I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10989

The Accident Externality from Trucking

Lucija Muehlenbachs
Stefan Staubli
Ziyan Chu

SEPTEMBER 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10989

The Accident Externality from Trucking

Lucija Muehlenbachs
University of Calgary and Resources for the Future

Stefan Staubli
University of Calgary, CEPR, IZA, and NBER

Ziyan Chu
Resources for the Future

SEPTEMBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Accident Externality from Trucking*

How much risk does a heavy truck impose on highway safety? To answer this question, we look at the rapid influx of trucks during the shale gas boom in Pennsylvania. Using quasi-experimental variation in truck traffic, we isolate the effect of adding a truck to the road. We find an additional truck raises the risk of a truck accident – and, at an even higher rate, the risk of nontruck accidents. These accidents pose an external cost in cases in which the truck is not found liable, not fully insured, or not directly involved. We show this external cost is capitalized in the insurance market: car insurance premiums of other road users increase when trucks are added to the road.

JEL Classification: G22, H23, I18, Q58, R41

Keywords: externality, trucking, hydraulic fracturing, traffic fatalities

Corresponding author:
Stefan Staubli
University of Calgary
Department of Economics
2500 University Dr. N.W.
Calgary, AB, T2N 1N4
Canada
E-mail: sstaubli@ucalgary.ca

* We are especially grateful to Alan Krupnick for his feedback and support throughout this project. We also thank Kathy Baylis, Dave Byrne, Laura Grant, Sumeet Gulati, Sarah Jacobson, Corey Lang, Shanjun Li, Leslie Martin, Michelle Megna, Tim Moore, Sharon Pailler, Robert Ranieri, Chi Man Yip, seminar participants at the London School of Economics, the University of Calgary, the University of Melbourne, and the University of Verona, and conference attendees of the Association of Environmental and Resource Economics annual meeting. We also thank Carinsurance.com for sharing data on insurance premiums and the Bechtel Foundation and the Smith Richardson Foundation for funding.

1 Introduction

Transportation of goods by heavy truck is ubiquitous. Trucks carry the largest share of goods by weight in the United States, truck driving has become the most common occupation in the majority of US states, and the industry accounts for 1 percent of US gross domestic product.[1,2,3] However, this mode of transportation comes with some risk: although heavy trucks are only 4 percent of all highway vehicles, they account for 12 percent of transportation fatalities.[4]

In this paper, we estimate the accident externality imposed by adding one truck to the road. Vickrey (1968) pointed out that merely by being on the road, regardless of fault, a vehicle imposes an accident externality, since the accident would not have occurred had the driver chosen, for example, to take the train.[5] In the case of trucks, this externality may be even more extreme. By being longer and taller than a car, a truck obstructs drivers' view and could therefore also cause accidents between other vehicles on the road (e.g., a car drives into oncoming traffic while trying to pass a truck). Thus, the presence of a truck makes a road more dangerous for other drivers, but this risk would be not taken into consideration when making the decision of how much to truck.

Furthermore, in accidents involving a truck, the truck's weight, larger frame, height, wheelbase, breaking distance, and rigidity will affect the amount of damage inflicted on the other vehicle. [6] Therefore, additional trucks on the road could translate into not only more accidents but more severe accidents, which would also not be internalized when a shipper chooses how much to transport via truck. However, adding a heavy truck to a road might also lead to behavioral shifts. Similar to Peltzman's (1975) finding that people react to safety regulations (e.g., seatbelts) by driving less carefully, people could compensate for being surrounded by heavy trucks by driving more carefully,

[1] Trucks carried 67 percent, or 13 billion tons, in 2012 (US Department of Transportation, 2013)

[2] This includes delivery trucks. Map: The Most Common Job in Every State, NPR, Planet Money, February 5, 2015. Quoctrung Bui. http://www.npr.org/blogs/money/2015/02/05/382664837/map-the-most-common-job-in-every-state.

[3] Gross-Domestic-Product-(GDP)-by-Industry Data, US Department of Commerce, Bureau of Economic Analysis, 2013

[4] In 2014 there were 3,649 fatalities involving a truck weighing more than 10,000 pounds. In these accidents, only 18 percent of the fatalities were occupants of the large trucks (US Department of Transportation, 2016).

[5] The idea is that regardless of fault, the marginal damage caused by *each* vehicle is the full damage cost; however, current liability regimes recover damages from only the guilty party, making driving too cheap, and resulting in too many cars on the road.

[6] The effect of vehicle weight on safety is a rich area of research, with consequences for corporate average fuel economy standards. See for example, Crandall and Graham (1989); Li (2012); Jacobsen (2013); Anderson and Auffhammer (2014); Bento, Gillingham and Roth (2017). However, these studies are restricted to private vehicles, and do not extend to heavy haulers, which easily exceed 80,000 pounds.

resulting in fewer and less severe accidents. Or, drivers could buy larger cars to protect themselves, resulting in more severe accidents in the long run (Li, 2012).[7] Thus, the effect of an additional truck on the frequency and severity of accidents is ultimately an empirical question.

We identify how an additional truck affects the frequency of both truck accidents and nontruck accidents by exploiting an exogenous shock in truck traffic due to the shale gas boom. The transportation literature has examined the safety risks of heavy trucks, but with an eye toward identifying predictors of accident rates (e.g., safety management practices and driver or company characteristics). The majority of this literature exploits cross-sectional variation (see the survey by Mooren et al., 2014), which means that estimates could be biased if unobserved characteristics, such as road infrastructure, drive both the number of trucks as well as the number of accidents. In constrast, this study examines the more general question of the effect of adding a truck to the road, and use panel data with plausibly exogenous variation in trucks. We also estimate how the accident externalities are capitalized into insurance premiums.

Hydraulic fracturing for natural gas requires shipping vast quantities of water to and from a well site. Water is pumped into the well at high pressure, fracturing the shale rock to release its natural gas, and wastewater, consisting of salty water that flows to the surface with the gas (alongside any fracturing fluid that returns to the surface), is shipped away. The water is primarily transported using tanker trucks.[8] Extracting natural gas from one shale gas well may involve more than 2,000 truck trips.

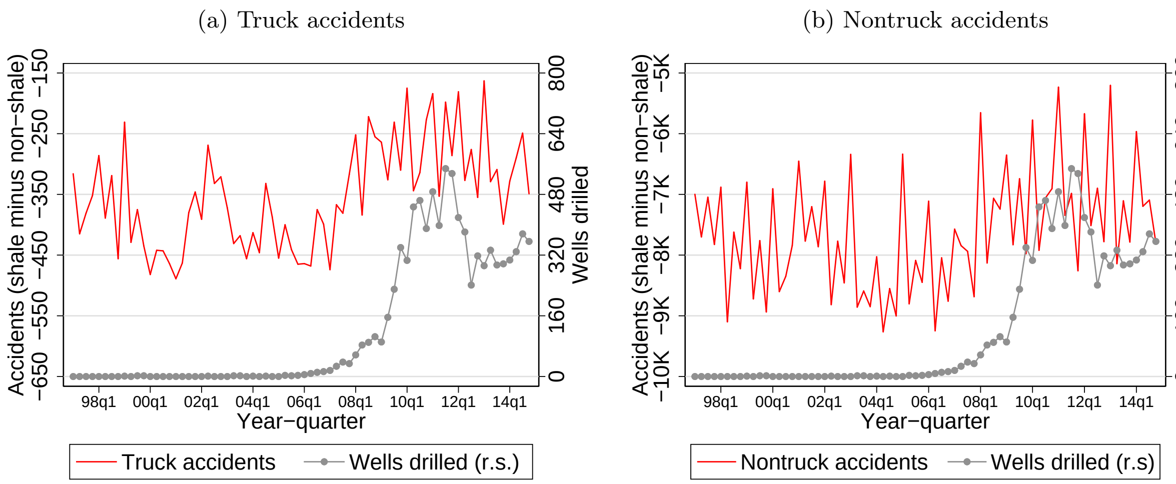
At the county level, one can see that with a shale boom comes an increase in both truck and nontruck accidents. Figure 1 shows the correlation between truck accidents, nontruck accidents, and wells drilled in Pennsylvania (where accident rates are expressed as the difference between counties that at some point in time have a shale well and those that do not). Graham et al. (2015), also using county-level changes in Pennsylvania, show an increase in traffic accidents in counties

[7] In this paper, we examine the short-run effect of a short-term influx of trucks in an area, and therefore do not examine vehicle purchasing decisions.

[8] Transporting water by truck is a costly endeavor and there are moves to transport more water via pipelines. The decision to pipe versus truck depends on water volumes, distances, pipeline right-of-way access, and water quality (IHS Energy Blogger, 2014). Although there is investment in pipelines (e.g., "Energy Firm Makes Costly Fracking Bet–on Water," *Wall Street Journal*, Russel Gold August, 13, 2013), it is still not very commonplace (e.g., "Water Pipelines Mostly a Pipe Dream in the Marcellus," *Pittsburgh Post-Gazette*, Anya Litvak, October 21, 2014). Furthermore, the water pipes transport only fresh water, not wastewater, which is transported via truck. Despite efforts to reuse wastewater, a significant portion is shipped for offsite disposal.

with shale gas wells.[9] However, these estimates are difficult to use to back out the safety risks of an additional truck. An increase in truck traffic could coincide with an unobserved county- or township-specific shock (e.g., a change in population) that is causing the number of accidents to rise. In the case of shale gas development, not only is the number of cars on the road changing, but so is the composition of drivers—specifically, there are more young male drivers (Wilson, 2016), a group that is statistically more likely to be in a crash (Massie, Campbell and Williams, 1995).[10]

Figure 1: Trends in accidents and wells drilled



Notes: Figure plots the difference between the total number of accidents in Pennsylvanian counties with shale and in those without.

To isolate the increase in truck traffic from other idiosyncratic shocks, we zoom in to an analysis at the road level. We use geographic information systems (GIS) to predict the most likely route of the trucks that transport water to and from wells in Pennsylvania, the US state that produces the most shale gas. We exploit the spatial and temporal variation in the location of wells, the location of water withdrawal points, and the location of wastewater disposal points. Hydraulic fracturing is concentrated over a very short period of time, and by knowing the drill date of wells, we can isolate the specific quarter in which a road is used by a large number of trucks. We can compare the rates

[9]Kalinin, Parker and Phaneuf (2017) also find an effect of the shale-boom on truck accidents, albeit through a different channel. Sand is an important ingredient in hydraulic fracturing, which has led to a sand-mining boom in Wisconsin. Examining township-level changes, Kalinin, Parker and Phaneuf (2017) find truck accidents are higher in townships after a sand mine is opened.

[10]We note that the in-migration of workers is smaller in Pennsylvania than other fracking states—Wilson (2016) finds the migration response in the northeastern US was almost eight times smaller than in North Dakota.

of accidents before and after trucks use the road while controlling for the general increase in traffic on all roads in the county. Specifically, we include county-year fixed effects, which capture the overall increase in traffic across all roads in the county-year (i.e., these would capture the increase in young men in the county, who will be dispersed over all roads in the county rather than use only the water-shipment routes). We include road-level fixed effects to get around the problem of selection bias–for example, that trucks might prefer roads with better safety infrastructure.

Using shale gas truck routes provides a unique setting in which a large influx of trucks is concentrated in a small area. If one were to estimate the effect of an observed increase in truck traffic, without knowing the source of the increase, it could be that the trucks are coming from the control roads. In the case of the water-hauling trucks, these are trucks brought to Pennsylvania in response to the rapid boom in shale gas. They are therefore new additions to the road, and are not the result of rerouting. Furthermore, traffic-count data are often spotty and continuously monitored on only very few roads. By predicting the most likely route that trucks will take, we can examine a wider range of roads, from local, neighborhood roads to divided highways.

Using traffic count data, our estimates imply that each shale gas well requires 2,800 truck trips for water withdrawal and wastewater disposal. Using accident data at the county level, we estimate that for each well, in the quarter and county in which it is drilled, there is an additional 0.25 accident involving heavy trucks and an additional 0.84 accident involving other vehicles (a 1 percent and 0.18 percent increase, respectively). This increase in accidents represents the accident risk from drilling a shale gas well in a county; it could be driven by additional trucks on the road, but also by changes in the types of drivers and/or cars on the road. Some counties have hundreds of wells drilled per quarter, and this estimate therefore can be used by county planners to anticipate the consequences of a shale boom for road safety at the county level. The increase in accidents means an increase in the number of fatalities and injuries as well, but since both truck and on-truck accidents increase, the average accident does not appear to be more severe.

