The Accident Externality from Trucking: Evidence from Shale Gas Development

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

February 2019

Abstract

The presence of a heavy truck on the road can impose an externality if accidents occur that would not have otherwise. Using the rapid influx of trucks during the shale gas boom in Pennsylvania, we obtain an estimate of the additional accidents that occur when a truck is added to the road. We find that trucks not only increase the number of truck accidents, but to an even larger extent, they increase the number of car-on-car collisions. We find suggestive evidence that this accident externality reverberates to even more road users through higher car insurance premiums.

Keywords: externality, trucking, hydraulic fracturing, traffic fatalities

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

*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, Hendrik Wolff, Chi Man Yip, seminar participants at London School of Economics, Stockholm School of Economics in Riga, University of British Columbia, University of Melbourne, University of Pennsylvania, and University of Verona, and conference attendees of the Association of Environmental and Resource Economics annual meeting and the University of Waterloo's Current Challenges in Environmental and Resource Economics Workshop. We also thank Carinsurance.com for sharing data on insurance premiums and the Bechtel Foundation and the Smith Richardson Foundation for funding. Addresses: Lucija Muehlenbachs, Department of Economics, University of Calgary *lmuehlen@ucalgary.ca.* Stefan Staubli, Department of Econmics, University of Calgary, *sstaubli@ucalgary.ca.* Ziyan Chu, Resources for the Future, Washington, DC *chu@rff.org.*

1 Introduction

Suppose, to pass a heavy truck, a car enters oncoming traffic and collides with another car. The accident occurred because the truck was on the road, but current liability regimes would never hold the truck responsible for any damages. These types of accidents are external to the decision of how much to truck, which would suggest we have too much trucking from society's perspective. Certainly, trucking is a ubiquitous form of transport: 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] Should the trucks pose an accident externality, the cost of the externality could fall on all other road users through higher car insurance premiums.

The aim of this paper is twofold. *First*, we study the causal effect of the safety risk that an additional truck on the road poses. To do so, we exploit the large and rapid influx of trucks transporting water for hydraulic fracturing in Pennsylvania, the state that produces the most shale gas in the US, and estimate the effect of an additional truck on the frequency and severity of accidents. The *second* aim of this paper is to quantify the monetary cost of these additional accidents for other road users. We exploit a novel data set that records the insurance premiums offered to new enrollees in Pennsylvania and study the casual effect of the truck traffic induced by shale on insurance premiums of nearby residents.

The economics literature on traffic accidents has focused on cars and light trucks. Vickrey (1968) first pointed out that by merely being on the road, a vehicle imposes an accident externality, since any accident that occurs would not have occurred had the driver chosen, for example, to take the train. This means that the marginal damage caused by *each* vehicle, regardless of fault, is the full damage cost, yet liability rules would never hold both cars responsible for the full damages. In the case of trucks, this externality is even more extreme, because by being long and tall, trucks can increase the risk of other types of accidents occurring between other road users (e.g., car-on-car collisions). Moreover, for those accidents that involve a truck, the truck's weight, larger frame, height, wheelbase, breaking distance, and rigidity will affect the amount of damage inflicted on

[1] Trucks carried 67 percent, or 13 billion tons, in 2012 (US Department of Transportation, 2013); the most common occupation includes the count of delivery truck drivers ("Map: The Most Common Job in Every State, NPR, Planet Money," February 5, 2015. Quoctrung Bui.); and the GDP estimate is from the Gross-Domestic-Product-(GDP)-by-Industry Data, US Department of Commerce, Bureau of Economic Analysis, 2013.

the other vehicle.[2] Therefore, additional trucks on the road could translate into not only more accidents, but more severe accidents. Alternatively, people could compensate for being surrounded by heavy trucks by driving more carefully, which would result in fewer and less severe accidents.[3] Thus, the effect of an additional truck on the frequency and severity of accidents is ultimately an empirical question.

The transportation literature has examined safety risks of heavy trucks, but with an eye toward identifying predictors of truck accident rates (e.g., safety management practices and driver or company characteristics). Most 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 contrast, we use panel data with plausibly exogenous variation in trucks to examine both the impact on truck accidents as well as car-on-car collisions.

The extent to which the accident externality is internalized depends on the underlying liability regime. In a collision, the negligent party internalizes all damages, although the size of the damages will increase with the number of parties involved in the accident. An externality then arises from accidents having higher costs, but costs which are not internalized by all parties, and this externality can materialize into higher insurance premiums for all road users, as shown by Edlin and Karaca-Mandic (2006) in the case of private vehicles. Following this logic, additional trucks on the road, even when non-negligent, will impose external accident costs. Furthermore, even in cases in which the truck is deemed negligent, the accident can also impose an externality on others if the truck is underinsured. Underinsured vehicles impose an external cost on other road users, through higher insurance premiums of those who obtain insurance, as shown in the context of private vehicles by Smith and Wright (1992) and Sun and Yannelis (2015). It is arguable that trucks on US roads are underinsured. Trucks are required to hold insurance, or a surety bond, to cover a minimum amount of liability (set by the Federal Motor Carrier Safety Administration, FMCSA), however, the current

[2]Studies on the effect of vehicle weight on safety (Crandall and Graham, 1989; Li, 2012; Jacobsen, 2013; Anderson and Auffhammer, 2014; Bento, Gillingham and Roth, 2017) have focused on private vehicles, and not the heavy haulers of this study, which easily exceed 80,000 pounds.

[3]Or, outside the scope of this paper, in the long run, drivers could buy larger cars to protect themselves, which could result in more severe accidents with other road users (Li, 2012).

liability minimum was set over thirty years ago, in 1985 at $750,000.[4] Discussions about raising the limit bring objections from small businesses.[5] We expect the underinsurance of catastrophic crashes, combined with an increase in the number of accidents overall, imply that more trucks on the road leads to higher car insurance rates.

Our estimation strategy relies on the vast quantities of water used and produced when hydraulically fracturing shale formations for natural gas. Water is pumped into a well at high pressure, fracturing the shale rock to release its natural gas, and wastewater is produced, consisting of salty water that flows to the surface with the gas (alongside any fracturing fluid also returning to the surface). Freshwater and wastewater are primarily transported using tanker trucks—with one well requiring 800 to 2,400 one-way trips.[6] We exploit the spatial and temporal variation in the location of wells, water sources, and waste destinations. We use geographic information systems (GIS) to predict the most likely route that trucks take to haul water to and from a well.

The shale gas routes provide a unique setting in which a large influx of trucks is concentrated in a small area over a short period of time (typically less than 90 days). This has advantages for our identification strategy. 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 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. A resource boom is not only associated with trucks, but also other

[4] 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 have differing conclusions on the frequency with which crashes exceeded the liability limits. 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).

[5] 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 Transportation, 2016). 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/.

[6] These estimates are from New York State Department of Environmental Conservation (2011); Abramzon et al. (2014); Gilmore, Hupp and Glathar (2013). Transporting water by truck is a costly endeavor and there are moves to transport more water via pipeline. 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. Appendix Section A.6 shows our estimates are constant over time.

confounding factors, such as an influx of young male drivers. Our estimation strategy allows us to isolate the impact of a truck *per se* by comparing, over time, the specific roads used by shale gas trucks to similar roads in the vicinity not used by shale gas trucks.

First, we examine the county-wide changes in accidents as more wells are drilled. Our estimates imply that each shale gas well is associated with an increases of 2,800 additional trucks. Using accident data at the county level, we estimate that for each well, in the quarter and county in which it is drilled, there are an additional 0.171 accidents involving heavy trucks and an additional 0.815 accidents involving other vehicles (a 1 percent and 0.26 percent increase, respectively).[7] Increased accident rates following shale booms have been documented in Pennsylvania, Wisconsin, and North Dakota (Graham et al., 2015; Kalinin, Parker and Phaneuf, 2017; Xu and Xu, 2018).[8] These estimates provide insights into the costs of shale gas development, useful for county planners and relevant to the literature on the economic impacts of hydraulic fracturing (see review, Mason, Muehlenbachs and Olmstead, 2015), but they cannot be used to back out the safety risks of adding a truck to the road. The increase in truck traffic could coincide with an unobserved county shock (e.g., a change in population) that also increases the number of accidents. 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).[9]

To isolate the increase in accidents from adding a truck to a road, we rely on an identification strategy that zeros in on individual road segments. We examine changes on the routes predicted to be used by trucks to similar roads in the vicinity not used by trucks. On the truck routes, we find a statistically significant increase in truck traffic within the range of previous reports' predictions of shale gas truck traffic. In contrast, we don't see an increase in the number of cars on these roads, suggesting that we can isolate the impact of a truck from the general increase in traffic associated with a shale boom. We examine separately local-neighborhood/rural roads from main-arterials/highways. We find that adding a truck to a local/rural road results in more truck

[7]Note, the increase in accidents means an increase in the number of fatalities and injuries as well, but since both truck and car-on-car accidents increase, the average accident does not appear to be more severe.

[8]In the case of Wisconsin, Kalinin, Parker and Phaneuf (2017) find an increase in truck accidents following a sand-mining boom which was spurred by the shale boom.

[9] Th 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.

accidents, but also more car accidents that don't directly involve a truck. Similarly, in the case of highways, we find that adding a truck to a highway increase the number of car accidents, but has no effect on the number of truck accidents. We find evidence that the increase in car accidents is driven by a change in driving behavior. Specifically, the share of accidents attributed to aggressive driving, speeding, and changing lanes is significantly higher on truck roads, while the characteristics of drivers and cars are unchanged.

Using the road-level estimates, we calculate the number of accidents that occur with each kilometer a truck drives on a Pennsylvania road. In our sample of local and rural roads, one truck accident occurs annually for every 128 trucks (and in our sample of highways we don't estimate additional truck accidents). We estimate a larger increase in car accidents: one car accident occurs annually for every 15 trucks on local and rural roads and for every 10 trucks on highways. An important caveat to these estimates is external validity: the risk of a truck will depend on road characteristics, traffic volumes, and the alternative routes available for cars. Our estimates apply to a sample of mostly rural Pennsylvanian roads and may not extend to other settings. For example, the differential effects of a truck on highways versus local roads suggests that trucks might pose a smaller risk in more urban areas than what is estimated in this paper. However, while these estimates should not be applied to other states, the qualitative finding, that trucks can increase the number of accidents between other road users, is a previously undetected finding that is important to consider when designing public policy around trucking.

We also estimate an impact on car insurance premiums using a unique data set of the premiums offered by six national carriers to a representative new insurance enrollee (a single 40-year-old male who commutes 12 miles to work each day in a new Honda Accord and has a clean driving record and good credit). We find that areas exposed to shale gas truck traffic see insurance rates increase for representative new enrollees. Specifically, the average truck-traversed zip code saw an increase in annual insurance premiums of $2.92, with the most-traversed seeing an increase of $30.40. The increase responds to the location of the truck routes, not the location of the wells. Translating the impact to an estimate per kilometer driven by a truck, a year's worth of truck miles would increase this new enrollee's insurance premium by 6 cents. While this estimate is small, the full cost of the truck would entail multiplying this increase by all new enrollees in the zip code (assuming that the increase for new enrollees is the same as the increase for this representative male). If the driving

population of Pennsylvania re-enrolled, this cost would mean one truck imposes an external cost of $8,469 in higher aggregate insurance premiums.

