# Appendix A: Methodological Issues in Quantitative Environmental Inequality Research

As the environmental justice movement grew in impact, so did debate over how to best assess and quantify environmental inequality. Environmental inequality research traditionally focuses on the communities disproportionately exposed to the environmental harms of capitalist production (Bullard 1990). However, a new line of research focuses on the facilities disproportionately responsible for toxic emissions and the communities surrounding facility locations (Grant, Trautner, Downey, and Thiebaud 2010). In sum, there is a methodological debate regarding quantitative environmental justice analysis. Much of the debate centers on determining emission sources, measuring proximity to environmental risks and defining the unit of analysis (Liu 2001).

In this appendix, I argue that "bringing the polluters back in" to environmental inequality analysis (Grant et al. 2010) is critical to understand how environmental inequality is produced. There are both broad and narrow purposes for this appendix. The broad goal is to discuss methodological debates among quantitative environmental justice scholars. The narrower purpose is to make an argument for the methodological approach to environmental inequality used throughout this monograph.

## A.1. Determining Emission Point Sources

Environmental inequality research traditionally involves point source pollution. Point source pollution is pollution that can be attributed to a primary point source, such as an industrial production facility. Common point source pollution data comes from the Environmental Protection Agency Toxic Release Inventory (EPA TRI) and the Environmental Protection Agency Greenhouse Gas Reporting Program (EPA GHGRP). However, there are limitations associated with using these data sources, particularly as it pertains to the oil and gas extraction industry.

### A.1.1. Environmental Protection Agency Toxic Release Inventory (EPA TRI)

The EPA TRI was created by under the Emergency Planning and Community Right to Know Act signed into law on October 17, 1986 by President Reagan. To assist preparedness for chemical spills, the EPA TRI provides communities with information on chemicals used at some industrial production facilities. Facilities must meet three criteria to be required to report to the EPA TRI: (1) it must be within a specific industrial sector, (2) it must employ 10 or more full-time employees, and (3) it must handle 25,000 pounds of chemicals or more within the year. The oil and gas extraction industry is exempt from reporting.

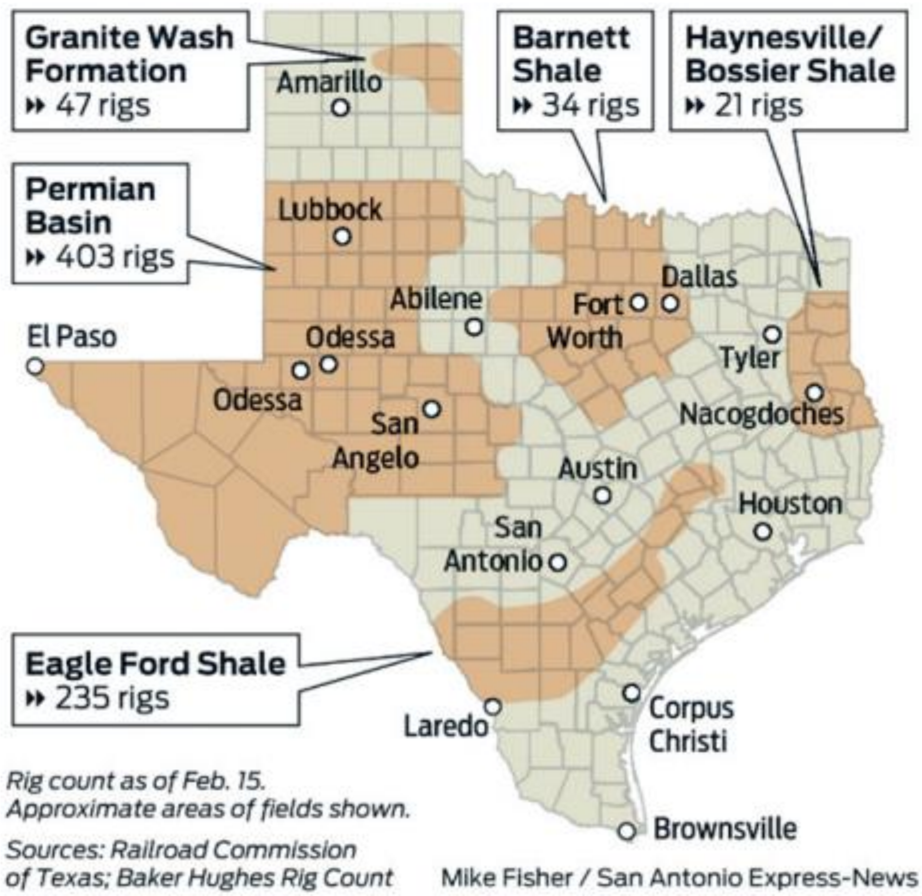
### A.1.2. Environmental Protection Agency Greenhouse Gas Reporting Program (EPA GHGRP)

