Hoggin' the Road: Pork Slaughterhouses Increase Heavy Truck Traffic, Reducing Road Quality

Ben Blemings[*], Margaret Bock[†], and Alexandre R. Scarcioffolo[‡]

This Version: November 2, 2020

PRELIMINARY DRAFT:

PLEASE DO NOT CITE, SHARE, OR CIRCULATE

Abstract

Although businesses are economically beneficial in some regards, they can also bring negative externalities. One such negative externality that has received substantial attention is heavy truck traffic, which is associated with hydraulic fracturing and agriculture. However, estimates of slaughterhouses on truck traffic and truck traffic on road damage using observational data have not been based on a causal framework. This paper appeals to cost minimizing producers and persistent natural regional advantages to exploit historic hog density as quasi-random variation in current pork slaughterhouse location. Estimates suggest 1 pork slaughterhouse increases truck traffic by 21.9 percent. Furthermore, estimates suggest that an additional pork slaughterhouse increases road roughness by 9-32 percent of a standard deviation. This is the first time the effect of trucks on road roughness has been estimated causally in a nationwide sample, making the estimates useful for the policy discussion around slaughterhouses, truck traffic resulting from heavy industry, and optimal tax estimates in structural urban models.

Keywords: Truck Traffic, Road Damage, Slaughterhouse

JEL Classification: R41, R42, Q10, N52

[*]Benjamin Blemings, Ph.D. Candidate, Department of Economics, West Virginia University, Morgantown WV 26506. E-mail: bbleming@mix.wvu.edu

[†]Margaret Bock, Ph.D. Candidate, Department of Economics, West Virginia University, Morgantown WV 26506. E-mail: mbock1@mix.wvu.edu

[‡]Alexandre R. Scarcioffolo, Assistant Professor, Department of Economics and Finance, Georgia College & State University, Milledgeville, GA 31061. E-mail: alexandre.scarcioffolo@gcsu.edu

1 Introduction

Throughout the 19th century in the US, there was sustained economic growth which was correlated with rising trends in industrialization and urbanization (Kaldor, 1967).[1,2] While industrialization and urbanization were positively correlated with economic growth, they also brought a host of negative externalities to cities, such as water pollution (Leon, 2008, pg. 5), air pollution (Hanlon, 2016), crime (United States President's Commission on Law Enforcement & Administration of Justice & Katzenbach, 1967, pg. 5, 17), and congestion (Glaeser, 2014, pg. 3). This period of history demonstrates negative trade-offs accompanying economic advantages of nearby businesses.

Modern industries, such as coal-powered energy generation, hydraulic fracturing, and agriculture also generate negative externalities.[3] Hydraulic fracturing causes road accidents due to truck traffic (Xu & Xu, 2020).[4] In response, at least one town in each major US shale formation (Bakken, Marcellus, Barnett) has attempted to ban heavy truck traffic from nearby fracking operations.[5]

Similar to fracking, the hog industry is also associated with heavy truck traffic.[6] Consider the case of Great Falls, Montana; the city attempted to prohibit the opening of a 3,000 acre meat processing facility near the city in 2019. According to residents, major reasons to cancel the facility was additional trucks would cause pollution, traffic, and road damage.[7] Residents are correct trucks are used by the hog industry; trucks are the preferred mode of transportation for hogs (Food and

[1]For instance, US GDP per capita grew by 216% from 1800 to 1900 (Bolt, Inklaar, de Jong, & van Zanden, 2018).

[2]Economists attempted to explain these trends by studying firm location decisions (Hotelling, 1929; Marshall, 1890), which has evolved today towards the study of agglomeration (Glaeser & Gottlieb, 2009).

[3]Agriculture is associated with water overuse (Pfeiffer & Lin, 2012), while coal and hydraulic fracturing cause air pollution (Hanlon, 2016; Howarth, Ingraffea, & Engelder, 2011).

[4]Over just one year of average fracking well development, (Abramzon, Samaras, Curtright, Litovitz, & Burger, 2014, pg. 312) estimated there would have been an additional 35,686 heavy truck trips across the state of New York had fracking development continued in the state.

[5]These concerns have been highlighted in popular media. Land owners in North Dakota were concerned about increased truck traffic from fracking wells (Associated Press, 2017); a Texas resident expressed concern about truck traffic from fracking, noting one of her friends was killed by a fracking industry truck (The Texas Tribune, 2014); Pennsylvania residents note the problems associated with the increased truck traffic due to fracking (The Allegheny Front, 2017); A 2014 study notes that not only are there external costs, such as pollution, associated with the truck traffic from fracking, there are also large direct roadway consumption costs imposed on transportation agencies (Abramzon et al., 2014).

[6]Other negative externalities of the agriculture sector include water contamination from fertilizer and pesticides, wildlife habit damage, emission of gases, soil erosion, among others (Pretty et al., 2000). For instance, livestock farms such as pig farms are known for its unbearable stench that have given rise to lawsuits from homeowners in North Carolina (Hellerstein & Fine, 2017).

[7]Source: https://www.nrdc.org/stories/these-montanans-dont-want-industrial-slaughterhouse-their-backyard?fbclid=IwAR0Diss7fr2Rp1R9jNpHVQBDPCLUSh_ck21DbpWTqrONKOjI3k3FMOj9cJk

Agriculture Organization of the United Nations, 2001) and Brodersen (2015) argues more than 2,000 trucks transport a total of 1,000,000 swines daily in the US.[8]

The question is how much truck traffic is attributable to additional pork slaughterhouses? Furthermore, do pork slaughterhouses and resulting truck traffic have negative consequences for road quality? The approach is uniquely positioned to answer these questions, because the data is nationwide and attention is paid to causal inference issues unlike dated experiments (Eschwege, 1979), other observational work (Bai et al., 2007), or accounting exercises (Abramzon et al., 2014).

Not accounting for endogenous slaughterhouse location, the naïve OLS results find no effect of pork slaughterhouses on truck traffic, seemingly contradicting concerns of the citizens of Great Falls. However, pork slaughterhouses having no effect on truck traffic is inconsistent with 1,000,000 hogs being moved daily in the United States (Brodersen, 2015). The leading explanations for the insignificant OLS results are both endogenous business location and difficult to quantify omitted variables related to differing regulations.

To address endogenous business location and uncover defendable causal estimates of pork slaughterhouses on truck traffic, historical hog density is hypothesized to be associated with current pork slaughterhouse location. Intuitively, cost-minimizing pork slaughterhouses locate near areas with many hogs, because transport costs associated with dead pigs is estimated to be $100 million annually (Fitzgerald, Stalder, Matthews, Schultz Kaster, & Johnson, 2009). Prior theoretical work (Picard & Zeng, 2005) argues transportation costs are under-acknowledged in the agriculture industry and empirically agricultural inputs have been used as an instrument for agricultural activity (McArthur & McCord, 2017). Agricultural firms have been shown to co-locate (Porter, 2003), further supporting the economic intuition for this hypothesis. Also, the historic hog data is measured prior to several important vertical integration periods in the hog industry, the construction of the interstates, or the growth in car usage.

Digitizing an old map to create data on hog density in 1948, hog density is strongly positively correlated with 2019 USDA-inspected pork slaughterhouses. In the preferred model, an additional 5,000 hogs in 1948 increases the number of pork slaughterhouses in a CBSA by 0.017. This is significant at 99% confidence and the Kleibergen-Paap first stage F-statistic that historic hog density = 0 is 50.58. As expected, the reduced form estimates are positive: an additional 5,000

[8] Section 2.1 discusses the role of truck transport in the hog industry.

hogs in 1948 increases 2017 truck miles by 0.4%. The main, second-stage, result is an additional 2019 pork slaughterhouse, due to 1948 hog density, leads to a truck traffic increase of 21.9%.

An assumption required to interpret treatment estimates as causal is that historic hog density only affects road conditions through pork slaughterhouses; several analyses support the validity of this assumption. If historic hog density causes differences in land use, agricultural agglomeration, or development of alternative transportation modes, and those in turn affect road conditions, the exclusion restriction fails. We condition on this set of controls and find that the results remain statistically significant and become larger in magnitude.

The slaughter industry is also closely associated with the rail industry which quickly proliferated across the United States in the later 1800's. After digitizing more maps, hog density is observed every 20 years going back to 1840, before rail became a widespread transportation mode. When 1840 hog density is used as an instrument in place of 1948 hog density, the results are unchanged. This is a strong indication of the validity of the exclusion restriction.

Lastly, analysis focuses on how quasi-experimental assignment of truck traffic affects road quality, specifically road roughness. The point estimates imply an additional pork slaughterhouse increases the distance a vehicles' suspension travels over 1 mile by between 2.228 inches and 7.709 inches. This translates into an additional daily truck mile causing a 0.046 increase in inches travelled by a vehicles' suspension over 1 mile.

Estimates lose significance at conventional levels in the preferred specification and we are inclined to believe that this not due to data or method problems, but is a limitation of being constrained by a linear functional form (Lochner & Moretti, 2015). To relax linearity, we examine differences in the cumulative distribution functions (CDFs) of truck miles, road roughness, and car miles, using Kolmogorov-Smirnov tests and another test for equality of distributions proposed by Goldman and Kaplan (2018), by top and bottom hog quartile. The evidence shows significant differences in the truck miles and road roughness distributions by 1948 hog quartile which are significant with p-values of 0. The average difference in the distributions, 0.2, is the same for truck miles and road roughness. Additionally, there is no difference in the distributions of car miles.

This paper moves the literature on the negative truck externalities of business production forward. Prior literature depends on experimental, descriptive, or accounting comparisons, but these approaches are unlikely to be interpretable as causal evidence or would be expensive to replicate.[9] To improve estimates, we appeal to cost-minimizing producers as a relevant way to instrument for industry location, which is a challenging task (Ellison, Glaeser, & Kerr, 2010).

Contrary to naïve approaches, the results show that pork slaughterhouses do increase truck traffic. Given the economic benefit of business activity, understanding how much additional truck traffic is important. It adds to the knowledge of the overall cost/benefit calculations necessary for a thorough evaluation of the impact of slaughterhouses.

Finally, a more general issue is the impact of trucks on roads. Commonly, structural models are used that incorporate parameters on how trucks affect road roughness (Parry, 2008).[10] This paper combines recent literature in agriculture (McArthur & McCord, 2017), agglomeration (Ellison et al., 2010), identification using maps and roads quality studies (Duranton & Turner, 2011) to estimate a parameter that defensibly represents a causal effect of trucks on road roughness. For future structural models that wish to incorporate the impact of trucks on road quality, the causal estimation of this parameter is crucial. While we acknowledge there are likely differences between hog trucks and trucks used in other industries such as fracking, we hope that our estimates can be useful for the policy discussion about the heavy trucks also resulting from fracking.

Below, relevant hog industry information is discussed and the data is described. Then, the cost-minimizing behavior that underlies the relationship between hog density and slaughterhouses is modeled, the hypothesis tests shown, and our identification assumptions discussed. Finally, main results are shown along with falsification and support for the identifying assumptions and proposed mechanisms.

2 Hog Industry Overview

In order to understand the research design and data used by this paper, some basic aspects of the hog industry must be explained. First, evidence is provided that live hogs are transported via truck. Next, the cost-reducing logic of why slaughterhouses locate close to hog-raising areas is explained.

[9] Extrapolating from Newbery (1988), this experiment would cost over \$933 million in 2020 dollars to replicate today.

[10] Additional negative externalities are pollution, congestion, and road accidents (Parry, 2008). In the EU, for example, trucks are often taxed both for diesel fuel consumption as well as the distance they travel on trips (Luechinger & Roth, 2016; Mandell & Proost, 2016). Other tax schemes multiply road travel taxes by truck axle.

Finally, vertical integration in the hog industry is described, importantly noting that integration has not affected the natural regional advantages that were present in the 1950s.

2.1 The Hog Industry Uses Trucks

The preferred mode of transporting live hogs is via trucks. The two main reasons are to reduce financial loss from damaged goods and animal welfare (Food and Agriculture Organization of the United Nations, 2001).[11] Hogs are temperamental animals, easily stressed by changing environments.[12] In addition, pigs cannot stand and balance well while moving (Aradom & Gebresenbet, 2013), cannot regulate body temperature (Huynh, Aarnink, & Verstegen, 2005), and pigs easily get motion sickness (Randall & Bradshaw, 1998). Because of environmental stress, the quality of the pork meat has been shown to degrade as stress levels increase (Gajana, Nkukwana, Marume, & Muchenje, 2013).[13] [14]

Although most hogs are likely transported by truck to slaughter, there is a possibility that the pork products are transported from the slaughterhouse by transport modes other than truck. Similar to the supply chain described in (Bai et al., 2007, pg. 33), pork products could be moved to their final domestic destinations by either truck or rail; this would imply rail transport is a complement to truck transport in the pork industry. To isolate the relationship between pork slaughterhouses

[11] Under US law, livestock are not allowed to be in a transport vessel for more than twenty-eight consecutive hours. For more temperamental livestock, such as hogs, constant confined conditions limit the consecutive hours of viable transport.

[12] To ease the stress of hogs, some slaughter facilities have implemented rules which govern the truck routes to and around slaughterhouses. (Miranda-De La Lama, Villarroel, & María, 2014, pg. 11) notes that some slaughterhouses force trucks to take detours to avoid wait times at the slaughterhouse and to keep the truck moving as to not cause stress for the hogs. This policy is similar in the way airports redirect planes around the surroundings of the airport.

[13] When a hog becomes stressed, it leads to lower lactic acid production in the animal. The lower lactic acid production causes a higher pH in the meat; higher pH levels are associated with lower meat quality (Gajana et al., 2013).

[14] In 2011, more than 111 million market weight pigs (roughly 130 kgs) were transported to slaughterhouses in the US (Johnson et al., 2013). Brodersen (2015) goes far to argue that on average there are more than 2,000 trucks transporting a total of 1,000,000 swines daily in the US: "*It is estimated that approximately 90% of the market pigs are moved offsite at weaning, and approximately 40% are moved again from nursery to another site for finishing. Also, accounting for movement of 3 million cull sows and 3 million replacement gilts makes a total of approximately 270 million pig movements on 260 workdays per year = over one million pigs on the road per day! If the 120 million market pigs alone are loaded 180 head per truck on 260 workdays, that is approximately 2,500 truckloads per day. Add another 500 trucks for weaned pigs, cull sows, and breeding stock and there are approximately 3,000 trucks moving pigs per day. You can work the numbers differently, but it's still a lot of trucks with pigs traveling through the country every day.*"

and truck miles travelled, non-passenger rail miles are controlled for in fully specified models to mitigate the potential bias from complementary transportation modes.