While we provide an estimate of the accident externality from trucking, this estimate does not include other, yet-to-be measured costs, such as the added stress of driving on roads with trucks. Similarly, there are a host of other documented externalities associated with trucking, such as the costs of congestion, pavement damages, noise, energy security, and local and global pollution (for example, Parry, 2008, Austin, 2015, He, Gouveia and Salvo, Forthcoming, and Cohen and Roth, 2018). It is also important to note that trucking has economic benefits. While to the best of our knowledge these benefits have not been quantified for trucking specifically, large benefits from transportation infrastructure have been documented, from reducing trade costs and increasing productivity, income, manufacturing, and land values (Ghani, Goswami and Kerr, 2016; Donaldson and Hornbeck, 2016; Donaldson, 2018).

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.

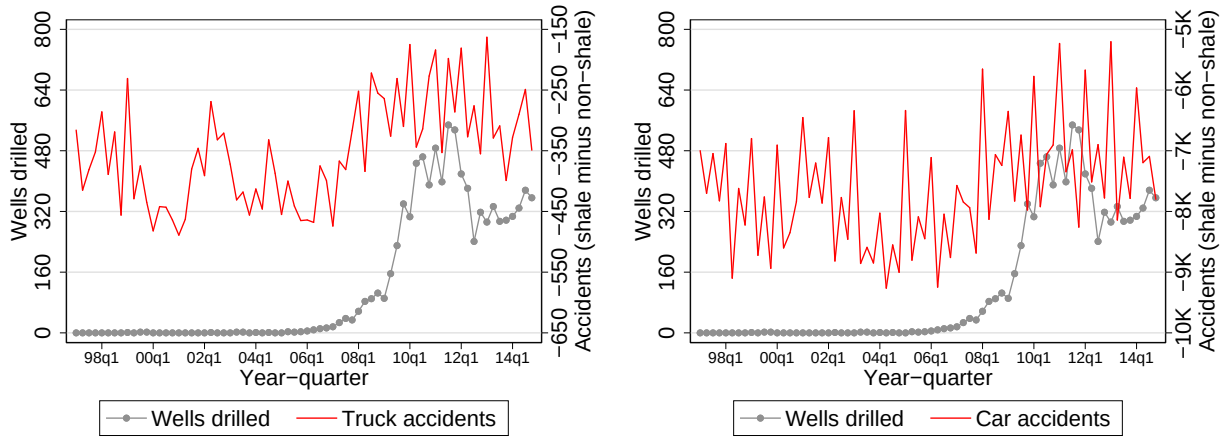
2 Background and data

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). 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.

Indeed, if we look at the count of accidents that occur in counties with shale gas wells, we can see that with a shale boom, both truck accidents and car-on-car accidents increase. Figure 1 shows

the correlation between truck accidents, car-on-car 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).

This increase in accidents represents the 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. To isolate the increase in truck traffic from other idiosyncratic shocks, we zoom in to the road level. Using GIS (described below), we predict the most likely route that the trucks take. Because hydraulic fracturing is concentrated over a short period of time, we can compare the rates of accidents before and after trucks use a given road segment relative to similar road segments not used by trucks, while controlling for the general increase in traffic on all roads in the county.



(a) Truck accidents

(b) Car accidents

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.

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, which provides insights into the impacts of shale gas development on road safety, and one at the road-segment level, which provides insights into the impacts of trucks on road safety.

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 vehicle counts from 2004 to 2014. Traffic count data must be handled with caution. Some observations are imputed by PennDOT, either by repeating the same traffic counts across different years, or by inflating using estimates of population growth.[10] In the years when traffic is 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 shale truck traffic was the heaviest (discussed later when interpreting the coefficients). Despite these shortcomings, we nonetheless obtain a shale-gas-truck count that is 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 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 completion occurs on average 80 days after the drill date.[11]

[10] 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.

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

Water withdrawal and waste 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.[12] 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.[13]

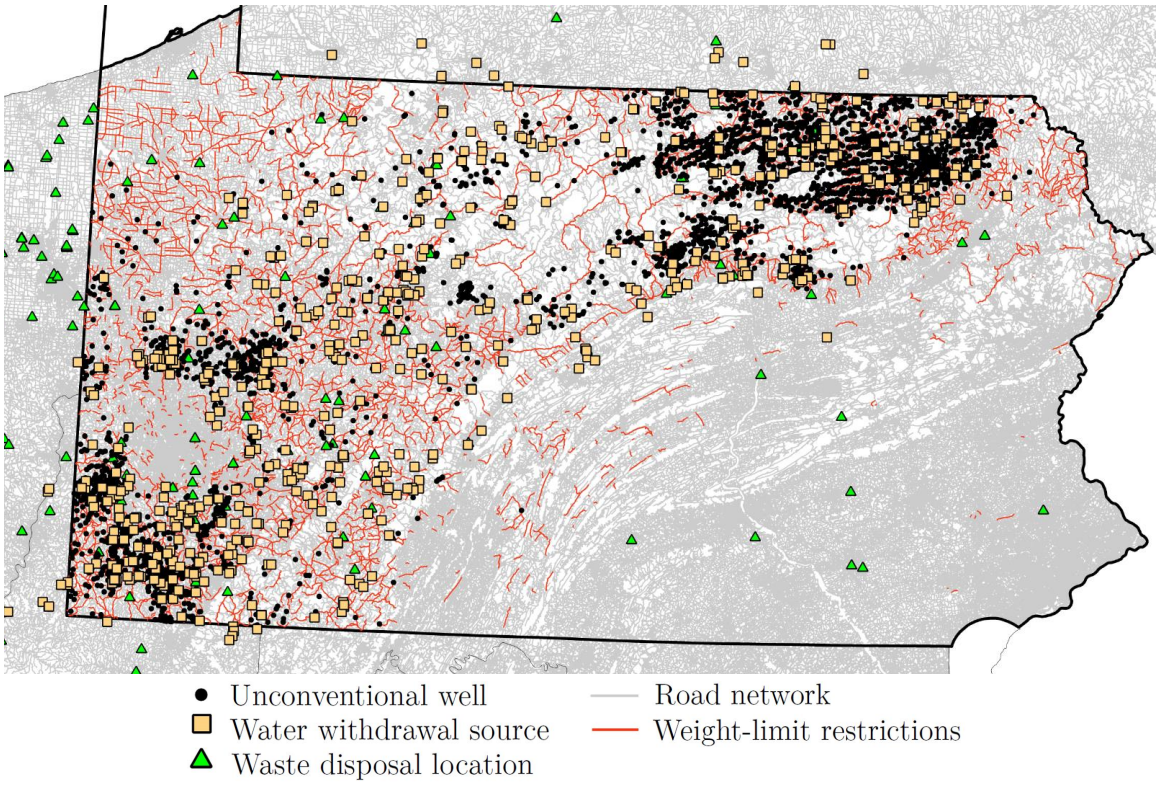


Figure 2: Wells, waste disposal, water withdrawal, and weight-limit restrictions in Pennsylvania, 2014

[12]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.

[13]There are more waste disposal sites even farther away than 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.

2.2 Construction of truck traffic routes

We use GIS to predict the most likely transportation route that the trucks take to get from a water withdrawal point to a well and from a well to a waste-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, but there are penalties on roads with lower speed limits. We assigned impedances on each road depending on the speed limit of the road type.[14] Private roads that are used for service vehicles and unpaved dirt trails that require four-wheel drive are included in the GIS work to connect wells to the road network, but otherwise, in our analysis we drop these roads.[15]

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.[16] 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 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.

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

[15] Including these roads increases the size of our sample by 18% but only .02% of all accidents occur on these roads.

[16] 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 are $13,000 to $23,000 per well (Abramzon et al., 2014).

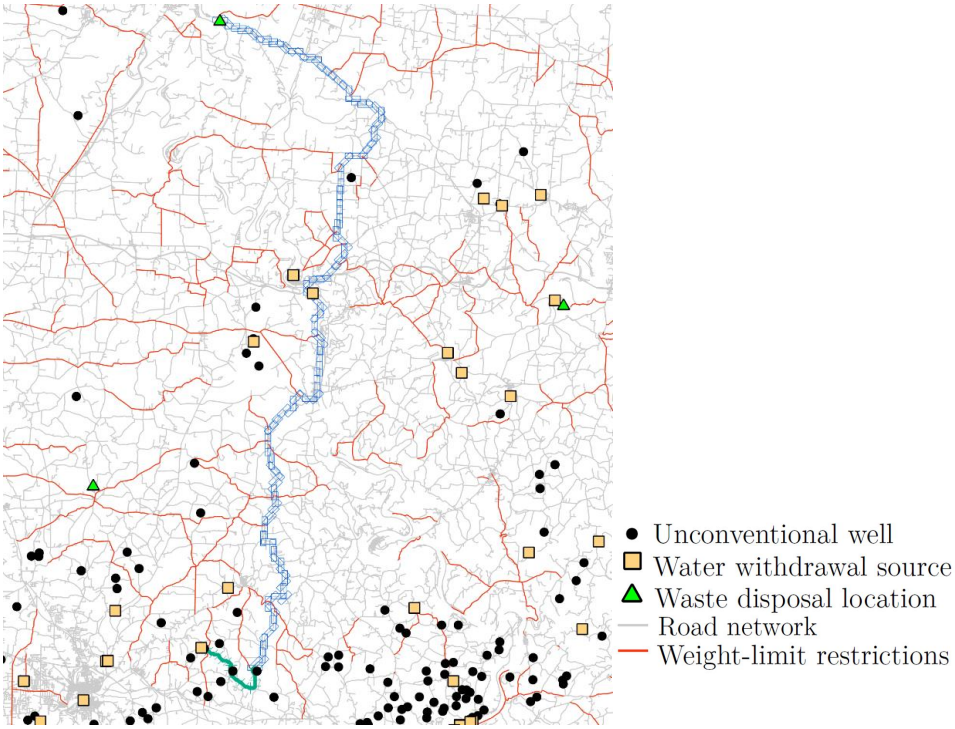


Figure 3: Example route from water withdrawal location to well to waste disposal site

2.3 Counting road use for water withdrawal and waste 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, we assume that the points that were approved in 2009 were also available in earlier years.[17] Because reported completion dates are on average 80 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. To obtain a road segment-year-quarter observation, we sum the total number of wells that are predicted to use the road segment in the year-quarter.

