

Do Methane Mitigation Incentives Intensify Livestock Production? Evidence from California, 2016-2025

Varun Magesh^{1,2†}, Claire M. Morton^{1,3†}, Qinfan Lyu⁴,
Jamie Harding⁴, Frank C. Curriero⁵, Cara Wychgram⁵,
Sara N. Lupolt^{4,6,7}, Keeve E. Nachman^{4,6,7}, Elena Eneva¹,
Daniel E. Ho^{1*}

^{1*}Stanford University RegLab, 559 Nathan Abbott Way, Stanford,
94305, California, USA.

²Harvard Law School, 1585 Massachusetts Avenue, Cambridge, 02138,
Massachusetts, USA.

³University of California, Berkeley, Department of Statistics, Evans
Hall, University Dr., Berkeley, 94720, California, USA.

⁴Johns Hopkins University, Center for a Livable Future, 111 Market
Place, Baltimore, 21202, Maryland, USA.

⁵Johns Hopkins Bloomberg School of Public Health, Department of
Epidemiology, 615 North Wolfe Street, Baltimore, 21205, Maryland,
USA.

⁶Johns Hopkins Bloomberg School of Public Health, Department of
Environmental Health and Engineering, 615 North Wolfe Street,
Baltimore, 21205, Maryland, USA.

⁷Risk Sciences and Public Policy Institute, Johns Hopkins Bloomberg
School of Public Health, 615 North Wolfe Street, Baltimore, 21205,
Maryland, USA.

*Corresponding author(s). E-mail(s): deho@stanford.edu;

†These authors contributed equally to this work.

Abstract

Public subsidies for methane mitigation infrastructure aim to combat climate change, but they may perversely incentivize polluting industries' growth. We

assess biogas capture incentives in California using facility data and satellite imagery in a matched sample design. Results show incentive anticipation and receipt increases facility growth by 860 (SE 200) mature dairy cows relative to matched controls over three years, suggesting that methane reduction incentives inadvertently drive facility expansion.

1 Main

Methane is a potent greenhouse gas with a global warming potential approximately 28 times greater than carbon dioxide over a 100-year period [1]. Global methane levels are rising steadily [1]. Food system emissions make up one third of global anthropogenic greenhouse gas emissions; 35% of these food system emissions are from methane [2]. Ruminant digestion and manure management from cattle, sheep, and other livestock are responsible for the vast majority of these methane emissions [3].

An emerging policy response to rising methane emissions from industrial and agricultural sources is public subsidy of investments in methane mitigation and capture infrastructure [4–8]. The World Bank is administering a \$234 million fund to issue grants and financing to abate oil & gas methane emissions through infrastructure maintenance and construction [8]. In the United States, the Inflation Reduction Act allocated over \$1.36 billion for methane reduction activities [5]. California has implemented policies to mitigate agricultural methane sources by creating substantial financial incentives for anaerobic dairy methane digesters, positioning the state as a leader in agricultural methane capture policy. In 2024 alone, biogas producers (nearly all dairy farms) received \$483 million through biogas credit sales [6]. Other states in the United States have adopted similar policies [7].

Prior policy analyses warn or conjecture that incentivization of digesters causes expansion of the largest dairy farms by making additional cows more profitable [9–12]. Yet to date, these concerns have been driven by anecdotal evidence [9–12].

We provide the first empirical statistical estimate of the relationship between mitigation incentives and the expansion of emitting facilities using evidence from dairy farms in California. We proceed by (1) introducing relevant policies, (2) describing our research approach, (3) analyzing farm expansion, and (4) estimating expansion-related methane emissions.

California has created two key policies to incentivize agricultural methane capture: an up-front grant program and a continuing crediting program. California’s Dairy Digester Research and Development Program (DDRDP) is administered by the California Department of Food & Agriculture (CDFA) and provides up-front grants for dairy methane mitigation infrastructure. California’s Low Carbon Fuel standard (LCFS) is administered by the California Air Resources Board (CARB) and provides credit incentives per unit of captured methane. Both policies incentivize anaerobic dairy digesters, a methane capture technology that traps methane emitted by so-called ‘lagoons’ of dairy manure. Digesters require a large up-front investment with smaller operational costs. They produce biogas that generates LCFS credits when it is used as

transportation fuel, either through direct combustion or by generating power for electric vehicles. The DDRDP provides grants for dairies to construct anaerobic digesters each year, with facilities receiving an average of \$1,621,590 to begin building a digester up to two years after the grant is issued [13]. The LCFS provides credits per unit of methane used in the state’s transportation supply [14]. CARB treats dairy farms’ high methane emissions as an allowable baseline, so methane captured by digesters has a very low carbon intensity score [6]. Producers of fuels with higher carbon intensity scores are obligated by law to purchase credits given to low-carbon fuel producers, creating a market for these credits [14]. This creates a financial incentive for dairies to create and capture additional units of methane. On one estimate, methane credit sales make up approximately 40% of the revenue generated from each dairy cow on a farm with a digester [15].

