed demographic and economic information at the school district level. The highest median income is in MSA school districts within non-oil counties, with an average of $55,100 (see Table 1). The highest non-white ratio (18 percent) is also observed in these school districts. Finally, we control for other businesses that are not covered by the TRI Program to account for local amenities. From CPB data, we obtain the number of all businesses at the school district level. We then deduct the number of TRI-type firms from the total number of businesses to compute the number of other businesses.

2.5 Identification Strategy and Instrument

The number of potentially polluting firms and oil and gas production/revenue can evolve simultaneously because of unobserved geographical characteristics or because companies from both oil/gas and other industries respond to policies implemented by local governments. To deal with this issue and accurately estimate the impact of oil and gas production on local environmental quality, we use an Instrumental Variable (IV) approach.

One approach to address the endogeneity of resource booms is to classify counties based on geological characteristics such as reserves of oil and gas (Michaels, 2010). We identify the major oil and gas basins in Texas, i.e., the Permian Basin in West Texas, the Eagle Ford shale formation in South Texas, the Barnett shale formation in North Texas, and the Haynesville/Bossier shale formation in East Texas. Counties located in any part of one of these basins are basin counties whereas the others are non-basin counties. We then create an indicator, D_i, that equals 1 if this school district is in an oil/gas basin county.

The school district oil and gas production/revenue is instrumented by the interaction between D_i and year indicators. The year indicator variable within the interaction allows us to capture the timing of the booms or changes in the world oil and gas prices (James, 2017; Feyrer et al., 2017).

Due to the uneven distribution of oil and gas resources in Texas, there is enough variation between school districts to identify the local effect of oil and gas production/revenue. Moreover, the location of the oil resources in Texas does not directly affect our dependent variables and vice-versa because the boundaries of Texan counties were defined before the discovery of oil and, thus, are not based on the presence of oil resources. The only possible indirect impact of oil and gas basins on these dependent variables must be through current oil and gas extraction.

3 Empirical analysis

3.1 Economic Activity and Environmental risk

To examine the impact of an oil boom on local economic activity associated with potential polluters, we estimate an empirical model that takes the following form:

$$y_{it} = \beta \; log \; p_{it} + s_{it}'\gamma + z_i'\delta + \tau_t + \varepsilon_{it} \tag{1}$$

Our dependent variable (y) is the log of the number of TRI-type firms, oil-based TRI-type firms, or non-oil TRI-type firms in a given school district i in a given year t. The explanatory variables can be divided into four groups: school district-level oil and gas production or oil and gas revenue (p, instrumented by the interaction between D_i and year indicators); school district-level characteristics (s) that vary with time, such as median income and population; time-invariant school district attributes (z) such as number of roads; and year effects (τ). The last term ε_{it} is an error term.

We estimate this model using two different approaches. The first approach is a linear IV regression specification. Estimation results are presented in Table 2 and Table 4 for MSA and non-MSA school districts, respectively. However, our dependent variables are left-censored. In this case, a linear model may provide inconsistent estimates of the parameters. It will also predict values of the dependent variables below zero. Therefore, we estimate equation (1) using a censored IV specification. Estimation results are presented in Table 3 and Table 5 for MSA and non-MSA school districts, respectively.

For both specifications, oil and gas operations have a significant and positive effect on the number of TRI-type firms in MSA and non-MSA school districts. A higher level of oil and gas production or revenue attracts more potential polluters. Not surprisingly, in all cases, the effect of oil and gas production or revenue is larger for TRI-type firms in TRI NAICS codes related to the oil and gas industry. The effect on the number of non-oil TRI-type firms is still positive in all specifications for both MSA and non-MSA school districts. However, the effect is statistically significant for MSA school districts only (in the censored IV regression). This suggests the existence of spillover effects in terms of economic activity. The fact that this effect is present in MSA school districts only might be due to the attractiveness (in terms of infrastructure, proximity to consumers, etc.) of these areas compared to remote rural neighborhoods in non-MSA school districts.

In MSA and non-MSA school districts, the presence of other businesses is positively associated with the number of TRI-type firms (oil-based or not). This is also the case for median income, except for oil-based TRI-type firms in MSA school districts. The percentage of non-white residents has a positive correlation with the number of TRI-type firms in MSA school districts only. This is consistent with the environmental justice literature (De Silva et al., 2016). Finally, a larger number of roads in a school district attracts more potential polluters.