2.2 Why Slaughterhouses Locate Close to Hogs

Slaughterhouses locating close to hogs is cost-saving. Gosálvez, Averós, Valdelvira, and Herranz (2006) finds a positive association between travel distance from farm to slaughter and hog mortality. As the transportation distance via heavy truck increases, the mortality of hogs also increases. Pig deaths by transit have been estimated to cost the US pork industry \$50 to \$100 million annually (Fitzgerald et al., 2009, pg. 1156). If not dead from transit, the stress and "damage" caused by long transport times has been found to decrease pork meat quality (Gajana et al., 2013; Guàrdia et al., 2005).

2.3 Evolution of the US Hog Industry Since 1950

2.3.1 Vertical Integration

A major theme of the hog industry over the last 70 years is increased vertical integration. Throughout the 1990s, the pork industry saw rapid use of contracting production. (Key, 2004, pg. 255) explains that "the share of hog production under contract increased from about 18% in 1990, to about 28% in 1995, to almost 60% in 2000." Several reasons exist for why the industry has become more vertically integrated. A general reason is due to economies of scale. Several factors, related to economies of scale, that have changed include: industrialization leading to demand for factory-like development with highly specialized labor (Rhodes, 1995), advances in gene practices (Wang, Huang, & Zhao, 2017, pg. 2793), selective breeding led to shorter weaning and bigger litters (Kemp, Da Silva, & Soede, 2018), disease control (Chen, Madson, Miller, & Harris, 2012; Colomer, Margalida, & Fraile, 2020), and improved farm biosecurity practices (Pig Progress, 2015).[15] In sum, hog farms were housing bigger litters with higher survival rates that were produc-

[15] Hogs were traditionally bred to be large and produce fatty cuts of meat; advances in gene manipulation allowed the hog industry to cater to consumer demand for leaner cuts of pork (Wang et al., 2017, pg. 2793). The development of hog vaccines for H1N1 and other swine illnesses (Chen et al., 2012; Colomer et al., 2020) specifically decreased hog infection rates.

ing highly demanded cuts of meat. These advances lead to greater economies of scale and vertical integration of the hog industry followed.

2.3.2 Locational Persistence

In the US, hog farms and producers traditionally located close to corn farms because pigs eat corn (McBride & Key, 2013). Often, farmers would raise both corn and hogs to save on corn transport costs. This explains the large concentration of hogs in the Corn Belt region of the US as shown in Figure 1.[16] Over time, as corn transportation became cheaper due to increased railroad construction and usage, some hog producers began to move out of this traditional hog production area of the country. This fact, as well as the increased use of environmental and land use regulations in Corn Belt states, pushed out less profitable hog production to other areas, including North Carolina and western states (Roe, Irwin, & Sharp, 2002). This relatively small dispersion of economic activity can be seen in (Pork Checkoff, 2015, pg. 85).

The hog production that did remain in the Corn Belt transformed from small farms to larger, more integrated operations for reasons discussed in the above section, increasing capacity while decreasing the number of individual operations. As noted in (Key, 2004, pg. 255), between 1994 and 1999, "the number of U.S. hog farms fell by more than 50%, from over 200,000 to less than 100,000, while the hog inventory remained relatively stable." Despite farms leaving the region, in 2009, the Corn Belt still contained 76% of all feeder-to-finish farms across the United States Department of Agriculture (USDA)'s Agricultural Resource Management Survey (ARMS).

Although some areas in the US that were not traditional hog producers saw increases in hog production, the characteristics of locational persistence seen in Corn Belt producers has been documented in these less traditional areas as well (Roe et al., 2002). In other words, despite the reason for a slaughterhouse's initial location in the US, whether it be corn production or county regulations, it has been found that hog production attracts other hog production by means of "external economies of scale due to localization economies" (Roe et al., 2002).[17]

Therefore, based on the importance of agglomeration in the hog industry found in Roe et al. (2002), it would be reasonable to assume that areas that had hog production outside of the Corn

[16] Definitions of the "Corn Belt" in the US include the states of Ohio, Indiana, Illinois, Iowa, and Missouri.

[17] This has also been noted in some research reports (MacDonald, Ollinger, Nelson, & Handy, 2000).

Belt in 1948, as shown in Figure 1, albeit small areas, attracted the initial producers that left the Corn Belt region in more recent times.

3 Data

The three key variables come from three sources. Additional data for controls come from three additional sources. The sources and brief descriptions of the data are offered below. All data is aggregated to the core-based statistical area (CBSA).

3.1 Digitized Historic Hog Density Map

Figure 1 illustrates the source of the historic hog data; this map was from a series of livestock and crop maps printed in a *World Geography* textbook published in 1948 (Bradley, 1948).[18] This scanned map was digitized and georeferenced, transforming it into a usable shapefile format. Essentially, through this process, the number of hogs in 1948 in a (2019) CBSA can be calculated. The digitized data is presented in Figure 2.

3.2 USDA Inspected Pork Slaughterhouses

The slaughterhouse data comes from the United States Department of Agriculture (USDA), Food Safety and Inspection Service (FSIS). The data are restricted to only slaughterhouses that process pork products. Slaughter volume is only categorized in the data as 1-5, with 1 being the smallest slaughter volume and 5 being the largest. In the main analysis, no distinction is made between slaughterhouses with comparably larger or smaller slaughter volumes. Figure 2 shows the overlap of current pork slaughterhouses and 1948 hog density.

[18] This book can be found in open-source format here: https://openlibrary.org/books/OL6028680M/World_geography

3.3 Federal Highway Performance Monitoring System: Vehicle Miles Travelled and International Roughness Index

3.3.1 Vehicle Miles Travelled

Data on CBSA road networks come from the Highway Performance Monitoring System (HPMS). HPMS data is published yearly by the Federal Highway Administration (FHWA). The 2017 data includes the International Roughness Index (IRI) and Annual Average Daily Traffic (AADT) separately for cars and trucks. For truck AADT, we analyze data from the HPMS definition of "combination" trucks. In the HPMS, combination trucks are defined as vehicle classifications 8-13.[19] The data are restricted to only two lane roads.

The data are collapsed into CBSA-level measures of road roughness and truck traffic. Because of the variation in length of road segments, the IRI and AADT of truck traffic for each individual road segment are weighted by the percent each road segment represents in its CBSA road network; measurements from longer road segments in a CBSA's network, therefore, receive more weight than shorter road segments. Specifically, the weight is calculated:

$$w_{ij} = \frac{L_{ij}}{C_j} \tag{1}$$

where w is the weight for road segment i in CBSA j, L is the length of the segment i in CBSA j, and C is the total length of CBSA j's road network. Multiplying w_{ij} by the IRI and AADT of trucks for segment i in CBSA j produces length-weighted, by CBSA, IRI and truck miles travelled. Summing the weighted values by CBSA gives the average IRI and average AADT of trucks at the CBSA level in 2017.

3.3.2 International Roughness Index

The IRI is considered to be a measure of the general slope of a road segment. Over one mile of road, the number of inches of suspension travel a typical car would experience is the formal source of the IRI measure. (Duranton, Nagpal, & Turner, 2020, pg. 6). As IRI increases, the slope or

[19] Trucks with four-or-less axles, single-trailer trucks through seven-or-more axles, and multi-trailer trucks are all considered. See http://onlinemanuals.txdot.gov/txdotmanuals/tri/classifying_vehicles.htm for a visual representation.

roughness of the road also increases. Dips, ruts, and potholes could all increase the IRI of a road segment. A large IRI measure is associated with poorer overall pavement quality.

3.4 Third Factors Which Are Potentially Related to Pork Slaughterhouses And/Or Truck Miles Travelled

The analysis also controls for other factors that could be related to both truck miles travelled and pork slaughterhouses: related industries, land use, and alternative transportation modes. First, county-level data on number of establishments and employees of related industries are sourced from County Business Patterns (CBP) data and aggregated to the CBSA level.[20] Next, data on land use comes from European Space Agency satellites.[21] We use this data to control for the amount of urban land use and the amount of rainfed cropland in each CBSA (à la Turner, Haughwout, and Van Der Klaauw (2014) and Donaldson and Storeygard (2016)). Finally, it is possible that railroad travel is used as either a complement or a substitute for hog transportation.[22] The number of non-passenger rail miles in a particular CBSA are calculated from the railroad shapefile from Homeland Infrastructure Foundation-Level Data (HIFLD) using geographic information systems (GIS). Table 1 shows descriptive statistics of the aforementioned variables at the CBSA level.

3.5 Summary Statistics

Table 1 displays the summary statistics for key variables at the CBSA level. Panel A shows the full sample, while Panels B and C split the sample by the mean 1948 hog count. As shown in Panel B, CBSAs with more than 23.03 hog points in 1948 have more slaughterhouses and truck miles than

[20] These industries include: support crop establishments, support for animal production, non-slaughter meat processing, agricultural equipment manufacturing, other farm material wholesalers, and grain/field bean wholesalers. We focus on these industries, because (Porter, 2003, pg. 564) finds that agriculture does not agglomerate next to non-agriculture industries.

[21] The data includes the number of pixels of a certain type of land use in a county as classified by satellite imagery and an algorithm.

[22] It would be a substitute if hogs were moved by rail, instead of truck. However, truck transport for hogs is the preferred mode (Food and Agriculture Organization of the United Nations, 2001). It would be a complement if pigs were trucked into slaughterhouses, then the processed meat was transported out of slaughterhouses via rail. To our knowledge this relationship has not been explicitly studied yet, although some studies study how whether person transportation modes are complements or substitutes (e.g. Hall, Palsson, and Price (2018)).

areas with less than the mean hog count. CBSAs with greater hog points in 1948 also have higher average IRI (rougher roads) and more crop land cover.

3.6 Maps

The following subsections describe the visual evidence supporting the proposed first stage and reduced form relationships. The relationship between 1948 hog density and 2019 pork slaughterhouses is mapped in Figure 2 and the relationship between 1948 hog density and 2017 AADT of trucks is mapped in Figure 3.

3.6.1 1948 Hogs on Slaughterhouse Locations

Figure 2 combines the pork slaughterhouse data and Figure 1 in order to evaluate the correlation between 1948 hogs and 2019 pork slaughterhouses. If places with more hogs in 1948 have more slaughterhouses in 2019, this would be consistent with a strong first stage. A good example is Nevada in Figure 2e. There are no hogs in 1948 in the eastern part of the state and there are no 2019 pork slaughterhouses in the eastern part of the state. In southwest Nevada, there are hogs in 1948 and a slaughterhouse in 2019. In southwest Texas, there were no hogs in 1948 and there is only 1 pork slaughterhouse in 2019. However, in eastern Texas, there were hogs in 1948 and there are pork slaughterhouses. In general, this relationship is true across most, if not all, states.

3.6.2 1948 Hogs on Truck Miles Travelled

Figure 3 shows the correlation between hog points in 1948 and truck miles travelled in 2017. In each panel, the density of 1948 hog points is displayed in addition to the 2017 AADT of trucks at the CBSA level. Less AADT of trucks is represented by orange shades while more AADT of trucks is represented by blue shades. Darker shades of both colors represent a higher density of hogs in 1948. CBSAs with the darkest shade of blue and the lightest shade of orange would suggest a positive reduced form coefficient.

Throughout all regions of the US, Figure 3 confirms that areas with more hogs in 1948 generally have high levels of truck AADT in 2017. For example, Panel 3d shows a greater density of hog points in 1948 up the Atlantic Coast is associated with larger truck AADT in 2017. Also, in the traditional hog raising area shown in Figure 3a, this positive relationship between hogs points and truck AADT seems to hold strongly as well.

4 Model

The relationship of interest is how pork slaughter affects truck traffic, shown in Equation 2.

$$Truck\ Traffic = f(Slaughter)\ + u. \tag{2}$$

The effect of interest is $\frac{dTruck}{dSlaughter}$, which can be estimated using the following equation via OLS:

$$Truck_m = a + \beta Slaughterhouses_m\ + X_m\gamma +\ u_m. \tag{3}$$[23]

Here, $\beta = \frac{\partial Truck}{\partial Slaughter}$ is the point estimate. However, it is unlikely that β represents a causal effect without additional strong assumptions, such as conditional independence.[24]

The two main threats to interpreting β as causal are reverse causality and omitted variables. If there is a control, x', that is not in X_m, thus omitted, and correlated with both trucks and slaughterhouses, then β is a biased estimate of $\frac{dTruck}{dSlaughter}$.[25] An omitted variable that is challenging to collect data on and control for in the regression is many potentially overlapping policies and regulations. For example, inspection laws in the Unites States vary considerably across states .[26]

Even assuming omitted variables were accurately measured, available, and were modeled appropriately (i.e. correct ones chosen and specified correctly), it is no guarantee that conditional independence is satisfied. Another acute possibility is that firms consider the transportation infrastructure when making location decisions, such as whether a location is easily accessible by truck. If this is the case, then β is endogenous due to reverse causality and not informative about the causal effect of slaughterhouses on truck traffic.

[23] This abstracts away from slaughter being continuous and number of slaughterhouses being discrete.

[24] One could just build slaughterhouses according to a lottery and thus the assignment would mimic a randomized controlled trial.

[25] $\widehat{\beta} \neq \beta$.

[26] Some states do not require inspections of slaughter facilities, some require just state inspections, and other require federal inspections. These regulations also vary by the type of meat that is processed. See `https://nationalaglawcenter.org/state-compilations/meatprocessing/` for more details.

4.1 Theory

To recover causal estimates of β, quasi-random assignment of slaughterhouses addresses omitted variables and endogeneity. We appeal to cost-minimizing producers who optimally avoid high transport costs (McArthur & McCord, 2017; Picard & Zeng, 2005) and the persistence of natural regional advantages for certain industries (Glaeser & Gottlieb, 2009; Porter, 2003). This theory is useful in establishing plausible conditions under which a causal estimate of β can be recovered. In section 5, we test the predictions of the model.

4.1.1 Economic Behavior Underlying "First Stage"

In the simple model, firms' objective is to minimize costs as a function of transporting hogs and other factors (such as labor and capital), which can be considered a numeraire input. The objective function of the representative firm is

$$\min_{D,I} C(Distance\ to\ Hogs,\ Other\ Inputs) = p_D D + p_I I.$$

We set the price of the numeraire input to 1. Firms are subject to a target output level, $\overline{Q}$, and a production function. Assuming a Cobb-Douglass production function ($Q = D^{\alpha} I^{1-\alpha}$, with $0 < \alpha < 1$) and an interior solution, the optimal distance is

$$D^* = \frac{\overline{Q}}{(\frac{p_D(1-\alpha)}{\alpha})^{1-\alpha}}$$

4.1.2 Price of Distance and Hog Density

Importantly the optimal location for slaughter, D^*, depends on the price of transportation, p_D. If businesses were randomly assigned and one was in the low price of distance location, for a fixed target output, this business with the low price of distance would be able to reach a lower isocost line. For a location with many hogs, the price of distance is likely to be lower. Therefore, businesses prefer to locate there.

