**External Economies of Shipping Energy Fuels**

Karen Clay, Carnegie Mellon University and NBER
Akshaya Jha, Carnegie Mellon University
Nicholas Muller, Carnegie Mellon University, Middlebury College, and NBER
Randy Walsh, University of Pittsburgh and NBER

Our proposed research focuses on the shipment of liquid energy fuels via railways and pipelines in the contiguous U.S. to examine market failures associated with the energy infrastructure. In particular, we propose an examination of how crude-by-rail shipments and energy fuel pipelines have affected residential housing prices in markets proximal to energy infrastructure. In addition, we will estimate the monetary damage associated with air pollution emissions from crude-by-rail and pipeline shipments by combining locomotive and pipeline emissions rates with an extant integrated assessment model. These two lines of research address the ***Energy Infrastructure and Policy*** topic area. The bulk of this proposal focuses on crude-by-rail shipments from the Bakken Formation as a proof of concept. However, the methods proposed herein apply to all transportation routes.

The proposed research will use existing, detailed spatial datasets that document existing rail lines and pipeline routes. For example, we will use rail routes that are used to ship crude from the Bakken formation to refineries. (Figure 1 shows the precise route take from Chicago to Newport News, VA along with a measure of population density near the rail line.) In addition, we will use similar datasets to map existing rail routes that are not used to ship crude. We will exploit the DataQuick Information Systems database which consists of housing assessor and transaction data over the (2000 to 2013) time period. A comparison of the implicit price of rail lines not used for crude transport with that of routes known to ship crude will facilitate an estimate of the causal effect of crude by rail shipments on housing prices. Specifically, we will use a difference in difference approach to identify the effect of the externality on homeowners. This tack exploits the fact that crude by rail has risen rapidly over time and has been more concentrated on some rail routes than others.

The same detailed spatial data will facilitate an analysis of the health impacts from emissions produced by line haul locomotives and pipelines as crude is transported from extraction sites (the Bakken e.g.) to refineries. This line of inquiry will link emissions, by county of emission, to changes in air quality, human exposure, and monetary damage through the AP2 integrated assessment model (Muller, 2011; 2014).

## Background

The Bakken field produced modest amounts of oil beginning in the 1950s, but large-scale production did not begin until 2007. As a result of rapid increases in production (Figure 2), North Dakota is currently the third largest producer of oil after Texas and the federal offshore region (PADD3).[1] The EIA notes "Since then [2007], advances in drilling methods and technology, a better understanding of the geology of the Bakken, higher crude oil prices, and the formation's large size and number of wells all have contributed to higher production and to the potential for continued future growth."[2]

---

[1] PADD stands for Petroleum Administration Defense District. There are five PADDs and the Bakken region is located in PADD2. http://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbblpd_m.htm

[2] http://www.eia.gov/todayinenergy/detail.cfm?id=17391

We focus here on the Bakken both because of the supply shock and because it represented 70 percent of crude by rail movements in 2014.[3] According to the EIA, "Total CBR movements in the United States and between the United States and Canada were more than 1 million barrels per day (bbl/d) in 2014, up from 55,000 bbl/d in 2010." [4] In the early years of the Bakken field, the shock appears to have been addressed through some combination of excess capacity in pipelines and rail and crowd-out or delay of lower valued goods.[5] As oil production grew, constraints posed by infrastructure caused a significant divergence in prices between Bakken and West Texas Intermediate during some periods, particularly in early 2012 (Figure 2).

**Proposed Approach**

To evaluate market failures associated with current energy infrastructure we propose to pursue two lines of inquiry. The first will estimate hedonic property models to test whether energy infrastructure used for the transport of crude oil has a measureable impact of residential housing prices. The second will use integrated assessment modelling to estimate the impacts that emissions from line haul locomotives and pipelines have on exposed populations.