Our research approach collates pre-existing datasets to locate dairies and identify those with digesters, statistical matching techniques to identify comparable dairies without digesters, and nine years of human-labeled satellite imagery to measure the growth of treated and control dairies. We collate data from California air, water, and agricultural regulatory agencies, the United States Department of Agriculture, and Cal-FF, a satellite imagery-based dataset of California factory farms [13, 16–18]. We use these records to identify dairies with and without digesters, then apply statistical matching methods to select an appropriate control group from the set of dairies without digesters. Control group facilities are selected to have comparable size and scale to digester facilities before they are treated — before digester planning begins. We annotate nine years (2016-2025) of quarterly satellite imagery that shows the infrastructure footprint of treated and control facilities and the way that it changes over time. We annotate 134 facilities with digesters (every digester planned after 2018) and 103 comparable control facilities that do not build digesters. The 237 facilities we annotate are approximately 17% of all 1,345 confirmed dairies in California. We annotate each facility once per quarter from Q1 2016 to Q1 2025, annotating a total of 8,769 images. Fig. 1 depicts our research process from data preparation to annotation to analysis.

We analyze dairy expansion by using annotated satellite data to compare changes between treated and untreated facilities. We use the barn area of each facility as the outcome of interest because barn area closely correlates with the number of mature dairy cows at a facility. For this analysis, we define the treatment date for each facility as the date a facility begins anticipating digester construction. This is because we want to capture the effect of anticipatory expansion; facilities may grow before a digester is complete knowing that they will earn greater returns in the near future. For most facilities, we use the date that their DDRDP construction grant application was approved, usually 3 years before the digester is operational.

We perform the analysis using the treatment effect estimation approach developed by Callaway and Sant’Anna for difference-in-difference causal estimates in settings with staggered treatment over time and never-treated units [19]. We convert barn area to estimated herd size. We use a regression of reported mature dairy cow numbers from the California Dairy and Livestock Database [20] against the barn area observed

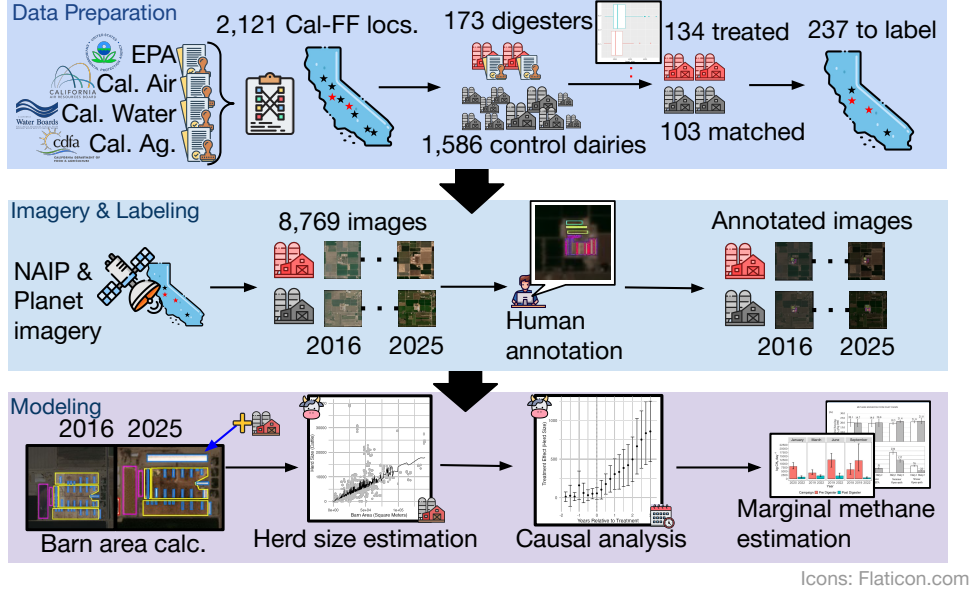


Fig. 1 An overview of our research process. We combine administrative records (EPA, USDA, and California state agencies) and Cal-FF, a dataset of California factory farms collected from satellite data, to identify dairies with and without digesters. We match comparable facilities to build a control group, then annotate thousands of images from 2016–2025 to observe facility growth over time. These data feed into causal models that estimate the relationship between digester adoption and facility scale and methane emissions.

in our annotated facility data to estimate the conversion factor between barn area and mature dairy herd size.

We estimate that the presence of a digester and attendant incentives are associated with an increase in barn area by 5,398 (SE 1,423) square meters three years after digester construction is anticipated. This is equivalent to 860 (SE 200) additional mature dairy cows housed three years after digester construction is anticipated. Fig. 2 shows our estimated treatment effect. This effect is consistent with a causal relationship between anticipated digester incentives and dairy expansion. We identify a strong comparable set of control facilities, demonstrate that they are similar to treated facilities prior to treatment, and show that treated facilities diverge and expand more quickly when they begin anticipating digester construction.

We estimate the methane emissions impact of this effect using a bottom-up approach based on expected methane emissions per additional cow. We use EPA-methodology measurements of methane emissions on California dairy farms [21], estimates of the CO₂-equivalence of methane emissions [2], and measurements of the methane reductions attributable to dairy digesters [22] to estimate the net increase in methane attributable to farm expansion caused by dairy digester incentives. We propagate uncertainties in inputs and coefficients using linear error propagation theory as implemented in Python’s `uncertainties` package [23].