Responding to the passage of the FY2008 Consolidated Appropriations Act, the EPA established the GHGRP. Starting in 2010, the EPA began to collect greenhouse gas emissions data from all facilities and automotive fleets that emit 25,000 metric tons of carbon dioxide equivalent or more per year. While the oil and gas extraction industry is required to report, information is not for specific facility points. Instead, information is for oil and gas extraction facility operations across entire shale plays, which is a very large geographic area. This provides little information about the specific place where pollution occurs.

### A.1.3. Problems with Environmental Protection Agency Data Sources

Information submitted to the EPA is limited, as specific industries are exempt from reporting to the EPA, the EPA fails to collect information on small producers, and information collected by the EPA on the oil and gas extraction industry is for large geographic areas, not specific points. The example below demonstrates limitations.

Figure A.1: Example Operator Oil and Gas Extraction Facility Span Across Texas Counties




The figure above shows Baker Hughes (a large oil and gas extraction company) oil rig counts for the various shale plays in Texas. Rigs are the mechanical devices used to extract oil and gas at a lease (i.e., it is a machine used at a oil and gas extraction facility). Emissions from these facilities would not be submitted to the EPA TRI because the oil and gas extraction industry is exempt. Baker Hughes would be required to submit a record to the EPA GHGRP for each of the separate shale formations if greenhouse gas emissions from rigs within the shale emit 25,000 tons or more of greenhouse gases. For example, if the 21 rigs located in the Haynesville/Bossier Shale were estimated to emit 23,000 tons of greenhouse gases, it would not be required to report. On the other hand, say the 235 rigs in the Eagle Ford Shale emitted 25,000 tons or more of greenhouse gasses, Baker Hughes would be required to submit a single report for all 235 rigs. Information is not broken down to the 235 facilities spread across the shale play. The specific locations of the facilities where emissions are occurring are not even collected. As such, using the EPA GHGRP, we cannot tell which communities are living near the oil and gas extraction facilities where pollution occurs.

## A.2. Measuring Community Environmental Risks

Two key issues among sociologists quantifying environmental risks are: (1) how to measure proxitity to risk, and (2) how to best estimate emission magnitude.

