RFF REPORT

The Accident Externality from Trucking

Lucija Muehlenbachs, Stefan Staubli, and Ziyan Chu

SEPTEMBER 2017; REV. JANUARY 2021

RESOURCES FOR THE FUTURE

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

January 5, 2021

Forthcoming, *Regional Science and Urban Economics*

Abstract

A heavy truck imposes an externality if its presence causes traffic accidents for which it is not held liable. We obtain an estimate of the increase in accidents that occur when a truck is added to a road, using quasi-experimental variation in the influx of truck traffic during the shale gas boom in Pennsylvania. We find evidence that adding trucks to a road is relatively safe for the trucks themselves, but less so for other cars on the road. We find an increase in the number of car accidents when there is an additional truck on the road. While we find an increase in car-accident counts, the additional accidents on city streets and rural roads are not more severe, and on highways even are less severe. We find suggestive evidence that the accident externality of trucking reverberates to even more road users through higher car insurance premiums.

Keywords: externality, trucking, hydraulic fracturing, insurance premiums

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 Carleton University, 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

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] Does the presence of trucks result in more accidents between other vehicles? And if so, is everyone paying through higher car insurance premiums?

The aim of this paper is twofold. *First*, we estimate the accident externality caused by additional trucks on the road. 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. *Second*, we quantify the monetary cost of the accident externality for other road users. We exploit a novel data set that records the insurance premiums offered to new enrollees in Pennsylvania to estimate the causal effect trucks have on insurance premiums.

Vickrey (1968) first pointed out that by merely being on the road, a vehicle imposes an accident externality. Vickrey gives the example that a two-car accident would not have occurred had the non-negligent party opted to take a train, so even the *non-negligent* party's marginal damages are equal to the total damages. Yet liability rules don't hold both vehicles responsible for the full damages, and therefore driving involves unpriced externalities. Examples are easy to envision in the case of trucks: suppose to pass a heavy truck, a car enters oncoming traffic and collides with another car. The accident would have occurred because the truck was on the road, but liability rules would never hold the truck accountable for the damages. Moreover, in the case of trucks, the accident externalities could be even more pronounced than in the case of cars, for two reasons. First, for those accidents that involve a truck, the truck's weight, larger frame, height, wheelbase, braking distance, and rigidity will increase the amount of damage inflicted on the other vehicle.[2] The additional damages from features of trucks would not be internalized in cases in which the truck is not at fault. Second, the simple existence of a long and tall object on the road can require

[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.

[2] Nehiba (2020) exploits a new dataset of truck weights and estimates a strong impact of weight on the severity of accidents.

more difficult maneuvering of other drivers. Trucks as objects on the road can cause other types of accidents to occur, for example, car-on-car collisions. On the other hand, it is possible that drivers will 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.

Little work has been done on estimating the accident externalities of trucking (Nehiba, 2020). 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), and has exploited 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 on car-on-car collisions. Our difference-in-differences estimation strategy relies on the short-lived, large quantities of water used and produced when hydraulically fracturing shale formations for natural gas. Water is pumped into wells 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. 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 first challenge to our identification strategy is that a resource boom

[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).

is not only associated with trucks, but also other confounding factors. In the case of shale gas development, not only is the number of cars on the road changing, but also is the composition of drivers—specifically, there are more young male drivers (Wilson, 2020), a group that is statistically more likely to be in a crash (Massie, Campbell and Williams, 1995).[4] We provide an estimation strategy that isolates the impact of the trucks themselves by comparing, over time, the specific roads used by shale gas trucks to similar roads in the vicinity not used by shale gas trucks.[5]

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, 2020).[6] But we note that these estimates are not 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. Our estimation strategy centers on individual roads: we compare changes on roads predicted to be used by trucks to changes on similar roads, in the same county, that are predicted not to be used by trucks, such that we obtain an estimate that is isolated from the shale boom at the county level. On our predicted truck routes, we find a statistically significant increase in truck traffic that falls within the range of previous reports' predictions of shale gas truck traffic (we estimate each well requires 641-988 one-way truck trips). In contrast, we estimate the number of cars on the truck routes is the same, suggesting that we can isolate the impact of a truck from the general increase in traffic associated with a shale boom. Importantly, we examine local roads separately from highways (with local roads being city streets or rural roads with a single lane of traffic in either direction).

Depending on the subsample we find zero to small effects on the number of accidents involving a truck. This finding is consistent with Nehiba (2020), who using data on truck counts across the U.S. also finds trucks on the road cause only a small increase in the occurrence of accidents involving trucks. Similarly, in our paper, adding a truck to a highway does not result in a statistically significant increase in truck accidents. Adding a truck to a local road, has a small increase of

[4] The in-migration of workers is smaller in Pennsylvania than other fracking states—Wilson (2020) finds the migration response in the northeastern US was almost eight times smaller than in North Dakota.

[5] In the Appendix we also provide estimates that capture the overall impact of a well, including the impact from the resource boom.

[6] 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). 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.

one truck accident occurring annually for every 93 trucks. More striking are our findings on the occurrence of car accidents when trucks are added to the road. We estimate that one car accident occurs annually for every 6 trucks on highways and one car accident occurs annually for every 7 trucks on local roads. The increased risk to other vehicles is a finding previously unreported in the literature. We don't find evidence that the results are driven by differences in the characteristics of drivers and cars on the truck routes. What appears to change is driving behavior. Specifically, the share of accidents attributed to aggressive driving, speeding, and changing lanes is significantly higher on truck routes, including those routes far from wells.

The risk of a truck will depend on truck characteristics, 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, we find differential effects of a truck on highways compared to smaller roads, which suggests that the impact of trucks in more urban settings may be different from what is estimated in this paper. Furthermore, oil field exemptions from limits on driving duration might result in the truck drivers in our sample to be more fatigued than in other industries.[7] As one might expect, fatigue has been identified as a contributing factor of vehicle accidents (Williamson et al., 2011; Smith, 2016), and therefore our estimates could be larger than in the case of heavy trucks. Moreover, our estimates may not extend to other shale regions, because of differences in road characteristics, traffic volumes, and alternative routes across shale states (of which there are over 30), even though they experience similar compositional shifts in truck traffic. Xu and Xu (2020) find that hydraulic fracturing in the Bakken formation in North Dakota lead to more truck traffic and truck accidents, but not more non-truck accidents. As they point out, North Dakota's population density, and consequently traffic density, is much lower than Pennsylvania's. With lower traffic density, the probability of a car-on-car crash triggered by a truck, is much lower. Nonetheless, the qualitative finding, that trucks can increase the number of accidents between other road users, is important to consider when designing public policy involving trucking (e.g.,Leard et al., 2015, Cohen and Roth, 2017, or Lichtman-Sadot, 2019).

[7] The Federal Motor Carrier Safety Administration's regulations for the hours of service of drivers (§ 395.1) has exceptions for drivers servicing the operations of natural gas and oil industry. After a consecutive eight days a driver must take a 34 hour break; for oil and gas, the break is 24 hours. Waiting time at well sites is also not counted as on-duty time, which would in other industries count towards the maximum allowed driving period of 14 consecutive hours (US Federal Motor Carrier Safety Administration, n.d.).

We obtain a dollar value of the externality using a unique data set of the car insurance premiums offered by six national carriers to a representative new insurance enrollee.[8] We expect that trucks on the road will have an outsized impact on car insurance premiums for two reasons. First, as pointed out by Vickrey (1968), the additional costs caused by additional parties on the road are not accounted for in the liability regime. With more expensive truck accidents, the external costs have the potential to be larger, materializing into even larger increases in car insurance premiums. Second, even in cases in which the truck is deemed negligent, the accident can also impose an externality if the truck does not carry sufficient insurance, raising the costs of those who are fully insured (Smith and Wright, 1992). It is arguable that trucks on US roads are underinsured because current required liability was set over thirty years ago, in 1985 at $750,000.[9] While we find more accidents, they are not more severe, and in the case of highways are even less severe. So it is an empirical question as to whether more trucks on the road leads to higher car insurance premiums. In examining insurance premiums, we find that areas exposed to shale gas truck traffic see insurance premiums 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. Importantly, 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 new enrollee insurance premiums by 8 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. We don't have the count of new enrollees in a given year, but if we assume that the number of vehicles registered in Pennsylvania in a year are re-enrolled, then our estimate would imply one truck imposes an external cost of $8,520 in higher aggregate insurance premiums.

[8] The data represent the rates offered to 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.

[9] 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). 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). Discussions about raising the limit bring objections from small businesses. 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/.