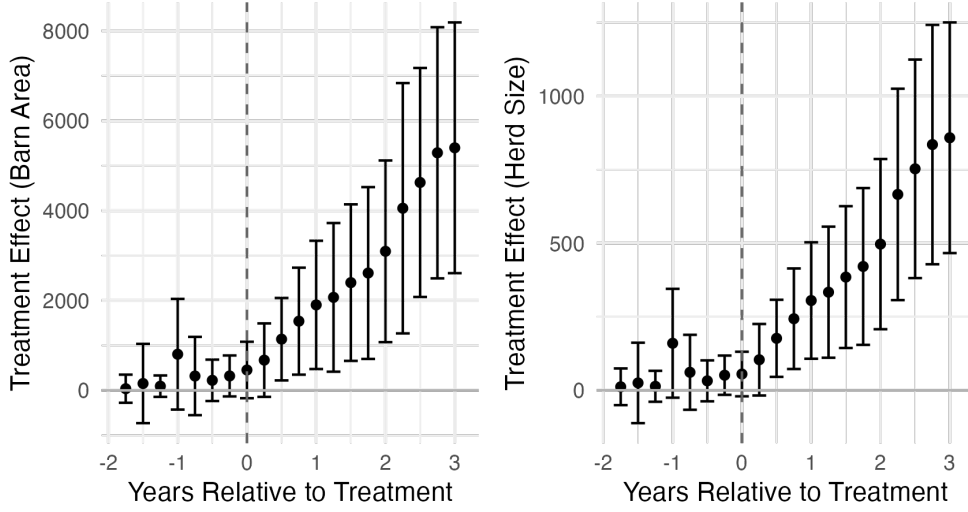


Fig. 2 Treatment effect by quarter, estimated using Callaway and Sant’Anna’s method for differences-in-differences causal estimates in settings with staggered treatment over time and never treated units. The treatment effect for barn area is shown on the left, and the treatment effect for herd size, calculated using an empirical conversion factor with uncertainty propagation, is shown on the right. The uncertainty bars shown are 95% confidence intervals.

Using this approach, we estimate that a dairy that is 860 (SE 200) cows larger generates 5,866 (SE 1,899) additional metric tons of CO₂-equivalent methane emissions per year. For comparison, the mean DDRDP application claims an emissions reduction of 17,563 metric tons of CO₂ equivalent methane per year. One limitation of our estimate is that the conversion factors we use are limited to methane emissions at the dairy itself. They do not include the impact of growing additional cow feed — land use change to support feed production, the feed production itself, and the transport of that feed to the farm. These are important sources of lifecycle emissions. Even without taking this into account, our approximation shows that facility expansion causes emissions that undercut around one third of claimed savings.

Our findings have direct implications for policymakers designing methane mitigation incentives. We provide the first empirical evidence that California’s dairy digester incentives—totaling nearly half a billion dollars in 2024 alone—increase growth of the largest emitters. Even if digesters were adopted at every U.S. animal farm and eliminated all manure-management methane emissions, this would only reduce total U.S. agricultural greenhouse gas emissions by 11% — enteric emissions from livestock and soil management make up the bulk of methane emissions [12]. Our results suggest that current incentive structures undercut these modest potential gains. Treated dairies add an estimated 860 additional mature dairy cows three years after anticipating digester construction. This expansion generates approximately 5,866 additional metric tons of

CO₂-equivalent emissions per year. These findings indicate that without explicit safeguards against facility expansion, climate incentive programs risk subsidizing growth of the very facilities they aim to regulate.

2 Methods

We collate data from California air, water, and agricultural regulatory agencies, the U.S. Department of Agriculture, and Cal-FF, a satellite imagery dataset of California factory farms [13, 16–18]. Using these records, we identify dairies with and without digesters and apply statistical matching to select controls with comparable baseline size and scale, before they begin planning to build digesters. We then annotate nine years (2016–2025) of quarterly satellite imagery showing the infrastructure footprint of treated and control facilities over time: 134 digester facilities (every digester planned after 2018) and 103 matched controls. We annotate a total of 8,769 images. Figure 1 depicts our research process from data preparation through annotation to analysis.

2.1 Administrative data collation

Our data sources build from Cal-FF, a dataset created using human-labeled satellite imagery, as ground truth information about the location and size of facilities in 2017, prior to all treatment events [17]. We connect several other sources of administrative data to Cal-FF facilities, including:

1. The California Dept. of Food & Agriculture’s Dairy Digester Resource and Development Program (DDRDP) data [13].
2. The California Air Resource Board’s California Agricultural Digester Dataset (CADD) [20].
3. The EPA’s AgSTAR dataset of dairy digesters [18].

We began by building a hand-curated dataset of AgSTAR facilities and identified facility addresses by reviewing associated permit documents and facility records (e.g., air permits from LCFS application package, facility information from the CADD). Then we linked those addresses against Cal-FF.