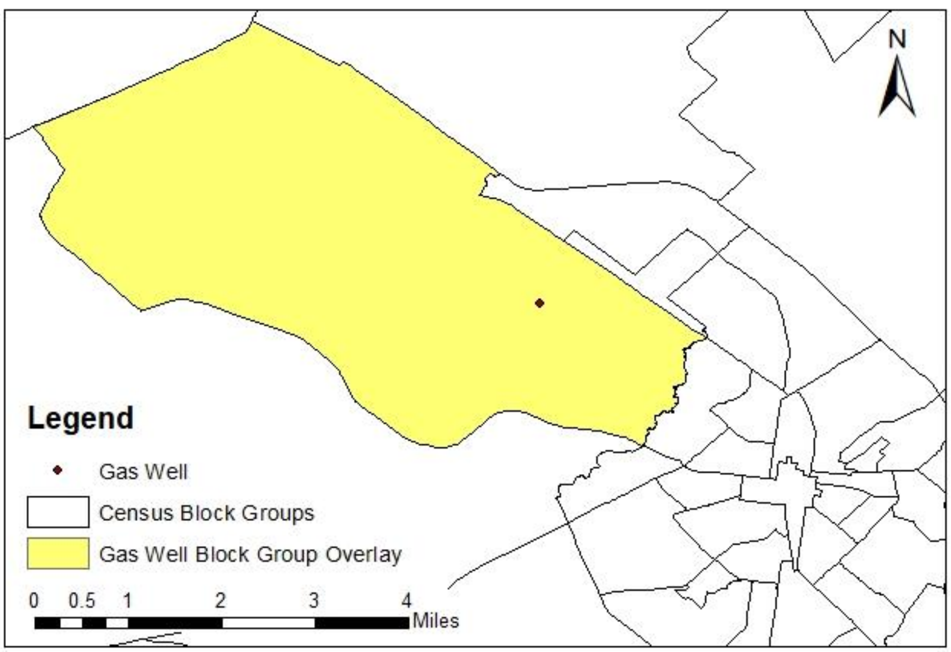
### A.2.1. Measuring Proximity

There are two key approaches to measuring proximity to environmental risks: the traditional unit hazard coincident approach, and the more modern, distance based approach.

#### *A.2.1.1. Unit Hazard Coincident Approaches*

Traditional environmental inequality analysis relied upon a unit hazard coincident approach. Classical studies examined whether or not a locally unwanted land use (LULU) was located within community boundaries (Bullard 1990; Mohai and Bryand 1992). In essence, this approach quantifies the characteristics of the immediate community in which the toxic facility is located, in comparison to those of the population not in the same immediate area. The spatial relationship between the facility and the community is determined by overlaying community boundaries and facility points to determine the community in which the facility is located.

Figure A.2: Example Unit Hazard Coincident Approach

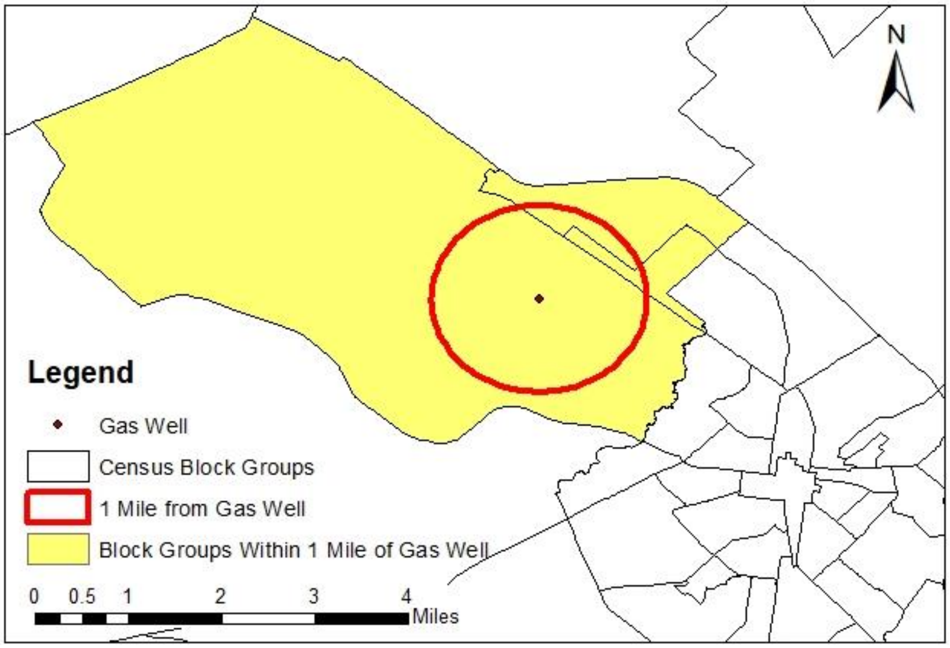



However, the unit hazard coincident approach is problematic, especially for facilities located near boundary lines. The community effected by toxic facilities often goes beyond the man-made boundary in which the facility is located.

#### *A.2.1.2. Distance Based Approaches*

More modern approaches to quantifying environmental inequality use geographic information technologies to determine the communities surrounding toxic facilities. The example below demonstrates a simple boundary intersection distance based approach. This approach quantifies the characteristics of communities surrounding facilities by determining the communities whose boundaries are within a specific distance from the facility and aggregating community data.

Figure A.3: Example Distance Based Approach




When comparing unit hazard coincidence and distance based approaches, Mohai and Saha (2007) find distance based approaches are robust and provide more precise estimates of communities exposed.