Edlin and Karaca-Mandic (2006) estimate the external cost of one car on existing car insurance premia to be between $1,725-$3,239. Our estimate is larger not only because we are looking at trucks and not cars, but we are also examining new insurance plans, which might be more elastic than existing plans.

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, Leard et al., 2015, He, Gouveia and Salvo, 2019, and Cohen and Roth, 2017). We would be remiss to only discuss the costs without discussing the economic benefits of trucking. 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. Section 4 reports our empirical findings on traffic and accidents and 5 reports our empirical findings on the insurance premiums. Section 6 concludes.

2 Background and data

Truck traffic induced by shale gas development is 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).[10] 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 the shale boom, accidents increase. Figure 1 shows the correlation between traffic 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).

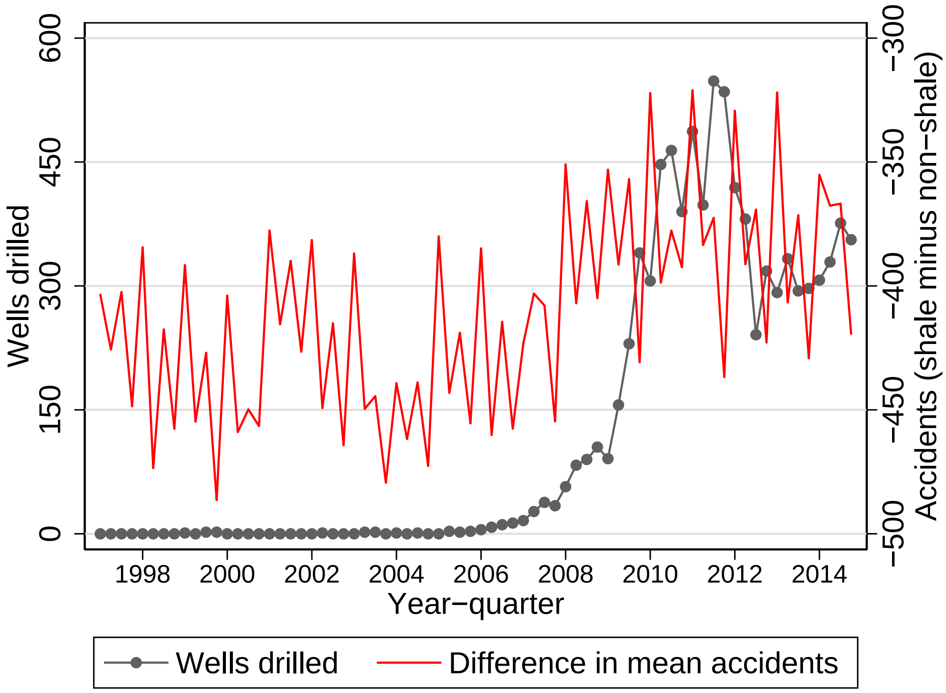


Figure 1: Trends in accidents and wells drilled

Notes: Figure plots the number of wells drilled in a quarter along with the difference between the mean number of accidents in Pennsylvanian counties with shale and in those without.

This increase in accidents represents the risk from drilling a shale gas well in a county; it could be driven by additional cars and 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

[10] These estimates are from New York State Department of Environmental Conservation (2011); Abramzon et al. (2014); Gilmore, Hupp and Glathar (2014). 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*, Russell 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.

zoom in to the road level. Using GIS (described in section 2.2), 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 relative to similar roads not used by trucks, while controlling for the general increase in traffic on all roads in the county.

2.1 Description of data sources

For our analysis, we combine data from several sources and construct a sample 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 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.[11] 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.

[11] 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.

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.[12]

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. 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]

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. For the road network, we use the TIGER dataset from the US Census Bureau. The TIGER dataset breaks up the US road network into 630 thousand road segments that have an average length of .628km. The TIGER dataset designates segments by different road types (primary roads, secondary roads, and

[12] 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).

[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.

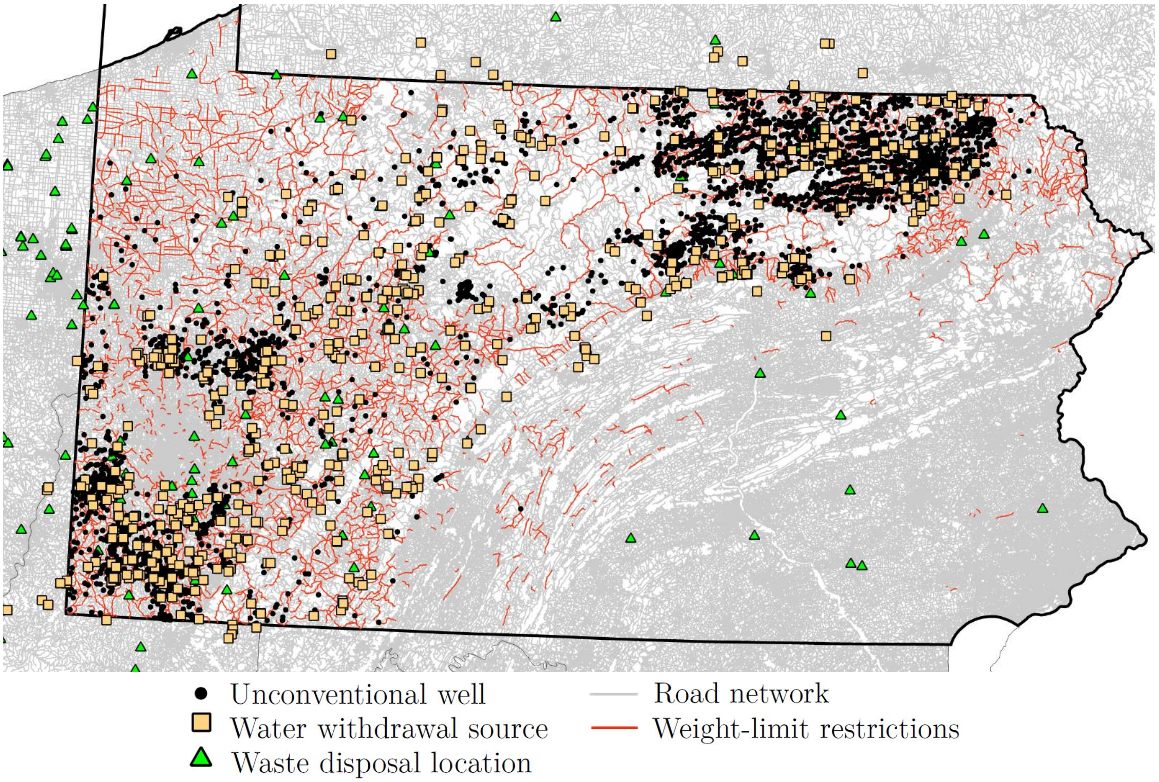


Figure 2: Wells, waste disposal, water withdrawal, and weight-limit restrictions in Pennsylvania, 2014

local neighborhood roads).[14] We calculate the "least cost" route, in which trucks would take the shortest distance, and we add penalties to roads with lower speed limits. We assigned impedances on each road depending on the speed limit of the road type.[15] 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.[16]

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 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 tons. Vehicles weighing more than the posted limit can

[14]The TIGER definitions for road types are the following. Primary roads are generally divided limited-access highways with interchanges and ramps; secondary roads are main arteries, with one or more lanes of traffic in each direction, with at-grade intersections with other roads and driveways; local neighborhood roads are paved non-arterial street, road, or byways with usually a single lane of traffic in each direction.

[15]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.

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

drive on the roads if they obtain a permit, by providing a security bond that can be used to repair the roads.[17] 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.