After performing this linkage, we found that 5 facilities that had digesters weren’t present in Cal-FF, either because they were fully constructed after 2017 (Cal-FF’s cutoff) or because of omissions in Cal-FF. We exclude the one facility constructed post-2017 from our dataset. It is of great interest that an entirely new facility was constructed and immediately implemented a digester, but it is hard to define its construction as an instance of facility expansion, so we exclude it for this analysis. For facilities omitted from Cal-FF, we manually created Cal-FF-like entries to include them. We linked CADD, an administrative dataset of California dairies, against Cal-FF to create a list of confirmed dairy facilities that did not have digesters [20].

2.2 Defining treatment as time of digester anticipation

We are interested in understanding the impact of the entire bundle of incentives that are available to facilities that construct a dairy digester, including grant and credit

programs at both the state and federal level. One of the most important and lucrative incentives is California’s Low Carbon Fuel Standard (LCFS), which increases the revenue a farmer generates per cow, providing a strong direct incentive to expand herd size [10]. The vast majority of digesters in California were built in response to this policy. These digesters were planned after 2017, when the legislature and CARB’s intent to incentivize anaerobic digesters became clear.

We want to capture the impact of the incentive on the herd sizes of dairy facilities. Facilities can expand in anticipation of getting a digester; when facilities know they will have a digester in the future, they may choose to invest in infrastructure to keep more cows, knowing their investments will pay off down the line.

In order to capture this anticipatory expansion, we define ‘treatment’ as the exact point when the paths of treated and untreated facilities separate: when the treated facilities confirm their plans to build a digester in the future. For many facilities, we have this information thanks to administrative records regarding a facility’s DDRDP grant status. Most facilities only construct digesters with grants, and the majority of facilities whose grants are denied do not go on to construct digesters. Based on this information, we estimate the ‘anticipation date’ for a facility as the date that the DDRDP grant is decided. For facilities without this data, we estimate the anticipation date as the date that the digester becomes operational, minus three years. This is based on mean anticipation-to-operational time that we observe in the populated DDRDP data.

We annotate every California facility that we are aware of that has evidence of digester construction.¹ This includes many facilities with digesters predating the LCFS policy, facilities that have been granted funds to build a digester but haven’t yet, and a handful of facilities that have digesters we noticed in aerial imagery but no administrative records of their existence. We label a total of 179 facilities with any evidence of digester construction.

2.3 Matching facilities with digesters to comparable control facilities

Only a subset of these facilities — those anticipated after 2017 — meet the treatment criteria we’ve laid out above. For each facility meeting our treatment criteria, we identify a comparable control facility which has similar characteristics in the years preceding treatment. We label these comparable control facilities so that we can compare the expansion rate of facilities with and without digesters. We match treated and control facilities based on available administrative data to decide which control facilities to label for our analysis. This matching enables us to label only the control facilities that are good comparisons for the treated digester facilities. We define treated facilities as facilities that have not cancelled their grants for digesters as of July 2025, and anticipated a digester on or after January 1, 2018. This restriction is based on the fact that information about LCFS incentives for captured methane only became available in 2018, so facilities that applied for digesters prior to that date possibly did not do so

¹There is one facility in Northern California that we examined, but did not annotate with bounding boxes. It is a very small, non-traditional dairy that has had a digester since the early 2000s, and did not have large concentrated barns to meaningfully annotate.

based on the LCFS incentives whose effect we seek to measure. Facilities that anticipated a digester prior to 2018 are not included as treatment or controls ($n = 39$). This left us with a total of 134 facilities to match to our 1,166 control facilities.

In order to apply this restriction, we obtain the timeline of a facility’s approval from California’s Dairy Digester Research & Development Program, which maintains a list of granted applications and publishes reports including application status timelines. For facilities that did not go through the DDRDP, we estimate that they anticipated digester construction three years before their digester became operational, based on the average construction timeline of DDRDP facilities. We use the MatchIt library in R [24] to implement 1:1 optimal matching based on Mahalanobis distance. This allows us to estimate the average treatment effect on the treated (ATT). We choose a matched analysis rather than simply a regression adjusted analysis because of the potential for matching to account for nonlinear relationships that linear regression does not (Stuart, 2010) (though the two are non-exclusive [25]), as well as the labor involved in labelling control facilities for further analysis. We match based on facility-year of treated and control facilities. For treated facilities, we consider the year immediately prior to the first year a digester was anticipated. We match based on: exact year, herd size in that year and the three previous years (CADD), and the number of buildings, building area, and total facility footprint in 2017, the year that policy incentives were announced (Cal-FF). We use linear interpolation to infer herd size when it is unavailable from CADD for a given facility-year. We allow control facilities to be matched in any year they are operational. This means that some control facilities are matched multiple times in different years. A total of 103 unique control facilities were matched to 134 treated facilities. Our standardized differences in covariate means all have absolute value under 0.5, the heuristic threshold suggested by Rubin (2001) [26].

