COLUMBIA | SIPA
Center for Environmental Economics and Policy

CEEP WORKING PAPER SERIES
WORKING PAPER NUMBER 13

NOVEMBER 2020

Livestock Plants and COVID-19 Transmission

Charles A. Taylor, Christopher Boulos, and Douglas Almond

https://ceep.columbia.edu/sites/default/files/content/papers/n13.pdf

Livestock Plants and COVID-19 Transmission*

Charles A Taylor[1], Christopher Boulos, and Douglas Almond[1,2]

[1] *School of International and Public Affairs, Columbia University*

[2] *Department of Economics, Columbia University*

November 2020

Abstract

Policy responses to the 2019 novel coronavirus (COVID-19) outbreak must strike a balance between maintaining essential supply chains and limiting the spread of the virus. Our results indicate a strong positive relationship between livestock processing plants and local community transmission of COVID-19, suggesting that these plants may act as transmission vectors into the surrounding population and accelerate the spread of the virus beyond what would be predicted solely by population risk characteristics. We estimate the total excess COVID-19 cases and deaths associated with proximity to livestock plants to be 236,000–310,000 (6–8% of all US cases) and 4,300–5,200 (3–4% of all US deaths) as of July 21, with the vast majority likely related to community spread outside these plants. The association is found primarily among large processing facilities and large meatpacking companies. In addition, we find evidence that plant closures attenuated county-wide cases and that plants that received permission from the USDA to increase their production line speeds saw more county-wide cases. Ensuring both public health and robust essential supply chains may require an increase in meatpacking oversight and potentially a shift toward more decentralized, smaller-scale meat production. *Keywords: COVID-19, Supply chains, Livestock, Agriculture, Public health*

*Preprint submitted to PNAS: not final version

1 Introduction

Among the many challenges posed by the COVID-19 outbreak, maintaining essential supply chains, while mitigating community spread of the virus, is vital to society. Using county-level data as of July 21, 2020, we test the relationship between one such type of essential activity, livestock processing, and the local incidence of COVID-19 cases. We find that the presence of a slaughtering plant in a county is associated with 4–6 additional COVID-19 cases per thousand, or a 51–75% increase from the baseline rate. We also find an increase in the death rate by 0.07–0.1 deaths per thousand people, or 37–50% over the baseline rate. Our estimates imply that excess COVID-19 infections and deaths related to livestock plants are 236,000–310,000 (6–8% of all US cases) and 4,300–5,200 (3–4% of all US deaths), respectively, with the vast majority occurring among people *not* working at livestock plants.

We further find the temporary closure of high-risk plants to be followed by lower rates of COVID-19 case growth. We also find that smaller, decentralized facilities do not appear to contribute to transmission, and that plants that received permission from the USDA to increase their production line speeds saw more county-wide cases. Our associations hold after controlling for population risk factors and other potential confounders, such as testing rates. Although lacking a natural experiment to cement causality, we employ a combination of empirical tools—including an event study, instrumental variables, and matching—to support our findings.

The centrality of livestock processing to local economies and national food supplies implies that mitigating disease spread may take an economic toll. Understanding the public health risk posed by livestock processing is essential for assessing potential impacts of policy action. However, generating case data attributable to livestock plants is challenging: contact tracing in the US is decentralized and sporadic, and there may be incentives for companies and government bodies to obscure case reporting (Subbian et al. 2020; Foley 2020; Leah

Douglas 2020; Novack and Sophie 2020; Mayer 2020). Our study represents an attempt to address this gap in knowledge.

1.1 Heterogeneity in COVID-19 patterns

The disease burden of COVID-19 is not uniformly distributed across the global population. Certain conditions appear to influence the degree to which people spread the virus. Some contexts and social behaviors are believed to lead to superspreading events that disproportionately affect local populations (Hamner 2020; Bouffanais and Lim 2020). Previous studies have explored links between the incidence of COVID-19 cases and a range of demographic and environmental factors, such as age, occupation, income, race, intergenerational mixing, temperature, and humidity (Dowd et al. 2020; Barbieri et al. 2020; Borjas 2020; Sajadi et al. 2020; O'Reilly et al. 2020; National Academies of Sciences and Medicine 2020). Social, commercial, and industrial activities are also believed to affect transmission, for which reason countries worldwide have implemented a range of economic and social distancing measures (Dowd et al. 2020; Anderson et al. 2020; Ebrahim et al. 2020; Lewnard and Lo 2020; Moghadas et al. 2020; Kraemer et al. 2020; Wells et al. 2020; Roser et al. 2020). In the US, some industries are exempted from shelter-in-place orders and have remained operational due to their necessity to meet basic societal needs (US Department of Homeland Security 2020). We investigate the relationship between transmission and one such activity, livestock processing.

1.2 COVID-19 and livestock plants

The livestock and poultry processing industry is an essential component of the global food supply chain. In the US, it is a large industry, employing 500,000 people. It is also highly concentrated: the largest four companies in beef, pork, and poultry processing capture 55–85% of their respective markets (Dyal 2020; MacDonald and McBride 2009; Hendrickson et al. 2017; Inc. 2020; Wohlgenant 2013; Ward 2002). This degree of concentration stands in contrast to the European Union (EU), for example, where the top 15 meat companies represent 28% of the region's meat production (European of Food, Agriculture, and Tourism Trade Unions 2011).

Over the decades, the livestock and poultry processing industry in the US has consolidated its operations into fewer, larger plants, in which meat production per plant has increased threefold since 1976 (Hendrickson 2015; Skerritt et al. 2020). Today, 12 plants produce over 50% of the country's beef, and 12 others similarly produce over 50% of the country's pork (2020; National Pork Board 2019). Early in the COVID-19 pandemic, livestock processing plants worldwide experienced spikes in infections, facing shutdowns that disrupted meat and dairy supplies (Mano 2020; Busvine 2020; Scott and Chandler 2020; Hirtzer and Freitas 2020). In the US, reports of COVID-19 spreading within the livestock processing industry led to increased attention and updated safety guidance by the CDC (Dyal 2020). Several plants were forced to shut down until, among other factors, a federal executive order invoked the status of livestock processing as 'critical infrastructure' for national security and mandated that these plants remain open (Mason and Polansek 2020; Order 2020).

Work routines in livestock processing have several characteristics that make plants susceptible to local outbreaks of respiratory viruses. The CDC includes among potential risk factors: long work shifts in close proximity to coworkers, difficulty in maintaining proper face covering due to physical demands, and shared transportation among workers (Dyal 2020). Previous research has proposed occupational exposure to livestock animals as a driver of viral spread, although an experimental study did not find pigs or chickens to be susceptible to the SARS-CoV-2 virus associated with COVID-19 (VanderWaal and Deen 2018; Myers et al. 2006; Schlottau et al. 2020; Yang et al. 2020).

Increases in production line speeds due to technological enhancements as well as policy changes have also been hypothesized to exacerbate COVID-19 transmission (Thompson and Berkowitz 2020; Mayer 2020). Among those we investigate are USDA waivers on poultry-production line-speed limits for plants with strong commercial production practices and microbial control (US Department of Agriculture 2018).[1]

The indoor climate of livestock facilities may increase transmission risk. To preserve meat after slaughter, processing areas are maintained at 0-12 °C (Cano-Muñoz and Muñoz 1991), and such low temperatures have been linked to increased COVID-19 risk (Carleton and Meng 2020; Zuber and Brüssow 2020). Though these rooms are kept at 90-95% relative humidity to prevent meat from drying and losing weight, the low absolute humidity at near-freezing temperatures may encourage the transmission of airborne viruses such as influenza (Shaman and Kohn 2009; Koep et al. 2013; Deyle et al. 2016). Moreover, studies have suggested that industrial HVAC systems used to cool and ventilate meat storage and processing facilities may further the spread of pathogenic bioaerosols, a proposed COVID-19 transmission route (Beck et al. 2019; Asadi et al. 2020; Mittal et al. 2020; Borak 2020; Zuber and Brüssow 2020).

Workers' socioeconomic status and labor practices may also contribute to infection and transmission. Among frontline meat processing workers in the US, 45% are categorized as low income, 80% are people of color, and 52% are immigrants, many of whom are undocumented and lack ready access to healthcare and other worker protections that could facilitate COVID-19 prevention and treatment (Fremstad et al. 2020; Compa 2004; Kandel and Parrado 2005). In addition, employees at these facilities may face incentives to continue working even while sick through company policies on medical leave and attendance bonuses (Dyal 2020; Mayer 2020; Grabell 2020). In addition, through consolidation over the decades, the meatpacking industry has potentially increased its monopsonistic power

[1] The CEOs of Wayne Farms and Tyson Foods—both granted waivers in April 2020—are, respectively, Chairman of the National Chicken Council (the body that initially lobbied for the line speed waivers) and a public advocate for the poultry industry, buying full-page newspaper ads in April stating that the food supply chain was 'broken.'

over labor markets, which has been linked to greater work hazards (Constance et al. 2013; MacDonald et al. 2000; Viscusi 1980).

2 Materials and Methods

Our analysis used a county-level dataset of COVID-19 cases and deaths from the New York Times, based on reports from state and local health agencies (The New York Times 2020). Included in counts are both confirmed and probable deaths, as categorized by states. The five county boroughs of New York City were grouped into one unit. We limited the analysis to the continental US. Our baseline model specification takes the following form:

$$outcome_i = \beta * livestock_i + \theta * controls_i + \alpha_s + \epsilon_i \tag{1}$$

where $outcome_i$ is the COVID-19 case or death rate in county i, β is the coefficient of interest, $controls_i$ is a vector of county-level covariates, α_s is a dummy for fixed effects in state s, and ϵ_i is the error term.

Covariate data include county-level race, ethnicity and age structure data from the US Census and mean county-level income data from the US Bureau of Economic Analysis (SEER Program, National Cancer Institute, NIH 2020; US Bureau of Economic Analysis 2020). Data on nursing home populations, incarcerated populations, uninsured populations, average household size, and work commuting methods come from the 2014-2018 American Community Survey (US Census Bureau 2019c; 2019e; 2019b; 2019d). Data on manufacturing establishments come from the American Economic Survey (2019a). Data on the number of frontline workers are derived from CEPR data (Fremstad et al. 2020), transforming from the Public Use Microdata Area level to the county level assuming even allocation. The freight index is from the FHA's Freight Analysis Framework (US Department of Transportation 2020) using the variable AADTT12, the annual average daily truck traffic in 2012, which we sum across all listed highways in a given county. Data on state-level social distancing policy come from a dataset synthesizing news articles tracking these policy measures (NASHP Staff 2020; Gershman 2020; Lee et al. 2020).

Locations and characteristics of livestock processing facilities come from the USDA's Food Safety and Inspection Service (FSIS) (USDA Food Safety and Inspection Service 2020a). Beef and pork livestock plants were filtered to include plants with volume of all processed products greater than one million pounds per month (Categories 4 and 5), which account for the vast majority of US production. Poultry livestock were filtered to include plants with volumes greater than 10 million pounds per month (Category 5) because that category alone accounts for the vast majority of US production. County-level mobility data were made accessible to COVID-19 researchers by Google (Google LLC 2020). County-level COVID-19 testing data come from a dataset gathered from 31 state health agencies. Data on line speed waivers come from the USDA FSIS (USDA Food Safety and Inspection Service 2020b). Data on plant closures and opening dates come from a novel dataset assembled from various local news reports, building on a dataset from the Midwest Center for Investigative Reporting (Chadde 2020). Historical livestock production data are from the 1959 USDA census of agriculture, accessed via the Inter-University Consortium for Political and Social Research (Haines et al. 2018).

3 Results

We find a strong relationship between proximity to livestock plants and the incidence of COVID-19 over time. Fig. 1 plots average COVID-19 case and death rates over time by whether there is a large livestock facility in a given county relative to rates in counties at varying distances from a plant. In both cases, we see an increasing divergence in outcomes beginning in early April based on livestock-plant proximity.

Fig. 1 does not account for county-level differences in terms of density and demographics. In Table 1, we estimate the relationship between livestock plants and COVID-19 incidence as of July 21, 2020 using regression models that control for potential confounding variables, including county-level measures of income; population density and its square; the timing of the first case; the proportions of elderly people; uninsured people, frontline workers, and people using public transportation; racial and ethnic characteristics; average household size; local freight traffic; and populations of nursing homes and prisons. We find that livestock plants are associated with an increase in COVID-19 cases by approximately 4 per thousand people, representing a 51% increase over the July 21 baseline rate of 8 per thousand. Likewise, death rates increase by 0.07 per thousand, or 37% over the county baseline of 0.2 deaths per thousand. The results are robust both nationally and when only considering variation within states after including state fixed effects. We also use an alternate specification with a binary measure of whether a county has one or more livestock plants. Such counties are associated with 6 additional cases per thousand, or a 75% increase over the baseline, as well as 0.1 additional deaths per thousand, or 50% over the baseline county death rate.[2]. In addition, COVID-19 appears to arrive earlier in counties with livestock plants (Table 3).

3.1 Heterogeneity by facility type, size, operations, and company

We now present potential characteristics of livestock facilities that might contribute to these observed relationships with the COVID-19 case and death rate.