First, we will estimate hedonic property models to explore whether crude-by-rail shipments and associated risks to physical property, the environment, and human health are capitalized into residential housing prices. This approach will rely on the same, detailed spatial data of rail lines discussed above. However, since rail lines that are not used for crude shipments also likely produce externalities reflected in housing markets, we will need to include all rail lines in the model to attempt to isolate the effect of energy fuel shipments on prices. These data exist, and a map of all rail lines is shown in figure 3. In addition, this aspect of our research will employ the DataQuick Information Systems dataset which PI Walsh has obtained access to. These data span the (2000 to 2013) time period and they contain housing Characteristics, assessed values and location for all properties in 1542 Counties. For a subset of these Counties (620) we also have sales data. We will also include spatially-detailed measures of pipelines to test whether pipelines affect property values.

To estimate the air pollution externalities, our second approach will use the AP2 integrated assessment model and route information to evaluate the relative impacts of moving oil from the Bakken region to refineries via railroad or pipelines. Because the model is calibrated to the county level, it is able to detect differences in exposure risk due to the specific path of rail networks relative to pipeline networks. This will be a critical factor in determining the relative risk of the two energy transportation modes. The AP2 model estimates the impacts of air pollution on human health, crops, timber, and man-made materials. Further, AP2 can be used to estimate the effects of emissions from evaporation due to oil spills on these outcomes for a range of spill-related scenarios.

**Data**

Our analysis requires a number of different types of data.

---

[3] http://www.eia.gov/todayinenergy/detail.cfm?id=20592#tabs_Slider-5

[4] http://www.eia.gov/todayinenergy/detail.cfm?id=20592#tabs_Slider-5

[5] http://www.nytimes.com/2014/10/09/us/as-trains-move-oil-bonanza-delays-mount-for-other-goods-and-passengers.html?_r=0http://www.nytimes.com/2014/08/26/us/grain-piles-up-waiting-for-a-ride-as-trains-move-north-dakota-oil.html

*Crude Production Data*
The EIA tracks monthly and annual production for the Bakken region. Production data by county are also available from the state of North Dakota. We are not yet sure which one we will use, since it is not clear how much value added will be gained from moving to the county level. Further, in the immediate vicinity of extraction, oil may not be shipped out via pipeline or rail from the county in which it is produced.

*Rail and Pipeline Data*
The investigators have already gathered data on the location of rail networks used to transport crude oil across the United States. Figures 1 and 3 show a map of rail networks. PI Walsh has developed a database including rail lines, at grade crossings (which require speed reductions by trains), and populations located proximal to rail lines which facilitates a detailed exposures analysis. Similarly detailed spatial databases on main pipelines are publicly available.

*Data on Routes Crude Oil Takes to Refineries*
We plan to identify the routes that Bakken crude takes to refineries with data provided by Genscape (a commercial vendor) which monitors movements of crude oil out of the Bakken region. Genscape publishes daily aggregate movements of oil out of Bakken by rail and pipeline on Bloomberg. Carnegie Mellon has a Bloomberg terminal, and we have verified that we can access these data. Genscape has additional subscription products that provide significantly finer levels of detail. The most important one for our purposes is the Petrorail report (http://www.genscape.com/solutions/oil/petrorail-report#tabs-Overview_panel). The Petrorail website states: "Genscape monitors 93% of all crude-by-rail loadings in the Bakken, as well as 68% of Bakken-related pipeline capacity." Genscape has sent us a list of the pipelines and rail terminals that they monitor out of Bakken. From our preliminary discussion with Genscape, we know that oil produced from the Bakken is loaded on a rail terminal in North Dakota and it is shipped to Chicago on line owned by either BNSF or Canadian Pacific. Then, in Chicago, the crude is offloaded to either CSX or Norfolk Southern for the trip to refineries in Philadelphia and New Jersey.

*Emission Rates*
These data are from the U.S. Environmental Protection Agency's National Emission Inventory (NEI). The USEPA publishes the NEI every three years and the most recent year available is 2011. The authors will use the 2014 inventory once it becomes available. Pollutants encompassed include: volatile organic compounds (VOCs), sulfur dioxide ($SO_2$), nitrogen oxides ($NO_x$), fine particulate matter ($PM_{2.5}$), and ammonia ($NH_3$). The NEI data for locomotives is provided at the county level of detail. For locomotives, emissions encompassed are due to combustion of fuel in transit. These emissions are a function of locomotive age, work load, and speed. We are exploring the feasibility of incorporating information on locomotive age structure (for example: see http://www.stb.dot.gov/econdata.nsf/f039526076cc0f8e8525660b006870c9?OpenView).