Note that, in Table 2 and Table 4 (linear IV specification), all our models pass the Weak-identification test (F-test reported in the tables). In Table 3 and Table 5 (censored IV specification), we report the p-value for the Hausman test. For all TRI-type and oil-based TRI firms (columns 1, 2, 4, and 5), we can reject the null hypothesis that both censored and IV censored regressions produce consistent results. The Hausman test of columns 3 and 6 (non-oil TRI-type firms) shows that IV and non-IV regressions yield similar results.

As a robustness check, we also estimate equation (1) using the Poisson Quasi-Maximum Likelihood (PQML) method with year fixed effects, which allows us to account for the count structure of our data. Note that in this case, our dependent variable is the number of TRI-type firms, oil-based TRI-type firms, or non-oil TRI-type firms in a given school district i in a given year t. Compared to the standard Poisson estimation, the PQML estimation does not assume that the data are distributed with the mean equal to the variance of the event count. The data need not even come from a Poisson process and may be either under or over-dispersed. It requires only that the conditional mean function is correctly specified. As shown in Table 6, the IV Poisson regression results are very similar to the results of the

Censored IV specification (Tables 3 and 5).[8]

Given these findings, one could question whether our results are driven by school district-level infrastructure and other demographic characteristics. Hence, we estimate a parsimonious empirical model controlling only for number of other businesses and year effects. We use the number of other businesses to control for the size/scale of the school district. We present the linear IV regression results in Tables A.3 (for MSAs) and A.4 (for non-MSAs) and the IV Poisson regression results in Table A.5. The interpretation of the findings in these tables is qualitatively the same as for the results we discussed in Tables 2, 4 and 6.[9]

3.2 Actual Environmental Costs

As shown in the previous section, oil abundance increases economic activity associated with potential polluters. The next question is whether the higher potential environmental risk resulting from this activity leads to a higher actual environmental cost. To investigate this question, we estimate equation (1), where y_{it} is either the number of TRI polluters or the proportion of TRI polluters (defined as the number of TRI polluters divided by the number of TRI-type firms).[10] The number of school districts that have TRI polluters is a small fraction of all school districts: over 50

[8]One advantage of the PQML estimator is that it allows for fixed effects (unlike censored regressions). However, given that we have a short sample period (5 years) and the within variation for most variables is relatively small, taking school district and time fixed effects eliminates all the variation. For example, on average, the mean of the number of TRI-type firms in an MSA school district is 27 with a standard deviation of 1. For non-MSA school districts, the mean of the number of TRI-type firms is 8, on average, with a standard deviation less than 1.

[9]We also estimate these specifications using a censored IV regression technique. The results are qualitatively similar to what we observe in Tables 3 and 5. In the interest of brevity, we don't report those results, but we can provide them upon request.

[10]The results for production and revenue are very similar. This is why, in this section, we focus only on oil and gas revenue.

percent of MSA school districts and over 75 percent of non-MSA school districts do not have TRI polluters. The average proportion of TRI polluters in MSA school districts is 0.08 compared to 0.05 in non-MSA school districts.

As in the previous section, when the dependent variable y_{it} in equation (1) is the number of TRI polluters, we use three empirical approaches: a linear IV specification, a censored IV specification and an IV Poisson specification. When the dependent variable y_{it} is the proportion of TRI polluters, we use a Wooldridge's two-step probit model (Wooldridge, 2010). The results in Table 7 indicate that oil and gas revenue has an impact on the total number of TRI polluters and the proportion of TRI polluters in MSA school districts only.[11]

Beyond these observations of interest to us, we see that the median income and the non-white ratio have a positive effect on the number and proportion of TRI polluters. The number of other businesses has a positive impact on the number of TRI polluters, but a negative effect on the proportion of TRI polluters. The number of rail roads positively affects the number and proportion of TRI polluters, while the number of roads matters only for the number of TRI polluters. As before, we estimate a parsimonious empirical model controlling only for number of other businesses and year effects. We present these IV regression results in Table A.6.

The coefficients presented for the censored regressions in Table 3 (all columns) and Table 7 (columns 3 and 4) are the average marginal effects. For an average school district in an MSA, a 1% increase in oil and gas revenue implies an increase

[11]In columns 1 and 2 (linear specification), we show that both models pass the weak-identification test. In columns 3 and 4, we report the Hausman test and show that we cannot reject the null hypothesis of no endogeneity. For the fractional probit, we use the Chi-square Wald test of exogeneity. We can reject the null hypothesis of no endogeneity in column 8, but not in column 7 (note that the test statistic in column 7 is very close to the critical value).