[2] In line with the literature, we find COVID-19 incidence to be strongly associated with population density, average household size, the timing of the first confirmed case, and the proportion of a county's population who are public transit commuters, elderly, Black, Hispanic, in a nursing home, uninsured, or institutionalized (Fig. 3

3.1.1 Facility type

We first looked at the relationship between reported cases and the type of animal slaughtered or processed. We found that beef, pork, and poultry plants each show a significant relationship with COVID-19 cases and deaths, with pork plants showing the greatest measured magnitude of the three in cases and beef plants showing the greatest magnitude in deaths (Table 4). As seen in the map in Fig. 2, pork and beef plants are well distributed throughout the US, and although poultry plants are relatively concentrated in the Southeast US, they are found across 10 states. Overall, the wide geographic distribution of facilities by type mitigates concerns of this being a regional phenomenon.

3.1.2 Facility size

We next investigated whether there are differential relationships with COVID-19 transmission based on the size of processing facilities. Livestock facility data are gathered from the USDA FSIS. Table 5 categorizes beef, pork, and poultry plants by order of magnitude based on the pounds per month processed: large (Category 5; over 10 million), medium (Category 4; over 1 million), and small (Category 3; over 100,000 and under 1 million). Each size category was sufficiently represented, with 349 small plants, 126 medium plants, and 225 large plants. Very small plants (Categories 1 and 2), which are often niche providers, were excluded.[3]

We found the relationship between livestock plants and COVID-19 transmission to be most pronounced among the largest plants, whose presence in a county is associated with a 35% higher COVID-19 case rate relative to the average coefficient for livestock plants shown in Table 1. Small and medium-sized plants were generally not found to have significant relationships with local COVID-19 transmission, suggesting that the scale of produc-

[3] In our main analyses, we include Category 4 and 5 pork and beef facilities and Category 5 poultry facilities (which comprise 57% of total poultry plants); see the Materials and Methods section for a full discussion.

tion is an important variable for industry leaders and policymakers to consider.

3.1.3 Production-line speeds

We next examined whether there is a relationship between local COVID-19 transmission and plant operating procedures. We collected data on whether a poultry plant had been granted a waiver from the USDA permitting production line processing speeds of 175 birds per minute, up from the statutory limit of 140. Waivers were first issued to 20 poultry plants in 2012 as part of a pilot study to test self-monitoring of safety. It was then expanded in September 2018 to allow all poultry plants the opportunity to apply for these waivers. A faster production line can result in both workers spending their days in closer quarters and facing greater difficulty in maintaining PPE compliance. These factors may increase the likelihood of viral transmission.

Out of the 120 poultry plants in our sample, 48 plants currently have waivers, 16 of which were issued in 2020.[4] An analysis of the relationship between line speed waivers and local COVID-19 incidence suggests, though with less precise estimates, that waivers predict increases in county-level case rates double those in counties with non-waiver poultry plants (Table 6). Among plants issued a waiver in 2020, the relationship is even greater in magnitude. This finding suggests a potential pathway between a livestock plant's operating procedures and COVID-19 transmission.

3.1.4 Facility operator

We next looked at differential relationships with COVID-19 by company. The relationship between local COVID-19 incidence and medium and large plants (FSIS Categories 4 or 5) owned or operated by some of the largest US processors (National Beef, JBS, Tyson Foods, Cargill, Smithfield) and their subsidiaries is presented in Table 7. These magni-

[4] Among counties with poultry plants, those with and without waivers appear similar in their average characteristics, reducing waiver selection concerns. The exception is that waivers counties have lower proportions of Black residents and prison populations, factors associated with increased COVID-19 risk.

tudes are shown in Fig. 4: the strongest relationship is found with National Beef, whose indicated relationship with COVID-19 case rates is approximately five times greater in magnitude than that of other livestock facilities. However, all of the large companies appear to have larger coefficients than the baseline. Aside from Smithfield, the relationship with deaths is positive, albeit less significant, which may be due to small sample size.[5]

3.2 Behavioral change

If livestock facilities are driving higher COVID-19 incidence, and if livestock processing is an essential industry, we would expect people in livestock plant counties to work more compared to those in non-livestock counties in response to lockdowns related to COVID-19. To this end, we employed county-level mobility data made available by Google for COVID-19 researchers. We constructed a baseline measure of average time-use change before and after March 13, 2020, the date the US declared a national disaster in relation to COVID-19 and two days after the WHO declared COVID-19 a pandemic.

We then examined how the presence of livestock plants varied with time spent working and engaging in shopping and recreation. We controlled for the same demographic and location-based covariates as in other models. We found that the presence of livestock plants is strongly associated with more time spent at work (Table 8). This association is relative to the baseline behavior change across all other counties, indicating that people in livestock plant counties are working more (or cutting back on work less) than people in other counties. Meanwhile, there is a lesser and imprecise relationship for retail and recreation activities, which may contribute to viral spread. This finding supports the notion that livestock plants, rather than unrelated changes in behavior in these same counties, are the more likely vehicle of COVID-19 transmission.[6]

[5] In our collected sample, the number of facilities per company varies: National Beef has only seven plants in seven counties, whereas Tyson Foods has 80 plants across 69 counties. The other companies fall somewhere in between.

[6] It is possible that additional time spent working, and thus out of the house, may explain some of the additional time spent on retail activities (e.g., gas stations or workday meals).

3.3 Plant shutdowns

Many livestock plants were temporarily shut down to halt the spread of COVID-19. In such cases, we would expect the dynamics of caseloads and deaths over time to vary negatively with the timing of shutdown, after a lag. Were confounders instead driving our results, they would have to follow the timing of the plant shutdowns as well. This helps argue against purely static confounders, such as highway connectedness or fraction of the population that is Hispanic.

Using a dataset tracking whether and when livestock plants closed, Fig. 5 presents an event study comparing the change in weekly COVID-19 case rates before and after closure, averaged across counties with plants that closed and counties with plants with no evidence of closure. Among livestock plants in our sample, we have the dates of closures that occurred in 26 counties, or 10% of counties with plants. The mean closure time was nine days. Some closed for a day or two for cleaning and disinfection, while others closed for longer periods while revising their operating procedures and monitoring staff. On the other hand, many plants remained open due to a perceived lack of risk, and others remained open despite significant local outbreaks.

In this event study, we examined case growth (weekly log difference), following the structure of previous analysis (Hsiang et al. 2020), as well as change in case rates. In addition, we performed pre-policy matching across the two groups based on percent case growth in the two weeks prior to shutdown. In doing so, we selected the top quartile of growth rates among the 233 counties with livestock plants that did not have a plant shutdown. We took this step to maximize comparability between the two groups, as we observed that preclosure growth in cases was, on the whole, greater in plants that closed (Fig. 6).

Coefficients are plotted from a panel regression, where counties (categorized as either having or not having a plant closure) are interacted with the weekly event index, both in terms of percent growth in cases (Fig. 5 A and C) and the change in case rates per 1000 (Fig.

5 B and D). This model controls for state-level social distancing and stay-at-home policy and includes a fixed effect for each county, thereby isolating within-county variation in timing (among counties with plant closures).

Fig. 5 shows that plant closures occurred in counties experiencing high growth in COVID-19 cases, as might be expected. Within one week of closure, however, the growth rate in shutdown counties reverted to the pre-policy growth rate from a higher peak compared to non-shutdown counties in the same time. By week 2, growth rates between the two categories, highly divergent in week 1, were roughly equal. By weeks 3–4, average growth rates in shutdown counties were in fact lower than even counties without plants. This lag structure for cases aligns with the fact that COVID-19 incubation periods may last for up to 14 days (Baud Baud et al. 2020).

The lower sustained COVID-19 growth rate postclosure suggests that plant closures have some relationship with COVID-19 transmission, which in turn suggests some relationship between plant-level activity and community disease spread within the county. Given that the average closing period was only nine days, it is unclear whether the plant closures themselves reduced COVID-19 transmission rates or whether closures resulted in plants taking more COVID-19 precautions (e.g., implementing enhanced safety protocols). It is also true that locales initially experiencing growth spikes will likely revert to average growth rates over time. However, the speed with which growth rates rose and fell in shutdown counties suggests that some closure-related mechanism is likely at play. And while shutdown counties have higher cumulative COVID-19 caseloads on average, this is likely because closures occur too late to suppress community spread outside of these plants.

4 Robustness

4.1 COVID-19 testing

Next we address concerns that these results primarily reflect differences in testing. Places with more testing tend to have more confirmed COVID-19 cases than places with less testing (mechanically). There does not appear to be a national database on county-level testing, so we compiled data from 31 states that have livestock facilities and testing data at the county level. Table 9 shows that while testing is positively associated with COVID-19 incidence, the relationship to livestock facilities remains large and significant. In a second specification, we add the positivity rate (total cases divided by total tests) as a further control. The magnitude of the livestock coefficients are of a similar magnitude to those in the baseline model in Table 1. However, these estimates are not directly comparable because of the smaller sample size of counties with testing data (1,773 counties across the 31 states).

4.2 Manufacturing activity

It is possible that a certain type of work similar to livestock processing—but not livestock processing itself—is driving the spread of COVID-19. To test this, we controlled for the county-level number of manufacturing establishments and share of income from manufacturing. We found that the relationship between livestock plants and COVID-19 incidence remained largely stable, meaning that it is not explained by a correlation with manufacturing (Table 10). While there is no obvious relationship with the number of manufacturing establishments, the coefficient for manufacturing share of income is positive and statistically significant, implying that manufacturing may be associated with higher COVID-19 incidence. Such a relationship is plausible given that, like livestock processing, employees in the manufacturing sector may work in close proximity and that many manufacturing activities are considered essential to supply chains.

4.3 Dropping counties distant from livestock plants

Another potential concern is that counties very far from livestock plants have lower population densities and different demographic make-ups than counties nearer these plants. Correspondingly, there is a risk that incorporating these counties into our analysis may introduce bias into our livestock plant estimates. An analysis omitting counties more than 100km from a county with a livestock plant shows a relationship with livestock facilities greater in magnitude than the base specification, indicating that our findings are robust to this risk and, perhaps, somewhat conservative (Table 11).

4.4 Dependent variable transformations

To address concerns about a skewed outcome variable, we used the natural log and inverse hyperbolic sine of the dependent variable and found a consistently positive but smaller magnitude relationship between livestock plants and increased COVID-19 case and death rates (Table 12).

4.5 Alternative statistical approaches to confounding

Above, we have shown the robustness of multivariate regression results to various confounders—demographic, geographic, and behavioral—and sample selection criteria. Additionally, we have shown that the dynamics over time of COVID-19 cases and deaths vary with the timing of livestock plant shutdowns.

Here, we present results of additional statistical methods used to explore the relationship between livestock plants and COVID-19 cases and deaths in the cross-section. The methods we use to help address potential bias and endogeneity concerns are: IV analysis, propensity score matching, and nearest neighbor matching. We note that the 259 counties in our sample with livestock plants differ in important ways from those without plants. We construct a balance table comparing counties with and without livestock plants (Table 13).

Counties with plants have higher population density, a lower proportion of elderly people, higher proportions of Black and Hispanic people, and larger household sizes. Income levels, by contrast, are similar. Each particular statistical method adjusts for these baseline differences in different ways. To preview, we find the observed relationship with COVID-19 incidence to be robust to all three approaches.

4.5.1 IVs

First, we employed an IV approach using historical livestock agricultural production data. The selection of this instrument is motivated by meat processors' need to minimize costs of transporting livestock supply when selecting the location of plants. In the first stage, we regressed the current number of livestock plants in each county on the county's livestock production value in 1959 in terms of animals sold, as derived from the USDA census. Note this includes only agricultural operations and not livestock processing. We believe this is a strong instrument given the facts that most of the interstate highway system was constructed during the 1960s, most currently operating livestock processing plants were built in the 1970s or later, and livestock agricultural operations in 1959 appear unlikely to affect current public health outcomes.

In the second stage, we regressed COVID-19 incidence on this predicted value of livestock plants, as well as the other covariates in the primary specification. The first stage in the IV analysis, presented in Table 14, shows that the instrument is highly relevant with an the F-statistic far above Stock and Yogo's 10% maximal bias threshold (Stock and Yogo 2002. The overall IV results in Table 15 show the relationship between livestock facilities and COVID-19 case and death rates to be even stronger for each outcome except the within-state death rate, which is of comparable magnitude but less precisely estimated. We note that the IV approach restricts identifying variation to that attributable to livestock agriculture proximity, thereby reducing statistical power.

4.5.2 Propensity score and nearest neighbor matching

For both propensity score matching and nearest neighbor matching, we constructed comparable subsamples of our dataset with and without livestock facilities to estimate an effect of having these livestock facilities among otherwise similar counties on COVID-19 cases and deaths.

For propensity score matching, we first predicted the probability that a county has at least one livestock facility (binary value) using a binomial regression that includes all the covariates from our primary model specification in Table 1, as well as their quadratic terms to increase model flexibility. We then confirmed that observations were relatively balanced across covariates within each propensity score quartile (Table 16. This suggests that the propensity score is indeed balancing the multidimensional covariates. In a second step, we used this predicted probability (i.e., the propensity score) as a control in a regression of COVID-19 incidence on livestock plants. The idea here is that the propensity score helps account for bias in the location of livestock plants.