Focusing on individual road segments, we obtain an estimate of the risk that a truck *per se* presents. After controlling for changes at the county-year, we do not see an increase in regular traffic on predicted water-hauling routes, but we do see an increase in truck traffic. The road-level estimates suggest that an additional year of truck-miles-traveled results in an additional 0.0001 to 0.0006 truck accident. The impact on nontruck accidents is even larger: an additional year of truck-miles-traveled results in an additional 0.003 to 0.01 accident not involving a truck. But we also find that the type of road matters for safety. Adding a truck to a divided highway with interchanges (ramps) is relatively safe, whereas adding a truck to a neighborhood road, rural road, or city street poses a risk. The highest risk comes from adding a truck to a road interchanges (i.e., a road that is intersected by at-grade driveways and other roads). This suggests that short-haul deliveries (e.g., from online shopping with home delivery) could pose more danger than long-haul truck transport.

The extent to which the accident externality from trucking is internalized depends on the underlying liability regime. In the current legal regime, only the negligent party is held liable for damages. Therefore, the risk of a liability-free accident is external to the decision of how much to truck. This includes accidents in which the truck is involved as well as accidents involving only other road users (i.e., those that occur because the truck is on the road, even though they do not directly involve the truck itself). This accident externality could materialize in the insurance market, making insurance rates higher for everyone on the road.[11] Furthermore, even when the truck is deemed the negligent party, an externality could still exist if the truck is underinsured, in which uninsured damages would be transferred to insurance rates of other drivers.[12] It is arguable that the trucks on US roads are underinsured. The Federal Motor Carrier Safety Administration (FMCSA) regulates the required liability that trucks have to carry. The current required liability was set in 1985 at $750,000.[13] Discussions about raising the limit bring objections from small businesses.[14] The industry is primarily made up of small operators; in 2015 the United States had 550,000 trucking companies, with an average of 20 trucks per company (US Department of

[11] Indeed, Edlin and Karaca-Mandic (2006) examine the hypothesis that the presence of a car on the road poses an externality on insurance rates, and find that states with higher traffic density have higher average insurance costs and rates.

[12] Indeed, uninsured and underinsured drivers pose an externality on insurance premiums (Smith and Wright, 1992; Sun and Yannelis, 2015)

[13] In 2013, a federal bill was introduced to raise the minimum to $4.422 million (H.R. 2730). However, the bill did not pass, and instead an amendment was passed prohibiting any increase to the liability limit during fiscal year 2015 (H.R. 4745); to date the limit remains the same. Several government and industry reports on the frequency with which crashes exceeded the liability limits have reached differing conclusions. A government report found that only 1 percent of the of truck crashes exceed the limit (3,300 of 330,000 total crashes) (US Department of Transportation, 2013), and a report by the American Trucking Association found that only 1.4 percent of accidents exceed $500,000. A report by the Trucking Alliance, however, found the limit was inadequate for 42 percent of the claims (Simpson, 2014).

[14] For a flavor of these concerns, we direct the reader to the comments section of a trucking magazine (reader discretion advised): http://www.overdriveonline.com/fmcsa-current-insurance-minimums-for-carriers-inadequate-new-rule-coming/.

Transportation, 2016).

Therefore, we expect that the underinsurance of catastrophic crashes, combined with an increase in the number of accidents, will lead to higher car insurance rates for those living near heavy truck traffic. We test this using a unique data set of insurance premiums offered by six national carriers for a hypothetical new insurance enrollee. We find that insurance rates increase more in areas exposed to shale gas truck traffic. Specifically, a zip code in the vicinity of a road that is used to transport water for one well increases insurance premiums by $.01 per year. Although small, this is applied to all new enrollees, and some zip codes are near many road segments traversed by many wells. The estimates imply that for every kilometer driven by a truck in a zip code, average annual car insurance premiums of new enrollees will increase by 6 cents for a typical truck's annual kilometers.

Our paper proceeds as follows. Section 2 provides background on shale gas development and describes our data. Section 3 describes our identification strategy of first estimating the impact of shale gas development on traffic counts and then on accident levels, at both the county and road-segment level. Section 4 reports our empirical findings on traffic and accidents and 5 reports our empirical findings on the insurance rates. Section 6 concludes.

2 Background and data

Truck traffic induced by shale gas development has become a major concern for local residents (Theodori, 2009), policymakers (Rahm, Fields and Farmer, 2015), and industry (Krupnick, Gordon and Olmstead, 2013). One shale gas well could require 800 to 2,400 truck trips and 20 wells could be drilled on one well pad (New York State Department of Environmental Conservation, 2011; Abramzon et al., 2014; Gilmore, Hupp and Glathar, 2013). Multiple truck trips are needed to transport equipment, including the drilling rig, pipe to construct the well, and sand used to prop open the water-induced fractures. However, most of the truck trips involve water trucks; 2 million to 4 million gallons of freshwater and fracturing fluids are pumped into each well to create the fractures and 10 to 70 percent of this volume may flow back to the surface, along with formation brine (Veil, 2010). The waste fluids are then collected for reuse, recycling, or disposal.

2.1 Description of data sources

For our analysis, we combine data from several sources and construct two separate samples, one at the county-level and one at the road-segment level. We now describe the different data sources.

Accidents We obtained detailed information on all motor-vehicle crashes in Pennsylvania from the Crash Reporting System (CRS) maintained by PennDOT. We have crash reports from 1997 to 2014 with information on the type of vehicles involved and the latitude and longitude of the accidents. The CRS data set covers more than 2 million crashes, 23,827 of which resulted in one or more fatalities. Importantly, this data set also has information on accidents that did not result in a fatality, which is an advantage over the national Fatality Accident Reporting System (FARS). Accidents must be reported if at least one motor vehicle was involved and there was an injury or death and/or damage to the vehicle that prevented it from being driven. Given that less serious crashes would thus not be reported in the data, we potentially underestimate the crash frequency.

Traffic counts PennDOT also collects data on traffic counts, providing annual truck and nontruck counts from 2004 to 2014. Traffic count data must be handled with caution. Some observations are are imputed, either by repeating the same traffic counts across different years, or after inflating using PennDOT's estimates of population growth.[15] In the years when traffic measured, only a 24-hour snapshot of time is used, and a "day-of-week-by-month" factor is applied to calculate the average daily count for the year. The 24-hour period might not coincide with the quarter that the water truck traffic was the heaviest (discussed later when interpreting the coefficients). In spite of these shortcomings, we nonetheless obtain shale-gas-truck counts that are comparable to estimates reported in the literature.

Shale gas wells We obtained the latitude and longitude of all 8,848 unconventional wells drilled in Pennsylvania as of the end of December 2014 from the Pennsylvania Department of Environmental Protection (PADEP) and the Pennsylvania Department of Conservation and Natural Resources (PADCNR). We have information on the "spud" date (i.e., date that drilling commenced) and the

[15] We exclude observations that appear to be imputed (i.e., when both the count of vehicles and the count of trucks remains exactly the same for more than one year, we keep only the first year; or if both increase but the percentage change in both truck traffic and nontruck traffic are the same..

date drilling was completed. Information on the timing of drilling is important because truck traffic to and from a well is particularly concentrated around the drill date. Most water is used within 45 days of completion and most completion dates are 80 days after the drill date.[16]

Water withdrawal and wastewater disposal points We obtained data from PADEP on the location of approved water withdrawal sources for hydraulic fracturing, including the approval date and the expiration date. In 2009 there were 240 approved withdrawal points, but by 2014 there were 1,124. From PADEP we also know the specific waste disposal location used by each well. Wells are required to report all waste shipments, giving us the universe of shipments.[17] We have 41,625 unique waste shipments from unconventional wells for which we know the location of the well, the location of the disposal point, and the quantity shipped. These shipments were to 233 distinct locations (including industrial waste treatment plants, municipal waste treatment plants, landfills, reuse, and injection disposal wells). The withdrawal and disposal locations in and near Pennsylvania are depicted in Figure 2.[18]

2.2 Construction of truck traffic routes

We use GIS to predict the most likely transportation route that the trucks would take to get from a water withdrawal point to a well and from a well to a wastewater disposal location. We use the TIGER road network from the US Census Bureau. The road network is made up of 630 thousand road segments. We calculate the "least cost" route, in which trucks would take the shortest distance; however, we put penalties on roads with lower speed limits. We assigned impedances on each road depending on the speed limit of the road type.[19] We use the full road network to predict GIS routes, but in our regression analysis we exclude the categories of private roads that are used for

[16] We construct our variables around the spud date and not the completion date because, although spud dates are available for all wells, few have a completion date, even when completed (for only 20 percent of producing wells is a completion date listed).

[17] In the analysis we also include data on the shipments of solid waste (which encompassed 20 percent of the waste shipment data). Assuming 7.3 Bbl/ton of sludge, we converted the solid waste into the same unit as the wastewater.

[18] There are more waste disposal sites even farther away than those depicted in the map. Although some waste is shipped as far as Utah, Michigan, and Idaho, the majority of the waste leaving Pennsylvania goes to Ohio, New York, and West Virginia.

[19] Weighting by typical speed limits of the road types, primary roads and highways were assigned the least impedance of 1, secondary roads were assigned an impedance of 1.18, local roads, 1.86, and trails and private roads, 4.33.

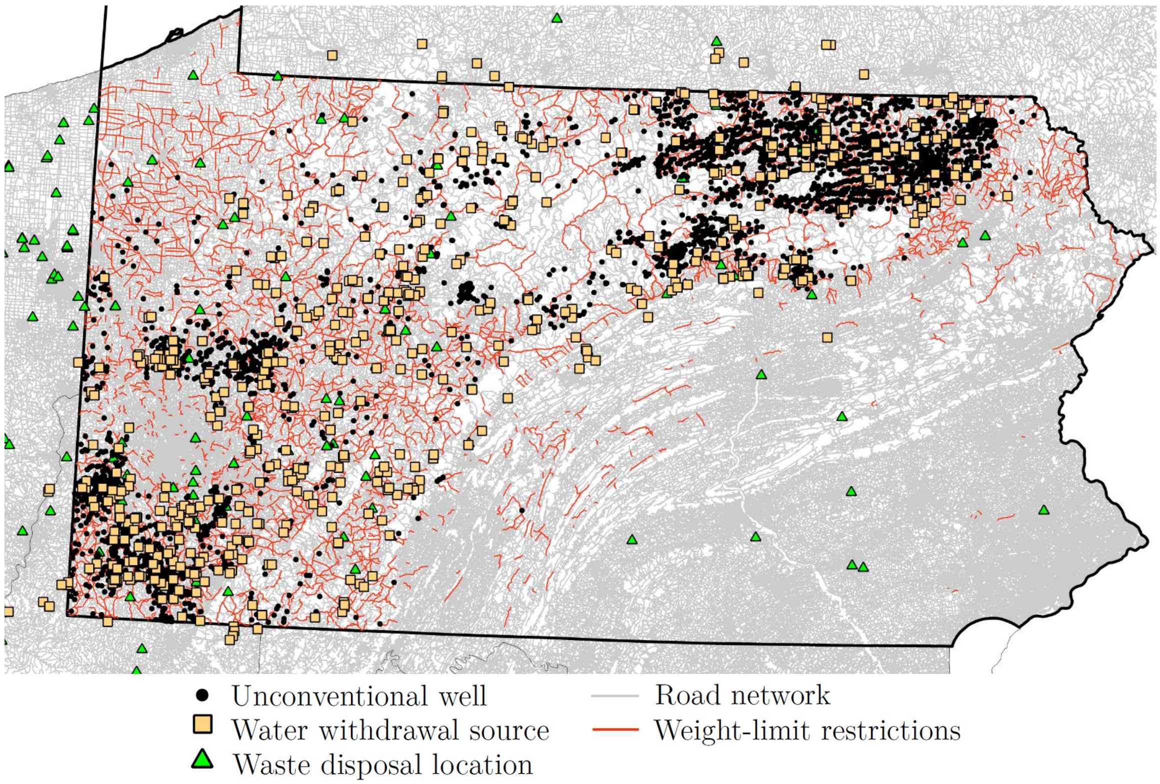


Figure 2: Wells, waste disposal, water withdrawal, and weight-limit restrictions in Pennsylvania, 2014

service vehicles and unpaved dirt trails that require four-wheel drive.[20]

The roads that trucks are allowed to use change over time because roads can be restricted by vehicle weight limits. Communities can protect themselves from the road damage induced by trucks by imposing weight restrictions on certain roads. The posted weight limit is typically 10 tons, and water-hauling trucks are typically over 40. Vehicles weighing more than the posted limit can drive on the roads if they obtain a permit, by providing a security bond that can be used to repair the roads.[21] We obtained data on which segments were posted and/or bonded as well as the start and expiration dates from the Pennsylvania Department of Transportation (PennDOT). Primary highways cannot be weight-limit restricted, but approximately 11,369 miles of secondary roads in Pennsylvania have posted weight restrictions, of which 4,619 have been posted since 2008. We

[20]Including these roads increases the size of our sample by 18% but only .02% of all accidents occur on these roads. A robustness check in the Appendix shows that including these road types has little effect on coefficient estimates (Table A6).