in the number of TRI-type firms (non-oil related) of 0.012% (see column 3 in Table 3) and an increase in the number of TRI polluters of 0.016% (see column 3 in Table 7). Over our sample period, oil and gas revenue in those school districts has increased by 41%. The average MSA school district had 25 TRI-type firms and 2.35 TRI polluters in 2010. As a result, the oil and gas boom in Texas has attracted 0.123 new TRI-type firms (non-oil related) in the average MSA school district between 2010 and 2014. It has also increased the number of firms reporting a release to the TRI by 0.0154. Even though the magnitude of those effects seems small, we have to bear in mind that school districts are relatively small areas (especially in MSAs) and even one additional TRI polluter might generate adverse environmental effects. The literature has indeed shown that the TRI-listed chemicals have serious effects on human health (Currie et al., 2015; Currie and Schmieder, 2009).

This suggests that, even though oil booms in Texan counties brought in more economic activity, they also resulted in a degradation of local environmental quality (measured as an increase in reported toxic releases) in MSA school districts. This last result raises some environmental justice concerns (Hamilton, 1995; De Silva et al., 2016) because MSAs seem to bear more environmental costs than non-MSAs and they are more densely populated areas. Any chemical release from these additional polluting firms would adversely affect a larger number of people than in rural areas.

4 Conclusion

In this paper, we analyze the impact of oil booms on local environmental quality using school district-level data from Texas between 2010 and 2014. Because school districts are relatively small areas, they constitute a good proxy for the locality adjacent to any potentially polluting firm located in the school district. To deal with the potential endogeneity between our dependent variables and oil and gas production, we use an IV approach in which school district oil and gas revenue is instrumented by the interaction of an indicator that equals 1 if this school district is in an oil basin county and an indicator of year.

We show that an increase in oil and gas revenue attracts more potentially polluting firms from various sectors covered by the TRI Program (i.e., firms that use toxic chemicals, but not necessarily report releases). While this might be seen as a positive impact in terms of economic activity, we also find that the proportion of firms that actually report a release is higher in school districts experiencing an oil boom. This negative environmental effect is stronger in MSA school districts, which are also more densely populated.

The pollutants covered by the TRI Program are toxic chemicals that pose a serious threat to human health and the environment. Our analysis, therefore, suggests that encouraging oil and gas exploitation might lead to substantial local environmental degradation. Given the recent oil discoveries in the US Gulf of Mexico and the existence of important shale deposits around the world, the issue of the local impacts of oil abundance is politically relevant. If the attainment of greater environmental quality is a policy goal, serious thought should then be given to complementary environmental regulations or to new regulations on compensation schemes designed to offset the costs of a higher environmental burden.

Acknowledgments

We thank James Cust and the participants at the 2020 EAERE virtual conference for their very helpful comments. We also thank the editor and two anonymous referees for very useful recommendations in improving the manuscript. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. The authors declare no competing interests.

References

[1] Allcott, H., and Keniston, D. (2018). Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America. *The Review of Economic Studies*, 85 (2): 695-731.

[2] Aizer, A., Currie, J., Simon, P., and Vivier, P. (2018). Do Low Levels of Blood Lead Reduce Children's Future Test Scores? *American Economic Journal: Applied Economics*, 10(1): 307-41.

[3] Bartik, A. W., Currie, J., Greenstone, M., and Knittel, C. R. (2019). The local economic and welfare consequences of hydraulic fracturing. *American Economic Journal: Applied Economics*, 11(4): 105-55.

[4] Black, D., McKinnish, T., and Sanders, S. (2005). The economic impact of the coal boom and bust. *The Economic Journal*, 115(503): 449-476.

[5] Borge, L., Parmer, P., and Torvik, R. (2015). Local natural resource curse?. *Journal of Public Economics*, 131: 101-114.

[6] Brown, J. P. (2014). Production of natural gas from shale in local economies: A resource blessing or curse? *Economic Review*, Federal Reserve Bank of Kansas City, 1-29.

[7] Caselli, F., and Michaels, G. (2013). Do Oil Windfalls Improve Living Standards? Evidence from Brazil. *American Economic Journal: Applied Economics*, 5(1): 208-238.

[8] Currie, J., and Schmieder, J. F. (2009). Fetal exposures to toxic releases and infant health. *American Economic Review*, 99(2): 177-83.

[9] Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2): 678-709.

[10] De Silva, D.G., Hubbard,T.P., and Schiller, A.R. (2016). Entry and Exit Patterns of "Toxic" Firms, *American Journal of Agricultural Economics*, 98(3): 881-909.

[11] De Silva, D.G., McComb, R.P., and Schiller, A.R. (2020). Do localities benefit from natural resource extraction.*Energy Journal*, 41(5): 185-211.