For nearest neighbor matching, we use the *MatchIt* package in R to restrict the sample to similar treated and control groups. The matching occured using a nearest neighbor algorithm based on predicting the livestock binary variable with the covariates in our primary specification. To ensure an adequate sample size, we allowed the algorithm to match two non-plant counties to every one county with a livestock plant. We found the resulting 774 county subsample to be well balanced (Table 17).[7] Table 18 consolidates the results and includes outputs from Table 1 for reference. In this analysis, coefficients for both case and death rates remain of a similar magnitude and level of significance.

[7] A balance table for the entire sample is shown in Table 13.

4.6 Community spread beyond livestock plants

COVID-19 transmission likely extends beyond the county containing the livestock plant. Fig. 7 expands our main analysis to include neighboring counties grouped by distance band, as charted in Fig. 1 and visualized in the map in Fig. 3. We found evidence of a relationship between livestock plants and increased COVID-19 case rates up to to 150km away from a plant, further supporting the notion of community spread beyond the immediate work context.[8]

To validate and contextualize our findings, we first estimated the total excess cases and deaths related to livestock plants implied by our results. For one set of estimates, we multiplied the plant-level coefficient for excess cases and deaths related to livestock plants by the total number of plants and the average population per plant to arrive at a national total. Taking a second approach, we used a binary measure for whether a county has one or more livestock plants and multiplied this coefficient by the county-level mean population and number of counties with livestock plants. The estimates resulting from this exercise are, respectively, 236,000–310,000 cases and 4,300–5,200 deaths. A summary of this calculation is shown in Table S19.

Next, we estimated the share of cases among livestock employees relative to total excess cases in an attempt to determine the share of excess cases that may be occurring outside the livestock plants. We used the CDC's state-level aggregate count of livestock workers testing positive for COVID-19 as of May 31 across 26 states (Waltenburg et al. 2020). Comparing this to state-level case data as of May 31, we found that livestock workers represented 2.7% of cases in these states. Using this ratio to estimate the total number of infected livestock workers among all of the cases observed in these states on July 21, we arrived at an estimate of 35,635 infected workers, approximately 7% of the industry's entire

[8] We present summary statistics by distance band in Tables 19 – 21. The average number of counties in each band increases with distance. There is a clear positive relationship between COVID-19 cases and deaths in relation to livestock facilities, and the county-level mean case rate varies directly with a county's proximity to a neighboring county with a livestock facility.

employee base. Using our calculation of 236,000–310,000 cases nationwide due to livestock plants, we estimated that livestock workers represented 12–15% of these excess total cases. In other words, for every worker infected at a livestock plant, between 7 and 8 local non-workers were ultimately infected by the end of the sample period, underscoring the high potential for community spread.

5 Discussion

Angrist and Krueger (Angrist and Krueger 1999) noted that “one should always be wary of drawing causal inferences from observational data." We know of no random assignment design that could address our research question and thereby yield the most reliable path to causal inference. The best we can do here is provide an unusually broad array of observational evidence. This includes (but is not limited to) ruling out the most obvious confounders, a cross-sectional IV, and the event study analysis leveraging shutdown timing. A still more compelling natural experiment would leverage explicit and exogenous variation that drives livestock plant shutdowns, i.e., an IV for the shutdowns or their timing. Unfortunately, we know of no such identifying variation.

Readers may disagree on whether our array of analyses has isolated a causal effect. Given this, and in order to be conservative, we avoid causal language throughout our text so as not to overstate the “hardness" of our method (Akerlof 2020). This avoidance and caution stands in contrast to other recent, impactful work on COVID-19.

Still, we believe our array of analyses constitutes the best feasible approach to shed light on the role of livestock processing plants in the US COVID-19 pandemic. For a question of this importance, we believe there is no “harder" method available (2020). As policymakers and industry leaders seek to preserve vital food supply chains while mitigating the pandemic's spread, evidence on the potential scope of the issue is particularly valuable, as well as assessment of the relationship between temporary plant shutdowns and subsequent

COVID-19 growth dynamics.

Although our estimate that 6–8% of US COVID-19 cases are associated with livestock plants may appear high, it is important to recall that high levels of geographic heterogeneity in COVID-19 incidence can be explained by some combination of individual behavior, government policy, social distancing compliance, and economic activity: the US, for example, has 4% of the world's population but approximately a quarter of all cases and deaths. When narrowing the geographic focus, we can imagine the distribution of COVID-19 incidence to be similarly clustered, if not even lumpier.

Kansas provides a telling example of the outsized role of livestock facilities: as of July 20, 3,200 out of 23,300 state cases (14%) were directly linked to meatpacking (Kansas Department of Health and Environment 2020). For context, there are 17,200 employees in the animal slaughtering industry in Kansas (**aes**), or 0.6% of the state's population, suggesting that livestock plants had a relationship of a magnitude closer in scale to our own estimates (Kansas' estimate is 23x of the industry's labor footprint). Although the figure we are estimating in our study (6–8% of all US cases out of a national livestock workforce of 0.15%, or a multiplier of 40–53x) is larger, we believe that this finding is plausible considering follow-on community spread; Kansas' official tally, though evidently aided to some degree of contract tracing, is reportedly hampered by lags in hiring staff and legislative actions that have inhibited tracing efforts (Mitchell 2020). That is, the figure we have calculated could in fact be more complete than the Kansas figure in capturing the spread resulting from livestock plants.

Our analysis of individual meatpacking companies may present an opportunity to explore how differences in corporate structure and operating practices may account for their differential public health outcomes. In particular, the evidence that shutting down plants temporarily may be related to decreases in COVID-19 case growth presents a potentially powerful transmission mitigant. In addition, the positive relationship between COVID-19 transmission and production line speed waivers issued to poultry plants, particularly those during the 2020 pandemic, is notable given that these waivers are intended for plants with safe commercial production practices and microbial control.[9] This finding suggests a need for additional examination of this program.

An implication of this study is some aspects of large meat processing plants render them especially susceptible to spreading respiratory viruses. One potential explanation is that large plants simply entail more activity and employ more people. Because these plants provide a central location for moving products, it is plausible that a linear increase in the potential infected within the plant would entail a nonlinear response, owing to the complex and exponential nature of disease transmission dynamics (Grassly and Fraser 2008). Another driver may be the large physical spaces where processing occurs. Larger rooms tend to be louder and thus require more shouting (Borak 2020, and they may require stronger climate control, which, as we note in our introduction, may aggravate COVID-19 spread. A larger space that employees must navigate in reaching their workstations may also increase the number of workplace interactions.

More broadly, the finding that meatpacking plants may contribute to high levels of community spread underscores the potential negative public health externalities generated by the industry, which may be attributable to industrial concentration, operating practices, and labor conditions. Complicating this matter from an economic standpoint is the supply chain choke point created by large plants disrupted by COVID-19, causing food shortages, driving up prices, and incurring substantial upstream and downstream economic losses. Cataloguing and addressing the underlying factors that produced this systemic risk in the first place could not only strengthen the US food system in the face of COVID-19 and future disruptions but also help illuminate analogous weak points in other industries and supply chains.

[9] In contrast, some plants receiving waivers recent Occupational Safety and Health violations (Thompson and Berkowitz 2020).

References

Akerlof, George A. 2020. "Sins of Omission and the Practice of Economics." *Journal of Economic Literature* 58 (2): 405–18.

Anderson, Roy M, Hans Heesterbeek, Don Klinkenberg, and T Déirdre Hollingsworth. 2020. "How will country-based mitigation measures influence the course of the COVID-19 epidemic?" *The Lancet* 395 (10228): 931–934.

Angrist, Joshua D, and Alan B Krueger. 1999. "Empirical strategies in labor economics." In *Handbook of labor economics,* 3:1277–1366. Elsevier.

Asadi, Sima, Nicole Bouvier, Anthony S Wexler, and William D Ristenpart. 2020. *The coronavirus pandemic and aerosols: Does COVID-19 transmit via expiratory particles?*

Barbieri, Teresa, Gaetano Basso, and Sergio Scicchitano. 2020. "Italian workers at risk during the Covid-19 epidemic." *Available at SSRN 3572065.*

Baud, David, Xiaolong Qi, Karin Nielsen-Saines, Didier Musso, Léo Pomar, and Guillaume Favre. 2020. "Real estimates of mortality following COVID-19 infection." *The Lancet Infectious Diseases* 20 (7): 773. ISSN: 1473-3099. https://doi.org/10.1016/S1473-3099(20)30195-X. https://doi.org/10.1016/S1473-3099(20)30195-X.

Beck, Samuel H, Alejandro Castillo, Kerry A Kinney, Alexander Zuniga, Zahra Mohammad, Ronald E Lacey, and Maria D King. 2019. "Monitoring of Pathogenic Bioaerosols in Beef Slaughter Facilities Based on Air Sampling and Airflow Modeling." *Applied Engineering in Agriculture* 35 (6): 1015–1036.

Borak, Jonathan. 2020. "Airborne Transmission of COVID-19." *Occupational Medicine* 70 (5): 297–299. ISSN: 0962-7480. https://doi.org/10.1093/occmed/kqaa080. eprint: https://academic.oup.com/occmed/article-pdf/70/5/297/33506386/kqaa080.pdf. https://doi.org/10.1093/occmed/kqaa080.

Borjas, George J. 2020. *Demographic determinants of testing incidence and COVID-19 infections in New York City neighborhoods.* Technical report. National Bureau of Economic Research.

Bouffanais, Roland, and Sun Sun Lim. 2020. "Cities-try to predict superspreading hotspots for COVID-19." *Nature* 583 (7816): 352–355. https://doi.org/10.1038/d41586-020-02072-3.

Busvine, Douglas. 2020. "Coronavirus spread accelerates again in Germany." *Reuters,* https://www.reuters.com/article/us-health-coronavirus-germany-cases/coronavirus-spread-accelerates-again-in-germany-idUSKBN22M019.

Cano-Muñoz, G, and Germán Cano Muñoz. 1991. *Manual on meat cold store operation and management.* 92. Food & Agriculture Org.

Carleton, Tamma, and Kyle C Meng. 2020. "Causal empirical estimates suggest COVID-19 transmission rates are highly seasonal." *medRxiv.*

Chadde, Sky. 2020. *Tracking Covid-19's impact on meatpacking workers and industry.* https://investigatemidwest.org/2020/04/16/tracking-covid-19s-impact-on-meatpacking-workers-and-industry/.

Compa, Lance A. 2004. "Blood, sweat, and fear: Workers' rights in US meat and poultry plants." *Human Rights Watch.*

Constance, Douglas H., Francisco Martinez-Gomez, Gilberto Aboites-Manrique, and Alessandro Bonanno. 2013. "The Problems with Poultry Production and Processing." In *The Ethics and Economics of Agrifood Competition,* edited by Harvey S. James Jr., 155–175. Dordrecht: Springer Netherlands. ISBN: 978-94-007-6274-9. https://doi.org/10.1007/978-94-007-6274-9_8. https://doi.org/10.1007/978-94-007-6274-9_8.

Deyle, Ethan R., M. Cyrus Maher, Ryan D. Hernandez, Sanjay Basu, and George Sugihara. 2016. "Global environmental drivers of influenza." *Proceedings of the National Academy of Sciences of the United States of America* 113 (46): 13081–13086. ISSN: 0027-8424. https://doi.org/10.1073/pnas.1607747113. eprint: https://www.pnas.org/content/113/46/13081.full.pdf. https://www.pnas.org/content/113/46/13081.

Dowd, Jennifer Beam, Liliana Andriano, David M Brazel, Valentina Rotondi, Per Block, Xuejie Ding, Yan Liu, and Melinda C Mills. 2020. "Demographic science aids in understanding the spread and fatality rates of COVID-19." *Proceedings of the National Academy of Sciences of the United States of America* 117 (18): 9696–9698.

Dyal, Jonathan W. 2020. "COVID-19 Among Workers in Meat and Poultry Processing Facilities—19 States, April 2020." *MMWR. Morbidity and Mortality Weekly Report* 69.

Ebrahim, Shahul H, Qanta A Ahmed, Ernesto Gozzer, Patricia Schlagenhauf, and Ziad A Memish. 2020. *Covid-19 and community mitigation strategies in a pandemic.*

European of Food, Agriculture, and Tourism Trade Unions. 2011. *Putting meat on the bones: a report on the structure and dynamics of the European meat industry.* http://www.meat-workers.org/sites/default/files/documents/EFFAT_PuttingMeatOnTheBones_EN.pdf.

Foley, Ryan J. 2020. *Outbreak at Iowa pork plant was larger than state reported.* https://apnews.com/85a02d9296053980ea47eba97f920707.

Fremstad, Shawn, Hye Jin Rho, and Hayley Brown. 2020. "Meatpacking Workers are a Diverse Group Who Need Better Protections," https://cepr.net/meatpacking-workers-are-a-diverse-group-who-need-better-protections/.

Gershman, Jacob. 2020. *A Guide to State Coronavirus Reopenings and Lockdowns.* https://www.wsj.com/articles/a-state-by-state-guide-to-coronavirus-lockdowns-11584749351.

