Shipments of Oil by Rail: Economic Implications for Safety and Safety-Related Investments

Eric Thompson, Ph.D.
Professor of Economics
Department of Economics
University of Nebraska-Lincoln

Aemal Khattak, Ph.D.
Professor
Department of Civil Engineering
University of Nebraska-Lincoln

Amirfarrokh Iranitalab
Ph.D. Candidate
Department of Civil Engineering
University of Nebraska-Lincoln

Collin Greer
Graduate Student
Department of Economics
University of Nebraska-Lincoln

A Report on Research Sponsored by

University Transportation Center for Railway Safety (UTCRS)

University of Nebraska-Lincoln

July 2018

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 26-1121-0018-013	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Shipments of Oil by Rail: Economic Implications for Safety and Safety-Related Investments		**5. Report Date** July 2018 **6. Performing Organization Code**
7. Author(s) Eric Thompson, Aemal Khattak, Amirfarrokh Iranitalab, and Collin Greer		**8. Performing Organization Report No.** 26-1121-0018-013
9. Performing Organization Name and Address Nebraska Transportation Center University of Nebraska-Lincoln Prem S. Paul Research Center at Whittier School 2200 Vine Street Lincoln, NE 68583-0853		**10. Work Unit No.** **11. Contract or Grant No.** DTRT13-G-UTC59
12. Sponsoring Agency Name and Address University Transportation Center for Railway Safety The University of Texas Rio Grande Valley 1201 W University Drive Edinburg, Texas 78539		**13. Type of Report and Period Covered** Final Report (October 2016 – June 2018) **14. Sponsoring Agency Code**
15. Supplementary Notes		

16. Abstract

Fracking technology has allowed for significant expansion of oil production in regions with limited oil pipeline capacity, such as the Bakken formation in North Dakota. These regions must compensate for the increase in oil production by expanding oil by rail shipments, specifically utilizing tanker cars. However, oil by rail shipments to the Eastern and Southern United States had declined in recent years. The present study utilizes U.S. Energy Information Administration forecasts of Bakken oil production and oil by rail shipments through the year 2040. A linear programming model was developed to estimate the volume of state-to-state oil by rail shipments by analyzing release incident rates. In addition, the present study assesses the growth of rail release incident costs and explores how changes in release incident costs impact the economic feasibility of rail-related safety investments. The data implied a positive correlation between oil production and oil by rail shipments. Annual release incidents costs for hauling oil by rail will rise significantly in the future due to projected increases in oil production and rising release incident rates.

17. Key Words		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified.	**20. Security Classif. (of this page)** Unclassified.	**21. No. of Pages** 71	**22. Price**

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

List of Figures and Tables

List of Tables

Abstract

Fracking technology has allowed for significant expansion of oil production in regions with limited oil pipeline capacity, such as the Bakken formation in North Dakota. These regions must compensate for the increase in oil production by expanding oil by rail shipments, specifically utilizing tanker cars. However, oil by rail shipments to the Eastern and Southern United States had declined in recent years. The present study utilizes U.S. Energy Information Administration forecasts of Bakken oil production and oil by rail shipments through the year 2040. A linear programming model was developed to estimate the volume of state-to-state oil by rail shipments by analyzing release incident rates. In addition, the present study assesses the growth of rail release incident costs and explores how changes in release incident costs impact the economic feasibility of rail-related safety investments. The data implied a positive correlation between oil production and oil by rail shipments. Annual release incidents costs for hauling oil by rail will rise significantly in the future due to projected increases in oil production and rising release incident rates.

Executive Summary

Fracking technology has allowed for a significant expansion of oil production in regions with limited oil pipeline capacity, such as the Bakken formation in North Dakota. While oil by rail shipments increased significantly in the Bakken region as oil production expanded, rail shipments of oil to the Eastern and Southern United States have declined. The potential Keystone XL Pipeline project would further erode the need for oil by rail shipments to the Southern and Eastern United States. However, oil by rail shipments to the Western United States are expected to continue into the future.

This study by the University of Nebraska-Lincoln Bureau of Business Research and Department of Civil Engineering utilizes U.S. Energy Information Administration forecasts of Bakken oil production and oil by rail shipments from the Bakken region through the year 2040. A 26 percent increase in Bakken oil production is forecasted for the 2015 to 2040 period.

A linear programming model was developed by research team members from the University of Nebraska-Lincoln Department of Civil Engineering to estimate the volume of state-to-state oil by rail shipments throughout the United States. Based on production forecasts and forecasts for oil refining capacity, oil by rail shipments to the Western United States are expected to grow by 15 percent from 2015 to 2040.

Data on release incidents and estimated train-miles of oil by rail shipments from state to state were used to estimate a release incident rate for oil by rail shipments. A release incident is a qualified[1]

[1] According to the Pipeline and Hazardous Materials Safety Administration (PHMSA), the reported incidents are either reported through telephone within 12 hours after occurrence for more severe incidents or through a written notice within 30 days for other incidents. The incidents that require telephonic notice include cases where: 1) as a direct result of a hazardous material a person is killed or injured requiring admittance to a hospital, the general public is evacuated for one hour or more, a major transportation artery or facility is closed or shut down for one hour or more, or the operational flight pattern or routine of an aircraft is altered; 2) fire, breakage, spillage, or suspected radioactive contamination occurs involving a radioactive material; 3) fire, breakage, spillage, or suspected contamination occurs involving an infectious substance other than a regulated medical waste; 4) a release of a marine pollutant occurs in a quantity exceeding 119 gallons for a liquid or 882 pounds for a solid; 5) a situation exists of such a nature that, in the judgment of the person in possession of the hazmat, it should be reported; or 6) during transportation by aircraft, a fire, violent rupture, explosion or dangerous evolution of heat occurs as a direct result of a battery or battery-powered device. Other incidents include: 1) an unintentional release of a hazmat during transportation including loading, unloading and temporary storage related to transportation; 2) a hazardous waste is released; 3) an undeclared shipment with no release is discovered; or 4) a specification cargo tank 1,000 gallons or

occurrence causing the release of hazardous materials from a train. On average, there were 8.6 release incidents per 1 million train-miles traveled during the 2007 to 2015 period. By 2040, these rates are forecast to rise to 14.6 release incidents per 1 million-train-miles traveled.

These release incident rates are applied to 2015 and 2016/17 oil by rail shipments from the Bakken region and projected 2040 oil by rail shipments from the Bakken region. The total annual external costs of release incidents during Bakken region oil by rail shipments was $25.5 million during 2015. The cost fell significantly to $8.9 million per year during 2016/17 as new pipeline capacity and market conditions reduced Bakken oil by rail shipments to the East Coast and Gulf Coast, although shipments to the West Coast did not decline. Costs would fall to an estimated $4.2 million per year if the completion of the Keystone XL Pipeline eliminated the need for Bakken oil by rail shipments to the East Coast and Gulf Coast. Annual release incidents costs for hauling Bakken oil by rail will rise to an estimated $8.9 million (in real dollars) by 2040, due to projected increases in Bakken oil production and rising release incident rates.

Research results also were used to develop benefit costs estimates for 4 potential viaduct projects in North Dakota and Minnesota, along routes where oil by rail shipments take place. The study found that the benefits of these oil by rail investments were less than the costs of construction. The additional external costs of oil by rail shipments was not found to significantly increase benefit costs ratios in the 4 potential viaduct projects.

---

greater containing any hazmat that received structural damage to the lading retention system or damage that requires repair to a system intended to protect the lading retention system, and did not have a release.

Chapter 1 Introduction

Fracking technology has allowed for a significant expansion of oil production in regions with limited oil pipeline capacity, such as the Bakken formation in North Dakota. For example, crude oil carloads rose from less than 10,000 in the United States to nearly 500,000 in 2014 (American Association of Railroads, 2016). The rapid growth in shipments of this flammable material has implications for safety in terms of the number and severity of rail release incidents, both at and away from at-grade rail-highway intersections. There is a need to assess the resulting growth of rail release incident costs, including under alternative scenarios for the development of oil pipelines in the region. In addition, there is a need to analyze how changes in release incident costs impact the economic feasibility of rail-related safety investments.

The next section of this report examines the methodological approach to the study, including a discussion of the four project tasks, using the example of the Bakken formation. Section 3 is a literature review regarding rail release incidents from hauling hazardous materials in general and oil in particular. Sections 4 through 7 reports on the results of the four research tasks.

Chapter 2 Methods

This research project is composed of 4 tasks: 1) estimating current and future oil by rail shipments from the Bakken region, 2) estimating oil by rail release incidents rates and the frequency, severity and non-injury costs of release incidents, 3) estimating the economic costs, including the costs of fatalities, injuries and non-injury costs, of rail release incidents due to oil by rail shipping from the Bakken region, and 4) modeling how changes in oil by rail shipments impacts the benefit cost ratio of potential highway safety projects, using the example of four viaduct projects. This section provides more detail about the approach to each task.

Estimates of the volume of Bakken oil by rail shipments are based in part on production estimates for regions of the country through 2040 from the *Annual Energy Outlook* of the Energy Information Administration. The forecasts allow for modeling through the year 2040. Projections of oil refining capacity in regions of the country will also influence oil by rail shipments from the Bakken. A linear programming model is developed to estimate the relationship between oil production, oil refining, and state-to-state shipments of oil by rail. Task 1 also considers how the volume of oil by rail shipments would be impacted by completion and opening of the Keystone XL Pipeline project.

Task 2 is to estimate release incident rates for oil by rail shipping as well as the severity and non-injury costs of oil by rail release incidents. Shipment volumes between states from the linear programming model are combined with the distance between states to estimate the total barrel-miles of oil by rail shipment each year, and to predict the total number of release incidents. Estimated release incidents are divided by estimated volumes to yield release incident rates. Statistical models can estimate the severity of release incidents as well as the non-injury related costs of release incidents such as damage to rail property, damage to other vehicles, cleanup costs, and other costs.

The third task expands on Task 2, estimating the economic cost of oil by rail shipments. Fatality and non-fatal injury release incident rates are used to estimate the human cost of release incidents. The incidence of fatalities and injuries are combined with data on the statistical value of a human life and the quality of life costs of non-fatal release incidents. Results are summed to create an estimate for the cost per release incident. To determine the annual cost of release incidents, per release incident costs are combined with information on the annual number of release incidents resulting from oil by rail shipments from the Bakken region. Annual economic costs from release incidents are compared for 2015 and 2040.

Finally, information gathered in Task 3 on the cost of release incidents is used to show how an increase in oil by rail shipments influences the economic feasibility of rail safety investments, specifically investments in viaducts to replace at-grade rail crossings. The research team selects four potential viaduct projects and conducts a benefit cost analysis of each project based on the number of train crossing per day, the number of lanes and annual average daily traffic (AADT) at the crossed highway. Benefits include avoided release incidents at the at-grade crossing and avoided wait times as trains cross the highway multiple times per day. The economic costs of release incidents and wait times are estimated based on Federal Railroad Administration and U.S. Department of Transportation data.

Chapter 3 Literature Review

This review examines existing literature on the safety-related aspects of shipping hazardous materials by rail, including crude oil. The reviewed literature is divided into two sub-sections: hazardous materials and crude oil.

3.1 Hazardous Materials

The majority of literature pertaining to shipment of hazardous materials is focused on different approaches to risk analysis. Because large hazardous materials accidents are relatively rare, railroads cannot effectively manage safety improvement efforts solely in response to the causes of specific accidents and risk-based approaches are needed to better understand predictive factors for conditions that can cause a release (Barkan et al, 2013). In rail transport of hazardous materials, risk is generally defined as a multiplication of derailment rate of a hazardous materials tank car, traffic exposure, conditional probability of release of a derailed hazmat car and the consequence of a tank car release (Liu et al, 2013a).

In a 1983 study, Nayak and co-authors (Nayak, et al., 1983) presented a number of methods for quantifying the probability of existence and severity of impacts of hazardous materials in accidents in rail transportation. This included development of measures for accident rates based on track class, severity of an accident based on accident speed, and the probability and mean amount of release based on accident speed. Finally, a method to estimate the impacts on people and property of the release of hazardous material was proposed in this study.

In a descriptive study (Ogero, et al., 2006), 1,932 accidents reported from the beginning of the 20th century to July 2004 around the world that involved the transportation of hazardous materials by road and rail were investigated. More than half of the accidents happened on roads (63%). The most frequent type of accidents were releases (78%), followed by fires (28%), explosions (14%), and gas clouds (6%). The most frequent initiating event of accidents turned out to be an impact or collision between vehicles, derailment of trains and trucks crashing. More than half of the accidents did not cause any fatalities. Among fatal incidents the number of deaths was frequently between 1 and 10. It seemed, given the occurrence of an incident, the consequences will probably be slightly more severe if it occurs on a railway rather than on a road. Moreover, if there is an evacuation (which is rather unusual) the number of people involved is usually between 101 and 1000 (29%).

Zografos and Androutsopoulos (2008) presented a decision support system for assessing alternative distribution routes in terms of travel time, risk and evacuation implications, while coordinating the emergency response deployment decisions with the hazardous materials routes. The proposed system was supposed to work towards alternative hazardous materials distribution routes recognition, in terms of cost and risk minimization, specification of the hazardous materials-related locations for first-response emergency service, and determination of evacuation paths from the impacted area to shelters and estimation of the evacuation time. The proposed system was implemented and evaluated in the industrialized area of Thriasion Pedion of Attica, Greece.

Rail routes are determined by economic concerns such as route length and travel time, while rail shipments of hazardous materials expose the population near the routes to the possibility of an accident resulting in a spill. In a 2007 study, Glickman et al. considered an alternate routing strategy that takes release incident risks into account from a macroscopic perspective. Rail transport risk was quantified by estimating the expected population that resides within a given radius of the location of a train accident and then using a weighted combination of cost and risk and generated alternate routes. The results showed that in some cases the alternate routes achieve significantly lower risk measures than the practical routes at a small incremental cost. The authors concluded that the issue of rail rerouting deserves more attention due to the situations in which risk can be reduced without lengthening the route substantially.

In another study (Liu, et al., 2013b) derailments, as the most common type of freight-train accidents in the United States were analyzed. Zero-truncated negative binomial regression model was developed to estimate the conditional mean of train derailment severity, in terms of number of derailed train-cars. Considering that the mean is not the only measure describing data distribution, a quantile regression model was also developed to estimate derailment severity at different quantiles. Combining the two models resulted in a better understanding of train derailment severity distribution.

Liu et al. (2013a) developed an integrated risk reduction framework, considering the cost-effectiveness of an individual risk reduction strategy, their interactive effects and optimal integration. They formulated hazmat risk management as a multi-attribute decision analysis problem and a negative binomial regression model was developed to estimate accident-cause-specific railcar derailment probability. Then they developed a Pareto-optimality approach to determine the lowest risk that can be achieved at a specific level of investment. Understanding the risk-and-cost relationship led to development of a decision model to determine the optimal investment. To illustrate the methodology, they analyzed two types of risk reduction strategies (cost-effectiveness of broken rail prevention and tank car safety design enhancement) and their optimal combination under a budget constraint.

A quantitative environmental risk analysis of rail transportation of a group of light, non-aqueous-phase liquid chemicals, was proposed in another study (Saat et al, 2014). The Hazardous Materials Transportation Environmental Consequence Model was used along with a geographic information system analysis of environmental characteristics, to develop probabilistic estimates of exposure to different spill scenarios along the North American rail network. The risk analysis considered the cost for cleaning up the soil contamination due to chemicals, route-specific probability distributions of soil type and depth to groundwater, annual traffic volume, rail-car release incident rate, and tank car safety features, to estimate the nationwide annual risk of transporting each product. According to the authors, the proposed approach can enable more effective management of the environmental risk of transporting hazardous materials.

Barkan et al. (2013) conducted a study as an effort to identify the causes of train accidents that can result in a tank car release of hazardous materials. Railroad derailment data were analyzed to identify the conditions most likely to lead to a release accident. The objective of the study was to identify proxy variables that can be used as performance measures. The results of the regression analysis showed that the speed of derailment and number of derailed cars were highly associated with hazardous materials release. Some accident causes were much more likely to lead to release conditions than others. Accident prevention strategies to reduce these causes were identified as more likely to reduce the risk of major railroad hazardous materials release accidents.

The broad objective of a research project (Bing et al., 2015) was to examine the causes and consequences of hazardous materials releases following railroad train accidents. The objectives of the project included: understanding the chain of events that could lead to a hazardous material release and the subsequent consequences; understanding the key risk metrics that quantify the likelihood and severity of each event along the chain of events, and comment on the significance of the metrics for hazmat transportation safety; performing initial evaluations of a specified set of risk reduction actions. The identified chain of events was comprised of: 1) Freight train accident due to an infrastructure or equipment defect, failure of signal or communications equipment, human error, or miscellaneous cause; 2) One or more freight cars derailed in accident, as a function of accident type, train speed, etc.; 3) Hazardous materials tank cars among derailed cars; 4) Derailed hazmat tank car releases following a train accident; and 5) Harm to people, property, and/or the environment. Also, the authors concluded that the best risk-reduction strategy might be to prevent train accidents from occurring in the first place, as opposed to mitigating the severity of events later in the chain which will also reduce the risks of other accidents on the territory.