[12] EPA (2011). Toxic Chemical Release Inventory Reporting Forms and Instructions. *Revised 2011 Version.* EPA 260-R-10-001.

[13] Feyrer, J., Mansur, E. T., and Sacerdote, B. (2017). Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review*, 107(4): 1313-34.

[14] Hamilton, J.T. (1995). Testing for Environmental Racism: Prejudice, Profits, Political Power? *Journal of Policy Analysis and Management*, 14: 107-132.

[15] Jacobsen, G. D. (2019). Who wins in an energy boom? Evidence from wage rates and housing. *Economic Inquiry*, 57(1): 9-32.

[16] James, A. (2015). US state fiscal policy and natural resources. *American Economic Journal: Economic Policy*, 7(3):238-57.

[17] James, A. (2017). Natural resources and education outcomes in the United States. *Resource and Energy Economics*, 49: 150-164.

[18] Keller, W., and Levinson A. (2002). Pollution Abatement Costs and Foreign Direct Investment Inflows to U.S. States. *Review of Economics and Statistics*, 84: 691-703.

[19] Marchand, J. (2012). Local labor market impacts of energy boom-bust-boom in Western Canada. *Journal of Urban Economics*, 71(1): 165-174.

[20] Marchand, J., and Weber, J. (2019). The Local Effects of the Texas Shale Boom on Schools, Students, and Teachers. *Working Paper No. 2017-12*.

[21] Michaels, G. (2010). The long term consequences of resource-based specialisation. *The Economic Journal*, 121(551): 31-57.

[22] Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, 105(12): 3633-59.

[23] U.S. EIA, (2015). Texas Profile State Profile and Energy Estimates (https://www.eia.gov/state/?sid=TX).

[24] U.S. EIA, (2019). Drilling Productivity Report (https://www.eia.gov/petroleum/drilling/pdf/dpr-full.pdf).

[25] Weber, J. G. (2012). The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming. *Energy Economics*, 34(5): 1580-158.

[26] Weber, J. G. (2014). A decade of natural gas development: the makings of a resource curse. *Resource and Energy Economics*, 37: 168-183.

[27] Wang, H. (2018). An Economic Impact Analysis of Oil and Natural Gas Development in the Permian Basin. *Working Paper*.

[28] Wooldridge, J.,M. (2010). Econometric Analysis of Cross Section and Panel Data. The MIT Press.

Table 1: School district level summary statistics

Variable	School District					
	Sample		MSA		Non-MSA	
	Oil	Non oil	Oil	Non oil	Oil	Non oil
Number of Schools	3.43	6.42	5.10	14.51	2.93	3.31
Population[a]	6.52	17.66	11.98	51.92	4.92	6.46
Number of Students[a]	1.28	3.46	2.53	9.95	0.92	1.16
University Ratio	0.12	0.14	0.12	0.16	0.12	0.13
Oil and gas revenue[b]	2.45	0.18	2.11	0.25	2.58	0.09
Oil and gas production[c]	8.10	0.88	8.31	1.32	8.16	0.42
Income[d]	4.51	4.87	4.88	5.51	4.40	4.30
Number of TRI type firms	10.93	20.94	19.94	62.32	8.15	9.20
Number of oil TRI firms	3.28	2.29	6.15	5.65	2.38	1.17
Number of non oil TRI firms	7.66	18.65	13.79	56.66	5.77	8.03
Number of TRI polluters	0.73	1.97	1.24	4.09	0.57	0.82
Number of other businesses	123.91	338.38	234.20	1,055.22	91.41	121.95
Nonwhite ratio	0.14	0.14	0.13	0.18	0.14	0.11
Unemployment rate	0.07	0.07	0.07	0.07	0.06	0.07
Number of rail roads	7.08	12.87	11.52	25.72	5.75	8.04
Number of roads	18.75	20.90	21.77	26.94	17.86	18.39
Area (in Km^2)	949.84	551.51	654.60	366.62	1,046.74	804.87
House rental ratio	0.24	0.25	0.24	0.26	0.24	0.24

[a] in 1,000, [b] in \$100 million, [c] in billions of KwH, and [d] in \$10,000.