### A.2.2. Measuring Emissions

Sociologists use various approaches to estimate the magnitude of industrial facility toxic emissions. While emission models are most commonly used, another approach is to use direct metering devices.

#### *A.2.2.1. Emission Models*

EPA GHGRP greenhouse gas emissions are estimated using a variety of different models. Some estimate emissions by examining fuel-specific data. Others simply multiply a default emission and heat factor by the amount of fuel used to estimate carbon dioxide, methane, and nitrous oxide emissions.

#### *A.2.2.2. Metering Devices*

Some facilities employ continuous monitoring systems located on flare stacks, which monitor toxic emission concentration and flow rate. While this provides the most precise emission estimates, it is costlier to implement.

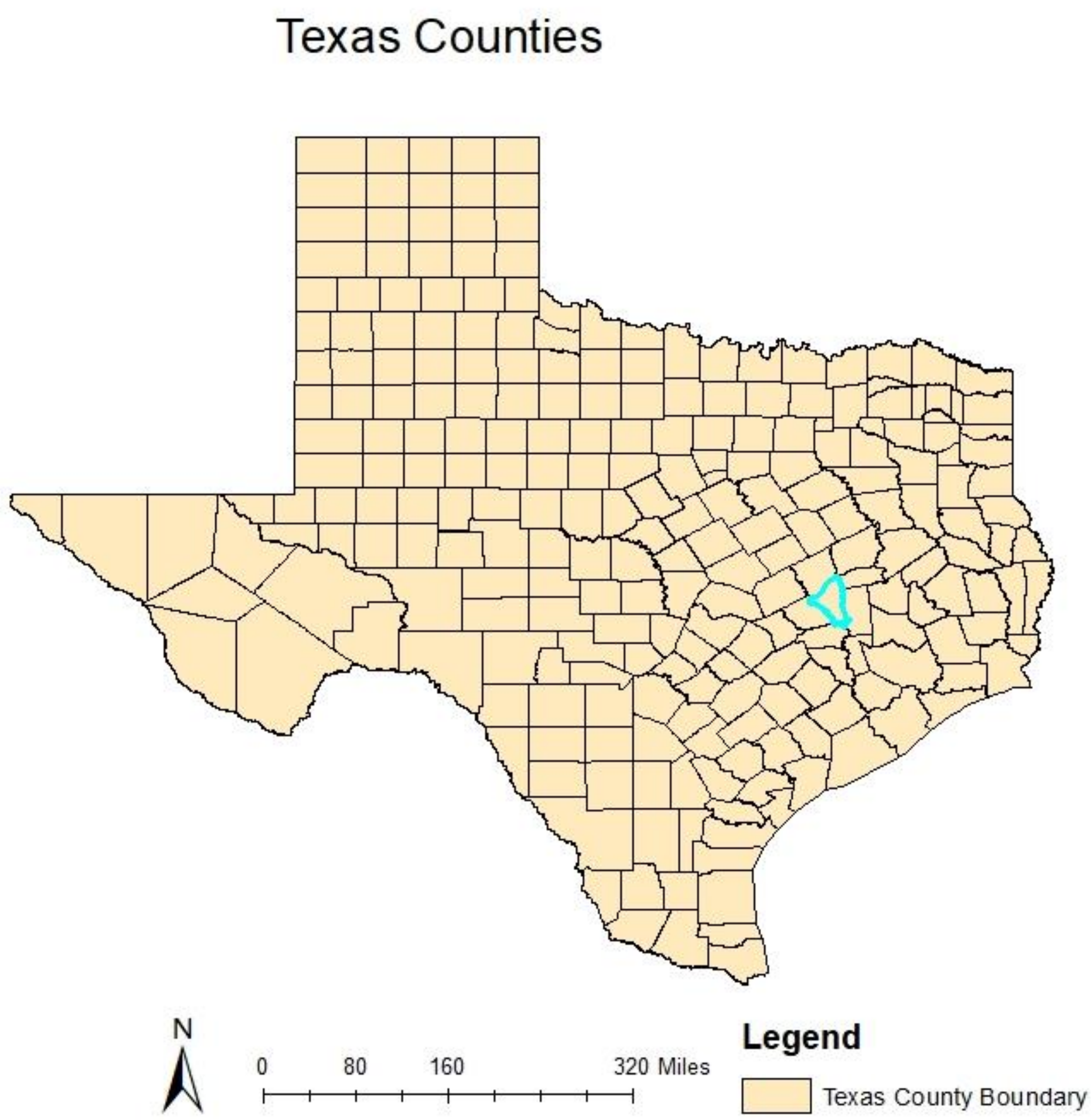
## A.3. Defining the Unit of Analysis in Community-Level Research