In the case of waste disposal, we know how much waste was shipped as well as the location of where it was shipped. Waste quantities were reported to PADEP annually from 2004 to 2009

[17] 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.

and semi-annually from 2010 to 2014.[18] 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, so that both withdrawal and shipment correspond to the trucks needed for one well.[19]

3 Identification strategy and estimation sample

We provide two different estimation strategies, each estimating different types of impacts. The first strategy estimates the county-wide aggregate safety impacts from variation in shale gas development. These estimates will capture both the effects of a change in truck traffic as well as any unobserved county-specific shifts from an influx of workers and wealth in the area. As shown by Weber (2012), Fetzer (2014), Jacobsen (2016), Bartik et al. (2016), Maniloff and Mastromonaco (2017), and Feyrer, Mansur and Sacerdote (2017), the local income and employment shocks from shale gas booms are sizeable. The second estimation strategy provides an estimate of the average change in accidents on the shale gas truck routes. This estimate will isolate the impact of an additional truck on the number of accidents under the assumption that. after controlling for time-specific and road-specific factors, roads not used by shale gas trucks provide a good counterfactual for roads used by trucks.

We exploit the temporal and spatial variation in the location of shale gas wells, water withdrawal locations, and disposal locations in the Marcellus shale region. Separately we estimate the effects of shale gas development on traffic counts, and then the effects 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. 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.

[18] 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.

[19] 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).

3.1 Strategy to identify the impact of shale gas development at the county 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. We compare outcomes in counties before and after wells are drilled, in relation to changes in the remaining counties. This comparison can be implemented with the following fixed effects regression:

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

where y_{ct} is the outcome variable of interest: when examining traffic flows, it is county c's average-daily-traffic count in year t (traffic counts are reported by road segment as annual averages, and we take the average across all roads in the county). When examining accidents, y_{ct} is the number of accidents (and t signifies year-quarter). The main coefficient of interest is α, representing the change in the mean of the outcome variable from drilling an additional well.

The identifying assumptions are that the locations of the wells are determined independently from changes in traffic and accidents and that there are no spillover effects from treatment counties to control counties.[20] 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 across treatment facilities.[21] 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 the road segment.

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

[21] 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).

3.2 Strategy to identify the impact of shale gas development at the road-segment level

Here we examine the same outcome variables but by road segment. Our treatment variable, Truck Routes$_{st}$ is constructed from the GIS prediction of the least-cost route between wells, water withdrawal, and waste disposal points and represents the number of wells predicted to use the segment to reach a water withdrawal or disposal point in the quarter (rescaled for every 10 wells). Using data by road segment, s, we examine traffic counts and accident counts, y_{st}:

$$y_{st} = \alpha \text{Truck Routes}_{st} + \lambda_s + \delta_t + \mu_{ct} + \varepsilon_{st} \tag{2}$$

Our main coefficient of interest is α, which represents the change in our outcome variable when a road is used to connect wells to a water source or waste disposal location, relative to the change in control roads. We include road-segment fixed effects, λ_s, to capture time-invariant differences in traffic and accidents across different roads. When examining accident counts, our data are at the quarterly level, so δ_t represent year-quarter fixed effects, capturing statewide seasonal road conditions, safety trends over time, and macroeconomic shocks. Also included in the accident regressions are county-by-half-year fixed effects, μ_{ct}, capturing countywide changes in traffic and accidents within a six-month period. These control for county-wide boomtown effects, such as an influx of young male drivers, affecting all roads in the county-half-year, but are not concentrated in the particular quarter on the particular road used by trucks. Our data on traffic counts are at the annual level, and therefore in the traffic regressions we only include μ_{ct}, representing county-year fixed effects. To deal with the concern that there might be spillover effects of treatment roads on our control roads, in all specifications, we drop control roads that are within 500 meters of a truck route.

The reason to zoom in to the road-segment level is to obtain an estimate of the increase of accidents on shale gas truck routes, absent other boomtown impacts. However, one potential concern is that some these truck roads may also be used by workers needing to get to the wells, particularly when there is only one road to access a well.[22] Therefore, the assumption that a

[22] We are less worried about workers driving to and from the withdrawal and disposal points, because trucks are parked at the well site.

comparison of segments used and not used by trucks over time, after controlling for county-half-year and road-segment fixed effects, can isolate the effect of trucks per se is more plausible the farther away from a well. We therefore include an additional regressor to capture the fact that routes nearest the wells are likely to have heavier traffic. Specifically, we allow for a differential impact on truck routes within a certain distance of a recently drilled well (drilled in the quarter or previous quarter). Hence, the regression we estimate is:

$$y_{st} = \alpha \text{Truck Routes}_{st} + \beta \text{ Truck Routes} \times \text{I(Near well)}_{st} + \lambda_s + \delta_t + \mu_{ct} + \varepsilon_{st} \tag{3}$$

In our main specification we use 2km to designate whether a route is near a well. However, we also show results from different regressions, each with a different distance used to designate whether a route is near a well. If we are indeed isolating a truck effect, increasing this distance should change the magnitude of the coefficient β, but should not change the magnitude of α.

Another potential worry is that the influx of trucks could result in new developments on the treated roads (for example, restaurants or gas stations) which would also increase the number of cars and/or accidents on the road. If so, our estimates would be capturing the long-run impact of adding trucks to the road, including changes in infrastructure in response to the trucks. For the majority of routes this is not likely to be the case, because the routes are used for such a short time (typically less than a quarter). However, for the routes used by many different wells over a longer time horizon, then part of our estimate could be driven by new developments. We note that this is more likely to be the case on arterials and highways than on local-neighborhood and rural roads.

3.3 Estimation sample

Table 1 compares the average characteristics across treatment and control roads, before and after the shale boom (pre and post 2007). The first two columns show the sample of roads that at some point in time are traversed by a truck. Compared to all other roads in the state, these roads have more traffic as well as more accidents of any type. Recognizing that the impact of a truck will be different depending on the type of road, we divide the sample into two groups of roads: (1) main arterials and highways (these are primary and secondary roads, which include main arterials that have one or more lanes of traffic in each direction) and (2) local-neighborhood and rural roads

(these are tertiary roads, which include city streets and rural roads, with usually a single lane in each direction). The treated segments are more likely to be classified as main arterials/highways.

Because the summary statistics indicate that the truck routes differ in observable ways, we also construct a control group that is more comparable to the treatment group, following a strategy similar to Kline and Moretti (2013). Specifically, using pre-2007 characteristics, we estimate a probit model of the probability of being a traversed road and then drop roads with a predicted probability of being traversed in the bottom 75 percent.[23] The last two columns of Table 1 show the mean of the trimmed control group, pre- and post-shale boom.

Table 1: Summary Statistics

	Traversed roads		All other roads		Trimmed other roads	
	Mean Pre	Mean Post	Mean Pre	Mean Post	Mean Pre	Mean Post
A. Quarterly accident data:						
Truck accidents	.013	.015	.002	.002	.006	.006
Car accidents	.132	.162	.029	.041	.092	.104
Accidents with a fatality	.0022	.0024	.0003	.0004	.0012	.0011
Accidents with an injury	.077	.083	.017	.022	.053	.056
Truck routes	.001	1.138	.000	.000	.000	.000
I(Highway)	.149	.163	.035	.035	.160	.166
I(Local or rural road)	.851	.837	.965	.965	.840	.834
Obs.	2,035,809	1,155,999	18,502,211	10,826,145	3,667,023	2,122,76)
B. Annual traffic counts:						
Truck count	542	733	361	467	368	512
Car count	5,652	6,841	4,957	6,172	4,311	5,656
Obs.	69,018	96,516	124,080	128,115	51,190	46,211

Notes: All data are by year-quarter (1997-2014) except traffic counts, which are annual (2004-2014). "Traversed roads" are roads that were at some point in time used by at least one well to access a water withdrawal or disposal site; "All other" are all other roads never traversed; "Trimmed other" keeps only the other roads whose prediction of traversal, based on pre-shale (i.e., pre-2007) characteristics, is in the upper 75 percentile. "Pre" refers to pre-2007 and "Post" refers to post-2007.

In our main specifications, we use the trimmed control group, to reduce the difference between control and treatment roads. However, as we show in the Appendix (Table A1) the results are similar if we are even more conservative and restrict the sample to only the traversed roads (such that the control group consists of roads that are traversed at some time in the past or future).

[23] We match on pre-2007 characteristics of county-average total population, indicators for the three road types, the average vehicle accident rate, truck accident rate, fatality rate, and injury rate, and average annual daily vehicle and truck traffic counts.

4 Effects on Traffic Counts and Accidents

Impact of shale gas development at the county level. From the county-level analysis we find that drilling a well in the county-year increases the average daily truck count on the average segment by 0.79 trucks (first column of Table 2). This effect represents a 0.18 percent increase relative to the baseline truck traffic in counties that ever have a well. Car traffic also increases, with each additional well the average segment in a county has 7.54 more cars per day, or 0.16 percent of the average car traffic in treated counties. We can translate the coefficient on truck traffic into a prediction of the number of trucks associated with a shale gas well. The county-level estimates imply that in the year-county in which a well is drilled there are an additional 2,798 trucks.[24] This county-level estimate includes all trucks, those transporting sand and equipment, or trucks associated with the broader economic boom.

The last two columns of Table 2 show the effect of an additional well on the frequency of accidents in a county-quarter. The water and waste hauling trucks are concentrated in a short period of time (less than 90 days). When we estimate a per truck impact on accidents, we therefore assume that the annual traffic increase happens during the treatment quarter and our analysis on accident counts is presented at the quarterly level. Each well results in an additional .171 truck accidents in the county-quarter, and an additional .815 car accidents. The increase in accidents seen at the county level is a combination of all additional trucks on the road, as well as any county-wide changes in the number and demographics of drivers.

Impact of shale truck traffic at the segment level. At the segment-level (Table 3), we see that when a road is predicted to be used by a well, the average daily truck count increases by 2.093, implying an estimate of 764 waste and water trucks per well.[25] We note that the segment-level estimate (764) is less than county-wide estimate (2,798) but just at the range of government reports (800-2,400) (New York State Department of Environmental Conservation, 2011). The larger

[24] 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.

[25] The average daily truck count increase is the coefficient estimate divided by 10 multiplied by 365. 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.

Table 2: Impact of shale gas development on traffic and accidents at the county level

	Annual-average daily traffic count		Quarterly accident count	
	Truck	Car	Truck	Car
Wells	.79***	7.53***	.171***	.815***
	(.27)	(2.29)	(.023)	(.228)
Implied %-effect	.18***	.16***	1.01***	.27***
	(.06)	(.05)	(.14)	(.07)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No
Year-quarter FE	No	No	Yes	Yes
R^2	.39	.52	.93	.98
Obs.	663	663	4,824	4,824

Notes: Dependent variables are annual-average of daily truck count on segments, annual-average daily car count (i.e., non-truck count), quarterly count of accidents with a truck, and count of accidents between cars only. Traffic counts are by county-year, averaged across a county's segments (2004-2014) and accident counts are the total in the county-year-quarter (1997-2014). Wells are the count of wells drilled in the county-year (or county-year-quarter in the case of accidents). Robust standard errors are clustered by county.
*** Statistically significant at the 1% level; ** 5% level; * 10% level.

county-level estimate includes other types of trucks that are not concentrated on the water/waste routes. We do not see a statistically significant increase in car traffic on the truck routes (Column 2). The average daily car count increases by 1.708 when a road is used by a well, representing a 0.03 percent increase relative to the baseline. This effect is an order of magnitude smaller than the relative increase in truck traffic (0.3 percent), providing evidence that our country-half-year controls capture the increase in cars on the road.