2.4 Labeling satellite & aerial imagery over time

For each facility treated facility and its match, we collect empirical information about the facility’s growth and infrastructure for a nearly ten-year period spanning early 2016 through early 2025. Each facility was hand-labeled by researchers and data labelers with experience working with factory farm satellite imagery. For each facility, we capture four different kinds of infrastructure: barns, lagoons, digesters, and feedlots. For barns and digesters, which are constructed and destroyed over longer time scales, we also include construction and destruction modifiers that allow labelers to mark infrastructure that is in transition. We provided each labeler with a detailed set of annotation criteria, included as supplementary material, to identify different elements of dairy farms and resolve edge cases (like calf hutches, which we decided are not barns, or how to code digesters that are destroyed and reconstructed) in a consistent way. Fig. 3 shows an example of imagery depicting a facility changing over time with annotations completed by a labeler.

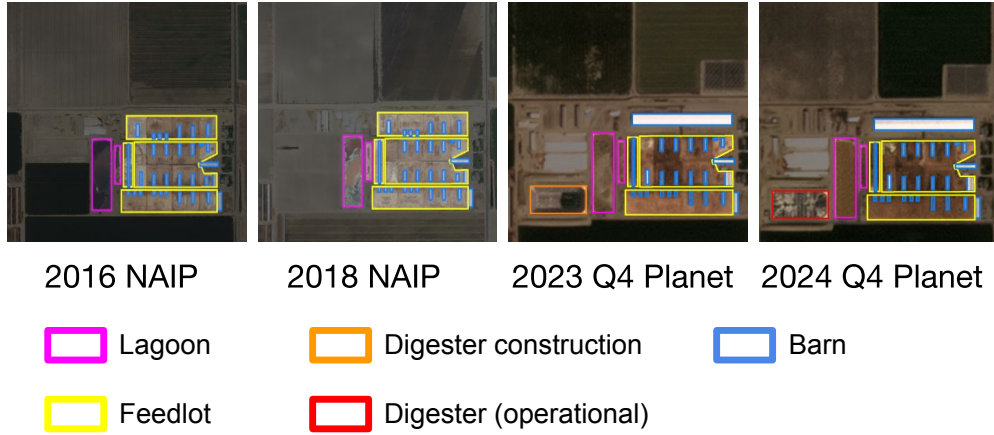


Fig. 3 An example of annotations drawn by human labelers on images over time, capturing facility expansion. Note the appearance of the digester to the bottom left and a large white barn in the center in 2023 and 2024.

2.4.1 Imagery sources

We drew on imagery from two different sources: Planet and National Agricultural Imagery Program (NAIP) data hosted by Google Earth Engine. The Planet data provided high time resolution and the NAIP data provided high imagery resolution.

Planet captures daily satellite images of the Earth’s surface. We used Planet’s quarterly basemap product, which is the best-quality composite of images taken over a three-month period. Planet’s quarterly basemaps are available on a highly regular schedule, allowing us to see the very latest developments at a facility, and to observe when changes occur with high and consistent time resolution. The resolution of planet images is lower, approximately 3m/pixel. While this sometimes makes identifying finer details difficult, most CAFO infrastructure shows up distinctively in the Planet imagery, especially since CAFOs in California tend to be quite large.

USDA’s NAIP program captures regular aerial imagery of rural and agricultural areas in the United States. Most places in California have NAIP flyover images from 2022, 2020, 2018, and 2016. Our NAIP data has submeter pixel resolution.

For each facility selected in the matching process, we captured 41 images of the facility’s bounding box plus a 1 kilometer buffer. We captured 37 Planet images, one for each quarter from Q1 2016 through Q1 2025. We captured 4 NAIP images, one for each available year within the period.

2.4.2 Labeling process

Images were labeled by authors and labelers at CloudFactory, an external data annotation service. All labelers had previous experience with CAFO imagery. Labelers used a tool called Roboflow to label images. The service provided quick and easy ways to create polygon and bounding box bounds. Importantly, the service allows you to repeat

previous labels on a new image. We took advantage of this feature to label one image, and then apply the labels throughout time, requiring new labels to be drawn only when the underlying facility changed. We labeled images in reverse chronological order to facilitate accurate labeling over time leveraging this feature. The one exception to the order was that we first labeled the latest high-resolution NAIP image (usually 2022). This helped us label barns and note smaller details in the facility. Then, we labeled in reverse chronological order, starting from Q1 2025 of Planet imagery and iterating back in time to Q1 2016. Other NAIP images appeared in the regular reverse chronological order. Labelers were allowed to move back and forth between images, allowing them to use the Planet imagery to detect changes, and the NAIP imagery to confirm the finer features of those changes. Labelers were also encouraged to observe the facility’s location in Google Earth Pro. Labelers used Google Earth Pro to access alternate sources of high-resolution satellite imagery from specific points in time to resolve ambiguous low-resolution features.

Overall, our analysis includes 8,769 annotated images of 237 different CAFO facilities in California.

We post-process annotations to associate them with the correct facilities and clean geometries. Raw annotations can contain labels for multiple facilities and invalid geometries. We post-process annotations by correctly orienting and buffering polygon geometries. We then connect each polygon with the right facility by looking for overlap with Cal-FF facilities. We apply a simple DBSCAN-like algorithm which iteratively links unmatched polygons with known neighborhoods established by Cal-FF until all polygons are associated with the correct facility.