Using this intuition and two cases of p_D, we can illustrate the first hypothesis. There are two types of locations, one dense in hogs and one sparse in hogs, and the price of distance is dependent on location:

$$p_D = \begin{cases} h & if\ Location = Sparse\ Hogs \\ l & if\ Location = Dense\ Hogs. \end{cases} \tag{4}$$

where *h* stands for high cost and *l* stands for low cost, so that $h > l$. Taken literally, this means that an additional mile away from the input, hogs, costs more for a slaughterhouse in a sparse location. At a more abstract level, it is possible that p_D is higher in sparse areas because of search/bargaining costs or fewer supplier options, leading to market power and price markups.

4.1.3 Endogenous Slaughterhouse Location

In this model, firms observe whether current hog density is sparse or dense and price of distance, both of which may be related to infrastructure quality. Firms optimally locate in dense areas with a low price of distance so that lower isocost lines are obtainable, but this makes location decisions endogenous with respect to infrastructure quality and characteristics, the dependent variables of interest. This endogeneity makes the OLS β from Equation 3 uninformative.

4.1.4 Historic Hog Density, Contemporary Hog Density, and Pork Slaughterhouses

To estimate a causal effect, historic hog density is leveraged as quasi-random variation in current pork slaughterhouse location decisions. Our hypothesis is that corr(historic hogs, contemporary hogs) $>$ 0, due to persistent natural regional advantage (Porter, 2003). By predicting contemporary hog density, historic hog density predicts areas with lower costs. Via the transitive property, corr(historic hogs, l) $>$ 0 and corr(historic hogs, pork slaughterhouses) $>$ 0.

4.2 Econometric Method

Our hypothesis testing builds on the theoretical framework above. First, we test the relationship between historic hogs and contemporary pork slaughterhouses. Then, we examine whether a predicted pork slaughterhouse has non-zero effects on truck miles travelled.

4.2.1 First Stage

We do not observe the price of distance, but it is not necessary to in our framework. We observe historic hog density and pork slaughterhouses, so we can test the relationship between historic hog density and contemporary pork slaughterhouses using the following equation

$$(5) \qquad 2019\ Pork\ Slaughterhouses_c = \alpha + \phi_1 1948\ Hogs_c + X_c\gamma + \mu_S + \lambda_m + \epsilon_c,$$

in which *c* stands for core-based statistical area, *m* stands for metro/micro area, *S* for state, and μ_S is a vector of state fixed effects. The theoretical model predicts that $\phi_1 > 0$. The vector of state fixed effects means that effects are estimated within state and estimates are not an artifact of specific states being prolific hog producers, like Iowa.

4.2.2 Second Stage

In the second stage, we test the relationship of interest: how do contemporary pork slaughterhouses affect contemporary infrastructure outcomes? To do this, the quasi-random assignment from the first stage is leveraged in the following equation

$$(6) \qquad Truck\ Miles_c = a + \beta_1 Pork\ Sl\widehat{augh}terhouses_c + X_c\eta + \mu_S + \lambda_m + u_c,$$

in which the subscripts are the same and $Pork\ Sl\widehat{augh}terhouses$ are the predicted values from Equation 5. Due to the necessity of trucks for pork slaughterhouse operation, it is expected that $\beta_1 > 0$.

4.3 Identification

For β_1 to represent a causal effect of pork slaughterhouses on truck miles, several conditions must be satisfied. Under 2 conditions, relevance and exclusion, β_1 is a causal effect. With 3 additional conditions, β_1 represents the heterogeneous local average treatment effect (LATE).

4.3.1 Relevance

Relevance means that the instrument, historic hog density, predicts the treatment, pork slaughterhouses, with adequate precision. The theory details that persistent natural regional advantage leads to a strong positive correlation between historic hog density and contemporary pork slaughterhouses. We support relevance in multiple ways: visually, non-parametrically, and parametrically.

In the parametric analysis, one rule-of-thumb that has been used is the F-statistic that $\phi_1 = 0$ must be sufficiently large (i.e. F>10). However, recent literature suggests that F $>$ 10 may not be sufficient. As an additional check, since our equation is just identified (1 instrument, 1 endogenous variable), we also report Anderson-Rubin confidence intervals which are efficient for potentially weak instruments.

4.3.2 Exclusion

The exclusion restriction means that historic hog density affects contemporary road characteristics, such as truck traffic, only through its relationship with contemporary pork slaughterhouses. This assumption, that historic hog density only affects truck traffic through pork slaughterhouses is shown by (1) in Figure 4. This assumption is reasonable, because hog density in 1948 was measured before the interstate was built and automobile travel was widespread. The instrument hogs is measured in 1948 and pork slaughterhouses is measured in 2019, a 71 year lag.[27,28] With specific attention on the proposed instrument in the current work, not only is there a long lag between the instrument and the outcome variable (69 years), but the hog industry did not start to become more centralized until more recently. This is similar to the arguments made in Duranton and Turner (2011) and other literature using historic instruments for contemporary outcomes (Banerjee & Iyer, 2005; Glaeser, Kerr, & Kerr, 2015; Nunn & Wantchekon, 2011).

The question here is slightly different from literature like Duranton and Turner (2011), since the instrument for business location is an input at a substantial lag. Recent literature has leveraged distance to contemporary agricultural inputs as an instrument (McArthur & McCord, 2017).

[27] Duranton and Turner (2011) uses a 1947 road map to instrument for road construction in a year as early as 1983 – a 46 year lag.

[28] Other instruments are used as well such as exploration routes in the early 1800s and planned railroad routes in 1898.

This instrument is similar, except it leverages a significant lag like the regional/urban literature (Duranton & Turner, 2011) to add to the likelihood excludability is reasonable.

Conditional Exclusion A forceful point made in Duranton and Turner (2011) is conditional exclusion can address potential violations. We have identified several ways through which exclusion could be potentially violated and control for them in X in both equations. Perhaps the most obvious one is that other industries may co-locate that are inputs or outputs for hogs or pork slaughter, so we control for the number of employees in related agriculture industries. We focus on agricultural industries, because Porter (2003) finds that not many other industries co-locate with agriculture. The next issue is alternative modes of transportation are associated with slaughter, primarily rail. If additional hogs lead to more rail transport, then exclusion is violated, so we control for the miles of non-passenger rail in a county. Finally, it is possible that hog density leads to differences in land use, so we control for satellite classification of urban and crop land use.

Figure 4 diagrams conditional exclusion also. By including the described controls, as shown in (2), we no longer have to assume historic hog density does not affect them. This in turn increases the likelihood that (1) is reasonable and exclusion is satisfied.

4.3.3 Independence, SUTVA, Monotonicity

The three additional assumptions required to interpret β_1 as a heterogeneous LATE are independence of the instrument and the outcome, the stable unit treatment value assumption (SUTVA), and monotonicity. Independence is likely satisfied as hogs in 1948 are as good as randomly assigned with respect to current truck traffic and road roughness. Monotonicity means that dense hog areas cannot lead to less contemporary pork slaughterhouses. These areas would be considered defiers and call into question the estimated effects. A way this might be violated is if hogs cause externalities that lead to legislation banning pork slaughterhouses, however non-parametric analysis suggests that denser hog areas are very unlikely to lead to reductions in pork slaughterhouses.

One final concern is SUTVA. SUTVA is violated if one area being treated by pork slaughterhouses affects the potential outcomes of a neighboring location. To reduce the likelihood of SUTVA violations, the unit of analysis is core-based statistical areas which are larger than counties, in general. This also follows the reasoning and unit of analysis used by Duranton and Turner (2011).

5 Results

First binscatters are shown to non-parametrically examine the data for the expected relationships. Next, OLS and 2SLS results are compared. Afterwards, car miles are used as falsification and exclusion is examined by adding theoretically important control variables and using older measures of hog density. Then, we split the slaughterhouse data into different sizes by slaughter volume and allow slaughter volume to enter as an additive total instead of a discrete count of establishments. To conclude the relationship between slaughterhouses, truck traffic, and road roughness is examined via OLS, 2SLS, and tests for equality in distributions.

5.1 Binned Scatter Plots of Reduced Form and First Stage

Figure 5 shows the binned scatter plots of the first stage (Equation 5) and reduced form. Figure 5a shows how an additional 5000 hogs is related to contemporary slaughterhouses in a CBSA. As expected, there is a positive, monotonic-like relationship between historic hog density and contemporary pork slaughterhouses. This relationship also appears strong in that it is highly likely that the slope depicted is statistically greater than zero. In Figure 5b, there is also a positive relationship between hog points in 1948 and log truck miles, showing evidence of a necessary, expected reduced form effect.[29] Both linear fits are influenced by the high leverage point at the right of the graph with many hogs, but removing this point would not drastically impact the slope of either fitted line. The underlying data is consistent with the expected results.

5.2 OLS

We briefly present results from using OLS from Equation 3 in Table 2. In column 1, pork slaughterhouses do not have statistically significant effect on truck miles travelled. Adding state fixed effects (column 2) results in a statistically significant positive effect, but after including controls

[29]This conclusion is unchanged when using truck miles.

this effect goes to 0.[30] Furthermore, using interacted state and whether a CBSA is a micro or metro area fixed effects is unable to recover a statistically significant effect. Overall, significance of OLS results are sensitive to fixed effects and whether controls are included.

5.3 2SLS, CBSA

Tables 3 presents the main results from estimating the two stage approach described in Equations 5 and 6 using CBSA level data. The table tests for instrument relevance using the F-statistic from testing the null hypothesis that Hog Density in 1948 = 0 in Equation 5. The F-statistics are well above in 10 in all panels, suggesting the instrument is adequately strong.

Table 3 shows results of the two-stage model with different combinations of fixed effects and standard error clustering. Panel A, using no other explanatory variables and heteroskedastic-robust standard errors, shows an additional pork slaughterhouses increase truck AADT by 37.5%. Panel B adds a dummy for the type of CBSA – metropolitan or micropolitan – as well as state-clustered standard errors. The magnitude of the second stage coefficient falls to a 19.3% increase, but is still positive and statistically significant. Metropolitan statistical areas in general have a high traffic of trucks, hence adding metropolitan fixed effects allows us to better isolate the impact of pork slaughterhouse on AADT truck traffic.

Panel C includes an interaction of state and metro fixed effects as well as state-clustered standard errors. Again, pork slaughterhouses, through their relationship with 1948 hog density, significantly increase the AADT of trucks. The estimates in all three panels are larger than the OLS results because 2SLS leverages psuedo-random variation while OLS suffers from potential endogenous location decisions. Also, unlike the naïve OLS results, the regression coefficients' significance does not depend on whether state fixed effects are included, making the 2SLS results less fragile to alternative specifications.

The preferred specification is shown in Panel C, which uses state X metro FE's. This specification is preferred, because while the minimum population in a metro area is uniform, it is highly likely that the difference between micro and metro areas differ by state. For example, CBSAs in

[30] State fixed effects control for the likely possibility that truck miles and/or pork slaughterhouses are correlated within states.

eastern states are usually much smaller in terms of land area and more densely populated compared to CBSAs in the western states.

5.4 Evidence Related to Exclusion Restriction

The primary threat to considering the estimates a causal effect is violations of the exclusion restriction, meaning that historic hog density is correlated with the error term, u_c in Equation 6. First, we adjust the required condition by adding controls, so that the requirement that $Cov(1948\ Hogs, u_c \mid S, M) = 0$ becomes

$$Cov(1948\ Hogs, u_c \mid S, M, X) = 0, \tag{7}$$

in which X are the theoretical elements of u_c that 1948 hogs are correlated with. If Equation 7 is satisfied, then conditional exclusion is satisfied and estimates can be given a causal interpretation. The second strategy for assessing the plausibility that historic hog density is excludable is using measurements of hog density at longer lags to sever the connection between hog density and variables correlated with second stage outcomes.

5.4.1 2SLS, CBSA w/ Controls

Table 4 extends the main 2SLS results by systematically adding control variables. By including controls, it is more likely that the exclusion restriction, which is necessary for a causal interpretation, is satisfied. The justification for each set of controls is that each may be correlated with outcome and the instrument.

Panel A includes controls for the number of employees in industries related to the pork slaughter industry.[31] The effect of an additional pork slaughterhouse on truck AADT in a CBSA is positive and significant at 90% confidence level. Panel B includes controls for the types of land cover in the CBSA (either urban or crop land). Again, our results are significant, showing our findings are not driven by urban-rural divide in land use. Additionally, the upward bias from omitting

[31] The related industries are: support for crop establishments, support animal production, non-slaughter meat by-product processing,agricultural equipment manufacturing, other farm material wholesalers, grain/bean wholesalers, and farm supply merchant wholesalers.

rail miles (Panel D) is not consistent with rail being a substitute for transporting hogs.[32] Our results, no matter what controls are added, are statistically significant, highlighting that the empirical conclusions are not sensitive to which controls are included.

5.4.2 Older Hogs

To provide evidence the results are not sensitive to violations of exclusion, hog density from years before 1948 is used. Figure A.1 displays historic hog densities from 1840-1920 that are digitized to create this data. This bolsters the argument for exclusion, because it less likely that more historic instruments violate exclusion due to the longer lag. Additionally, several potential confounders, such as trains, were far less established as far back as 1840.

Figure 6 displays the 2SLS second-stage regression coefficients of the effect of pork slaughterhouses on truck traffic for different ways of including the several years of instruments available.[33] The first coefficient in Figure 6 is the Wald estimate from Panel C of Table 3. The interval depicted beside it is the Anderson-Rubin 95% percent confidence interval. This interval is efficient and robust to weak instruments in the case of just-identified 2SLS.[34]

When we use 1840 hog density as an instrument in place of 1948 hog density, the results remain virtually unchanged. The slaughter industry is closely associated with the rail industry which quickly proliferated across the United States in later 1800's. Using a measure of hog density before the advent of widespread rail transportation infrastructure reinforces our identification strategy and exclusion restriction arguments for using an historic input to instrument for current industry location. Using 1860 and 1880 hog density as an instrument, the results are no longer statistically significant. It is likely that this loss of significance is due to the large-scale disruption happening across America at this time due to the American Civil War (1861-1865) and subsequent Reconstruction period.

[32] Since pork slaughterhouses locate near rail to transport the product after it has been processed, the correlation between pork slaughterhouses and rail miles is positive. That means that the sign of the bias is entirely dependent on the direction of the correlation between truck miles travelled and non-passenger rail miles. If rail and trucks were substitutes, the correlation would be negative and the coefficient would be biased downwards, but the data is consistent with rail and trucks being complements.

[33] Figure A.2 shows the over-identified estimates.

[34] In the case of very weak instruments, A-R confidence intervals may not be able to find one or both bounds.