The last two columns in Panel B show the segment-level approach to estimate the accident increase on roads used for shale gas wells. Specifically, we find .00033 more truck accidents in the segment-quarter when a road is used by one well (a 2.5 percent increase in truck accidents). Car accidents also increase on the truck routes (an additional .00169 car accidents per well, or a 1.2 percent increase in car accidents).

Impact of trucks by road type and distance from the well. We next investigate the impact of trucks on road safety by road types separately, acknowledging that there will be different impacts depending on whether the road is a main arterial/highway or a local-neighborhood/rural road. Furthermore, we also test the robustness of our estimates by including controls for proximity

Table 3: Impact of shale-gas trucks on traffic and accidents at segment level

	Annual-average daily traffic count		Quarterly accident count	
	Truck	Car	Truck	Car
Truck routes	20.93***	17.08	.0033**	.0169***
	(6.18)	(19.02)	(.0016)	(.0060)
Implied %-effect	3.03***	.30	25.32**	12.05***
	(.90)	(.33)	(11.97)	(4.26)
Segment FE	Yes	Yes	Yes	Yes
County-year FE	Yes	Yes	No	No
County-half-year FE	No	No	Yes	Yes
Year-quarter FE	No	No	Yes	Yes
R^2	.71	.80	.59	.81
Obs.	93,182	93,182	9,044,927	9,044,927

Notes: Dependent variables are annual-average of daily truck count on segments, annual-average daily car count (i.e., non-truck count), quarterly count of accidents with a truck, and count of accidents between cars only. Observations are by segment. Traffic counts are by segment-year (2004-2014) and accident counts are by segment-year-quarter (1997-2014). "Truck routes" are the count of wells (in counts of 10) in the year (in the case of traffic) and year-quarter (in the case of accidents) that are predicted to use the road segment in the year or quarter. Robust standard errors are clustered by road-segment.

*** Statistically significant at the 1% level; ** 5% level; * 10% level.

to a well. These controls capture the combined accident effect of trucks and workers driving to the well.

Table 4 shows regressions by road type, with and without controls for well proximity, revealing the importance of controlling for well proximity: for most regressions, after allowing for a differential effect on the routes within 2km of a well, our estimate of the accident externality of trucks becomes smaller. As expected, across highways (Panel A) and local/rural roads (Panel B), we estimate a similarly sized increase in truck traffic. Also, similar to Table 3, we estimate a small, statistically insignificant increase in cars on both types of roads: a statistically insignificant daily increase of 14.06 (.2%) cars on highways and 34.54 (.7%) cars on local and rural roads. On highways, we do not detect an increase in the number of accidents involving a truck but we do detect an increase in car collisions that do not involve a truck. When adding a truck to a local or rural road, we detect both more truck and car collisions.

Whether we can attribute our estimates to the impact of trucks alone, will depend on how well we are controlling for other changes on these routes. If the coefficient on truck routes remains constant, even far from the well itself, then this would provide evidence that the estimate is not confounded by changes associated with the well (e.g., cars needing to access the well pad). Figure 4

Table 4: Segment-level impacts on traffic and accidents, by road type

	Annual-average daily traffic count				Quarterly accident count			
	Truck		Car		Truck		Car	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Highways								
Truck routes	17.69**	17.90**	18.45	14.06	.0054	-.0015	.0378**	.0231*
	(7.72)	(8.78)	(22.74)	(25.40)	(.0041)	(.0029)	(.0153)	(.0119)
Truck routes*I(Near well)		-.08		1.94		.0471**		.1006**
		(1.05)		(3.99)		(.0199)		(.0430)
Implied %-effect	2.34**	2.37**	.30	.23	6.50	-1.85	4.76**	2.91*
	(1.02)	(1.16)	(.36)	(.41)	(5.00)	(3.51)	(1.92)	(1.50)
R^2	.67	.67	.76	.76	.60	.60	.81	.81
Obs.	37,512	37,512	37,512	37,512	1,420,374	1,420,374	1,420,374	1,420,374
B. Local and rural roads								
Truck routes	27.43***	25.92***	39.95	34.54	.0007***	.0005**	.0051**	.0043*
	(9.26)	(9.46)	(29.03)	(30.05)	(.0002)	(.0003)	(.0023)	(.0024)
Truck routes*I(Near well)		2.83		10.12		.0021***		.0087*
		(2.59)		(9.33)		(.0008)		(.0047)
Implied %-effect	4.30***	4.06***	.76	.66	102.34***	73.97**	21.24**	17.78*
S.E.	(1.45)	(1.48)	(.55)	(.57)	(35.20)	(37.13)	(9.54)	(10.00)
R^2	.75	.75	.85	.85	.10	.10	.59	.59
Obs.	55,670	55,670	55,670	55,670	7,624,553	7,624,553	7,624,553	7,624,553

Notes: Dependent variables are the annual-average daily truck count and daily car (non-truck) count and the segment-year-quarter log of the count of accidents with a truck, count of accidents between cars only. Traffic regressions include segment fixed effects and county-year fixed effects. Accident regressions include fixed effects for segment, county-half-year, and year-quarter.

"Truck routes" are the count of wells (in counts of 10) in the year-quarter that are predicted to use the road segment. "Truck routes*I(Near well)" are the counts within 2km of a recently drilled well.

Panel A: Subsample of roads classified as primary or secondary: main arteries that have one or more lanes of traffic in each direction.

Panel B: Subsample of roads classified as tertiary: local neighborhood roads, rural roads, and city streets.

Implied %-effect is calculated using the coefficient on Truck routes.
Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

shows coefficients from separate regressions, differing by the distance indicating whether a route is "near" a well (these coefficients are also displayed in Appendix Table A2). In the case of highways (Panels a and b), we find that the size of the estimate on truck routes hardly varies with the distance we use to control for proximity. Even when we separately control for truck routs within 5km of a well, we find a positive and statistically significant increase in car accidents on routes farther still. This suggests that adding a truck to a highway has no impact on truck-related accidents and appears to increase car collisions. On the other hand, we find a large significant increase in both truck and car accidents close to a well, and this effect declines with the distance from a well. Thus,

these coefficients likely capture something that is related to well access but not solely related to trucks.

For local and rural roads, we also find that the size of the estimate for truck routes remains constant, regardless of the distance we use to control for proximity. The estimates suggest that adding a truck to a local or rural road increases the incidence of truck-related accidents as well as car collisions. As in the case of highways, we find that the effect is significantly larger closer to a well, indicating that proximity to a well per se has an effect on accidents that is not related to trucks.

Calculating an accident estimate, per-truck. Using the estimate of the increase in the number of trucks (Table 4, Column 2) and the estimate of the increase in truck accidents (Column 6), we can calculate a per truck estimate.

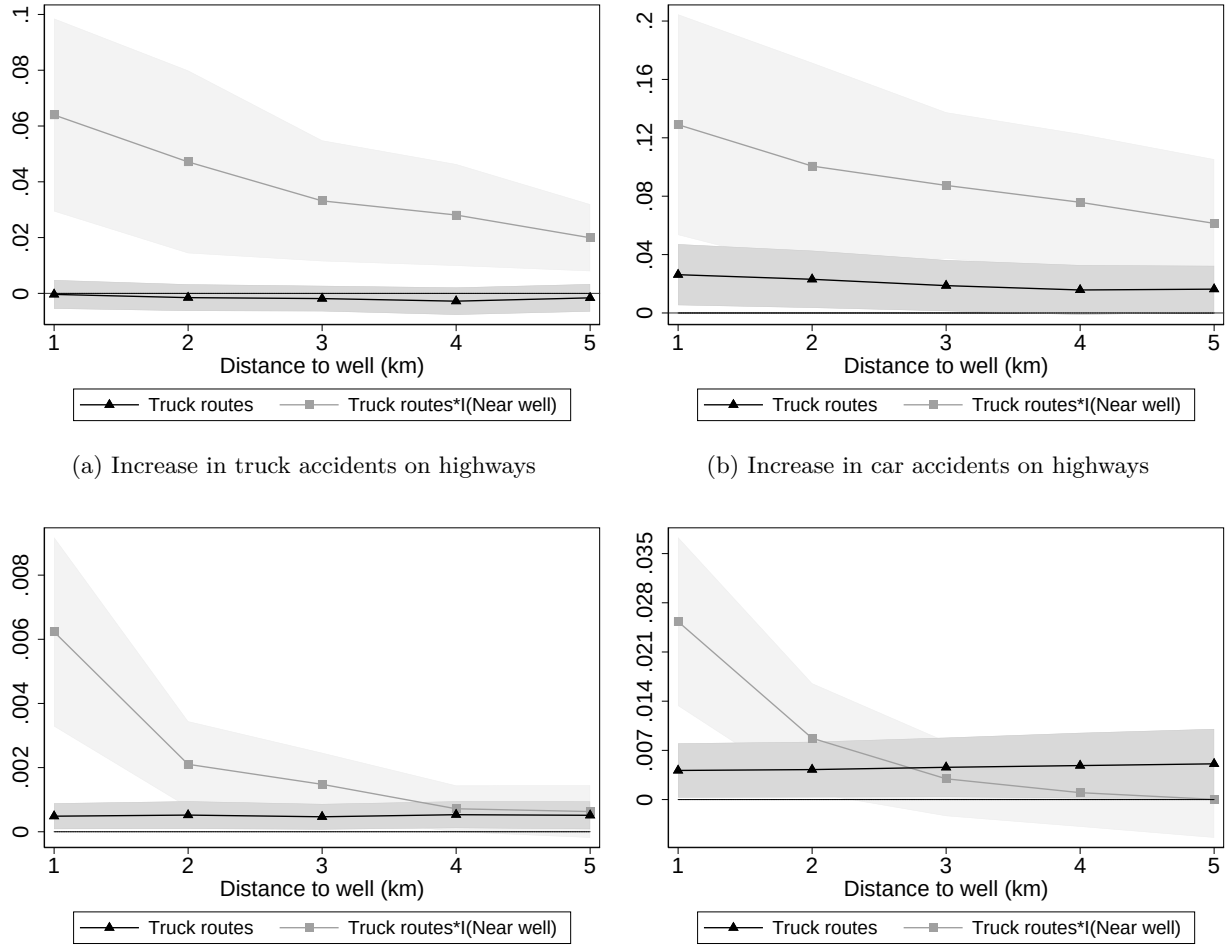
If the typical kilometers a truck drives in a year were to occur only on local and rural roads, the estimates imply annually one truck accident for every 128 trucks.[26] On the other hand, an additional truck has no significant impact on the number of truck accidents on highways.