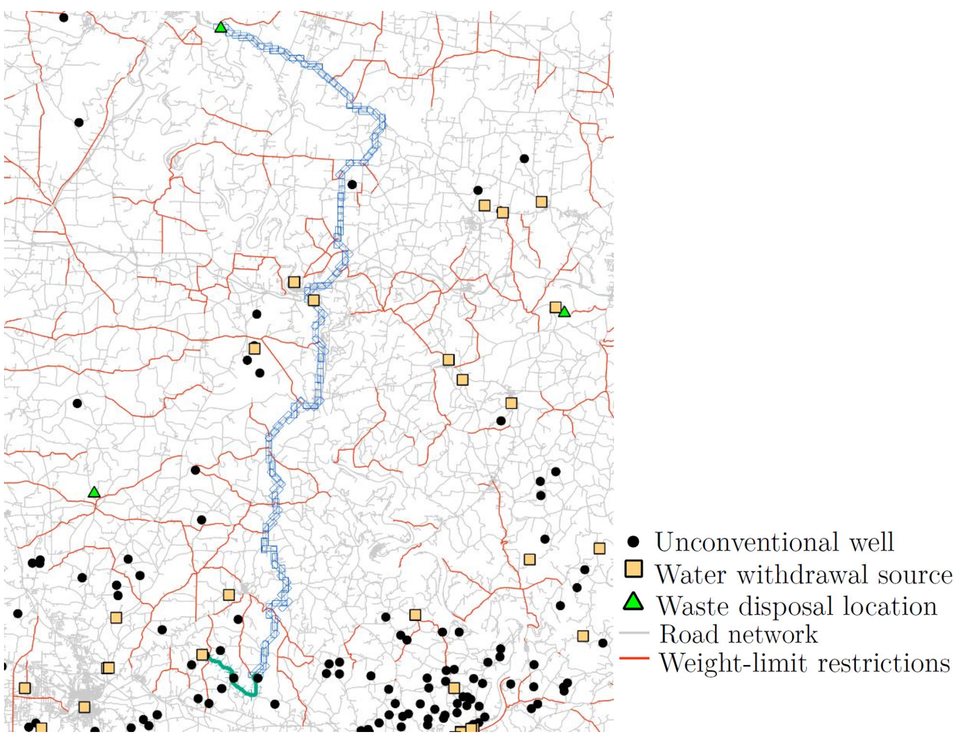


Figure 3: Example route from water withdrawal location to well to waste disposal site

[17]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).

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.[18] 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 and semi-annually from 2010 to 2014.[19] 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.[20]

2.3 Estimation sample

A criticism of empirical studies on accident risk is that road types are often not distinguished in the analysis (Dickerson, Peirson and Vickerman, 2000): "it seems inappropriate to assume that drivers' behaviour, and thus accident rates, do not differ between, say, motorways and small side streets." And that implications exist from mileage on highways being associated with less risk than mileage on undivided streets (Janke, 1991). Recognizing that the impact of a truck will most likely be different depending on the type of road, we divide the sample into two groups of roads: (1) main arterials and highways (these include both primary roads which are usually divided, limited access

[18] 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.

[19] To divide the annual into sub-year intervals, 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.

[20] Eighty percent of the waste shipments are of wastewater, reported in barrels. Twenty percent of the waste shipments are solids reported in tons (comprised of drill cuttings, flowback sand, frac-fluid waste, and general oil and gas waste). We convert tons to barrels using the number of barrels of crude oil in a ton (7.3 Bbl/ton, Team, 2003). This conversion will likely over estimate the number of trucks needed for much of the solid waste, because crude oil is lighter than sludge (5.7Bbl/ton Speight and Arjoon, 2012), or solid shale rock itself ($\tilde{2}$.2Bbl/ton, Manger, 1963).

roads distinguished by the presence of interchanges; and secondary roads, which are main arterials that have one or more lanes of traffic in each direction, may or may not be divided and have grade intersections) and (2) local-neighborhood and rural roads (these are tertiary roads, which are paved, non-arterial streets, including city streets and rural roads, usually a single lane in each direction). Table 1 compares the average characteristics across road types and treatment (truck routes) and control roads, before and after the shale boom (pre and post 2007). We don't have traffic data for all roads, but of those measured, highways have more truck and car traffic than local roads. We do have accident data on all roads, and we see that highway segments have more truck and car accidents. Truck routes are roads that at some point in time are a predicted truck route for a well. Compared to all other roads in the state, these roads have more accidents of any type, both in the pre- and post- period, but even more so in the post- period.

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 a road being a truck route and then exclude roads with a predicted probability in the bottom 25 percent.[21] The last two columns of Table 1 show the mean of the trimmed control group, pre- and post-shale boom.

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 A2), results are similar if we include the full sample or if we are even more conservative and restrict the sample to only roads that at one point and time are truck routes (such that the control group consists of roads that are used by wells at some time in the past or future).[22]

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

[22]Specifically, using the full sample or the most conservative sample, the coefficient point estimates are not statistically different, but we lose some statistical significance when using the most restrictive sample, such that our estimates are only significant at the 10% level.

Table 1: Summary Statistics

	Truck routes		Control roads		Trimmed control roads	
	Mean Pre	Mean Post	Mean Pre	Mean Post	Mean Pre	Mean Post
A. Sample of highways:						
Quarterly truck accidents	.083	.084	.035	.034	.033	.031
Quarterly car accidents	.772	.835	.545	.598	.488	.521
Quarterly accidents with a fatality	.013	.013	.006	.006	.006	.006
Quarterly accidents with an injury	.452	.431	.325	.338	.285	.286
No. of observations (accident data)	302,749	188,410	645,202	380,541	605,530	362,894
No. unique segments (accident data)	7,606	7,606	16,196	15,642	15,168	14,818
Annual daily average truck count	623	832	592	811	582	771
Annual daily average car count	6,195	7,539	6,739	9,146	6,614	8,503
No. of observations (count data)	34,511	51,135	18,601	23,623	17,533	22,234
No. unique segments (count data)	6,732	6,969	9,010	7,523	8,477	7,186
B. Sample of local and rural roads:						
Quarterly truck accidents	.0006	.0009	.0003	.0004	.0003	.0004
Quarterly car accidents	.020	.031	.011	.021	.010	.017
Quarterly accidents with a fatality	.0003	.0004	.0001	.0002	.0001	.0002
Quarterly accidents with an injury	.011	.015	.006	.011	.006	.008
No. of observations (accident data)	1,733,060	967,589	17,857,009	10,445,604	12,777,817	7,206,040
No. unique segments (accident data)	43,475	43,471	447,554	434,741	320,013	310,899
Annual daily average truck count	461	622	320	389	264	319
Annual daily average car count	5,109	6,055	4,643	5,500	3,632	4,163
No. of observations (count data)	34,507	45,381	105,479	104,492	81,411	74,989
No. unique segments (count data)	18,637	18,774	79,489	58,231	60,368	42,360

Notes: All data are by year-quarter (1997-2014) except traffic counts, which are annual (2004-2014). Here "Truck routes" are roads that were at some point in time used by at least one well to access a water withdrawal or disposal site; "Control roads" are all other roads never traversed; "Control roads in trimmed sample" drops control roads whose prediction of traversal, based on pre-shale (i.e., pre-2007) characteristics, is in the bottom 25 percentile. "Pre" refers to pre-2007 and "Post" refers to post-2007.

3 Identification strategy

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.

We examine traffic and accident outcomes by individual road segments. 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}:

$$y_{st} = \alpha \text{Truck Routes}_{st} + \beta \text{ Truck Routes} \times \text{I(Near well)}_{st} + \lambda_s + \mu_{ct} + \varepsilon_{st} \tag{1}$$

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. Importantly, also included in the accident regressions are county-by-year-quarter fixed effects, μ_{ct}. These control for changes in traffic accidents affecting all roads in the county-year-quarter, but are not concentrated in the particular quarter on the particular road used by trucks, including changes in weather from quarter to quarter and county-wide boomtown effects, such as an influx of young male drivers. 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 1,000 meters of a truck route.[23]

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 of these truck routes may also be used by workers needing to get to the wells, particularly when there is only one road to access a well.[24] Therefore, the assumption that a comparison of segments used and not used by trucks over time, after controlling for county-year-quarter 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, Truck Routes $\times$ I(Near

[23]Our results are similar if we include all roads as well as if we drop only those within 500m. In the Online Appendix (Table A3) we show results when we don't drop any control roads and when we drop control roads within 500 meters of the truck route.

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

well), 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). 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.

4 Effects on Traffic Counts and Accidents

4.1 Main Results

Table 2 presents estimates of regression (1) using as outcome variables traffic counts and accidents. Acknowledging that there will be different impacts depending on whether the road is a main arterial/highway or a local-neighborhood/rural road, we estimate the impacts of trucks by road types separately.

Changes in traffic counts on truck routes. In our sample of highways, when a segment is used by one well, we see the annual average increase by 1.755 more trucks per day,[25] which would imply each well uses 641 waste and water trucks.[26] We estimate a similarly sized increase in truck

[25]The independent variables in Table 2 are in counts of 10 wells, so dividing the coefficient on highway truck routes (Panel A, Column 1) by 10 gives 1.755.