[21]Typical bonds are $6,000 per mile of unpaved road and $12,500 per mile of paved road. As an aside, the estimates of road damages, $13,000 to $23,000 per well, (Abramzon et al., 2014) are less than the bonds.

calculate the different routes for different years, using the road's weight-limit and bonding status at the beginning of the year. We do not allow trucks to traverse weight-limit posted roads, unless the road is listed as bonded. The decision to post a weight limit on a road is based on preventing road damage and not accident risk, or of specific importance to our identification strategy: the weight-limits are exogenous to the expectation of future accident risk. Figure 2 depicts the roads that are weight-limit restricted (i.e., posted and not-bonded) as of the end of our sample period. Interestingly almost all of the posted roads overlie the Marcellus formation (not depicted), indicating the influx of trucks following shale gas development.

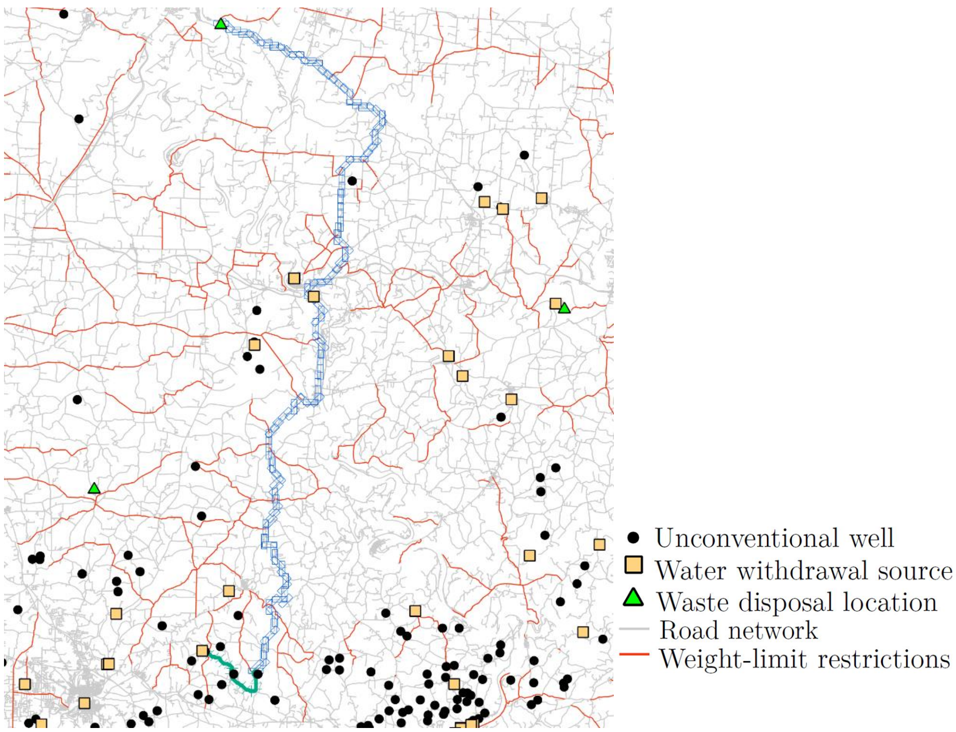


Figure 3: Example route from water withdrawal location to well to waste disposal site

2.3 Counting road use for water withdrawal and disposal

We assume that the wells use the nearest (in least-cost terms) approved water-withdrawal source. Since the water withdrawal data start in 2009, so we assume that the points that were approved in 2009 were also available in earlier years.[22] Because reported completion dates are on average 80

[22]Although not in our data, approvals were also required before 2009, see Abdalla and Drohan (2009). Nonetheless, of the 8,848 wells in our sample, only 507 were drilled before 2009.

days after drilling begins, we only count a road as connecting a well to a water withdrawal source one quarter after the well is drilled (later in the paper, we show evidence that this is a reasonable assumption). To obtain a segment-year-quarter observation, we sum the total number of wells that are predicted to use the segment in the year-quarter.

In the case of waste disposal, we know how much waste was shipped as well as the location where it was shipped. Waste quantities were reported to PADEP annually from 2004 to 2009 and semi-annually from 2010 to 2014.[23] We construct the count of water withdrawal connections to be comparable to the count of waste disposal counts, such that each count signifies one well. The same well can have multiple shipments to different waste disposal locations. We therefore rescale each shipment quantity so that total shipments over the lifetime of a well sum to one. This way, one connection for water withdrawal can be interpreted as the number of trucks needed for one well, and one connection to a waste shipment location can be interpreted as the number of trucks needed to dispose solid and liquid waste from one well.[24]

To investigate spillover effects on nearby roads, we also identify the roads that are within 1m to 500m of a predicted route (as well as 501m to 1,000m), and we create variables for the spatial lags of the predicted routes (if a road is adjacent to two predicted routes, it is assigned the sum of the two).

2.4 Summary statistics

Table 1 reports summary statistics for the county-level and road-level data. The first panel shows that shale counties and nonshale counties are quite different overall. Over the whole sample period, shale counties have on average fewer traffic accidents, fewer fatalities, and fewer injuries than nonshale counties. However, these counties have also much less traffic in general (looking at the segment-average traffic count). Across all years, including nondrilling years, an average of three

[23] To divide the annual data into half-years, we examine the distribution of half-year waste shipments as a function of half-years since the well was drilled. We then divide the annual data into half-years using this empirical distribution (55 percent of the waste is estimated to fall in the first half-year and 45 percent in the second). To disaggregate into the quarter, we divide the half-year observations into equal halves across the quarters. Waste shipment data in 2007 are likely incomplete; there are only 10 percent of the number of observations as there are in 2006. Therefore, we do not include 2007 in our estimation; however, when it is included, our results are qualitatively and quantitatively similar.

[24] Unfortunately, we cannot exploit differences in the weight of the trucks depending on whether they are driving to a water withdrawal site, empty, or returning full to the well because for only a portion of the data do we know the direction of the road of the accident (most accidents are located using latitude and longitude coordinates).

wells were drilled per quarter in shale counties. But this number obscures heavy drilling in some county-year-quarters (e.g., the maximum count in a county-year-quarter is 116 wells). Zooming in to the road-level data (second panel, Table 1), segments are categorized into those that at some point in time are traversed by a truck (for either water withdrawal or water disposal) and all other roads. The shale-truck-traversed roads over the whole sample period have more accidents, fatalities, and injuries, and are overall distinct from all other roads.

Table 1: Summary statistics

	County level			
	Shale counties		**Nonshale counties**	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Quarterly data:				
Heavy truck accidents	17	(20)	36	(31)
All other accidents	305	(484)	696	(710)
Fatalities	4	(4)	7	(7)
Injuries	229	(341)	563	(688)
# Wells	3	(11)	0	(0)
Observations	2,808		2,016	
Annual data:				
Segment-average heavy-truck count	339	(203)	521	(364)
Segment-average non-truck count	3,556	(2,000)	5,201	(3,649)
Total population	134,695	(232,477)	216,556	(276,992)
Observations	286		451	
	Road segment level			
	Roads traversed by		**All other roads**	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Quarterly data:				
Heavy truck accidents	.136	(1.013)	.032	(.551)
All other accidents	.013	(.222)	.001	(.057)
Fatalities	.002	(.058)	.000	(.022)
Injuries	.109	(.902)	.026	(.556)
# Water-withdrawal connections	.022	(.294)	.000	(.000)
# Waste-disposal connections	.335	(2.326)	.000	(.000)
I(Primary road)	.047	(.211)	.018	(.132)
I(Secondary road)	.103	(.303)	.017	(.130)
I(Local road)	.851	(.356)	.965	(.184)
Observations	3,673,440		33,394,392	
Annual data:				
Heavy truck count	704	(1,371)	443	(923)
All other traffic count	5,871	(6,511)	5,796	(7,389)
Observations	107,746		367,602	

Notes: All data are by year-quarter (1997-2014) except traffic counts, which are annual (2004-2014). Shale counties are counties in which at least one shale gas well was drilled over the sample period. "Roads traversed by shale trucks" were used by at least one well to access a water withdrawal or disposal site.

3 Identification strategy

Our identification strategy exploits temporal and spatial variation in the location of shale gas wells, water withdrawal locations, and water disposal locations in the Marcellus Shale region. In a first set of regressions, we estimate the effects of shale gas development on traffic counts, and in a second set of regressions we estimate the effects of shale gas on traffic accidents. The combination of these two outcome variables allows us to rescale the traffic accident estimates into an accident per-additional truck estimate.[25]

We begin by looking at aggregated effects at the county level and then zoom in to examine the road-segment level. The county-level analysis captures both the effect of a change in shale-gas truck traffic as well as any unobservable county-specific shifts from an influx of workers and wealth in the area. On the other hand, the road-segment analysis isolates the pure effect of truck traffic under plausible assumptions. Thus, a comparison of the county and road-segment levels allows us to disentangle these two channels.

In our main specifications, we use data from all counties (or roads) in Pennsylvania. Yet, given the potential differences in observable and unobservable characteristics across shale and nonshale counties (and traversed and non-traversed roads), in the Appendix we restrict the sample to only those counties (or roads) that were at some time affected. We show our results are qualitatively and quantitatively similar. This is perhaps not surprising because in our regressions we include a broad set of fixed effects that account for these differences.

3.1 County-level analysis

The county-level analysis compares outcomes in counties before and after they have shale gas wells, in relation to changes in outcomes in counties that do not. This comparison can be implemented with the following fixed effects regression:

$$y_{ct} = \alpha Wells_{ct} + \lambda_c + \delta_t + \varepsilon_{ct}, \tag{1}$$

[25] This is akin to calculating our own IV-estimate using different samples, but not estimating these together because the traffic counts are measured at the annual level and for many fewer roads.

where $Wells_{ct}$ is the count of shale gas wells drilled in county c at time t, λc are county fixed effects to control for permanent county-specific differences, δ_t are time fixed effects to control for common macroeconomic conditions, and y_{ct} is the outcome variable of interest. In the case of traffic counts, y_{ct} is county c's average daily traffic count over all segments in the traffic-count sample and t signifies year, because traffic data are only reported annually. In the case of traffic safety, y_{ct} is the number of accidents involving a truck, the number of accidents not involving a truck, the number of accidents with a fatality, or the number of accidents with an injury, and t signifies the year-quarter. The main coefficient of interest is α, representing the change in the mean level of the outcome variable from drilling an additional well.