2.5 Herd size estimation

We seek to estimate the effect of digester anticipation on herd size, because each additional cow contributes to facility GHG emissions. In order to estimate this effect, given the lack of CARB CADD data after 2022, we regress herd sizes (before and during 2022) on observed infrastructure footprint. We model the herd size of mature dairy cows simply as:

$$\text{herd_size}_y = \beta_0 + \beta_1 \cdot \text{barn_area}_{q,y} + \beta_2 \cdot \text{feedlot_area}_{q,y} \quad (1)$$

where herd_size_y is the mature cow herd size of a facility in year y , β_0 is a common intercept shared by all facilities, $\text{barn_area}_{q,y}$ is the barn area of a facility in quarter q of year y , and $\text{feedlot_area}_{q,y}$ is the feedlot area of a facility in quarter q of year y .

We choose to focus this regression on mature cow herd size for two reasons. First, we found that this quantity was more stable and tightly correlated with observed facility features. Second, mature dairy cow herd size is a good estimate of the total number of cows processed by the dairy system, since most other kinds of cows in dairies are either calves that become mature, or older cows that were once mature dairy cows.

Herd size may also vary with year (based on market conditions), lagoon area, and digester area. We chose to exclude these parameters from our herd size estimates because we wanted to use herd size estimates to measure the treatment effect of digester anticipation on herd size, and quantities like year and digester area may be endogenous to the treatment effect.

2.6 Statistical analysis of facility expansion

We analyze our matched, labeled dataset using a difference-in-differences approach proposed in Callaway and Sant’Anna (2021) [19]. First, we normalize time such that treated facilities and their corresponding control facilities have time $t = 0$ in the quarter they receive treatment (first quarter of digester anticipation). Post-treatment quarters correspond to $t \geq 0$, and pre-treatment quarters correspond to $t < 0$. We include only facilities with complete labeled data from 8 quarters pre-treatment to 12 quarters post-treatment. This effectively excludes pairs of facilities that were treated after 2022, since they do not have complete post-treatment data. This does not exclude facilities for insufficient pre-treatment data, since all facilities in 2018 and onwards and we have data from early 2016 and onwards.

Following Callaway and Sant’Anna [19], we estimate $ATT_{g,t}$, the average treatment effect on the treated (ATT) for facilities treated at calendar time g , evaluated at calendar time t , for $g, t \leq \mathcal{T} := \text{maximum calendar time}$. For all $g \neq 0, t$, $ATT_{g,t}$ is estimated using doubly robust estimators, where control units are considered to have $g = 0$, since they never received treatment. Since we are interested in the effect of receiving a digester at various event-times $e := t - g$, we aggregate the $ATT_{g,t}$ to produce ATT_e for $e \in [-8, \dots, 12]$ as shown in equation 2.

$$ATT_e = \sum_g \mathbf{1}(g + e \leq \mathcal{T}) P(G = g | G + e \leq \mathcal{T}) ATT_{g,g+e} \quad (2)$$

We calculate bootstrap standard errors clustered by facility id. When our outcome Y is herd size, which we estimate using a linear model, we propagate standard errors through 100 bootstrap replicates resampling the outcomes from a normal distribution centered at the original outcome, with standard error from the regression, and recalculating the ATT_e s on the new data. We sum the mean multiplier bootstrap variance across simulations with the variance of estimates between simulations to obtain our propagated variance estimate.

Our confidence intervals are then determined by our estimate $\pm 1.96 * SE$, where SE is either the multiplier bootstrap standard errors or, for our herd size outcome, the propagated standard errors.

2.7 Methane emission calculations

We extract baseline methane emission rates from the Arndt dataset [21], which provides measurements for two California dairy facilities across summer and winter seasons using both U.S. EPA and open-path measurement methods. The dataset contains methane emissions data in grams per animal unit (AU; a measure of live animal weight) for three facility scopes: whole facility, animal housing, and manure storage. For our calculations, we use empirical open-path measurements and use only the animal housing and manure storage components, as these allow us to differentially apply digester impacts. We model measurement uncertainty by calculating the maximum deviation between the reported mean and the upper/lower confidence bounds for each emission estimate, then propagate this uncertainty through all subsequent calculations using the first-order **uncertainties** package.

Our per-cow annual CO2-equivalent emissions calculation follows:

$$E_{cow} = [M_{manure} \cdot (1 - \eta_{digester}) + M_{housing}] \cdot \alpha_{AU} \cdot \gamma_{GWP} \cdot 365 \cdot 10^{-6} \quad (3)$$

where E_{cow} is the annual CO₂-equivalent emissions per cow (metric tons CO₂e/year), M_{manure} is the baseline manure storage methane emissions (g CH₄/AU/day) [21], $\eta_{digester}$ is the empirically observed digester methane reduction efficiency (0.82 ± 0.16) [22], $M_{housing}$ is the animal housing methane emissions (g CH₄/AU/day) [21], α_{AU} is the animal unit to cow conversion factor (1.4 cows/AU) [21], γ_{GWP} is the methane global warming potential (28 g CO₂e/g CH₄), 365 converts daily to annual emissions, and 10^{-6} converts grams to metric tons. We apply the digester reduction factor exclusively to manure storage emissions. Our final estimate averages results across all four dairy-season combinations for each measurement method, providing a representative per-cow annual climate impact for facilities operating digesters, which we then multiply by the estimated treatment effect on herd size to determine total additional emissions attributable to farm expansion incentivized by digester programs. Measurement uncertainty in each quantity was propagated using the `uncertainties` package’s implementation of first-order error propagation.