Table 2: IV Regression results for TRI-type firms in MSA school districts

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.028***	0.055***	0.014			
	(0.009)	(0.018)	(0.009)			
Log of oil and gas production				0.022***	0.043***	0.011
				(0.007)	(0.014)	(0.007)
Log of income	0.270***	-0.082	0.384***	0.278***	-0.066	0.388***
	(0.081)	(0.120)	(0.086)	(0.080)	(0.118)	(0.085)
Log of population	0.184	-0.734***	0.337**	0.181	-0.740***	0.336**
	(0.151)	(0.177)	(0.149)	(0.151)	(0.176)	(0.149)
Log of number of other businesses	0.649***	1.040***	0.545***	0.651***	1.044***	0.546***
	(0.149)	(0.177)	(0.146)	(0.149)	(0.177)	(0.146)
Non white ratio	0.455***	0.434*	0.452***	0.463***	0.451*	0.456***
	(0.124)	(0.257)	(0.135)	(0.124)	(0.254)	(0.135)
Unemployment rate	-0.934*	-3.046***	-0.263	-0.919*	-3.017***	-0.256
	(0.498)	(0.970)	(0.564)	(0.492)	(0.962)	(0.561)
Log number of rail roads	-0.002	-0.013	0.003	-0.003	-0.014	0.003
	(0.012)	(0.024)	(0.013)	(0.012)	(0.024)	(0.013)
Log number of roads	0.041**	0.065**	0.019	0.043**	0.069**	0.020
	(0.018)	(0.033)	(0.020)	(0.018)	(0.033)	(0.020)
Log of land area	0.016	0.039	0.020	0.023	0.052	0.023
	(0.022)	(0.038)	(0.023)	(0.022)	(0.036)	(0.023)
House rental ratio	0.110	0.397	0.317	0.116	0.408	0.320
	(0.261)	(0.259)	(0.299)	(0.261)	(0.259)	(0.299)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,227	2,227	2,227	2,227	2,227
R^2	0.941	0.500	0.933	0.941	0.500	0.933
Weak identification F - test	26.43	26.43	26.43	31.09	31.09	31.09

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Censored IV regression results for TRI-type firms in MSA school districts

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.027***	0.049***	0.012**			
	(0.005)	(0.007)	(0.006)			
Log of oil and gas production				0.021***	0.038***	0.010**
				(0.004)	(0.006)	(0.004)
Log of income	0.262***	-0.003	0.321***	0.269***	0.008	0.325***
	(0.030)	(0.045)	(0.033)	(0.030)	(0.045)	(0.033)
Log of population	0.178***	-0.586***	0.361***	0.176***	-0.590***	0.360***
	(0.028)	(0.041)	(0.031)	(0.028)	(0.041)	(0.030)
Log of number of other businesses	0.629***	0.836***	0.503***	0.630***	0.837***	0.504***
	(0.027)	(0.039)	(0.029)	(0.027)	(0.039)	(0.029)
Non white ratio	0.440***	0.309***	0.397***	0.449***	0.323***	0.401***
	(0.066)	(0.095)	(0.072)	(0.066)	(0.095)	(0.072)
Unemployment rate	-0.904***	-1.925***	-0.255	-0.890***	-1.907***	-0.249
	(0.249)	(0.372)	(0.276)	(0.248)	(0.373)	(0.275)
Log number of rail roads	-0.002	-0.016*	-0.007	-0.002	-0.017**	-0.007
	(0.006)	(0.008)	(0.006)	(0.006)	(0.008)	(0.006)
Log number of roads	0.040***	0.053***	0.027***	0.042***	0.056***	0.028***
	(0.009)	(0.013)	(0.009)	(0.009)	(0.013)	(0.009)
Log of land area	0.016*	0.009	0.017*	0.022***	0.021*	0.020**
	(0.008)	(0.012)	(0.009)	(0.008)	(0.012)	(0.009)
House rental ratio	0.106	0.144	0.190***	0.112*	0.154	0.192***
	(0.066)	(0.099)	(0.073)	(0.066)	(0.100)	(0.073)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,227	2,227	2,227	2,227	2,227
Log likelihood	-593.851	-2,010.388	-834.101	-598.412	-2,024.923	-834.403
Hausman Test	0.016	0.006	0.413	0.024	0.007	0.417
Left-censored observations	0	252	80	0	252	80

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: IV Regression results for TRI-type firms in non-MSA school districts