For examining traffic counts, the dependent variable is expressed in levels because one well should increase traffic by the same amount, regardless of the original traffic counts. However, in the case of accidents, we have no conjecture on the functional form. It is plausible for an additional truck to increase the accident rate more on roads that already have a lot of accidents. Indeed, when the residuals of the level-level regression are compared with a log-level regression, the residuals are more normal in the case of a log-level regression and we therefore estimate our accident regressions using a log-level functional form. Specifically, we use the inverse hyperbolic sine (IHS) transformation, which is similar to transforming with a logarithm but allows for observations of zeros.[26] The coefficients can still be interpreted as a percentage change in the outcome variable (Burbidge, Magee and Robb, 1988).[27]

The identifying assumptions are that the locations of the wells and water sites are determined independently from changes in traffic and traffic safety and that there are no spillover effects from treatment counties to control counties.[28] The first assumption is likely satisfied because well location is primarily based on geology, water withdrawal points are based on stream management, and waste disposal locations depend on the chemical concentration of the waste and cost differentials

[26]Allowing for observations of zeros is useful because many road segments do not see any accidents in a year-quarter. In the case of county-level observations, there are no observations of zeros, but for consistency with the segment-level estimates, we use the IHS transformation. In the Appendix, we show that results using a $\log(y+1)$ transformation are qualitatively and quantitatively similar.

[27]The inverse hyperbolic sine is typically defined as follows: $\sinh^{-1}(y) = ln(y + \sqrt{y^2+1})$. A more general version has a scaling parameter; here it is set to one.

[28]Spillover effects from treatment to control counties is a violation of the so-called stable unit treatment value assumption (SUTVA; Rubin, 1980).

across treatment facilities.[29] The second assumption is more critical because water withdrawal and disposal points are not always located in the same county where the well is drilled, thus increasing truck traffic in neighboring counties. Such spillover effects into neighboring counties would lead to a downward bias in our estimates. As described in detail in the next section, spillover effects should be less of a concern in our second identification strategy, in which the unit of analysis is at the road segment.

3.2 Road segment-level analysis

Here we examine 630 thousand road segments s over time t to estimate changes after roads are used by shale gas trucks. We examine the same outcome variables, y_{st}, as in the county-level analysis, first examining the increase in truck traffic counts and nontruck traffic counts, and second, examining the number of accidents involving trucks, not-involving trucks, resulting in a fatality, and resulting in an injury. We estimate:

$$y_{st} = \sum_{d \in D} \alpha_d Withdrawal_{d,st} + \sum_{d \in D} \beta_d Disposal_{d,st} + \lambda_s + \mu_{ct} + \delta_t + \varepsilon_{st} \tag{2}$$

where $D = \{$0m, 1-500m, 501-1000m$\}$.

Treatment variables are constructed from the GIS prediction of the least-cost routes between wells, water withdrawal, and disposal points. $Withdrawal_{0,st}$ is the number of wells that use the segment as a water withdrawal route in t; and $Disposal_{0,st}$ is the number of wells that use the segment as a waste shipment route in t. We designate these counts with the subscript 0 to indicate that they are the counts of wells on the GIS-predicted routes (distance $d = 0$ meters). It is possible that passenger car traffic may try to avoid roads with shale gas truck traffic because of the increased risk of an accident. To examine such avoidance behavior, we construct similar variables to capture being near a predicted route. Specifically, the well counts at distance $d =$ 1-500m are the same well counts as on the predicted route at time t but assigned to any road that is within 1m to 500m

[29] Landowners have some leeway on whether wells will be drilled on their property; in Pennsylvania minerals are typically owned by landowners. We would worry if these owners' decisions about wells depended on their expectations of where future accidents might increase, but this is not likely. The location choice is also determined by the drilling companies; however, these multimillion dollar wells optimize where shale resources are the richest, and "hot spots" of more valuable natural gas liquids in the Marcellus Shale are not uniformly distributed (e.g., see the clustering of wells in Figure 2).

of the predicted route. We expect that the farther the distance from the predicted route, the less influence the truck traffic will have on accident outcomes. These additional regressors are also useful in the case that trucks are not taking the exact route predicted in GIS.

We include road-segment fixed effects, λ_s, to capture time-invariant differences in traffic and accidents across different roads. In the case of the accident outcomes, which are at the quarterly level, we include year-quarter fixed effects, δ_t, to capture statewide seasonal road conditions, and macroeconomic shocks. Importantly, we also include county-by-year fixed effects, μ_{ct}, which capture annual, countywide changes in traffic and accidents. We are assuming that boomtown effects, such as an influx of young male drivers, will affect all roads in the county and will not be concentrated in the particular quarter on the particular road that is predicted to be used by trucks. Then, by including county-by-year fixed effects, we can isolate the effect of the increase in trucks from other boomtown effects. Thus, the main parameter of interest α_d^0 estimates the safety level of a road connecting a well to a water withdrawal point, relative to safety levels when the road did not have this connection, as well as relative to changes in control roads.

4 Results

4.1 Effect on traffic counts

From the county-level analysis we find that drilling a well in the county-year increases daily truck traffic on the average segment by 0.79 truck (Panel A, Table 2). This effect represents a 0.22 percent increase relative of the baseline truck traffic in counties that ever have a well. The estimate implies that each well increases truck counts by 2,798 in the county-year it is drilled.[30] The daily nontruck traffic increases by 7.54 cars for each additional well, or 0.15 percent of the average traffic count in treated counties.

Turning to Panel B, Table 2, we see that when a road is predicted to be used for water withdrawal by a well, the average daily truck count increases by 4.65, implying an estimate of 1,697 trucks per

[30] To get this number, we first multiply the estimated coefficient by the number of segments in a county, which gives us the *county-wide* increase in daily truck traffic per well. We then divide this number by the average number of segments a truck traverses in a county, so as not to double count the same truck traversing more than one segment. Finally, since the estimated coefficient is a *daily* increase, we must multiply by 365 days to get an estimate of the total number of trucks in the year. Because the truck trips are concentrated over a short period of time, we can use the total truck count over the year to be the same as the total truck count over the quarter.

Table 2: Effect on traffic counts

	Heavy-truck count (1)	Nontruck count (2)
Panel A: County level		
Wells	0.79***	7.54***
	(0.27)	(2.29)
Year fixed effects	Yes	Yes
County fixed effects	Yes	Yes
Mean of dep. var. (treated)	354	4,814
R^2	0.39	0.52
N	663	663
Panel B: Segment level		
Withdrawal 0m	46.5*	99.7
	(26.1)	(115.4)
1-500m	21.4	57.2
	(29.8)	(113.8)
501-1000m	60.5	95.7
	(37.9)	(169.2)
Disposal 0m	22.1***	23.1
	(6.4)	(19.3)
1-500m	12.6***	18.3
	(4.1)	(15.5)
501-1000m	1.4	18.1
	(3.1)	(15.1)
Segment fixed effects	Yes	Yes
Year-county fixed effects	Yes	Yes
Mean of dep. var. (treated)	550	5,485
R2	0.73	0.82
N	189,716	189,716

Notes: Dependent variable is the annual average daily truck count (Column 1) and daily nontruck count (Column 2).
Panel A: Observations are by county and year (2004-2014). Wells are the count of wells drilled in the county-year. Robust standard errors are clustered by county.
Panel B: Observations are by segment and year (2004-2014). Withdrawal 0m is the number of wells predicted to use the road segment to reach a water withdrawal source (in counts of 10). We also repeat these counts for the nearby roads (those within 1-500m and within 501-1000m of the treatment road). Disposal 0m is the number of wells (in counts of 10) that are predicted to use the road segment in that year to reach a water disposal location (e.g., landfill, waste water treatment plant). Robust standard errors are clustered by road-segment.
*** Statistically significant at the 1% level; ** 5% level; * 10% level.

well for water withdrawal.[31] Further focusing on the withdrawal routes, we don't see a statistically significant increase on nearby roads (i.e., on roads within 1-500m or 501-1000m of the predicted withdrawal route) and we also do not see an increase of nontruck traffic (column 2), implying that our county-year controls are indeed controlling for the general increase in regular vehicle traffic associated with a shale gas boom. In the case of the water disposal routes, we see an average daily increase of 2.21 trucks on the predicted route, as well as an increase of 1.26 trucks on adjacent roads (within 1-500m of the predicted route). These coefficients would translate to 1,266 trucks per well

[31] The average daily truck count increase is the coefficient on Withdrawal 0m divided by 10. The daily increase is estimated using data generated from random draws of portions of the year and so we must multiply it by 365 days to get an estimate of the total number of truck trips needed for water withdrawal.

for waste disposal (806 on the predicted route and 459 on nearby roads). Including county-year fixed effects also controls for the increase in nontruck traffic, and there is no statistically significant increase of additional cars on the predicted route. We note that the per-well estimate of trucks from the segment-level estimation (2,963 when adding the withdrawal and disposal estimates) is similar to the county-level estimate (2,798) as well as estimates in government reports (2,40) (New York State Department of Environmental Conservation, 2011). Comparing across the segment-level and county-level estimates, we don't expect that we would have different results for truck traffic counts (i.e., controlling for changes in demographics should not affect truck counts); however, we do expect to see differences across accident rates.

4.2 Effect on accidents

Table 3 Panel A shows the effect of an additional well on the frequency of accidents in a county-quarter. Each well results in a 1 percent increase in the number of accidents involving a truck (an additional 0.17 truck accident in the county-quarter), and a 0.18 percent increase in nontruck accidents (an additional 0.55 car accident). By dividing these estimates by the average number of road segments in a county (7,430), we find that the *per segment* number of truck and nontruck accidents increase by 0.000023 and 0.000196, respectively. The increase in accidents seen at the county level is a combination of more trucks on the road, as well as potentially any county-wide changes in the demographics of drivers.

As shown in Panel B, the segment level estimation approach yields smaller estimates. Specifically, we find that the number of accidents involving a truck increases by 0.253 percent when a road is used to bring water to a well and by 0.016 percent when a road is used to ship liquid and solid waste away from a well. The effect of truck accidents on the withdrawal routes is larger than that on disposal routes. This could be driven by the fact that withdrawal routes serve more trucks (as estimated from the traffic regressions) as well as by road type (the disposal routes involve more highways, which we later show are safer, in Section 4.2.1). Multiplying these estimates with the segment average, we find that each well increases the number of truck accidents on a segment by 0.000001 ($= 0.00253 \times 0.0038 + 0.00016 \times 0.0038$).

Similarly, each well increases the number of nontruck accidents by 0.293 percent on water withdrawal segments and by 0.038 percent on water-disposal segments. Nontruck accidents also increase on roads near disposal routes, perhaps reflecting that drivers switch their driving routes to avoid heavy trucks. Together these estimates imply an additional 0.000173 nontruck accidents in the segment quarter. This increase in nontruck accidents suggests that trucks impose an additional externality of causing accidents between other vehicles. Because trucks are not directly involved, these additional accidents would not be internalized in the decision to transport by truck, and could result in higher insurance premiums for all vehicles on the road (explored later).

At the county level, we also see that accidents with fatalities increase by 0.6 percent (an additional 0.02 accident per well). Using a value of a statistical life (VSL) of \$9.1 million, as suggested by the US Environmental Protection Agency, the estimate implies each well imposes a fatality cost of \$202,020. Similarly, accidents with injuries increase by 0.19 percent (or an additional 0.31 accident per well). An upper-bound estimate of the cost of these accidents with injuries is \$338,580 per well.[32] This number is informative for county planners anticipating shale gas development, to anticipate the strain local emergency medical services.

The segment-level results also show an increase in accidents with fatalities, though not a statistically significant increase, possibly because we lack the necessary statistical power; for a given treated road segment, in a year-quarter, fatal accidents are rare (0.00067 accident per year-quarter-segment). We see an increase in accidents with injuries when a road is used for water withdrawal and a smaller decrease on the roads used for disposal. Remember that we see an increase in the number of *both* truck and nontruck accidents, and so the severity of the average crash might not be increasing—and indeed it appears to be constant.[33]

Using the estimates of the absolute increase in accidents as well as the absolute increase in trucks (as predicted from the traffic regressions, Section 4.1), we find that for the average kilometers a truck drives in a year, there are an additional 0.0001 to 0.0006 truck accidents.[34] If each accident reached the current liability limit of \$750,000, the actuarially fair insurance rate would range from \$0.0009/km to \$0.0064/km, both more than an order of magnitude smaller than the current average insurance rates of \$0.057/km.[35]

Similar calculations for the increase in nontruck accidents show that the presence of a truck on the road results in an even larger increase in accidents between other road users. A typical year's worth of truck kilometers will result in an additional 0.003 to 0.01 accidents in which the truck is not directly involved.[36] These are not necessarily severe accidents; at the same time that a

[32]Using an estimate of \$1.1 million, which is for critical injuries including medical costs and lost productivity (Blincoe et al., 2014); this is an upper bound because not all injuries are critical.