Much debate revolves around the ideal unit of analysis when conducting community-level environmental inequality research. Commonly used units of analysis (from largest geographic scale to smallest geographic scale) include county boundaries, zip code boundaries, Census tract boundaries, and Census block group boundaries.

### A.3.1. County Boundaries

Counties are very large geographic areas commonly used in environmental inequality analysis. Below is a map of all Texas counties, with Brazos County highlighted. Brazos County is at the upper edge of the Eagle Ford Shale.

Figure A.4: Texas County Boundaries in 2012




In 2012, there were 254 counties in Texas. The mean county size was 1062 square miles with a standard deviation of 658 square miles.

### A.3.2. Zip Code and Census Tract Boundaries

Since counties are so large, zip codes and Census tracts are more commonly used units of analysis in quantitative environmental inequality research. Zip codes and Census tracts are smaller geographic areas in comparison to counties, but still spread across a large geographic area. While

Census tracts are contained within counties, zip codes can spread across counties. Below is a map of all zip codes and Census tracts within in Brazos County, with a single Census tract in West Downtown Bryan highlighted.

Figure A.5: Texas Zip Code and Census Tract Boundaries in 2012

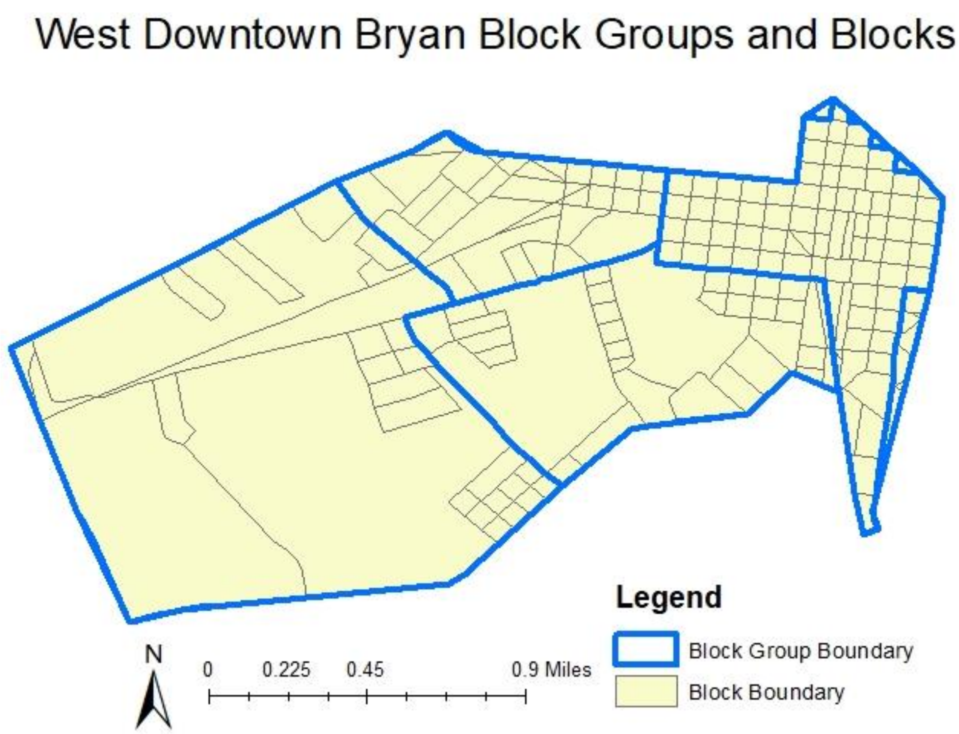

In 2012, there were 2024 zip code areas in Texas. The mean zip code area was 115 square miles with a standard deviation of 198 square miles.