To understand the size of these estimates, it is useful to look at how this would translate into a truck's insurance premium. If each truck accident was the fault of the truck and reached the current liability limit of $750,000, the actuarially fair insurance rate in this case would be $0/km for highways to $0.074/km for local/rural roads. The current average insurance rate of $0.057/km falls in this range.[27] The estimate for local/rural roads ($0.074/km) is larger in size than the CO_2 emissions costs ($0.043/km).[28] However, both of these estimates are smaller than the health costs associated with a truck's local air pollution. He, Gouveia and Salvo (Forthcoming) show that in São Paulo, NOx concentrations result in one hospitalization per year for every 10-20 trucks, and one death per year for every 100-200 trucks.

[26] The coefficient on truck routes suggests 0.00005 more truck accidents on rural roads after connecting one well. Dividing by the absolute increase in trucks (946 trucks) and the average length of a segment (.537km), 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), this results in 0.0078 truck accidents per year of truck driving on local and rural roads.

[27] 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 $0.028/km, whereas companies with fewer than 5 trucks pay $0.075/km. The average current rate is lower than the local/rural road estimate because this average would cover trucks traveling on all types of roads and also not all accidents would reach the full liability.

[28] Using the Federal Highway Administration's average fuel economy for heavy trucks (8.5 kilometers per gallon of gasoline equivalent), 0.0088 tons of CO_2 emissions per gallon, and a social cost of carbon dioxide of $42/ton.



(a) Increase in truck accidents on highways

(b) Increase in car accidents on highways

(c) Increase in truck accidents on local and rural roads

(d) Increase in car accidents on local and rural roads

Figure 4: Estimates depending on distance from a well

Notes: The figures present the coefficients on Truck routes, α, and on Truck routes*I(Near well), β, from separate regressions of equation (3). For each distance indicated, a regression is run, differing by the distance used to determine whether a route is near a well, I(Near well). Subfigures (a) and (b) uses the subsample of highways; (c) and (d) uses the subsample of local and rural roads. Subfigures (a) and (c) show the change in the number of truck accidents in the segment-year-quarter on truck routes and truck routes near wells; (b) and (d) show the change in the number of car accidents. Shaded areas represent 90% confidence intervals. Coefficients and standard errors are also presented in Table A2.

Parry (2008) provides an estimate for the optimal tax structure to account for externalities associated with trucking fuel use, such as local and global pollution, as well as externalities associated with kilometers traveled, such as congestion, truck accidents, pavement damage, noise (the optimal

tax includes a diesel fuel tax of \$.69/gallon and a per-kilometer tax of \$.04/km to \$.20/km).[29] We reveal an additional externality, that of more car accidents (Table 4), and therefore the optimal tax will be larger than previously thought. Using the estimate of the increase in car accidents (Column 8), the presence of a truck results in an even larger increase in the number of accidents between other road users. On local and rural roads, annually one car accident occurs for every 15 trucks. On highways annually one car accident occurs for every 10 trucks.

Impact on accident severity. Above we find more accidents involving trucks on local/rural roads and more car accidents on both highways and local/rural roads. With more accidents involving a heavy truck, we might expect to see more severe accidents. However, it appears that across all these new accidents, we don't see more injuries or fatalities (Table 5). In the case of adding a truck to a highway, we even estimate a decrease in the number of injuries, suggesting that the additional accidents are less severe. This could be because of the reduced speed on highways with trucks. Although we do not detect an increase in injuries and fatalities, the increase in the number of accidents still has a cost. To get a measure of the costs, we turn to data on insurance premiums (Section 5).

Placebo regression. We test the identifying assumption that trends in accidents on control roads provide good counterfactuals for trends in treatment roads in absence of treatment. Most wells were drilled between 2007 and 2014. We test for differential trends in accidents prior to shale gas drilling by recoding the observations so that, falsely, the roads are used eight years earlier, and we run the same regressions using the data from 1997 to 2006. If the trends are similar between treated and control roads in the absence of shale gas drilling, then we would expect the point estimates to be statistically insignificant. Indeed, the coefficients in this placebo test are statistically insignificant with the exception of injuries on highways (Table 6). We estimate a statistically significant negative impact on highways used by trucks, suggesting that our estimates for injuries could be downward biased. This could explain why we estimate fewer injuries on highways used by trucks (Table 5).

[29] The current federal diesel tax rate is \$.244/gallon and Pennsylvania's state diesel tax rate is \$.64/gallon. Pennsylvania does not have a tax per vehicle-miles-traveled, while other states do (Kentucky for example charges \$.02/km driven by heavy trucks and Oregon charges up to \$.18/km depending on the truck's weight and number of axels). Registration fees in Pennsylvania vary by class, from \$62 per year to \$1,664 per year.

Table 5: Injuries and fatalities

	Minor	Moderate	Major	Any injury	Fatality
A. Highways					
Truck routes	-.0250***	-.0075***	-.0027***	-.0155**	-.0002
	(.0052)	(.0019)	(.0010)	(.0062)	(.0006)
Truck routes*I(Near well)	-.0044	-.0161**	.0017	.0136	.0006
	(.0112)	(.0072)	(.0053)	(.0172)	(.0025)
Mean dep. var.	.3	.1	.028	.44	.013
R^2	.66	.43	.19	.75	.11
Obs.	1,420,374	1,420,374	1,420,374	1,420,374	1,420,374
B. Local and rural roads					
Truck routes	.0000	-.0001	-.0002	.0011	-.0001
	(.0010)	(.0003)	(.0002)	(.0015)	(.0000)
Truck routes*I(Near well)	.0020	.0010	-.0006	.0025	.0005
	(.0024)	(.0011)	(.0004)	(.0030)	(.0004)
Mean dep. var.	.0082	.003	.00088	.013	.00032
R^2	.37	.17	.05	.47	.03
Obs.	7,624,553	7,624,553	7,624,553	7,624,553	7,624,553

Notes: Dependent variables are, respectively, the count of accidents with one or more minor injury, moderate injury, major injury, any injury, or fatality. All regressions include fixed effects for segment, county-half year, and year-quarter. Panel A: Subsample of roads classified as primary or secondary: main arteries that have one or more lanes of traffic in each direction. Panel B: Subsample of roads classified as tertiary: local neighborhood roads, rural roads, and city streets. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table 6: Placebo test: Fictitious treatment dates

	Truck	Car	Injury	Fatal
A. Highways				
Truck routes	-.0016	.0007	-.0099**	.0002
	(.0027)	(.0072)	(.0047)	(.0007)
Truck routes*I(Near well)	-.0081	-.0127	-.0207*	-.0008
	(.0104)	(.0195)	(.0116)	(.0026)
Mean dep. var.	.087	.8	.47	.014
R^2	.61	.82	.76	.12
Obs.	797,002	797,002	797,002	797,002
B. Local and rural roads				
Truck routes	-.0001	-.0020	-.0014	-.0001
	(.0001)	(.0015)	(.0009)	(.0001)
Truck routes*I(Near well)	.0004	-.0017	-.0008	-.0002
	(.0006)	(.0033)	(.0024)	(.0003)
Mean dep. var.	.00062	.021	.012	.00031
R^2	.14	.64	.52	.05
Obs.	4,246,748	4,246,748	4,246,748	4,246,748

Notes: Treatment variables are given fictitious dates (specifically, all treatment variables are recoded to have occurred 8 years prior). Sample therefore covers 1997-2006. All regressions include fixed effects for segment, county-half year, and year-quarter. Panel A: Subsample of roads classified as primary or secondary: main arteries that have one or more lanes of traffic in each direction. Panel B: Subsample of roads classified as tertiary: local neighborhood roads, rural roads, and city streets. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Testing for differences in the type of accident. Above we use variation in shale gas truck traffic to estimate the impact of adding a truck to a road, but to do so, we are making the assumption that this variation is not correlated with unobservables that increase the number of accidents. We cannot definitively test whether this assumption holds, but we can look for evidence that it is violated. We would see differences in driver and vehicle characteristics on truck routes if the county-half-year fixed effects were not sufficiently controlling for boomtown impacts. We therefore run equation (3), but use as outcome variables the fraction of accidents occurring in a

Table 7: Share of accidents in a segment-quarter, by characteristic and road type

	Highways				Local or rural roads			
A. Driver characteristics								
	Shale Lic.	Male<25	Alcohol	Unbelted	Shale Lic.	Male<25	Alcohol	Unbelted
Truck routes	-.0002	-.0018	-.0005	-.0007	.0002	-.0034	-.0095	-.0059
	(.0007)	(.0027)	(.0017)	(.0023)	(.0012)	(.0070)	(.0058)	(.0060)
Truck routes*I(Near well)	.0008	.0083**	.0013	.0038	.0008	.0058	.0176	.0239
	(.0014)	(.0037)	(.0026)	(.0031)	(.0039)	(.0160)	(.0145)	(.0198)
Mean of dep. var.	.0046	.23	.1	.17	.0025	.27	.14	.19
R^2	.07	.08	.10	.10	.27	.23	.26	.25
Obs.	271,555	271,555	271,555	271,555	147,019	147,019	147,019	147,019
B. Accident characteristics								
	Aggressive	Speeding	Changing	Tailgating	Aggressive	Speeding	Changing	Tailgating
Truck routes	-.0039	-.0064**	.0025*	-.0015	.0254***	.0154**	.0023*	-.0008
	(.0033)	(.0026)	(.0015)	(.0014)	(.0083)	(.0077)	(.0014)	(.0021)
Truck routes*I(Near well)	.0008	.0013	-.0015	.0033**	-.0037	-.0194	-.0019	.0077
	(.0046)	(.0046)	(.0015)	(.0017)	(.0188)	(.0203)	(.0018)	(.0052)
Mean of dep. var.	.58	.25	.042	.06	.54	.28	.0065	.028
R^2	.15	.21	.18	.15	.29	.35	.21	.20
Obs.	271,555	271,555	271,555	271,555	147,019	147,019	147,019	147,019

Notes: Dependent variable in each column is the share of accidents in a segment-quarter with a characteristic listed in the column heading. Additional characteristics can be found in Appendix Table A4. Shale license refers to share of accidents in the segment-quarter that involve a driver with a license from Arkansas, Louisiana, Oklahoma, or Texas (the states outside of Pennsylvania producing the most shale gas). All specifications include fixed effects for segment, county-half-year, and year-quarter. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

segment-quarter that are attributed to accidents with particular characteristics.[30]

Consistent with our specification isolating a truck-only effect, we do not find a statistically significant difference on truck routes in the share of accidents by driver and car characteristics. Truck routes are similar in characteristics associated with a shale boom (such as drivers with licenses from the main shale producing states, men younger than 25, Table 7, cars registered in shale states, fatigued drivers, or share of luxury vehicles, Table A4). We do see some statistically