[33]In the Appendix we estimate accident severity. By examining the crashes that do occur, we test whether they are more likely to have a fatality or injury if they are on a predicted route. This specification does not find that accident severity is increasing; although we would expect that a truck accident would be more severe, the number of both truck and nontruck accidents is increasing, so on average, severity remains the same (Table A7).

[34]We see $0.00253 \times 0.0038 = 9.6 \times 10^{-6}$ more accidents on withdrawal routes and $0.00016 \times 0.0038 = 6.1 \times 10^{-7}$ more on disposal routes. Dividing by the absolute increase in trucks on the 0m routes (1,697 and 806 trucks) and the average length of a segment (.665km), we get the per kilometer risk. Scaling up by the number of kilometers a truck typically travels in a year, 79,060 km (US Department of Transporation, 2013), a truck results in 0.0001 truck accidents on disposal routes and 0.0006 truck accidents on withdrawal routes.

[35]The 2015 average insurance rate includes liability and cargo premiums (American Transportation Research Institute, 2016). Companies with larger fleets pay lower insurance (by self-insuring, using higher deductibles, and relying on umbrella policies); for example, companies operating more than 1,000 trucks pay 2.8 cents per km, whereas companies with fewer than 5 trucks pay 7.5 cents per km.

[36]Using coefficients on the 0m predicted routes for withdrawal and disposal respectively.

year's worth of truck driving increases the number of accidents with a fatality by 0.00002 (though statistically insignificantly different from zero), it increases the number of accidents with an injury by 0.003.[37] Calculating the monetary effect of these additional fatalities and injuries results in an externality estimate of $3,210 per truck year. Although this does not include an estimate of property damages, it is comparable in size to the CO_2 emissions costs per truck-year ($3,406).[38]

Table 3: Impact on accident frequency

	IHS(truck) (1)	IHS(vehicle) (2)	IHS(fatal) (3)	IHS(injury) (4)
Panel A: County level				
Wells	.0101***	.0018***	.0060***	.0019***
	(.0017)	(.0006)	(.0008)	(.0005)
County fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	17	305	3.7	162
R^2	.90	.99	.70	.98
N	4,824	4,824	4,824	4,824
Panel B: Segment level				
Withdrawal 0m	.0253***	.0293***	.0021	.0116**
	(.0055)	(.0056)	(.0019)	(.0048)
1-500m	.0005	.0019	-.0001	.0007
	(.0006)	(.0016)	(.0003)	(.0012)
501-1000m	.0010	.0014	.0004	.0016
	(.0007)	(.0017)	(.0004)	(.0012)
Disposal 0m	.0016***	.0038***	-.0002	-.0028***
	(.0006)	(.0011)	(.0002)	(.0010)
1-500m	.0000	.0013***	.0000	.0003**
	(.0001)	(.0002)	(.0000)	(.0002)
501-1000m	.0001	.0018***	.0000	.0009***
	(.0000)	(.0003)	(.0000)	(.0002)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	.0038	.048	.00067	.026
R^2	.42	.65	.10	.60
N	34,493,677	34,493,677	34,493,677	34,493,677

Notes: Dependent variables are the inverse hyperbolic sine (IHS) transformation of the count of accidents with a truck (Column 1), count of accidents not involving any trucks (Column 2), count of accidents with one or more fatalities (Column 3), and count of accidents with one or more injuries (Column 4).
Panel A: Observations are by county-year-quarter (1997-2014). Wells are the count of wells drilled in the county-year-quarter. Robust standard errors are clustered by county.
Panel B: Observations are by segment-year-quarter (1997-2014). Withdrawal 0m is the count of wells in the year-quarter that are predicted to use the road segment to connect to a water withdrawal source (in counts of 10). We repeat these counts on nearby roads (within 1-500m and 501-1000m). Disposal 0m is the count of wells (in counts of 10) in the year-quarter that are predicted to use the road segment to connect to a disposal location (e.g., landfill, waste water treatment plant). Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

[37]Summing the coefficients across withdrawal and disposal routes, even though statistically insignificant in the case of fatalities.

[38]Using the Federal Highway Administration's average fuel economy for heavy trucks (5.3 miles per gallon of gasoline equivalent), 0.0088 tons of CO_2 emissions per gallon, and a social cost of carbon dioxide of $42/ton.

4.2.1 Effect on accidents by road type

One advantage of predicting traffic routes is that we can examine a more extensive set of roads than had we relied on traffic count data. This allows us to examine heterogeneity across roadtypes.

Table 4: Impact on accident frequency by road type

	Primary	**Secondary**	**Local**	**Primary**	**Secondary**	**Local**
	(1)	(2)	(3)	(4)	(5)	(6)
	IHS(truck accidents)			IHS(Nontruck accidents)		
Withdrawal 0m	-.0474*	.0936***	.0051***	-.0046	.0689***	.0178***
	(.0262)	(.0206)	(.0014)	(.0274)	(.0211)	(.0036)
1-500m	.0057	-.0041	.0002	-.0328	.0228	.0016
	(.0102)	(.0107)	(.0002)	(.0213)	(.0217)	(.0012)
501-1000m	.0049	.0130	.0003	.0531**	.0376	-.0009
	(.0122)	(.0156)	(.0003)	(.0259)	(.0346)	(.0010)
Disposal 0m	-.0004	.0055**	.0006***	.0060**	-.0032	.0056***
	(.0016)	(.0022)	(.0002)	(.0027)	(.0035)	(.0011)
1-500m	-.0007	-.0005	.0001**	-.0015	-.0011	.0017***
	(.0007)	(.0013)	(.0000)	(.0017)	(.0030)	(.0002)
501-1000m	-.0002	.0030*	-.0000	.0076	.0063	.0016***
	(.0017)	(.0016)	(.0000)	(.0051)	(.0040)	(.0003)
Segment fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	.076	.043	.00036	.38	.72	.013
R^2	.59	.34	.06	.63	.80	.37
N	711,540	883,194	32,898,943	711,540	883,194	32,898,943

	Primary	**Secondary**	**Local**	**Primary**	**Secondary**	**Local**
	(1)	(2)	(3)	(4)	(5)	(6)
	IHS(fatal accidents)			IHS(accidents with injuries)		
Withdrawal 0m	-.0059	.0108	.0001	-.0018	.0347*	.0072***
	(.0129)	(.0080)	(.0005)	(.0221)	(.0201)	(.0026)
1-500m	-.0047	-.0036	.0001	-.0232	.0215	.0002
	(.0034)	(.0038)	(.0002)	(.0165)	(.0168)	(.0008)
501-1000m	.0087	.0022	.0002	.0348*	.0551**	-.0010
	(.0083)	(.0062)	(.0003)	(.0188)	(.0244)	(.0007)
Disposal 0m	-.0000	-.0008	-.0000	-.0003	-.0125***	.0021**
	(.0006)	(.0008)	(.0001)	(.0025)	(.0033)	(.0009)
1-500m	-.0001	.0008	.0000	-.0007	-.0026	.0007***
	(.0002)	(.0007)	(.0000)	(.0013)	(.0025)	(.0002)
501-1000m	-.0001	.0001	.0000	.0026	.0050	.0007***
	(.0003)	(.0007)	(.0000)	(.0039)	(.0033)	(.0002)
Segment fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	.0045	.012	.00015	.21	.4	.0069
R^2	.16	.12	.03	.58	.74	.29
N	711,540	883,194	32,898,943	711,540	883,194	32,898,943

Notes: Columns in each panel represent separate regressions on subsamples based on road type (primary, secondary, and local roads). Dependent variables are the inverse hyperbolic sine (IHS) transformation of counts of accidents with one or more trucks (Panel A, Columns 1-3), accidents no involving truck vehicles (Panel A, Columns 4-6), accidents with one or more fatalities (Panel B, Columns 1-3), and accidents with one or more injuries (Panel B, Columns 4-6). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table 4 shows that adding trucks to primary roads could be deemed relatively safe—these are divided, limited-access highways, with interchanges. Adding a truck to a primary road does not appear to increase accident frequency—we even estimate a reduction in truck accidents, statistically significant at the 10 percent level. We do estimate an increase in nontruck accidents on disposal routes, but the coefficient is very small. In contrast, secondary roads are the most dangerous. These are main arteries that have one or more lanes of traffic in each direction (which may or may not be divided) and at-grade intersections with many other roads and driveways. The secondary roads are about two times more dangerous for trucks than local roads, which include local neighborhood roads, rural roads, and city streets (usually a single lane in each direction).

4.3 Robustness

4.3.1 Differing treatment dates

Our identification strategy rests on the assumption that after the inclusion of fixed effects, road segments that are not used by shale gas trucks in the county-quarter are a good counterfactual for treated roads in the absence of treatment. To shed light on this assumption, we include a set of lead and lagged variables, designated by τ. The lead effects will detect whether treatment and control road segments have differential trends prior to commencement of drilling (in the case of withdrawal) or shipping (in the case of waste disposal). The estimated coefficients on the lead effects should not be significantly different from zero if the common-trend assumption is valid. The lagged effects will measure how the effect of truck traffic on accidents persists after the drilling (or shipping).

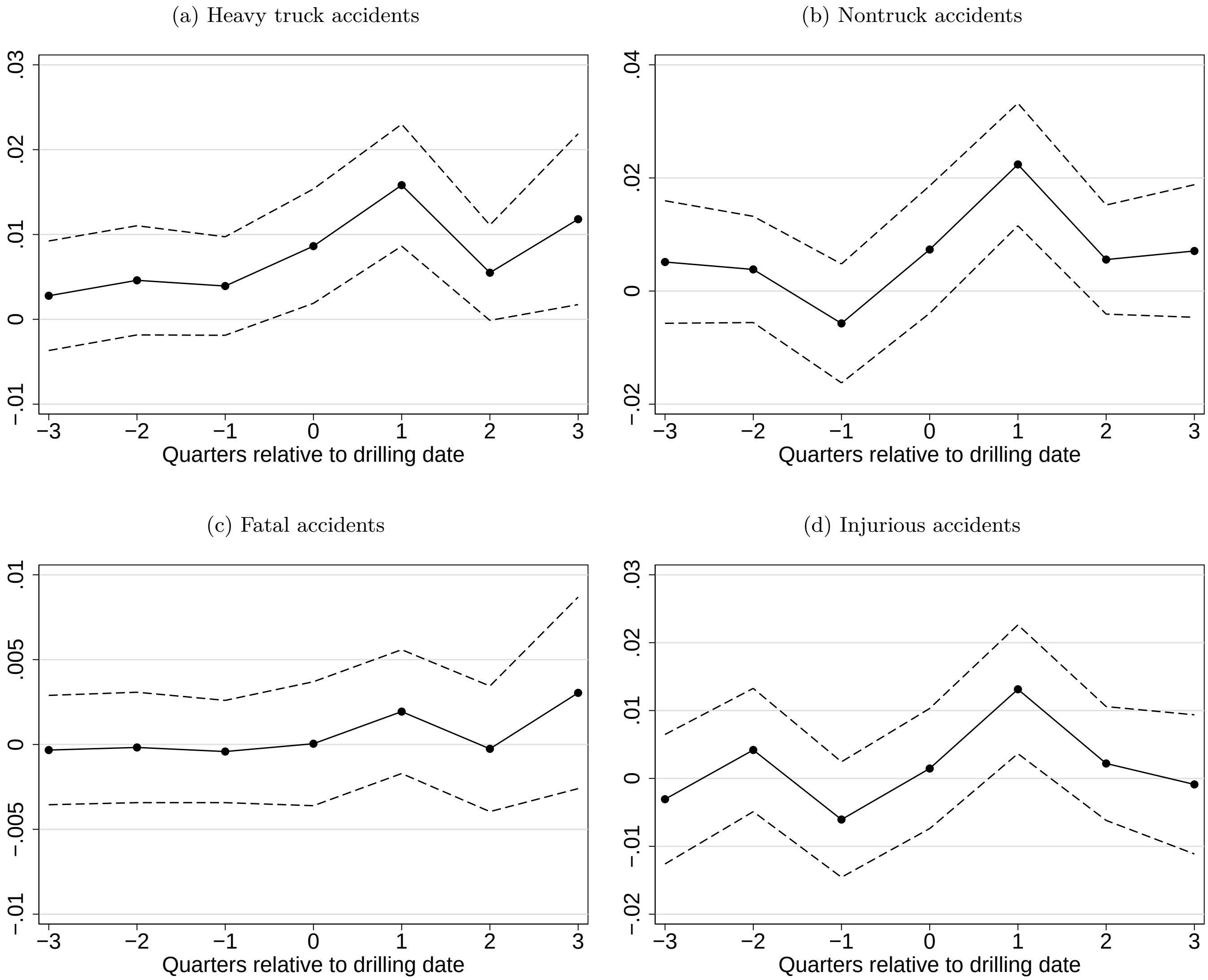
$$y_{st} = \sum_{\tau=-3}^{3} \sum_{d \in D} \alpha_d^\tau Withdrawal_{d,st}^\tau + \sum_{\tau=-3}^{3} \sum_{d \in D} \beta_d^\tau Disposal_{d,st}^\tau + \lambda_s + \delta_t + \mu_{ct} + \nu_{rt} + \varepsilon_{st} \quad (3)$$