Google LLC. 2020. *Google COVID-19 Community Mobility Reports.* https://www.google.com/covid19/mobility/. Accessed July 13, 2020.

Grabell, Michael. 2020. *What Happens If Workers Cutting Up the Nation's Meat Get Sick?* https://www.propublica.org/article/what-happens-if-workers-cutting-up-the-nations-meat-get-sick.

Grassly, Nicholas C, and Christophe Fraser. 2008. "Mathematical models of infectious disease transmission." *Nature Reviews Microbiology* 6 (6): 477–487.

Haines, Michael, Price Fishback, and Paul Rhode. 2018. *United States Agriculture Data, 1840 - 2012.* https://doi.org/10.3886/ICPSR35206.v4. https://doi.org/10.3886/ICPSR35206.v4.

Hamner, Lea. 2020. "High SARS-CoV-2 attack rate following exposure at a choir practice—Skagit County, Washington, March 2020." *MMWR. Morbidity and Mortality Weekly Report* 69. https://doi.org/10.15585/mmwr.mm6919e6.

Hendrickson, Mary, Philip H Howard, and Douglas Constance. 2017. "Power, food and agriculture: implications for farmers, consumers and communities." *Consumers and Communities.*

Hendrickson, Mary K. 2015. "Resilience in a concentrated and consolidated food system." *Journal of Environmental Studies and Sciences* 5 (3): 418–431.

Hirtzer, Michael, and Tatiana Freitas. 2020. "US Could Be Weeks From Meat Shortages With Shutdowns Spreading." *Bloomberg.com,* https://www.bloomberg.com/news/articles/2020-04-24/meat-threats-grow-with-first-brazil-shutdown-u-s-turkey-halt?sref=q8seIhDd.

Hsiang, Solomon, Daniel Allen, Sébastien Annan-Phan, Kendon Bell, Ian Bolliger, Trinetta Chong, Hannah Druckenmiller, et al. 2020. "The effect of large-scale anti-contagion policies on the COVID-19 pandemic." *Nature,* ISSN: 1476-4687. https://doi.org/10.1038/s41586-020-2404-8. https://doi.org/10.1038/s41586-020-2404-8.

Inc., Tyson Foods. 2020. *Tyson Foods Facts.* Tyson Foods Inc. https://ir.tyson.com/about-tyson/facts/default.aspx.

Kandel, William, and Emilio A Parrado. 2005. "Restructuring of the US meat processing industry and new Hispanic migrant destinations." *Population and Development Review* 31 (3): 447–471.

Kansas Department of Health and Environment. 2020. *2019 Novel Coronavirus Summary.* https://www.coronavirus.kdheks.gov/DocumentCenter/View/1125/Historical---August-3?bidId=.

Koep, Tyler H, Felicity T Enders, Chris Pierret, Stephen C Ekker, Dale Krageschmidt, Kevin L Neff, Marc Lipsitch, Jeffrey Shaman, and W Charles Huskins. 2013. "Predictors of indoor absolute humidity and estimated effects on influenza virus survival in grade schools." *BMC infectious diseases* 13 (1): 71.

Kraemer, Moritz U. G., Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, et al. 2020. "The effect of human mobility and control measures on the COVID-19 epidemic in China." *Science* 368 (6490): 493–497. ISSN: 0036-8075. https://doi.org/10.1126/science.abb4218. eprint: https://science.sciencemag.org/content/368/6490/493.full.pdf. https://science.sciencemag.org/content/368/6490/493.

Leah Douglas, May 14. 2020. *As more meatpacking workers fall ill from Covid-19, meat companies decline to disclose data.* https://thefern.org/ag_insider/as-more-meatpacking-workers-fall-ill-from-covid-19-meat-companies-decline-to-disclose-data/.

Lee, Jasmine C., Sarah Mervosh, Yuriria Avila, Barbara Harvey, and Alex Leeds Matthews. 2020. *See How All 50 States Are Reopening (and Closing Again).* https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html.

Lewnard, Joseph A, and Nathan C Lo. 2020. "Scientific and ethical basis for social-distancing interventions against COVID-19." *The Lancet. Infectious diseases.*

MacDonald, James M, and William D McBride. 2009. "The transformation of US livestock agriculture scale, efficiency, and risks." *Economic Information Bulletin,* no. 43.

MacDonald, James M., Michael Ollinger, Kenneth E. Nelson, and Charles R. Handy. 2000. "CONSOLIDATION IN US MEATPACKING." (2000), Agriculture Economic Report Number 785, nos. 1473-2016-120772, 42. https://doi.org/10.22004/ag.econ.34021. http://ageconsearch.umn.edu/record/34021.

Mano, Ana. 2020. "Nine meat plants in southern Brazil face COVID-19 outbreaks." *Reuters,* https://www.reuters.com/article/us-health-coronavirus-brazil-meatpackers-idUSKBN22C2J8.

Mason, Jeff, and Tom Polansek. 2020. "Trump orders US meat-processing plants to stay open despite coronavirus fears," https://www.reuters.com/article/us-health-coronavirus-trump-liability/trump-orders-us-meat-processing-plants-to-stay-open-despite-coronavirus-fears-idUSKCN22A2OB.

Mayer, Jane. 2020. *How Trump Is Helping Tycoons Exploit the Pandemic.* https://www.newyorker.com/magazine/2020/07/20/how-trump-is-helping-tycoons-exploit-the-pandemic.

Mitchell, Conner. 2020. *Kansas' top health official sounds the alarm on COVID-19, predicts current trend line will 'steepen'.* https://www2.ljworld.com/news/state-region/2020/jul/01/kansas-top-health-official-sounds-the-alarm-on-covid-19-predicts-current-trend-line-will-steepen/.

Mittal, Rajat, Rui Ni, and Jung-Hee Seo. 2020. "The flow physics of COVID-19." *Journal of Fluid Mechanics* 894:F2. https://doi.org/10.1017/jfm.2020.330.

Moghadas, Seyed M, Affan Shoukat, Meagan C Fitzpatrick, Chad R Wells, Pratha Sah, Abhishek Pandey, Jeffrey D Sachs, Zheng Wang, Lauren A Meyers, Burton H Singer, et al. 2020. "Projecting hospital utilization during the COVID-19 outbreaks in the United States." *Proceedings of the National Academy of Sciences of the United States of America* 117 (16): 9122–9126.

Myers, Kendall P, Christopher W Olsen, Sharon F Setterquist, Ana W Capuano, Kelley J Donham, Eileen L Thacker, James A Merchant, and Gregory C Gray. 2006. "Are swine workers in the United States at increased risk of infection with zoonotic influenza virus?" *Clinical infectious diseases* 42 (1): 14–20.

NASHP Staff. 2020. *Each State's COVID-19 Reopening and Reclosing Plans and Mask Requirements.* https://www.nashp.org/governors-prioritize-health-for-all/.

National Academies of Sciences, Engineering, and Medicine. 2020. *Rapid Expert Consultation on SARS-CoV-2 Survival in Relation to Temperature and Humidity and Potential for Seasonality for the COVID-19 Pandemic (April 7, 2020).* Washington, DC: The National Academies Press. https://doi.org/10.17226/25771. https://www.nap.edu/catalog/25771/rapid-expert-consultation-on-sars-cov-2-survival-in-relation-to-temperature-and-humidity-and-potential-for-seasonality-for-the-covid-19-pandemic-april-7-2020.

National Pork Board. 2019. *Estimated Daily US Slaughter Capacity by Plant (head per day).* National Pork Board. National Pork Board.

Novack, Christopher Collins, and Sophie. 2020. *COVID-19 Cases Now Tied to Meat Plants in Rural Texas Counties Wracked with Coronavirus.* https://www.texasobserver.org/east-texas-coronavirus-chicken/.

O'Reilly, Kathleen M, Megan Auzenbergs, Yalda Jafari, Yang Liu, Stefan Flasche, and Rachel Lowe. 2020. "Effective transmission across the globe: the role of climate in COVID-19 mitigation strategies." *The Lancet. Planetary Health.*

Order, White House Executive. 2020. *Executive Order on Delegating Authority Under the DPA with Respect to Food Supply Chain Resources During the National Emergency Caused by the Outbreak of COVID-19.* https://www.whitehouse.gov/presidential-actions/executive-order-delegating-authority-dpa-respect-food-supply-chain-resources-national-emergency-caused-outbreak-covid-19/.

Roser, Max, Hannah Ritchie, Esteban Ortiz-Ospina, and Joe Hasell. 2020. "Coronavirus pandemic (COVID-19)." *Our World in Data.*

Sajadi, Mohammad M, Parham Habibzadeh, Augustin Vintzileos, Shervin Shokouhi, Fernando Miralles-Wilhelm, and Anthony Amoroso. 2020. "Temperature and latitude analysis to predict potential spread and seasonality for COVID-19." *Available at SSRN 3550308.*

Schlottau, Kore, Melanie Rissmann, Annika Graaf, Jacob Schön, Julia Sehl, Claudia Wylezich, Dirk Höper, Thomas C Mettenleiter, Anne Balkema-Buschmann, Timm Harder, et al. 2020. "Experimental Transmission Studies of SARS-CoV-2 in Fruit Bats, Ferrets, Pigs and Chickens." *The Lancet.*

Scott, Jason, and Ainslie Chandler. 2020. "An Australian Meatworks Is at the Center of a Virus Outbreak." *Bloomberg.com,* https://www.bloomberg.com/news/articles/2020-05-04/virus-outbreak-in-australian-meatworks-echoes-problems-in-u-s.

SEER Program, National Cancer Institute, NIH. 2020. *US Population Data.* Underlying data provided by National Center for Health Statistics. Accessed May 15, 2020 at https://seer.cancer.gov/popdata/download.html#19.

Shaman, Jeffrey, and Melvin Kohn. 2009. "Absolute humidity modulates influenza survival, transmission, and seasonality." *Proceedings of the National Academy of Sciences of the United States of America* 106 (9): 3243–3248.

Skerritt, Jen, Deena Shanker, and Michael Hirtzer. 2020. *Meat Shortages Reopen Costly Path to Smaller US Plants.* https://www.bloomberg.com/news/articles/2020-06-26/meat-shortages-reopen-costly-path-to-small-u-s-slaughterhouses.

Stock, James H, and Motohiro Yogo. 2002. *Testing for Weak Instruments in Linear IV Regression.* Working Paper, Technical Working Paper Series 284. National Bureau of Economic Research. https://doi.org/10.3386/t0284. http://www.nber.org/papers/t0284.

Subbian, Vignesh, Anthony Solomonides, Melissa Clarkson, Vasiliki Nataly Rahimzadeh, Carolyn Petersen, Richard Schreiber, Paul R DeMuro, et al. 2020. "Ethics and Informatics in the Age of COVID-19: Challenges and Recommendations for Public Health Organization and Public Policy." Ocaa188, *Journal of the American Medical Informatics Association,* ISSN: 1527-974X. https://doi.org/10.1093/jamia/ocaa188. eprint: https://academic.oup.com/jamia/article-pdf/doi/10.1093/jamia/ocaa188/33539766/ocaa188.pdf. https://doi.org/10.1093/jamia/ocaa188.

The New York Times. 2020. *Coronavirus (Covid-19) Data in the United States.* https://github.com/nytimes/covid-19-data.

Thompson, Shayla, and Deborah Berkowitz. 2020. *USDA Allows Poultry Plants to Raise Line Speeds, Exacerbating Risk of COVID-19 Outbreaks and Injury.* https://www.nelp.org/publication/usda-allows-poultry-plants-raise-line-speeds-exacerbating-risk-covid-19-outbreaks-injury/.

US Bureau of Economic Analysis. 2020. *Table SAINC5: Personal Income by Major Component and Industry.* Accessed May 15, 2020 at https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=1#reqid=70&step=1.

US Census Bureau. 2019a. *All Sectors: County Business Patterns by Legal Form of Organization and Employment Size Class for US, States, and Selected Geographies: 2018.* https://www2.census.gov/programs-surveys/cbp/data/2018/CB1800CBP.zip.

US Census Bureau. 2019b. *Commuting characteristics by sex, 2004-2018 American Community Survey 5-year estimates.* https://data.census.gov/cedsci/table?tid=ACSST1Y2018.S0801.

———. 2019c. *Group quarters population by group quarters type, 2004-2018 American Community Survey 5-year estimates.* https://data.census.gov/cedsci/table?q=group%5C%20quarters&tid=ACSST1Y2018.S2602.

———. 2019d. *Households and families, 2004-2018 American Community Survey 5-year estimates.* https://data.census.gov/cedsci/table?tid=ACSST1Y2017.S1101.

———. 2019e. *Selected characteristics of health care coverage in the United States, 2004-2018 American Community Survey 5-year estimates.* https://data.census.gov/cedsci/table?tid=ACSST1Y2018.S2701.

US Department of Agriculture. 2018. *Petition To Permit Waivers of Maximum Line Speeds for Young Chicken Establishments Operating Under the New Poultry Inspection System; Criteria for Consideration of Waiver Requests for Young Chicken Establishments To Operate at Line Speeds of Up to 175 Birds per Minute.* https://www.federalregister.gov/documents/2018/09/28/2018-21143/petition-to-permit-waivers-of-maximum-line-speeds-for-young-chicken-establishments-operating-under.