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.022***	0.052***	0.004			
	(0.007)	(0.009)	(0.008)			
Log of oil and gas production				0.018***	0.043***	0.003
				(0.006)	(0.007)	(0.006)
Log of income	0.291***	0.164*	0.260***	0.300***	0.186*	0.262***
	(0.093)	(0.099)	(0.100)	(0.092)	(0.098)	(0.100)
Log of population	0.018	-0.293***	0.135	0.017	-0.295***	0.134
	(0.089)	(0.087)	(0.091)	(0.089)	(0.087)	(0.091)
Log of number of other businesses	0.774***	0.540***	0.665***	0.774***	0.540***	0.665***
	(0.093)	(0.091)	(0.095)	(0.093)	(0.092)	(0.095)
Non white ratio	-0.156	-0.466*	0.001	-0.121	-0.382	0.006
	(0.197)	(0.269)	(0.218)	(0.190)	(0.263)	(0.212)
Unemployment rate	-0.143	-0.427	-0.208	-0.154	-0.453	-0.210
	(0.414)	(0.519)	(0.448)	(0.413)	(0.518)	(0.448)
Log number of rail roads	0.000	-0.014	0.017	-0.001	-0.016	0.017
	(0.015)	(0.026)	(0.016)	(0.015)	(0.026)	(0.016)
Log number of roads	0.056**	0.021	0.061**	0.056***	0.022	0.061**
	(0.022)	(0.030)	(0.025)	(0.022)	(0.030)	(0.025)
Log of land area	-0.043**	0.074**	-0.028	-0.040*	0.083**	-0.027
	(0.021)	(0.033)	(0.023)	(0.021)	(0.033)	(0.023)
House rental ratio	0.471**	0.689***	0.665***	0.471**	0.689***	0.665***
	(0.210)	(0.227)	(0.235)	(0.210)	(0.229)	(0.234)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,629	2,629	2,629	2,629	2,629	2,629
R^2	0.864	0.308	0.847	0.864	0.302	0.847
Weak identification F - test	68.69	68.69	68.69	72.25	72.25	72.25

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Censored IV regression results for TRI-type firms in non-MSA school districts

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.020***	0.039***	0.002			
	(0.003)	(0.003)	(0.003)			
Log of oil and gas production				0.017***	0.032***	0.001
				(0.003)	(0.003)	(0.003)
Log of income	0.219***	0.176***	0.151***	0.228***	0.191***	0.151***
	(0.036)	(0.041)	(0.037)	(0.035)	(0.041)	(0.037)
Log of population	-0.056**	-0.241***	0.093***	-0.057**	-0.246***	0.093***
	(0.026)	(0.030)	(0.027)	(0.026)	(0.030)	(0.027)
Log of number of other businesses	0.815***	0.435***	0.687***	0.816***	0.437***	0.687***
	(0.026)	(0.030)	(0.027)	(0.026)	(0.030)	(0.027)
Non white ratio	-0.101	-0.192*	0.013	-0.068	-0.128	0.015
	(0.094)	(0.105)	(0.097)	(0.092)	(0.102)	(0.095)
Unemployment rate	-0.085	-0.217	-0.245	-0.095	-0.236	-0.247
	(0.203)	(0.231)	(0.213)	(0.203)	(0.232)	(0.213)
Log number of rail roads	0.001	-0.013*	0.003	0.000	-0.014*	0.003
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Log number of roads	0.042***	0.012	0.036***	0.043***	0.012	0.036***
	(0.009)	(0.011)	(0.010)	(0.009)	(0.011)	(0.010)
Log of land area	-0.044***	-0.008	-0.012	-0.041***	-0.000	-0.012
	(0.009)	(0.010)	(0.009)	(0.008)	(0.010)	(0.009)
House rental ratio	0.287***	0.345***	0.249***	0.287***	0.341***	0.250***
	(0.082)	(0.095)	(0.091)	(0.083)	(0.095)	(0.091)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,629	2,629	2,629	2,629	2,629	2,629
Log likelihood	-1,093.696	-2,263.696	-1,301.546	-1,094.315	-2,279.457	-1304.508
Hausman Test	0.016	0.006	0.413	0.024	0.007	0.417
Left-censored observations	15	553	195	15	553	195

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: IV Poisson regression results for TRI-type firms in school districts

Variable	Number of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: School districts in MSAs						
Log of oil and gas revenue	0.025***	0.088***	0.015***			
	(0.003)	(0.013)	(0.003)			
Log of oil and gas production				0.020***	0.073***	0.012***
				(0.003)	(0.012)	(0.003)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,227	2,227	2,227	2,227	2,227
Panel B: School districts in non-MSAs						
Log of oil and gas revenue	0.019***	0.280***	-0.006			
	(0.004)	(0.064)	(0.004)			
Log of oil and gas production				0.015***	0.231***	-0.005
				(0.003)	(0.055)	(0.003)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,629	2,629	2,629	2,629	2,629	2,629

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

All regressions include log of income, log of population, log of number of other businesses, non white ratio, unemployment rate, log number of rail roads, log number of roads, log of land area, and house rental ratio.