where $D = \{$0m, 1-500m, 501-1000m$\}$

Figure 4 shows evidence that most of the water withdrawal happens the quarter after the well is drilled. Plotting the effect on the water withdrawal roads (α_0^τ) accident rates are the highest in the quarter after drilling (the quarter used in our main estimating equation for water withdrawal).[39] Although we know the waste shipment date, for completeness we also include leads and lags around the waste shipment date. All are statistically insignificant and small, with the exception of coefficients on nontruck accidents, which are statistically significant but almost an order of magnitude smaller than what we see on the withdrawal routes. This could be because one well can spread its shipments over many years, and so our waste shipment treatment is not as strong. Waste shipment coefficients are plotted in Appendix Figure A1 on the same axes as the withdrawal figures.

[39] Using the notation of equation (3) the variables in our main estimating equation (2) would be: $Withdrawal_{d,st}^1$, for the quarter after the well was drilled, and $Disposal_{d,st}^0$, for the quarter that the waste was shipped.

Figure 4: Effect of water-withdrawal on accident frequency



Notes: Coefficients on the leads and lags around the drill date of wells are plotted for the roads connecting a well to a water withdrawal location. Coefficients are from equation (3), α_0^τ. Each figure is from a different regression of the following outcome variables: (a) the count of accidents involving a truck; (b) the count of nontruck accidents; (c) the count of accidents with a fatality; and (d) the count of accidents with injuries. The dashed line represents standard errors.

Table 5: Placebo test: Fictitious treatment dates

	IHS(truck) (1)	IHS(vehicle) (2)	IHS(fatal) (3)	IHS(injury) (4)
Panel A: County level				
Wells	-0.0001	-0.0001	0.0010	-0.0002
	(0.0010)	(0.0004)	(0.0013)	(0.0005)
County fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	19	266	3.5	146
R^2	0.91	0.99	0.71	0.98
N	2,680	2,680	2,680	2,680
Panel B: Segment level				
Withdrawal 0m	-.0022	-.0078	.0005	-.0111***
	(.0027)	(.0042)	(.0016)	(.0042)
1-500m	.0002	-.0003	-.0002	-.0009
	(.0005)	(.0014)	(.0003)	(.0011)
501-1000m	-.0004	.0003	.0004	-.0001
	(.0004)	(.0012)	(.0003)	(.0011)
Disposal 0m	-.0009*	-.0046***	.0000	-.0060***
	(.0005)	(.0009)	(.0002)	(.0008)
1-500m	-.0000	.0000	.0000	-.0003*
	(.0001)	(.0002)	(.0000)	(.0002)
501-1000m	.0001	-.0000	.0001**	.0001
	(.0001)	(.0002)	(.0000)	(.0001)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	.0039	.044	.00068	.026
R^2	.45	.70	.12	.64
N	20,078,409	20,078,409	20,078,409	20,078,409

Notes: Specifications are the same as in Table 3 but treatment variables are given fictitious dates (specifically, all treatment variables are recoded to have occurred 8 years prior). Sample therefore covers 1997-2006. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

4.3.2 Placebo regression

We further test the identifying assumption that trends in accidents in control counties and control roads are good counterfactuals for treatment counties and treatment roads. Although including leads in equation (3) alleviates the concern that control and treatment counties have differential trends, to do so we are including an additional 12 variables in the estimating equation. Here we perform an additional test for differential trends in accidents prior to shale gas drilling. Most wells were drilled between 2007 and 2014. We recode the observations so that, falsely, the wells were drilled eight years earlier, and we run the same regressions using the data from 1997 to 2006. Most of the coefficients in Table 5 are statistically insignificant, and when significant, they are negative.

5 Valuation of the external costs to car insurance customers

The previous sections provide evidence that adding one truck to a road creates an accident externality: the number of accidents involving a truck, as well as those not-involving a truck, increase. The increase in accident risk, particularly if the trucks are underinsured, would imply that we should see an increase in insurance premiums. Here we estimate the external cost borne by all new insurers.

From CarInsurance.com, an online resource for consumers to find and compare car insurance policies, we obtained a unique data set of zip code level insurance rates available to the same hypothetical individual in 2012 and 2014. Specifically, the auto insurance quotes come from six large carriers, Allstate, Farmers, GEICO, Nationwide, Progressive and State Farm, and are based on insurance for a new Honda Accord driven by a single 40-year-old male who commutes 12 miles to work each day and has a clean driving record and good credit.[40] Importantly, the data are quotes for the same hypothetical person, which is an advantage over using population-average data on insurance premiums, in which any change in premiums could be driven by changes in the demographics of the drivers.

The data include all zip codes in Pennsylvania. We sum the GIS-predicted routes to obtain the total number of water connections within 25km of the centroid of a zip code (specifically, we sum, separately, the total Withdrawal 0m and Disposal 0m counts). According to a survey, 77 percent of accidents occur within 24km of one's residence.[41]

Table 6: Summary statistics at the zip code level

	Traversed zip codes		**Nontraversed zip codes**	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Average premium ($)	1076.3	(87.2)	1481.9	(472.2)
Δ in premium between 2012 and 2014 ($)	63.7	(30.2)	30.5	(84.0)
Withdrawal 0m (zip total in 10s)	22	(40)	0	(0)
Disposal 0m (zip total in 10s)	708	(1107)	0	(0)
Observations	2433		872	

Notes: Data are by zip code for 2012 and 2014. Average premium (dollars) is the zip code average quote obtained from six national insurance carriers for the same hypothetical 40-year-old male driver of a Honda Accord. Traversed zip codes are zip codes that have at have had at least one withdrawal or disposal connection over the sample period.

[40] Rates are for policy limits of 100/300/50 ($100,000 for injury liability for one person, $300,000 for all injuries and $50,000 for property damage in an accident) and a $500 deductible on collision and comprehensive coverage, including uninsured motorist coverage.

[41] https://www.progressive.com/newsroom/article/2002/may/fivemiles/

Across both years of data, the average insurance premium in Pennsylvania is lower in zip codes that are traversed by trucks (a $1,076 annual average premium compared with $1,482). The traversed zip codes saw a larger increase in premiums between the two years (a $64 increase versus a $31 increase). Similar to the preceding sections, we include all zip codes, but in the Appendix we show that results do not differ if we restrict our sample to only the traversed zip codes (Table A3).

We run a regression in which we regress the zip code's average insurance premium on the number of wells that use road segments near the zip code. The regression includes zip-code fixed effects to capture permanent level differences across zip codes, and year fixed effects to capture the general increase in premiums across the state.

The results show that adding one segment that is used as a withdrawal route from one well increases the average annual premium by 0.001 percent, or by 1 cent. In the case of waste disposal routes, the increase is even smaller, at 0.00002 percent or 0.02 cents. These are small increases compared with the general increase across the two years ($64, Table 7). However, when a zip code is traversed, it is often not only by one segment used by one well; the most heavy traversed zip code is estimated to see a $67 increase.[42] Furthermore, remember that this increase is applied to all drivers, and therefore even in the zip codes affected by a few cents, the effect is multiplied by all new insurees.

A year's worth of trucking in the vicinity of the zip code, will increase insurance premiums of new enrollees by $0.06.[43] For comparison, Edlin and Karaca-Mandic (2006) estimate that an additional car in a state will increase average insurance premiums by $0.00036 to $0.0014. Our larger estimate could arise because our estimate is concentrated within a zip code rather than dispersed across a state; because insurance plans for new enrollees could adjust faster than insurance plans of the existing population; and because a truck poses more risk by function of its size and typical kilometers traveled.

Although using the count of traversed-road-segments avoids some of the difficulties that arise in

[42] The most heavily traversed zip code has 3,640 withdrawal connections and 74,750 disposal connections

[43] To get the per truck-year increase in premiums, we first divide the coefficients by the number of trucks per segment (1,697 in the case of withdrawal trucks and 806 in the case of disposal trucks, on the 0m segments) and the number of kilometers in a segment (0.665km per segment). Note that to reach a water withdrawal point, one well on average traverses 11 segments in 25 km of the zip-code centroid; and to reach a disposal point, one well on average traverses 41 segments in 25km of the zip code. Therefore, so as not to count the same truck in the zip code more than once, we also divide the estimated coefficients by the number of segments the trucks traverse. Then we multiply by the annual average of kilometers traveled, 79,060 km. Combining withdrawal and disposal estimates we calculate a 6 cent increase=$0.01/11/1,697/0.665 \times 79,060 + 0.0002/41/806/0.665 \times 79,060$.

Table 7: Car insurance premiums on drilling activity

	Average premium	Vehicle theft
	(1)	(2)
	IHS(dollars)	IHS(count)
Withdrawal 0m (total, in 1000s)	.0118***	.2372
	(.0030)	(.2614)
Disposal 0m (total, in 1000s)	.0002***	-.0002
	(.0001)	(.0136)
Zip code fixed effects	Yes	No
Year fixed effects	Yes	Yes
County fixed effects	No	Yes
R^2	.99	.97
N	3,305	134
Mean of dep. var. (treated, in levels)	1,076	81

Notes: The dependent variable in column (1) is the inverse hyperbolic sine (IHS) transformation of the average insurance premium offered across six national insurance providers for the same hypothetical new insuree. Observations are by zip code and year (for the years 2012 and 2014). The dependent variable in column (2) is the IHS transformation of the number of vehicle thefts. Observations are by county and year (2012 and 2014). Variables are the total number of segments in 25km of the zip code centroid (or in the county in the case of thefts) connecting wells to withdrawal locations or disposal locations (in counts of 1,000). Robust standard errors are clustered by zip code (or county). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

the county-level estimates based on wells, unlike the segment analysis, we are still not controlling for general demographic changes that are correlated with the increase in traversed segments.[44] One concern is that we expect to see higher insurance premiums in areas with more vehicle theft, and previous literature shows that indeed, vehicle theft in shale-rich counties is higher during a boom (James and Smith, 2016). Looking at the same two years, we also examine how vehicle thefts are related to the number of traversed segments. We obtained data on the number of vehicle thefts at the county level from Pennsylvania's Uniform Crime Reporting System. The number of traversed roads is statistically insignificant in determining vehicle theft, albeit large in magnitude.

6 Conclusion

Internalizing the external cost that a truck imposes on other road users would require an ambitious revamping of the current liability regime. We estimate that the addition of a single truck to the road not only increases the number of accidents involving a truck, but also, because the presence of the truck makes the road less safe, increases the number of accidents between other road users. Although an insurance system has the potential to internalize accidents in which a truck is directly

[44]Population controls would be "bad controls," because they would change with our treatment variable.

involved, there are no mechanisms in place that would internalize the increase in accidents of other road users. And even when a truck is directly involved in the accident, the current insurance market does not necessarily internalize the external cost.