US Department of Homeland Security. 2020. *Advisory memorandum on identification of essential critical infrastructure workers during COVID-19 response.* https://www.cisa.gov/sites/default/files/publications/Version_3.0_CISA_Guidance_on_Essential_Critical_Infrastructure_Workers_1.pdf.

US Department of Transportation. 2020. *FAF4 Network Database and Flow Assignment: 2012 and 2045.* https://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf4/netwkdbflow/index.htm.

USDA Food Safety and Inspection Service. 2020a. *Meat, Poultry and Egg Product Inspection Directory.* https://www.fsis.usda.gov/wps/portal/fsis/topics/inspection/mpi-directory.

———. 2020b. *Salmonella Initiative Program Participants Table.* https://www.fsis.usda.gov/wps/wcm/connect/188bf583-45c9-4837-9205-37e0eb1ba243/waiver_table.pdf?MOD=AJPERES.

VanderWaal, Kimberly, and John Deen. 2018. "Global trends in infectious diseases of swine." *Proceedings of the National Academy of Sciences of the United States of America* 115 (45): 11495–11500. ISSN: 0027-8424. https://doi.org/10.1073/pnas.1806068115. eprint: https://www.pnas.org/content/115/45/11495.full.pdf. https://www.pnas.org/content/115/45/11495.

Viscusi, W. Kip. 1980. "Union, labor market structure, and the welfare implications of the quality of work." *Journal of Labor Research* 1 (1): 175–192. ISSN: 1936-4768. https://doi.org/10.1007/BF02685204. https://doi.org/10.1007/BF02685204.

Waltenburg, Michelle A, Tristan Victoroff, Charles E Rose, Marilee Butterfield, Rachel H Jervis, Kristen M Fedak, Julie A Gabel, Amanda Feldpausch, Eileen M Dunne, Connie Austin, et al. 2020. "Update: COVID-19 among workers in meat and poultry processing facilities—United States, April–May 2020." *MMWR. Morbidity and Mortality Weekly Report* 69. https://doi.org/10.15585/mmwr.mm6927e2.

Ward, Clement E. 2002. "A Review of Causes for and Consequences of Economic Concentration in the US Meatpacking Industry." *CAFRI: Current Agriculture, Food and Resource Issues* (2002-01-06), Number 3, nos. 519-2016-37619, 28. https://doi.org/10.22004/ag.econ.45696. http://ageconsearch.umn.edu/record/45696.

Wells, Chad R., Pratha Sah, Seyed M. Moghadas, Abhishek Pandey, Affan Shoukat, Yaning Wang, Zheng Wang, Lauren A. Meyers, Burton H. Singer, and Alison P. Galvani. 2020. "Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak." *Proceedings of the National Academy of Sciences of the United States of America* 117 (13): 7504–7509. https://doi.org/10.1073/pnas.2002616117. eprint: https://www.pnas.org/content/117/13/7504.full.pdf. https://www.pnas.org/content/117/13/7504.

Wohlgenant, Michael K. 2013. "Competition in the US meatpacking industry." *Annu. Rev. Resour. Econ.* 5 (1): 1–12.

Yang, Qiqi, Xiang Zhao, Philippe Lemey, Marc A. Suchard, Yuhai Bi, Weifeng Shi, Di Liu, et al. 2020. "Assessing the role of live poultry trade in community-structured transmission of avian influenza in China." *Proceedings of the National Academy of Sciences of the United States of America* 117 (11): 5949–5954. ISSN: 0027-8424. https://doi.org/10.1073/pnas.1906954117. eprint: https://www.pnas.org/content/117/11/5949.full.pdf. https://www.pnas.org/content/117/11/5949.

Zuber, Sophie, and Harald Brüssow. 2020. "COVID 19: challenges for virologists in the food industry." *Microbial Biotechnology,* https://doi.org/10.1111/1751-7915.13638. eprint: https://sfamjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1751-7915.13638. https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1751-7915.13638.

6 Figures

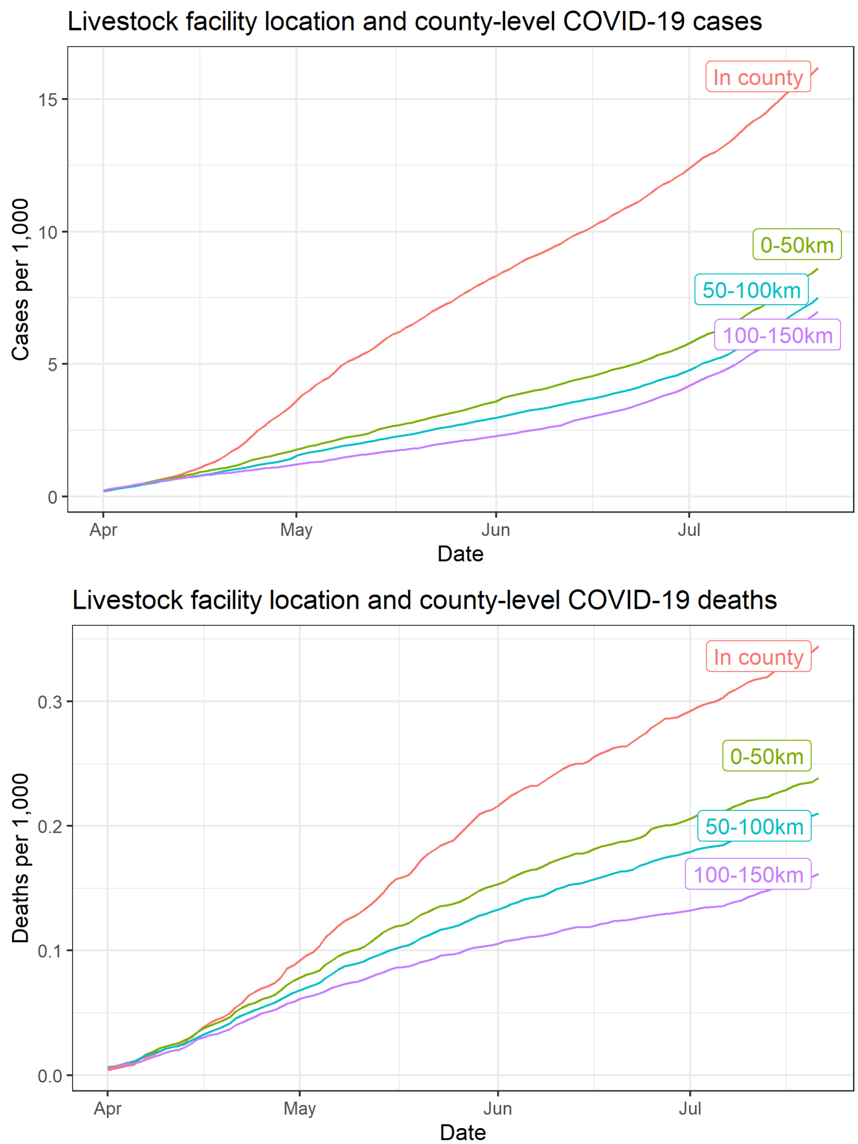

Figure 1: Mean county-level COVID-19 cases per thousand (top) and deaths per thousand (bottom) over time based on proximity to a livestock facility. The band '0–50km' excludes the county itself. Counties are categorized into non-overlapping, single categories based on the nearest facility (e.g., if a county contains a livestock facility and is within 50 km of another facility outside the county, the county is coded 'In county' and not '0–50 km'. A visualization map is included in Fig. 3

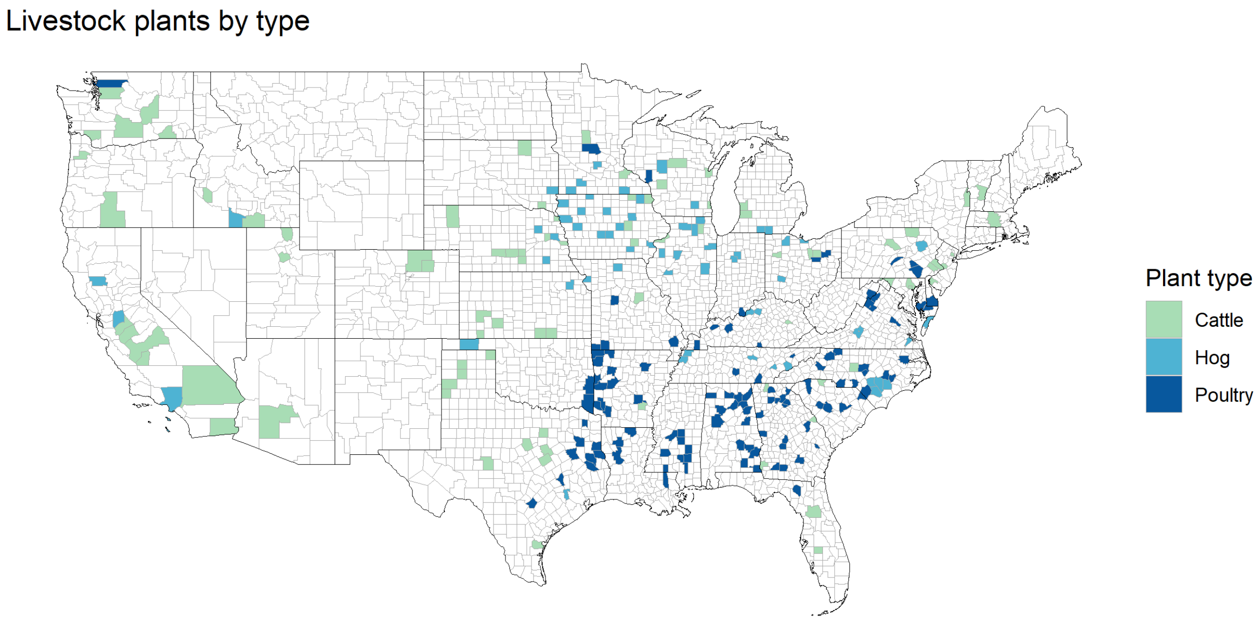

Figure 2: Shaded counties contain at least one beef or pork facility categorized by USDA FSIS as processing more than one million pounds per month (Categories 4 and 5) or at least one poultry facility categorized as processing more than ten million pounds per month (Category 5).

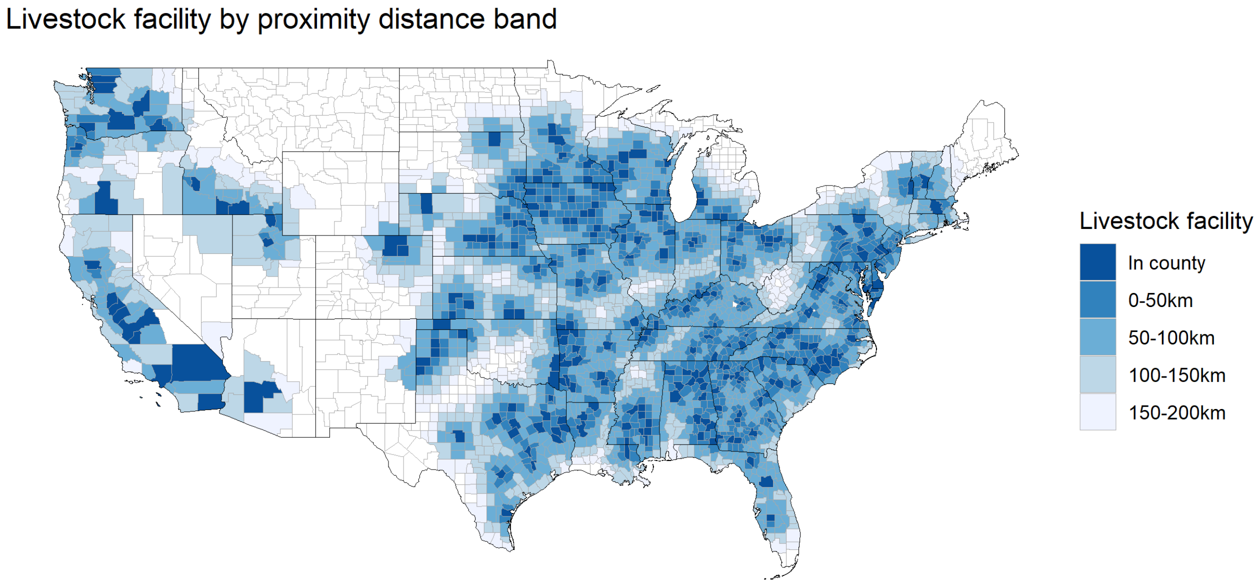

Figure 3: Map of counties in terms of proximity to livestock facilities based on county geographic centroids. The band ‘0–50km’ excludes the county itself. Counties are categorized into non-overlapping, single categories based on the nearest facility (e.g., if a county contains a livestock facility and is within 50km of another facility outside the county, the county is coded ‘In county’ and not ‘0–50km’.