[30]Driver characteristics include the share of accidents with a driver with a license from the largest shale producing states (Arkansas, Louisiana, Oklahoma, or Texas). Including Pennsylvania, 95% of production came from these states according to the EIA's production estimates for 2010. We also examine the share of accidents with a male driver under 25 years of age, share of accidents that are alcohol related, or have an unbelted driver or passenger. Other characteristics specific to the accident include the share of accidents with an aggressive driving indicator, speeding indicator, merging/changing-lanes indicator, or tailgating indicator. In the Appendix, Table A4, we show the share of accidents with vehicles registered in the four largest shale states, share of accidents involving a luxury vehicle (as defined by the make of car being an Acura, Audi, BMW, Buick, Cadillac, INFINITI, Jaguar, Land Rover, Lexus, Lincoln, Mercedez-Benz, Porsche, or Volvo), the average age of the vehicles in the accident, the share of accidents attributed to avoiding an object (or animal, pedestrian, or vehicle), the share of accidents with a distracted-driver, the share of accidents with a fatigue or asleep indicator, share of accidents with male drivers aged 25-50, or share of accidents with male drivers 50 years or older.

significant coefficients in the case of characteristics of driving behavior, specifically, the share of accidents attributed to aggressive driving, speeding, and changing lanes. We find highways used by trucks result in less speeding, perhaps driven by congestion or defensive driving near trucks. This can be safety enhancing because with higher speeds comes more accidents (van Benthem, 2015). When local and rural roads are used by trucks, we see the opposite: an increase in speeding and aggressive driving. This increase could be from drivers attempting to make up lost time behind a truck or that aggressive driving is needed in order to pass trucks on local roads. Consistent with trucks changing driving behavior, we see an increase in accidents associated with changing lanes across both types of roads. Also reassuringly we do not find an impact on other risky-driving indicators (such as tailgating, alcohol, not using a seatbelt, or in the Appendix, distracted-driving). On highways, vehicles are newer, but only by less than a month (less than one percent newer) and there are slightly fewer men aged 25-50 (two percent fewer).

Running similar regressions at the county level (Table A5), demonstrates the power of the county-half-year controls in the road-segment specification. When we look at shale gas development in general at the county level, we find that with each well drilled in the county, there are more accidents with drivers from shale states as well as vehicles registered in shale states. There are more males aged 25-50 as well as unbelted drivers. These characteristics do not show up in the road-segment specification, implying that control and treatment segments in the same county see similar increases. At the county level, similar to the road-segment level, vehicles are newer (but again by less than one percent) and counter-intuitive to what one might expect from a shale boom, there are fewer luxury vehicles (but only by a very small amount, of less than one percent).

5 Valuation of the external costs to customers of car insurance

The previous sections provide evidence that adding one truck to a road creates an accident externality: a truck on the road increases the number of car accidents not involving a truck. Furthermore, the increase in truck accidents on local and rural roads could also create an externality, especially if the trucks are underinsured. Here we look for evidence whether the accident externalities propagate into insurance premiums by looking at the change in premiums offered to a representative new enrollee of car insurance.

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. 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.[31] Using these data, we make the assumption that changes to this representative driver's insurance rates will shed light on changes likely happening to other individuals (e.g., with different ages, genders, or cars). Importantly, the data are quotes for the same hypothetical person, which is an advantage over using population-average data on existing 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. Using our GIS-predicted routes we calculate the total number of segments used by trucks within 25km of the centroid of a zip code (the total number of segment well-connections in a year-zip code), given that most accidents occur within 25km of one's residence.[32]

Table 8: Summary statistics at the zip-code level

	Traversed zip codes		Nontraversed zip codes	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Average premium ($)	1076.3	(87.2)	1481.9	(472.2)
Δ in premium between 2012 to 2014 ($)	63.7	(30.2)	30.5	(84.0)
Truck routes (total, in 1000s)	7.30	(11.28)	0	0
Wells	23.92	(47.88)	0	0
Obs.	2,433		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.

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). But the

[31] 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.

[32] Abdalla et al. (1997) show that most casualties in accidents occur within 25km of one's residence and according to an insurance company survey, 77 percent of accidents occur within 24km of one's residence. https://www.progressive.com/newsroom/article/2002/may/fivemiles/

traversed zip codes saw a larger increase in premiums between the two years (a $64 increase versus a $31 increase).

We run a regression in which we regress the zip code's average insurance premium on the number of segments used by wells in and around the zip code (within 25km of the center of the zip code). The regression includes the count of wells drilled in the county-year, to capture impacts from local shale gas production that don't necessarily have to do with the number of segments traversed by trucks, as well as 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 coefficient on truck routes represents the dollar change in premiums when 1,000 segments are used in or near the zip code (the count increases for every segment used by every well in the quarter). When one road segment of a truck route, used by one well, is in or near a zip code, the insurance premiums increase by less than a penny, however, when a zip code is near a truck route, it is often not only near one segment used by one well; the most heavily traversed zip code is near 76 thousand segment/well connections, which translates to a premium increase of $30. Furthermore, remember that this increase could be applied to all new insurance enrollees, making aggregate costs larger. If we assumed that all Pennsylvanians over the age of 20 saw the predicted increase on their insurance premiums, then this would aggregate to an externality of $18 million dollars in 2014 from shale gas development.[33]

We can also use the coefficient estimate to calculate a per-truck estimate. We estimate that annual insurance premiums of new enrollees increases by $0.06 per truck.[34] 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 for three reasons: our treatment is concentrated within a zip code rather than dispersed across a whole state; our outcome variable is the insurance plan offered to a new enrollee, which will adjust faster than the average insurance

[33]To obtain this number we use county-level population estimates from the US Census. This will result in an upper bound of the total increase in insurance rates. We assume that all car insurance holders see an increase in their premiums, not only the new enrollees. We assume that all individuals over the age of 20 hold car insurance. And we assume that the insurance increase from a truck route in 25km of a zip code is similar to the insurance increase from a truck route in a county. These assumptions all inflate the size of the cost estimate. Using these numbers, with the number of truck routes in 2014, the cost of insurance across all of Pennsylvania increased by $18 million dollars (with Allegheny county, Washington county, and Westmoreland county seeing the largest increases).

[34]To get the per-truck-year increase in premiums, we first divide the coefficients by the number of trucks per segment (764) and the number of kilometers in a segment (0.665km per segment). Then we multiply by the annual average of kilometers traveled, 79,060 km. The estimate implies a 6 cents increase ($0.0004/764/0.665 \times 79,060$).

plan of all existing insurance contracts; and our treatment is an additional truck, which will pose more risk than a car by function of its size and typical kilometers traveled. The magnitude of the costs will depend on how many new enrollees see insurance increases. Using county-level population counts again we can get an upper bound on this estimate. If we assume all Pennsylvanians over the age of 20 see an insurance increase, then one truck would cause an externality of $8,469 in aggregate.[35]

Table 9: Car Insurance Premiums on Truck Routes

	Average Premium (Dollars)	Uninsured (Share)	Luxury (Share)	Veh. Age (Years)	Veh. Thefts (County Count)
Truck routes (total, in 1000s)	.401***	-.000	.000	-.003	.710
	(.131)	(.000)	(.000)	(.002)	(.892)
Wells	-.011	-.000	-.000***	-.003***	.071
	(.047)	(.000)	(.000)	(.001)	(.113)
Year FE	Yes	Yes	Yes	Yes	Yes
Zip code FE	Yes	Yes	Yes	Yes	No
County FE	No	No	No	No	Yes
R^2	.99	.86	.99	.93	1.00
Obs.	3,305	2,965	2,965	2,965	134
Mean of dep. var.	1,183	.0397	.125	9.52	81.2

Notes: Dependent variables are (1) the average insurance premium offered across six national insurance providers for the same hypothetical new insuree; (2) the zip code's share of accidents with one or more uninsured driver; (3) the zip code's share of accidents with one or more luxury vehicles; (4) the average age of vehicles in accidents in the zip code; (5) the number of vehicle thefts in the county.
In the first four columns, observations are by zip code and year (for the years 2012 and 2014) and truck routes are the count of segments used within 25km of the zip code centroid (in count of 1000s). In the last column, observations are by county and year (2012 and 2014) and truck routes are the count of segements used in the county (in count of 1000s). Robust standard errors are clustered by zip code in columns 1-4 column (or county the last column). *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Interestingly, the count of truck routes is more important than the count of the wells themselves; the coefficient on the count of wells drilled in and around the zip code is statistically insignificant. This implies that the impacts from shale gas are felt beyond the location of the wells themselves. Pennsylvania collects an impact fee from wells, which is then distributed back to communities, with a larger share going back to municipalities in which wells are drilled. This estimate implies that it is appropriate to spread the impact fees to municipalities without wells but with truck routes.

Although we are using the count of traversed-road-segments and controlling for wells drilled nearby, one might be concerned that the increase in premiums is driven by something other than

[35]To get the increase per truck, we divide the Pennsylvania-wide cost, of $18 million, by the total number of truck routes (scaled by how many are driven in a year by one truck) as well as the number of trucks per well.

trucks. Sun and Yannelis (2015) find that the presence of uninsured drivers on the road increases car insurance premiums (they find a one percent increase in uninsured drivers increases premiums by $27, over an average annual premium of $2,356). However, uninsured drivers are not likely the driving cause of our finding because we do not find a statistically significant increase in the share of accidents involving uninsured drivers on the truck routes. Another pathway for increased premiums is if nearby accidents involve more expensive vehicles and accident damages are therefore larger. To examine this pathway, we look at the share of accidents involving luxury vehicles, as well as the average age of the vehicles in collisions, but we do not see a statistically significant difference in the zip codes with many truck routes.[36] Another possible pathway for increased premiums is if there are more vehicle thefts, 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 as the zip code regressions, we also examine how vehicle thefts in a county are related to the number of traversed segments in a county and find they are statistically insignificant.[37]

6 Conclusion

We find that the addition of a single truck to the road not only increases the number of accidents involving a truck, but also 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 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.

For example, if a car has the misfortune to crash into a truck, total damages will be larger than had the car crashed into an equally sized car. These damages will fall on the negligent party, in this case the car and not the truck, and therefore would not be internalized into the decision of how much to truck. In the case that the truck is the negligent party, current liability limits are low enough to allow for the possibility of a judgment-proof firm. Trucks must carry insurance, or post a surety bond, to cover accidents costing $750,000, a limit that has not grown with inflation

[36] Similar to the county regressions on the type of accident (Table A5), near wells (not truck routes) we see fewer luxury vehicles and also newer vehicles.

[37] Vehicle theft counts at the county level were obtained from Pennsylvania's Uniform Crime Reporting System.

over the past 30 years. If accident costs are more than a trucking company's assets, the possibility of bankruptcy could mean these costs wouldn't be fully internalized. Accordingly, the accident externalities associated with trucks on the road appear to be dispersed across other road users through higher insurance premiums.