If the other vehicle is the negligent party, the truck will not be held responsible for damages. Although this study was not designed to estimate the additional cost of hitting a truck versus a car, Anderson and Auffhammer (2014) present clear evidence that striking a vehicle of heavier weight imposes more damage. Therefore, negligent drivers who have the misfortune to hit a truck are then responsible for larger damages—and these additional damages would not be considered in a firm's decision of how much to truck.

If the truck is the negligent party, the current liability limits are low enough to allow firms to be judgment proof. Trucks must carry insurance, or post a surety bond, to cover accidents costing $750,000, a limit that has not grown with inflation over the past 30 years. If an accident costs more than the liability limit and is also more than the trucking company's assets, the possibility of bankruptcy would mean that the accident costs would again not be internalized.

We find evidence that these external costs affect the insurance premiums of all drivers on the road. The information technology revolution could provide means to tax heavy trucks for the kilometers that they drive. However, ideally, this tax would be by road-type kilometers driven, because we find the risks vary by what type of road the truck drives on: highways with interchanges are the least dangerous, highways without interchanges the most, and local roads are moderate. With automated, driverless trucks, keeping track of the kilometers driven by road-type would not be arduous, and even a kilometer-by-road-type tax could become feasible. With such a tax, the decision of how much to truck and on which roads, would then be made in consideration of the external accident costs. While it is possible that automated trucking will reduce the frequency of truck accidents, it would not necessarily reduce other accident externalities. This paper points to the large externality of additional accidents, not directly involving trucks, occurring when trucks are on the road. Without similar safety advances in the other vehicles on the road, crashes in the proximity of automated trucks will remain a large externality associated with trucking.

References

Abdalla, Charles, and Joy R. Drohan. 2009. "Water Withdrawals for Development of Marcellus Shale Gas in Pennsylvania, Pennsylvania State University, Marcellus Education Fact Sheet."

Abramzon, Shmuel, Constantine Samaras, Aimee Curtright, Aviva Litovitz, and Nicholas Burger. 2014. "Estimating the consumptive use costs of shale natural gas extraction on Pennsylvania roadways." *Journal of Infrastructure Systems*, 20(3): 06014001.

American Transportation Research Institute. 2016. "An analysis of the operational costs of trucking: a 2016 update."

Anderson, Michael L, and Maximilian Auffhammer. 2014. "Pounds that kill: the external costs of vehicle weight." *The Review of Economic Studies*, 81(2): 535–571.

Bento, Antonio, Kenneth Gillingham, and Kevin Roth. 2017. "The Effect of Fuel Economy Standards on Vehicle Weight Dispersion and Accident Fatalities." NBER Working Paper.

Blincoe, Lawrence J., Ted Miller, Eduard Zaloshnja, and B.A. Lawrence. 2014. "The economic and societal impact of motor vehicle crashes, 2010, Report No. DOT HS 812 013, US Department of Transportation, National Highway Traffic Safety Administration Washington, DC."

Burbidge, John B, Lonnie Magee, and A Leslie Robb. 1988. "Alternative transformations to handle extreme values of the dependent variable." *Journal of the American Statistical Association*, 83(401): 123–127.

Crandall, Robert W, and John D Graham. 1989. "The effect of fuel economy standards on automobile safety." *Journal of Law and Economics*, 97–118.

Edlin, Aaron S, and Pinar Karaca-Mandic. 2006. "The accident externality from driving." *Journal of Political Economy*, 114(5): 931–955.

Gilmore, Kevin R, Rebekah L Hupp, and Janine Glathar. 2013. "Transport of hydraulic fracturing water and wastes in the Susquehanna river basin, Pennsylvania." *Journal of Environmental Engineering*, 140(5): B4013002.

Graham, Jove, Jennifer Irving, Xiaoqin Tang, Stephen Sellers, Joshua Crisp, Daniel Horwitz, Lucija Muehlenbachs, Alan Krupnick, and David Carey. 2015. "Increased traffic accident rates associated with shale gas drilling in Pennsylvania." *Accident Analysis & Prevention*, 74: 203–209.

IHS Energy Blogger. 2014. "Haul or transfer? Water transportation strategies for unconventional oil and gas development." *IHS Markit.*

Jacobsen, Mark R. 2013. "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior." *American Economic Journal: Applied Economics*, 5(3): 1–26.

James, Alexander, and Brock Smith. 2016. "There will be blood: crime rates in shale-rich US counties." *Journal of Environmental Economics and Management.*

Kalinin, Alexey, Dominic Parker, and Daniel Phaneuf. 2017. "Effects of Local Regulation on Neighboring Jurisdictions: Evidence from Mining Ordinances."

Krupnick, Alan, Hal Gordon, and Sheila Olmstead. 2013. "Pathways to Dialogue: What the Experts Say about the Environmental Risks of Shale Gas Development." Resources for the Future.

Li, Shanjun. 2012. "Traffic safety and vehicle choice: quantifying the effects of the 'arms race' on American roads." *Journal of Applied Econometrics*, 27(1): 34–62.

Massie, Dawn L, Kenneth L Campbell, and Allan F Williams. 1995. "Traffic accident involvement rates by driver age and gender." *Accident Analysis & Prevention*, 27(1): 73–87.

Mooren, Lori, Raphael Grzebieta, Ann Williamson, Jake Olivier, and Rena Friswell. 2014. "Safety management for heavy vehicle transport: A review of the literature." *Safety science*, 62: 79–89.

New York State Department of Environmental Conservation. 2011. "Revised Draft Supplemental Generic Environmental Impact Statement on the Oil, Gas and Solution Mining Regulatory Program."

Peltzman, Sam. 1975. "The effects of automobile safety regulation." *The Journal of Political Economy*, 677–725.

Rahm, Dianne, Billy Fields, and Jayce L Farmer. 2015. "Transportation Impacts of Fracking in the Eagle Ford Shale Development in Rural South Texas: Perceptions of Local Government Officials." *Journal of Rural & Community Development*, 10(2).

Rubin, Donald B. 1980. "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment." *Journal of the American Statistical Association*, 75: 591–593.

Simpson, Andrew G. 2014. "Drive Underway to Increase Truckers' Minimum Insurance Limits." *Insurance Journal.*

Smith, Eric, and Randall Wright. 1992. "Why is automobile insurance in Philadelphia so damn expensive?" *American Economic Review*, 756–772.

Sun, Stephen Teng, and Constantine Yannelis. 2015. "Quantifying the Premium Externality of the uninsured." *Journal of the European Economic Association.*

Theodori, Gene L. 2009. "Paradoxical perceptions of problems associated with unconventional natural gas development." *Southern Rural Sociology*, 24(3): 97–117.

US Department of Transporation. 2013. "Annual Vehicle Distance Traveled in Miles and Related Data by Highway Category and Vehicle Type, Table VM-1, Federal Highway Administration."

US Department of Transportation. 2013. "Examining the Appropriateness of the Current Financial Responsibility and Security Requirements for Motor Carriers, Brokers, and Freight Forwarders-Report to Congress, Federal Motor Carrier Safety Administration."

US Department of Transportation. 2016. "Pocket Guide to Large Truck and Bus Statitistics."

Veil, John A. 2010. "Final report: Water management technologies used by Marcellus shale gas producers." Environmental Science Division, Argonne National Laboratory for the US Department of Energy, Office of Fossil Energy, National Energy Technology Laboratory.

Vickrey, William. 1968. "Automobile accidents, tort law, externalities, and insurance: an economist's critique." *Law and Contemporary Problems*, 33(3): 464–487.

Wilson, Riley. 2016. "Moving to Economic Opportunity: The Migration Response to the Fracking Boom."

A Appendix (For Online Publication)

Here we provide information on supplementary regressions referred to in the main text.

A.1 Restricting to only-treated samples

In the paper our regression samples include all counties (zip codes or roads). Here we show that our results are robust to restricting the sample to those that are treate at some point in time. This is an expected finding because the nontreated counties have no variation in treatment and will then be captured by the fixed effects. Table A1 shows results for the traffic regressions, Table A2 shows results for our accident regressions, and Table A3 shows results for our insurance premium regressions.

Table A1: Effect on traffic counts

	Heavy-truck count	Nontruck count
	(1)	(2)
Panel A: County level		
Wells	0.71**	3.24**
	(0.30)	(1.51)
Year fixed effects	Yes	Yes
County fixed effects	Yes	Yes
Mean of dep. var. (in levels)	354	3,543
R^2	0.37	0.59
N	386	386
Panel B: Segment level		
Withdrawal 0m	47.7*	106.7
	(24.6)	(109.5)
1-500m	19.4	60.2
	(28.3)	(108.2)
501-1000m	58.3	98.7
	(36.2)	(160.7)
Disposal 0m	20.2***	20.3
	(6.1)	(18.5)
1-500m	10.9***	15.4
	(4.0)	(15.1)
501-1000m	-0.4	15.0
	(3.0)	(14.8)
Segment fixed effects	Yes	Yes
Year-county fixed effects	Yes	Yes
Mean of dep. var. (in levels)	524	5,220
R^2	0.73	0.82
N	77,935	77,935

Notes: Specifications replicate Table 2 except that we restrict the sample to those counties that at some point in time have a well (Panel A) or segments that at some point in time connect a well to a water withdrawal location or waste disposal location (Panel B).

Table A2: Robustness: Accident regressions using only ever-treated sample

	IHS(truck) (1)	IHS(vehicle) (2)	IHS(fatal) (3)	IHS(injury) (4)
Panel A: County level				
Wells	0.0098***	0.0023***	0.0054***	0.0028***
	(0.0018)	(0.0005)	(0.0009)	(0.0005)
County fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	17	305	3.7	162
R^2	0.87	0.99	0.61	0.98
N	2,808	2,808	2,808	2,808
Panel B: Segment level				
Withdrawal 0m	.02528***	.02906***	.00211	.01151**
	(.00551)	(.00562)	(.00186)	(.00483)
1-500m	.00046	.00165	-.00011	.00064
	(.00060)	(.00164)	(.00027)	(.00119)
501-1000m	.00092	.00130	.00040	.00160
	(.00067)	(.00167)	(.00037)	(.00120)
Disposal 0m	.00164***	.00364***	-.00020	-.00273***
	(.00056)	(.00108)	(.00020)	(.00097)
1-500m	.00002	.00109***	.00003	.00034*
	(.00006)	(.00025)	(.00003)	(.00017)
501-1000m	.00006	.00155***	.00002	.00084***
	(.00005)	(.00034)	(.00002)	(.00022)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	.0038	.047	.00068	.026
R^2	.47	.66	.10	.60
N	13,845,617	13,845,617	13,845,617	13,845,617

Notes: Specifications replicate Table 3 except that we restrict the sample to those counties that at some point in time have a well (Panel A) or segments that at some point in time connect a well to a water withdrawal location or waste disposal location (Panel B).

Table A3: Robustness: Car insurance premiums using only shale zip codes

	Average premium	Vehicle theft
	(1) IHS(dollars)	(2) IHS(count)
Withdrawal 0m (total, in 1000s)	.0210***	.2281
	(.0030)	(.2664)
Disposal 0m (total, in 1000s)	-.0001*	.0002
	(.0001)	(.0140)
Zip code fixed effects	Yes	No
Year fixed effects	Yes	Yes
County fixed effects	No	Yes
R^2	.97	.96
N	2,433	122
Mean of dep. var. (in levels)	1,076	81

Notes: Specifications replicate Table 7 except that only those zip codes that are traversed at some point in time are used. Table 7 uses the full sample.

A.2 Treatment effect is time constant

Shale gas extraction has rapidly grown and changed over time. Therefore, it is possible that recent infrastructure for water pipes or increased recycling of wastewater on site could mean that our estimates would be smaller in recent years. To examine this conjecture, we include in our specification interactions with an indicator for occurring after 2011. Table A4 shows that our treatment is not different in the years after 2011; coefficients on the interaction terms are statistically insignificant (with the exception of small negative terms on the 501-1000m waste disposal routes.

Table A4: Robustness: Are impacts constant over time?