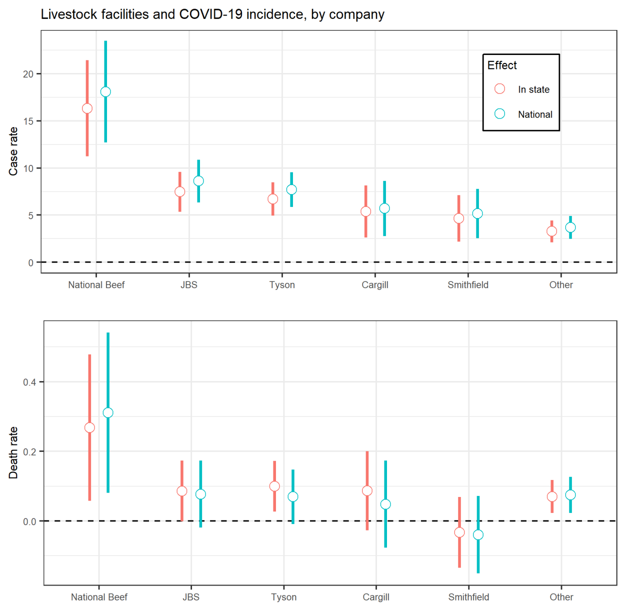

Figure 4: Relationship between COVID-19 cases and livestock plants owned or operated by large meatpacking companies. Coefficients are firm fixed effect coefficients plotted from Table 7. Error bars represent 95% confidence intervals.

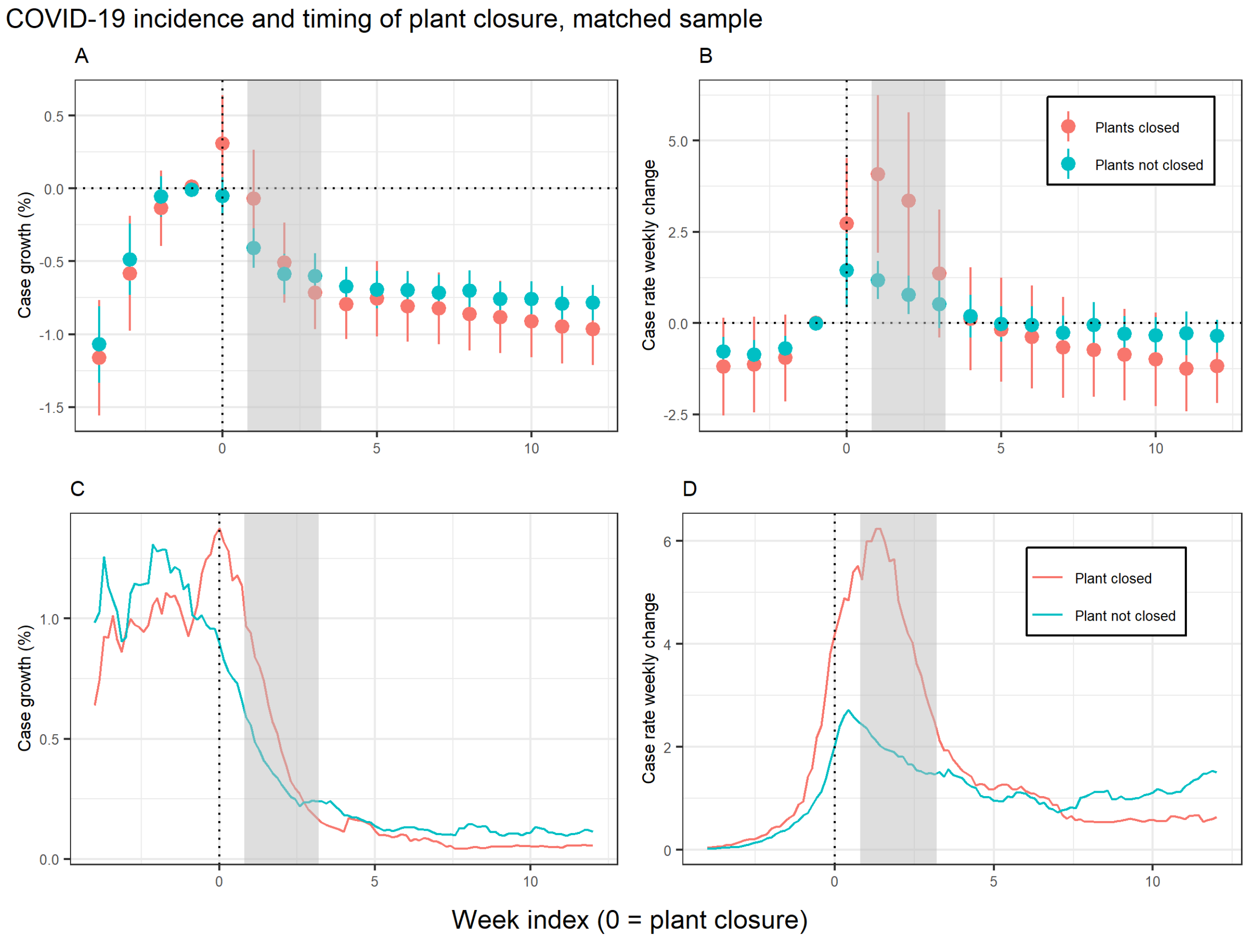

Figure 5: Graphs match COVID-19 pre-trends of control group (green lines) to counties with plant shutdowns (red lines) based on percent growth in cases (weekly log difference) in the two weeks prior to shutdown. Selected counties are in the top quartile of growth rates among the 233 counties with livestock plants that did not have a plant shutdown. For non-shutdown counties, week 0 is assigned to the mean shutdown date, April 22, 2020. Panels A and B plot coefficients from a panel regression where counties are interacted with the weekly event index in terms of percent growth in cases (A) and change in case rates per 1000 (B). Estimates are relative to the baseline trend across all counties. One week prior (week -1) is omitted as the reference level. Models control for stay-at-home orders at the state level and include a fixed effect for each county. Error bars reflect a 95% confidence interval. Panels C and D are daily line charts of the mean values of each group in terms of percent case growth and change in case rate, respectively. Grey shaded bars reflect the estimated period when the effect of closing a plant would have been reflected in cases (1-3 weeks after) given that incubation periods may last up to 14 days) (Baud 2020).

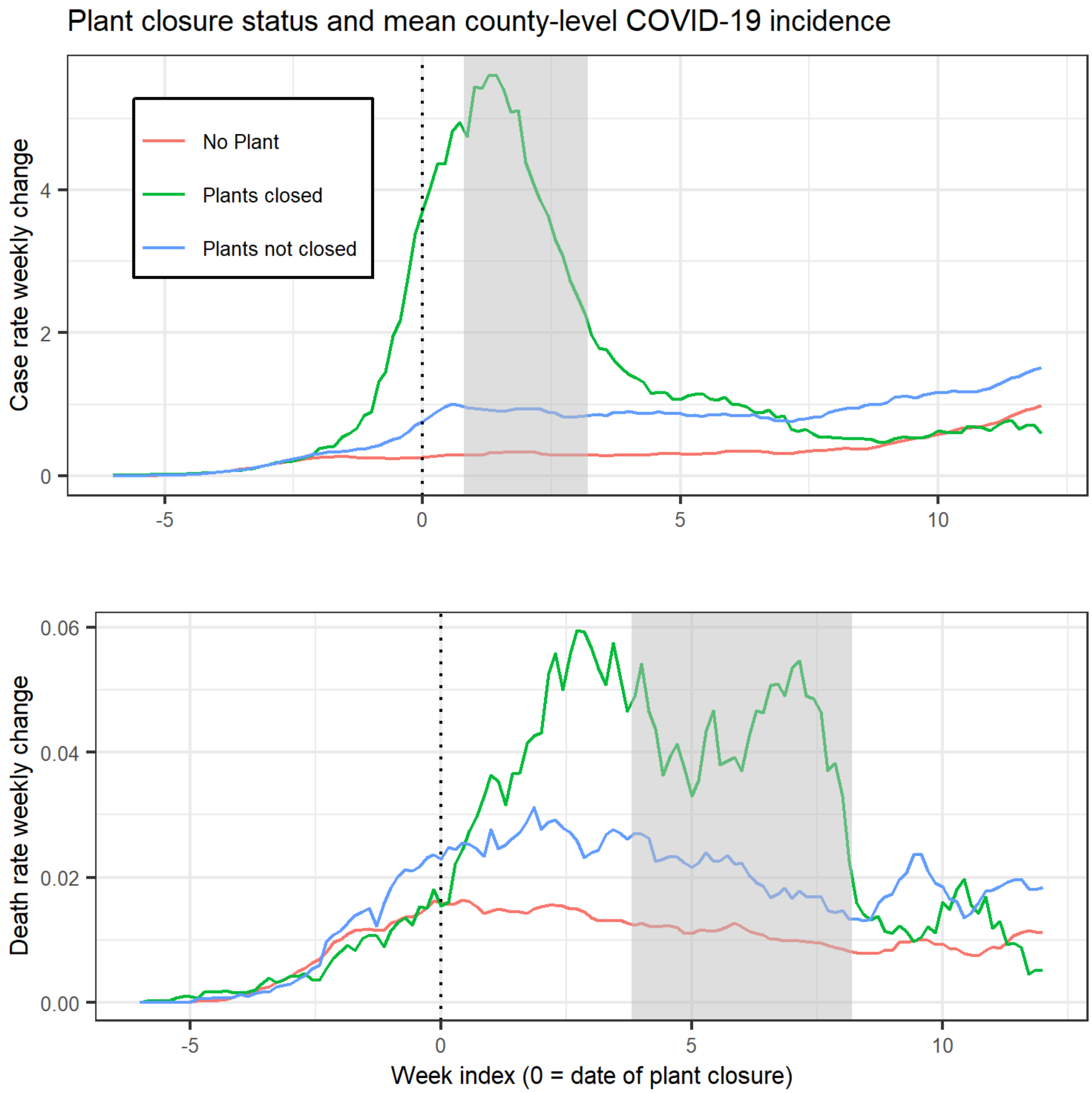

Figure 6: Comparisons of change in weekly county-level COVID-19 case and death rates across counties with livestock plants that temporarily shut down due to COVID-19 concerns, counties with plants that did not shut down, and counties without plants. For non-shutdown counties, week 0 is assigned to be the median shutdown date, April 22, 2020. Grey shaded bars reflect the estimated period when the effect of closing a plant would have been reflected in cases (1-3 weeks after) given that incubation periods may last up to 14 days, and the time between symptom onset and death (4-8 weeks after) (Baud 2020).

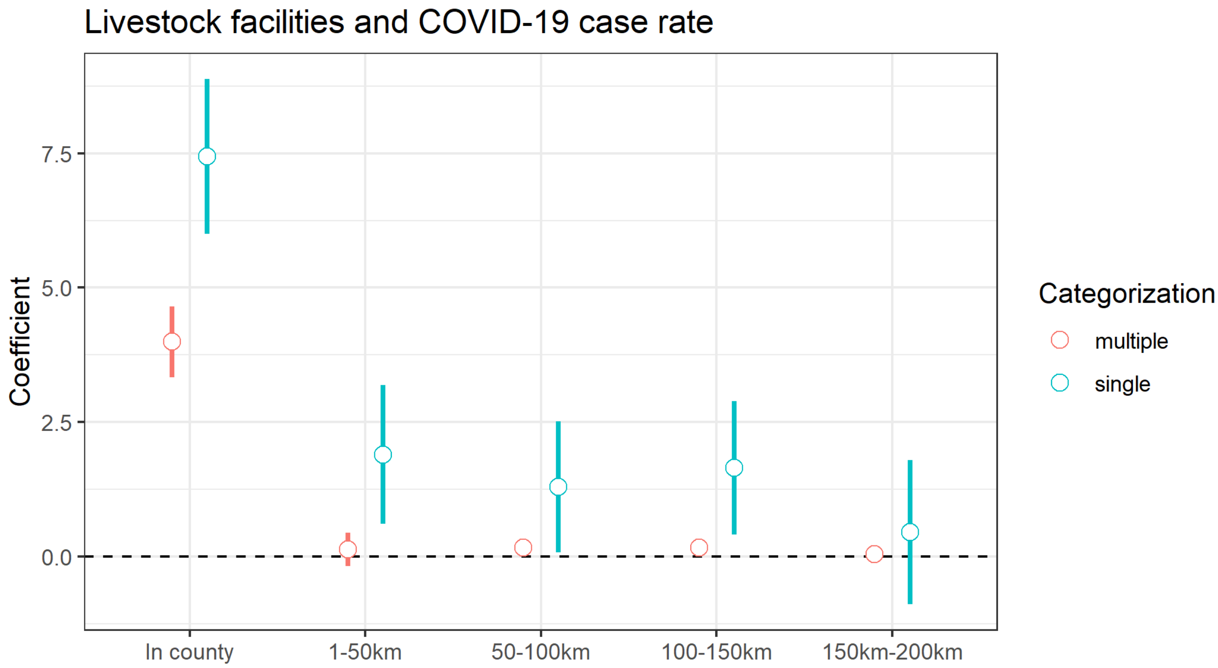

Figure 7: Relationship between COVID-19 incidence and livestock plants by distance band on cases per thousand. Coefficients plotted are based on a model that groups counties by distance from the geographical centroid of counties with a livestock facility, in addition to standard controls, state fixed effects, and standard errors are clustered at the state level. 'Single categorization' means that bands are non-overlapping: if a county contains a livestock facility and there is also a facility outside the county within 50km, the county is coded 'In county' and not '0-50km'. 'Multiple categorizations' means that counties can have facilities at multiple distance bands and thus be counted in several groups. For example, a county located 25km away from one county with a livestock facility and 75km from another will be included in both Categories 2 and 3. Error bars reflect a 95% confidence interval.