Table 7: Regression results for TRI polluters

Variable	Log of TRI-polluting firms				TRI-polluting firms		Proportion of TRI-polluting firms	
	IV		Censored-IV		IV-Poisson		Fractional Response IV Probit	
	MSA	Non-MSA	MSA	Non-MSA	MSA	Non-MSA	MSA	Non-MSA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log of oil and gas revenue	0.033*	0.007	0.016*	-0.002	0.085***	-0.013	0.045***	-0.002
	(0.019)	(0.007)	(0.008)	(0.004)	(0.028)	(0.014)	(0.015)	(0.010)
Log of income	0.422***	0.227***	0.342***	0.240***	1.630***	1.713***	0.295***	0.446***
	(0.126)	(0.066)	(0.053)	(0.039)	(0.204)	(0.241)	(0.112)	(0.129)
Log of population	0.027	0.008	0.020	0.037	-0.254*	0.281**	0.172*	0.267***
	(0.114)	(0.044)	(0.047)	(0.028)	(0.149)	(0.133)	(0.091)	(0.071)
Log of number of other businesses	0.079	0.103**	0.115**	0.082***	0.513***	0.455***	-0.319***	-0.209***
	(0.106)	(0.046)	(0.045)	(0.028)	(0.141)	(0.148)	(0.083)	(0.071)
Non white ratio	1.246***	0.151	1.002***	0.275***	2.122***	2.370***	1.091***	0.995***
	(0.330)	(0.214)	(0.100)	(0.071)	(0.221)	(0.549)	(0.192)	(0.301)
Unemployment rate	-1.084	-0.196	-0.829*	-0.317	4.391**	-1.602	-2.137**	-1.124
	(0.891)	(0.328)	(0.450)	(0.229)	(1.724)	(1.292)	(0.910)	(0.801)
Log number of rail roads	0.275***	0.144***	0.180***	0.072***	0.629***	0.471***	0.332***	0.279***
	(0.033)	(0.022)	(0.009)	(0.005)	(0.041)	(0.034)	(0.021)	(0.026)
Log number of roads	0.035	0.021	0.039***	0.042***	0.027	0.248***	-0.041	0.060
	(0.035)	(0.021)	(0.014)	(0.010)	(0.052)	(0.072)	(0.035)	(0.045)
Log of land area	0.015	-0.043**	0.014	-0.022***	-0.095*	-0.082**	-0.044*	-0.034
	(0.038)	(0.021)	(0.014)	(0.008)	(0.051)	(0.036)	(0.026)	(0.028)
House rental ratio	0.168	0.508***	-0.034	0.008	-0.310	-0.631	-0.390	-0.814***
	(0.271)	(0.163)	(0.123)	(0.092)	(0.361)	(0.484)	(0.239)	(0.263)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,629	2,227	2,629	2,227	2,629	2,227	2,629
R^2	0.479	0.350						
Log likelihood			-1,786.499	1,263.253			-7,469.00	-8,553.00
Hausman Test			0.785	0.934				
Weak identification F - test	26.43	68.69						
Wald test of exogeneity χ^2							6.17	1.40
Left-censored observations			1,253	1,998				

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix

Table A.1: Variable Descriptions

Variable	Description
Number of Schools	Number of schools in the school district
Population	School district level total population
Number of Students	Total number of students in the school district
University Ratio	Percentage of the population who holds a university degree in the school district
Number of TRI type firms	School district level number of TRI type firms
Number of oil TRI firms	School district level number of TRI type firms that belong to one of the NAICS codes listed in Table A.2
Number of non oil TRI firms	School district level number of TRI type firms that do not belong to one of the NAICS codes listed in Table A.2
Number of TRI polluters	School district level number of firms that reported a release above the EPA threshold (25,000 pounds) to the TRI in at least one year
Oil and gas revenue	Total market value of oil and gas production at school district level
Oil and gas production	Total production of oil and gas in kwh at school district level
Number of other businesses	Number of firms that do not belong to a NAICS code covered by the TRI Program in the school district
Median income	School district level median income in $
Non white ratio	School district level share of non white population
Unemployment rate	School district level unemployment rate
Number of roads	We use the U.S. Census Bureau's Census Feature Class Codes (CFCC) to identify roads. These road maps are provided by ESRI Data & Maps (2000) and we combine them with maps of school districts boundaries. We use all major highways to small roads that provide access to businesses, facilities, and rest areas along limited-access highways
Number of rail roads	As in roads we use the U.S. Census Bureau's Census Feature Class Codes (CFCC) and ESRI Data & Maps (2000) to identify rail roads. We use all major and minor rail tracks identified by ESRI Data & Maps
Area	School district level land area in square kilometers
House rental ratio	Number of rented houses divided by the total number of owned houses