3.2 Crude Oil

Due to dramatic growth of shipment of crude oil by rail in the past decade in the United States (9,500 carloads in 2008, growing to 407,761 carloads in 2013 according to American Association of Railroads), and the consequent increase in the number of accidents occurred involving trains that ship crude oil, a number of recent studies have focused on this matter.

A literature survey study (Lord et al., 2015) was conducted to examine the publicly available data pertaining to chemical and physical properties of tight crude oils as they relate to potential combustion events in the rail transport environment. The literature and data sources that were reviewed included recent reports on Bakken crude properties commissioned by the American Fuel & Petrochemical Manufacturers, North Dakota Petroleum Council, and U.S. Department of Transportation Pipeline and Hazardous Materials Safety Administration, and data from the U.S. Strategic Petroleum Reserve. The authors listed the key findings of this literature-based investigation as:

- Due to significant variability in criteria and procedures utilized in selection, acquisition, and analysis of crude oil samples, the available data are of insufficient quality to enable a meaningful comparison of crude oils—either to each other or against a designated standard.
- In addition to variability due to sampling and analysis methods, variability may also be introduced through crude oil conditioning, storage, and transport.
- Currently used methods for assignment of crude oil transportation hazard classification and packing group are often inadequate.

> The United Stated Department of Transportation's PHMSA (Pipeline and Hazardous Materials Safety Administration) regulates the transport of hazardous materials by all modes within the U.S. The Hazardous Materials Regulations require the shipper to properly classify hazardous material prior to offering it for transportation. In three basic steps, the shipper of a hazardous material must: 1. Properly identify all the hazards of the material 2. Determine which of the nine hazard classes characterizes the hazards associated with the material 3. Assign each material to a packing group, if applicable. HMR has nine hazard classes, and most crude oils fall into Class 3: flammable liquids. Other possible hazard classes for petroleum hydrocarbons include flammable gas (Class 2) or non-hazardous. Once assigned to a hazard class, materials may then be assigned to packing groups (PG), from great danger (PG I) to minor danger (PG III).

- Relationships between crude oil properties and probability or severity of combustion events in rail car spill scenarios have not been established.
- There is some literature consensus that vapor pressure of a "flammable liquid" is a property of interest, as the vapor phase evolved from a liquid actually burns rather than the liquid phase.

However, general lack of uniformity in methods and QA/QC across industry makes comparison of crude oil vapor pressure difficult.

- Bakken crude is a light, sweet oil that exhibits a statistically higher true vapor pressure than the slightly heavier, blended sweet and sour oils that are stored at the U.S. Strategic Petroleum Reserve (SPR).
- Numerous combustion events can occur from an accident involving hydrocarbons and hydrocarbon mixtures including crude oils, with severity dependent on the amount of fuel involved, surrounding infrastructure, and environment.

No single parameter defines the degree of flammability of a fuel, rather, multiple parameters are relevant. PHMSA issued a Safety Alert on January 2, 2014 to notify emergency responders, shippers, carriers, and the public that recent derailments and resulting fires indicate that the sweet crude oil transported from the Bakken region of North Dakota may be more flammable than traditional heavy crude oil. Looking at the properties of Bakken region crude oil in report to congress (Andrews, 2014), the author concluded that based on factors such as volatility, corrosivity, hydrogen sulfide content, and composition/concentration of the entrained gases in the material, this crude oil might be more flammable than regular crude oil. Regarding this fact, the author suggests some considerations about the tank cars used for shipment of crude oil from Bakken region and also some policy considerations.

In a project (Lee et al., 2015), a panel of leading experts on oil chemistry, behavior, and toxicity reviewed the available literature relevant to potential oil spills in Canadian marine waters, lakes, waterways, and wetlands. The review examined spill impacts and oil spill responses for the full spectrum of crude oil types and included scientific literature, key reports and selected oil spill case studies, including tanker spills, an ocean rig blowout, pipeline spills, and train derailments. The panel also consulted industry, government and environmental stakeholders across Canada. The results uncovered that dozens of crude oil types transported in Canada exist along a chemical continuum, from light oils to bitumen and heavy fuels, and the unique properties of each of these oil types determine how readily spilled oil spreads, sinks, disperses, and impacts aquatic organisms and what proportion ultimately degrades in the environment. Despite the importance of oil type, the panel concluded that the overall impact of an oil spill depends mainly on the environment and conditions where the spill takes place and the time lost before remedial operations.

Chapter 4 Task 1: Current and Projected Shipments

The first task is to predict current and future oil by rail shipments produced in the Bakken region. The initial step is to examine current and future production of oil in the Bakken and refinery demand from states around the nation through the year 2040. The second step is model what future Bakken oil production, future production in other parts of the country, and future demand for crude oil from refineries implies about oil by rail shipments. The third step is to model how current and future oil by rail shipments would change in the presence of the Keystone XL Pipeline project.

4.1 Current and Future Oil Production and Refining

The U.S. Energy Information Administration does not provide a specific forecast for Bakken oil production or shipments but does provide current state production data and multi-state projections which can be used to develop estimates for the Bakken region. State production in North Dakota and Montana is used as the estimate of production in the U.S. portion of the Bakken region. Historic state oil production data are available from the U.S. Energy Information Administration (2018a). There was an estimated 1.225 million barrels of production per day in the Bakken region in 2015, of which 1.117 million barrels are from North Dakota alone.

Oil production forecasts are available for six multi-state districts. Figure 4.1 reprints a figure developed by the Energy Information Administration which shows the six regions and production projections through 2040. The projections are from *2017 Annual Energy Outlook* and include forecast for the Dakotas/Rocky Mountain region which contains the U.S. portion of the Bakken region (U.S. Energy Information Administration, 2017a). Most of the Bakken region is in North Dakota. Comparing the 2017 average daily production between the Rocky Mountains/Dakotas and the Bakken it seems that the Bakken regions accounts for about 67 percent of Rocky Mountains/Dakotas region production (U.S. Energy Information Administration, 2017c). That report also notes that the Bakken region is responsible for most of the growth in oil production occurring in the Rocky Mountain/Dakotas.

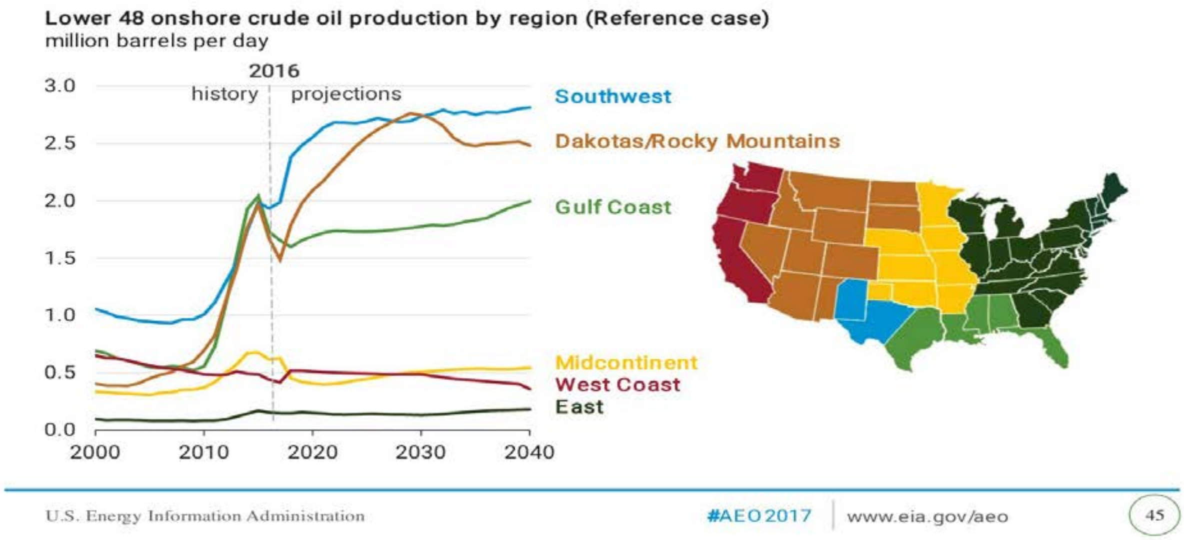

Figure 4.1 Production regions and forecast

Our baseline estimate is that the Bakken region will continue to account for 67 percent of oil production in the Dakotas/Rocky Mountain region (see Appendix 1). Bakken oil production would reach a peak of 1.760 million barrels per day in 2029 under this baseline assumption and 1.582 million barrels per day in 2040.

Figure 4.2 shows estimated Bakken production over time under the baseline assumption and other scenarios where Bakken oil product accounts for between 55% and 95% of regional production. Production dips after 2015 in reaction to low international oil prices but then rises steadily through 2029 before declining through 2050. If a 95 percent assumption were used instead, Bakken production would peak at more than 2.5 million barrels per day in 2029 and be 2.3 million barrels per day in 2040.

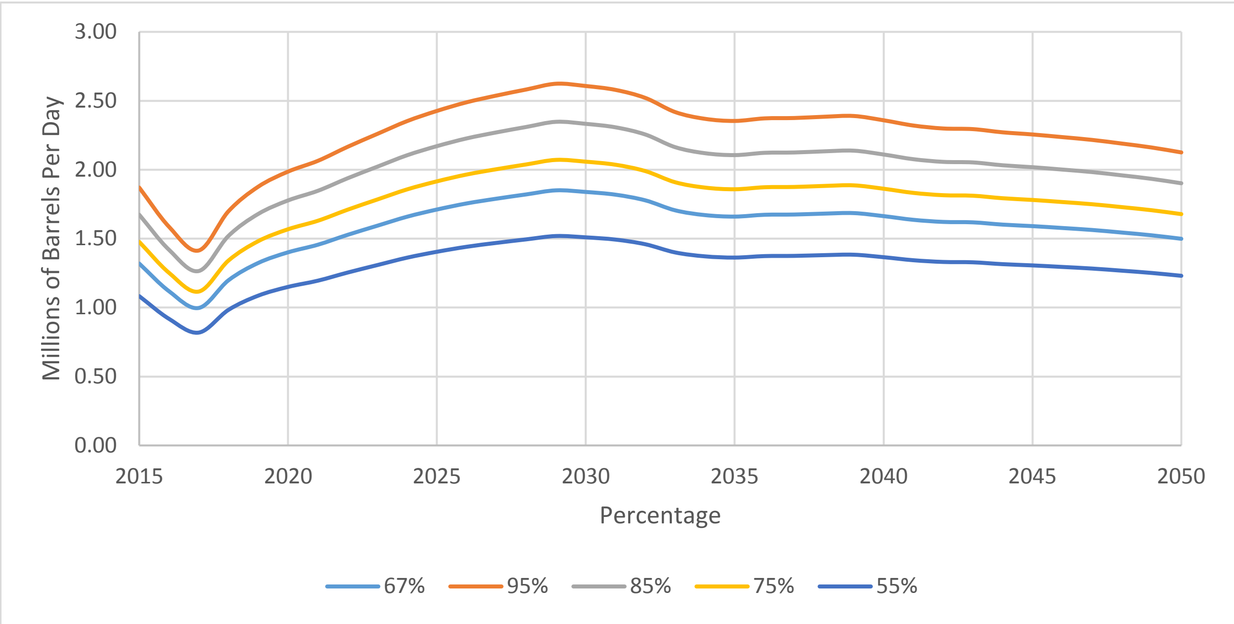

Figure 4.2 Estimated Bakken percentage of Rocky Mountain region production. Source: Author's calculations based on U.S. Energy Information Administration, 2017a. Annual Energy Outlook 2017

This approach can be used to forecast oil production in all 48 continental U.S. states and state data can be used to develop production estimates for 5 Petroleum Administration Defense Districts (PADDs). Figure 4.3 below shows the 5 PADDs for the continental United States. It is important to redefine production projections for these PADDs since data on oil shipping within the United States is organized by PADDs. In particular, the United States Department of Energy provides current estimates of oil shipments between PADDs overall and by mode, whether pipeline, rail, tanker, or barge.

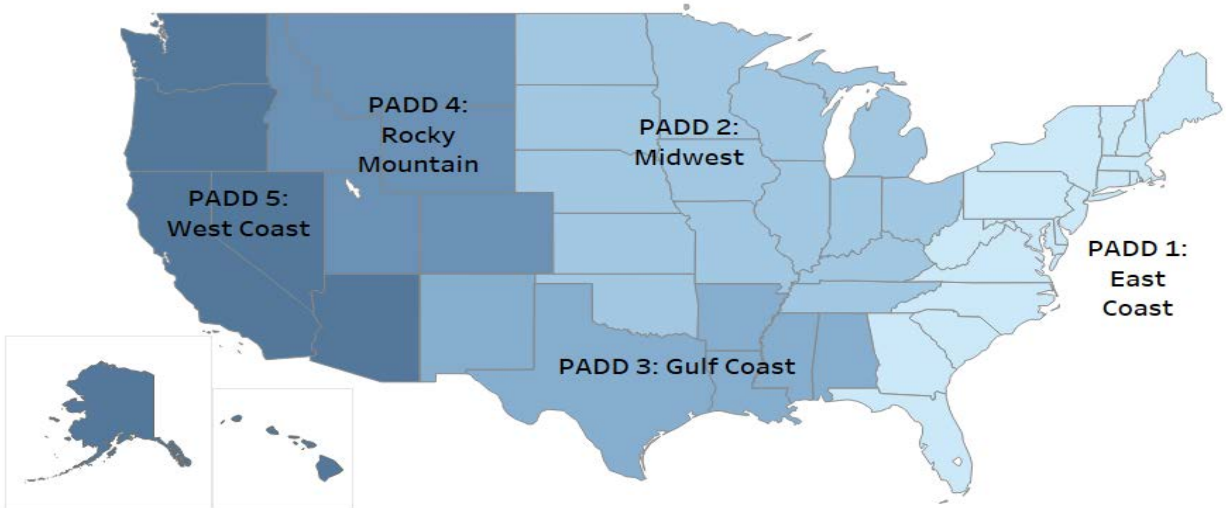

Figure 4.3 PADD definitions

For example, Figure 4.4 shows recent shipments of oil by rail from PADD 2, which contains North Dakota, to three other PADDs during recent years. Results show that as Bakken oil production ramped up after 2010, oil by rail shipments rose rapidly to the East Coast (PADD 1). Oil by rail shipments declined beginning in early 2016 as oil prices declined. Shipments to the Gulf Coast (PADD 3) similarly declined. Notably, oil by rail shipments to the West Coast (PADD 5) remained fairly constant; in other words, did not decline from 2013 levels.

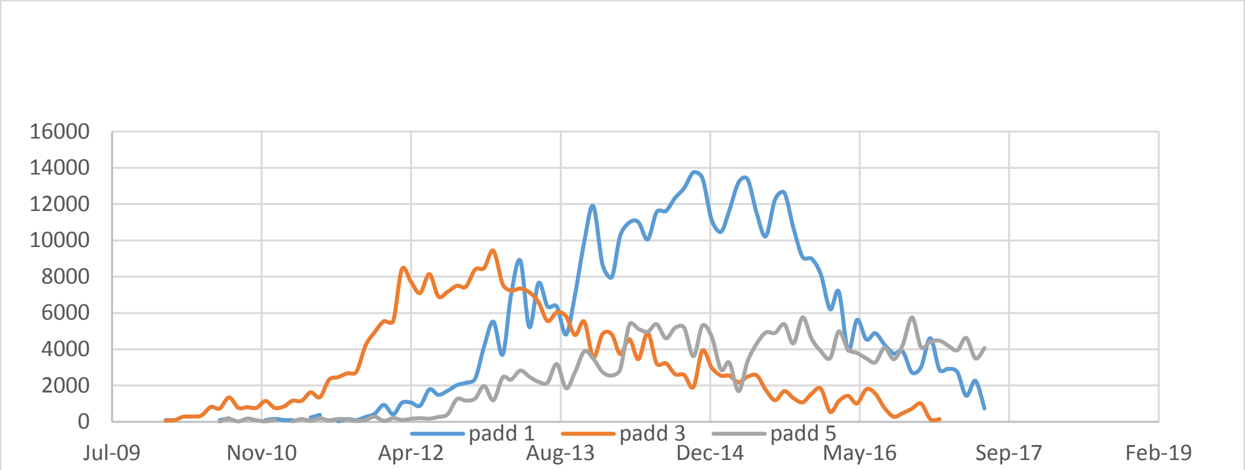

Figure 4.4 Oil by rail shipments from PADD 2. Source: U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker (ship), Barge, and Rail between PAD Districts.*

Future oil by rail shipment patterns will depend on future demand from oil refineries in the 5 PADDs, as well as the presence of pipeline capacity or other factors which influence the economic feasibility of oil by rail shipments. There is a need to estimate oil by rail shipments for PADDs and their component states through 2040 to compare with oil production.

Unfortunately, the U.S. Energy Information Administration does not provide a forecast through 2040 for how much crude oil will be consumed. The agency, however, does provide historic data on crude oil refining capacity, which the research team used as a proxy to forecast crude oil demand (U.S. Energy Information Administration 2017b). Specifically, historic refining capacity for each PADD was modeled as a function population to establish whether demographic forecasts could be used to forecast oil capacity and refining.

Crude oil refining capacity is strongly correlated with population growth in three PADD regions, PADDS 2-4. PADDs 1 and 5 do not follow any obvious pattern looking at population or regional oil production. Table 4.1 shows the relationship for between PADD 2 and PADD 5. There is a positive and statistically significant relationship between population and oil refining capacity in PADD 2 but not in PADD 5.