*Housing Data*
We currently have assessor data (Housing Characteristics, Assessed Values and Lat/Long) for all properties in 1542 Counties. For a subset of these Counties (620) we also have Sales Data from 2000 through roughly 2013. The data are from DataQuick Information Systems and are available due to PI Walsh's association with the Social Science Research Center at Duke University. A sample of coverage of the housing data is shown in figure 4.

***Deliverables***

We plan to deliver two types of calculations. First, we will report the results from the regression analyses that estimates the impact of crude-by-rail shipments on residential real estate prices. Second, we will report the marginal social costs ($/bbl) due to shipments of crude-by-rail derived using the AP2 integrated assessment model. A comparison of the results from the two approaches will provide an interesting useful robustness check on both methods. In addition, armed with the various databases collected as part of this project, we can report populations at risk, and the value of property at risk due to accidents and fire.

Figure 1:

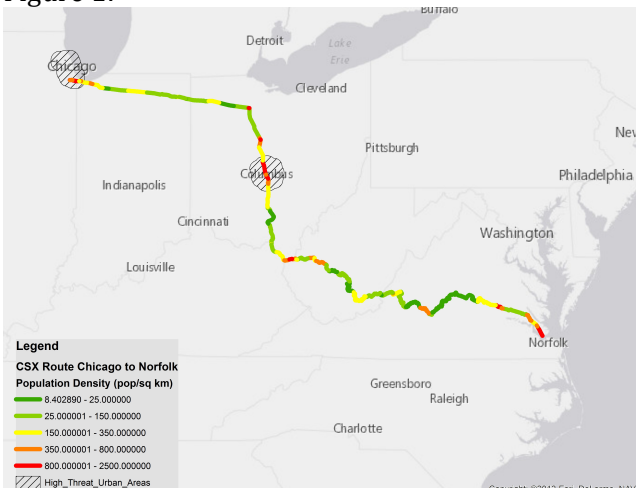
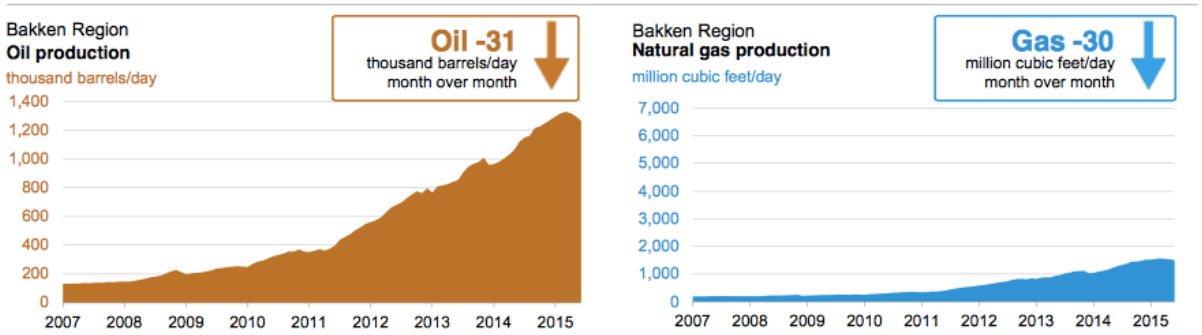



Figure 2: Oil and Gas Production in the Bakken Region.