[26]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.

traffic on local/rural roads. In the sample of local/rural roads, road segments used by one well see 2.708 more trucks per day, which would imply each well uses 988 waste and water trucks, which is in the range of government reports (800-2,400) (New York State Department of Environmental Conservation, 2011).[27]

Key to our identification strategy, on truck routes we do not see an increase in car traffic on either highways or local roads (Column 2). The coefficients imply that one well using a highway is correlated with a statistically insignificant 1.384 more cars per day, and one well using a local road is correlated with a statistically insignificant 3.970 more cars per day. The car increase represents a statistically insignificant .2 to .8 percent increase in car traffic relative to the baseline. This effect is an order of magnitude smaller than the relative increase in truck traffic (a statistically significant 2 to 4 percent), providing evidence that our county-year-quarter controls capture the increase in cars on the road.[28]

Including controls for proximity to a well (Truck routes*I(Near Well)) captures the combined accident effect of trucks and workers driving to the well. If we are capturing most of the shale truck traffic to/from a well, we should expect to see no significant difference in truck traffic on the routes nearest to the wells. If there are workers needing to get to the wells with cars, we should expect to see more cars on the routes nearest to the wells. We don't find more trucks on the routes nearest to the wells or more cars nearest the wells. Our small and statistically insignificant estimates of cars could be (1) because there is no difference on the roads nearest the wells, or (2) because the limited coverage of traffic count data limits the power we have to detect a difference. Thirty percent of the truck routes are near a well (I(Near well)>0), however, in our sample of truck routes with traffic counts, only 16% are near a well. We discuss what these two different interpretations would mean for our accident counts below.

Changes on accident counts. Now we turn to the count of truck (Column 3) and car (Column 4) accidents in a quarter. We find that adding a truck to a highway does not increase the

[27]How do these compare to a regression if we used county-level observations? Comparing the estimates to county-level estimates found in the Appendix (section A), we note that the segment-level estimate of the number of trucks per well (641-988) is less than county-wide estimate (2,798). The larger county-level estimate includes other types of trucks that are not concentrated on the water/waste routes.

[28]If we used county-level observations, instead of segment-level while controlling for counties, we would indeed estimate an increase in car traffic. 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 (Appendix section A).

Table 2: The impact of trucks on traffic counts and accidents, by road type

	Average daily traffic count		Quarterly accident count	
	Truck	Car	Truck	Car
	(1)	(2)	(3)	(4)
A. Highways				
Truck routes	17.55**	13.84	-.0018	.0240**
	(8.90)	(25.72)	(.0029)	(.0120)
Truck routes*I(Near well)	-.05	1.88	.0472**	.1000**
	(1.06)	(4.05)	(.0199)	(.0430)
Segment FE	Yes	Yes	Yes	Yes
County-year-quarter FE	No	No	Yes	Yes
County-year FE	Yes	Yes	No	No
Mean of dep. var.	757	6,240	.08	.79
R^2	.67	.76	.60	.81
N	37,330	37,330	1,436,809	1,436,809
B. Local and rural roads				
Truck routes	27.08***	39.70	.0006**	.0068***
	(9.97)	(31.74)	(.0003)	(.0025)
Truck routes*I(Near well)	2.72	9.62	.0022***	.0083*
	(2.78)	(9.87)	(.0008)	(.0048)
Segment FE	Yes	Yes	Yes	Yes
County-year-quarter FE	No	No	Yes	Yes
County-year FE	Yes	Yes	No	No
Mean of dep. var.	640	5,256	.0007	.02
R^2	.75	.85	.11	.59
N	53,505	53,505	22,321,018	22,321,018

Notes: Observations are by road segment-year in the case of traffic counts and by road segment-quarter in the case of accidents.
Dependent variables are the annual-average daily truck count and daily car (non-truck) count and the segment-year-quarter 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 and county-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.

number of truck accidents but does increase car accidents.[29] Specifically, when a highway is used

[29] On highways truck accidents are unchanged on the truck routes (Panel A, Column3) but accidents involving other vehicles (non-trucks) increase (Panel A, Column 4).

to connect one well, we see an increase in car accidents of .3%.[30] Compared to highways, we find that adding a truck to a local/rural road, is much more dangerous. We find an increase in both more truck and car accidents on local/rural roads. When a local road connects a well, we see a 8.6% increase in truck accidents and a 3.4% increase in car accidents.[31]

Our traffic estimates suggest that we should not see more accidents nearer to the well (Truck routes*I(Near Well)), because we did not estimate more traffic nearer to the well. However, we do find an increase in both car and truck accidents nearest to the well. How to interpret these conflicting results? If our estimate of no increase in traffic counts nearest to wells is correct, then the increase in accidents would suggest that there is something specific about shale gas traffic (e.g., compositional changes of riskier drivers).[32] In this case our truck traffic results come with the caveat that they are not generalizable beyond shale gas truck traffic. However, given that we have many fewer segments in the traffic-count dataset, it is likely that our traffic counts are under-powered, and that we see more accidents near wells, reiterates the importance of controlling for (Truck routes*I(Near Well)) so that we don't include the accidents nearest to the well when trying to isolate a truck effect.

How to interpret the size of the effects? Using the estimate of the increase in the number of trucks (Table 2, Column 1) and the estimate of the increase in truck accidents (Columns 3), we can calculate a per-truck estimate truck accident rate. On highways we estimate that an additional truck does not have a statistically significant impact on the number of truck accidents. On local and rural roads, we estimate an additional truck increases the number of truck accidents. Using the typical kilometers a truck drives in a year, and assuming they were to occur only on local and rural roads, the estimates imply annually one truck accident for every 83 trucks.[33]

How to interpret the size of our truck accident estimates? First, we can ask whether our accident

[30] Truck routes is in counts of 10 wells, so the accident increase in levels is had from dividing the coefficient by 10 (we see an increase of .00240 car accidents from one well). On average segments have .79 accidents per quarter (so the implied effect is an increase of .3%=(.00240/.79)*100).

[31] In levels, a local road used by one well results in .00006 more truck accidents in a quarter and .00068 more car accidents in a quarter. In percentage terms, the percent increase in truck accidents is larger given far fewer baseline-truck accidents (local roads have on average .0007 truck accidents per quarter, compared to .08 on highways).

[32] We explore the composition of the accidents in more detail in Table 4.

[33] The coefficient on truck routes in Column 3 suggests 0.00006 more truck accidents on rural roads after connecting one well. Dividing by the estimate of the absolute increase in trucks on local roads (988 trucks) and the average length of a local-road segment (.537km), we get the per kilometer risk. Scaling up by the number of kilometers a truck typically travels in a year, 109,685 km (US Department of Transporation, 2013), this results in 0.012 truck accidents per year of truck driving on local and rural roads.

rates reflect current truck insurance premiums. If each truck accident was the fault of the truck and reached the current liability limit of $750,000, the actuarially fair insurance premium would be $0/km for highways to $0.085/km for local/rural roads. The average truck insurance premium, including liability and cargo, was $0.057/km (American Transportation Research Institute, 2016), which falls in the range of our estimates. 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.

Second, we could compare how our estimate of accident costs compare to pollution costs from trucks. The estimate for local/rural roads ($0.085/km) is larger in size than the CO_2 emissions costs ($0.05/km).[34] However, both of these estimates are smaller than the health costs associated with a truck's local air pollution. He, Gouveia and Salvo (2019) 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. And then including our estimate of the increase in car accidents, we have suggestive evidence that the presence of a truck results in an increase in the number of accidents between other road users. On local and rural roads, annually one car accident occurs for every 7 trucks and on highways annually one car accident occurs for every 6 trucks.[35] The car accidents would not be internalized through truck insurance premiums, and therefore would be reason to implement a tax. 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

[34] Using the average fuel economy for heavy trucks (specifically, from the Federal Highway Administration, the average miles traveled per gallon of fuel consumed for combination trucks of 5.8 miles/gallon, or 9.3 km/gallon, `https://www.fhwa.dot.gov/policyinformation/statistics/2013/vm1.cfm`), .0113 tons of CO_2 emissions per gallon of diesel (or 10.21 kilograms per gallon according to the EPA's Emission Factors for Greenhouse Gas Inventories), and a social cost of carbon dioxide of $42/ton.

[35] Using the coefficient of accident count, coefficient on truck counts, road- specific segment lengths (.537km in the case of local roads and 2.57km in the case of highways), and the average annual kilometers driven by trucks.

\$.04/km to \$.20/km).[36] We reveal an additional externality, that of more car accidents (Table 2), and therefore the optimal tax will be larger than previously thought.

4.2 Placebo and Robustness Tests

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. In this section, we present results from a placebo test and various robustness tests to provide evidence that our identification strategy likely isolates the effect of trucks.

Trucking route placebo regression. We test the identifying assumption that trends in traffic and 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 prior to shale gas drilling by recoding the observations so that, falsely, the roads are used in earlier years. We run the traffic count regressions using a shift of three years, since the data start in 2005, over a sample from 2005 to 2011. And we run the accident regressions using a shift of eight years, since those data start in 1997, over a sample 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 3).