2.8 Methodological limitations & DDRDP analysis

One important limitation of our work is a potentially unobserved selection effect; facilities with an unobserved intent to expand may select into digester construction. To control for this potential selection effect, we implement a supplementary analysis in which we compare matched facilities that applied for, but were rejected from, DDRDP grants to facilities that applied for and received these grants. This allows us to quantify the effect of digester construction among only facilities that sought to construct digesters.

We matched 25 unique facilities that applied for, but were rejected from, DDRDP grants to 25 unique facilities that applied for and were accepted for the grants. Some facilities applied for the DDRDP multiple times and were accepted after initial rejection. For these facilities, we consider the outcome of the final DDRDP application and we require matched facilities to have the same final DDRDP funding round. Since all potential matched facilities received a digester, and were thus labelled, we matched based on the previous year of labelled barn area data for each facility, as well as the feedlot area at time of last application. Using this matched dataset, we performed the same analysis as in the main text. We observed that the digester anticipation did not have a significant effect on barn area or herd size (Fig. 4), but the mean effects were similar in direction and magnitude to the analysis that did not restrict to facilities that applied to the DDRDP. We attribute the lack of statistical significance to the restricted sample size of our DDRDP analysis.

Supplementary information. We report the full labeling criteria document as supplementary information. This document was provided to all labelers to ensure consistent labeling of facilities over time.

Acknowledgements. We are grateful to Allie Wainer Jazzmin Fenwick, Jackline Ogutu, Daniel Kipkosgei, James Gichuki, Nyambe Mwangelwa, Erin Maneri, Suchita

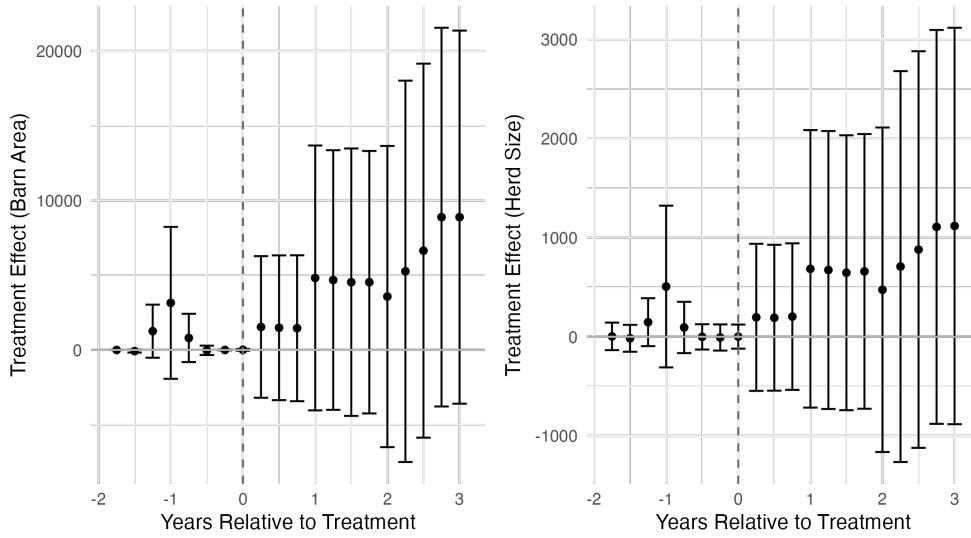


Fig. 4 Treatment effect by quarter for facilities that applied to DDRDP grants, estimated using Callaway and Sant’Anna’s method for differences-in-differences causal estimates in settings with staggered treatment over time and never treated units. Because we only observe 25 rejected facilities that never go on to build a digester, the uncertainty of our estimates is large. The treatment effect for barn area is shown on the left, and the treatment effect for herd size, calculated using an empirical conversion factor with uncertainty propagation, is shown on the right.

Sainju, members of the Stanford RegLab and the Johns Hopkins University Center for a Livable Future for annotation, feedback, and helpful conversations.

Declarations

This work was supported by the Silicon Valley Community Foundation.