Table A.2: Oil based TRI NAICS codes

TRI NAICS	Description
211111 :	Crude Petroleum and Natural Gas Extraction
211112 :	Natural Gas Liquid Extraction
212112 :	Bituminous Coal Underground Mining
211130 :	Natural Gas Extraction
324xxx :	Petroleum and Coal Products Manufacturing
424710 :	Petroleum Bulk Stations and Terminals

Table A.3: IV Regression results for TRI-type firms in MSA school districts: alternate specification

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.028**	0.044**	0.012			
	(0.011)	(0.021)	(0.012)			
Log of oil and gas production				0.022**	0.034**	0.010
				(0.008)	(0.016)	(0.009)
Log of number of other businesses	0.861***	0.380***	0.906***	0.861***	0.381***	0.906***
	(0.015)	(0.027)	(0.017)	(0.015)	(0.027)	(0.017)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,227	2,227	2,227	2,227	2,227
R^2	0.935	0.451	0.925	0.935	0.447	0.925
Weak identification F - test	20.60	20.60	20.60	24.27	24.27	24.27

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: IV Regression results for TRI-type firms in non-MSA school districts: alternate specification

Variable	Log of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Log of oil and gas revenue	0.022***	0.058***	0.004			
	(0.007)	(0.010)	(0.008)			
Log of oil and gas production				0.019***	0.048***	0.003
				(0.006)	(0.008)	(0.006)
Log of number of other businesses	0.809***	0.259***	0.838***	0.809***	0.260***	0.838***
	(0.017)	(0.024)	(0.021)	(0.017)	(0.024)	(0.021)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,629	2,629	2,629	2,629	2,629	2,629
R^2	0.857	0.220	0.839	0.857	0.208	0.839
Weak identification F - test	68.12	68.12	68.12	70.08	70.08	70.08

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.
Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: IV Poisson regression results for TRI-type firms in school districts: alternate specification

Variable	Number of TRI-type firms					
	All TRI	Oil TRI	Non oil TRI	All TRI	Oil TRI	Non oil TRI
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: School districts in MSAs						
Log of oil and gas revenue	0.020***	0.074***	0.014***			
	(0.004)	(0.021)	(0.004)			
Log of oil and gas production				0.016***	0.059***	0.011***
				(0.003)	(0.017)	(0.003)
Log of number of other businesses	0.935***	0.905***	0.938***	0.936***	0.911***	0.939***
	(0.015)	(0.083)	(0.011)	(0.015)	(0.086)	(0.011)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,227	2,227	2,227	2,227	2,227
Panel B: School districts in non-MSAs						
Log of oil and gas revenue	0.020***	0.321***	-0.004			
	(0.004)	(0.051)	(0.004)			
Log of oil and gas production				0.017***	0.294***	-0.004
				(0.004)	(0.051)	(0.003)
Log of number of other businesses	0.877***	0.542***	0.947***	0.877***	0.537***	0.947***
	(0.009)	(0.028)	(0.009)	(0.009)	(0.028)	(0.009)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,629	2,629	2,629	2,629	2,629	2,629

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Regression results for TRI polluters: alternate specification

Variable	Log of TRI-polluting firms				TRI-polluting firms		Proportion of TRI-polluting firms	
	IV		Censored-IV		IV-Poisson		Fractional Response IV Probit	
	MSA	Non-MSA	MSA	Non-MSA	MSA	Non-MSA	MSA	Non-MSA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log of oil and gas revenue	0.055**	0.005	0.056***	-0.000	0.053***	0.004	0.073***	-0.002
	(0.024)	(0.007)	(0.008)	(0.004)	(0.019)	(0.016)	(0.014)	(0.010)
Log of number of other businesses	0.320***	0.192***	0.312***	0.194***	0.738***	1.219***	0.050***	0.228***
	(0.028)	(0.023)	(0.010)	(0.008)	(0.023)	(0.046)	(0.017)	(0.026)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,227	2,629	2,227	2,629	2,227	2,629	2,227	2,629
R^2	0.227	0.219						
Log likelihood			-2072.469	-1430.831			-7,572.00	-8,679.00
Hausman Test			0.000	0.239				
Weak identification F - test	20.60	68.12						
Wald test of exogeneity χ^2							15.25	1.79
Left-censored observations			1,253	1,998				

Stock-Yogo weak ID test critical values: 10% maximal IV size 16.38.

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1