Varying the distance to wells. If the coefficient on truck routes remains relatively constant, even far from the well itself, then this would provide evidence that our estimate on truck routes is not confounded by changes associated with the well (e.g., cars needing to access the well pad). Figure 4 shows coefficients from separate regressions, differing in the distance defining "near" a recently drilled well. On the roads nearest to the wells, we find a large significant increase in both truck and car accidents, and this effect declines with the distance from a well. Thus, these controls likely capture something that is related to well access but not solely related to trucks. Including these controls, we find our estimates of accidents on truck routes vary little, regardless of

[36] 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.

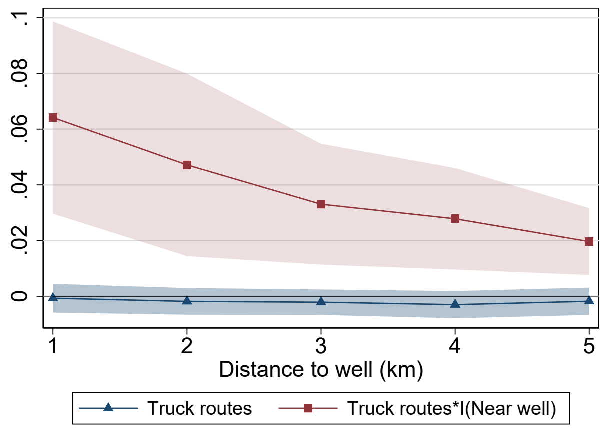
Table 3: Placebo test: Fictitious treatment dates

	Average daily traffic count		Quarterly accident count	
	Truck (1)	Car (2)	Truck (3)	Car (4)
A. Highways				
Truck routes	-13.07	-36.63	-.0015	-.0008
	(9.07)	(30.46)	(.0027)	(.0072)
Truck routes*I(Near well)	-1.82	1.07	-.0082	-.0119
	(1.99)	(9.36)	(.0103)	(.0194)
Mean of dep. var.	766	6,445	.09	.80
R^2	.73	.82	.61	.82
Obs.	23,635	23,635	799,234	799,234
B. Local and rural roads				
Truck routes	-1.15	35.74	-.0001	-.0033**
	(14.87)	(53.03)	(.0001)	(.0015)
Truck routes*I(Near well)	-2.96	17.1354	.0003	-.0016
	(5.74)	(22.37)	(.0006)	(.0033)
Mean of dep. var.	708.98	5681.44	.00	.02
R^2	.81	.91	.11	.59
Obs.	62,520	62,520	12,145,816	12,145,816

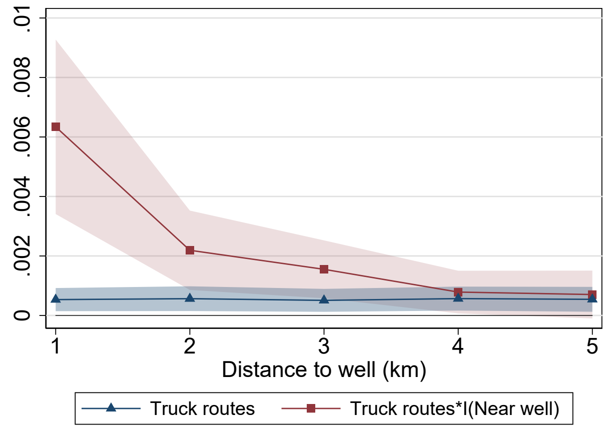
Notes: Regression specifications are the same as those for Table 2, difference is that treatment variables are given fictitious dates. Specifically, treatment variables are shifted 3 years prior in the traffic sample, since the traffic data start in 2004 (traffic sample therefore covers 2004-2011). Treatement variables are shifted by 8 years prior in the accident data, since the accident data start in 1997 (accident sample therefore covers 1997-2006). Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

the distance used to control. The most variation comes from car accidents on highways, but they are nonetheless pretty consistent (with coefficients fluttering between being statistically significance and ranging from .0168 to .0271 in size).

A. Truck accidents

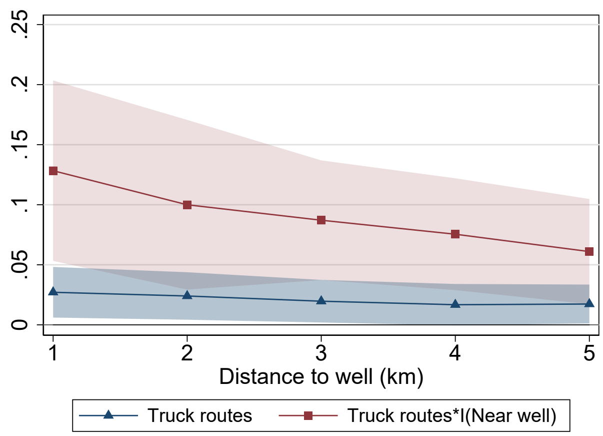


(a) Highways

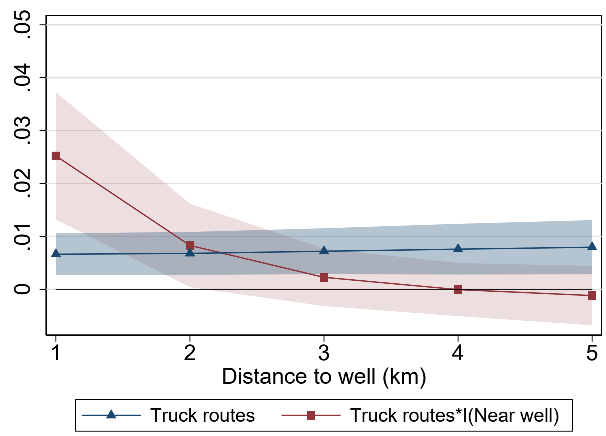


(b) Local and rural roads

B. Car accidents



(c) Highways



(d) Local and rural roads

Figure 4: Truck and car accidents depending on distance from a well

Notes: Figures present the coefficients on Truck routes, α, and on Truck routes*I(Near well), β, from separate regressions of equation (1) on truck and car accidents. 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 (c) use the subsample of highways and subfigures (b) and (d) the subsample of local and rural roads. Shaded areas represent 90% confidence intervals.

Testing for compositional changes on truck routes. To use the variation in shale gas truck traffic to estimate the impact of adding a truck to a road, 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 run equation (1), but use as outcome variables the fraction of accidents occurring in

Table 4: Share of types of accidents in a segment-quarter, by road type

A. Driver characteristics	Highways				Local or rural roads			
	Shale Lic.	Male<25	Alcohol	Unbelted	Shale Lic.	Male<25	Alcohol	Unbelted
Truck routes	-.0001	-.0023	.0001	-.0011	-.0000	.0002	-.0100*	-.0067
	(.0007)	(.0027)	(.0017)	(.0022)	(.0011)	(.0066)	(.0053)	(.0056)
Truck routes*I(Near well)	.0007	.0079**	.0012	.0040	.0019	.0059	.0147	.0302*
	(.0014)	(.0037)	(.0026)	(.0031)	(.0037)	(.0151)	(.0126)	(.0174)
Mean of dep. var.	.0046	.23	.1	.17	.0022	.27	.14	.2
R^2	.08	.08	.10	.10	.16	.17	.20	.19
Obs.	273,485	273,485	273,485	273,485	208,334	208,334	208,334	208,334
B. Accident characteristics								
	Aggressive	Speeding	Changing	Tailgating	Aggressive	Speeding	Changing	Tailgating
Truck routes	-.0044	-.0054**	.0025*	-.0022	.0278***	.0193***	.0022*	-.0014
	(.0033)	(.0026)	(.0015)	(.0014)	(.0078)	(.0071)	(.0012)	(.0019)
Truck routes*I(Near well)	.0004	.0007	-.0016	.0034**	-.0063	-.0225	-.0030	.0087*
	(.0046)	(.0045)	(.0015)	(.0016)	(.0175)	(.0193)	(.0024)	(.0046)
Mean of dep. var.	.58	.25	.042	.061	.56	.29	.0065	.03
R^2	.15	.22	.18	.15	.23	.31	.16	.18
Obs.	273,485	273,485	273,485	273,485	208,334	208,334	208,334	208,334

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

a segment-quarter that are attributed to accidents with particular characteristics.[37]