5.5 Car Miles Falsification

Table 5 presents a falsification exercise using logged car AADT as the dependent variable instead of logged truck AADT. The proposed mechanism suggests that pork slaughterhouses should only affect truck AADT and not car AADT because trucks are used to transport hogs. In Panel A, this seems to be the case: the effect of pork slaughterhouses on logged car AADT is not significant. Adding state by metro fixed effects and state clustered standard errors actually yields a significant, positive result with 90% confidence. A possible reason for the significant results in Panel B could be the increased car travel of employees of related industries when a pork slaughterhouse is in a CBSA. However, this effect goes away in Panel C when adding agglomeration controls such as employees in related industries, land usage and non-passenger rail miles. That car miles are not significantly associated with pork slaughterhouses is consistent with the proposed mechanisms.

5.6 Heterogeneity by Slaughterhouse Volume

Table 6 shows 2SLS results of pork slaughterhouses on truck AADT at the CBSA level, but separates the analysis by the slaughter volume of the slaughterhouses. In the USDA data, a slaughterhouse is given a slaughter volume classification of 1 - 5. Most of the slaughterhouses in the data are either a 1 or a 2 (n = $>$ 400/550) in terms of slaughter volume. If it is the transportation of hogs by trucks that drives the increase in truck traffic in our main analysis, then we should expect to see smaller or no results if only the smallest slaughterhouses, in terms of slaughter volume, are considered.

Panel A displays the 2SLS results for the smallest slaughter volume facilities (volume = 1, physical size = very small). An additional 5,000 hogs in 1948 increases the number of slaughter volume = 1 pork slaughterhouses by 0.004 and is significant at 95 percent confidence. Despite this, the F-statistic that ϕ_1 = 0 is 5.054 (relatively weak) and the second stage Wald coefficient is not statistically significant. However, the A-R confidence interval shows that this is statistically significant with a minimum lower bound effect size of 18.9 percent increase in truck traffic.

Panel B displays the 2SLS results for the next largest slaughter volume facilities (volume = 2, physical size = very small). An additional 5,000 hogs in 1948 increases the number of slaughter volume = 2 pork slaughterhouses by 0.005 and is significant at 99 percent confidence. Historic hog density predicts more larger slaughterhouses than smaller slaughterhouses and this prediction is more precise. Due to historic hog density more larger slaughterhouses and more precisely, the F-statistic that historic hog density is 0 in the first stage is 20.97. Correspondingly, the second stage Wald estimate is statistically significant at 95 percent confidence. Furthermore the lower bound of the A-R confidence interval is also larger than lower bound of the A-R confidence interval in Panel A, which is consistent with the expectation of larger effects of larger slaughterhouses on truck miles.

Panel C takes a different strategy to break down discrete number of slaughterhouses into a more continuous measure of intensity of production. Ideally, the exact number of hogs slaughtered would be observable, however in the publicly available USDA data only a 1-5 measure of slaughter intensity is known. If the difference between a slaughter volume of 1 to 2 is the same as the difference between 4 to 5, then it is appropriate to sum over all pork slaughterhouses in a given CBSA. We cannot confirm that summing volume of all slaughterhouses is strictly appropriate, but believe this is an important piece of evidence given the potential heterogeneity across pork slaughterhouses. The results show that an additional unit of slaughter volume of increases truck miles traveled by 9.6 percent, an effect that is significant at 99 percent confidence. Overall, being more careful to account for differences in slaughter volume is consistent with the proposed mechanisms.

5.7 Road Roughness as Outcome

Next we focus on the second research question: do slaughterhouses affect road roughness? First naïve OLS results are shown. Then, the same 2SLS strategy is applied. Finally, the section concludes with an analysis that compares the distributions of pork slaughterhouses, truck miles, road roughness, and car miles by top and bottom historic hog quartile.

5.7.1 Naïve OLS

Figure 7 shows the naïve estimates for the effect of truck AADT on IRI. A higher value of IRI is associated with poorer road quality. Despite adding controls and various fixed effects, the OLS results provide an unintuitive result. OLS estimation yields significant results, but suggests that increased truck traffic *improves* road quality. Previous studies (Abramzon et al., 2014; Bai et al.,

2007) and experiments (Highway Research Board, 1962) suggests increased truck traffic decreases road quality. It is clear that naïve estimations suffer from bias.

5.7.2 2SLS

Table 8 presents estimates of how pork slaughterhouses and truck traffic affect road roughness. Panel A includes only state fixed effects. The reduced form (column 1) is statistically significant at 99 percent confidence and estimates that 5,000 additional hogs in 1948 increases contemporary international roughness index by 0.15. The first stage remains strong and the second stage estimates that an additional pork slaughterhouse increases the inches a vehicles' suspension travels over 1 mile by 7.7 inches, an effect that is statistically significant at 99 percent confidence.

Panel B adds metro/micro fixed effects and the reduced form (column 1) is only one-third the size of Panel A; it is no longer statistically significant at conventional levels. Nevertheless, the second stage coefficient in column 3 remains statistically significant at 90 percent confidence, implying that an additional pork slaughterhouse increases IRI by 3.07. In the preferred specification, interacted state and metro FE's, the reduced form is not close to significant and this causes the second stage to be insignificant also. It could be considered that the preferred specification loses significance on the main result, so we return to this question using a different strategy.

Panel D uses truck miles as the endogenous treatment of interest. The first stage results show that an additional 5,000 hogs in 1948 leads to an additional 3.2 daily truck miles, an effect which is statistically significant at 90 percent confidence. The first stage F is only 3.729, but the second stage Wald is found to be statistically significant at 95 percent and the A-R confidence interval finds significance above 99 percent confidence. The coefficient can be interpreted as an additional daily truck mile traveled leads to a 0.046 increases in IRI. Similar to Panel C, when the preferred specification is used in Panel E, the second stage coefficient loses statistical significance.

5.7.3 Cumulative Distribution Functions

A marked difference between the 2SLS results in Table 8 and the naïve OLS results presented in Table 7 is that 2SLS can recover the expected positive sign for point estimates. However, Table 8 sometimes loses statistical significance on how pork slaughterhouses and truck miles travelled affect road roughness depending on the specification. In return for losing some direct interpretability,

Figure 7 tests for differences in distributions of outcomes by historic hog quartile without functional form assumptions.

First Stage Figure 7a shows the "distributional version" of the first stage hypothesis that hog quartile affects pork slaughterhouse location. The Kolmgorov-Smirnov test rejects that the distributions are equal at 99 percent confidence. A more modern test for differences in distributions fails to reject equality of distribution. Failure to reject may be due to the Goldman and Kaplan (2018) test becoming conservative in the presence of ties in the distributions, which is the case with the discrete nature of count of slaughterhouses (Goldman & Kaplan, 2018).

The distribution of slaughterhouses has many 0's and is skewed to the right. But, while 80 percent of the bottom hog quartile has no slaughterhouses, only 60 percent of the top hog quartile has no slaughterhouses. Since mean slaughterhouses is about 0.46, at the mean of slaughterhouses, the difference in the distributions is 0.2. There is some evidence that the distribution of contemporary slaughterhouses is different for the top and bottom historic hog density quartiles.

Falsification One might be concerned that locations with many hogs develop differently than locations without many hogs. One way that this would be apparent in the data is if the distribution of car miles by historic hog quartile are different. We test for differences in distributions in Figure 7b. Neither the K-S or the Goldman and Kaplan (2018) test finds significant differences between the car miles distributions by hog quartile. Moreover, in Figure 7b, the distributions often overlap and there is very little visible space between them, confirming the intuition of the tests and difference from the first stage in Figure 7a.

Reduced Form on Truck Miles and Road Roughness Our final two tests for equality of distributions focuses on the main outcomes of interest, truck traffic and road roughness. Figure 7c shows the tests for equality of distributions of log truck miles by historic hog density. Both the K-S and Goldman and Kaplan (2018) tests for equality of distributions reject equality at 99 percent confidence. The average vertical distance between the distributions is calculated as 0.19/0.21 and this is also roughly the difference at the mean of log truck miles, 5.93 and the difference at the mean pork slaughterhouses.

Next, the differences in distributions of IRI are tested in Figure 7d. Both tests reject equality of distributions at 99 percent confidence. Furthermore, the average vertical distance between the distributions is 0.2 which is exactly the same as the first stage and reduced form with truck traffic. This is the first evidence of a causal relationship between truck traffic and road roughness that has been estimated and shown in observational data. Finally, Figure 7d offers a potential explanation for sometimes insignificant results in Table 8: a long right tail of the distribution of IRI for lower hog quartiles. There is a city in the bottom hog quartile with an IRI of 207.4, which is nearly a whole standard deviation above the max of IRI in the upper hog quartile.

6 Conclusion

This paper investigates the causal effect of pork slaughterhouses on truck traffic and truck traffic on road roughness. Naïve results indicate no effect of slaughterhouses on truck traffic, which is suspect given the necessity of trucks for slaughterhouse operation. To circumvent endogenous business location making OLS estimates uninformative, we appeal to a model where cost-minimizing slaughter firms prefer high density hog areas because these areas have lower transportation costs.

Historic hog density is strongly correlated with current pork slaughterhouses. Using two-stage least squares, pork slaughterhouses increase truck miles travelled, but do not affect car miles travelled. By including measures of co-locating industries, alternative transportation modes, and land use and using hog density as far back as 1840, exclusion is supported as valid suggesting causal interpretations are reasonable. Historic hogs also more precisely predicts larger volume slaughterhouses.

Using OLS suggests greater truck miles decrease road roughness, but this makes no sense from an engineering perspective. While the precision of 2SLS estimates of pork slaughterhouses and truck miles on road roughness are sensitive to which fixed effects are included, they are of the expected sign (positive). Furthermore, non-parametric analysis suggests that there are statistically different distributions of truck traffic and road roughness by how many hogs existed in that CBSA in 1948. Furthermore, the average difference in the distributions is almost exactly the same for both trucks and road roughness.

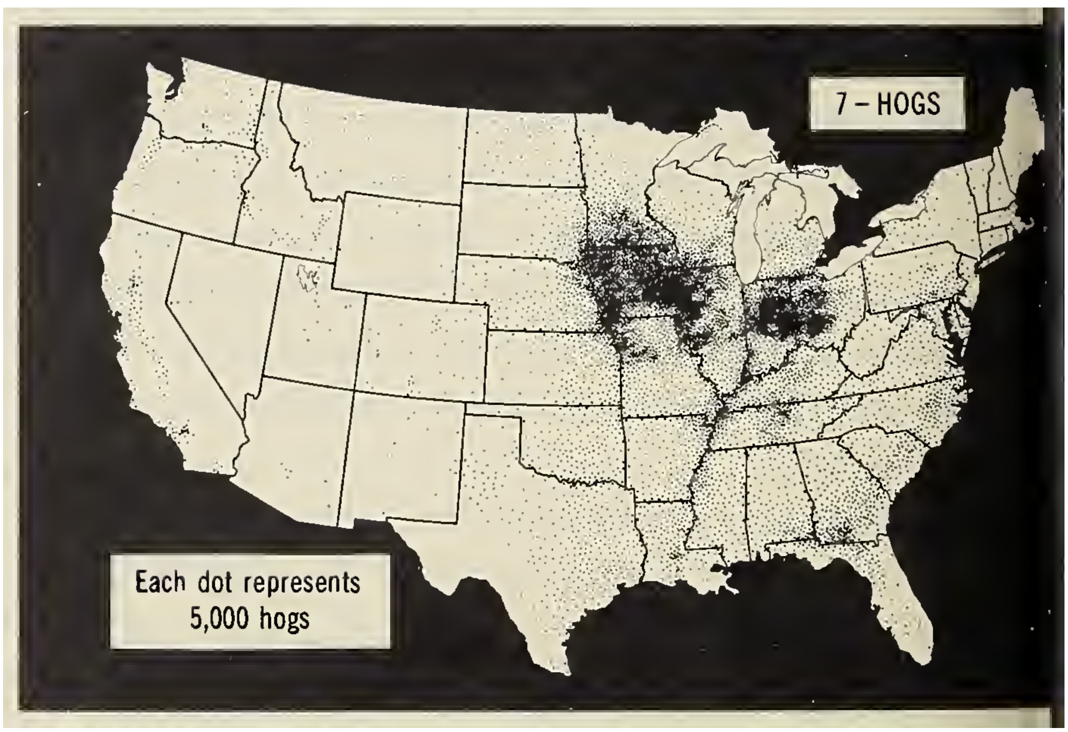
The first contribution of this paper is combining literatures on agricultural inputs (McArthur & McCord, 2017) and using historic measurements (Duranton & Turner, 2011) as a basis for quasi-random variation in contemporary agriculture business location. We document that by not using this design, the results are inconsistent with basic background on pork slaughterhouse operation. The extension of these two designs is invaluable in recovering causal estimates consistent with industry background.

The second contribution of this paper is estimating the causal effect of additional pork slaughterhouses on truck traffic. This is critical for policy-making related to pork slaughterhouses, because estimates that do not address causal issues (like endogenous business location, omitted variables, etc.) suggest no negative truck externalities of pork slaughterhouses. If policy-making is based on these null results or other descriptive work, then it may lead to sub-optimal policies which harm individuals such as those in Great Falls, Montana.

The third contribution is estimating the causal effect of pork slaughterhouses and truck traffic on road roughness in a nationwide sample. Truck traffic is responsible for several negative externalities that harm all who use the road network. Of these externalities, such as pollution and accidents, an important externality is road infrastructure damage. Not only do damaged roads increase vehicle repair costs for drivers, but they increase maintenance repair bills to towns and cities across the US (Abramzon et al., 2014).

The most general policy parameter is how road roughness responds to truck traffic. Significant resources have already been expended to reveal information on this parameter (Eschwege, 1979), but trucks and roads have changed since then and the same experiment would be very costly to re-do. Despite the importance of the road roughness parameter for structural urban models about truck taxes and policy-making in regards to heavy industry such as agriculture and fracking, prior parameters calculated/estimated from observational data (Abramzon et al., 2014; Bai et al., 2007) are not informed by modern causal inference methods. This paper fills this gap and although some estimates are not precise, documents how historic hog density is positively related to pork slaughterhouses, truck traffic, and road roughness in a causal framework in nationwide data.

Figure 1: Hog Density in 1948



Note: Figure shows hog density in 1948. The map comes from the 1948 textbook *World Geography*.