Source: http://www.eia.gov/petroleum/drilling/pdf/bakken.pdf

Figure 3: Railroad Network

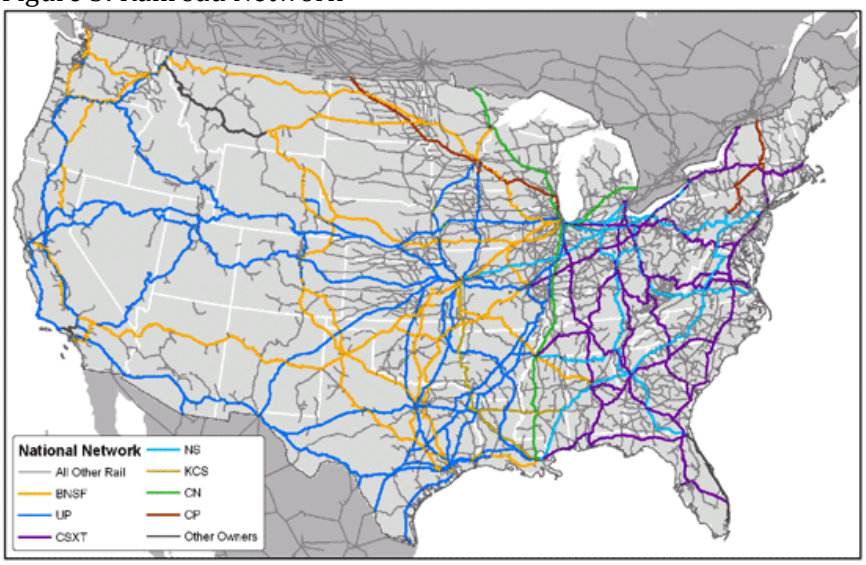
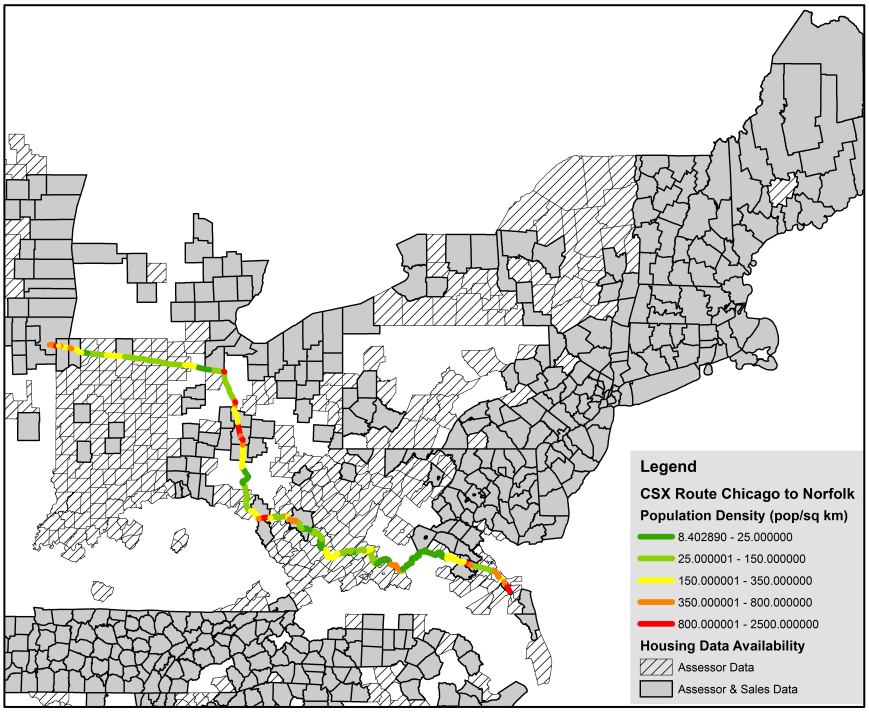



Figure 4: Housing Data Coverage in the Northeastern U.S.




References:
Muller, N.Z. 2014. "Boosting GDP growth by accounting for the environment: Including air pollution and greenhouse gas damages increases estimated U.S. growth." N.Z. Muller. Science. August 22nd, 2014, Vol. 345 no. 6199 pp. 873-874: DOI:10.1126/science.1253506.
Muller, N.Z. 2011 "Linking Policy to Statistical Uncertainty in Air Pollution Damages." N.Z.Muller. The B.E. Press Journal of Economic Analysis & Policy. Vol. 11 : Iss. 1 (Contributions), Article 32.

Disclosure statement: Karen Clay, Akshaya Jha, Nicholas Muller, and Randy Walsh have no financial or other interests that might bear on the proposed work and have no ties to the energy industry.