Truck routes are similar in characteristics associated with a shale boom (e.g., males younger than 25, cars registered in the four-largest-shale-producing states, fatigued drivers, or share of luxury vehicles). Exceptions are in directions one might not expect, and are small. Specifically, we find fewer alcohol related accidents (but only a .7% decrease that is statistically significant at the 10% level, Table 4), and fewer males aged 25 to 50 (but only a .2% decrease, Table A4). We

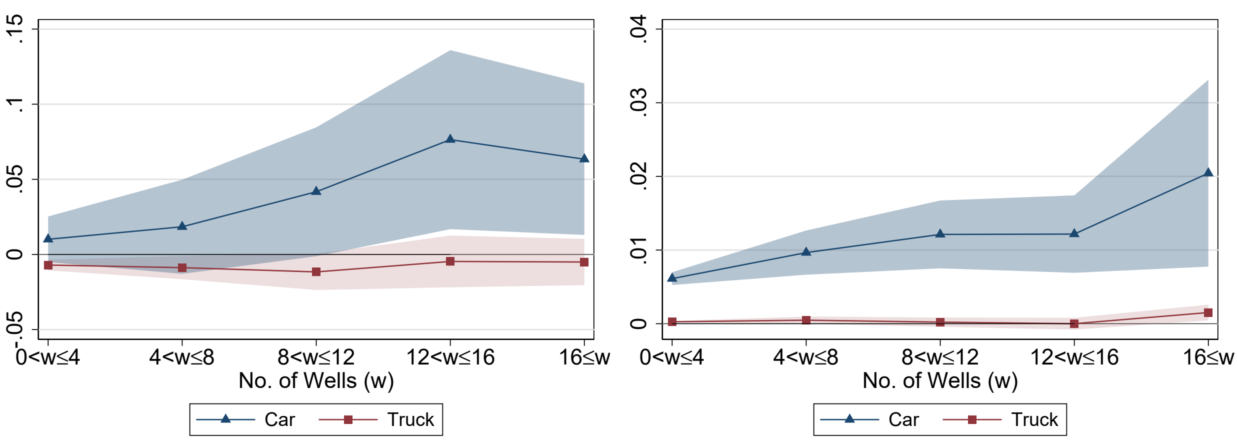
[37] Driver characteristics include the share of accidents with a driver with a license from the largest shale producing states (Arkansas, Louisiana, Oklahoma, or Texas). Including Pennsylvania, 95% of production came from these states according to the EIA's production estimates for 2010. We also examine 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.

do see some statistically significant coefficients in the case of characteristics of driving behavior. We find highways used by trucks result in less speeding (.2% less), 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 the share of accidents attributed to aggressive driving (.5%) and speeding (.7%). 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 (3.3% on local roads and .6% on highways). We do not find an impact on other risky-driving indicators (such as tailgating, not using a seatbelt, or in the Appendix, distracted-driving). On highways, vehicles are newer, but only by less than a month (.1% newer). Nearest to the well, we do detect an increase in young men and tailgaters, reiterating the importance of controlling for the roads nearest to the wells.

Importantly, we note that when we examine characteristics at the county-level (Appendix Section B.3), we find boom counties indeed see different driver characteristics, which further demonstrates the power of the county-year-quarter controls in the road-segment specification. Running similar regressions as Table 4 but 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 (Table A5). 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).

Is the relationship between trucks and accidents nonlinear? The accident risk of a truck could depend on how many other trucks are on the road. For example, it could be that the incremental accident risk from a truck declines with each additional truck if traffic speed is reduced and people drive more carefully in the presence of many tucks. However, it could also be that the incremental accident risk increases with each truck if congestion results in more people attempting to pass the trucks. We examine whether the relationship between trucks and accidents is nonlinear by splitting treated segments into five groups, roughly evenly spaced groups: segments connecting up to four wells, segments connecting four to eight wells, segments connecting eight to 12 wells, segments connecting 12 to 16 wells, and segments connecting more than 16 wells (with segments connecting no wells as the reference category).[38] We then estimate equation (1) where we replace Truck Routes by separate indicators for each group.

Figure 5 plots the estimated coefficients on the indictors, for highways and local and rural roads separately. On highways we find evidence for a nonlinear effect of trucks on car accidents: adding an additional truck only causes more car accidents if many other trucks are already present. In contrast, on rural roads adding an additional truck increases the accident risk for cars independent of the number of trucks present. The car accident risk appears to increase with the number of trucks, but we cannot rule out that the point estimates for the different groups are statistically the same. Consistent with the main estimates, we find additional trucks have little effect on truck accidents on highways, but do increase truck accidents on local and rural roads (although the estimated effects are small in magnitude).



(a) Highways

(b) Local and rural roads

Figure 5: Truck and car accidents by intervals of well counts

Notes: Each figure presents the coefficients from a regression that includes separate dummies for an increasing number counts that a road segment is used as a well's truck route.

[38]Taking an increment of four wells results in a similar number of treated segments in each group (except for the first group, which contains more treated segments because most segments only connect one or two wells.)

Table 5: The impact of trucks on injuries and fatalities

	Minor	Moderate	Major	Any injury	Fatality
A. Highways					
Truck routes	-.0255***	-.0075***	-.0027***	-.0151**	-.0001
	(.0052)	(.0019)	(.0010)	(.0062)	(.0006)
Truck routes*I(Near well)	-.0047	-.0162**	.0016	.0124	.0005
	(.0112)	(.0072)	(.0053)	(.0171)	(.0025)
Mean dep. var.	.31	.1	.029	.44	.013
R^2	.66	.43	.19	.75	.11
Obs.	1,436,809	1,436,809	1,436,809	1,436,809	1,436,809
B. Local and rural roads					
Truck routes	.0000	-.0001	-.0002	.0021	-.0001
	(.0009)	(.0003)	(.0002)	(.0015)	(.0001)
Truck routes*I(Near well)	.0021	.0010	-.0006	.0022	.0005
	(.0024)	(.0011)	(.0004)	(.0030)	(.0004)
Mean dep. var.	.0083	.003	.00088	.013	.00033
R^2	.32	.14	.05	.41	.03
Obs.	22,321,018	22,321,018	22,321,018	22,321,018	22,321,018

Notes: Observations are by road segment-quarter. 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 and county-year-quarter. Robust standard errors are clustered by road-segment. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

Impact of trucks on accident severity. On local roads, because we find more accidents involving a heavy truck, we might expect to see more severe accidents. However, we don't see more injuries or fatalities on local roads (Table 5). In the case of adding a truck to a highway, we even estimate a decrease in the number of injuries. This could be because of the reduced speed on highways with trucks (Table 4). Even though average severity went down (highways) or remained the same (local roads) the count of accidents increased, and accidents have cost. To get a measure of the costs, we turn to data on insurance premiums (Section 5).

5 Valuation of the accident externality of trucking: evidence from car insurance premiums

The previous sections provide evidence that adding one truck to a road increases the number of car accidents not involving a truck, which would be classified as an external cost of trucking. Another source of externalities is suggested by the estimated increase in truck accidents on local and rural roads, especially if the trucks are underinsured. Here we look for evidence whether these accident externalities propagate into insurance premiums by looking at the data on the change in premiums offered to a representative new enrollee of car insurance over time.

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 premiums 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.[39] Using these data, we make the assumption that changes to this representative driver's insurance premiums 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.[40]

Table 6: Summary statistics at the zip-code level

	Traversed zip codes		Nontraversed zip codes	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Average premium ($)	1076.3	(87.2)	1481.9	(472.2)
Δ in premium between 2012 to 2014 ($)	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 truck route (withdrawal or disposal connection) over the sample period.

[39] 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.

[40] 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% of accidents occur within 24km of one's residence. https://www.progressive.com/newsroom/article/2002/may/fivemiles/

Across both years of data, the average insurance premium in Pennsylvania is lower in zip codes that are traversed by trucks (a \$1,076 annual average premium compared with \$1,482). But the 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 (each time a segment connects a well in the quarter, the count increases). 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, in aggregate, the increase is not necessarily trivial; the most heavily traversed zip code is near 76 thousand segment/well connections, which would translate 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 with registered vehicles saw the predicted increase on their insurance premiums, then this would aggregate to an externality of \$14 million dollars in 2014 from shale gas development.[41]

We can also use the coefficient estimate to calculate a per-truck estimate. We estimate that annual insurance premiums of new enrollees increases by \$0.08 per truck.[42] 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

[41] To get a Pennsylvania-aggregate estimate, we first obtained the count of vehicles registered in each county in Pennsylvania in 2014 (Pennsylvania Department of Transportation, 2014). We multiplied each county's number of vehicles registered by the size of the county's treatment (county's registered vehicles*0.000401*county's total truck routes), and then summed over all counties in Pennsylvania (to get the total of \$14 million). Note that this is an upper bound because not all registered vehicles are new insurance enrollees, and so not all would see an increase in insurance. We also use the zip-code-level estimate for a truck route and assume it is the same as a county estimate.