Table 1: Descriptive Statistics, By 1948 Hog Count, CBSA

Panel A: Aggregate

	Mean	Median	SD	Min	Max	Count
Hog Points in 1948	23.03	16.25	22.80	0.0	186.3	911
Pork Slaughterhouses	0.46	0.00	0.88	0.0	9.0	911
Truck Miles	1149.07	791.32	1170.59	13.6	11914.0	911
Ln(Truck Miles)	6.61	6.67	0.99	2.6	9.4	911
Ln(Car Miles)	5.93	6.01	0.75	2.5	7.9	911
IRI	101.96	100.03	23.55	26.1	207.4	911
Non-Passenger Rail Miles	163.23	104.35	206.42	0.0	2480.5	911
Urban Land Cover Pixels (90X90)	2003.05	485.00	5827.37	16.0	72893.0	911
Crop Land Cover Pixels (90X90)	20190.09	6899.00	34156.36	1.0	328431.0	911
Amount of Land	4.69e+09	2.60e+09	5.91e+09	2.8e+08	7.1e+10	911
Metro Area	0.41	0.00	0.49	0.0	1.0	911

Panel B: 1948 Hogs > Mean

	Mean	Median	SD	Min	Max	Count
Hog Points in 1948	45.43	39.09	23.65	23.2	186.3	333
Pork Slaughterhouses	0.66	0.00	1.16	0.0	9.0	333
Truck Miles	1195.42	887.13	1069.71	51.1	6489.4	333
Ln(Truck Miles)	6.72	6.79	0.90	3.9	8.8	333
Ln(Car Miles)	5.89	6.01	0.76	2.5	7.8	333
IRI	104.68	102.31	21.87	38.1	185.3	333
Non-Passenger Rail Miles	216.55	129.08	272.96	1.2	2480.5	333
Urban Land Cover Pixels (90X90)	2905.47	599.00	7491.35	81.0	69527.0	333
Crop Land Cover Pixels (90X90)	29113.44	16947.00	32868.87	64.0	213704.0	333
Amount of Land	4.28e+09	2.33e+09	5.56e+09	6.8e+08	7.1e+10	333
Metro Area	0.49	0.00	0.50	0.0	1.0	333

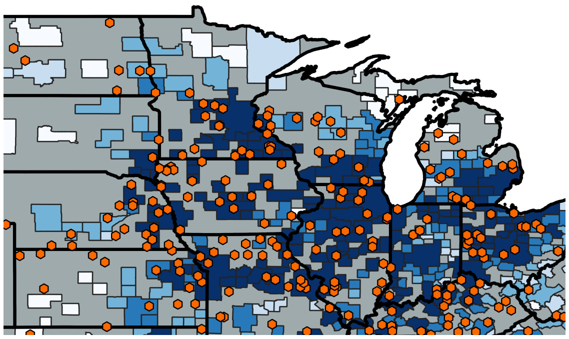
Panel C: 1948 Hogs < Mean

	Mean	Median	SD	Min	Max	Count
Hog Points in 1948	10.12	9.70	6.44	0.0	23.0	578
Pork Slaughterhouses	0.34	0.00	0.65	0.0	4.0	578
Truck Miles	1122.37	765.90	1225.05	13.6	11914.0	578
Ln(Truck Miles)	6.55	6.64	1.04	2.6	9.4	578
Ln(Car Miles)	5.95	6.00	0.75	3.1	7.9	578
IRI	100.40	98.64	24.35	26.1	207.4	578
Non-Passenger Rail Miles	132.52	88.70	147.47	0.0	1276.4	578
Urban Land Cover Pixels (90X90)	1483.15	441.00	4530.41	16.0	72893.0	578
Crop Land Cover Pixels (90X90)	15049.13	3523.50	33851.19	1.0	328431.0	578
Amount of Land	4.93e+09	2.69e+09	6.09e+09	2.8e+08	5.5e+10	578
Metro Area	0.37	0.00	0.48	0.0	1.0	578

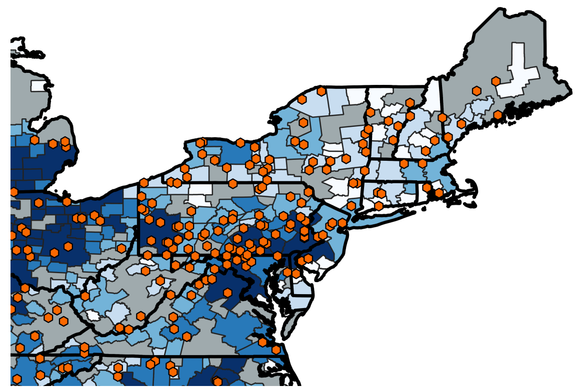
Notes: Core-based statistical area (CBSA) descriptive statistics for hog points, slaughterhouses, and vehicle miles travelled are displayed above. 1 hog point in 1948 = 5,000 hogs. IRI = International Roughness Index. Panel A shows the descriptive statistics for the entire 911 CBSA sample. Panel B includes only those CBSAs with 1948 hog points greater than the sample mean, while Panel C includes those CBSAs with 1948 hog points less than the sample mean. SD = standard deviation.

Figure 2: Overlap of 1948 Hog Density and 2019 Pork Slaughterhouses by Region, CBSA Level

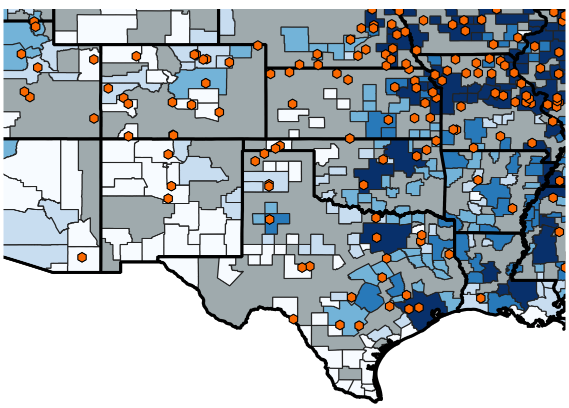
Panel A: Midwest North



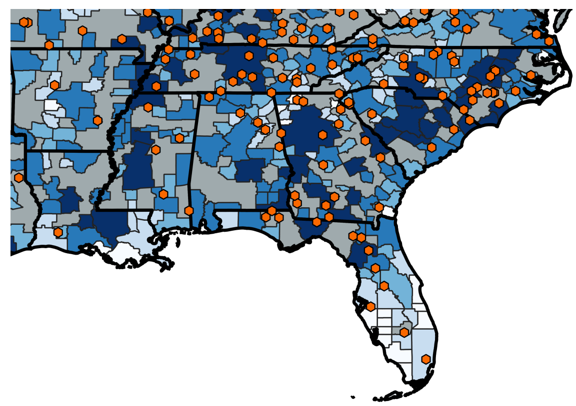
Panel B: East North



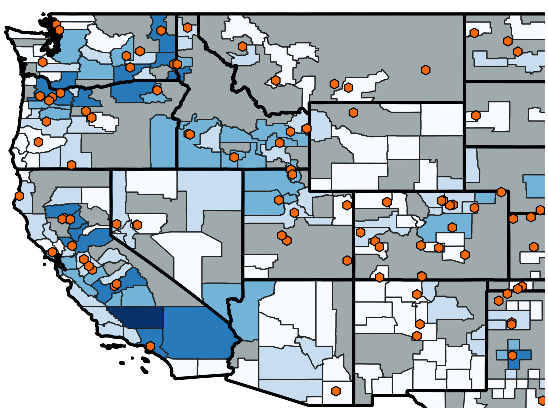
Panel C: Midwest South



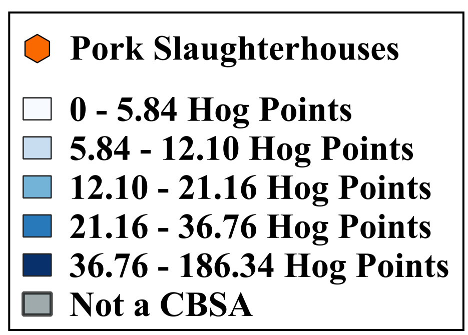
Panel D: East South



Panel E: Western States

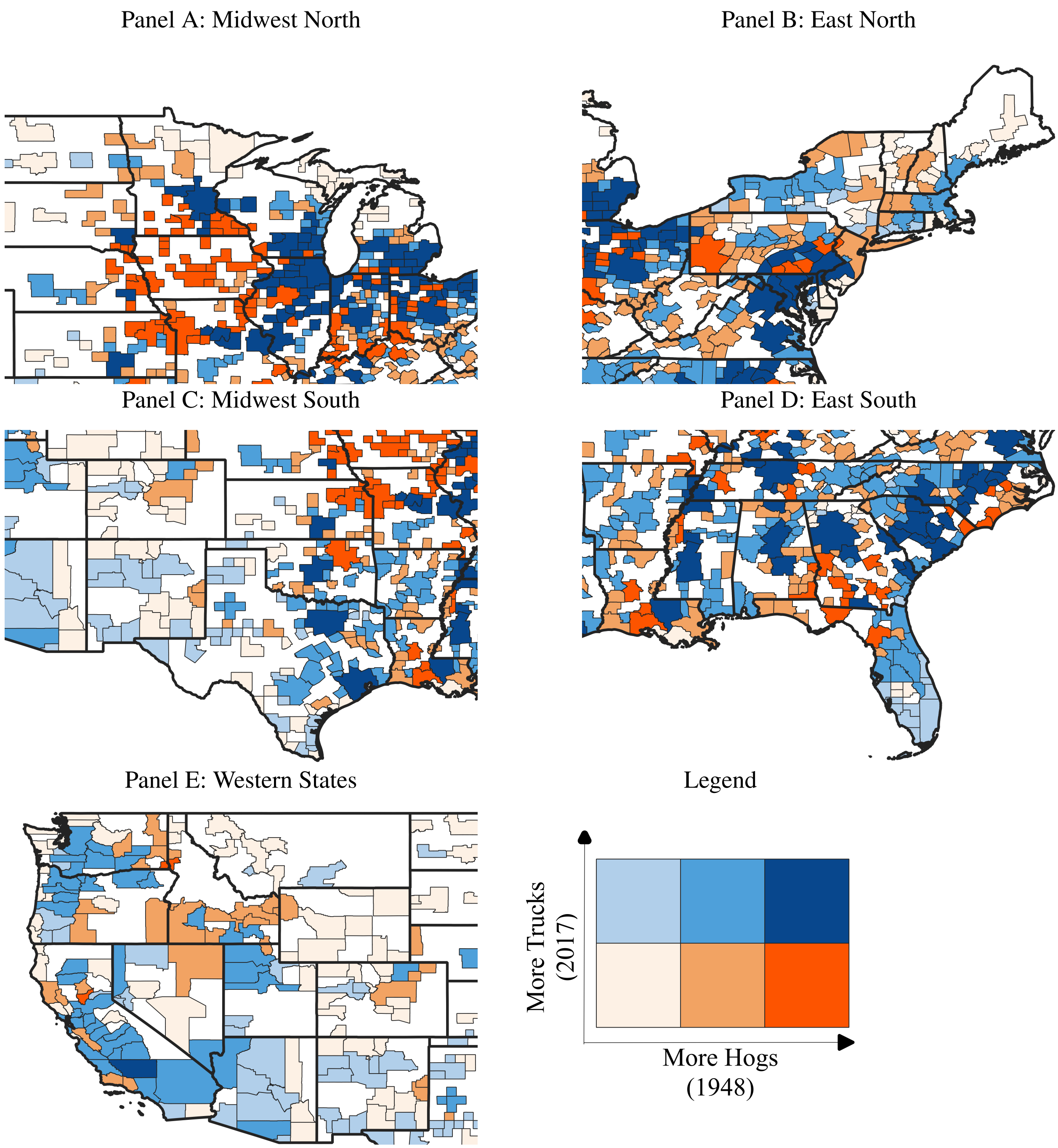


Legend



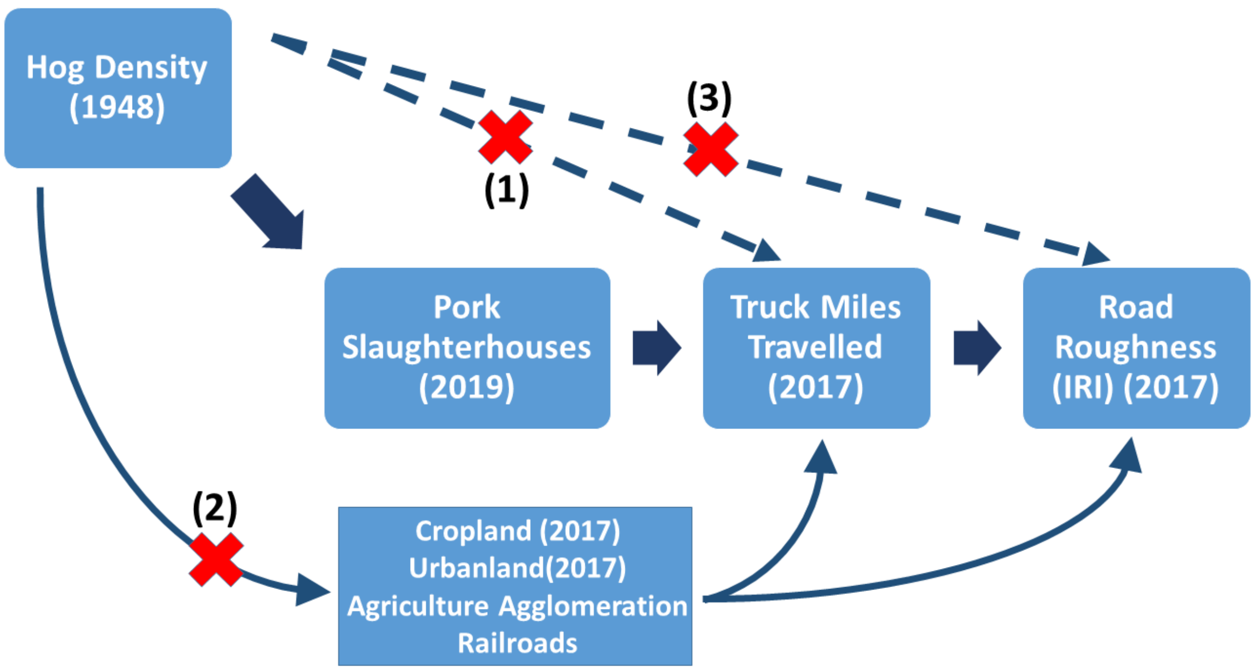
Note: Figure visually shows the first stage relationship between 1948 hogs and 2019 pork slaughterhouses at the CBSA level. An orange point represents one pork slaughterhouse in 2019. Darker shades of blue are associated with more hog points in Figure 1. Gray areas are not considered to be part of a metropolitan or micropolitan area.

Figure 3: Overlap of 1948 Hog Density and 2017 Truck Miles by Region, CBSA Level



Note: Figure visually shows the reduced form relationship between 1948 hogs and 2017 AADT of trucks aggregated to the CBSA level. Hog groups: $0 < 7.47$; $7.47 \leq 32.69$; > 32.69. Truck groups: $0 < 791$; ≥ 791. White areas are not in a CBSA.