Table 4.1: Regressions of Refining Capacity and Population

	PADD 5 Refining Capacity	**PADD 2 Refining Capacity**
Constant	2,828,037***	-138,473
	(206,407)	(324,5923)
PADD Population	0.0043	0.0475***
	(.0038)	(0.0041)
R-squared	0.052	0.847
Adjusted R-squared	0.012	0.841
Number of Observations	26	26

Standard Errors are reported in parenthesis

*** indicates significance at the 99% level

Regression results such as those reported in Table 4.1 were used to forecast refining capacity in PADDS 2-4. Forecasts of refining capacity depend on population forecasts. Population forecasts were available from the University of Virginia (Weldon Cooper Center for Public Services, 2018). The University of Virginia produces their population forecasts by extending past population growth combined with expected changes in the population of people of child bearing age. Figure 4.5 shows the capacity forecast for PADD 2. PADDS 1 and 5 were extended in order to forecast their trend to 2040.

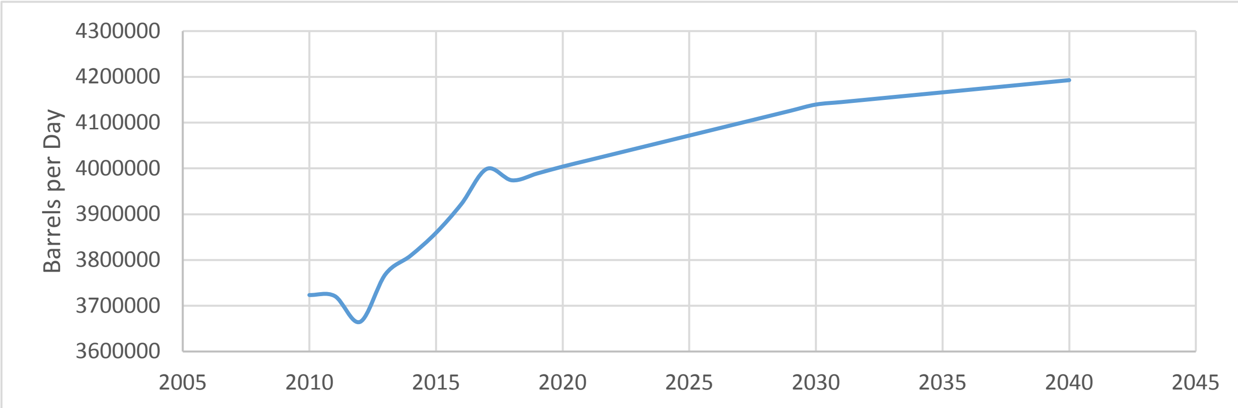

Figure 4.5 PADD 2 oil refining capacity. Source: U.S. Energy Information Administration, 2017b. Petroleum and Other Liquids

The next step was to forecast refining capacity for each state. To make state forecasts, each state's share of their PADD's refining capacity in 2017 was assumed to remain the same throughout the forecasted period. This is the same assumption that was made for state oil production. For example, Oklahoma's oil refining capacity was about 13% of PADD 2's refining capacity in 2017. Figure 4.6 shows the state forecast.

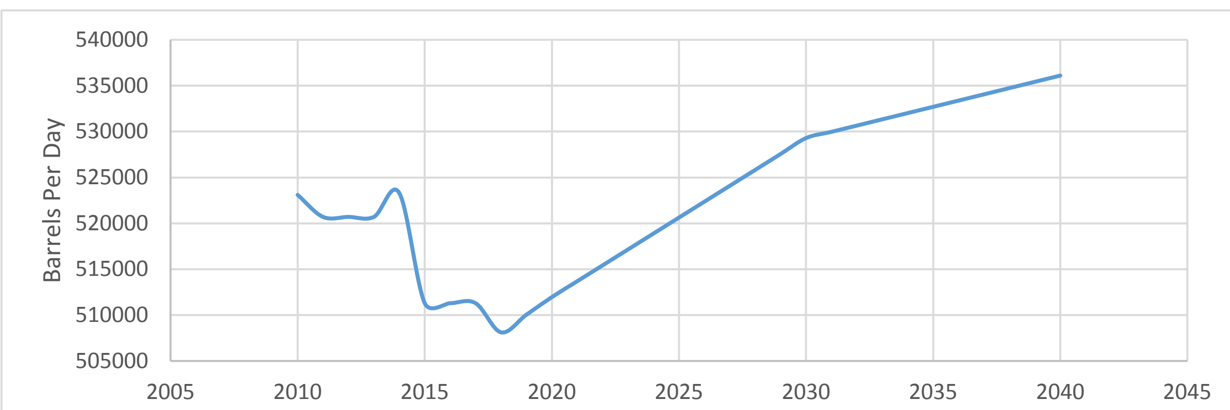

Figure 4.6 Oklahoma refining capacity. Source: U.S. Energy Information Administration, 2017b. Petroleum and Other Liquids

4.2 Implications for Oil by Rail Production

There is a significant need for crude oil transportation within the continental United States, especially given current and projected differences in the location of oil production and oil refining. The U.S. Energy Information Administration already tracks shipments among the 5 PADDs both in aggregate and by pipeline, rail and other modes (such as trucking to nearby pipelines). Data on oil by rail shipments from PADD 2 in recent years were presented in Figure 4.4. Similarly, Figure 4.7 shows oil pipeline shipments from PADD 2 in recent years. Oil by pipeline shipments to the Gulf Coast (PADD 3) rose consistently throughout the period.

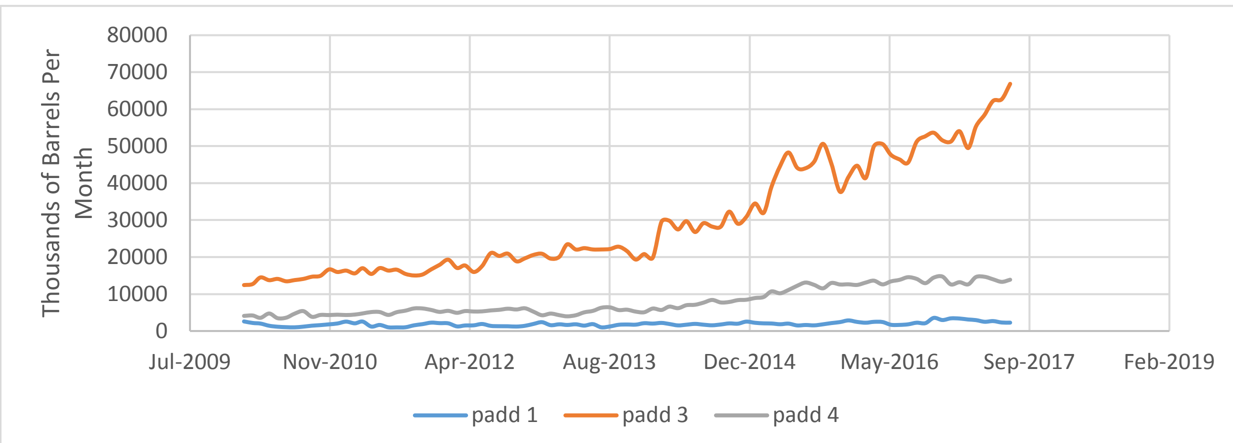

Figure 4.7 Oil by pipeline shipments from PADD 2. Source: U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker (ship), Barge, and Rail between PAD Districts*.

Table 4.8 shows the disposition of oil production in PADD 2 in recent years. The share of oil production is presented for oil refining within PADD 2 and oil transportation out of PADD 2 by mode. A small share, less than 10%, of PADD 2 production, is refined within the region. The proportion shipped by pipeline increased in 2014, when the 100,000 barrel per day Butte Pipeline expansion was completed. The proportion shipped by rail rose initially with Bakken production, but begin to fall in 2014. The share of oil shipped by rail was greater than the share shipped by pipeline during the 2013 to 2015 period. The addition of new pipeline capacity in 2014, and declining demand for oil by rail shipments due to market conditions, caused shipments by pipeline to grow to 80 percent of production by 2017.

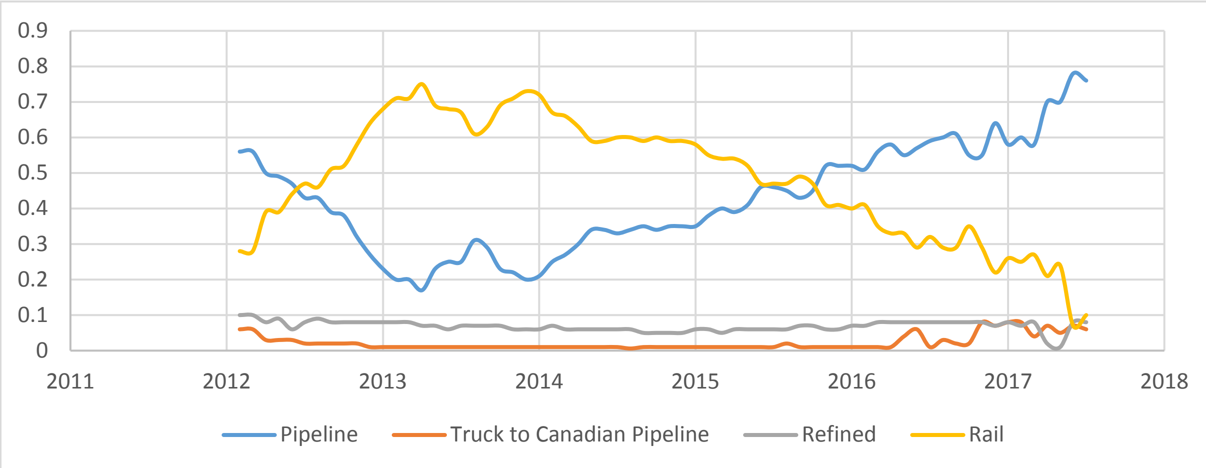

Figure 4.8 Portion of PADD 2 oil disposition by model. Source: U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker (ship), Barge, and Rail between PAD Districts*.

These findings are consistent with information available from regulatory authorities within the state of North Dakota. In particular, the North Dakota Pipeline Authority provides historical estimates of how much oil was transported by rail in North Dakota which averages to about 31% of the oil transported in North Dakota over the last 2 years (Kringstad, 2017). The North Dakota Pipeline Authority also provides rail and pipeline transportation capacity information from 2007-2017 and a projected oil transportation capacity for the next three years (North Dakota Pipeline Authority, 2017).

Based on the projections for the production of oil and the transportation capacity of oil, oil transportation will not be restricted due to lack of capacity. Projected oil production is less than the combined rail and pipeline capacity in all years. Pipeline capacity is less than the projected oil production, so some rail must be used. The recent decline of rail usage to transport oil from the Bakken region can be explained by increased pipeline capacity and usage for oil destined for the Gulf Coast. However, oil by rail has not generally declined in terms of shipments to the West Coast. No oil pipeline routes extend directly from the Bakken to the West Coast.

Table 4.2 below provides specific estimates of oil by rail shipments from the Bakken region (North Dakota and Montana) to other destination states. Bakken-to-state shipping estimates are provided for the recent year of 2015 and the year 2040. These Bakken-to-state shipping estimates were developed utilizing a linear programming model, which is described in detail in the next section. The model provides state-to-state oil by rail shipping patterns for both recent years and projected into the future. Table 4.2 shows model estimates of annual oil by rail shipments from the Bakken (North Dakota and Montana) to states within each PADD. Estimated oil by rail shipments in 2015 were 264.8 million barrels per year. About half of oil by rail shipments were to the East Coast (PADD 1), in particular to New Jersey, Pennsylvania, and Virginia. One third of shipments went to the Gulf Coast (PADD 3) states of Louisiana and Texas. Shipments of 50.0 million barrels per year go to the West Coast (PADD 5); in particular, to the state of Washington. There were no modeled shipments to Rocky Mountain States (PADD 4).

Model estimates for 2015 benefited from data on PADD to PADD oil by rail shipments. In particular, state-to-state shipment estimates were required to sum PADD to PADD totals. Oil by rail shipment estimates, therefore, reflected shipments from PADD 2 to PADD 1, PADD 3, and PADD 5 reported by the U.S. Energy Information Administration. Projections for 2040 were not required to conform to PADD to PADD shipment data, which are not known for 2040, allowing the model to choose state-to-state shipping activity to minimize shipping costs. Without the restriction, the linear programming model estimated that only the Bakken to PADD 5 shipments would be present in 2040. Shipments are up slightly due to rising production in the Bakken region from 2005 to 2040.

Table 4.2 Model Estimates for Annual Oil by Rail Shipments to PADD Districts

Bakken to PADD Oil By Rail Shipments (Thousands of Barrels Per Year)		
	2015	2040
PADD 1 (New Jersey, Pennsylvania, Virginia)	128,834	0
PADD 3 (Louisiana, Texas)	85,889	0
PADD 4 (none)	0	0
PADD 5 (Washington)	50,063	59,318
Total	264,786	59,318

Source: U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker (ship), Barge, and Rail between PAD Districts*.

4.3 Oil by Rail Shipments with the Keystone XL Pipeline

Historical data for PADD 2 and the Bakken region reveal that oil by rail shipments are impacted by oil pipeline capacity. In particular, oil by rail activity dropped with the opening of the Butte pipeline expansion in 2014. The addition of the north-south Keystone XL Pipeline through Nebraska would increase the potential to ship oil by pipeline to the Gulf Coast (PADD 3), including Louisiana and Texas. Specifically, the Keystone XL Pipeline would connect Hardisty, Alberta to Steele City, Nebraska with the goal of delivering oil from the US and Canada to the Gulf coast. The Keystone XL Pipeline would also connect with east to west pipelines, including in Steele City, Nebraska, facilitating shipments to the East Coast (PADD 1), including New Jersey, Pennsylvania, and Virginia. The Keystone XL Pipeline would not increase the capacity for pipeline shipments to the West Coast (PADD 5).

The Keystone XL Pipeline project, therefore, would provide significant competition for oil by rail shipments to the East Coast (PADD 1) and the Gulf Coast (PADD 3). There would be a significant reduction in current oil by rail shipments to these two PADDs. However, no significant disruptions would be expected to oil by rail shipments to the West Coast (PADD 5). Note that this implies that the Keystone XL Pipeline project would not necessarily impact projected oil by rail shipments in 2040. Table 4.3

shows how the completion of the Keystone XL Pipeline project would be expected to change current oil by rail shipments. Current shipments are based on model estimates for 2015 from Table 4.2. Table 4.3 assumes that shipments to PADD 1 and PADD 3 would be completely eliminated with the availability of the Keystone XL Pipeline. The total annual loss in oil by rail shipments would be 214.7 million barrels per year. Oil by rail shipments to PADD 5, however, would not be impacted.

Table 4.3 Current Oil by Rail Shipments With and Without the Keystone XL Pipeline

	Current Bakken to PADD Oil By Rail Shipments (Thousands of Barrels Per Year)		
	Without Keystone XL Pipeline	With Keystone XL Pipeline	Change
PADD 1 (New Jersey, Pennsylvania, Virginia)	128,834	0	-128,834
PADD 3 (Louisiana, Texas)	85,889	0	-85,889
PADD 4 ()	0	0	0
PADD 5 (Washington)	50,063	50,063	0
Total	264,786	50,063	-214,723

Source: U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker (ship), Barge, and Rail between PAD Districts*.

Chapter 5 Task 2: Release Incident Rates and the Frequency, Severity, and Non-Injury Costs of Release Incidents

In this section, four statistical models were estimated for frequency, severity, and costs of rail-based crude oil release incidents in the United States. State-to-state volume of crude oil movement, distance, availability of other modes of transportation (pipelines and waterways), and number of available class I railroad companies were considered as predictor variables in the models. Using these models, the effects of the predictor variables on frequency, severity, and costs of rail-based crude oil release incidents were identified and quantified, and then these models were used for predicting these measures for the future through 2040 using the crude oil production-consumption projections.

5.1 Methods

The methods used in this section include an optimization formulation for approximating state-to-state crude oil movement in the US, and statistical modeling techniques used for identification of effective factors and prediction. These methods are introduced in this sub-section.

5.1.1 Linear Program (LP) for U.S. State-to-State Crude Oil Movements

Energy Information Administration (EIA) reports the movement of crude oil in the U.S. based on Petroleum Administration for Defense Districts (PADDs), which are geographic aggregations of the 50 States and the District of Columbia into five districts. Based on the accessible information regarding annual state production of crude oil, annual state refining capacity, state-to-state transportation distance, unit-price of crude oil transportation for different modes (rail, pipe and water), and the PADD-to-PADD movement of crude oil by transportation mode information, a Linear Program (LP) was formulated to approximate annual state-to-state volume of crude oil movement. This LP is presented in equation set 1.

$$Min\ Z = \sum_{i=1}^{51}\sum_{j=1}^{51}\sum_{m=1}^{3} x_{ij,m} d_{ij} c_m \qquad (1)$$

subject to:

$(i)\ \sum_{j=1}^{51}\sum_{m=1}^{3} x_{ij,m} = P_i$ for $i = 1$ to $i = 51$

$$(ii)\ \sum_{i=1}^{51}\sum_{m=1}^{3} x_{ij,m} \le R_j \qquad \text{for } j = 1 \text{ to } i = 51$$

$$(iii)\ \sum_{i\epsilon N_k}\sum_{j\epsilon N_l}\sum_{m=1}^{3} x_{ij,m} = T_{kl,m} \qquad \text{for } k = 1 \text{ to } k = 5 \text{ and } l = 1 \text{ to } l = 5$$

$$(iv)\ x_{ij,m} \ge 0 \qquad \text{for } i,j = 1 \text{ to } i,j = 51 \text{ and } m = 1\ to\ 3$$

In this formulation $x_{ij,m}$ is the volume of crude oil movement from state *i* to state *j* by transportation mode *m*, d_{ij} is the defined distance from state *i* to state *j*, c_m is the cost per unit volume per unit distance of transportation of crude oil by mode *m*, P_i is the annual crude oil production of state *i*, R_j is the annual petroleum refinery capacity for state *j*, N_k is a set of states that belong to PADD *k*, and $T_{kl,m}$ is the annual volume of crude oil movement from PADD *k* to PADD *l* by transportation mode *m*.