[42] To get the per-truck-year increase in premiums, we first divide the coefficients by the number of trucks per segment (814, or the average between the highway and local road estimate) and the number of kilometers in a segment (0.665km per segment). Then we multiply by the annual average of kilometers traveled (109,685 km). The estimate implies a 8 cents increase ($0.000401/814/0.665 \times 109,685$).

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 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 vehicle registration counts again we can get an upper bound on this estimate. If we assume all vehicles registered in Pennsylvania see an insurance increase, then one truck would cause an externality of \$8,520 in aggregate.[43]

Table 7: The impact of trucks on car insurance premiums

	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
Mean of dep. var.	1,183	.0397	.125	9.52	81.2
R^2	.99	.86	.99	.93	1.00
Obs.	3,305	2,965	2,965	2,965	134

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. The far reach of truck traffic results in an externality that is atypical of findings in the literature on the burden of shale gas development. Typically, the negative externalities surrounding shale gas production have been felt by those living nearby (e.g., from risks to drinking water, Muehlenbachs,

[43]To get the increase per truck, we divide the Pennsylvania-wide cost, of \$14 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.

Spiller and Timmins, 2015, Hill and Ma, 2017, air quality, Zhang et al., 2020, or local earthquakes Cheung, Wetherell and Whitaker, 2018, Burnett, Mothorpe and Jaume, 2018) and the positive externaltites have been felt by those farther away (e.g., from lower greenhouse gas emissions or electricity prices from the shift from coal to gas-fired electricity, Linn and Muehlenbachs, 2018, Brehm, 2019, Coglianese, Gerarden and Stock, 2020). Shale trucks transporting water and waste far from the well is an instance of a negative externality extending over a larger area. Pennsylvania collects an impact fee from wells (Black, McCoy and Weber, 2018), 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 trucks. Smith and Wright (1992) describe how the presence of uninsured drivers on the road increases car insurance premiums. 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.[44] 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, 2017). 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.[45]

[44] 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.

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

6 Conclusion

We estimate the life of a shale gas well requires 641 to 988 one way truck trips for hauling of water to and from the well. The shale boom in Pennsylvania experienced a large increase in truck traffic, when at the height of drilling in 2011, close to 2000 wells drilled were drilled. The truck traffic associated shale drilling is short lived, unlike other more permanent increases, for example, from the opening of a warehouse, which is estimated to be on the order of 500 trucks per day (Bluffstone and Ouderkirk, 2007).

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 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 advances in GIS navigation, keeping track of the kilometers driven by road-type would not be arduous, and even a kilometer-by-road-type tax could be 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. Combining our findings with previous findings that welfare would be improved with weight-based taxes (He, 2016; Cohen and Roth, 2017; Nehiba, 2020), the first-best solution would be a tonnage-by-kilometer-by-road-type tax.

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.

Allcott, Hunt, and Daniel Keniston. 2018. "Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America." *The Review of Economic Studies*, 85(2): 695–731.

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

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

Black, Katie Jo, Shawn J McCoy, and Jeremy G Weber. 2018. "When externalities are taxed: The effects and incidence of PennsylvaniaâĂŹs impact fee on shale gas wells." *Journal of the Association of Environmental and Resource Economists*, 5(1): 107–153.

Bluffstone, Randall A, and Brad Ouderkirk. 2007. "Warehouses, trucks, and PM2. 5: Human health and logistics industry growth in the eastern Inland Empire." *Contemporary Economic Policy*, 25(1): 79–91.

Brehm, Paul. 2019. "Natural gas prices, electric generation investment, and greenhouse gas emissions." *Resource and Energy Economics*, 58: 101106.

Burnett, J Wesley, Christopher Mothorpe, and S Jaume. 2018. "Earthquake alley: Unconventional oil and gas development, induced seismic activity, and housing price impacts." International Association for Energy Economics.

Cheung, Ron, Daniel Wetherell, and Stephan Whitaker. 2018. "Induced earthquakes and housing markets: Evidence from Oklahoma." *Regional Science and Urban Economics*, 69: 153–166.

Coglianese, John, Todd D Gerarden, and James H Stock. 2020. "The effects of fuel prices, environmental regulations, and other factors on US Coal Production, 2008-2016." *The Energy Journal*, 41(1).

Cohen, Linda R., and Kevin D. Roth. 2017. "A Second-best Dilemma: Freight Trucks, Externalities, and the Dispatch Effect." *Working Paper.*

Dickerson, Andrew, John Peirson, and Roger Vickerman. 2000. "Road accidents and traffic flows: an econometric investigation." *Economica*, 67(265): 101–121.

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.

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. 2014. "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, Jen Z. 2016. "Heterogeneous responses and differentiated taxes: evidence from the heavy-duty trucking industry in the US."

He, Jiaxiu, Nelson Gouveia, and Alberto Salvo. 2019. "External effects of diesel trucks circulating inside the Sao Paulo megacity." *Journal of the European Economic Association*, 17(3): 947–989.

Hill, Elaine, and Lala Ma. 2017. "Shale gas development and drinking water quality." *American Economic Review*, 107(5): 522–25.

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

Jacobsen, Grant D, and Dominic P Parker. 2016. "The economic aftermath of resource booms: evidence from boomtowns in the American West." *The Economic Journal*, 126(593): 1092–1128.

James, Alexander, and Brock Smith. 2017. "There will be blood: Crime rates in shale-rich US counties." *Journal of Environmental Economics and Management*, 84: 125–152.

Janke, Mary K. 1991. "Accidents, mileage, and the exaggeration of risk." *Accident Analysis & Prevention*, 23(2-3): 183–188.

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 G Gordon, and SM Olmstead. 2013. "Pathways to Dialogue: What the experts say about the environmental risks of shale gas development (in Resources for the Future Report."

Leard, Benjamin, Joshua Linn, Virginia McConnell, and William Raich. 2015. "Fuel costs, economic activity, and the rebound effect for heavy-duty trucks." *Resources for the Future Discussion Paper*.

Lichtman-Sadot, Shirlee. 2019. "Can public transportation reduce accidents? Evidence from the introduction of late-night buses in Israeli cities." *Regional Science and Urban Economics*, 74: 99–117.

Linn, Joshua, and Lucija Muehlenbachs. 2018. "The heterogeneous impacts of low natural gas prices on consumers and the environment." *Journal of Environmental Economics and Management*, 89: 1–28.

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.

Manger, G Edward. 1963. "Porosity and bulk density of sedimentary rocks." *Contributions to Geochemistry, Geological Survey Bulletin, 1144-E.*

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

Marchand, Joseph. 2012. "Local labor market impacts of energy boom-bust-boom in Western Canada." *Journal of Urban Economics*, 71(1): 165–174.

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.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The housing market impacts of shale gas development." *American Economic Review*, 105(12): 3633–59.

Nehiba, Cody. 2020. "Taxed to death? Freight truck collision externalities and diesel taxes." *Regional Science and Urban Economics*, 85: 103577.

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.

Pennsylvania Department of Transportation. 2014. "Report of Registrations for Calendar Year 2014." *http://www.dot.state.pa.us/public/dvspubsforms/BMV/Registration%20Reports/ReportofRegistration2014.pdf*, (accessed July 10, 2020).

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, Austin C. 2016. "Spring forward at your own risk: daylight saving time and fatal vehicle crashes." *American Economic Journal: Applied Economics*, 8(2): 65–91.

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

Speight, James G, and Karuna K Arjoon. 2012. *Bioremediation of petroleum and petroleum products.* John Wiley & Sons.

Team, USGS Uinta-Piceance Assessment. 2003. "Petroleum systems and geologic assessment of oil and gas in the Uinta-Piceance Province, Utah and Colorado." US Geological Survey.

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."

US Federal Motor Carrier Safety Administration. n.d.. *FMCSA Regulations, Part 395, Hours of Service of Drivers, §395.1.*

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.

Weinstein, Amanda L, Mark D Partridge, and Alexandra Tsvetkova. 2018. "Follow the money: Aggregate, sectoral and spatial effects of an energy boom on local earnings." *Resources Policy*, 55: 196–209.

Williamson, Ann, David A Lombardi, Simon Folkard, Jane Stutts, Theodore K Courtney, and Jennie L Connor. 2011. "The link between fatigue and safety." *Accident Analysis & Prevention*, 43(2): 498–515.