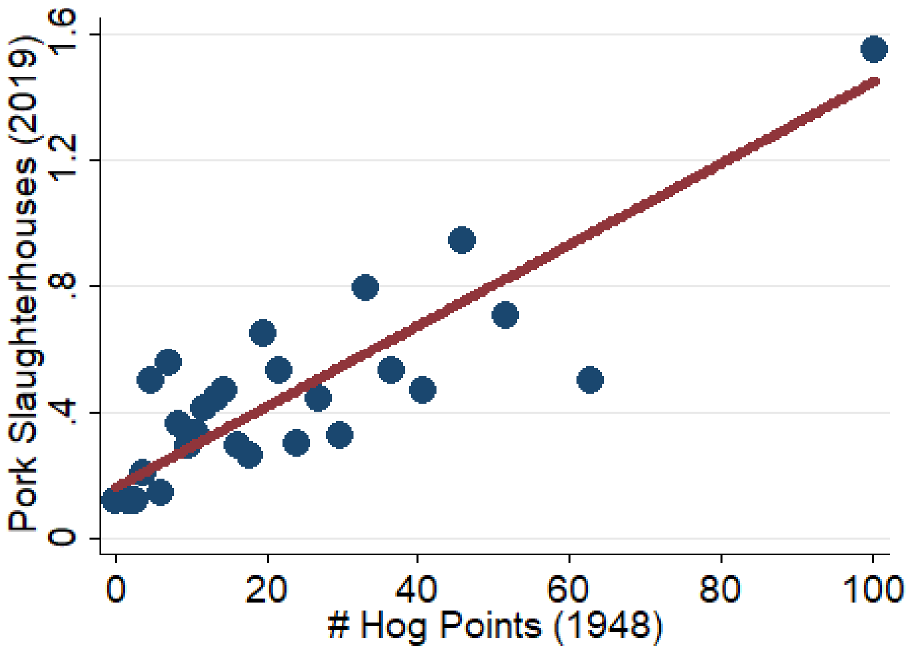
Figure 4: Identification Strategy: Directed Acyclic Graph Illustration



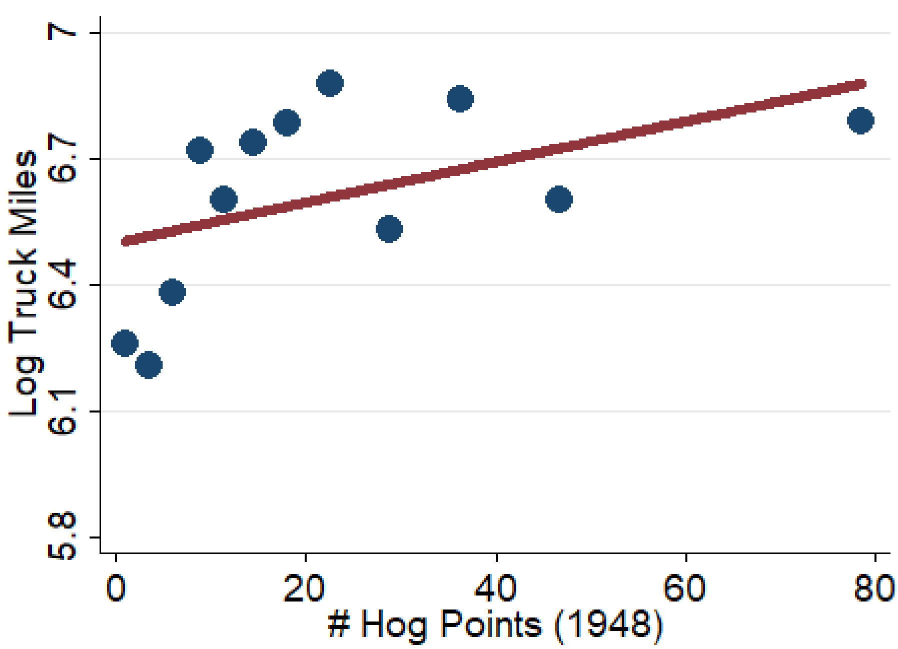
Notes: Figure shows a visual representation of the identification strategy using directed acyclic graphs (DAGs) (Pearl, 1998). Solid arrows represent observed data, dashed lines indicate unobserved data.

Figure 5: Non-Parametric Relationships Between Key Variables, CBSA Level

Panel A: First Stage: 1948 Hog Points on Current Pork Slaughterhouses



Panel B: Reduced Form: 1948 Hog Points on Log Truck Miles



Note: N = 911. Binsreg (Cattaneo, Crump, Farrell, & Feng, 2019) used. Procedure selects number of bins to optimize a bias-variance tradeoff. Equal number of observations in each bin. No covariates controlled for.

Table 2: OLS Estimates of Pork Slaughterhouses on Truck Vehicle Miles Travelled, CBSA

	(1)	(2)	(3)	(4)
Pork Slaughterhouses	0.03	0.10**	0.02	0.01
	(0.03)	(0.04)	(0.04)	(0.05)
Observations	911	911	911	911
State Fixed Effects		X	X	X
Standard Errors	Robust	State-Clustered	State-Clustered	State-Clustered
Controls Included			X	X
StateXMetro FE				X

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is logged AADT of trucks at the CBSA level. Column 1 employs heteroskedastic-robust standard errors. Column 2 uses state-clustered standard errors and adds state fixed effects. Column 3 adds controls for related industries, type of land cover, and non-passenger rail miles. Column 4 adds state by metro fixed effects.

Table 3: 2SLS Estimates of Slaughterhouses on Average Daily Truck Miles, CBSA Level

	(1)	(2)	(3)
	Reduced Form	First Stage	Second Stage
Dependent Variable	Truck Miles	Pork Slaughterh.	Truck Miles
Panel A: Single Independent Variable			
Hog Points in 1948	0.005***	0.013***	
	(0.001)	(0.002)	
Pork Slaughterhouses			0.375***
			(0.127)
Kleibergen-Paap F			28.00
Observations	911	911	911
Panel B: Metro FE, State FE, State Cluster			
Hog Points in 1948	0.003**	0.017***	
	(0.001)	(0.002)	
Pork Slaughterhouses			0.193**
			(0.087)
Kleibergen-Paap F			49.91
Observations	911	911	911
Panel C: State FE X Metro FE, State Cluster			
Hog Points in 1948	0.004**	0.017***	
	(0.002)	(0.002)	
Pork Slaughterhouses			0.219**
			(0.089)
Kleibergen-Paap F			50.58
Observations	911	911	911

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All panels display 2SLS regressions at the CBSA level. Columns 1, 2, and 3 display the reduced form, first stage, and second stage respectively. Dependent variable is logged AADT of trucks at the CBSA level. Panel A examines the impact of pork slaughterhouses on logged AADT of trucks. Panel B includes a dummy for CBSA type – metropolitan or micropolitan – and adds state fixed effects and state-clustered standard errors. Panel C examines the effect of pork slaughterhouses on logged AADT of trucks by adding a state by metro fixed effects and state-clustered standard errors.

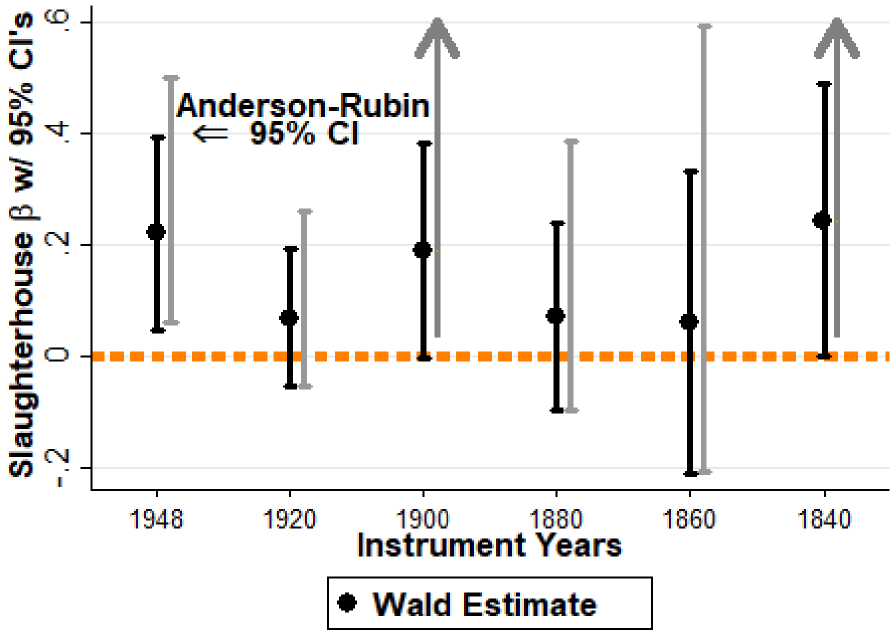
Table 4: 2SLS Estimates of Slaughterhouses on Average Daily Truck Miles, CBSA Level

Dependent Variable Panel A: Controls for # of Employees in Related Industries	(1) Reduced Form Truck Miles	(2) First Stage Pork Slaughterh.	(3) Second Stage Truck Miles
Hog Points in 1948	0.003 (0.002)	0.012*** (0.002)	
Pork Slaughterhouses			0.267* (0.161)
Kleibergen-Paap F			38.43
Observations	911	911	911
Panel B: Controls for (Satellite) Type of Land Cover Data			
Hog Points in 1948	0.003* (0.002)	0.012*** (0.002)	
Pork Slaughterhouses			0.293* (0.163)
Kleibergen-Paap F			34.93
Observations	911	911	911
Panel C: Controlling for Related Industries and Land Cover			
Hog Points in 1948	0.003 (0.002)	0.010*** (0.002)	
Pork Slaughterhouses			0.321* (0.193)
Kleibergen-Paap F			31.30
Observations	911	911	911
Panel D: Controlling for Non-Passenger Rail Miles			
Hog Points in 1948	0.003 (0.002)	0.010*** (0.002)	
Pork Slaughterhouses			0.292* (0.158)
Kleibergen-Paap F			20.92
Observations	911	911	911
Panel E: Controlling for Rail, Related Industry, and Land Cover			
Hog Points in 1948	0.003 (0.002)	0.009*** (0.002)	
Pork Slaughterhouses			0.362* (0.202)
Kleibergen-Paap F			19.12
Observations	911	911	911

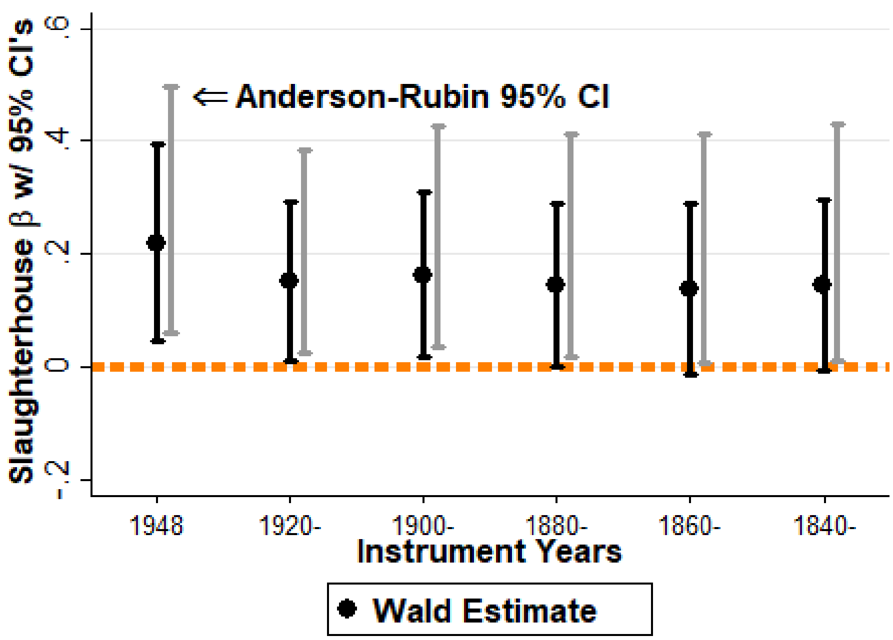
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. State-clustered standard errors in parentheses. All panels display 2SLS regressions. Columns 1, 2, and 3 display the reduced form, first stage, and second stage respectively. Dependent variable is logged AADT of trucks at the CBSA level. Each panel adds a new control variable. Panel A controls for the number of employees in related agricultural industries. Panel B controls for type of land cover (urban or crop land). Panel C controls for both related industries and type of land cover. Panel D controls for non-passenger rail miles. Finally, Panel E controls for related industries, land cover type, and non-passenger rail miles.

Figure 6: Treatment Effect by Year of Instrument(s)

Panel A: Just-Identified



Panel B: Hogs Summed Over Included Years, Just-Identified



Note: Estimates are β_1 from 2SLS regressions, according to Equations 5 and 6. The difference along the x-axis is which years of historical hog density are included in the instrument. In Panel A, only just-identified equations are used. In Panel B, the instruments in each year are added over indicated years. Arrows indicate no upper-limit on the AR Confidence intervals (Anderson, Rubin, et al., 1949).

Table 5: 2SLS Estimates of Slaughterhouses on Average Daily Car Miles, CBSA Level

	(1)	(2)	(3)
	Reduced Form	First Stage	Second Stage
Dependent Variable	Car Miles	Pork Slaughterh.	Car Miles
Panel A: Ln(Car Miles), Single Independent Variable			
Hog Points in 1948	0.000	0.013***	
	(0.001)	(0.002)	
Pork Slaughterhouses			0.033
			(0.083)
Kleibergen-Paap F			28.00
Observations	911	911	911
Panel B: Ln(Car Miles), StateXMetro FE, State Cluster			
Hog Points in 1948	0.002	0.017***	
	(0.001)	(0.002)	
Pork Slaughterhouses			0.096*
			(0.058)
Kleibergen-Paap F			50.58
Observations	911	911	911
Panel C: Ln(Car Miles), StateXMetro FE, State Cluster, Controls			
Hog Points in 1948	-0.000	0.009***	
	(0.001)	(0.002)	
Pork Slaughterhouses			-0.000
			(0.149)
Kleibergen-Paap F			19.12
Observations	911	911	911

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. All panels display 2SLS regressions at the CBSA level. Dependent variable is logged AADT of cars. Columns 1, 2, and 3 display the reduced form, first stage, and second stage respectively. Panel A examines the impact of pork slaughterhouses on logged AADT of cars. Panel B includes state by metro fixed effects and state-clustered standard errors. Panel C also includes state by metro fixed effects and state-clustered standard errors, but also adds controls for related industries, land cover type and non-passenger rail miles.