The objective function *Z* represents the total costs of movement of crude oil among all the 50 United States and the District of Columbia (D.C.). Constraints *(i)* assure all the annual produced crude oil is moved to refineries, constraints *(ii)* keep the annual volume of crude oil moved to each state less than or equal to the annual refining capacity of the state, constraints *(iii)* satisfy the PADD-to-PADD crude oil movement by transportation mode among states, and constraints *(iv)* assure the non-negativity of the movement volumes. For 51 origins, 51 destinations, 3 modes of transportation (rail, pipe and water), and 5 PADDs, the LP includes 7803 decision variables, 153 type *(i)* equality constraints, 153 type *(ii)* inequality constraints, 60 type *(iii)* equality constraints, and 7803 non-negativity constraints.

The assumptions of this formulation include: c_ms are equal all over the U.S. and throughout time, regardless of origins and destinations; all the produced crude oil is shipped in the same year to refineries; and cost of transportation is the only factor that impacts transportation mode and destination choice.

5.1.2 Statistical Modeling Techniques

As was mentioned, state-to-state frequency, severity, and costs of crude oil release rail incidents were statistically modeled on a set of predictor variables. The severity measures were the number of released tank cars and quantity of released crude oil. The predictor variables were state-to-state volume of crude oil shipment, distance of shipment, availability of other modes of transportation, the number of competing class I railroads, and quadratic forms and interaction terms of volume and distance. Mixed-effects Negative Binomial Regression (MNBR) was used for modeling frequency and number of released tank cars, and Mixed-effects Ordered Logit Models (MOLM) were used for modeling categorized quantity of release and total monetary costs. These methods are introduced in this sub-section.

5.1.3 Mixed-effects Negative Binomial Regression (MNBR)

Frequency of incidents and number of released tank cars are count response variables. Popular regression models for count data are Poisson regression models, as a class of generalized linear models (GLM) (Bilder and Loughin, 2014). These models do not account for overdispersion, meaning that there is more variability to the counts than what the models assume there is (Cox 1983). Negative Binomial Regression (NBR) is often used as an alternative to the Poisson regression to account for overdispersion. NBRs assume a loglinear relation between the count response variable and the predictor variables.

Let $V_1, V_2, \ldots, V_n$ denote an independent and identically distributed sample of unit mean gamma random variables with shape parameter α; that is $f(v_1) \propto v_1^{\alpha-1} e^{-\alpha v_1} I(v_1 > 0)$. Suppose the i^{th} count Y_i has a Poisson distribution with mean $v_i\mu_i$ conditional on v_i, therefore $Y_i|v_i \sim Poisson(v_i\mu_i)$. The counts are then marginally independent negative binomial variables with mass functions given by equation 2, where $y \in \{0, 1, 2, \ldots\}$ (Booth, et al., 2013).

$$\Pr(Y_i = y; \alpha, \mu_i) = \frac{\Gamma(\mathrm{y} + \alpha)}{\Gamma(\alpha) y!} \left(\frac{\alpha}{\mu_i + \alpha}\right)^{\alpha} \left(\frac{\mu_i}{\mu_i + \alpha}\right)^{y} \tag{2}$$

If μ_i is related to a set of predictor variables, denoted by vector $\boldsymbol{x}_i$, while β_0 and $\boldsymbol{\beta}$ are the model constant and the vector of model coefficients, respectively, the NBR loglinear equation will be as equation 3.

$$\log(\mu_i) = \beta_0 + \boldsymbol{x}_i{}'\boldsymbol{\beta} \text{ or } \mu_i = e^{\beta_0 + \boldsymbol{x}_i{}'\boldsymbol{\beta}} \tag{3}$$

In this study a potential three-level correlation among the observations because of presence of grouping among them was possible. These grouping levels were pairs of origin-destination states, pairs of origin-destination PADDs, and the year incidents occurred in. In other words, there could be a correlation among the frequency of incidents and number of released tank cars that occurred between the same pairs of states and/or PADDs in different years, or the incidents that occurred in the same year, regardless of their origin and destinations. One way to account for multilevel grouping is inclusion of random effects to the above NBR (Bilder and Loughin, 2014; Booth et al, 2003), resulting in Mixed-effects Negative Binomial Regression (MNBR), as in equation 4.

$$\log(\mu_i) = \beta_0 + \boldsymbol{x}_i{}'\boldsymbol{\beta} + b_{0i} + \boldsymbol{x_i}'\boldsymbol{b_i} \text{ or } \mu_i = e^{\beta_0 + \boldsymbol{x}_i{}'\boldsymbol{\beta} + b_{0i} + \boldsymbol{x}_i{}'\boldsymbol{b}_i} \tag{4}$$

In this equation b_{0i} is the random parameter for the model constant (for observation *i*) and it is assumed to have a Normal distribution with mean 0 and unknown variance. $\boldsymbol{b_i}$ is the vector of random parameters for some or all of the predictor variables' coefficients (for observation *i*), and they are also assumed to follow Normal distributions with mean 0 and unknown variances. These variances are estimated along with the fixed effects. These models were estimated using the R function *glmer.nb()* from the *lme4* package (Bates, et al., 2014).

5.1.4 Mixed-effects Ordered Logit Models (MOLM)

Ordered Logit Models (OLM), is a tool for modeling ordinal categorical response variables on a set of predictors. If the probability of category *i* of the *J* categories of the response variable is π_i, then cumulative probability for category *j* of *Y* is $P(Y \leq j) = \pi_1 + \pi_2 + \cdots + \pi_j$ and $P(Y \leq J) = 1$. The log-odds of cumulative probabilities is, then, as equation 5 (Bilder and Loughin, 2014).

$$\log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \log\left(\frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_J}\right) \tag{5}$$

OLM assumes this log-odds of cumulative probabilities is a linear function of the predictors and the slope of this relationship is the same regardless of the category (Bilder and Loughin, 2014; Agresti and Kateri). The OLM model is stated as equation 6.

$$\log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \beta_{j0} - \boldsymbol{x}_i{}'\boldsymbol{\beta} \tag{6}$$

In this equation vector $\boldsymbol{x}_i$ is a set of predictor variables, β_{j0} is the model constant for the response category j and $\boldsymbol{\beta}$ is the vector of coefficients.

Like the MNBR model, these models can account for the potential three-level correlation among quantity and cost of release incidents as a result of grouping by inclusion of random effects in the OLM models, leading to mixed-effects ordered logit models (MOLM). Equation 7 shows MOLM. In this equation b_0 and $\boldsymbol{b}$ are defined as in equation 4 (Christensen, 2011).

$$\log\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \beta_{j0} - \boldsymbol{x}_i'\boldsymbol{\beta} + b_0 - \boldsymbol{x_i}'\boldsymbol{b} \quad (7)$$

Function *clmm()* from the R package *ordinal* was used for estimating MOLMs (Christensen, 2011).

5.1.5 Model Interpretation Tools

Percentage Change (PC) and Odds Ratios (OR) were used as model interpretation tools for MNBR and MOLM, respectively. PC is defined as the percentage change in the mean response that results from a *c*-unit change in a predictor x_i (holding other predictor variables constant) *(13)*. In MNBR, PC for x_i equals $100(e^{c\beta_i} - 1)$, if only the main effects of x_i is used in the model, and equals $100(e^{c\beta_i + c\beta_i' x_i} - 1)$, if the quadratic form of x_i is also in the model (β_i' is the coefficient of the quadratic term). OR for MOLMs is defined as the change in the odds of $Y > j$ versus $Y \leq j$, corresponding to a *c*-unit change in a predictor, x_i. In case of inclusion of only the main effects of x_i in the model, OR equals $e^{c\beta_i}$, and equals $e^{c\beta_i + c\beta_i' x_i}$, if the quadratic form is included. In this study PCs and ORs are reported in terms of point estimates, along with 95% Confidence Intervals (CI). Parametric bootstrap and Wald CIs were calculated for PCs and ORs, respectively (Bilder and Loughin, 2014; Agresti and Kateri).

5.2 Data

The data used in this study was comprised of several datasets obtained from different sources. These included U.S. crude oil release rail incidents data, state production of crude oil, crude oil wells and refineries locations, state capacity of crude oil refining, PADD-to-PADD data of crude oil movement by water, pipe and rail, U.S. class I railroads maps, and U.S. crude oil pipeline and waterway maps. These datasets and the final variables are presented in this sub-section.

Transportation distance was used as a cost factor in the LP (equation 1) and as a predictor that affected frequency and severity of incidents. In this study, this distance was defined as the geodesic (the shortest path between two points on an ellipsoid) distance between each state's origin points to all the states' destinations points. Origin points of each state were defined as the geometric centroid of the crude oil wells in that state, and the destination points of each state was defined as the geometric centroid of the refineries located in that state. Origin/destination was considered as the geometric centroid of the state, if there were not any wells/refineries located in that state. The location information of 2016 U.S. oil and gas wells and 2017 U.S. refineries were obtained from *FracTracker (2017)* and the *Energy Information Administration (EIA) (2017a)*, respectively. The geometric centroids of oil and gas wells and refineries in each state (origins and destinations), and the distances from all origins to all destinations were calculated using the geographic information system software *ArcGIS* version 10.5.1. The LP introduced in the methodology section was solved for ten years (2007-2016) to approximate the state-to-state crude oil movement volumes. The LP's input data was obtained from different sources. Distance (d_{ij}) was defined and was assumed to be constant in throughout the ten years. It was sufficient to consider the cost of moving crude oil by mode m, c_m, in a relative manner. Based on contacts with different crude oil carriers the costs of moving crude oil by rail was assumed 7.15 times as large as pipe and 5 times as large as water (c_{rail}= 5.0, c_{pipe}= 0.7 and $c_{water} = 1.0$). Despite the existence of spatial and temporal variations in these ratios, they were assumed constant in this study, as the LP was relatively insensitive to changes of these values (less than 1% changes in the output) due to consideration of constraint *(iii)* which assures the correct share of modes. Annual crude oil production (P_i), annual petroleum refinery capacity (R_j) and the annual PADD-to-PADD volume of crude oil movement were also obtained from EIA (Energy Information Administration, 2017b) for 2007-2016.

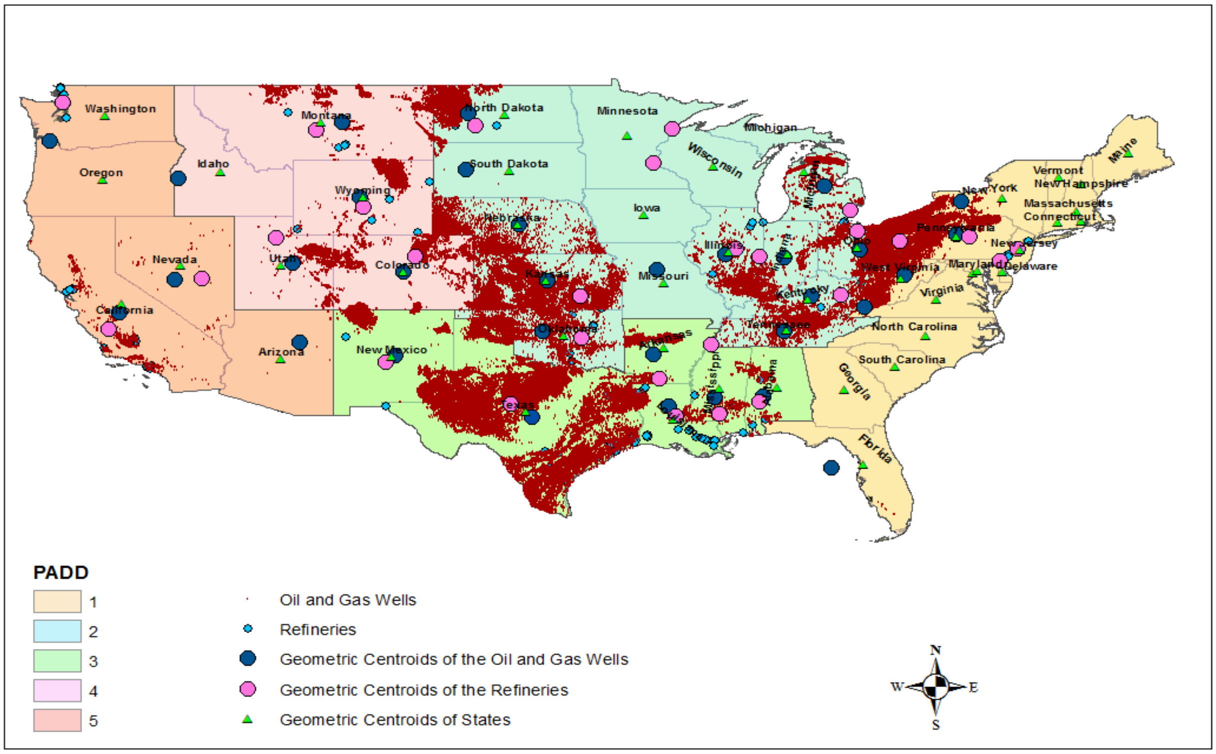

Figure 5.1 Map of the United States, PADDS, oil and gas wells, refineries, and their geometric centroids (from various sources mentioned in the text)

Two variables were defined to capture the possible effects of availability of other modes or other class I railroad companies on frequency, severity, and costs of release incidents. This was based on the hypothesis that in case of availability of pipelines and/or waterway for movement of crude oil, the railroad companies may try to decrease their price to stay a competitive mode by decreasing their costs, leading to a lower level of safety. Also, the larger the number of competing class I railroad companies are available between the origin and destination, similar intention may result in cheaper but less safe transportation. A binary variable accounted for availability of other modes based on the petroleum pipelines and waterways for petroleum movement maps, obtained from EIA (2017a). A continuous variable captured the number of available class I railroads between origins and destinations, based on the class I railroad maps available from Association of American Railroads (2017).

Ten-year data (2007-2016) of crude oil release incidents from trains in the U.S. was extracted from the PHMSA incident database by the Incident Reports Database Search tool (Pipeline and Hazardous Materials Safety Administration, and Office of Hazardous Materials Safety, 2017). It included 460 release incidents, 680 released tank cars, 1,738,926 gallons of released crude oil and $65,608,355 total damages. Total damages included carrier/property damage, response/clean-up costs, evacuation costs, injuries/fatalities, and roadway closure (costs of evacuation were assumed $250 per person-day *(8)*, monetary costs of not-hospitalized injury as the only type of injury/fatality that occurred in the dataset was assumed $62,500 per injury (Iranitalab and Khattak, 2017), and roadway closure was assumed to cost $218,000 per day (Erkut, et al., 2007; Mallela and Sadavisam, 2011). This dataset included the origin and destination of movement of each train that was involved in the release incidents. Using this information, the annual frequency of incidents, number of tank cars, quantity of crude oil released and total costs for each pair of states (with at least one incident) were extracted. Pairs of states with larger-than-zero approximated crude oil movement volumes were added to the dataset with zero for frequency and severity of incidents. Volumes and other variables (distance, other modes and other class I railroad companies) were also added. The final dataset was comprised of 318 rows. Each row was a pair of states with positive volume of crude oil exchange in one of the years 2007-2016. The summary of the variables is presented in Table 5.1.

Table 5.1 Variables and Descriptive Statistics of the Final Dataset

Variable	Variable Type	Values and Statistics
		Response Variables
Frequency	Count	Min = 0, Max = 17, Mean = 1.3648, Var. = 5.6772
Number of Tank Cars	Count	Min = 0, Max = 35, Mean = 2.0440, Var. = 19.6511
Quantity Released	Continuous (gallons)	Min = 0, Max = 475176.00, Mean = 5451.20, Var. = 1.93E+09
	Categorical	Categories: = 0, 0 < ≤100, 100 < ≤10000, >10000 Ratios: 0 (45.77%), 1 (47.34%), 2 (04.07%), 3 (02.82%)
Total Costs	Continuous (2016 U.S. Dollar)	Min = 0, Max = 25,632,806, Mean = 205,669, Var. = 2.71E+11
	Categorical	Categories: = 0, 0 < ≤15000, 15000 < ≤100000, >100000 Ratios: 0 (56.43%), 1 (33.86%), 2 (04.39%), 3 (05.33%)
		Predictor Variables
Volume (*volume*)	Continuous (1000 barrels)	Min = 0, Max = 1.54E+05, Mean = 1.75E+04, Var. = 9.35E+08
Distance (*distance*)	Continuous (miles)	Min = 67.03, Max = 2384.39, Mean = 742.0607, Var. = 2.20E+05
Other Modes (*omodes*)	Dichotomous	Yes (38.99%), No (61.01%)
Number of Class I Railroad Companies (*railroads*)	Count	Min = 0, Max = 3, Mean = 1.3648, Var. = 0.8066

The variances of the two continuous response variables (quantity released and total costs) were relatively large. This was due to small portion of extremely large values relative to the other values in these two variables, which could cause biased estimates if a linear regression model was utilized

(Nachtsheim, et al., 2004). Natural logarithm or a root transformation were possible solutions for this issue. However, as logarithm of zero is not computable and model interpretation of a root transformed response variable is not as conclusive, an ordinal categorization of these variables was preferred. Categorization also alleviated the effects of possible inconsistency and inaccuracies in reporting and approximating costs and quantities. The thresholds of the categories were determined based on the variables' dispersion between maximum and minimum values and abating the effects of the very large values without excluding them.