References

- [1] Mar, K. A., Unger, C., Walderdorff, L. & Butler, T. Beyond CO2 equivalence: The impacts of methane on climate, ecosystems, and health. *Environmental Science & Policy* **134**, 127–136 (2022). URL <https://www.sciencedirect.com/science/article/pii/S1462901122001204>.
- [2] Crippa, M. *et al.* Food systems are responsible for a third of global anthropogenic ghg emissions. *Nature Food* **2**, 198–209 (2021).
- [3] Reisinger, A. *et al.* How necessary and feasible are reductions of methane emissions from livestock to support stringent temperature goals? *Philosophical*

- [4] European Commission. Methane abatement partnership roadmap. Policy Document, European Commission, Directorate-General for Energy, Brussels, Belgium (2024). URL https://energy.ec.europa.eu/document/download/1978e73b-0158-4593-87a5-c051bc0ec714_en?filename=Methane%20Abatement%20Partnership%20Roadmap.pdf. Launched at COP29, Baku, Azerbaijan.
- [5] U.S. Environmental Protection Agency. Methane emissions reduction program (2025). URL <https://www.epa.gov/inflation-reduction-act/methane-emissions-reduction-program>. Webpage describing the Methane Emissions Reduction Program under the Inflation Reduction Act.
- [6] Smith, A. Low carbon fuel standards. Personal webpage (2025). URL <https://www.aaronsmithhagecon.com/low-carbon-fuel-standards>. Aaron Smith, Professor of Agricultural and Resource Economics, University of California, Berkeley.
- [7] Witcover, J. & Murphy, C. Oregon’s clean fuels program: A review and status update. *Transportation Research Record* (2021). URL <https://journals.sagepub.com/doi/10.1177/0361198120972394>.
- [8] About GFMR. URL <https://www.worldbank.org/en/programs/gasflaringreduction/about>.
- [9] Cullenward, D. California’s low carbon fuel standard. Policy Paper, Kleinman Center for Energy Policy, University of Pennsylvania, Philadelphia, PA (2024). URL <https://kleinmanenergy.upenn.edu/wp-content/uploads/2024/10/KC-Paper-16-Californias-Low-Carbon-Fuel-Standard.pdf>. Kleinman Center Paper Series.
- [10] Smith, A. Are Manure Subsidies Causing Farmers to Milk More Cows? (2023). URL <https://agdatanews.substack.com/p/are-manure-subsidies-causing-farmers>.
- [11] Kaeding, D. Report: Digesters encourage farm expansion and make pollution worse in Kewaunee County. URL <https://www.wisfarmer.com/story/news/2024/06/13/did-5-kewaunee-co-cafos-expand-after-building-methane-digesters/74053556007/>.
- [12] Kim, B. F. *et al.* Deconstructing the manure digester and biogas controversy: an examination of the public health, environmental justice, climate, and economic implications. *Current Environmental Health Reports* (2025). In press.
- [13] Agriculture, C. D. o. F. . CDFA - OARS - Dairy Digester Research & Development Program. URL <https://www.cdca.ca.gov/oars/ddrdp/>.

- [14] Low Carbon Fuel Standard \textbar California Air Resources Board. URL <https://ww2.arb.ca.gov/rulemaking/2024/lcfs2024>.
- [15] Smith, A. What’s Worth More: A Cow’s Milk or its Poop? URL <https://agdatanews.substack.com/p/whats-worth-more-a-cows-milk-or-its?r=i2qe&triedRedirect=true>.
- [16] California Dairy & Livestock Database (CADD) \textbar California Air Resources Board. URL <https://ww2.arb.ca.gov/resources/documents/california-dairy-livestock-database-cadd>.
- [17] Magesh, V. *et al.* Cal-ff: A comprehensive dataset of factory farms in california compiled using computer vision and human validation. *Nature Scientific Data* (2025). Forthcoming.
- [18] United States Environmental Protection Agency. AgSTAR Data and Trends (2014). URL <https://www.epa.gov/agstar/agstar-data-and-trends>.
- [19] Callaway, B. & Sant’Anna, P. H. C. Difference-in-differences with multiple time periods. *Journal of Econometrics* **225**, 200–230 (2021).
- [20] California Dairy & Livestock Database (CADD) \textbar California Air Resources Board. URL <https://ww2.arb.ca.gov/resources/documents/california-dairy-livestock-database-cadd>.
- [21] Arndt, C. *et al.* Short-term methane emissions from 2 dairy farms in california estimated by different measurement techniques and us environmental protection agency inventory methodology: A case study. *Journal of Dairy Science* **101**, 11461–11479 (2018).
- [22] Rodriguez, M. V. *et al.* Anaerobic Digester Installation Significantly Reduces Liquid Manure Management CH₄ Emissions at a California Dairy Farm. *GCB Bioenergy* **17**, e70047 (2025). URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/gcbb.70047>.
- [23] Lebigot, E. O. Uncertainties: a Python package for calculations with uncertainties. <http://pythonhosted.org/uncertainties/>. Accessed: 2025-09-30.
- [24] Ho, D. E., Imai, K., King, G. & Stuart, E. A. Matchit: Nonparametric pre-processing for parametric causal inference. *Journal of Statistical Software* **42** (2011).
- [25] Ho, D. E., Imai, K., King, G. & Stuart, E. A. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis* **15**, 199–236 (2007).

- [26] Rubin, D. B. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology* **2**, 169–188 (2001). URL <https://doi.org/10.1023/A:1020363010465>.