In 2012, there were 5313 Census tracts in Texas. The mean tract area was 51 square miles with a standard deviation of 210 square miles.

### A.3.3. Census Block Group and Block Boundaries

Census block groups are much smaller than Census tracts and zip codes, and Census blocks are even smaller than block groups. Below is a map of the Census tract in West Downtown Bryan (which was highlighted in the previous map) and the Census blocks and block groups within the tract.

Figure A.5: Texas Block Group and Block Boundaries in 2012




As you can see, Census block groups and blocks are much smaller than tracts. The American Community Survey (ACS) provides detailed five-year community demographic and economic estimates at the block group level to the public. Block-level ACS community estimates are restricted and unreliable, as the ACS was sampled at a higher level of geography.

In 2012, there were 15,799 Census block groups in Texas. The mean block group area was 17 square miles, with a standard deviation of 103 square miles.

In 2012, there were 917,499 Census blocks in Texas. The mean block group area was .3 square miles with a standard deviation of 1.74 square miles.

## A.4. Conceptual Limitations of Community-Level Research

In sum, environmental inequality research primarily focuses on the community as the unit of analysis. Electronically metered, distance based, block group-level approaches provide the most precise estimates of communities most affected by facility toxic emissions. However, with the community as the primary focus of environmental inequality, little attention is paid to how variation between facilities relates to environmental inequality. Instead, much of environmental inequality research conceptualizes all toxic facilities as the same. However, rural sociology research on disproportionality demonstrates a significant variation among facility emission rates within an industry (Freudenberg 2006). It is critical to use the facility, which is the producer of environmental inequality, as the unit of analysis so that we can better understand why some facilities pollute at a higher rate than others.

## A.5. Bringing the Facility Into Environmental Inequality Research

### A.5.1. Prior Research

Much of environmental inequality research focuses on the community while overlooking toxic facilities themselves. However, because they are the producers of environmental inequality, Grant, Trautner, Downey and Thiebaud (2010) shift the focus to the industrial facility. Using novel fuzzy set qualitative comparative analysis and the Environmental Protection Agency's Risk-Screening Environmental Indicators of 2,053 chemical industry plants in 2002, they find that community and facility characteristics combine to produce disproportionate pollution emissions. Facilities in Census tracts that are more black, more Latinix, or have a greater percentage of the population employed in manufacturing, and facilities that have more employees or are branch plants are more likely to have highly risky emissions (Grant et.al 2010).

While Grant et. al (2010)'s research involving the industrial facility in environmental inequality analysis is a critical advancement, it is methodologically limited in four key ways. First, the research relies upon the Environmental Protection Agency's Risk-Screening Environmental Indicators (RSEI) model. The RSEI model evaluates the environmental risk of a facility using information about chemicals reported to the EPA TRI, together with factors about the chemical's toxicity and potential for human exposure. As described earlier, the EPA TRI is limited, as the only companies required to report are those that emit 25,000 tons of chemicals annually and have 10 or more full time employees. Since most organizations are small (Granovetter 1984), failing to include small organizations in their analysis limits the scope of their findings. Second, the research relies on a less precise method of measuring community proximity to risk. While Grant et. al (2010) use a unit hazard coincident approach to determining the characteristics of communities surrounding facilities, as described earlier, a distance based approach is much more precise. Third, the research relies on community information at the Census tract-level, which is a much larger geographic unit of analysis than the Census block group. Forth, Grant uses fuzzy set qualitative comparative analysis methods to determine which combinations of facility and organizational characteristics best explain facility emissions. While this is an innovative method to conduct exploratory research, it is not theory driven, it is data driven. As such, it is more difficult to differentiate genuine relationships from spurious ones.

### A.5.2. My Approach

I overcome previous limitations by taking a different approach. The approach taken in this project is inspired by Grant et. al (2010)'s work, but makes up for methodological limitations to studying the oil and gas extraction industry. While Grant et. al (2010) relies upon the EPA TRI and RSEI model to determine emission point sources and emission volumes, my research uses the Texas Railroad Commission well surface location coordinates for all producing oil and gas extraction facilities and electronically metered oil and gas production and disposition volumes. Additionally, I employ a distance based, block group method to determine the characteristics of communities most effected by oil and gas extraction facility venting and flaring volumes. Finally, I use a theory driven, quantitative regression model at the oil and gas extraction facility-level to determine how the characteristics of communities surrounding facilities are related to disproportionate emissions. While my approach is briefly described below, a detailed description of specific terms and models is in the subsequent appendix.