Wilson, Riley. 2020. "Moving to economic opportunity: the migration response to the fracking boom." *Journal of Human Resources*, 0817–8989R2.

Xu, Minhong, and Yilan Xu. 2020. "Fraccidents: The impact of fracking on road traffic deaths." *Journal of Environmental Economics and Management*, 102303.

Zhang, Ruohao, Huan Li, Neha Khanna, Daniel Sullivan, Alan Krupnick, and Elaine Hill. 2020. "Satellite Detection of Air Pollution: Air Quality Impacts of Shale Gas Development in Pennsylvania." *Working Paper.*

A Boomtown Impacts

In this section, we present our estimates for 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. Indeed, the local income and employment shocks from oil and gas booms are sizeable (see for example, Marchand, 2012, Weber, 2012, Weinstein, Partridge and Tsvetkova, 2018, Jacobsen and Parker, 2016, Maniloff and Mastromonaco, 2017, Feyrer, Mansur and Sacerdote, 2017, and Allcott and Keniston, 2018).

Strategy to identify boomtown impacts. 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:

$$y_{ct} = \alpha Wells_{ct} + \lambda_c + \delta_t + \varepsilon_{ct}, \tag{2}$$

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.[1] 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.[2] 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.

Main results. 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 A1). 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.[3] 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 A1 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

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

[2] Landowners have some leeway on whether wells will be drilled on their property; in Pennsylvania minerals are most often 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] 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.

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.

Table A1: The boomtown impacts on traffic and accidents

	Annual-average daily traffic count		Quarterly accident count	
	Truck	Car	Truck	Car
Wells	.79***	7.54***	.171***	.815***
	(.27)	(2.29)	(.023)	(.228)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	No	No
Year-quarter FE	No	No	Yes	Yes
Mean of dep. var.	440	4,814	17	305
R^2	.39	.52	.93	.98
Obs.	663	663	4,824	4,824

Notes: Observations are by county-year in the case of traffic counts (2004-2014) and by county-quarter in the case of accidents (1997-2014. 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. 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, of which Pennsylvania has 67. *** Statistically significant at the 1% level; ** 5% level; * 10% level.

B Additional Tables

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

B.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 A2 shows results do not change across different control groups.

B.2 Exploring spill-over roads

The increased truck traffic on truck routes might mean nearby roads are also affected, violating SUTVA. In our main specification we deal with this concern by omitting any control roads that are within 1,000 meters of a truck route. Our results are similar with or without dropping roads within 1,000 meters, though Table A3 shows some evidence of attenuation bias when we do not exclude roads within 1,000 meters (e.g., our estimates of truck counts are smaller when we include all control roads).

B.3 Are the accidents on truck routes 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 4). 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-year-quarter 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).

Table A2: Robustness: Regressions using different control groups, by road type

	Average daily traffic count		Quarterly accident count	
	Truck (1)	Car (2)	Truck (3)	Car (4)
A. Highways, full sample				
Truck routes	17.58**	14.06	-.0018	.0245**
	(7.12)	(20.59)	(.0029)	(.0120)
Truck routes*I(Near well)	-.05	1.87	.0472**	.0996**
	(.85)	(3.24)	(.0199)	(.0430)
Mean of dep. var.	774	6,324	.08	.79
R^2	.65	.74	.62	.86
Obs.	33,738	33,738	1,493,100	1,493,100
B. Highways, sample of only ever-traversed roads				
Truck routes	13.29*	6.49	-.0013	.0235*
	(7.39)	(21.33)	(.0030)	(.0127)
Truck routes*I(Near well)	.21	2.58	.0467**	.1007**
	(.86)	(3.32)	(.0200)	(.0430)
Mean dep. var.	775	6,325	.08	.79
R^2	.65	.74	.66	.81
Obs.	15,687	15,687	483,733	483,733
C. Local and rural roads, full sample				
Truck routes	26.95***	39.85*	.0006**	.0075***
	(6.57)	(21.00)	(.0003)	(.0025)
Truck routes*I(Near well)	2.73	9.67	.0022***	.0079
	(1.84)	(6.53)	(.0008)	(.0048)
Mean of dep. var.	705	5,590	.00	.02
R^2	.74	.83	.07	.48
Obs.	86,608	86,608	30,512,233	30,512,233
D. Local and rural roads, sample of only ever-traversed roads				
Truck routes	23.08***	27.85	.0005*	.0043*
	(7.13)	(22.90)	(.0003)	(.0024)
Truck routes*I(Near well)	2.27	7.07	.0020**	.0085*
	(1.80)	(6.56)	(.0008)	(.0047)
Mean dep. var.	705	5,592	.00	.02
R^2	.78	.82	.10	.51
Obs.	16,836	16,836	2,663,233	2,663,233

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 and county-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 A3: Robustness: Regressions including different nearby roads

	Average daily traffic count		Quarterly accident count	
	Truck	Car	Truck	Car
	(1)	(2)	(3)	(4)
A. Highways, include spillover roads				
Truck routes	14.71*	12.35	-.0008	.0226**
	(8.43)	(24.20)	(.0027)	(.0111)
Truck routes*I(Near well)	.15	2.27	.0478**	.1008**
	(1.04)	(3.95)	(.0197)	(.0427)
Mean of dep. var.	754	6,259	.08	.78
R^2	.67	.76	.59	.81
Obs.	41,689	41,689	1,525,858	1,525,858
B. Highways, drop roads within 500m				
Truck routes	17.93**	14.96	-.0016	.0234**
	(8.78)	(25.43)	(.0029)	(.0118)
Truck routes*I(Near well)	-.09	1.87	.0472**	.1004**
	(1.05)	(4.00)	(.0199)	(.0429)
Mean of dep. var.	756	6,248	.08	.79
R^2	.67	.76	.60	.81
N	38,153	38,153	1,451,863	1,451,863
C. Local and rural roads, include spillover roads				
Truck routes	26.11***	46.58	.0006**	.0066***
	(8.96)	(28.93)	(.0002)	(.0024)
Truck routes*I(Near well)	3.28	11.34	.0022***	.0086*
	(2.65)	(9.28)	(.0008)	(.0048)
Mean of dep. var.	625.46	5,226	.00	.02
R^2	.76	.87	.08	.51
N	118,916	11,8916	24,353,696	24,353,696
D. Local and rural roads, drop roads within 500m				
Truck routes	27.67***	44.78	.0006**	.0067***
	(9.62)	(30.84)	(.0003)	(.0024)
Truck routes*I(Near well)	2.96	10.86	.0022***	.0084*
	(2.75)	(9.74)	(.0008)	(.0048)
Mean of dep. var.	637	5259	.00	.02
R^2	.76	.87	.08	.52
N	104,252	104,252	23,000,666	23,000,666

Notes: Table 2 in the main text shows results excluding control roads that are within 1,000 meters of a truck route. Panels A and C show results without excluding any roads near the truck routes. Panels B and D excludes only roads with 500 meters of the shale routes. Robust standard errors are clustered by road-segment. *** 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	-.0016	-.0002	-.0085***	.0001	-.0019	.0011	-.0115	-.0058
	(.0021)	(.0009)	(.0032)	(.0026)	(.0057)	(.0024)	(.0075)	(.0059)
Truck routes*I(Near well)	-.0014	.0010	.0051	.0013	-.0034	-.0018	.0215	-.0077
	(.0028)	(.0014)	(.0042)	(.0034)	(.0083)	(.0060)	(.0207)	(.0106)
Mean of dep. var.	.08	.025	.38	.23	.08	.019	.33	.18
R^2	.11	.09	.10	.09	.19	.17	.17	.18
Obs.	273,485	273,485	273,485	273,485	208,334	208,334	208,334	208,334
B. Accident characteristics								
	Reg. Shale	Luxury	Veh. Age	Avoiding	Reg. Shale	Luxury	Veh. Age	Avoiding
Truck routes	-.0007	.0011	-.0859**	-.0005	-.0023*	.0058	-.1008	.0041*
	(.0009)	(.0019)	(.0374)	(.0007)	(.0012)	(.0044)	(.0894)	(.0023)
Truck routes*I(Near well)	.0001	.0040	-.0103	.0002	.0017	-.0077	.2282	-.0026
	(.0012)	(.0028)	(.0536)	(.0010)	(.0043)	(.0074)	(.2000)	(.0048)
Mean of dep. var.	.0073	.11	7.4	.026	.003	.1	8	.032
R^2	.09	.10	.18	.10	.16	.17	.24	.18
Obs.	273,485	273,485	273,485	273,485	208,334	208,334	208,334	208,334

Notes: This table provides additional outcome variables from those found in Table 4. 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 4 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.