5.3 Modeling Results

This sub-section presents the results of the four models that were estimated for frequency, severity, and costs of crude oil release rail incidents. These models considered response variables' relationship with the predictors between each pair of states with positive crude oil transportation volume: 1) the number incidents with released crude oil; 2) the number of tank cars that released crude oil; 3) the quantity of crude oil released; and 4) the total monetary costs of crude oil release.

In all four models, the predictors included the volume and distance of crude oil shipment between pairs of states as continuous variables, availability of other modes of transportation as a binary variable (yes/no), number of available class I railroads as an integer variable (0-7), and quadratic and interaction terms for volume and distance variables. Three grouping factors were considered in the models: year; origin-destination state pairs; and origin-destination PADD pairs. All the four main variables (volume, distance, other modes, and railroad companies) were used in the models regardless of their statistical significance, while the inclusion of quadratic and interaction terms, and the grouping factors were decided based on Corrected Akaike Information Criteria (AICc) values (Cavanaugh, 1997).

The estimated coefficients, likelihood ratio (LR) test p-values and estimated standard deviations of random effects for the intercepts are presented in Table 5. The quadratic form of volume was significant in all models, while the quadratic form of distance and the interaction of distance and volume did not contribute to any of the models in terms of AICc and were excluded. The contribution of three grouping factors varied among the models, which led to different random effects specifications. Random effects for variables other than the intercept did not contribute to the models' AICc. The estimated equations for the frequency, tank cars, quantity, and costs models are presented in Equations 2 to 5, respectively. In these equations $\hat{\mu}_i$ is the estimated frequency of crude oil rail incidents, $\hat{\mu}_t$ is the estimated number of tank cars released crude oil, $\hat{P}(Y_q \leq j)$ is the estimated probability of amount of crude oil release falling in a category equal to or smaller than category *j*, $\hat{P}(Y_c \leq j)$ is the estimated probability of costs of crude oil release falling in a category equal or smaller than category *j*, X_1 is the amount of crude oil shipped between a pair of states in thousand barrels per year, X_2 is the geodesic distance between a pair of states in miles, X_3 is the availability of modes other than rail (pipeline/water) between a pair of states, X_4 is the number of available class I railroad companies between a pair of states, *e* is the base of natural logarithm and $N(\mu, \sigma^2)$ denotes a normal distribution with mean of μ and variance of σ^2. The other parameters correspond with their definitions in the methodology section.

$$\hat{\mu}_i = e^{-2.76761+0.00003X_1-0.000000000168{X_1}^2+0.00131X_2+0.31678X_3+0.61329X_4+\hat{b}_s+\hat{b}_p},$$
$$\hat{b}_s \sim N(0, 0.8090^2), \hat{b}_p \sim N(0, 0.2076^2). \quad (2)$$

$$\hat{\mu}_t = e^{-2.79799+0.00004X_1-0.000000000207{X_1}^2+0.00148X_2+0.29936X_3+0.58017X_4+\hat{b}_s},$$
$$\hat{b}_s \sim N(0, 1.0840^2). \quad (3)$$

$$\frac{\hat{P}(Y_q \leq j)}{1-\hat{P}(Y_q \leq j)} = e^{\hat{\beta}_{j0}-0.00009X_1+0.000000000462{X_1}^2-0.00508X_2-1.77976X_3-1.00642X_4+\hat{b}_s+\hat{b}_p+\hat{b}_y},$$
$$\hat{b}_s \sim N(0, 3.34268^2), \hat{b}_p \sim N(0, 1.88159^2), \hat{b}_y \sim N(0, 0.07415^2), \quad (4)$$
$$\hat{\beta}_{00} = 0, \hat{\beta}_{10} = 5.711, \hat{\beta}_{20} = 13.773, \hat{\beta}_{30} = 15.263.$$

$$\frac{\hat{P}(Y_c \leq j)}{1-\hat{P}(Y_c \leq j)} = e^{\hat{\beta}_{j0}-0.00008X_1+0.000000000400{X_1}^2-0.00271X_2-1.18009X_3-0.55749X_4+\hat{b}_s+\hat{b}_p},$$
$$\hat{b}_s \sim N(0, 1.623^2), \hat{b}_p \sim N(0, 1.397^2), \quad (5)$$
$$\hat{\beta}_{00} = 0, \hat{\beta}_{10} = 4.433, \hat{\beta}_{20} = 8.206, \hat{\beta}_{30} = 9.068.$$

Percentage change (PC) and odds ratios (OR) were used for interpretation of MNBR and MOL, respectively. Confidence intervals (CI) were calculated for these two measures, along with point estimates to assist with model interpretation. As the quadratic form of the variable *volume* was included in the final models, PC or OR for this variable was a function of itself, while they were independent for other variables. Therefore, point estimates and 95% CIs for PCs and ORs are presented in Table 5.2 for variables *distance*, *omodes*, and *railroads*, while for *volume* these measures are illustrated in Figure 5.2, corresponding to a range of values for *volume*. Parametric bootstrap and Wald 95% CI's were calculated for MNBR and MOL models, respectively. The value of *c* in calculating PC and OR for *volume*, *distance*, *omodes*, and *railroads* were assumed as 1000, 100, 1, and 1, respectively.

Inclusion of "0" in CI's for PC's and inclusion of value of "1" in CI's for OR's denote the lack of evidence towards the statistical significance of the variable's effects on or association with the response variable. With 95% confidence and holding all variables constant except the variable being interpreted, the models can be interpreted as follows:

For each 100-mile increase in the distance of crude oil shipment by rail between pairs of states, the frequency of crude oil release incidents increased by 7.78% to 19.08%. This change in distance led to 8.00% to 21.10% increase in the number of released tank cars. Corresponding with the 100-mile increase in distance, the odds of increase in quantity released from one of the predetermined levels to a higher level changed by 1.31 to 2.11 times. This change also resulted in 14% to 51% positive change in the odds of increase in total costs from any level to a higher level.

The models indicated a lack of evidence for the existence of any impacts from availability of modes of transportation (other than rail) from the origin states to the destination states on frequency or severity of crude oil release incidents. However, the number of class I railroad companies between states were found to be statistically significant in the frequency and tank car models. Each one-unit increase in the number of class I railroad companies resulted in an increase in the frequency of crude oil release incidents by 39.45% to 129.55% and in the number of release tank cars by 27.50% to 129.39%. There was not any sufficient evidence for the effects of this variable on quantity released or total costs.

Table 5.2 Estimation Results of the Four Incident Frequency and Severity Models

Model Components		Variables	Frequency Model			Tank Cars Model			Quantity Released Model			Total Costs Model		
			Coefficient	LR Test p-value		Coefficient	LR Test p-value		Coefficient	LR Test p-value		Coefficient	LR Test p-value	
Fixed Effects	Main Effect	Intercept	-2.76761			-2.79799			5.711, 13.773, 15.263			4.433, 8.206, 9.068		
		Volume	0.00003	0.00000	***	0.00004	0.00000	***	0.00009	0.00022	***	0.00008	0.00001	***
		Distance	0.00131	0.00000	***	0.00148	0.00000	***	0.00508	0.00000	***	0.00271	0.00005	***
		Other Modes	0.31678	0.14582		0.29936	0.23414		1.77020	0.06852	.	1.20350	0.07169	.
		Railroad	0.61329	0.00000	***	0.58017	0.00005	***	1.00220	0.09064	.	0.67910	0.13693	
	Quadratic and Interaction Terms	Volume^2	-1.68E-10	0.00046	***	-2.07E-10	0.00015	***	-4.62E-10	0.01367	*	-4.00E-10	0.00129	***
		Distance^2	—	—	—	—	—	—	—	—	—	—	—	—
		Volume*Dist.	—	—	—	—	—	—	—	—	—	—	—	—
Standard Deviation of Random Effects for the Intercept		States	0.809			1.084			3.34268			1.623		
		PADDs	0.2076			—			1.88159			1.397		
		Year	—			—			0.07415			—		

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

—: Not used in the model due to not contributing to the model according to AICc

Table 5.3 5% Confidence Intervals and Point Estimates of PC's and OR's

Variables	Frequency Model			Tank Cars Model			Quantity Released Model			Total Costs Model		
	Percentage Change			Percentage Change			Odds Ratios			Odds Ratios		
	Point Estimate	Lower Bound of CI	Upper Bound of CI	Point Estimate	Lower Bound of CI	Upper Bound of CI	Point Estimate	Lower Bound of CI	Upper Bound of CI	Point Estimate	Lower Bound of CI	Upper Bound of CI
Distance	14.01	7.78	19.08	15.92	8.00	21.10	1.66	1.31	2.11	1.31	1.14	1.51
Other Modes	37.27	-9.11	104.88	34.90	-18.24	117.16	5.93	0.84	42.05	3.25	0.85	12.43
Railroad	84.65	39.45	129.55	78.63	27.50	129.39	2.72	0.82	9.10	1.75	0.84	3.63

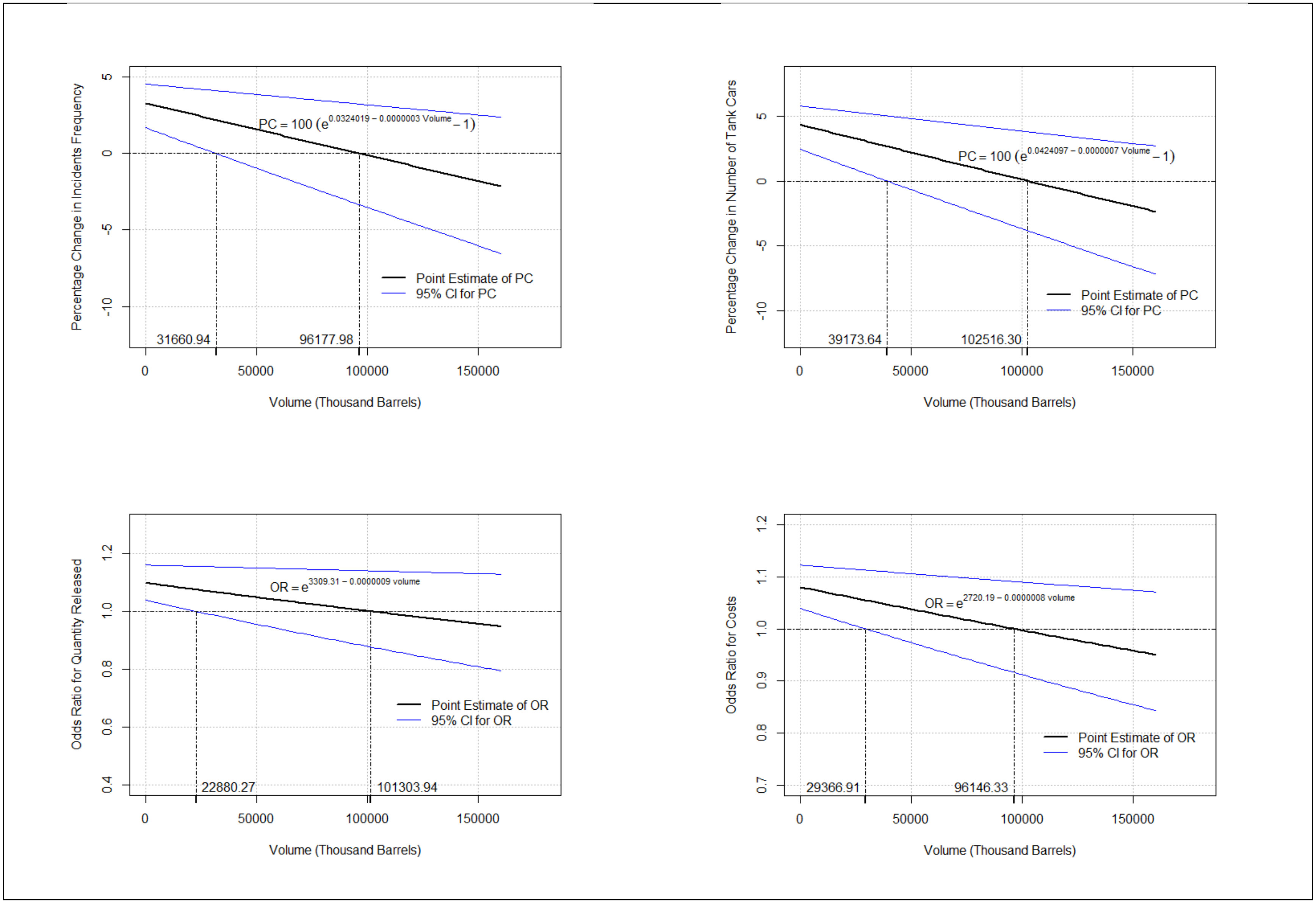

Figure 5.2 95% confidence intervals and point estimates of PC and OR for *volume* in the four estimated models.

Figure 5.2 shows how point estimates and CI's for *volume*'s PC and OR change as a function of *volume* itself. For each 1-million-barrel per year increase in the shipment of crude oil between a pair of states, sufficient evidence was found regarding the response variables limited to restricted volumes. The frequency of incidents increased by variable amounts less than 5%, up to a volume point of approximately 32 million barrels. The number of tank cars released crude oil increased by a value between 0 to 6%, up to a volume point of approximately 40 million barrels. Quantity released, and total costs increased by less than 20% and 15%, respectively, and for up to approximately 23 million and 29 million barrels per year. The accurate amount of change can be calculated for all possible values of volume using the PC and OR estimated equations reported in Figure 5.2. With 95% confidence, there was not sufficient evidence found for effects of *volume* on the response variables for values higher than the ones mentioned. In all four models, approximately after 100 million barrels, increase in *volume* was identified to decrease frequency and aggregate severity of incidents, without statistical significance. This may be due to lack of sufficient observation in this volume range, relative to lower values for volume.

Prediction for Future

In this sub-section, using the LP and models estimated in the previous sub-section, and also the crude oil production and refining capacity projection for future throughout 2040 (presented in chapter 4), frequency, severity, and costs of crude oil rail-based release incidents are predicted. In solving the LP, the values for $T_{kl,m}$ were adjusted for each future year based on the projected production of the origins and refining capacity of the destinations. These predictions are presented in Table 5.4. For example, the predictions for year 2025 is as follows: there will be 57 crude oil release cases; 70 tank cars will release crude oil; in 19 O-D pairs there will be no release; in 42 O-D pairs the aggregate quantity of release will be less than 100 gallons; in 1 O-D pair the aggregate quantity of release will be between 100 to 10,000 gallons; in 37 O-D pairs there will be no costs of release; in 24 O-D pairs the aggregate costs of release will be less than $15,000, and in 1 O-D pair the aggregate costs of release will be over $10,000.

Table 5.4 Predictions for Frequency, Severity and Costs of Crude Oil Release Incidents in the Future

Year	Release Incidents	Tank	Quantity Released (gallons)				Costs ($)			
			0	0 < < 100	100 < < 10000	> 10000	0	0 < < 15000	15000 < < 100000	> 100000
2018	13	20	5	12	0	0	11	6	0	0
2019	11	19	5	12	0	0	11	6	0	0
2020	13	18	5	12	0	0	11	6	0	0
2021	14	20	5	12	0	0	11	6	0	0
2022	14	21	5	12	0	0	11	6	0	0
2023	15	22	5	12	0	0	11	5	0	1
2024	15	22	5	12	0	0	11	5	0	1
2025	16	24	5	12	0	0	11	5	0	1
2026	16	24	5	12	0	0	11	5	0	1
2027	16	25	5	12	0	0	11	5	0	1
2028	17	25	5	11	1	0	11	5	0	1
2029	17	26	5	11	1	0	11	5	0	1
2030	17	26	5	11	1	0	11	5	0	1
2031	17	27	5	11	1	0	11	5	0	1
2032	17	27	5	12	0	0	11	5	0	1
2033	17	27	5	12	0	0	11	6	0	0
2034	17	27	5	12	0	0	10	7	0	0

2035	17	26	5	12	0	0	10	7	0	0
2036	16	25	5	12	0	0	10	7	0	0
2037	16	24	5	12	0	0	10	7	0	0
2038	16	23	5	12	0	0	10	7	0	0
2039	16	23	5	12	0	0	10	7	0	0
2040	15	23	5	12	0	0	10	7	0	0

5.4 Types and Consequences of Crude Oil Release

This section investigates the effects of characteristics of crude oil, tank cars, and release incidents on type and consequences of crude oil release from trains. Then, the effects of types and consequences of release on the monetary costs of such incidents are identified and quantified. The results will be helpful in determination and economic evaluation of safety countermeasures, considering their costs and benefits.