7 Tables

Table 1: Livestock facilities and county-level COVID-19 incidence

	COVID-19 incidence per 1000 as of 2020-07-21					
	...... Case rate			...... Death rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	4.49***	4.07***	5.98***	0.07***	0.07***	0.10***
	(0.88)	(0.80)	(1.14)	(0.02)	(0.02)	(0.02)
Plant count	Level	Level	Binary	Level	Level	Binary
Controls	X	X	X	X	X	X
State FE		X	X		X	X
Observations	3,032	3,032	3,032	3,032	3,032	3,032
R^2	0.36	0.45	0.46	0.27	0.42	0.42

*p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variable is COVID-19 cases (models 1 to 3) and deaths (models 4 to 6) per thousand. Livestock facility level is the sum of beef, pork, and poultry plants in the county. Livestock facility binary denotes a binary variable representing whether a county has at least one livestock plant. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in models 2, 3, 5, and 6. Standard errors are clustered at the state level.

Table 2: Livestock facilities and county-level COVID-19 incidence, all covariates

	COVID-19 incidence per 1000 as of 2020-07-21			
	...... Case rate		...... Death rate	
	(1)	(2)	(3)	(4)
Livestock facility	4.49*** (0.88)	4.07*** (0.80)	0.07*** (0.02)	0.07*** (0.02)
Income per capita (log)	−0.11 (1.06)	1.04 (0.71)	0.06 (0.06)	0.04 (0.03)
Density	−0.15 (0.67)	0.70 (0.57)	0.005 (0.04)	−0.01 (0.02)
Density-squared	0.02 (0.07)	−0.06 (0.05)	0.0003 (0.01)	0.01** (0.002)
Timing first case	−7.58*** (1.58)	−5.73*** (1.29)	−0.25*** (0.06)	−0.19*** (0.04)
Elderly proportion	−18.14 (12.13)	−16.05 (9.87)	0.91** (0.36)	1.10*** (0.32)
Black proportion	24.64*** (3.02)	19.27*** (2.57)	1.00*** (0.14)	0.93*** (0.17)
Hispanic proportion	12.76*** (3.85)	21.04*** (6.66)	0.06 (0.07)	0.18 (0.12)
Freight intensity	−0.19** (0.08)	−0.07 (0.07)	−0.001 (0.01)	0.01 (0.01)
Public transit proportion	0.22 (0.14)	0.18 (0.13)	0.03*** (0.01)	0.02*** (0.01)
Household size	2.89* (1.56)	1.41 (1.59)	0.21*** (0.07)	0.17*** (0.06)
Nursing home proportion	115.64** (47.73)	49.60 (30.79)	6.39*** (1.86)	4.49*** (1.48)
Prisoner proportion	15.44 (9.82)	22.13** (10.08)	−0.12 (0.23)	−0.08 (0.20)
Uninsured proportion	10.67 (8.27)	25.45*** (7.19)	−0.19 (0.23)	0.31 (0.25)
Frontline proportion	0.60 (1.24)	1.12 (1.24)	−0.04 (0.05)	−0.04 (0.05)
Plant count	Level	Level	Level	Level
Controls	X	X	X	X
State FE		X		X
Observations	3,032	3,032	3,032	3,032
R^2	0.36	0.45	0.27	0.42

*p<0.1; **p<0.05; ***p<0.01

Expanded presentation of baseline model in Table 1 with all control covariates shown.

Table 3: Livestock facilities and COVID-19 arrival timing

	Timing of COVID-19 emergence, as of 2020-07-21					
	Days to 1 case			Days to 10 cases		
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	-1.11^{**}	-1.12^{**}	-1.08	-5.12^{***}	-4.65^{***}	-6.71^{***}
	(0.44)	(0.42)	(0.66)	(1.03)	(1.08)	(1.54)
Plant	Level	Level	Binary	Level	Level	Binary
Controls	X	X	X	X	X	X
State FE		X	X		X	X
Observations	3,032	3,032	3,032	2,768	2,768	2,768
R^2	0.47	0.51	0.51	0.51	0.57	0.58

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Regression model estimating the relationship between a livestock facility and the number of days elapsed since January 1, 2020 (Julian day) until the first confirmed case (Models 1-3) and the 10th case (Models 4-6). Livestock facility level is the sum of beef, pork, and poultry plants in the county. Livestock facility binary denotes a binary variable representing whether a county has at least one livestock plant. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in Models 2, 3, 5, and 6. Standard errors are clustered at the state level.

Table 4: Livestock facility by type and county-level COVID-19 incidence

	COVID-19 incidence per 1000 as of 2020-07-21					
	...... Case rate			...... Death rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Beef plants	3.89**	3.68**	4.79**	0.11***	0.11***	0.14***
	(1.67)	(1.47)	(2.06)	(0.04)	(0.03)	(0.04)
Pork plants	6.58***	6.00***	7.26***	0.07*	0.07**	0.09**
	(2.28)	(1.82)	(2.40)	(0.04)	(0.03)	(0.04)
Poultry plants	3.68***	3.18***	3.38***	0.04**	0.05**	0.06**
	(0.71)	(0.71)	(0.83)	(0.02)	(0.02)	(0.03)
Plant count	Level	Level	Binary	Level	Level	Binary
Controls	X	X	X	X	X	X
State FE		X	X		X	X
Observations	3,032	3,032	3,032	3,032	3,032	3,032
R^2	0.36	0.46	0.46	0.27	0.42	0.42

Note: *p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variables are COVID-19 cases (Models 1-3) and deaths (Models 4-6) per thousand. For 'Level', beef plants, pork plants, and poultry plants are the number of the respective plants in a county. For 'Binary', dummy variables taking the value of one if a county contains, respectively, the respective plant. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State fixed effects included in Models 2, 3, 5, and 6. Standard errors are clustered at the state level.

Table 5: Livestock facilities and county-level COVID-19 cases, by size

	COVID-19 incidence per 1000 as of 2020-07-21			
	...... Case rate		...... Death rate	
	(1)	(2)	(3)	(4)
Livestock facility (small)	−0.32	−0.04	0.02	0.004
	(0.23)	(0.17)	(0.01)	(0.01)
Livestock facility (medium)	1.02	0.98	0.07	0.06*
	(0.87)	(0.84)	(0.04)	(0.04)
Livestock facility (large)	6.06***	5.48***	0.06***	0.07***
	(1.18)	(1.14)	(0.02)	(0.02)
Controls	X	X	X	X
State FE		X		X
Observations	3,032	3,032	3,032	3,032
R^2	0.37	0.46	0.27	0.42

Note: *p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variables are COVID-19 cases (Models 1-2) and deaths (Models 3-4) per thousand. Livestock facilities are the sum of beef, pork, and poultry plants in the county, split into three separate variables, small, medium, and large, which take the value of 1 for USDA FSIS Categories 3, 4, and 5, respectively. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in Models 2 and 4. Standard errors are clustered at the state level.

Table 6: Poultry facility and county-level COVID-19 incidence, by line speed waiver

	COVID-19 incidence per 1000 as of 2020-07-21			
	...... Case rate		...... Death rate	
	(1)	(2)	(3)	(4)
Poultry plant	2.422***	2.729***	0.049	0.057**
	(0.748)	(0.825)	(0.030)	(0.027)
Poultry plant:Waivers	2.426**		0.028	
	(1.022)		(0.029)	
Poultry plant:Waivers 2020		4.895**		0.021
		(2.113)		(0.046)
Plant count	Binary	Binary	Binary	Binary
Beef-pork controls	X	X	X	X
Controls	X	X	X	X
State FE	X	X	X	X
Observations	3,032	3,032	3,032	3,032
R^2	0.458	0.458	0.420	0.420

*p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variables are COVID-19 cases (Models 1-2) and deaths (Models 3-4) per thousand. Livestock plants (poultry, beef, and pork included separately) are indicated by a binary variable representing whether a county has at least one such plant. Beef and pork plants are controlled for but omitted from the output. Poultry plant:Waivers denotes the interaction of poultry plant counties and USDA line speed waivers. Poultry plant:Waivers 2020 is limited to waivers granted in 2020. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in all models. Standard errors are clustered at the state level.

Table 7: Livestock facilities and county-level COVID-19 cases, by company

	COVID-19 incidence per 1000 as of 2020-07-21			
	...... Case rate		...... Death rate	
	(1)	(2)	(3)	(4)
National Beef	18.09***	16.32**	0.31	0.27
	(6.62)	(6.09)	(0.21)	(0.21)
JBS	8.61***	7.47***	0.08	0.09*
	(2.41)	(2.48)	(0.05)	(0.05)
Tyson	7.71***	6.71***	0.07	0.10***
	(2.33)	(2.15)	(0.05)	(0.04)
Cargill	5.71**	5.40***	0.05	0.09
	(2.25)	(1.90)	(0.07)	(0.07)
Smithfield	5.16**	4.65***	−0.04	−0.03
	(2.20)	(1.63)	(0.05)	(0.03)
Other	3.68***	3.28***	0.08**	0.07**
	(0.78)	(0.76)	(0.03)	(0.03)
Plant count	Binary	Binary	Binary	Binary
Controls	X	X	X	X
State FE		X		X
Observations	3,032	3,032	3,032	3,032
R^2	0.37	0.46	0.27	0.42

Note: *p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variables are COVID-19 cases (Models 1-2) and deaths (Models 3-4) per thousand. Covariates include a factor variable of counties with a livestock plant (USDA FSIS category 4 and 5) by company ownership, implicitly a binary. In the rare occurrence that multiple companies have plants in the same county, the county is assigned to the company with less total plants. 'Other' refers to all counties with livestock plants not owned by any of the companies shown. The omitted category is counties without livestock plants. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in Models 2 and 4. Standard errors are clustered at the state level.

Table 8: Livestock facility and mobility patterns from Google

	Change in time use from pre-March 13 baseline			
	Workplace		Retail and recreation	
	(1)	(2)	(3)	(4)
Livestock facility	0.97***	0.82***	0.49*	0.50*
	(0.22)	(0.21)	(0.28)	(0.27)
Controls	X	X	X	X
State FE		X		X
Observations	2,826	2,826	2,721	2,721
R^2	0.33	0.41	0.20	0.27

Note: *p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variable is the average change in Google's mobility-based index of people's time spent visiting workplaces (Models 1-2) and engaging in retail and recreation activities (Models 3-4) in the four weeks following March 13, 2020, relative to a baseline level set prior to March 13, 2020. Livestock facility is the sum of beef, pork, and poultry plants in the county. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. Some counties do not have mobility estimates from Google during this time period due to lack of raw data. State-level fixed effects are included in Models 2 and 4. Standard errors are clustered at the state level.

Table 9: COVID-19 testing, livestock facilities, and COVID-19 incidence

	Dependent variable:							
	...... Case rate				...... Death rate			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Livestock facility	4.07***	4.30***	4.19***	4.19***	0.07***	0.06**	0.06**	0.06**
	(0.80)	(1.23)	(1.21)	(1.20)	(0.02)	(0.03)	(0.03)	(0.03)
Testing per 1000			0.01*	0.01*			0.0001**	0.0001**
			(0.003)	(0.003)			(0.0000)	(0.0000)
Positivity rate				0.86**				0.02**
				(0.38)				(0.01)
Controls	X	X	X	X	X	X	X	X
State FE	X	X	X	X	X	X	X	X
Observations	3,032	1,773	1,773	1,773	3,032	1,773	1,773	1,773
R^2	0.45	0.44	0.45	0.45	0.42	0.44	0.44	0.44

*p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county-level data from 31 states with livestock facilities and available data on county-level testing gathered from 31 state health departments. Dependent variables are COVID-19 cases (Models 1-4) and deaths (Models 5-8) per thousand in these states. Livestock facility is the sum of beef, pork, and poultry plants in the county. Testing per thousand represents the number of tests taken per thousand people in these states as of July 14, 2020. Positivity rate is total cases divided by total tests. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in all models. Standard errors are clustered at the state level.

Table 10: Manufacturing and county-level COVID-19 cases

	Dependent variable:					
	COVID-19 cases per 1000 as of 2020-07-21					
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	4.49***	4.35***	4.24***	4.07***	4.08***	3.97***
	(0.88)	(0.35)	(0.35)	(0.80)	(0.80)	(0.79)
Manufacturing establishments		0.002*			−0.0004	
		(0.001)			(0.001)	
Manufacturing income share			5.46***			4.23
			(1.43)			(2.67)
Controls	x	x	X	X	X	X
State FE				X	X	X
Observations	3,032	3,032	3,032	3,032	3,032	3,032
R^2	0.36	0.37	0.38	0.45	0.45	0.46

Note: *p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data. Dependent variable is COVID-19 cases per thousand. Livestock facility is the sum of beef, pork, and poultry plants in the county. Manufacturing establishments is the number of such establishments in a county, and manufacturing income share is a county's share of total income from manufacturing. Models 1 and 4 replicated from Table 1 for reference. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State fixed effects included in Models 4-6. Standard errors are clustered at the state level.