Table 6: 2SLS, Slaughterhouses on Truck Miles, by Slaughterhouse Slaughter Volume

Dependent Variable	(1) Reduced Form Truck Miles	(2) First Stage Pork Slaughterh.	(3) Second Stage Truck Miles
Panel A: Slaughter Volume = 1			
Hog Points in 1948	0.004**	0.004**	
	(0.002)	(0.002)	
Slaughterhouse (Volume = 1)			0.977
			(0.603)
Kleibergen-Paap F			5.054
Observations	911	911	911
AR Conf. Int. Set			[.189728, ...]
Ander-Rubin P			0.0135
Panel B: Slaughter Volume = 2			
Hog Points in 1948	0.004**	0.005***	
	(0.002)	(0.001)	
Slaughterhouse (Volume = 2)			0.757**
			(0.329)
Kleibergen-Paap F			20.97
Observations	911	911	911
AR Conf. Int. Set			[.197383, 1.91652]
Ander-Rubin P			0.0135
Panel C: Cumulative Additive Slaughter Volume			
Hog Points in 1948	0.004**	0.040***	
	(0.002)	(0.006)	
Discrete, Additive Slaughter Volume			0.096***
			(0.037)
Kleibergen-Paap F			45.17
Observations	911	911	911
AR Conf. Int. Set			[.027011, .213575]
Ander-Rubin P			0.0135

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. State-clustered standard errors in parentheses. All panels display 2SLS regressions at the CBSA level. Columns 1, 2, and 3 display the reduced form, first stage, and second stage respectively. Panel A examines the impact of pork slaughterhouses with a slaughter volume of "1" on AADT of trucks. Panel B examines the impact of pork slaughterhouses with a slaughter volume of "2" on AADT of trucks. Panel C sums all slaughter volumes of slaughterhouses in a given CBSA and examines the effect of cumulative slaughter volume on AADT of trucks. Anderson-Rubin confidence intervals and p-values are reported.

Table 7: Naive OLS of Trucks on Road Roughness

	(1)	(2)
Ln(Truck Miles)	-3.63***	-4.42***
	(0.91)	(0.91)
Observations	911	911
State Fixed Effects		X
Standard Errors	Robust	State-Clustered
Controls Included		X
StateXMetro FE		X

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Dependent variable is average International Roughness Index (IRI) measure at the CBSA level. Column 1 employs heteroskedastic-robust standard errors. Column 2 adds, in addition to state-clustered standard errors, state fixed effects, controls (related industries, land use and rail miles), and state by metro fixed effects.

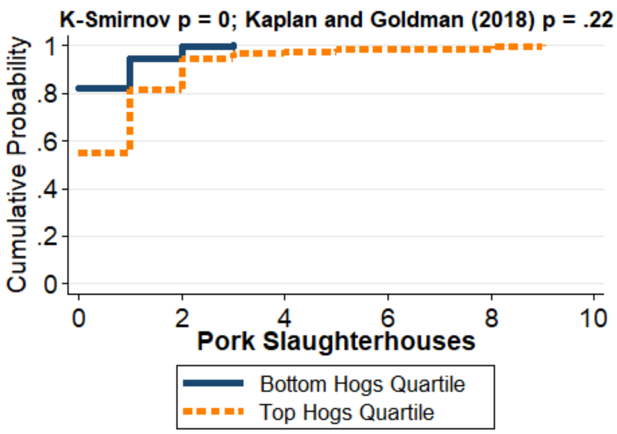
Table 8: 2SLS Estimates of Slaughterhouses and Truck Miles on Road Roughness, CBSA Level

	(1) Reduced Form	(2) First Stage	(3) Second Stage
Panel A: State FE's			
Hog Points in 1948	0.151***	0.020***	
	(0.042)	(0.003)	
Pork Slaughterhouses			7.709***
			(2.172)
Kleibergen-Paap F			51.47
Observations	911	911	911
Panel B: State FE, Metro FE			
Hog Points in 1948	0.054	0.017***	
	(0.033)	(0.003)	
Pork Slaughterhouses			3.070*
			(1.862)
Kleibergen-Paap F			33.49
Observations	911	911	911
Panel C: StateXMetro FE's			
Hog Points in 1948	0.039	0.017***	
	(0.035)	(0.003)	
Pork Slaughterhouses			2.228
			(1.896)
Kleibergen-Paap F			33.55
Observations	911	911	911
Panel D: State FE's, Truck Miles Ind. Var.			
Hog Points in 1948	0.151***	3.274*	
	(0.042)	(1.696)	
Truck Miles			0.046**
			(0.023)
Kleibergen-Paap F			3.729
Observations	911	911	911
AR Conf. Int. Set			[.019788, ...]
Ander-Rubin P			0.00301
Panel E: StateXMetro FE's, Truck Miles Ind. Var.			
Hog Points in 1948	0.039	1.700	
	(0.041)	(1.458)	
Truck Miles			0.023
			(0.030)
Kleibergen-Paap F			1.359
Observations	911	911	911
AR Conf. Int. Set			entire grid
Ander-Rubin P			0.319

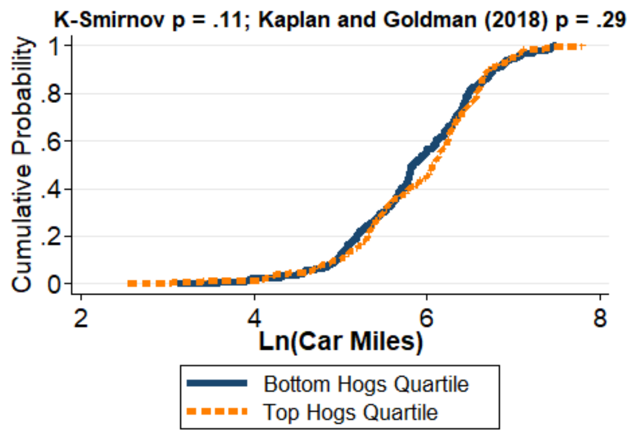
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. State-clustered standard errors in parentheses. All panels display 2SLS regressions at the CBSA level. Dependent variable is average International Roughness Index (IRI) at the CBSA level. Columns 1, 2, and 3 display the reduced form, first stage, and second stage respectively. Panel A examines the impact of pork slaughterhouses on IRI with state fixed effects. Panel B shows the impact of pork slaughterhouses on IRI with separate state and metro fixed effects. Panel C uses state by metro fixed effects. Panel D examines the impact of truck AADT on IRI with state fixed effects. Panel E uses state by metro fixed effects and truck AADT as the explanatory variable for IRI.

Figure 7: Cumulative Distribution Functions by Hog Quartile

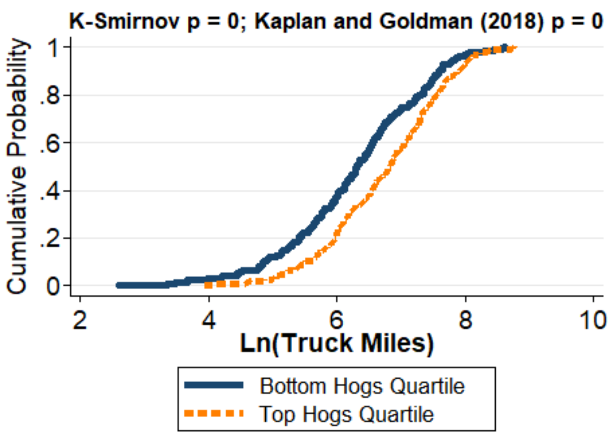
Panel A: First Stage: Pork Slaughterhouses by Hog Quartile



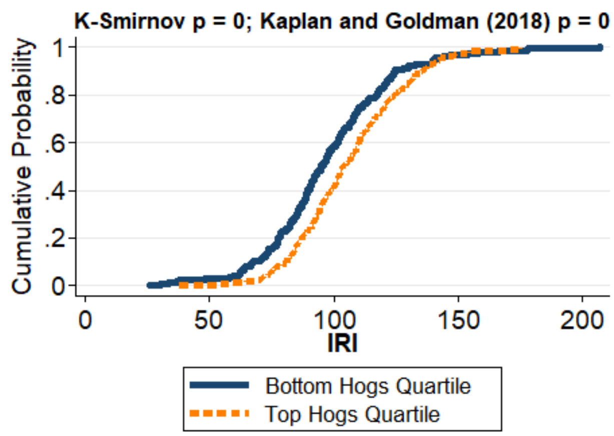
Panel B: Falsification: Log Car Miles by Hog Quartile



Panel C: Reduced Form: Log Truck Traffic By Hog Quartile



Panel D: Reduced Form: Road Roughness By Hog Quartile



Note: Kolmgorov-Smirnov and Goldman and Kaplan (2018) p-values of testing the null hypothesis that the distributions are equal. Bottom hogs quartile is the bottom quartile of hogs in 1948. Tops hogs is the top quartile of hogs in 1948. In Panel A, the distribution is of 2017 pork slaughterhouses. In Panel B, the distributions are of 2017 log car miles. In Panel C, the distributions are of 2017 log car miles. In Panel D, the distributions are of 2017 international (road) roughness index (IRI).

References

Abramzon, S., Samaras, C., Curtright, A., Litovitz, A., & Burger, N. (2014). Estimating the consumptive use costs of shale natural gas extraction on Pennsylvania roadways. *Journal of Infrastructure Systems*, *20*(3), 06014001.

Anderson, T., Rubin, H., et al. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics*, *20*(1), 46–63.

Aradom, S., & Gebresenbet, G. (2013). Vibration on animal transport vehicles and related animal behaviours with special focus on pigs. *Journal of Agricultural Science and Technology. A*, *3*(3A), 231.

Associated Press. (2017). *Commission delays decision on using river water in oilfields.* https://apnews.com/6e94ad14f7e042a5915ce5f64bcd8c09. (Accessed: June 2020)

Bai, Y., Oslund, P. C., Mulinazzi, T. E., Tamara, S., Liu, C., Barnaby, M. M., & Atkins, C. E. (2007). *Transportation logistics and economics of the processed meat and related industries in Southwest Kansas* (Tech. Rep.). University of Kansas Center for Research, Inc.

Banerjee, A., & Iyer, L. (2005). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review*, *95*(4), 1190-1213.

Bolt, J., Inklaar, R., de Jong, H., & van Zanden, J. L. (2018). *Real GDP per capita: 2011 US$.* (Maddison Project Database, version 2018, https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2018)

Bradley, J. H. (1948). *World Geography*. Boston: Ginn.

Brodersen, R. (2015). *One million pigs on the road every day.* https://www.aasv.org/shap/issues/v23n6/v23n6pm.html. (Accessed: June 2020)

Cattaneo, M. D., Crump, R. K., Farrell, M. H., & Feng, Y. (2019). On binscatter. *arXiv preprint arXiv:1902.09608*.

Chen, Q., Madson, D., Miller, C. L., & Harris, D. H. (2012). Vaccine development for protecting swine against influenza virus. *Animal Health Research Reviews*, *13*(2), 181–195.

Colomer, M. A., Margalida, A., & Fraile, L. (2020). Vaccination is a suitable tool in the control of aujeszky's disease outbreaks in pigs using a population dynamics P Systems Model. *Animals*, *10*(5), 909.

Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, *30*(4), 171–98.

Duranton, G., Nagpal, G., & Turner, M. (2020). Transportation infrastructure in the US. *NBER Working Paper Series*(w27254).

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616-52.

Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100*(3), 1195-1213.

Eschwege, H. (1979). *Truck weight and its effect on highways* (Tech. Rep. No. 10994). United States General Accounting Office.

Fitzgerald, R. F., Stalder, K. J., Matthews, J. O., Schultz Kaster, C. M., & Johnson, A. K. (2009). Factors associated with fatigued, injured, and dead pig frequency during transport and lairage at a commercial abattoir. *Journal of Animal Science*, *87*(3), 1156-1166.

Food and Agriculture Organization of the United Nations. (2001). *Guidelines for Humane Handling, Transport and Slaughter of Livestock.* http://www.fao.org/3/x6909e/x6909e00.htm. (Accessed: June 2020)

Gajana, C., Nkukwana, T., Marume, U., & Muchenje, V. (2013). Effects of transportation time, distance, stocking density, temperature and lairage time on incidences of pale soft exudative (pse) and the physico-chemical characteristics of pork. *Meat Science*, *95*(3), 520–525.

Glaeser, E. L. (2014). A world of cities: The causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association*, *12*(5), 1154–1199.

Glaeser, E. L., & Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, *47*(4), 983-1028.

Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *The Review of Economics and Statistics*, *97*(2), 498-520.

Goldman, M., & Kaplan, D. M. (2018). Comparing distributions by multiple testing across quantiles or cdf values. *Journal of Econometrics*, *206*(1), 143–166.

Gosálvez, L., Averós, X., Valdelvira, J., & Herranz, A. (2006). Influence of season, distance and mixed loads on the physical and carcass integrity of pigs transported to slaughter. *Meat Science*, *73*(4), 553 - 558.

Guàrdia, M., Estany, J., Balasch, S., Oliver, M., Gispert, M., & Diestre, A. (2005). Risk assessment of DFD meat due to pre-slaughter conditions in pigs. *Meat Science*, *70*(4), 709 - 716.

Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, *108*, 36–50.

Hanlon, W. W. (2016). Coal smoke and the costs of the industrial revolution. *NBER Working Paper*(w22921).

Hellerstein, E., & Fine, K. (2017). *A million tons of feces and an unbearable stench: Life near industrial pig farms.* https://www.theguardian.com/us-news/2017/sep/20/north-carolina-hog-industry-pig-farms. (Accessed: June 2020)

Highway Research Board. (1962). *The AASHTO road test: Pavement research* (Tech. Rep. No. HRB Special Report 61E). Washington, DC.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, *39*(153), 41–57.

Howarth, R. W., Ingraffea, A., & Engelder, T. (2011). Should fracking stop? *Nature*, *477*(7364), 271–275.

Huynh, T., Aarnink, A., & Verstegen, M. (2005). Reactions of pigs to a hot environment. In *Livestock environment VII, 18-20 May 2005, Beijing, China* (p. 544).

Johnson, A. K., Gesing, L. M., Ellis, M., McGlone, J., Berg, E., Lonergan, S. M., ... others (2013). 2011 and 2012 early careers achievement awards: Farm and pig factors affecting welfare during the marketing process. *Journal of Animal Science*, *91*(6), 2481–2491.

Kaldor, N. (1967). *Strategic factors in economic development*. New York State School of Industrial and Labor Relations.

Kemp, B., Da Silva, C. L. A., & Soede, N. M. (2018). Recent advances in pig reproduction: Focus on impact of genetic selection for female fertility. *Reproduction in Domestic Animals*, *53*(S2), 28-36.

Key, N. (2004). Agricultural contracting and the scale of production. *Agricultural and Resource Economics Review*, *33*(2), 255.

Leon, D. A. (2008). *Cities, urbanization and health.* Oxford University Press.

Lochner, L., & Moretti, E. (2015). Estimating and testing models with many treatment levels and limited instruments. *Review of Economics and Statistics*, *97*(2), 387–397.