### *A.5.2.1. Determining Emission Point Sources- Texas Railroad Commission Well Surface Location Coordinates*

While federal agencies do not collect the longitude and latitude coordinates on all oil and gas extraction wells, state agencies make this information available to the public, though sometimes at a cost. Wellbore surface location coordinates were obtained for all oil and gas wells from the Texas Railroad Commission. Prior to drilling in Texas, all companies are required to report wellbore surface locations to the Texas Railroad Commission. These wellbore surface locations were projected onto a map using a North American Datum 1983 State Plane Texas Central FIPS 4203 Feet. While wellbore surface locations are available to the public though a GIS viewer, this information is not available to be downloaded and used to conduct comprehensive geographic and statistical analysis. As such, these coordinates were obtained through several Public Information Act requests to the Texas Railroad Commission. The Texas Railroad Commission required several thousands of dollars in processing fee payments (requests to waive the fee were denied), which were paid by the Texas A&M Sociology Department Graduate Research Award Committee.

### *A.5.2.2. Measuring Proximity- Distance Based Approach*

I employ a boundary intersection distance based approach to quantify the characteristics of individuals and households living in Census block groups within one mile of the wellbore surface location. First, I projected wellbore surface locations and 2012 Census Topologically Integrated Geographic Encoding and Referencing (TIGER)/Line Shapefiles onto a map using a North American Datum 1983 State Plane Texas Central FIPS 4203 Feet. Then I drew a one-mile buffer around each wellbore surface location. Next, I overlaid the one-mile buffer with Census block group polygons. Finally, I aggregated the block group data by the unique oil and gas extraction facility identifier to quantify the characteristics of communities living in block groups within one mile of the facility. I chose a one-mile buffer over other distances, as Moahi and Saha (2007) find it provides a more precise estimate and better fits the environmental inequality hypotheses.

### *A.5.2.3. Measuring Emissions- Electronically Metered Volumes*

I obtained venting and flaring gas volumes in thousand cubic feet (at base pressure of 14.65 pounds per square inch and base temperature of 60 degrees Fahrenheit) from the Texas Railroad Commission. Statewide Rules 27, 54 and 58(b) require all operators submit monthly production reports for each oil and gas extraction facility (i.e., lease) on or before the last working day of the following month. Venting and flaring volumes obtained from thermal mass flow meters are required to be submitted on these monthly reports. For each oil and gas extraction facility, I aggregated the recorded volumes for each month of 2012. Electronically metered volumes are better than operator estimates, as it minimizes human error.

### *A.5.2.4. Defining Unit of Analysis- Oil and Gas Extraction Facility*

While much research relies on the community as the unit of analysis, I rely upon the oil and gas extraction facility. The oil and gas extraction facility involves one or more wellbore surface locations located on the same lease of land. My research involves all producing oil and gas extraction facilities in 2012.

## A.6. References


Bullard, Robert D. 1990. *Dumping in Dixie: Race, Class, and Environmental Quality.* Boulder, CO: Westview Press.

Freudenberg, William. 2006. "Environmental Degradation, Disproportionality, and the Double Diversion: Reaching Out, Reaching Ahead, and Reaching Beyond." *Rural Sociology* 71:3-32.

Granovetter, Mark. 1984. "Small is Bountiful: Labor markets and Establishment Size." *American Sociological Review* 49:323–334.

Grant, Don, Mary Trautner, Liam Downey and Lisa Thiebaud. 2010. "Bringing the Polluters Back In: Environmental Inequality and the Organization of Chemical Production." *American Sociological Review* 75(4):479-504.

Liu, Feng. 2001. *Environmental Justice Analysis: Theories, Methods and Practice.* Boca Raton, FL: CRC Press LLC.

Bryant, Bunyan, and Paul Mohai. 1992. *Race and the Incidence of Environmental Hazards: A Time for Discourse.* Boulder, Colorado: Westview Press.