Table 11: Livestock facilities and county-level COVID-19 incidence, subset counties within 100km of plants

	COVID-19 incidence per 1000 as of 2020-07-21					
	...... Case rate			...... Death rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	5.27***	4.30***	4.82***	0.10***	0.09***	0.10***
	(1.19)	(1.16)	(1.21)	(0.03)	(0.03)	(0.03)
Plant count	Level	Level	Binary	Level	Level	Binary
Controls	X	X	X	X	X	X
State FE		X	X		X	X
Observations	1,187	1,187	1,187	1,187	1,187	1,187
R^2	0.41	0.54	0.53	0.35	0.49	0.49

*p<0.1; **p<0.05; ***p<0.01

Relationship between livestock facilities and COVID-19 cases and deaths per thousand. Replicating the baseline model in Table1 but only includes counties within 100km of a livestock facility.

Table 12: Livestock facilities and county-level COVID-19 incidence, non-linear transforms

	COVID-19 incidence per 1000 as of 2020-07-21					
	...... Case rate			...... Death rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	4.07***	0.31***	0.32***	0.07***	0.06***	0.25***
	(0.80)	(0.05)	(0.05)	(0.02)	(0.01)	(0.05)
Mean rate	8.06	8.06	8.06	0.21	0.21	0.3
Plant count	Level	Level	Level	Level	Level	Level
Transform	None	IHS	Log	None	IHS	Log
Controls	X	X	X	X	X	X
State FE	X	X	X	X	X	X
Observations	3,032	3,032	3,032	3,032	3,032	2,080
R^2	0.45	0.63	0.62	0.42	0.44	0.42

*p<0.1; **p<0.05; ***p<0.01

Relationship between livestock facilities and COVID-19 cases and deaths per thousand. All models replicate the baseline model in Table 1 using nonlinear transformations. In relation to the dependent variable, ‘IHS’ denotes the use of an inverse hyperbolic sine transformation, and ‘Log’ denotes the use of a natural log (with zero values dropped).

Table 13: Balance table of counties with and without livestock plants

Covariate	No plant	Plant	t_value
Counties	$2,813$	259	
Case rate	7.20	16.20	9.99
Death rate	0.19	0.34	5.67
Income (log)	10.66	10.65	-0.68
Density	1.21	1.32	2.79
Elderly	0.13	0.12	-11.01
Black	0.10	0.13	3.43
Hispanic	0.09	0.15	5.56
Freight	0.39	0.94	1.99
Public transit	0.83	0.83	0.04
Household size	2.51	2.62	7.73
Nursing	0.01	0.01	-2.48
Prison	0.02	0.01	-2.49
Uninsured	0.10	0.11	2.36
Frontline	0.17	0.07	-17.89

Summary statistics of covariates used in all models. Columns ‘No plant’ and ‘Plant’ are mean covariate values of counties with and without livestock plants, and ‘t-value’ denotes the test statistic to discern whether there is a statistically significant difference.

Table 14: IV - first stage: livestock plant location and livestock sales in 1959

	Dependent variable:			
	Livestock facilities			
	(1)	(2)	(3)	(4)
Livestock sales 1959	0.015***	0.016***	0.019***	0.018***
	(0.002)	(0.002)	(0.002)	(0.002)
F-stat	87.3	37.1	70.8	16.5
Controls		X		X
State FE			X	X
Observations	3,032	3,032	3,032	3,032
R^2	0.076	0.109	0.123	0.144

Note: *p<0.1; **p<0.05; ***p<0.01

First stage IV regression model. Dependent variable is current number of livestock facilities in a county, which is the sum of beef, pork, and poultry plants. Livestock sales 1959 is the county-level sales of agricultural livestock products in 1959 from the USDA census. Models 2 and 4 include the same controls from the baseline specification, including income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects included in Models 3-4. Standard errors clustered at the state level.

Table 15: Livestock facilities and county-level COVID-19 incidence, IV

	Dependent variable:			
	...... Case rate		...... Death rate	
	(1)	(2)	(3)	(4)
Livestock facility	9.00***	6.12***	0.13*	0.06
	(2.80)	(1.43)	(0.07)	(0.06)
Controls	X	X	X	X
State FE		X		X
Observations	3,032	3,032	3,032	3,032
R^2	0.33	0.45	0.27	0.42

*p<0.1; **p<0.05; ***p<0.01

Regression model with an instrument for the presence of a livestock plant in a county using the county's livestock production value in 1959 in terms of animals sold. Livestock facility is the sum of beef, pork, and poultry plants in the county. Controls include income per capita (log), density (population per built-up land area) and density squared, number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in Models 2 and 4. Standard errors are clustered at the state level.

Table 16: Comparison of counties with and without livestock plants, by quartile of plant propensity score

t-value	q1	q2	q3	q4
Counties	758	758	758	758
Income (log)	−0.42	-0.62	2.62	-1.41
Density	-2.11	0.94	1.18	-1.21
Elderly	-0.89	-0.36	0.88	-1.36
Black	-4.42	2.32	0.35	-2.36
Hispanic	0.85	0.34	0.66	2.10
Freight	0.76	0.81	1.59	0.67
Public transit	-1.83	1.14	-0.18	-1.69
Household size	0.90	1.72	-0.89	1.43
Nursing	-0.25	0.14	-0.54	0.66
Prison	-0.14	1.42	-2.28	-1.09
Uninsured	-0.47	0.12	1.02	0.64
Frontline	-1.74	1.02	-3.14	0.74
Plant #	3	31	63	161

Balance table for propensity score matching analysis. T-value denotes the test statistic to discern whether there is a statistically significant difference in covariate values between counties with and without livestock plants, grouped by quartile of plant propensity score.

Table 17: Average values of subsample after using the nearest neighbor matching algorithm

Covariate	No plant	Plant	t_value
Counties	516	258	
Income (log)	10.66	10.65	-0.46
Density	1.30	1.32	0.54
Elderly	0.11	0.12	0.33
Black	0.13	0.13	-0.18
Hispanic	0.13	0.15	0.95
Freight	0.69	0.94	0.87
Public transit	0.79	0.84	0.29
Household size	2.60	2.63	1.15
Nursing	0.01	0.01	-0.43
Prison	0.01	0.01	-0.01
Uninsured	0.10	0.11	0.57
Frontline	0.07	0.07	0.43

Summary statistics of matching datasets constructed using a nearest neighbors algorithm. Matching performed here at a 2:1 ratio, reflected in there being twice the number of counties without a plant as there are counties with one. Columns 'No plant' and 'Plant' are mean covariate values of counties with and without livestock plants, and 't-value' denotes the test statistic to discern whether there is a statistically significant difference.

Table 18: Livestock facilities and county-level COVID-19 incidence, matching methods

	Dependent variable:					
	...... Case rate			...... Death rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Livestock facility	4.07***	3.95***	3.73***	0.07***	0.07***	0.08***
	(0.80)	(0.74)	(0.82)	(0.02)	(0.02)	(0.02)
Method	None	Propensity	NN	None	Propensity	NN
Controls	X	X	X	X	X	X
State FE	X	X	X	X	X	X
Observations	3,032	3,032	774	3,032	3,032	774
R^2	0.45	0.32	0.44	0.42	0.31	0.41

*p<0.1; **p<0.05; ***p<0.01

Regression model with cross-sectional county data using three different empirical models. Dependent variables are COVID-19 cases (Models 1-3) and deaths (Models 4-6) per thousand. Livestock facility is the sum of beef, pork, and poultry plants in the county. Method 'None' replicates Models 2 and 5 of the baseline specification in 1 for reference. Method 'Propensity' uses a propensity score matching model. Method 'NN' uses of a nearest-neighbor score matching model. Controls include income per capita (log), density (population per built-up land area) and density squared, the number of freight miles traveled, timing of first case (index of Julian day of first confirmed case), as well as proportions of the county population over the age of 70, Black, Hispanic, public transit commuters, uninsured, frontline workers, or in nursing homes or prisons. State-level fixed effects are included in all models. Standard errors are clustered at the state level.

Table 19: Counties by distance band from livestock facility (multiple categorizations)

	In county	0-50km	50-100km	100-150km	150-200km
Counties	259	832	1,816	2,219	2,422
Pop total	51,759,204	86,263,276	208,939,193	237,238,573	248,145,926
Pop avg	199,842	103,682	115,055	106,912	102,455
Density	1.32	1.37	1.33	1.30	1.28
Plants/county	1.24	0.28	0.16	0.13	0.12
Cases total	762,534	1,151,008	2,544,241	2,879,670	2,900,812
Deaths total	19,964	53,266	96,552	106,441	106,053
Case rate (county)	16.20	10.71	9.36	8.83	8.47
Case rate (pop)	14.73	13.34	12.18	12.14	11.69
Death rate (county)	0.34	0.28	0.25	0.23	0.22
Death rate (pop)	0.39	0.62	0.46	0.45	0.43

Table 20: Counties by distance band from livestock facility (single categorization)

	In county	0-50km	50-100km	100-150km	150-200km
Counties	259	711	1,027	517	219
Pop total	51,759,204	70,351,016	116,349,269	49,127,602	22,349,326
Pop avg	199,842	98,947	113,290	95,024	102,052
Density	1.32	1.36	1.31	1.16	1.05
Cases total	762,534	912,749	1,259,490	509,034	248,417
Deaths total	19,964	51,331	42,983	17,547	5,631
Case rate (county)	16.20	8.63	7.51	6.99	5.80
Case rate (pop)	14.73	12.97	10.83	10.36	11.12
Death rate (county)	0.34	0.24	0.21	0.16	0.14
Death rate (pop)	0.39	0.73	0.37	0.36	0.25

Table 21: Counties by distance band from livestock facility (single categorization), NYC dropped

	In county	0-50km	50-100km	100-150km	150-200km
Counties	259	710	1,027	517	219
Pop total	51,759,204	61,952,268	116,349,269	49,127,602	22,349,326
Pop avg	199,842	87,257	113,290	95,024	102,052
Density	1.32	1.34	1.31	1.16	1.05
Cases total	762,534	685,970	1,259,490	509,034	248,417
Deaths total	19,964	28,436	42,983	17,547	5,631
Case rate (county)	16.20	8.60	7.51	6.99	5.80
Case rate (pop)	14.73	11.07	10.83	10.36	11.12
Death rate (county)	0.34	0.24	0.21	0.16	0.14
Death rate (pop)	0.39	0.46	0.37	0.36	0.25

Summary statistics for counties grouped by distance from geographical centroid to the nearest livestock facility. The band '0–50km' excludes the county itself. 'Multiple categorizations' means that counties can have facilities at multiple distance bands and thus be counted several times. For example, a county located 25km away from one county with a livestock facility and 75km from another will be included in both Categories 2 and 3. 'Single categorization' means that bands are non-overlapping: if a county contains a livestock facility and there is also a facility outside the county within 50km, the county is coded 'In county' and not '0-50km'. Rows represent the number of counties in each band, total population, mean population by county, density by county, number of plants per county, cumulative COVID-19 cases and deaths, equal-weighted county-level mean case rate, population-weighted county-level mean case rate, equal-weighted county-level mean death rate, and population-weighted county-level mean death rate. Table 21 drops all counties within New York City from the summary.

Table 22: Summary of baseline COVID-19 effects as of 2020-07-21

Summary	Cases	Deaths
COVID-19 incidence	3,868,989	141,200
Avg rate (pop)	11.91	0.43
Avg rate (county)	7.96	0.20
Coef (level)	4.07	0.07
Coef (binary)	5.98	0.10
County percent effect (level)	0.51	0.37
County percent effect (binary)	0.75	0.50
Plants ——————————-		
Total plants	322	322
Counties with plants	259	259
Population of plant counties	51,759,204	51,759,204
Population (county avg)	199,842	199,842
Population per plant (county avg)	180,168	180,168
Impact ——————————-		
Total impact (level)	236,265	4,313
Total impact (binary)	309,572	5,228
Percent of total (level)	0.06	0.03
Percent of total (binary)	0.08	0.04

Table depicting the intermediate calculations used to arrive at estimates of the excess COVID-19 cases and deaths in the US attributable to livestock processing. COVID-19 incidence is the cumulative number of positive COVID-19 cases or deaths. Avg rate (pop) and avg rate (county) denote respectively the population-weighted and equal-weighted county-level means of cases and deaths. Coef (level) and coef (binary) denote respectively the calculated livestock plant coefficient in the level and binary model specifications. County percent effect presents the percentage increase in the calculated livestock plant coefficient over the equal-weighted county-level means. Total plants denotes the number of all plants counted in our dataset. Counties with plants represents the number of counties in our analysis that have at least one plant. Population of plant counties denotes the sum population of counties with plants. Population (county avg) is the total population in these counties divided by the number of counties. Population per plant (county avg) is calculated by dividing individual counties' populations per plant by the number of counties with at least one plant. Total impact (level) is the estimated number of excess cases or deaths, calculated by multiplying the level coefficient by the average population per plant, multiplied by the total plants. Total impact (binary) is the estimated number of excess cases or deaths, calculated by multiplying the binary coefficient by the average population per county, multiplied by the total number of counties. Percent of total for level and binary specifications are the ratio of the total impact calculated to the total number of cases or deaths.