We find suggestive evidence that the accident externalities associated with trucking increase the premiums offered to new insurance enrollees. Internalizing these external costs would require an ambitious revamping of the current liability regime or implementing a tax on the kilometers traveled by trucks. Several countries (Germany, Austria, and Poland) and US states (Kentucky, New Mexico, New York, and Oregon) already levy taxes for the distance traveled by heavy trucks and the information technology revolution should make it easier for other states to follow suit. 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 all accident externalities. Specifically, this paper reveals additional car collisions that occur because trucks are on the road. Unless there were similar safety advances for cars, crashes in the proximity of automated trucks would 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."

Abdalla, Ibrahim M, Robert Raeside, Derek Barker, and David RD McGuigan. 1997. "An investigation into the relationships between area social characteristics and road accident casualties." *Accident Analysis & Prevention*, 29(5): 583–593.

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.

Austin, David. 2015. "Pricing freight transport to account for external costs." *Congressional Budget Office Working Paper Series.*

Bartik, Alexander Wickman, Janet Currie, Michael Greenstone, and Christopher R Knittel. 2016. "The local economic and welfare consequences of hydraulic fracturing."

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

Cohen, Linda R., and Kevin D. Roth. 2018. "A Second-best Dilemma: Freight Trucks, Externalities, and the Dispatch Effect." *Working Paper.*

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

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, 108(4-5): 899–934.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics*, 131(2): 799–858.

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

Fetzer, Thiemo. 2014. "Fracking growth."

Feyrer, James, Erin T Mansur, and Bruce Sacerdote. 2017. "Geographic dispersion of economic shocks: Evidence from the fracking revolution." *The American Economic Review*, 107(4): 1313–1334.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr. 2016. "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing." *The Economic Journal*, 126(591): 317–357.

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.

He, Jiaxiu, Nelson Gouveia, and Alberto Salvo. Forthcoming. "External effects of diesel trucks circulating inside the Sao Paulo megacity." *Journal of the European Economic Association.*

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

Jacobsen, Grant. 2016. "Who wins in an energy boom? Evidence from wage rates and housing."

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."

Kline, Patrick, and Enrico Moretti. 2013. "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics*, 129(1): 275–331.

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.

Maniloff, Peter, and Ralph Mastromonaco. 2017. "The local employment impacts of fracking: A national study." *Resource and Energy Economics*, 49: 62–85.

Mason, Charles F, Lucija A Muehlenbachs, and Sheila M Olmstead. 2015. "The economics of shale gas development." *Annual Review of Resource Economics*, 7(1): 269–289.

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."

Parry, Ian WH. 2008. "How should heavy-duty trucks be taxed?" *Journal of Urban Economics*, 63(2): 651–668.

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."

van Benthem, Arthur. 2015. "What is the optimal speed limit on freeways?" *Journal of Public Economics*, 124: 44–62.

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.

Weber, Jeremy G. 2012. "The effects of a natural gas boom on employment and income in Colorado, Texas, and Wyoming." *Energy Economics*, 34(5): 1580–1588.

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

Xu, Minhong, and Yilan Xu. 2018. "Fraccidents: The Impact of Fracking on Road Traffic Deaths."

A Appendix (For Online Publication)

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

A.1 Exploring different control groups

In the paper our regression samples include the trimmed control roads. Here we show that our results are robust to using the full sample as well as restricting the sample to those segments that are treated at some point in time. Table A1 shows results do not change across different control groups.

A.2 Increasing the distance of "near" a well

Here we present, in tabular form, the coefficients in Figure 4. Table A2 shows results from separate regressions, each with different definitions of what is near a well (indicating roads that are within 1km of a recently drilled, up to 5km of a recently drilled well). In all cases the coefficient on truck routes is not statistically significantly different across different distances used. In the case of local and rural roads, the impact of being near a well dissipates outside of 3km (but this is not the case for highways).

A.3 Placebo test at the county level: Fictitious treatment dates

The paper reports a placebo test for the road-segment regressions. Here we present a placebo test for the county-level regressions. Table A3 reports county-level estimates when the treatment variables are given fictious dates (specifically, all treatment variables are recorded to have occurred 8 years prior and the sample therefore covers 1997-2006). Outcomes are the number of truck accidents, the number of car accidents, the number of fatalities, and the number of injuries in a county-quarter. For all outcomes in this table, the fictious number of wells drilled in a year-quarter has no significant impact, providing evidence for the validity of the common trends assumption.

A.4 Are the accidents different? Accident shares by additional characteristics

In the main part of the paper we examine whether the share of accidents of different types is different on treatment roads in the quarter used by wells (Table 7). Here we expand that analysis to include different outcome variables. Table A4) shows regressions using different outcome variables of the share of accidents in a segment-quarter that have one or more vehicle/driver with certain characteristics. Differences in observable characteristics could imply that our county-half-year fixed effects are not capturing all changes that influence the number of accidents on the treatment routes. We show most characteristics are statistically insignificantly different on the treated truck routes. Two exceptions are that, on highways, vehicles are newer (but only by less than a month, less than one percent newer) and men aged 25-50 are fewer (two percent fewer).

Table A5 shows similar regressions at the county-level. With each shale gas well drilled in the county-quarter, there is a larger share of accidents involving drivers with licenses from shale states and a larger share of accidents involving vehicles registered in shale states. There are also more males aged 25-50 as well as unbelted drivers. These characteristics do not show up as significantly different in the road-segment specification, implying that control and treatment segments in the county are similar. At the county level, similar to the road-segment regressions, vehicles are newer (but again by less than one percent) and counter-intuitive to what one might expect from a shale boom, there are fewer luxury vehicles (but only a very small difference, of less than one percent).

A.5 Treatment effect is time constant

Shale gas extraction has rapidly grown but has also changed over time. 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 post-2011. Table A6 shows that the interaction of post-2011 on our truck routes is not statistically significant. While this suggests that the impact is constant across years, the coefficient on truck routes is

larger when we include this additional regressor, so it is possible that shale gas development is less of a problem for road safety in recent years.

A.6 Impact of road curvature

Transportation research has found that road curvature reduces accidents (Haynes et al., 2007, 2008; Wang et al., 2009) or has a mixed impact (Milton and Mannering, 1998). We investigate whether the safety impacts of a truck depend on road curvature. We create a measure of road curvature by dividing the length of the road with its displacement (the Euclidian distance between the location of the beginning and the end of the road).[1] The ratio of length to displacement will be larger when the road is curvier. We construct an indicator for roads in the upper 75th percentile of this measure of curvature and find no differential impact of trucks on roads above and below the 75th percentile of the curvature distribution.

We also aggregated the road-specific curvature to the zip-code average and created a dummy for zip codes in the 75th percentile of road curvature. Table A8 shows that curvature also does not appear to have an effect on average premiums.

[1] A shortcoming with this measure of curvature is that it might be length/location dependent (e.g., areas with many short segments might appear straight, though when joined result in a curvy roadway).

Table A1: Robustness: Regressions using different control groups, by road type

	Truck	Car	Injury	Fatality
A. Highways, full sample				
Truck routes	-.0016	.0241**	-.0154**	-.0002
	(.0029)	(.0120)	(.0062)	(.0006)
Truck routes*I(Near well)	.0472**	.1000**	.0136	.0007
	(.0199)	(.0430)	(.0172)	(.0025)
Mean dep. var.	.083	.79	.44	.013
R^2	.62	.86	.84	.14
Obs.	1,509,882	1,509,882	1,509,882	1,509,882
B. Highways, sample of only ever-traversed roads				
Truck routes	-.0012	.0232*	-.0129**	-.0001
	(.0030)	(.0127)	(.0064)	(.0002)
Truck routes*I(Near well)	.0466**	.1008**	.0119	.0005
	(.0200)	(.0430)	(.0172)	(.0005)
Mean dep. var.	.083	.79	.44	.013
R^2	.65	.81	.74	.12
Obs.	488,935	488,935	488,935	488,935
C. Local and rural roads, full sample				
Truck routes	.0006**	.0074***	.0023	-.0001
	(.0003)	(.0025)	(.0015)	(.0001)
Truck routes*I(Near well)	.0022***	.0079	.0022	.0005
	(.0008)	(.0048)	(.0030)	(.0004)
Mean dep. var.	.0007	.024	.013	.00032
R^2	.07	.48	.37	.03
Obs.	31,341,355	31,341,355	31,341,355	31,341,355
D. Local and rural roads, sample of only ever-traversed roads				
Truck routes	.0005*	.0044*	.0011	-.0001*
	(.0003)	(.0024)	(.0015)	(.0001)
Truck routes*I(Near well)	.0020**	.0086*	.0024	.0005
	(.0008)	(.0047)	(.0030)	(.0004)
Mean dep. var.	.0007	.024	.013	.00032
R^2	.09	.51	.39	.03
Obs.	2,738,784	2,738,784	2,738,784	2,738,784

Notes: This table shows results with different control roads (in tables in the paper, control roads are those that are similar to treatment roads based on pre-shale-boom characteristics). All regressions include fixed effects for segment, county-half year, and year-quarter. Panels A and C show results from the full sample of roads: control roads are all other roads. Panels B and D show results from the subsample of only those roads that, at some point in time, are traversed: control roads are those that either earlier or later were traversed. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A2: Robustness: Different definitions of near well, by road type

	Regressions varying distance from wells:				
	1km	2km	3km	4km	5km
A. Highways: Truck accidents					
Truck routes	-.0004	-.0015	-.0019	-.0028	-.0016
	(.0031)	(.0029)	(.0028)	(.0030)	(.0030)
Truck routes*I(Near well)	.0640***	.0471**	.0332**	.0281**	.0199***
	(.0210)	(.0199)	(.0132)	(.0111)	(.0073)
Mean dep. var.	.083	.083	.083	.083	.083
R^2	.60	.60	.60	.60	.60
Obs.	1,420,374	1,420,374	1,420,374	1,420,374	1,420,374
B. Highways: Car accidents					
Truck routes	.0262**	.0231*	.0187*	.0158	.0163*
	(.0127)	(.0119)	(.0106)	(.0103)	(.0097)
Truck routes*I(Near well)	.1290***	.1006**	.0874***	.0758***	.0614**
	(.0459)	(.0430)	(.0304)	(.0284)	(.0267)
Mean dep. var.	.79	.79	.79	.79	.79
R^2	.81	.81	.81	.81	.81
Obs.	1,420,374	1,420,374	1,420,374	1,420,374	1,420,374
C. Local and rural roads: Truck accidents					
Truck routes	.0005**	.0005**	.0005*	.0005**	.0005*
	(.0002)	(.0003)	(.0002)	(.0003)	(.0003)
Truck routes*I(Near well)	.0062***	.0021***	.0015**	.0007	.0006
	(.0018)	(.0008)	(.0006)	(.0004)	(.0005)
Mean dep. var.	.0007	.0007	.0007	.0007	.0007
R^2	.10	.10	.10	.10	.10
Obs.	7,624,553	7,624,553	7,624,553	7,624,553	7,624,553
D. Local and rural roads: Car accidents					
Truck routes	.0041*	.0043*	.0046*	.0048*	.0051*
	(.0023)	(.0024)	(.0026)	(.0028)	(.0030)
Truck routes*I(Near well)	.0253***	.0087*	.0030	.0010	.0000
	(.0073)	(.0047)	(.0032)	(.0029)	(.0033)
Mean dep. var.	.024	.024	.024	.024	.024
R^2	.59	.59	.59	.59	.59
Obs.	7,624,553	7,624,553	7,624,553	7,624,553	7,624,553

Notes: Table displays coefficients found in Figure 4.