Methods

This sub-section introduces the statistical approaches used in this section by first discussing multinomial response models, which are used for modeling types and consequences of release of crude oil, and then discussing continuous outcome models, which are used in modeling post-release costs.

5.4.1 Multinomial Response Models for Types and Consequences of Release

Two models were developed for types of crude oil release and consequences of crude oil release from trains. The response variables in these models were multinomial (multicategory) and indicated the type or consequence of crude oil release. Figure 5.3 shows the outcomes of the release incidents based on their frequencies in the dataset. One multinomial model with gas dispersion (with and without spillage), spillage (with and without gas dispersion), and simultaneous gas dispersion and spillage as categories of the response variable was estimated. The categories for the consequences multinomial model included fire, explosion and none. Multinomial logit modeling of these response variables utilized predictor variables that included characteristics of crude oil, tank car, and release incidents.

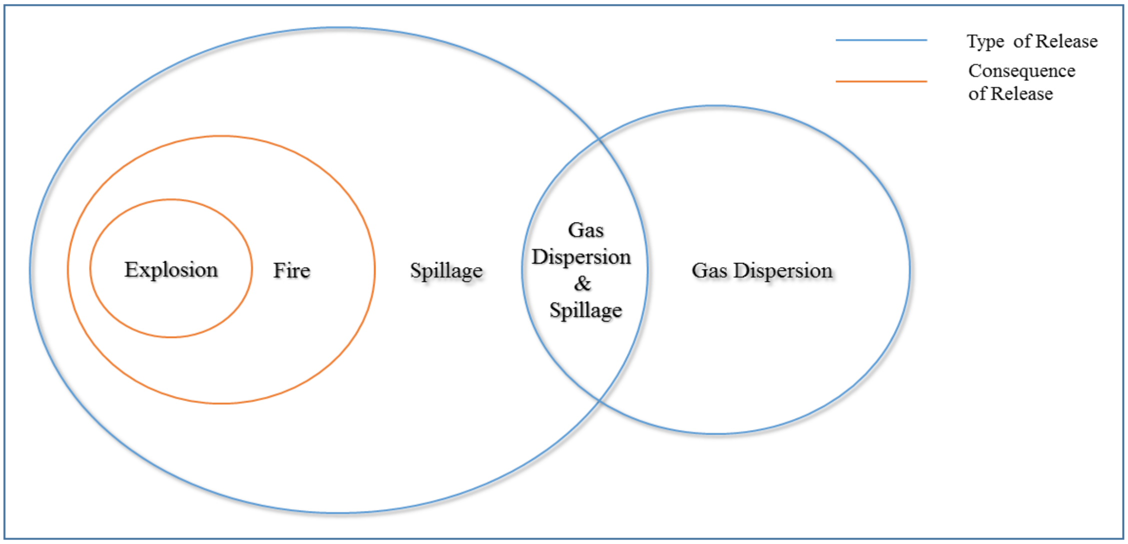

Figure 5.3 Types and consequences of crude oil release

Multinomial logit models are common multi-category nominal response models originally introduced for modeling choice behavior as a categorical outcome in economics (McFadden, 1980). They can model other categorical response variables and are sometimes known as multinomial regression or baseline logit models in those cases (Bilder and Loughin, 2014). Assuming the *n*th observation with a categorical outcome i, the utility function for this outcome is:

$$U_{in} = V_{in} + \varepsilon_{in} = \alpha_i + \beta_i X_{in} + \varepsilon_{in} \quad (6)$$

In Equation 6, V_{in} is the observable part of the utility function for *n*th observation with a categorical outcome *i*, α_i is an intercept, β_i is a vector of model coefficients, X_{in} is a vector of observable factors that influence the outcome, and ε_{in} is an error term that accounts for unobserved effects. An assumption of the error terms being independent and identical distributed with a generalized extreme value distribution results in the multinomial logit model as shown below (Train, 2002; Ben-Akiva and Lerman, 1985).

$$P_n(i) = \frac{e^{V_{in}}}{\sum_j e^{V_{jn}}} = \frac{e^{\boldsymbol{\alpha_i+\beta_i X_{in}}}}{\sum_j e^{\boldsymbol{\alpha_j+\beta_j X_{jn}}}} \quad (7)$$

In Equation 7, $P_n(i)$ is the probability of occurrence of outcome *i* for observation *n*. Like section 3, model interpretation utilized odds ratios, with similar definition and interpretation.

5.4.2 Continuous Response Model for Post-Release Costs

Total damage costs, as mentioned in section 3, included carrier/property damage, response/clean-up costs, evacuation costs, injuries/fatalities, and roadway closure. Linear regression models were used *(28)* with costs as continuous response variables, and types and consequences of release, and two other factors as predictor variables. The objective of estimating these models was testing whether the types and consequences of release of crude oil significantly affect the post-release costs, and if so, to quantify the effects.

5.5 Release and Predictor Data

The similar PHMSA-based data introduced in section 3 was used in the modeling. Table 5.5 presents the variables and their respective statistics.

The *Bakken* variable indicated whether the crude oil was shipped from the Bakken region and should have been categorized as light sweet crude oil or not. This variable was formed based on the origin state of the shipment (North Dakota or Montana). The packing group information was available in the dataset. Packing group I, II, and III represent great, medium, and minor danger, respectively. The criteria for assigning packing group for crude oil is based on flash point and initial boiling point of the crude oil, which shippers should obtain through laboratory tests (Class 3-Assignment, 2010). Information regarding tank head puncture resistance system and tank insulation was extracted from the tank car specification marking (Specifications for Tank Cars, 2010), that was available in the dataset. Tank head puncture resistance system is capable of sustaining coupler-to-head impacts of the relative speed of 18 mph, usually accomplished by the installation of separate head shields or full-head tank jackets made of

1/2-inch-thick steel on each end of the tank car. Tank insulation is used to moderate the temperature of crude oil during transportation (Maty, 2017).

Table 5.5 Variables and Their Statistics

Variable	**Names**	**Values and Statistics**
Response Variables		
Type of Release	Type	Spillage (86.21%), Gas Dispersion (08.93%), Both (04.86%)
Consequence of Release	Cons	Fire (07.21%), Explosion (07.21%), None (85.58%)
Predictor Variables		
Bakken Crude Oil	Bakken	0 = No (51.72%), 1 = Yes (48.28%)
Packing Group	pack.group	I (51.88%), II (30.41%), III (17.71%)
Tank Head Puncture Resistance System	punc.res	0 = No (90.28%), 1 = Yes (09.72%)
Tank Insulation	Insulated	0 = No (95.45%), 1 = Yes (04.54%)
Tank Design Pressure (psi)	Dsgnpress	mean = 107.97, variance = 3207.14
Quantity Released (gallon)	quant.rel	mean = 2994.55, variance = 56620119
Non-Accident Release (NAR)	Nar	0 = No (20.53%), 1 = Yes (79.47%)

Federal Railroad Administration (FRA) provides the definition of Non-Accident Releases (NARs) as “the unintentional release of a hazardous material while in transportation, including loading and unloading while in railroad possession, that is not caused by a derailment, collision, or other rail related accident. NARs consist of leaks, splashes, and other releases from improperly secured or defective valves, fittings, and tank shells, and include venting of non-atmospheric gases from safety relief devices.” NARs were detected in the data based on the provided narrations.

Post-release costs are introduced below. As was mentioned, these costs included carrier/property damage, response/clean-up costs, evacuation costs, injuries/fatalities, and roadway closure (costs of evacuation were assumed $250 per person-day (Saat, et al., 2014), monetary costs of not-hospitalized injury as the only type of injury/fatality that occurred in the dataset was assumed $62,500 per injury (Iranitalab and Khattak, 2017), and roadway closure was assumed to cost $218,000 per day (Erkut, et al., 2007; Mallela and Sadavisam, 2011). The minimum, maximum, mean, and standard deviation of the costs were $0, $25,330,322, $146,792, and $1,365,787, respectively (rounded to the nearest dollar).

5.6 Modeling Results

This section presents the estimated statistical models. These include multinomial logit models capturing the impacts of crude oil, tank car design and incident characteristics on type and consequence of release of crude oil in a train incident and a linear regression model quantifying the effects of type and consequence of release of crude oil on the post-release costs.

Table 5.6 p-values of the LR Test in the Release Type and Release Consequence Models

Variables		Type of Release		Consequence of Release	
Crude Oil Characteristics	bakken	0.00062	***	0.09105	.
	pack.group	0.00101	**	0.00000	***
Tank Car Characteristics	punc.res	0.00000	***	0.07486	.
	insulated	0.00198	**	—	
	dsgnpress	0.48524		—	
Incident Characteristics	nar	0.00000	***	0.00000	***
	quant.rel	NA		0.00000	***

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
—: Not Used
NA: Not Applicable

5.6.1 Models for Types and Consequences of Release

The types of release that were available in the dataset included gas dispersion, spillage, and both. Two possible consequences of release were considered in this paper: fire, and explosion. The variables used in each model and the p-values of the Likelihood Ratio (LR) tests are presented in Table 5.6. Variable selection was based on the Corrected Akaike Information Criteria (Bilder and Loughin, 2014; Agresti and Kateri, YEAR).

Point estimates of the odds ratios and their 95% Confidence Intervals (CI) for the release type with “spillage” as the base level are presented in Table 5.7. With 95% confidence and subject to keeping all the other variables (rather than the variable being interpreted) constant, the model interpretations are as follows: the odds of gas dispersion vs. spillage, and both types of release vs. spillage change by an amount between 0.2856 to 0.9841 times, and 0.0530 to 0.5457 times, respectively, for the light sweet crude oil (from Bakken region). Packing group II decreased the odds of gas dispersion vs. spillage by 0.1405 and 0.6394 times relative to packing group I. These values were estimated as 0.0843 to 0.8416 for packing group III. Equipment of tank cars to puncture resistance system changed the odds of gas dispersion vs. spillage by an amount between 2.3509 to 9.6949 times, and both release types vs. spillage by an amount between 4.4941 to 33.6677 times. Insulation of the tank cars increased the odds of both release types vs. spillage only by 2.5800 to 20.2558 times. The odds of gas dispersion vs. spillage were increased by an amount between 1.7430 to 466.3863 times, for non-accident releases. Other than these effects, there was no sufficient evidence on the impacts of predictor variables on the types of release.

Table 5.7 Values of c, Point Estimates of Odds Ratios and Confidence Intervals for Odds Ratios in the Release Type Model

Variables		c	Gas			Both (Gas and Spillage)		
			Point Estimate	95% CI Lower Bound	95% CI Upper Bound	Point Estimate	95% CI Lower Bound	95% CI Upper Bound
Crude Oil Characteristics	Bakken	1	0.53012	0.28557	0.98409	0.17001	0.05297	0.54568
	pack.groupII	1	0.29975	0.14052	0.63941	1.08628	0.41387	2.85119
	pack.groupIII	1	0.26631	0.08427	0.84156	2.09962	0.76861	5.73554
Tank Car Characteristics	punc.res	1	4.77408	2.35091	9.69489	12.30062	4.49408	33.6677
	Insulated	1	1.32465	0.37406	4.69094	7.22909	2.57998	20.25583
	dsgnpress	25	0.97387	0.80125	1.18369	0.93484	0.77104	1.13345
Incident Characteristics	Nar	1	28.5114	1.74298	466.38631	11.01463	0.61996	195.69301

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 5.8 presents the odds ratios and 95% profile LR CI’s for the consequences of crude oil release model, with “none” as the base level. Again, with 95% confidence and holding all other the variables except the variable being interpreted constant, it can be said that packing group II, relative to packing group I increased the odds of explosion vs. no release consequence by an amount between 1.6065 to 158.9321 times. There was no sufficient evidence towards the existence of any impacts of packing group II on fire and packing group III on fire and explosion, relative to packing group I. Non-accident releases, relative to accident releases, decreased the odds of fire and explosion vs. no consequence, by amounts between 0.0040 to 0.1088 time, and 0.0002 and 0.0566 times, respectively. The odds of fire and explosion vs. no release consequence in a crude oil release incident increased for every 1000 gallon increase in quantity of release of crude oil by a percentage between 13.69% to 37.56% and 12.96% to

37.91%, respectively. Sufficient evidence was not available to support the existence of any effects of Bakken region crude oil, tank head puncture resistance system, tank car insulation, and tank car design pressure on fire or explosion.

Table 5.8 Values of c, Point Estimates of Odds Ratios and Confidence Intervals for Odds Ratios in the Release Consequence Model

Variables		C	Fire			Explosion		
			Point Estimate	95% CI Lower Bound	95% CI Upper Bound	Point Estimate	95% CI Lower Bound	95% CI Upper Bound
Crude Oil Characteristics	Bakken	1	4.03791	0.57541	28.33582	14.3892	0.47405	436.76913
	pack.groupII	1	0.52436	0.04996	5.50364	15.97908	1.60654	158.93206
	pack.groupIII	1	0.47702	0.07433	3.06148	0.08171	0.00278	2.40099
Tank Car Characteristics	punc.res	1	5.56493	0.86115	35.96186	0.79086	0.0295	21.20505
	Insulated	1	—	—	—	—	—	—
	Dsgnpress	25	—	—	—	—	—	—
Incident Characteristics	Nar	1	0.0208	0.00398	0.10879	0.00311	0.00017	0.0566
	quant.rel	1000	1.25059	1.13693	1.37561	1.2481	1.12956	1.37907

Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

—: Not Used

Post-Release Costs Model

A robust linear regression model was estimated at the incident level, using total costs as the response variables, and types of release and consequences of release as the predictor variables. The point estimates and 95% CI’s for the estimated coefficients, along with LR test p-values and standard errors are presented in Table 5.9. LR test results and CI’s indicate there was not enough evidence in the dataset to show that variations in types of release affected the damage costs directly. However, the estimated coefficients for fire and explosion, along with non-accident release variable and quantity released were statistically significant in the model. These variables changed damage costs by amounts between the upper and lower bounds of the CI's reported in Table 5.9.

Table 5.9 Robust Regression Results for Costs of Crude Oil Release from Trains

Coefficients	Point Estimate	Standard Error	LR Test p-value		95% CI Lower Bound	95% CI Upper Bound
(Intercept)	253274.38	936.22	—		251439.42	255109.33
Gas Dispersion	-291.41	454.96	0.52220		-1183.12	600.30
Spillage	303.96	563.48	0.58990		-800.44	1408.36
Fire	2072608.65	1205.57	0.00000	***	2070245.79	2074971.52
Explosion	13529080.97	2188.78	0.00000	***	13524791.03	13533370.91
Non-Accident Release	-251744.34	755.17	0.00000	***	-253224.45	-250264.23
Quantity Released	28860.20	192.79	0.00000	***	28709.20	29011.26

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

5.6.2 Model Implications for Accident Rates

Linear programming model outputs on state-to-state oil by rail shipment volumes, distances, and release incidents over the historic 2007 to 2015 period were used to estimate an historical release incident rate. Specifically, that programming model estimated both the volume of oil hauled by rail during a year between pairs of states as well as the hauling distance between the midpoint of each state and the predicted number of release incidents. The oil by rail release incident rate focused on the 2007 to 2015 period, given that shipping patterns changed significantly between 2015 and 2016. Data were combined to estimate that

4.80 trillion barrel-miles of oil hauled by rail during the 2007 to 2015 period between the lower 48 states. Given that each 120-car train can haul an estimated 97,140 barrels, this translates into 49.42 million trainload-miles of oil over the 10-year period.[2] There were 429 predicted release incidents over the 10-year period associated with the 49.42 million trainload-miles, implying that there are 8.68 release incidents per 1,000,000 miles of oil by train travel.[3]

At the end of our forecast period, in 2040, the linear programming models estimates that there will be 100.5 billion barrel-miles of oil hauled by rail between pairs of states, which is the equivalent of 1.03 million trainload-miles, assuming the typical length of trains does not change over time. There were an estimated 15.14 release incidents in 2040 (consistent with Table 5.4 above), implying that there will be 14.64 release incidents per 1,000,000 miles of oil by train travel in 2040.

[2] CSX the third largest freight rail company by revenue reports that their average train length is 6,833 feet (Johnson, 2017). A train car is about 5'5'' which means an average train has about 120 cars carrying freight. Sixty-nine percent of tank cars are the DOT-11 train car which can carry 809.5 barrels of oil. A train then with 120 tank cars can transport 97,140 barrels of oil.

[3] As the linear programming model produced estimates of state to state travel, it was determined that it was better to use release incident rates for all national oil by rail travel rather than release incident rates for oil hauled from the Bakken region (North Dakota and Montana). The release incident rate for oil hauled from the Bakken region would have been 9.70 accidents per million oil by rail train miles.

Chapter 6 Task 3: Economic Costs of Oil by Rail Release Incidents

Release incident rates in Task 2 can be utilized to estimate the total number and cost of release incidents generated by hauling oil by rail from the Bakken region. Current and predicted volumes of oil hauling by rail from Task 1 will be combined with information on the average barrels of oil hauled per train, trip length in miles and release incident rates per train-mile to predict the average annual number of oil by rail release incidents.

6.1 Number of Release Incidents Due to Shipment of Bakken Oil by Rail

Table 6.1 includes estimates of train-miles of Bakken oil shipped and predicted release incidents for the year 2015. Table 4.2 provided model shipment estimates from the Bakken region to other PADDs. There were 264.8 million barrels of oil hauled by rail from the Bakken region in 2015. The average trip length for Bakken oil by rail shipments in that year was 1,225 miles, given that oil was shipped to the East Coast and the Gulf Coast as well as to the West Coast. Multiplying this distance by 264.8 million barrels of oil suggests 324.4 billion barrel-miles of Bakken oil hauled by rail during calendar year 2015. Further, given 97,140 barrels per train, there were an estimated 3.34 million train-miles of oil by rail. Given a rate of 8.68 release incidents per 1 million train miles travelled, there were an estimated 28.99 release incidents involving trains hauling oil from the Bakken region in 2015. This estimate is reported in Table 6.1 below.

The introduction of the Keystone XL Pipeline is assumed to eliminate hauling of oil by rail to the East Coast (PADD 1) and the Gulf Coast (PADD 3). We assume that oil by rail shipping would not be limited to the West Coast (PADD 5). As seen in Table 4.2, the linear programming model estimates that oil by rail shipping to the West Coast was 50.063 million barrels per year in 2015. The average trip length to the West Coast was 942 miles, implying 0.49 million train-miles of oil by rail hauling per year and an estimated 4.21 release incidents.

Given the sharp decline in oil by rail shipment in recent years, an estimate was made regarding oil hauled by rail for the most recently available 12-month period (August 2016-July 2017) rather than the annual average from 2015. According to the North Dakota pipeline authority, there were 86.755 million barrels of oil shipped by rail from North Dakota during that time. This is approximately 24% of the total production of oil in North Dakota during that 12-month period, which is consistent with the data presented earlier in Figure 4.8. This estimate of 86.75 million barrels is used to estimate 92.48 million barrels of oil hauled per year for the entire Bakken region, given Montana oil production was 6.6% of North Dakota oil production during 2017 according the U.S. Energy Information Administration data. As seen in Table 4.4, by the 2016/2017 period a majority of Bakken oil by rail shipments are to the West Coast (PADD 5).

According to the results of the linear programming model, the average length of an oil by rail trip from North Dakota was 1,221 miles. Multiplying this distance by 92.48 million barrels of oil by rail hauling suggests 112.9 billion barrel-miles of Bakken oil hauled by rail during the 12-month period. Further, given 97,140 barrels per train there were an estimated 1.16 million train-miles of oil by rail hauling per 12 months. Given an release incident rate of 8.68 release incidents per 1 million train miles travelled, there are currently an estimated 10.09 release incidents per year involving trains hauling oil from the Bakken region. This estimate is reported in Table 6.1 below. As was assumed for 2015, levels of shipping to the West Coast were maintained to the 2016/17 period.

As was seen in Table 4.2, there would be an estimated 59.318 million barrel-miles of oil by rail hauling from the Bakken region in 2040, and that hauling would all be to the West Coast. The average trip length to the West Coast is 942 miles, implying 0.58 million train-miles of oil by rail hauling per year and an estimated 8.42 annual release incidents. These figures are reported in Table 6.1. That estimate would apply both with and without the Keystone XL Pipeline project.

Note that the estimated release incidents from 2040 oil by rail shipments originating in the Bakken region is more than half of the estimate of total release incidents (15.14) due to all state-to-state shipments of oil by rail in 2040 (see Table 5.4). This finding makes sense given that the linear programming model projects that oil by rail shipments from the Bakken region will account for 55 percent of all state-to-state oil by rail shipments in 2040.

Table 6.1 Estimates for Annual Train Miles and Release incidents

	Annual Train-Miles and Release Incidents Bakken Oil by Rail			
	Without Keystone XL Project		With Keystone XL Project	
	Train-Miles	Release incidents	Train-Miles	Release incidents
Recent (2015)	3.34 million	28.99	0.49 million	4.21
Current (2016/17)	1.16 million	10.09	0.49 million	4.21
Future (2040)	0.58 million	8.42	0.58 million	8.42

Source: Author estimates using output of linear programming model

6.2 Average Costs per Release Incident

The next step is to estimate the annual economic cost of oil by rail release incidents. The economic cost of these release incidents can be estimated by considering the average property damage, clean-up, injury, and fatality costs of oil by rail release incidents. Once estimated, release incident costs per mile of oil by rail hauling can be used to estimate the total cost of hauling oil by rail from the Bakken region, given the estimate of release incidents in Table 6.1.

Given the estimate of 10.09 release incidents in 2016/17, the average cost per release incident can be multiplied by 10.09 to estimate the annual cost of oil by rail release incidents due to hauling from the Bakken region. The monetary costs per release incident were estimated in the previous section based on incidents during the 2007 to 2015 period. The combined carrier costs, property damage, response costs, and cleanup costs per rail release incident by trains hauling oil was $125,500.

Injury and fatality costs per release incident are based on Federal Railroad Administration, Office of Safety Analysis data on fatal and non-fatal injuries from railroad release incidents, including at grade rail-highway release incidents, during the 2014 to 2017 period (Office of Safety Analysis, 2018). There were 47,165 release incidents reported by the Office of Safety Analysis during the 2014 to 2017 period, with 18 percent occurring at highway-rail intersections. Of the 47,165 release incidents over that period, there were 3,118 fatalities for a fatality rate of 6.61%. This does not mean that 6.61% of accidents resulted in fatalities, as some accidents would have resulted in multiple fatalities. There were 3,197 non-fatal injuries for a non-fatal injury rate of 74.63%.

The value of a statistical life was $9.6 million in 2017 implying an average fatality costs of $634,600 per rail release incident. The non-fatal injury cost per release incident naturally depends on the severity of the set off injuries which occur. Detailed information on the severity of injuries is not available for rail release incidents so data for injuries in motor vehicle accidents were utilized instead (Blincoe, et al., 2015). Values for the year 2010 were updated to 2017 using the increase in the value of a statistical life between 2010 and 2017. The social value per non-fatal injury was estimated to be $162,200. This social cost per non-fatal injury implies an average non-fatal injury costs of $121,000 per rail release incident.

The total cost per rail release incident is $881,200, the sum of the non-injury costs is $125,500, the fatality costs per release incident is $634,600 and the non-fatal injury costs per release incident is $121,000. Note that this value does not include any highway vehicle delay costs at highway rail at-grade intersections. These costs could not be calculated given the number of such intersections crossed along rail routes. In particular, relevant AADT at those intersections and other required information are not available.

In Table 6.2, the $881,200 cost per release incident is applied to the estimate 28.99 release incidents per year in 2015 due to oil by rail hauling from the Bakken region. The result is a total annual cost of $25.546 million for that year. But that was in a period before shifting market conditions began to sharply curtail oil by rail shipments to the East Coast and Gulf Coast. These changes are reflected in total release incidents costs for the more recent 2016/2017 period. The estimate of 10.09 release incidents during this 12 month period due to hauling Bakken oil by rail to yield a total cost of $8.891 million. This cost estimate is included in Table 6.2 below. The 2015 cost would fall to $3.710 million if the Keystone XL Pipeline was in place and operating. The same could be assumed for 2016/17. Future annual costs assume that release incident costs do not change in real terms. Future annual costs would be $7.420 million per year and would not vary based on the Keystone XL Pipeline project.

Table 6.2 Estimates for Annual Release Incidents and Social Release Incident Costs

	Annual Release Incidents and Social Release Incident Costs Bakken Oil by Rail			
	Without Keystone XL Project		With Keystone XL Project	
	Release incidents	Costs	Release incidents	Costs
Recent (2015)	28.99	$25.546 million	4.21	$3.710 million
Current (2016/17)	10.09	$8.891 million	4.21	$3.710 million
Future (2040)	8.42	$7.420 million	8.42	$7.420 million

Source: Author estimates using output of linear programming model

Chapter 7 Task 4: Infrastructure Investments

The research will examine the benefits and costs of four viaduct projects. Descriptions are provided for each of the four viaducts. A cost benefit analysis of each project is then developed with both the presence and absence of oil by rail shipments.

Oil by rail shipments impact the need for viaduct projects. This review of potential viaduct projects in North Dakota and bordering areas in Minnesota reinforced this connection. For example, one of the four projects, a rail crossing project in Dickinson, North Dakota near the Bakken oil fields, was motivated in part by oil by rail shipments. In particular, the Dickinson 2035 Transportation Master Plan (2013) implies that they expect more rail traffic because of oil from the Bakken region. The Dickinson 2035 plan says that “a new crude oil rail loading facility was recently constructed west of Dickinson and has capacity to load two-unit trains daily.” The Minnesota Department of Transportation also believes that there will be an increase of crude oil by rail shipments which helps motivates the decision to choose projects in Minnesota.

In this chapter, we evaluate four specific projects using a cost benefit analysis in areas seeing increased rail traffic from oil shipments from the Bakken region of North Dakota and Montana. Two of the projects are in Minnesota near the North Dakota border, and two are in North Dakota. A description of the four projects is provided below, first for the Minnesota projects and next for the North Dakota projects (Minnesota Department of Transportation, 2018; U.S. Department of Transportation Federal Highway Administration, 2017). Key characteristics of the projects are listed below. Note that annual release incident rates were estimated for each project by the research team based on AADT, the number of highway lanes and the number of trains per day. Release incident rates estimations utilized an accident prediction website maintained by the Office of Safety Analysis of the Federal Railroad Administration.[4] Predicted accidents were assumed to equal predicted release incidents.

[4] The Federal Railroad Administration (FRA) is an agency within the U.S. Department of Transportation. One of the purposes of the FRA is to advise on and regulate train safety in the United States (Federal Railroad Administration,

7.1 Minnesota Projects

One Minnesota project is in East Grand Forks where a rail line crosses 2nd Ave NE. The 2nd Ave project is forecast to cost $14.931 million. The FRA says that the 2nd Ave crossing has a 2.0949% of being the site of a train collision every year and has an AATD of 7,400, and 10 trains per day (Federal Railroad Administration Office of Safety, 2016).

Another crossing in Moorhead, Minnesota on 21st St. is projected to cost $30.0 million and the FRA gives it a 10.8275% chance of being the site of a train collision (Fargo-Moorhead Metropolitan Council of Governments, 2014). Notably 21st has an AADT of 28,000 compared to 2nd Ave, which likely explains some of the differences in collision possibilities. There are 8 trains per day.

2018a). The FRA Office of Safety Analysis has an accident prediction website which gives the characteristics of highway-rail crossing in the United States in addition to giving the probability of an accident occurring at a crossing within a year (Federal Railroad Administration, 2018b). The FRA gives the formula that they use to create the accident prediction. The formula uses data like automobile traffic, train traffic, number of lanes, number of tracks, maximum allowed train speed, type of crossing warning device, and past reported accidents at the crossing among other characteristics (Federal Railroad Administration, 2007). The accident probability produced by the FRA's formula is a product of weighted values for each characteristic where the weights depend on the type of warning device at a crossing. Since the FRA provides their formula it can be used to see how changes to crossing may affect the probability of an accident. The FRA Office of Safety Analysis also provides formulas for the probability that an accident results in an injury or fatality which makes it easier to conduct a cost benefit analysis of making a change to the accident probability at a crossing (Federal Railroad Administration, 2007). The FRA also provides information on the average cost of property damage for a type of accident (injury, fatal, no injuries or fatalities) which also aids a cost benefit analysis (Brod, et. al, 2013).

7.2 North Dakota Projects

One project site in North Dakota is 42nd St. in Grand Forks, which is expected to cost $21.384 million and have a 2.4163% chance of being the site of a train collision in a year (North Dakota Department of Transportation, 2017). There is an AADT of 13,490 and 10 trains per day.

The other project in North Dakota being considered is Dickinson at a rail crossing on State Ave (City of Dickinson, 2013). The project has been completed at a cost of $32.4 million and could be a useful place to study as it is near the Bakken oil fields. There is a 10.7332% chance of being the site of a train collision in a year. There is an AADT of 7,385 and 14 trains per day.

7.3 Benefit Cost Analysis

Benefit cost analyses for the four projects compare the construction costs for each viaduct project with the present value of project benefits. Project benefits accrue over a 30 year period and are discounted over time at a 7% real interest rate, in line with the rate of return on investment earned in domestic stock markets. No annual costs are assumed once the project is completed. Annual benefits accrue primarily to the highway users who utilize each viaduct.

There are two classes of benefits to consider for at-grade rail separation projects. One is the reduction in release incidents. The release incident rate at crossings with grade separation falls to zero. The second benefit results from improved traffic flow. Cars and trucks no longer must wait for trains to cross a road, leading to valuable time savings. Indeed, the Dickinson Transportation Plan (2013) gives long wait times as a motivation factor in the construction of at-grade rail crossing separations.

Time savings is the primary benefit from improved traffic flow. Time is spent waiting at at-grade crossings and there are vehicle operating costs due to fuel consumption while idling. A special analysis of time and operating costs due to idling was developed by the Federal Highway Administration as part of its report *2002 Status of the Nation's Highways, Bridges and Transit: Conditions & Performance* (Office of Policy and Government Affairs, 2002). The report showed total time cost for all vehicles delayed at a train crossing as a single-train of a specific length crosses a highway given the AADT and number of lanes in the highway. For the highways in the 4 viaduct projects, the total delay time for the multiple vehicles waiting at a single-training crossing would be between 0.375 and 1.5 hours. Given the number of trains per day, there would be from 3.75 to 15.0 total hours of delay per day by all vehicles at each intersection. Data in the 2002 report also can be used to calculate that 0.836 gallons of gasoline or diesel and 0.054 gallons of motor fuel are used by vehicles for each 1 hour of delay.

Estimates for daily hours of delay can be used to estimate the time costs of delay. Daily hours of delay for each intersection can be multiplied by 365 days per year and by the average cost of delay per vehicle hour of $24.21 (which is the average hourly wage and benefits for individuals at work and 50% of the average hourly wage for individuals at leisure).[5] The annual value of time loss ranges from $33,000 per year for the East Grand Forks, Minnesota project to $133,000 per year at the Grand Forks, Minnesota project. Vehicle operating costs due to idling also should be considered. The daily hours of delay should be multiplied by 365 days per year and the $3.50 per hour vehicle operating costs of idling to yield the total annual vehicle operating (idling) costs.[6] Total annual time and operating (fuel) costs for idling at the four intersections are shown in Table 7.1.

[5] The average hourly vehicle cost of $24.21 is calculated based on information from Highway Statistics 2016 and the 2017 National Household Travel Survey from the United States Department of Transportation and the Occupation Employment Statistics Survey and National Compensation Survey of the United States Bureau of Labor Statistics. Highway Statistics 2016 was used to determine that truck traffic accounts for 12.4% of vehicle-miles on rural non-highway interstates, while the National Household Travel Survey was used to determine that household travel for work using automobiles, light trucks, motor cycles, etc. accounts for 18.21% of the vehicle-miles while leisure travel accounts for the remaining 69.34%. Trucks were assumed to have vehicle occupancy of 1 and the average hourly wage and benefits of truck drivers in North Dakota was $37.33 per hour in 2017. Automobiles and light trucks driven per work were assumed to vehicle occupancy of 1 and the average hourly wage and benefit for all workers was $33.88 per hour in 2017. Automobiles and light trucks driven for leisure have an average occupancy of 1.67 according to the 2017 National Household Travel Survey and an average hourly wage of $23.14. Fifty percent of the hourly wage is used for the travel time of persons away from work and was applied to each passenger.

[6] Average costs of $3/gallon were multiplied by 0.836 gallons of gasoline or diesel per hour of idling and average costs of $18.30/gallon retail for motor oil were multiplied by 0.054 gallons of motor oil per hour of idling.

Table 7.1 Benefit Components and Benefit and Cost Comparisons

City	Construction Costs	Annual Savings: Time and Vehicle Operating Costs	Annual Savings: Reduced Release Incidents	Present Value of Savings	Benefit Cost Ratio
East Grand Forks, MN	$14,931,000	$37,900	$29,300	$833,700	0.06
Moorhead, MN	$30,000,000	$121,400	$151,200	$3,382,400	0.11
Dickinson, ND	$32,400,000	$53,100	$149,900	$2,518,700	0.08
Grand Forks, ND	$21,384,000	$151,700	$33,700	$2,301,600	0.11

Source: Author calculations

Safety benefits arise because the grade separation (viaduct) eliminates release incidents and their costs. Current release incident rates, based on the number of trains and the highway AADT at the current at-grade intersection, are applied to per release incident costs to estimate the annual safety benefits.

As discussed earlier, the combined carrier costs, property damage, response costs, and cleanup costs per rail release incident by trains hauling oil was $125,500. Further, non-fatal injury and fatality costs per rail-highway at-grade interception release incident can be estimated based on Federal Railroad Administration, Office of Safety Analysis data on release incidents, fatalities, and non-fatal injuries from train release incidents and rail-highway release incidents during the 2014 to 2017 period (Office of Safety Analysis, 2018). There were 8,533 release incidents of these types reported by the Office of Safety Analysis during the 2014 to 2017 period. There were 1,022 fatalities for a fatality rate of 11.98%. There were 3,511 non-fatal injuries for a (non-fatal) injury rate of 42.08%.

The value of a statistical life was $9.6 million implying an average fatality costs of $1,149,800 per highway-rail release incident. The non-fatal injury cost per release incident naturally depends on the severity of the set of injuries which occur. Detailed information on the severity of injuries is not available for rail release incidents, so data for injuries in motor vehicle release incidents were utilized instead (Blincoe, et al., 2015). Values for the year 2010 were updated to 2017 using the increase in the value of a statistical life between 2010 and 2017. The social value per non-fatal injury was estimated to be $162,200. This social cost per non-fatal injury implies an average non-fatal injury costs of $68,300 per train alone or highway-rail release incident.

The total cost per rail release incident is $1,396,300, the sum of the non-injury costs ($125,500) of the fatality costs per release incident ($1,149,800), the non-fatal injury costs per release incident ($121,000). As shown in Table 7.1, the estimated annual savings from reduced release incidents ranges from $29,300 per year at the East Grand Forks, Minnesota viaduct to $151,200 per year at the Moorhead, Minnesota viaduct.

Annual delay, operating cost and release incident savings occur over the next 30 years. To compare these benefits with costs, Table 2 also shows the present value of these future savings using a 7% real discount rate and assuming the real value of annual benefits remains steady. The present value ranges from $833,700 in the case of the East Grand Forks, Minnesota viaduct to $3,382,400 in the case of the Moorhead, Minnesota viaduct.

Table 7.1 also includes a benefit to cost ratio for each of the four viaduct projects. The benefit cost ratio is calculated by dividing the present value of project benefits by project construction costs. Benefit cost ratios are low for the projects, ranging from 0.06 to 0.11, which is far below the neutral ratio of 1.0.

Table 7.2 shows how the benefit cost would change with increased future oil by rail activity. Production projections indicate that Bakken oil output would rise by 231,000 barrels per day on average during the 2018 to 2040 period compared to production in 2015. This information is combined with estimates that each 120 car oil train can carry 97,140 million barrels of oil, indicating that there would be approximately 2.5 additional trains per day carrying oil.

Benefit cost information is recalculated in Table 7.2 under the assumption that the number of daily trains rises by 2.5 at each intersection, even though not all trains would utilize each of these four intersections each day. Delay costs would rise proportionally with the number of trains and estimates of the annual probability of an accident also would rise and were recalculated using the website model of the Office of Safety Analysis of the Federal Railroad Administration. The benefit cost ratios in Table 7.2 rise very little compared to those presented in Table 7.1, implying that future increases in oil by rail hauling does not change the underlying benefit cost analyses of these 4 viaduct projects.

Table 7.2 Benefit Components and Benefit and Cost Comparisons with Increased Oil by Rail Shipments

Title	Construction Costs	Annual Savings: Time and Vehicle Operating Costs	Annual Savings: Reduced Release incidents	Present Value of Savings	Benefit Cost Ratio
East Grand Forks, MN	$14,931,000	$47,400	$54,900	$1,270,200	0.09
Moorhead, MN	$30,000,000	$159,300	$172,300	$4,114,600	0.14
Dickinson, ND	$32,400,000	$62,600	$153,400	$2,679,700	0.08
Grand Forks, ND	$21,384,000	$189,700	$67,800	$3,194,900	0.15

Source: Author calculations

Literature Cited

Agresti, A., and M. Kateri. Categorical Data Analysis. In *International Encyclopedia of Statistical Science* (M. Lovric, ed.), Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 206–208.

American Association of Railroads, 2015. *Moving Crude Oil Safely by Rail*. Available at: www.aar.org/BackgroundPapers/Moving Crude Oil Safely by Rail.pdf

Association of American Railroads, 2017: Freight Rail Works. 140,000-Mile Private Rail Network Delivers for America's Economy. http://archive.freightrailworks.org/network/. Accessed Jan. 11, 2017.

Andrews, A, 2014. Crude Oil Properties Relevant to Rail Transport Safety: In Brief. *Congressional Research Service Report*.

Barkan, C. P. L., C. T. Dick, and R. T. Anderson, 2013. Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1825, No. 1, pp. 64–74.

Bates, D., M. Mächler, B. Bolker, and S. Walker, 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Ben-Akiva, M. E., and S. R. Lerman, 1985. *Discrete choice analysis: theory and application to travel demand*. MIT press.

Bilder, C. R., and T. M. Loughin, 2014. *Analysis of categorical data with R*. CRC Press.

Bing, A., A. Dizon, J. Brickett, A. Papson, and L. O'Rourke, 2015. *Risk Evaluation Framework and Selected Metrics for Tank Cars Carrying Hazardous Materials*.

Blincoe, L.J, Miller, T.R., Zaloshnja, E. and Lawrence, B.A., 2015 (May). *The economic and societal impact of motor vehicle crashes, 2010 (revised)*. Report No. DOT HS 218 013, Washington, DC: National Highway Safety Traffic Administration.

Booth, J. G., G. Casella, H. Friedl, and J. P. Hobert, 2003. Negative binomial loglinear mixed models. *Statistical Modelling*, Vol. 3, 2003, pp. 179–191.

Brod, D., G. Weisbrod, C. Williges, S. Jones Moses, D. Gillen, C. Martland, 2013. *Comprehensive Costs of Highway-Rail Grade Crossing Crahses*, National Cooperative Highway Research Program, Report 755.

Cavanaugh, J. E, 1997. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Statistics & Probability Letters*, Vol. 33, No. 2 pp. 201–208.

Christensen, R. H. B. A., 2011. Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package. *Analysis*, Vol. 1, No. 1, pp. 1–18.

Christensen, R. H. B., 2015. Ordinal - Regression Models for Ordinal Data. *R package version*, Vol. 22, pp. 6–28.

City of Dickinson, 2013. *Dickinson 2035 Transportation Master Plan*, Available at Dickinsonplan.com (March).

Class 3-Assignment of Packing Group. 49 Code of Federal Regulations § 173.121, 2010.

Cox, D. R., 1983 Some remarks on overdispersion. *Biometrika*, Vol. 70, No. 1, pp. 269–274.

Energy Information Administration (EIA), 2017a. Layer Information for Interactive State Maps. https://www.eia.gov/maps/layer_info-m.php. Accessed Jan. 11, 2017.

Energy Information Administration, 2017b. Petroleum & Other Liquids. https://www.eia.gov/petroleum/data.php. Accessed Jan. 11, 2017.

Erkut, E., S. A. Tjandra, and V. Verter, 2007. Hazardous materials transportation. *Handbooks in operations research and management science*, Vol. 14, pp. 539–621.

FracTracker. Oil & Gas in the U.S. https://www.fractracker.org/map/national/us-oil-gas/. Accessed Feb. 11, 2017.

Fargo-Moorhead Metropolitan Council of Governments, 2014, *Long Range Transportation Plan,* (July).

Federal Railroad Administration, 2007. Railroad-Highway Grade Crossing Handbook - Safety | Federal Highway Administration. Accessed from https://safety.fhwa.dot.gov/ hsip/xings/com_roaduser /07010/sec03.cfm#b. Retrieved June 2018.

Federal Railroad Administration Office of Safety Analysis, 2016. *FRA Web Accident Prediction System (WBAPS).* Available at *https://safetydata.fra.dot.gov/ webaps/default.aspx*

Federal Railroad Administration, 2018a. *Railroad Safety Overview.* Accessed from https://www.fra.dot.gov/Page/P0010. Accessed June 2018.

Federal Railroad Administration, 2018b. *Web Accident Prediction System (WBAPS).* Accessed from https://safetydata.fra.dot.gov/webaps/default.aspx. Accessed June 2018.

Glickman, T. S., E. Erkut, and M. S. Zschocke, 2007. The cost and risk impacts of rerouting railroad shipments of hazardous materials. *Accident Analysis and Prevention*, Vol. 39, No. 5, pp. 1015–1025.

Iranitalab, A., and A. Khattak, 2017. Comparison of four statistical and machine learning methods for crash severity prediction. *Accident Analysis & Prevention*, Vol. 108, pp. 27–36.

Johnson, Eric M., 2017. "U.S. to Investigate Growing Length of Freight Trains as a Threat to Safety," *Insurance Journal* (December 7th). Available at: https://www.insurancejournal.com/news/national/2017/12/07/473390.htm

Khattak, Aemal and Eric Thompson, 2012. *Development of a Methodology for Assessment of Crash Costs at Highway-Rail Grade Crossings in Nebraska*, Mid-American Transportation Center, Report 25-1121-0001-422

Kringstad, Justin J., 2018. *North Dakota Pipeline Authority Monthly Update,* (September 15). Available at: northdakotapipelines.com/directors-cut/.

Lee, K., M. Boufadel, B. Chen, J. Foght, P. Hodson, S. Swanson, and A. Venosa, 2015. The Behaviour and Environmental Impacts of Crude Oil Released into Aqueous Environments. *The Royal Society of Canada, Ottawa, November*.

Liu, X., M. R. Saat, and C. P. L. Barkan, 2013a. Integrated risk reduction framework to improve railway hazardous materials transportation safety. *Journal of Hazardous Materials*, Vol. 260, pp. 131–140.

Liu, X., M. R. Saat, X. Qin, and C. P. L. Barkan, 2013b. Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. *Accident Analysis and Prevention*, Vol. 59, pp. 87–93.

Lord, D., A. Luketa, C. Wocken, S. Schlasner, T. Aulich, R. Allen, and D. Rudeen, 2015. Literature Survey of Crude Oil Properties Relevant to Handling and Fire Safety in Transport. *Sandia National Laboratories: Albuquerque, NM*.

McFadden, D, 1980. Econometric Models for Probabilistic Choice Among Products. *The Journal of Business*, Vol. 53, No. 3, pp. S13–S29.

Mallela, J., and S. Sadavisam, 2011. *Work Zone Road User Costs: Concepts and Applications*. US Department of Transportation, Federal Highway Administration.

Maty, A.D., 2017. *Field Guide to Tank Cars*. Association of American Railroads.

Minnesota Department of Transportation, 2018. About Railroads in Minnesota, Available at http://www.dot.state.mn.us/aboutrail/

Nachtsheim, C. J., J. Neter, M. H. Kutner, and W. Wasserman, 2004. Applied linear regression models. *McGraw-Hill Irwin*.

Nayak, P. R., D. B. Rosenfield, J. H. Hagopian, A. D. Lillie, and A. Park, 1983. Event Probabilities and Impact Zones for Hazardous Materials Accidents on Railroads.

North Dakota Pipeline Authority, 2017. Oil Transportation Table (June 5). Available at: northdakotapipilines.com/oil-transportation-table/.

North Dakota Department of Transportation, 2017. *2040 North Dakota State Rail Plan,* (November).

Office of Policy and Government Affairs, 2002. “Highway-Rail Grade Crossings,” Chapter 26 in *2002 Status of the Nation’s Highways, Bridges, and Transit: Conditions and Performance,* National Highway Administration. Available at https://www.fhwa.dot.gov/policy/2002cpr/

Office of Safety Analysis, 2018. Total Accidents/Incidents (2017 Preliminary). Available at http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/summary.aspx. Accessed on June 21, 2018.

Oggero, A., R. M. Darbra, M. Muñoz, E. Planas, and J. Casal, 20-6. A survey of accidents occurring during the transport of hazardous substances by road and rail. *Journal of Hazardous Materials*, Vol. 133, No. 1–3, pp. 1–7.

Pipeline and Hazardous Materials Safety Administration, and Office of Hazardous Materials Safety, 2017. Incident Reports Database Search. https://hazmatonline.phmsa.dot.gov/IncidentReportsSearch/ IncrSearch.aspx. Accessed Jan. 11, 2017.

Revzani, Ali, Matthew Peach, Andrew Thomas, Ricardo Cruz, and Walter Kemmsies, 2015. "Benefit-Cost methodologies for highway-rail grade crossing safety protocols as applied to transportation project prioritization processes," *Transportation Research Procedia*, 8: 89-120.

Saat, M. R., C. J. Werth, D. Schaeffer, H. Yoon, and C. P. L. Barkan, 2014. Environmental risk analysis of hazardous material rail transportation. *Journal of Hazardous Materials*, Vol. 264, pp. 560–569.

Specifications for Tank Cars. 49 Code of Federal Regulations § 179, 2012.

Train, K., 2002. Discrete Choice Methods with Simulation. *Cambridge University Press*, pp. 1–388.

U.S. Department of Transportation Federal Highway Administration, 2017. *Railroad-Highway Grade Crossing Handbook,* (August). Available at: https://safety.fhwa.dot.gov/hsip/xings/com_roaduser/07010/sec03.cfm#b

U.S. Energy Information Administration, 2017a. Annual Energy Outlook 2017. Eia.gov/aeo. (January 5).

U.S. Energy Information Administration, 2017b. Petroleum and Other Liquids, (June 21). Available at: https://www.eia.gov/dnav/pet/pet_pnp_cap1_a_(na)_8D0_BpCD_a.htm.

U.S. Energy information Administration, 2017c. North Dakota Field Production of Crude Oil (Thousand Barrels), (September 29). Available at: www.eia.gov/dnav/pet/hist/ LeafHandler.ashx?n=PET&s=mcrfpnd1&f

U.S. Energy Information Administration, 2018. Petroleum and Other Liquids. https://www.eia.gov/dnav/pet/pet_crd_crpdn_adc_mbbl_m.htm. (January 1).

U.S. Energy Information Administration, 2018b. *Movements by Pipeline, Tanker, Barge and Rail between PAD Districts*. Available at https://www.eia.gov/dnav/pet/pet_move_ptb_dc_R20-R10_mbbl_m.htm.

Weldon Cooper Center for Public Service, 2018. 50 States Population Projection. University of Virginia. https://demographics.coopercenter.org/united-states-interactive-map.

Zografos, K. G., and K. N. Androutsopoulos, 2008. A decision support system for integrated hazardous materials routing and emergency response decisions. *Transportation Research Part C: Emerging Technologies*, Vol. 16, No. 6, pp. 684–703.

Zhao, Shanshan, Aemal Khattak and Eric Thompson, 2015. “Safety and Economic A.ssessment of Converting Two-Way Stop-Controlled Intersections to Roundabouts on High Speed Rural Highways,” *Journal of Transportation Research Forum*, 54(1): 131-144.

Appendix A

The *Annual Energy Outlook* provides a forecast for six geographic regions of the continental United States but not by state (U.S. Energy information Administration, 2017a). State level data of past oil production is available from the EIA (U.S. Energy Information Administration, 2018). The research team ran regressions with the state level data against past data for regions in hopes of creating forecasts for individual states. However, since little past data on the forecasted geographic regions is available we had to combine four regions: the Southwest, Dakotas/Rocky Mountains, Gulf Coast, and the Midcontinent to create a region defined by state borders, as there is more historical data on state oil production. This means the three regions with the most production: the Southwest, Dakotas/Rocky Mountains and the Gulf Coast are combined with the Midcontinent to form one central region (labelled the Center region in our regressions).

Regressions that looked at percent change between the region and state or differences failed to produce reliable forecasts. Large states in our regressions would take up all the future production in our models. For example, the regression for North Dakota using a log-log model.

Table A.1 Regression Results for Historic North Dakota Data

	Log(North Dakota)
Constant	-9.96452*
	(4.92859)
Log(Center Region oil production)	1.48902***
	(0.35107)
R-squared	0.346018348
Adjusted R-squared	0.326783594
Number of Observations	36

Standard Errors are reported in parenthesis

*, *** indicates significance at the 90% and 99% levels respectively

North Dakota grows faster than the Center Region and combined with the other large producing states in the Center Region (Texas, New Mexico, Wyoming, Louisiana, Oklahoma, Colorado) takes up all the production predicted by the EIA for the Center Region.

The best forecast approach we have to prevent this problem is to take 2015 data on state level production and 2015 data on the regions and assume states maintain the same share of their region until 2040. As seen in Figure A.1.1 the two models (regression and proportional forecasting) are not very different, but the proportional model prevents the large states like North Dakota from ultimately taking up all the predicted oil production.

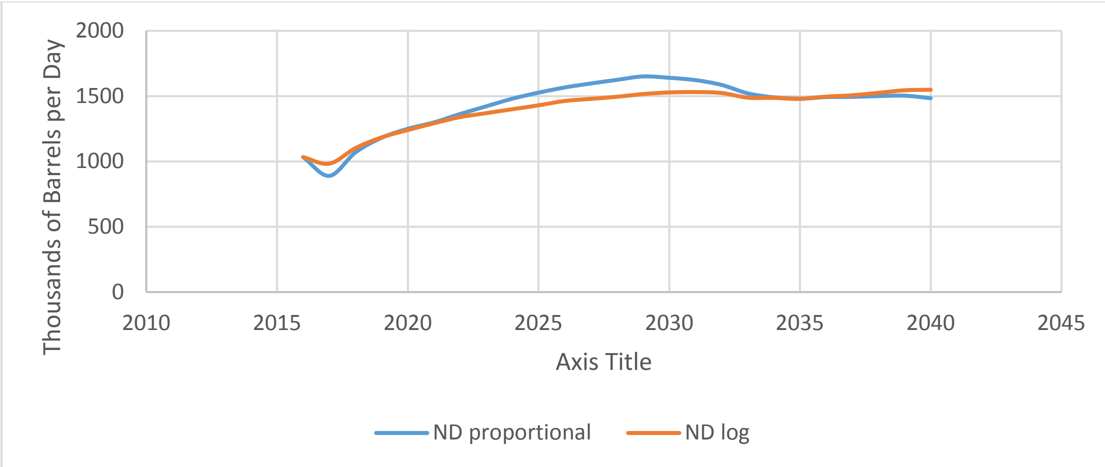

Figure A.1 Oil production. Source: Author estimates.