	Segment level			
	(1) IHS(truck)	(2) IHS(vehicle)	(3) IHS(fatal)	(4) IHS(injury)
Withdrawal 0m	.02917***	.02798***	.00294	.01253**
	(.00718)	(.00667)	(.00241)	(.00546)
Withdrawal 1-500m	.00006	.00305	-.00030	.00103
	(.00079)	(.00202)	(.00033)	(.00152)
Withdrawal 501-1000m	.00021	.00038	-.00007	.00127
	(.00081)	(.00215)	(.00048)	(.00172)
Disposal 0m	.00153	.00455***	.00051	-.00120
	(.00098)	(.00169)	(.00053)	(.00146)
Disposal 1-500m	.00009	.00150***	-.00003	.00040
	(.00010)	(.00038)	(.00004)	(.00027)
Disposal 501-1000m	.00007	.00240***	.00005	.00145***
	(.00008)	(.00052)	(.00005)	(.00035)
Post-2011*(Withdrawal 0m)	-.00983	.00319	-.00232	-.00283
	(.00854)	(.00957)	(.00381)	(.00919)
Post-2011*(Withdrawal 1-500m)	.00105	-.00284	.00044	-.00078
	(.00113)	(.00286)	(.00057)	(.00235)
Post-2011*(Withdrawal 501-1000m)	.00171	.00241	.00109	.00075
	(.00118)	(.00300)	(.00077)	(.00245)
Post-2011*(Disposal 0m)	.00018	-.00103	-.00095	-.00204
	(.00094)	(.00170)	(.00059)	(.00154)
Post-2011*(Disposal 1-500m)	-.00011	-.00030	.00008	-.00006
	(.00012)	(.00040)	(.00005)	(.00029)
Post-2011*(Disposal 501-1000m)	-.00002	-.00084*	-.00006	-.00078**
	(.00009)	(.00046)	(.00005)	(.00035)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	.0026	.043	.00052	.024
R^2	.42	.65	.10	.60
N	34,493,677	34,493,677	34,493,677	34,493,677

Notes: Specifications replicate Table 3 except that we include interactions with an indicator for post-2011. This should alleviate the concern that in recent years more water was shipped using pipelines or reused at the well site.

A.3 Different functional form

In the main paper we use the inverse hyperbolic sine transformation of accident counts when regressing accidents on shale trucking routes, because there are many segment-quarters with zero accidents. Here we show that the results are similar when using the $\log(y+1)$ transform instead.

Table A5: Robustness: Log-transform of dependent variables

	IHS(truck) (1)	IHS(vehicle) (2)	IHS(fatal) (3)	IHS(injury) (4)
Panel A: County level				
Wells	0.0092***	0.0017***	0.0048***	0.0019***
	(0.0015)	(0.0006)	(0.0006)	(0.0005)
County fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	25	468	4.9	259
R^2	0.91	0.99	0.72	0.98
N	4,824	4,824	4,824	4,824
Panel B: Segment level				
Withdrawal 0m	.01950***	.02323***	.00127	.00659
	(.00421)	(.00445)	(.00150)	(.00461)
1-500m	.00040	.00153	-.00012	-.00026
	(.00047)	(.00127)	(.00022)	(.00107)
501-1000m	.00073	.00116	.00036	.00115
	(.00052)	(.00129)	(.00031)	(.00106)
Disposal 0m	.00128***	.00306***	-.00022	-.00394***
	(.00044)	(.00086)	(.00016)	(.00091)
1-500m	.00001	.00098***	.00003	.00017
	(.00005)	(.00019)	(.00003)	(.00016)
501-1000m	.00004	.00137***	.00001	.00077***
	(.00004)	(.00026)	(.00002)	(.00020)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	.0026	.043	.00057	.035
R^2	.42	.65	.10	.57
N	34,493,677	34,493,677	34,493,677	34,493,677

Notes: Specifications replicate Table 3 except that we take the log of the dependent variable (after adding one, $\log(y+1)$). Table 3 uses the inverse hyperbolic sine transformation of the dependent variable.

A.4 Including all road types in the sample

In the paper, the estimation sample includes only primary, secondary, and local roads, excluding private roads that are used for service vehicles and unpaved dirt trails that require four-wheel drive. We use the full road network to predict GIS routes, since private roads are used to access wells. Excluding this category reduces our sample size by 18 percent but reduces the number of accidents by only 2 percent. Table A6 shows that our results are unchanged by including all road types.

Table A6: Robustness: Including all road segments

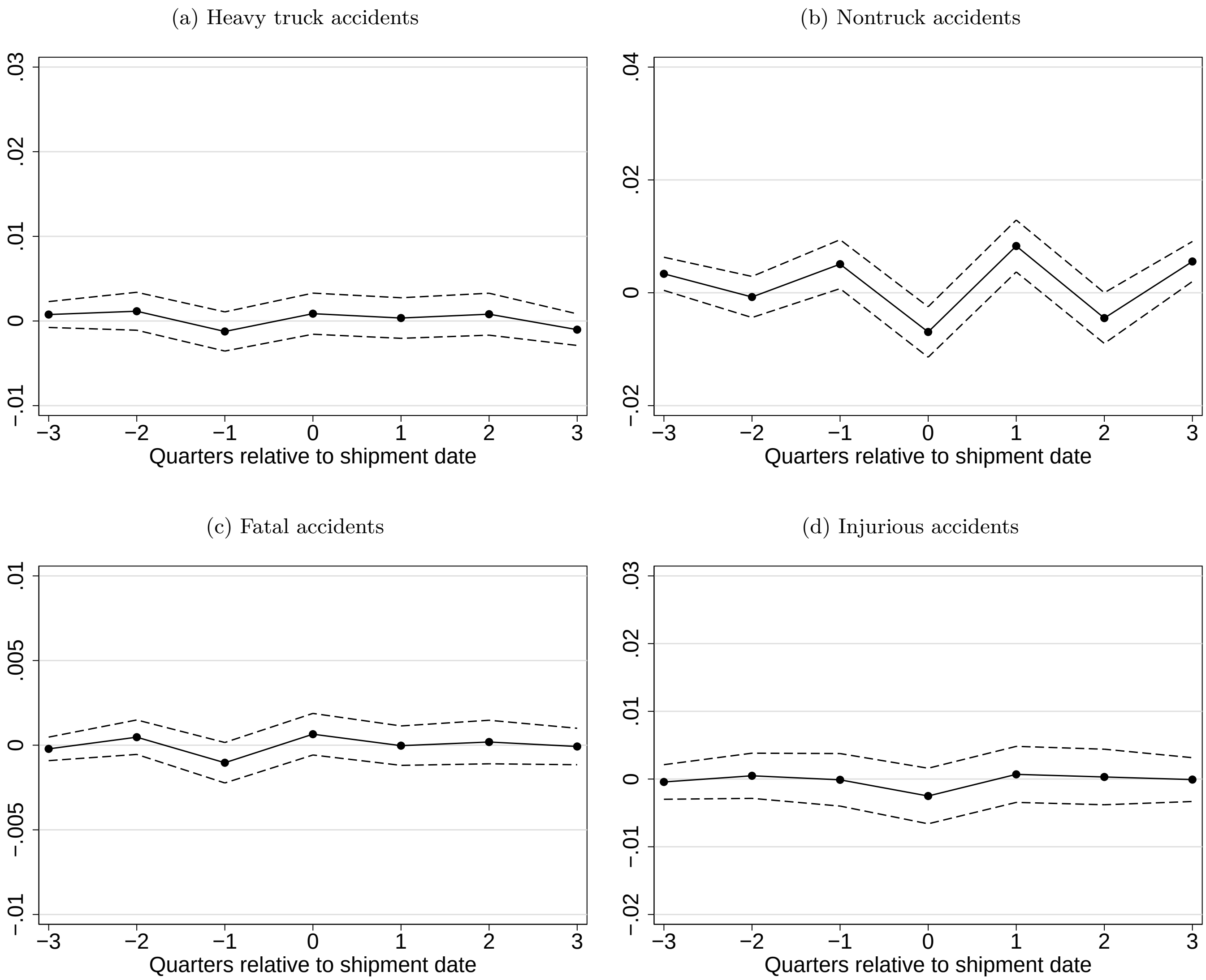
	Segment level			
	(1) IHS(truck)	(2) IHS(vehicle)	(3) IHS(fatal)	(4) IHS(injury)
Withdrawal 0m	.02225***	.02619***	.00187	.01039**
	(.00484)	(.00494)	(.00163)	(.00423)
1-500m	.00046	.00170	-.00009	.00059
	(.00047)	(.00130)	(.00021)	(.00094)
501-1000m	.00076	.00111	.00031	.00124
	(.00052)	(.00129)	(.00028)	(.00092)
Disposal 0m	.00156***	.00378***	-.00019	-.00251***
	(.00053)	(.00101)	(.00019)	(.00091)
1-500m	.00002	.00125***	.00002	.00036**
	(.00005)	(.00022)	(.00003)	(.00015)
501-1000m	.00005	.00152***	.00001	.00073***
	(.00004)	(.00028)	(.00002)	(.00018)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	.0021	.035	.00043	.02
R^2	.42	.65	.10	.60
N	42,205,846	42,205,846	42,205,846	42,205,846

Notes: Specifications replicate Table 3 except that we include all road segments in the analysis (instead of dropping dirt roads and private access roads).

A.5 Coefficients on disposal routes

The paper shows, by plotting the coefficient estimates on the withdrawal routes before and after drilling, that the water withdrawal traffic mostly occurs in the quarter after drilling. In the case of the waste disposal shipments we know the half-year that waste was shipped, but we nonetheless plot lags before and after this shipment date (equation 3). These are plotted on the same axes as the withdrawal route figures in the main text (Figure 4). They are much smaller in magnitude, and nearly always statistically insignificant. This is likely because one well can have shipments spread across many different quarters, and so the treatment is not as strong.

Figure A1: Effect of wastewater shipment on accident frequency



Notes: Plotted on the same axes as Figure 4 are the leads and lags of wells using the road for waste shipment. Coefficients from equation (3), β_0^τ, are plotted where τ represents the quarters before and after the shipment date. Each figure is from a different regression of the following outcome variables: (a) the count of accidents involving a heavy truck; (b) the count of nontruck accidents; (c) the count of accidents with a fatality; and (d) the count of accidents with injuries. Dashed line represents standard errors.

A.6 Accident severity

The paper shows that both truck accidents and accidents not involving trucks increase. If only *truck* accidents increased, we would expect that the average accident would become more severe. However, because nontruck accidents are also increasing, the average severity of an accident need not be increasing. And indeed, with more trucks, the average accident that occurs on a segment is not more severe.

Table A7: Accident severity: accident level analysis

	Panel B: Segment level			
	(1) Pr(acc. fatality)	(2) IHS(no. fatalities)	(3) Pr(acc. injury)	(4) IHS(no. injuries)
Withdrawal 0m	0.00104	0.00024	0.00584	0.00914
	(0.00300)	(0.00277)	(0.01171)	(0.01415)
1-500m	-0.00921	-0.00935*	0.00142	-0.02807
	(0.00607)	(0.00561)	(0.02735)	(0.03310)
501-1000m	0.00163	0.00318	0.01647	-0.00847
	(0.00795)	(0.00813)	(0.03038)	(0.03664)
Disposal 0m	-0.00005	-0.00008	-0.00519***	-0.00519***
	(0.00038)	(0.00034)	(0.00169)	(0.00192)
1-500m	0.00110**	0.00123**	-0.00485	-0.00811*
	(0.00053)	(0.00051)	(0.00387)	(0.00466)
501-1000m	0.00031	0.00030	0.00072	0.00230
	(0.00073)	(0.00068)	(0.00356)	(0.00426)
Segment fixed effects	Yes	Yes	Yes	Yes
County-year fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (in levels)	.011	.011	.53	.58
R^2	0.06	0.06	0.09	0.09
N	1,576,413	1,576,413	1,576,413	1,576,413

Notes: Observations are at the crash-level. Outcome variables are (1) whether the accident had one or more fatalities, (2) the inverse hyperbolic sine transformation of the number of fatalities, (3) whether the accident had one or more injuries, (4) the inverse hyperbolic sine transformation of the number of injuries.