Each column in each panel represents a separate regression, differentiated by the distance that is considered "near a well."

The first two panels are the sub-sample of highways and the last two panels are the sub-sample of local and rural roads. The dependent variable in Panels A and C are the count of collisions involving one or more trucks in the segment-quarter. The dependent variable in Panels B and D are the count of collisions not involving a truck in the segment quarter.

All specifications include segment, county-half-year, and year-quarter fixed effects.

Robust standard errors are clustered by road segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A3: Placebo test at county level: Fictitious treatment dates

	Truck	Car	Injury	Fatal
Wells	0.0197	-0.0972	-0.0362	0.0036
	(0.0173)	(0.2210)	(0.1174)	(0.0045)
County fixed effects	Yes	Yes	Yes	Yes
Year-quarter fixed effects	Yes	Yes	Yes	Yes
Mean of dep. var. (treated, in levels)	19	266	146	3.5
R^2	0.94	0.99	0.99	0.82
Obs.	2,680	2,680	2,680	2,680

Notes: Treatment variables are given fictitious dates (specifically, all treatment variables are recoded to have occurred 8 years prior). Wells are the count of wells drilled in the county-year-quarter. Robust standard errors are clustered by county. Sample therefore covers 1997-2006. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A4: Type of accident: Share of accidents in a segment-quarter, by road type

	Highways				Local or rural roads			
A. Driver characteristics								
	Distracted	Fatigue	Male 25-50	Male 50+	Distracted	Fatigue	Male 25-50	Male 50+
Truck routes	-.0023	-.0005	-.0092***	-.0000	-.0041	.0012	-.0092	-.0044
	(.0021)	(.0009)	(.0032)	(.0026)	(.0062)	(.0025)	(.0082)	(.0064)
Truck routes*I(Near well)	-.0008	.0009	.0047	.0010	-.0037	-.0017	.0221	-.0051
	(.0028)	(.0014)	(.0042)	(.0034)	(.0082)	(.0067)	(.0227)	(.0117)
Mean of dep. var.	.08	.025	.38	.23	.086	.019	.33	.18
R^2	.11	.09	.10	.09	.26	.23	.23	.23
Obs.	271,555	271,555	271,555	271,555	147,019	147,019	147,019	147,019
B. Accident characteristics								
	Reg. Shale	Luxury	Veh. Age	Avoiding	Reg. Shale	Luxury	Veh. Age	Avoiding
Truck routes	-.0007	.0006	-.0730*	-.0007	-.0018	.0026	-.0760	.0032
	(.0009)	(.0019)	(.0379)	(.0007)	(.0013)	(.0050)	(.0975)	(.0026)
Truck routes*I(Near well)	.0003	.0046	-.0114	-.0001	.0013	-.0056	.1970	-.0035
	(.0012)	(.0029)	(.0543)	(.0011)	(.0043)	(.0081)	(.2175)	(.0053)
Mean of dep. var.	.0073	.11	7.4	.026	.0034	.1	8.1	.03
R^2	.09	.10	.18	.09	.26	.23	.30	.22
Obs.	271,555	271,555	271,555	271,555	147,019	147,019	147,019	147,019

Notes: This table provides additional outcome variables from those found in Table 7. Dependent variables in each column are the share of accidents with one or more of the characteristics listed in the column headings, except for Veh. Age which refers to the average age of the vehicles involved in accidents. Reg. Shale refers to share of accidents in the segment-quarter that involve a car registered in Arkansas, Louisiana, Oklahoma, or Texas (the states outside of Pennsylvania producing the most shale gas). Shale Lic. refers to the share of accidents involving drivers with license from one of these states. All specifications include fixed effects for segment, county-half-year, and year-quarter. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A5: Type of accident: County-level share of accidents

A. Driver characteristics

	Shale Lic.	Alcohol	Unbelted	Distracted	Fatigue	Male<25	Male 25-50	Male 50+
Wells	.0003***	-.0001	.0002***	-.0001	.0001**	.0001	.0005***	-.0001
	(.0001)	(.0001)	(.0001)	(.0001)	(.0000)	(.0000)	(.0001)	(.0001)
Mean dep. var.	.0046	.12	.18	.078	.029	.24	.34	.21
R^2	.19	.37	.56	.54	.31	.38	.55	.24
N	4,824	4,824	4,824	4,824	4,824	4,824	4,824	4,824

B. Accident characteristics

	Aggressive	Speeding	Changing	Tailgating	Reg. Shale	Luxury	Veh. Age	Avoiding
Wells	.0001	.0001	.0000	.0001	.0005***	-.0002***	-.0077***	-.0000
	(.0002)	(.0002)	(.0000)	(.0001)	(.0001)	(.0001)	(.0015)	(.0001)
Mean dep. var.	.55	.31	.02	.039	.007	.088	8.3	.031
R^2	.51	.69	.54	.61	.29	.64	.66	.37
N	4,824	4,824	4,824	4,824	4,824	4,824	4,823	4,824

Notes: This table provides county-level regressions of the characteristics found in Tables 7 and A4. Dependent variables are the share of accidents in the county-year-quarter with one or more of the characteristics listed in the column headings, except for Veh. Age which refers to the average age of the vehicles involved in accidents. Wells are the count of wells drilled in the county-year-quarter. Reg. Shale refers to share of accidents in the segment-quarter that involve a car registered in Arkansas, Louisiana, Oklahoma, or Texas (the states outside of Pennsylvania producing the most shale gas). Shale Lic. refers to the share of accidents involving drivers with license from one of these states. All regressions include fixed effects for year-quarter and county. Robust standard errors are clustered by county. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A6: Robustness: Is the impact constant over time?

	Truck	Car	Injury	Fatality
A. Highways				
Truck routes	-.0032	.0334**	-.0076	.0015
	(.0066)	(.0158)	(.0095)	(.0016)
Truck routes*I(Post-2011)	.0023	-.0139	-.0106	-.0023
	(.0072)	(.0166)	(.0086)	(.0019)
Truck routes*I(Near well)	.0470**	.1012**	.0140	.0007
	(.0198)	(.0429)	(.0172)	(.0026)
Mean of dep. var.	.083	.79	.44	.013
R^2	.60	.81	.75	.11
Obs.	1,420,374	1,420,374	1,420,374	1,420,374
B. Local and rural roads				
Truck routes	.0010***	.0050**	.0009	.0001
	(.0004)	(.0024)	(.0016)	(.0001)
Truck routes*I(Post-2011)	-.0006	-.0009	.0002	-.0002
	(.0004)	(.0029)	(.0019)	(.0001)
Truck routes*I(Near well)	.0021***	.0087*	.0025	.0005
	(.0008)	(.0047)	(.0030)	(.0004)
Mean of dep. var.	.0007	.024	.013	.00032
R^2	.10	.59	.47	.03
Obs.	7,624,553	7,624,553	7,624,553	7,624,553

Notes: Specifications include an interaction with an indicator for post-2011. The coefficient is statistically insignificant, which should alleviate the concern that in recent years more water was shipped using pipelines or reused at the well site. All regressions include fixed effects for segment, county-half year, and year-quarter. Panel A: Subsample of roads classified as primary or secondary: main arteries that have one or more lanes of traffic in each direction. Panel B: Subsample of roads classified as tertiary: local neighborhood roads, rural roads, and city streets. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A7: Accidents on a measure of road curvature

	Truck	Car	Injury	Fatal
A. Highways				
Truck routes	-.0116	.0384*	-.0147	-.0004
	(.0076)	(.0221)	(.0157)	(.0009)
Truck routes*I(Near well)	.0705	.2340*	.0311	-.0042
	(.0614)	(.1338)	(.0414)	(.0044)
Truck routes*I(Curvy road)	.0129	.0207	.0186	.0016
	(.0097)	(.0395)	(.0174)	(.0014)
Mean dep. var.	.15	.66	.36	.0088
R^2	.69	.78	.72	.15
Obs.	601,947	601,947	601,947	601,947
B. Local and rural roads				
Truck routes	.0006**	.0054*	.0024	.0000
	(.0003)	(.0031)	(.0017)	(.0001)
Truck routes*I(Near well)	.0021***	.0090*	.0028	.0005
	(.0008)	(.0046)	(.0029)	(.0004)
Truck routes*I(Curvy road)	-.0002	-.0036	-.0043	-.0003**
	(.0006)	(.0041)	(.0031)	(.0002)
Mean dep. var.	.0007	.024	.013	.00032
R^2	.10	.59	.47	.03
Obs.	7,624,553	7,624,553	7,624,553	7,624,553

Notes: Coefficient added to capture the differential impact of trucks on curvy roads. I(Curvy road) is an indicator for the road segment being in the 75th percentile of curvature, as measured as the length of the road divided by its displacment. All regressions include fixed effects for segment, county-half year, and year-quarter. Panel A: Subsample of roads classified as primary or secondary: main arteries that have one or more lanes of traffic in each direction. Panel B: Subsample of roads classified as tertiary: local neighborhood roads, rural roads, and city streets. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Table A8: Car Insurance Premiums on a Measure of Road Curvature

	Average Premium (in $)
Truck routes (total, in 1000s)	.394***
	(.130)
Wells	-.011
	(.047)
Truck routes (total, in 1000s)*I(Curvy zip code)	.067
	(.210)
Year FE	Yes
Zip code FE	Yes
R^2	.99
Obs.	3,305
Mean of dep. var.	1,183

Notes: Dependent variable is the average insurance premium offered across six national insurance providers for the same hypothetical new insuree. I(Curvy zip code) refers to an indicator if the zip-code average of segment curvature is in the upper 75th percentile. Robust standard errors are clustered by zip code. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Appendix References

Haynes, R., A. Jones, V. Kennedy, I. Harvey, and T. Jewell (2007). District variations in road curvature in england and wales and their association with road-traffic crashes. *Environment and planning A 39*(5), 1222–1237.

Haynes, R., I. R. Lake, S. Kingham, C. E. Sabel, J. Pearce, and R. Barnett (2008). The influence of road curvature on fatal crashes in new zealand. *Accident Analysis & Prevention 40*(3), 843–850.

Milton, J. and F. Mannering (1998). The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. *Transportation 25*(4), 395–413.

Wang, C., M. Quddus, and S. Ison (2009). The effects of area-wide road speed and curvature on traffic casualties in england. *Journal of transport geography 17*(5), 385–395.