Luechinger, S., & Roth, F. (2016). Effects of a mileage tax for trucks. *Journal of Urban Economics*, *92*, 1 - 15.

MacDonald, J., Ollinger, M., Nelson, K., & Handy, C. (2000). Consolidation in US meatpacking. *USDA Agricultural Economic Report*(785).

Mandell, S., & Proost, S. (2016). Why truck distance taxes are contagious and drive fuel taxes to the bottom. *Journal of Urban Economics*, *93*, 1 - 17.

Marshall, A. (1890). *Principles of Microeconomics*. London: Macmillan.

McArthur, J. W., & McCord, G. C. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of Development Economics*, *127*, 133–152.

McBride, W. D., & Key, N. (2013). US hog production from 1992 to 2009: Technology, restructuring, and productivity growth. *USDA Economic Research Report*(158).

Miranda-De La Lama, G., Villarroel, M., & María, G. A. (2014). Livestock transport from the perspective of the pre-slaughter logistic chain: A review. *Meat Science*, *98*(1), 9–20.

Newbery, D. M. (1988). Road damage externalities and road user charges. *Econometrica: Journal of the Econometric Society*, 295–316.

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, *101*(7), 3221-52.

Parry, I. W. (2008). How should heavy-duty trucks be taxed? *Journal of Urban Economics*, *63*(2), 651 - 668.

Pearl, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods & Research*, *27*(2), 226–284.

Pfeiffer, L., & Lin, C.-Y. C. (2012). Groundwater pumping and spatial externalities in agriculture. *Journal of Environmental Economics and Management*, *64*(1), 16–30.

Picard, P. M., & Zeng, D.-Z. (2005). Agricultural sector and industrial agglomeration. *Journal of Development Economics*, *77*(1), 75–106.

Pig Progress. (2015). *Preventing 5 major pig diseases in a breeding herd.* https://www.pigprogress.net/Health/Articles/2015/10/Preventing-five-major-pig-diseases-in-the-breeding-herd-2685176W/. (Accessed:

June 2020)

Pork Checkoff. (2015). *Quick facts: The pork industry at a glance.* http://porkgateway.org/wp-content/uploads/2015/07/quick-facts-book1.pdf. (Accessed: June 2020)

Porter, M. (2003). The economic performance of regions. *Regional Studies*, *37*(6-7), 549–578.

Pretty, J. N., Brett, C., Gee, D., Hine, R., Mason, C., Morison, J., ... van der Bijl, G. (2000). An assessment of the total external costs of UK agriculture. *Agricultural Systems*, *65*(2), 113–136.

Randall, J. M., & Bradshaw, R. H. (1998). Vehicle motion and motion sickness in pigs. *Animal Science*, *66*(1), 239–245.

Rhodes, V. J. (1995). The industrialization of hog production. *Review of Agricultural Economics*, 107–118.

Roe, B., Irwin, E. G., & Sharp, J. S. (2002). Pigs in space: Modeling the spatial structure of hog production in traditional and nontraditional production regions. *American Journal of Agricultural Economics*, *84*(2), 259–278.

The Allegheny Front. (2017). *On health effects, blame the trucks, not the fracking?* https://www.alleghenyfront.org/on-health-effects-blame-the-trucks-not-the-fracking/. (Accessed: June 2020)

The Texas Tribune. (2014). *Dissecting Denton: How one city banned fracking.* https://www.texastribune.org/2014/12/15/dissecting-denton-how-texas-city-baned-fracking/. (Accessed: June 2020)

Turner, M. A., Haughwout, A., & Van Der Klaauw, W. (2014). Land use regulation and welfare. *Econometrica*, *82*(4), 1341–1403.

United States President's Commission on Law Enforcement & Administration of Justice, & Katzenbach, N. d. (1967). *The challenge of crime in a free society: A report by the President's Commission on Law Enforcement and Administration of Justice*. United States Government Printing Office.

Wang, Y.-F., Huang, J.-J., & Zhao, J.-G. (2017). Gene engineering in swine for agriculture. *Journal of Integrative Agriculture*, *16*(12), 2792–2804.

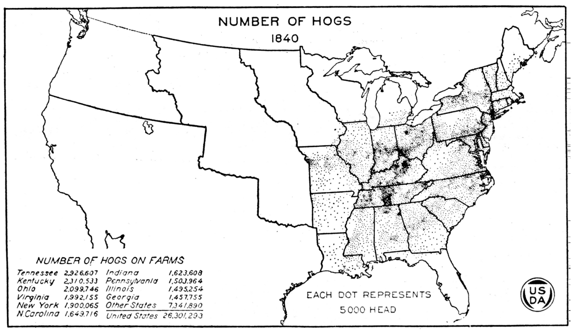
Xu, M., & Xu, Y. (2020). Fraccidents: The impact of fracking on road traffic deaths. *Journal of*

Environmental Economics and Management, *101*, 102303.

A Appendix

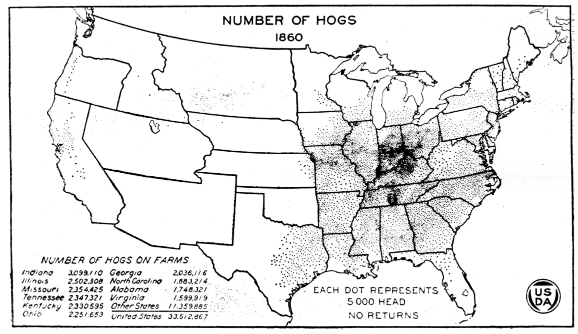
Figure A.1: Hog Density: 1840-1920

Panel A: 1840 Hog Points



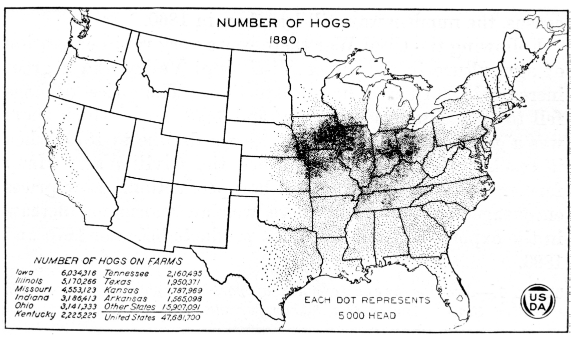
United States Production of Hogs, 1840
United States Department of Agriculture Yearbook 1922, Washington D.C.: Government Printing Office, 1922
Downloaded from *Maps ETC*, on the web at http://etc.usf.edu/maps [map #00157]

Panel B: 1860 Hog Points



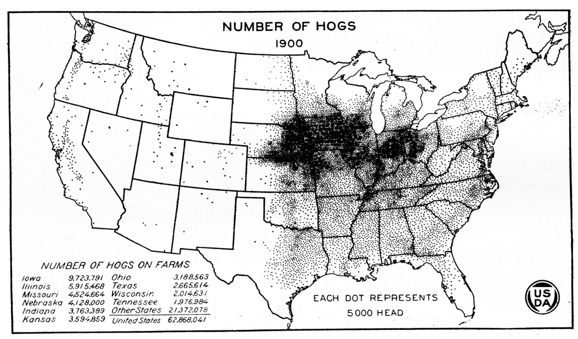
United States Production of Hogs, 1860
United States Department of Agriculture Yearbook 1922, Washington D.C.: Government Printing Office, 1922
Downloaded from *Maps ETC*, on the web at http://etc.usf.edu/maps [map #00158]

Panel C: 1880 Hog Points



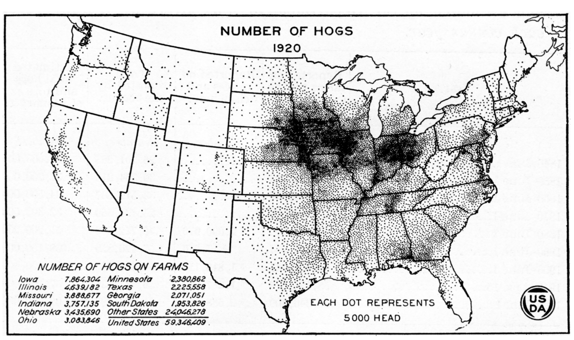
United States Production of Hogs, 1880
United States Department of Agriculture Yearbook 1922, Washington D.C.: Government Printing Office, 1922
Downloaded from *Maps ETC*, on the web at http://etc.usf.edu/maps [map #00159]

Panel D: 1900 Hog Points



United States Production of Hogs, 1900
United States Department of Agriculture Yearbook 1922, Washington D.C.: Government Printing Office, 1922
Downloaded from *Maps ETC*, on the web at http://etc.usf.edu/maps [map #00160]

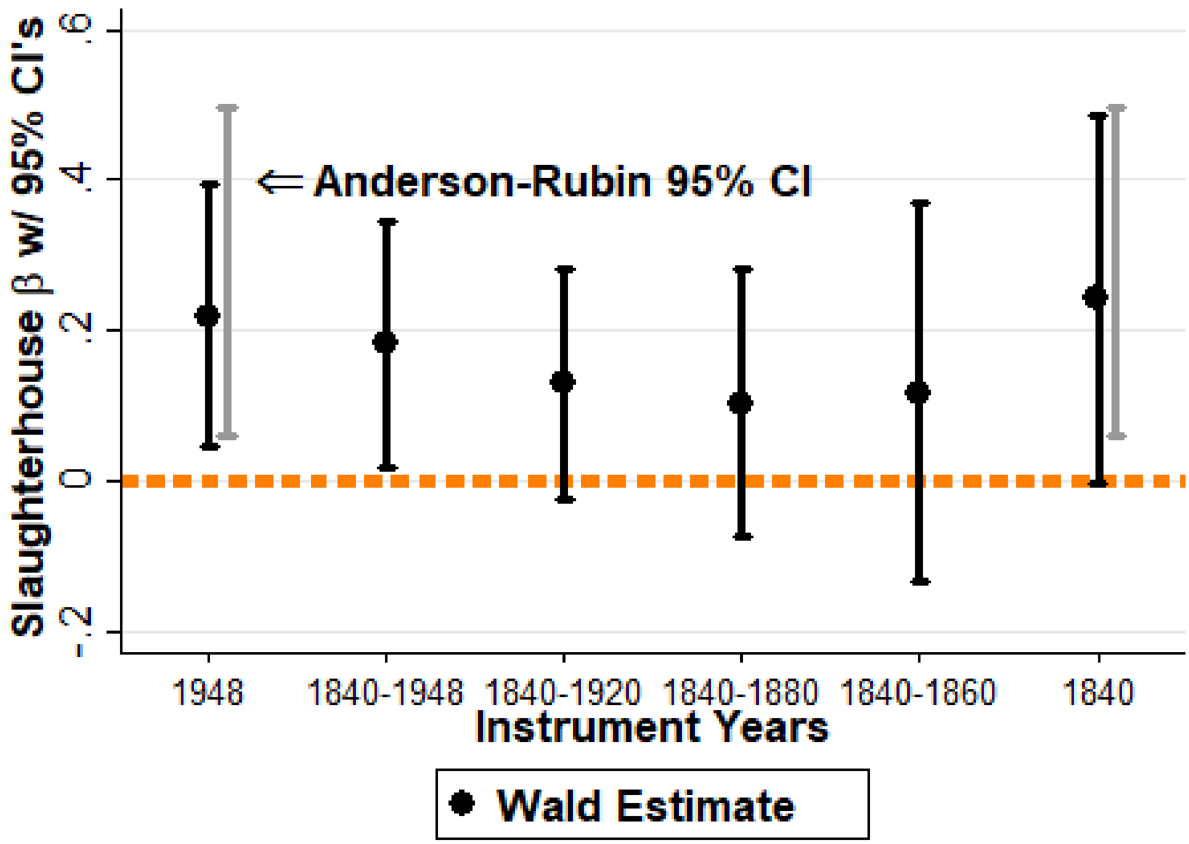
Panel E: 1920 Hog Points



United States Production of Hogs, 1920
United States Department of Agriculture Yearbook 1922, Washington D.C.: Government Printing Office, 1922
Downloaded from *Maps ETC*, on the web at http://etc.usf.edu/maps [map #00161]

Note: Maps of historic hog densities in the United States are shown in the Panels above. Each dot represents 5000 hogs. Maps are from the 1922 US Department of Agriculture Yearbook.

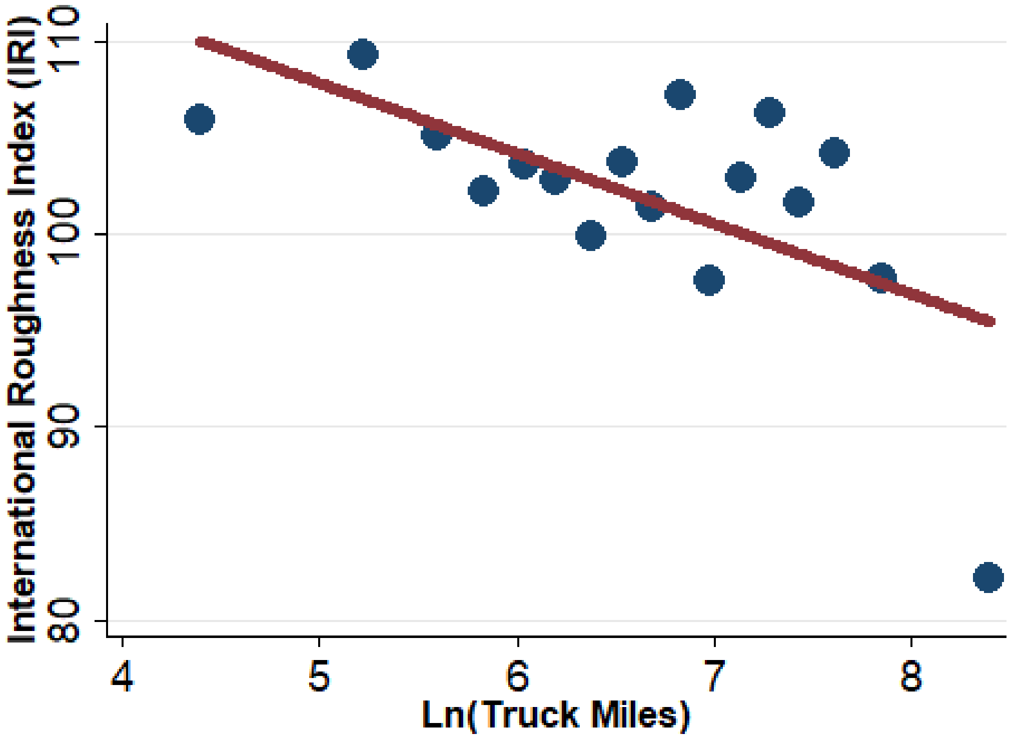
Figure A.2: Over-Identified



Note:

Figure A.3: Naive Binned Scatter and

Panel A: Trucks Negatively Associated with Road Roughness



Note: