Where's the Value? The Impacts of Sow Gestation Crate Laws on Pork Supply and Consumer Value Perceptions

Benjamin Blemings[*], Peilu Zhang[†], Clinton L. Neill[‡]

This Version: August 29, 2022; Most Recent: Link

Abstract

Despite varying farm animal welfare laws being enacted and considered, it is unclear and unknown whether they generate sufficient benefits to justify their costs. This study asks how sow gestation crate laws that dictate minimum crate size, a particular farm animal welfare law, impact meat production and the subsequent value that consumers place on meat produced under the laws. An observational analysis leverages the 8 states that have implemented laws regarding gestating sow crates in a differences-in-differences framework using staggered adoption robust estimators. We then perform a consumer preference experiment using a contingent valuation design with treatment effects to separate out the heterogeneous beliefs about animal welfare law effects. The results of the study show that the costs of such laws significantly reduce production, thereby increasing producer costs. However, consumers do value the corresponding pork products more in accordance with their animal welfare beliefs but not quality. Such information on costs and benefits should be considered in future laws relating to gestation crate bans.

Keywords: Farm Animal Welfare, Preference Experiment, Hog Production

JEL Classification: Q18, Q11, D24

[*]Benjamin Blemings, Ph.D., Dyson School of Applied Economics and Management, Cornell University, Warren Hall, 137 Reservoir Ave, Ithaca, NY 14850. E-mail: btb77@cornell.edu

[†]Peilu Zhang, Ph.D., Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14850. Email: pz88@cornell.edu

[‡]Clinton L. Neill, Ph.D., Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14850. Email: cln64@cornell.edu

1 Introduction

The implementation of increased animal welfare standards, both from private motivations and public regulation, is attributed to the value that the general public and meat consumers place on the animal itself. However, this perception of value can be confounded by an individual's preferences for improvements in welfare and quality of meat products (Hudson, 2010; Tonsor, Wolf, & Olynk, 2009). In addition, there are few papers that consider value from a producer or societal perspective in the terms of costs and benefits of improving animal welfare (Carlier & Treich, 2020). Instead, much of the literature focuses on the benefits as measured by consumers' preferences and WTP for products under different animal-friendly practices used in the production process (Carlsson, Frykblom, & Lagerkvist, 2007; Liljenstolpe, 2008; Lusk, Nilsson, & Foster, 2007; Lusk, Norwood, & Pruitt, 2006; Tonsor, Olynk, & Wolf, 2009; Tonsor, Wolf, & Olynk, 2009; Viske, Lagerkvist, & Carlsson, 2006). A smaller portion of literature focuses on the costs of such regulation (Olynk, Tonsor, & Wolf, 2010; Sumner, Matthews, Mench, & Rosen-Molina, 2010; Tonsor & Wolf, 2010). Thus, a study examining both costs and benefits is warranted, especially one that addresses the confounding nature of the quality and direct welfare aspects of the issue.

From these previous studies we cannot explicitly tell whether consumers care about animal welfare, care only about the quality of products under animal-welfare laws, or care about both (Carlier & Treich, 2020; Hudson, 2010). Thus, there is a need to better value enhancements in animal welfare from both the social/public good view and the costs/benefits of production. In order to tackle this problem, we combine causal analysis of the effect of farm animal welfare laws on production with stated preference information to determine the change in market value from animal welfare changes. This combination of methodology allows for a robust understanding of how animal welfare is valued both under different regulatory scenarios and under different consumer motivations. Therefore, we are able to assess both costs and benefits of animal welfare regulation within one study. This allows us to address the confounding problem brought forth by Hudson (2010), better separate consumers' preferences and beliefs about animal welfare, and determine the casual effects of welfare regulation on meat producers.

We focus our analysis on hog output (in pounds and head), producer gross income (in dollars), and a related pork product (chops). Specifically, we examine the impacts of several states enacting laws on the minimum size of sow gestation crates to determine the effects on supply side outcomes and use information treatments to determine the core of consumer preferences for pork products produced under such conditions. Our supply side analysis is one of the most unified analyses of sow gestation crate laws to date since it studies the average causal impacts of the law on supply outcomes across all adopting states. In regards to prior work on welfare information, most studies focus on the relationship between label information and consumption decisions (Chen, Liu, Jaenicke, & Rabinowitz, 2019; Edenbrandt & Lagerkvist, 2021; Lai, List, & Samek, 2020; Lin & Nayga Jr, 2022; Weaver & Finke, 2003). Another branch of research in this area explores the heterogeneity of consumers besides the correlation, and show differential effects of label information on consumers with different characteristics (Dörnyei & Gyulavári, 2016; Rimal, 2005; Van Der Merwe, Kempen, Breedt, & De Beer, 2010; Zhu, Lopez, & Liu, 2016). From the consumer side, our paper adds insights of how label information changes consumers' choices to the literature through the analysis of mechanisms.[1]

The results from both sets of analyses indicate that the benefits from the animal welfare laws do not necessarily outweigh the production costs to the producer. This is due to the heterogeneous effects of the laws on producers and the fact that consumers' beliefs about the benefits of gestation crates is a key component in driving willingness-to-pay differences. Moreover, producers, who incur the increased cost of such regulation, do not benefit from the addition of value for consumers. Thus, current and future implications for animal welfare regulation should consider these effects on supply and demand which may not be win-win across both sides of the market.

[1]Thus, we define 'value' from two perspectives: The output from producers - with particular interest on gross income; and the willingness-to-pay and motivations for purchasing pork products.

The remainder of this article is as follows: a discussion of the supply side analysis utilizing a difference-in-difference approach; a section on the consumer demand analysis utilizing a contingent valuation design with information treatments; a section on the policy implications, followed by concluding comments.

2 Supply Side Analysis

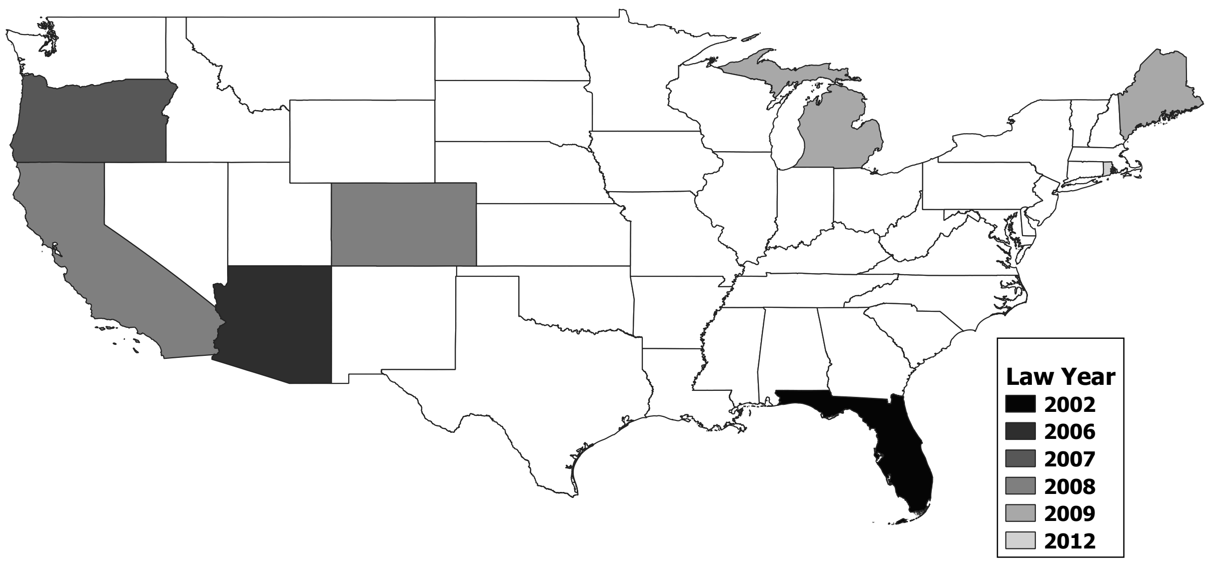
First, this paper estimates how a farm animal welfare (FAW) law, that dictates minimum sizes for gestating sow crates, affects supply and value. Data for state-level regulations on the size of gestating sow crates comes from the Animal Welfare Institute. Over the sample, the law has been introduced in 8 states which are shown in Figure 1 and details are given in Table A.1. There is very little variation in the laws, they all dictate the minimum size of gestation crates for sows.

Data on yearly state-level outcomes for hogs comes from USDA's NASS. As our first research question is how these laws influence the supply of pork, the first set of dependent variables are pounds sold, gross income, and production dollars.[2] Our next research question asks how consumers value the meat produced under these laws. There is a measurement of value from NASS, the value of intra-state farm home consumption; however, this may not reliably measure willingness to pay, so we report the value results in Section A.2. The lack of observational data on value is an additional motivation for conducting the choice experiment in Section 3.

This data is well-built for a difference-in-differences approach, since we observe several key variables for a long period and several law changes which occur over the length of this sample. The assembled data is a panel of 50 states from 1988-2020, with descriptive statistics being presented in Table A.2. As shown in Figure A.1, gross income and production dollars are nearly exactly equiv-

[2] Gross income and production income are deflated by the consumer price index and are in 2020 dollars.

Figure 1: Adoption of Sow Gestation Crate Laws



Note: Data comes from the Animal Welfare Institute. Year is the year the law was passed. Details of the laws can be seen in Table A.1.

alent with an unconditional correlation of .9976. The main results use producer's gross income as the main variable and production dollars results are reported in Section A.1.[3]

2.1 Method

The effect of the laws are estimated using two-way fixed effects specified as,

$$y_{st} = \beta_1 CrateLaw_{st} + \mu_s + \rho_t + e_{st}, \tag{1}$$

in which s stands for state and t stands for time. $CrateLaw$ is a binary variable that equals 1 if state s has passed the gestation crate law in year t or before year t. The state fixed effects, μ_s, accounts for time-invariant state level factors. The year fixed effects, ρ_t, account for trends in hog meat production and demand. The standard errors are clustered at the state-level to account for errors that are correlated within state and some forms of mis-specification.[4]

The theoretical prediction for β_1 is straightforward for supply. For the supply outcomes, pounds sold, gross income, and production dollars, β_1 is expected to be negative. Gestating sow crates are an input to production and the law dictates using larger crates, assuming the pre-legislation crates do not meet the new legislation. This would require using more inputs for the same amount of output and in a perfectly competitive market this could be seen as shifting the supply inwards due to the increasing price of inputs. For the value outcome, if consumers place higher value on more humanely raised hogs or believe that humanely raised meat is of higher perceived quality, β_1 will be positive.

Identification Attaching a causal interpretation to β_1 requires the parallel trends assumption to hold. In this context, it means that supply in states that introduced FAW laws would continue trend-

[3]Gross income is the total values of cash receipts plus the value of home consumption. The value of production is similar and also accounts for differences in inventory.

[4]We account for few treated clusters, 8, using the wild bootstrap.

ing on parallel paths in the absence of the passing of the law. We will use dynamic specifications to diagnose whether parallel trends is a plausible assumption. The dynamic specifications follow,

$$y_{st} = a + \sum_{j=2}^{J} \beta_j + \sum_{k=0}^{K} \beta_k + \mu_s + \rho_t + e_{st}. \tag{2}$$

Due to the omission of the coefficient for 1 year prior to the beginning of the policy, the β_j and β_k coefficients can be interpreted as the difference between the comparison states and the states with the law compared to their difference in the year prior to the adoption of the law. If the β_j coefficients are not distinguishable from zero, it means there are no average differences between treated and comparison states in pre periods which implies that the trends would remain parallel in the post periods in the absence of the intervention. We do not include time-varying covariates, because it is likely that many of the relevant ones, such as number of farms, are affected by the law and are therefore endogenous which is worse than an unconditional approach (Caetano, Callaway, Payne, & Rodrigues, 2022).[5]

Staggered Adoption Robust Estimators As shown in Figure 1, the first year a law was passed was in 2002 and the last year one passed was 2012. Since staggered adoption of laws can bias estimates if effects are heterogeneous, several new estimators which address this issue are also used. No time-varying controls are included, but extensive attention is paid to the heterogeneous impacts of the laws across separate cohorts of treated states using different aggregations of the estimates produced by the Callaway-Sant'Anna (CS) estimator (Callaway & Sant'Anna, 2021).

2.2 Static Results

As shown by column 1 of Panels A and B of Table 1, when the effect of the law is estimated with TWFE, the estimates are of the expected signs, negative, for supply outcomes of pounds sold

[5] The other issues with covariates are that farms from NASS do not breakdown by the type of animal making it a noisy measure of hog farms and that there are few states with few observations.

(Panel A) and gross income (Panel B). As shown in column 1 of Panel A, the TWFE estimate implies the law reduces the amount of hog sold by 228.91 million pounds. This is a reasonably sized effect which equals 15.4% of the standard deviation of pounds sold. Moreover, the effect is economically significant, representing 38.3% of the average amount of pounds of hog sold. The gestation crate law is also associated with a $92.11 million reduction in producers' gross income.

Table 1: Estimates of Effect of Gestation Crate Law on Pounds Sold and Producer Gross Income

Panel A: Pounds of Hog Sold	(1)	(2)	(3)
Gestation Crate Law	-228.91	-232.89*	-142.75*
	(141.69)	(135.36)	(83.77)
	[0.09]		
Observations	1634	1634	1650
TWFE	X	-	-
Imputation Method[a]	-	X	-
CS Method[b]	-	-	X

Panel B: Gross Income	(1)	(2)	(3)
Gestation Crate Law	-92.11	-93.36	-69.38
	(68.20)	(65.47)	(46.21)
	[0.17]		
Observations	1634	1634	1650
TWFE	X	-	-
Imputation Method[a]	-	X	-
CS Method[b]	-	-	X

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Cluster-robust standard errors, by state, in parentheses. All regressions include county and year fixed effects. Gestation Crate Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. In Panel A, the dependent variable is pounds of hog sold. In Panel B, the dependent variable is producer gross income. Column 1 estimates the law's effect using two-way fixed effects OLS. [a]Column 2 uses the imputation estimator of Borusyak, Jaravel, and Spiess (2021) and Gardner (n.d.).[b]Column 3 uses group-time effect aggregation from Callaway and Sant'Anna (2021). It uses never treated as comparison units, the doubly robust differences-in-differences estimator based on inverse probability of tilting and weighted least squares (Sant'Anna & Zhao, 2020), and shows the simple aggregation.

Both estimates are borderline statistically significant. Using conventional cluster-robust inference at the state level, the estimates for pounds sold is not significant above the 10% threshold. However, once the wild cluster bootstrap is used the estimate obtains statistical significance with a p-value of 0.09. Gross income is not statistically significant above 10% with either conventional cluster robust inference or with bootstrapping. However, it is worth noting that the estimate is economically significant at 9.5% of a standard deviation or 22.3% of the average hog gross income (Imbens, 2021).[6]

There are several factors that could be contributing to the estimates in column 1 of Table 1 being relatively imprecise. First, the outcomes are based on survey data and are therefore measured with error which leads to attenuation bias. Second, there is limited power from a limited number of observations overall, 1,650, and few states enacting the policy, 8. Finally, there is potential bias that arises from staggered adoption of the law and dynamic or heterogeneous effects. These limitations, particularly the second and third ones, are addressed in the following analyses.

2.3 Alternative Estimators

The staggered treatment of welfare laws can lead TWFE to place non-negligible positive weights on inappropriate comparisons. These comparisons can bias the estimates by placing too much weight on comparing late treated to early treated units.[7] As shown in the Goodman-Bacon (GB) Decompositions in Figure A.2, almost all of the weight in the TWFE regressions are put on favorable comparisons (triangle markers) which suggests a limited bias in the estimated effects for most outcomes. Furthermore, most of the estimated effects with substantial weight are near the estimated coefficient for TWFE.[8]

Next, we present the alternative estimators in the second and third columns of Table 1. Each of these estimators, the imputation estimator (Borusyak et al., 2021) and the CS estimator (Call-

[6]As shown in Table A.6, the results for production are very similar to those for gross income. As shown by Table A.7, the coefficient on the law's effect on value of home consumption, excluding inter-state trade, is also of the expected sign when estimated by TWFE, positive.

[7]Note that this is not an issue in a 2 period setup or if there are no heterogeneous effects across time or groups (De Chaisemartin & d'Haultfoeuille, 2020; Roth, Sant'Anna, Bilinski, & Poe, 2022). The comparison that is most desirable is comparing states that are treated to those that are never treated.

[8]One notable exception is the weights on value; as shown in Figure A.2d, the largest weighted comparisons have negative effect sizes, while the TWFE estimate is positive. This is indicative that there may be heterogeneity in the effect of the laws on value across treatment cohorts.

away & Sant'Anna, 2021), are robust to staggered timing with heterogeneous effects in order to see whether the results are different after accounting for staggered adoption of the law. As shown by columns 2 and 3 of Panel A, using the imputation estimator or the CS estimator leads to similar conclusions of the effect of the law on pounds sold. The imputation estimator has a similar effect size to TWFE, at -232.89, but the CS estimator is smaller at -142.75. Both are statistically significant with 90% confidence. The pattern of the CS estimator producing smaller effects holds for gross income. However, the law does not have a statistically significant impact on gross income when using the other estimators.[9]

2.3.1 Dynamic Effects

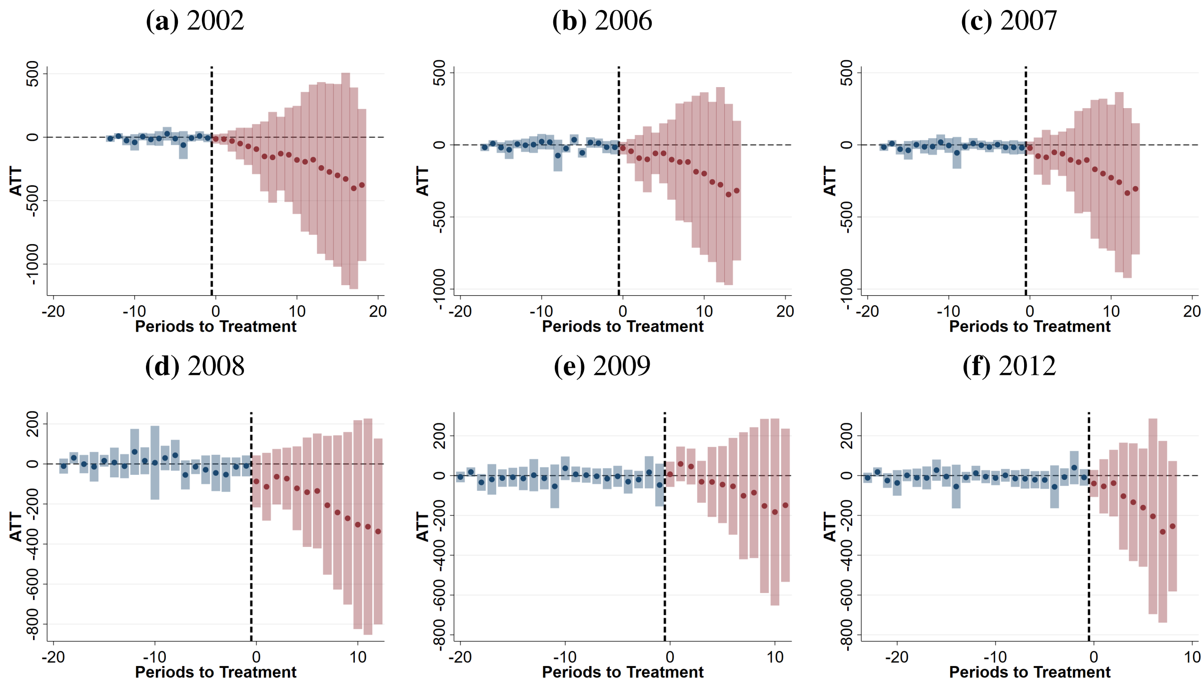
Before further investigating the heterogeneity of the treatment effects, we first determine whether there is evidence that these effects can be given a causal interpretation. This requires the parallel trends assumption which we investigate using the dynamic effects specification given in Equation 2. We use the event-study aggregation from CS (Callaway & Sant'Anna, 2021) and report these event studies separately by treatment cohort to minimize bias due to heterogeneity across cohorts.

As shown in the event studies for pounds sold which are presented in Figure 2, there are no discernible patterns in the pre-period coefficients. As shown in the event-studies for gross income which are presented in Figure 3, there are no discernible patterns in the pre-period coefficients although they are more volatile than in the pounds sold event studies.[10] The lack of apparent pre-trends in the figures is consistent with parallel trends holding which allows the estimates to have a causal interpretation. While it is true that states select whether and when to enact these laws, there

[9]The estimators make different assumptions and which estimator to prefer depends on those assumptions. In this case, the statistical significance is the same and the point estimates are not very different, so it's unnecessary to take a stand on an estimator.

[10]The event-studies for value, presented in Figure A.6, show no noticeable pre-period trends in value.

Figure 2: Event Studies by Cohort for Pounds Sold

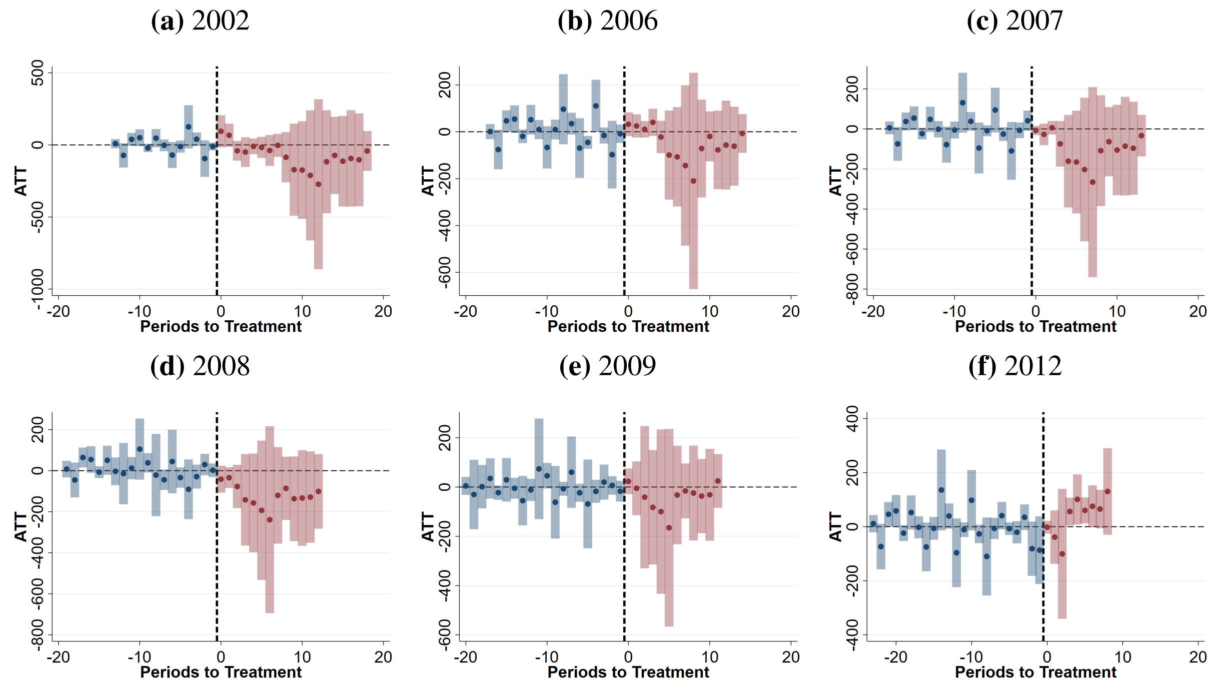


Note: The y-scale and x-scale differs across figures. These figures report coefficients from Equation 2 for the estimation of dynamic effects of the law change on pounds of hog sold. Each figure represents a different treatment cohort. The coefficients are obtained from the event-study aggregation of Callaway and Sant'Anna (2021). Coefficients represent the change in pounds of hog sold for states that adopt the gestation crate size law relative to non-adopting state in the years before and after the law passes, as compared with the year immediately prior to the law passing. Gestation crate law adoption dates comes from the Animal Welfare Institute and the pounds of hog sold comes from the USDA NASS Quick Stats. The whiskers represent 95% confidence intervals. The standard errors are clustered at the state-level and are calculated using the wild bootstrap.

is no evidence that their hog production or gross income are trending differently in the pre periods which suggests that the model assumptions are adequate to identify causal effects.

While there are some individually significant pre-period coefficients, there are never more than two in a row with the same sign. In contrast, the post-period coefficients exhibit sustained estimated differences that are different than 0 in the same direction. As shown in Figure 2, the coefficients are negative in every post period for nearly every treatment cohort and increasingly so, suggesting the negative impact on total output grows over time. For gross income (Figure 3), there is a period where post period coefficients decline, but eventually move back towards 0 which suggests that the harmful impacts of the law on producers' gross income (our measure of producer value) are somewhat, though not completely, mitigated after an adjustment period. The 2012 cohort follows

Figure 3: Event Studies by Group for Gross Income



Note: The y-scale and x-scale differs across figures. These figures report coefficients from Equation 2 for the estimation of dynamic effects of the law change on producer gross income. Each figure represents a different treatment cohort. The coefficients are obtained from the event-study aggregation of Callaway and Sant'Anna (2021). Coefficients represent the change in pounds of hog sold for states that adopt the gestation crate size law relative to non-adopting state in the years before and after the law passes, as compared with the year immediately prior to the law passing. Gestation crate law adoption dates comes from the Animal Welfare Institute and producer gross income comes from the USDA NASS Quick Stats. The whiskers represent 95% confidence intervals. The standard errors are clustered at the state-level and are calculated using the wild bootstrap.

this general pattern, but after a few years the readjustment upwards is large enough to become positive.

2.3.2 Heterogeneity By Cohort

There is likely to be some heterogeneity across treatment cohorts. This is perhaps unsurprising as later adopters can observe early adopters and attempt to mitigate any negative impacts on supply and there are differences in the sizes of the hog industry. Figure 4 embraces the heterogeneous impacts across treatment cohorts by estimating the effect of the sow gestation crate laws for each separate treatment cohort. As Figure 4 shows, there are negative impacts that are statistically significant at 10% for pounds sold and these are consistent across all cohorts. However, for gross income and production the laws have negative (and sometimes statistically insignificant) effects on production and gross income for all but the last treatment cohort of 2012.[11]

Ideally, it would be possible to understand the root cause of these heterogeneous impacts of the law. One explanation that is unlikely to generate these differences is the exact text of the laws which all require gestating sows to be able to stand up, lie down, and turn around. Thus, we turn our attention to pre-treatment variables for each of the states. One immediately obvious difference for the 2012 cohort, composed solely of Rhode Island, is that it is a mere fraction of the size of the other cohorts in terms of both amount of land farmed, number of farms, and size of hog industry. An alternative explanation that is consistent with the estimates is that the reduction in pounds sold has gotten less worse for each progressive cohort and the effect on gross income even changed for the final cohort which could hint to adaptation occurring.

2.4 County-Level

For several of the analyses presented so far, the results are often more imprecise than would be preferred to make strong conclusions. An alternative way to construct the dataset, which may address this shortcoming, is to perform the analysis at the county-level. This increases the unit of observation from 50 states to over 3,000 counties. However, this alternative construction of the dataset comes at several costs. The only outcome available at the county-level is hog production (measured in head), county-level data only exists every 5 years (1997-2017), there is often missing data, and it doesn't increase the number of states that introduced the law.[12]

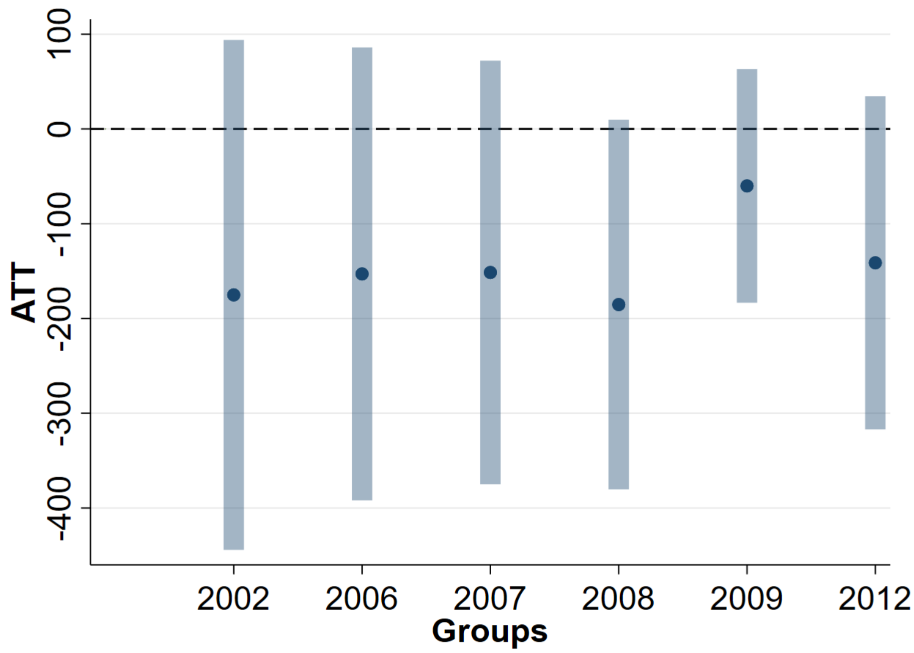
Nevertheless, we believe that examining the effect of the law on supply in this alternate way could offer more precise inference which corroborates the robustness of the state-level results. The county-level results are presented in Table 2, in which each panel handles missing outcomes

[11] The effects of the law on value differ the most by cohort. The law increased value for the 2006 and 2012 cohorts and these are both statistically significant. However, 4 treatment cohorts had increased value and 2 had decreased value. An interesting note is the 2012 cohort reduced pounds sold, but increased production, gross income, and value.

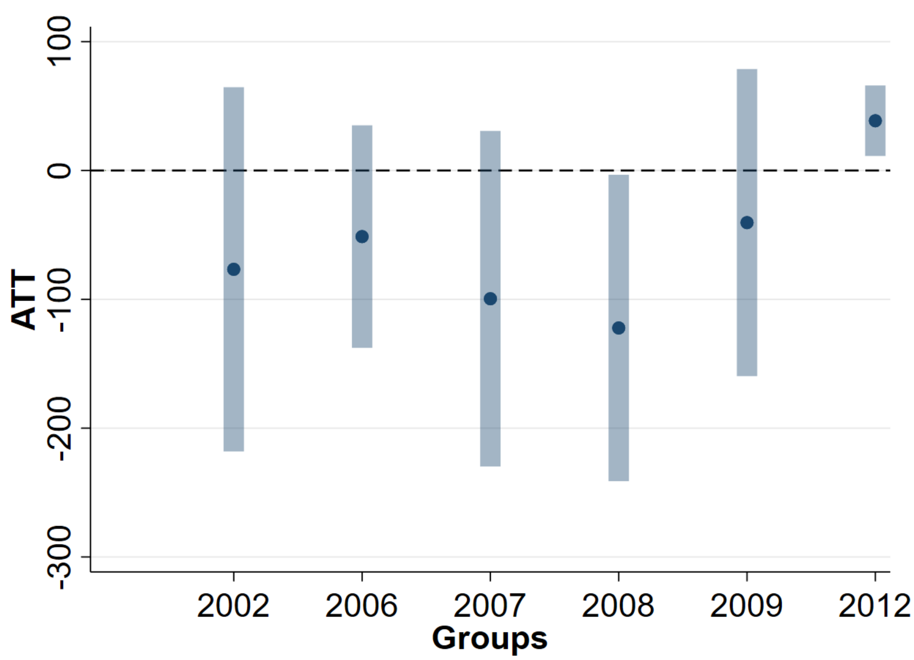
[12] Sumamry statistics are shown in Table A.3.

Figure 4: Cohort Level Estimates

(a) Pounds Sold



(b) Gross Income



Note: Estimates are group-level aggregates from Callaway and Sant'Anna (2021). Wild bootstrapped standard errors. 95% confidence intervals shown. Pounds sold, production, and gross income are in millions. Value is in thousands.

slightly differently. Panel A keeps all missing values as missing, Panel B converts all missing values to 0, and Panel C only reports county-years that were not in the USDA NASS data as zeroes with others as missing. As shown in column 1 of Table 2, the TWFE estimators preserve the expected sign and are also statistically significant at 90% confidence regardless of how the missing outcome data are handled. How the zeroes are handled also has little effect on the TWFE point estimates with them being between -18,780 and -14,403. As shown in the GB Decomposition in Figure A.3, the TWFE weights are placed nearly exclusively on appropriate comparisons and all point estimates are negative and relatively close to the aggregate point estimate.

In contrast to the TWFE estimate not providing more precise inference, column 2 of Table 2 shows that imputation estimator at the county-level does provide estimates that are significant above 95% confidence. Although the imputation estimator is always statistically significant above 95% confidence, the point estimate does vary more than TWFE from -24,414 to -14,880. The magnitude, relative to the standard deviation, is slightly smaller than at the state-level. The TWFE estimates are 5.28 to 5.95% of the standard deviation which is a result of going to a more granular geographic unit with higher variation and the different outcome measure (pounds sold at the state-level compared to head at the county-level).[13] Finally, we report the results with standard errors clustered at the county-level in Table A.5 which produces even more forgiving inference, albeit with likely too narrow confidence intervals due to treatment being assigned at the state level, with all coefficients statistically significant above 99% confidence.

3 Consumer Analysis

In order to identify the motivation of consumers' WTP, we implemented a between-subject design with different labeling (treatment) information to separate and compare the WTP for animal wel-

[13]The CS estimator is reported in Table A.4 and the results are never significant. Since we are doing this as a robustness check for precision and the imputation estimator is more efficient (Roth et al., 2022), it is appropriate to focus on the imputation estimator over the CS estimate.

Table 2: Estimates of Effect of Gestating Sows Law on Hogs (Measured in Head) Sold, County Level

Panel A: Naïve Missing		
	(1)	(2)
Gestating Sows Law	-18780.83*	-24414.04**
	(9745.02)	(9803.96)
	[0.07]	
Observations	11341	11552
TWFE	X	-
Imputation Method[a]	-	X
Panel B: Naïve Zero		
	(1)	(2)
Gestating Sows Law	-14403.45*	-14880.09**
	(7333.56)	(6952.37)
	[0.08]	
Observations	15726	15726
TWFE	X	-
Imputation Method[a]	-	X
Panel C: Merge Zero		
	(1)	(2)
Gestating Sows Law	-16602.70*	-21468.43**
	(8851.37)	(8962.49)
	[0.08]	
Observations	12656	12778
TWFE	X	-
Imputation Method[a]	-	X

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Cluster-robust standard errors, by state, in parentheses. All regressions include county and year fixed effects. The dependent variable is hogs sold, measured in head. Gestating Sows Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. In Panel A, all missing outcomes are left missing. In Panel B, all missing outcomes are replaced with zero. In Panel C, only missing values that arise due to no county-year in the USDA Quick Stats data are counted as zeroes. Column 1 estimates the law's effect using two-way fixed effects OLS. [a]Column 2 uses the imputation estimator of Borusyak et al. (2021) and Gardner (n.d.).

fare preferences and the preference of perceived quality that may also be associated with products under gestation crate regulations. We believe the identification of what motivates consumer's WTP is important for policy makers weighing the costs and benefits of related regulations.

We use a double-bounded contingent valuation approach to estimate WTP (Hanemann, Loomis, & Kanninen, 1991) in each treatment group. The contingent valuation method (CVM) is a standard approach to elicit people's WTP for non-market goods or services; it has been applied in fields such as environmental, health, food, and marketing (Diener, O'Brien, & Gafni, 1998; McCluskey, Grimsrud, Ouchi, & Wahl, 2003; McFadden & Train, 2017; Neill & Holcomb, 2019; Neill & Williams, 2016; Venkatachalam, 2004). There is remains some debate about its validity due mainly to hypothetical response bias, discrepancy between willingness to pay and willingness to accept, and the scope problem (Carson, 2012; Hausman, 2012). This argument is more of a concern for goods and services with no actual market transaction connection - which is not the case in this study, though we still employ cheap talk to reduce the amount of hypothetical bias in the experiment (Bishop & Heberlein, 2019).

The main purpose of our experiment is to compare the WTP across different treatment groups, instead of the absolute value of the WTP. Compared to studies which elicit the absolute value of WTP, the combination of CVM and a between-subject design to some extent mitigates the main issues of using CVM mentioned above in our experiment. The between-subject design also allows us to test mechanisms about how the treatment information affects the WTP, while being able to distinguish preferences from beliefs.

3.1 Design

The product of interest in our experiment are pork chops, a common cut of pork, with 4 price point variations ($4.0, $4.5, $5.0, $5.5). In each two-stage choice question, subjects were first asked to select between pork chops produced under the Gestating Crate Law (GCL) and pork chops which are not produced under the GCL. The price of pork chops under GCL is randomly picked from the 4 price variations; the price of pork chops which are not produced under the GCL is always $4.50 per pound. If the subject chooses the GCL pork chops in the first stage, the price of the GCL pork chops increases by $0.50 in the second stage; if the subject chose the non-GCL pork chops, the price of the GCL pork chops decreases by $0.50 in the second stage. The price of non-GCL pork chops remains at $4.50 in the second stage. In each question, we also presented an introduction of gestation crates and the GCL before subjects make selections in both the first and second stage.

The experiment was conducted online through Qualtrics.[14] Subjects were randomly assigned to one of the four groups (one baseline & three treatment groups). Each group only varies in the extra information provided in the choice question about the increase in the quality of products and the increase in animal welfare due to the law.

Table 3 shows the information summary of each treatment group. In the Baseline group, there was no extra information about the quality of products or animal welfare. In the "Quality & Welfare" group, we presented the information "Research has found that the such Gestation Crate Laws increase the quality of pork products, and increases the welfare of hogs/sows." In the "Quality" and "Welfare" group, the information provided was only about the quality of products and the welfare, respectively. Figure 5 is a screenshot of the choice question in the first stage in the "Quality" group. The information in the second stage only differs in the price changed based on subjects' choices in the first stage.

Table 3: Summary of Information Given to Treatment Groups

	Introduction of gestation crates and the law	Extra information	No. Observation
Baseline	Yes	N/A	268
Quality & Welfare	Yes	"Research has found that the such Gestating Crate Laws increase the quality of pork products, and increases the welfare of hogs/sows."	258
Quality	Yes	"Research has found that the such Gestation Crate Laws increase the quality of pork products."	243
Welfare	Yes	"Research has found that the such Gestation Crate Laws increase the welfare of hogs/sows."	270

Notes: Each subject was randomly assigned to one of the groups.

[14]In the survey, we included an attention-check question so that we could make sure that the experiment was not answered by a bot.

Figure 5: Screen Shot of Choice Question in the First Stage of "Quality" Group

Gestation crates (sometimes also called gestation stalls) refer to metal crates (approximately 7 feet long and 2 feet wide) that house female breeding stock (hogs/sows) in individually confined areas during an animal's four-month pregnancy. As of 2022, eight passed the laws against using gestation crates. Specifically, the law prohibits the use of crates in gestating hogs/sows so that they can stand up, lie down, or turn around freely.

Research has found that the such Gestation Crate Laws <u>increase the quality of pork products.</u>

Suppose the next time you go to the grocery store for purchasing pork chops you see the following two options. Which pork chop option would you choose to purchase?

Please answer as if you were actually making shopping decisions in the grocery stores.

Pork chops at $5.00 per pound. The pork is produced under the Gestating Sow Law (i.e., no use of gestation crates.)	Pork chops at $4.50 per pound. The pork is not produced under the Gestating Sow Law (i.e., use of gestation crates.)
○	○

At the end of the experiment, we asked participants' about their beliefs on animal welfare, welfare laws, and other benefits and costs of gestating sow laws. Demographic information including the state of residence was also elicited at the beginning of the experiment (for detailed instructions of the experiment, please refer to Appendix B). In total, 1039 observations were collected. As shown in Table 4, the average demographic characteristics are similar and not statistically different across groups suggesting randomization was successful in terms of identifying effects of the information treatments.

Table 4: Summary of Demographic Characteristics by Treatment Group

	Baseline	Quality & Welfare	Quality	Welfare
Age (median)	43	41	43	42
Gender	51.9% (F)	50.8% (F)	51.0% (F)	52.2% (F)
	48.1% (M)	49.2% (M)	49.0% (M)	47.8% (M)
Education(median)	2 year/Associates Degree	2 year/Associates Degree	2 year/Associates Degree	2 year/Associates Degree
Household Income(median)	\$60,001-\$65,000	\$60,001-\$65,000	\$60,001-\$65,000	\$50,001-\$55,000
Household Size(mean)	3	3	3	3
Pork products consumption (weekly)	2.5	2.5	2.5	2.4

Notes: Two-sided Mann-Whitney U-Tests were used to test the difference in Age, Household Income, Household Size, and Pork consumption between each two of the groups. χ^2 tests were used to test the difference in Gender and Education. We did not find any significant differences in any comparisons.

3.2 Methods

A double-bounded regression model, controlling for demographic characteristics and beliefs, is used to estimate the effects of the information treatments on WTP. The price in the first stage is p_1 and the price in the second stage is p_2. There are four possible outcomes in our methodology: (a) the individual chooses GCL products in the first stage and chooses non-GCL products in the second stage, $p_2 > p_1$; (b) the individual chooses GCL products in both stages, $p_2 > p_1$; (c) the individual chooses non-GCL products in the first stage and GCL products in the second stage, $p_2 < p_1$; (d) the individual chooses non-GCL products in both stages, $p_2 < p_1$. We can infer that in case (a), the individual's $p_1 < WTP < p_2$; in case (b), $p_2 < WTP < \infty$; in case (c), $p_2 < WTP < p_1$; in case (d), $0 < WTP < p_2$.

The double-bounded model by Lopez-Feldman (2012) assumes the WTP can be modeled as a linear function:

$$WTP_i(z_i; u_i) = z_i\beta + u_i \tag{3}$$

where z_i is a vector of explanatory variables, β is a vector of parameters and u_i is an error term $u_i \sim N(0,\ \sigma^2)$. The estimation of the double-bounded model captures the probability of individuals answering yes or no in the first and second question for the four cases, and uses maximum likelihood estimation to get estimates of β and σ for the WTP in Equation 3. The detailed estimation process, following Lopez-Feldman (2012), is shown in Appendix C.

Besides the demographic information, we include beliefs as explanatory variables in the model. We have a set of 11 belief questions at the end of the experiment, and the order of the 11 questions are randomized. Subjects used a 5-point Likert scale to address the belief questions (from "strongly disagree" to "strongly agree"). In order to reduce the issue of multicollinearity/dimensionality and create a more parsimonious model, we use principal component analysis with varimax rotation to conduct a factor analysis of responses to the belief questions before the estimation of the double-bounded model (Boxall & Adamowicz, 2002; Tonsor, Wolf, & Olynk, 2009). Based on loadings of indicators by the factor analysis, we suggest two factors that capture the major associations individuals make with the gestation crate laws. The first factor is labeled "pork products quality and animal welfare". The second factor is referred to as "benefits of gestation crates to consumers". Table 5 shows factor loadings of each belief question, and each question is assigned to the specific factor which has higher loadings. Then, a score of each factor is calculated for each individual and is used in the double-bounded model estimation as explanatory variables.

Table 5: Factor Analysis of Beliefs

Statement (belief question)	Factor1 loadings: pork products quality and animal welfare	Factor2 loadings: benefits of gestation crates to consumers
I believe a ban of gestation crates increases the quality of pork products.	[0.7210]	0.0127
I believe pork products generated using gestation crates are subject to more food safety risks.	[0.6891]	0.0257
Animals should not reside in confined areas like gestation crates.	[0.7107]	-0.1753
I believe a ban on gestation crates increases the welfare of hogs.	[0.7401]	-0.0151
I believe that natural pork is from animals not raised in gestation crates.	[0.6746]	0.0187
I rarely think about the use of gestation crates when purchasing hog products.	-0.0291	[0.5582]
I would be less likely to purchase pork products which are generated using gestation crates.	[0.7391]	-0.1110
I would likely pay more if pork products were generated under the ban of gestation crates.	[0.7152]	-0.0101
The use of gestation crates by farmers reduces pork prices.	0.1962	[0.6563]
The use of gestation crates increases productivity of hog operations.	-0.0043	[0.7863]
All things considered, the use of gestation crates by farmers is good for consumers.	-0.2783	[0.7310]

Notes: The brackets indicate the specific factor each belief item assigned to.

3.3 Results

We first show the estimation results of each treatment group separately in Table 6. The significant positive coefficient of Factor 1 for each group suggest that consumers perceiving the GCL with higher quality of products and higher animal welfare are willing to pay a higher premium for pork chops produced under the GCL. Individuals who believe the gestation crates generate benefits to consumers such as the lower price are WTP lower premium for pork chops produced under the GCL, as indicated by the negative coefficient of Factor 2 for each group.

We compare the mean WTP in Table 6 between each treatment group. We find the WTP of the "Quality" group is significantly lower than the WTP of the "Quality & Welfare" group and the "Welfare" group ($p = 0.039$, $p = 0.028$). There are no significant differences in the mean WTP between the baseline and the other groups, nor between the "Quality & Welfare" group and the "Welfare" group.

We further robustly compare the WTP of each group by setting treatment groups as independent variables in pooled models. Table 7 contains estimation results of pooled models with and without controls. We treat the Baseline and the Quality group as the base group, respectively. In both results, the "Quality & Welfare" and the "Welfare" treatments significantly increase consumers' WTP compared to the Baseline or the "Quality" group. We do not find any significant effects of the "Quality" treatment on WTP compared to the baseline. In pooled models, the coefficients of

Table 6: Censored regression results of each treatment group.

	Baseline	Quality & Welfare	Quality	Welfare
Constant	5.622*** (0.866)	7.425*** (0.868)	5.724*** (0.927)	6.312*** (0.999)
Age	-0.001(0.008)	-0.015^{**} (0.007)	-0.006 (0.008)	0.004 (0.008)
Male	0.309 (0.245)	-0.191 (0.234)	-0.256 (0.253)	0.113 (0.253)
Household Income	0.009 (0.018)	0.007 (0.016)	0.030* (0.017)	0.018 (0.017)
Education	0.054 (0.120)	0.070 (0.109)	-0.061 (0.120)	-0.069 (0.118)
Children in the Household	0.094 (0.310)	-0.245 (0.279)	0.124 (0.319)	-0.140 (0.347)
Household Size	-0.184^{*}(0.106)	-0.208^{**} (0.097)	-0.006 (0.116)	-0.073 (0.126)
Weekly Pork Consumption	0.041 (0.112)	0.011 (0.117)	0.034 (0.127)	-0.024 (0.112)
Resident of the GSL state	0.029 (0.256)	-0.043 (0.231)	-0.239 (0.258)	-0.057 (0.267)
Factor 1	0.495*** (0.123)	0.517*** (0.101)	0.454*** (0.118)	0.615*** (0.123)
Factor 2	-0.312^{***} (0.119)	-0.161 (0.114)	-0.347^{***} (0.122)	-0.214^{*} (0.126)
Mean WTP	$5.764 (0.131)	$5.978 (0.141)	$5.584 (0.126)	$6.010 (0.144)
95% WTP Conf. Interval	[$5.506, $6.022]	[$5.702, $6.253]	[$5.337, $5.831]	[$5.728, $6.291]
Log-likelihood	-297.935	-243.681	-281.985	-282.627
No. Obervation	268	258	243	270

Notes: Factor 1: Pork products quality and animal welfare. Factor 2: Benefits of gestation crates to consumers. Standard errors are presented in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

"household income" and "household size" become significant due to the increase in power. The results of the pooled model analysis are in line with our comparisons in Table 6. We argue that the increase in the WTP of pork products produced under the GCL is mainly from consumers' considerations about the animal welfare, instead of the quality of products, which is in line with the intention of the law.

3.3.1 Consumer Beliefs

Next, we address how our treatment information affect consumers' WTP. Understanding mechanisms behind our results above is important for labeling strategies of related regulations in the market. We first investigate whether the treatment information changes consumers' beliefs. The separation of beliefs and preferences is a growing part of the WTP literature (Howard, Roe, Interis, & Martin, 2020; Lusk, Schroeder, & Tonsor, 2014; Neill & Williams, 2016; Roe, Interis, & Howard, 2018). We did not find any significant differences in the distribution of each belief

Table 7: Censored regression results of pooled models.

	vs Baseline	vs Baseline	vs Quality group	vs Quality group
Q&W group	0.295* (0.158)	0.376** (0.154)	0.444*** (0.161)	0.454*** (0.156)
Welfare group	0.241 (0.154)	0.282* (0.150)	0.390** (0.157)	0.360** (0.153)
Quality group	-0.149 (0.155)	-0.078 (0.150)		
Baseline			0.149 (0.155)	0.078 (0.150)
Constant	5.705*** (1.114)	6.153*** (0.462)	5.556*** (1.116)	6.075*** (0.464)
Age		-0.005 (0.004)		-0.005 (0.004)
Male		-0.004 (0.123)		-0.004 (0.123)
Household Income		0.017** (0.008)		0.017** (0.008)
Education		-0.003 (0.058)		-0.003 (0.058)
Children in the Household		-0.058 (0.155)		-0.058 (0.155)
Household Size		−0.117** (0.055)		−0.117** (0.055)
Weekly Pork Consumption		0.016 (0.057)		0.016 (0.057)
Resident of the GSL state		-0.099 (0.126)		-0.099 (0.126)
Factor1		0.514*** (0.058)		0.514*** (0.058)
Factor2		−0.276*** (0.060)		−0.276*** (0.060)
Controls	No	Yes	No	Yes
Log-likelihood	-1177.065	-1116.243	-1177.065	-1116.243
No. Obervation	1039	1039	1039	1039

Notes: Standard errors are presented in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

between any two groups by the Kolmogorov-Smirnov test. This means the treatment information is not likely to change or affect people's original beliefs.

It is also possible that when the information is different from or "offend" consumers' original beliefs, the information may have negative effects on consumers' WTP. We test this mechanism by focusing on the "Quality" group. In our main results, the WTP of the "Quality" group is significantly less than it in the "Quality & Welfare" and the "Welfare" groups. Although the difference in WTP between the Baseline and "Quality" group is not significant, the WTP of the "Quality" group is less than it in the Baseline, and the coefficient of "Quality" group is negative in the pooled model estimation (column 2 in Table Table 7). It is counter-intuitive for us to infer that people do not like high-quality products. Thus, it is possible that due to the treatment information in the "Quality" group does not match people's related original beliefs, they exhibit a lower WTP. We

Table 8: Pooled Double-bounded model estimation for mechanism tests.

	vs Quality Treatment sub-group (a)	vs Quality Treatment sub-group (b)	vs Quality Treatment sub-group (c)	vs Quality Treatment sub-group (d)
Treat QW	0.405^{**} (0.162)	0.751^{**} (0.320)	0.275 (0.211)	0.579^{***} (0.184)
Treat W	0.314^{**} (0.159)	0.656^{**} (0.318)	0.179 (0.206)	0.495^{***} (0.184)
Baseline	0.033 (0.156)	0.380 (0.319)	-0.100 (0.204)	0.213 (0.180)
Constant	6.034^{***} (0.473)	5.969^{***} (0.587)	6.310^{***} (0.532)	5.961^{***} (0.486)
Age	-0.005 (0.004)	-0.004 (0.004)	-0.004 (0.004)	-0.004 (0.004)
Male	-0.013 (0.004)	0.083 (0.138)	0.049 (0.136)	0.030 (0.126)
Household Income	0.014^{**} (0.009)	0.014 (0.010)	0.010 (0.009)	0.018^{**} (0.009)
Education	0.005 (0.059)	0.013 (0.065)	0.034 (0.064)	-0.021 (0.060)
Children in the Household	-0.021 (0.157)	-0.160 (0.174)	-0.134 (0.170)	-0.058 (0.155)
Household Size	-0.114^{**} (0.055)	-0.159^{**} (0.062)	-0.141^{**} (0.060)	-0.128^{**} (0.057)
Weekly Pork Consumption	0.021 (0.058)	0.008 (0.064)	0.031 (0.064)	0.004 (0.058)
Resident of the GSL state	-0.059 (0.128)	-0.088 (0.142)	-0.099 (0.140)	-0.039 (0.130)
Factor1	0.512^{***} (0.060)	0.520^{***} (0.066)	0.527^{***} (0.065)	0.516^{***} (0.060)
Factor2	-0.271^{***} (0.473)	-0.250^{***} (0.067)	-0.241^{***} (0.068)	-0.245^{***} (0.065)
Log-likelihood	-1075.943	-870.407	-950.612	-995.077
No. Obervation	1009	826	915	920

Notes: Sub-group (a): in the "Quality" group, only subjects whose answers to the belief question are NOT "strongly disagree" or "disagree" are included. Sub-group (b): in the "Quality" group, only subjects whose answers to the belief question are "strongly disagree" or "disagree" are included. Sub-group (c): in the "Quality" group, only subjects whose scores of factor 2 are less than or equal to 0 are included. Sub-group (d): in the "Quality" group, only subjects whose scores of factor 2 are greater than 0 are included. Standard errors are presented in parentheses. $*p < 0.10, **p < 0.05, ***p < 0.01$.

only include the subjects whose answers to the belief question "I believe a ban of gestation crates increases the quality of pork products" are NOT "Strongly disagree" or "Disagree" in the pooled model estimation, i.e., the belief-matched group. However, the results do not change as shown in the column sub-group (a) of Table 8. The results are also robust when we only include the subjects whose answers to the belief question are "Strongly disagree" or "Disagree", i.e., the belief-against group, shown in the column sub-group (b) of Table 8. This suggests that the treatment information is not likely to affect WTP whether different or similar to consumers' original beliefs.

The third mechanism we test is the two factors we identified from the factor analysis. Factor scores indicate associations consumers make with the use of gestation crates. We do not find significant differences in scores of Factor1 ("quality and animal welfare") across groups. However, we find the average scores of Factor2 ("benefits of gestation crates to consumers") are higher in the "Quality & Welfare" and the "Quality" group compared to the Baseline or the "Welfare" group, as shown in Table 9. We further estimate the pooled model only including sub-groups in the "Quality" group based on scores of Factor2. Table 8 shows that the significant differences in WTP between the other groups and the "Quality" group we found in our main results disappear when we only include individuals in the "Quality" group whose Factor2 scores are less than or equal to 0 (shown in column sub-group c). In addition, the sign of the coefficient of the Baseline group becomes negative. We argue that the differences in WTP between other groups and the "Quality" group are mainly from individuals in the "Quality" group whose scores of Factor2 are greater than 0 (shown in column sub-group d). These results suggest that the quality treatment information make consumers associate the use of gestation crates with more "benefits to consumers", and then reduce their WTP for products produced under the ban of gestation crates.

Table 9: Mean Factor2 scores

	Baseline	Quality & Welfare	Quality	Welfare
Mean scores	-0.067	0.065	0.098	-0.084
p-values	B vs QW: 0.096	B vs Q: 0.081	W vs QW: 0.029	W vs Q: 0.027
p-values	B vs W: 0.651	QW vs Q: 0.901		

Notes: P-values are from Two-sided Mann-Whitney U-Tests.

In conclusion, our main results first show that consumers care about animal-welfare, and they are willing to pay more for products produced under regulations which increase animal welfare. We then show why treatment information related to only increasing the quality of products reduces WTP. When the treatment information is related to benefits of consumers including the quality of goods, it is not likely to affect WTP by changing consumers' related original beliefs or matching or "offending" consumers' original beliefs. The information may affect WTP by enhancing the "benefits to consumers" associations of consumers make with related regulations. In other words, the information increases the importance or weights consumers place on the factor "benefits to consumers" when they think about related regulations. Thus, consumers reduce their WTP for products produced under regulations which are good for animal welfare, but may increase the price or reduce the productivity, which are not good for consumers' benefits. Our results can be generalized for labeling strategies of products produced under animal-welfare related regulations, and then provide important marketing suggestions to business owners or policy makers.

4 Discussion

Analyzing the observational data in a causal framework allows for estimating both the costs (to production) and benefits (to income) of a prominent animal welfare law. The results from the consumer preference study then help us understand where the change in income is originating from in the causal inference analysis, and how a change in perceptions might change the value. This work has implications for the larger understanding of changes in value for animal welfare regulations based on consumer beliefs/perceptions. Moreover, this work can inform other states' policymakers on the potential changes in costs and benefits based on consumer perceptions before the implementation of such policies. This study addresses the sparse literature that examines both costs and benefits from animal welfare regulations across multiple states.

We take into account consumer beliefs and uncover that welfare information is the most effective for increasing WTP for pork related products. Yet, animal welfare laws have heterogeneous effects on producer value and income, but most importantly has negative impacts on quantity of hog sales. Further work should examine disaggregated effects of such policies and determine if the benefits truly outweigh the costs of the regulation.

4.1 Policy Implications

From this two-sided view of the FAW regulation, the benefits of such interventions are heterogeneous in their impacts on overall value. The intrinsic motivation of consumers for such regulation is more heavily embedded in increased animal welfare standards rather than other beliefs about quality of meat products as suggested by previous literature on vote-buy gaps (Kehlbacher, Bennett, & Balcombe, 2012; Tonsor, Wolf, & Olynk, 2009). This results lends itself to the idea that policy-makers should take a multifaceted approach to improving farm animal welfare as suggested by previous literature (see Clark, Stewart, Panzone, Kyriazakis, and Frewer (2017) for a larger analysis and discussion of these studies). Moreover future policy should take into account that while these regulations aim to increase the social welfare of the meat production system, the cost to producers comes in terms of significantly fewer pounds of hogs sold, and an initial decline in gross income that slowly recovers to pre-implementation of the regulations. As noted by prior work, producers are likely to bear the majority of costs when FAW are implemented due to the fact that consumer preferences are heterogeneous about the law and the products for which they may affect (Ortega & Wolf, 2018).

These policies are also not ubiquitous across states in terms of application. For example, these regulations only specify that hogs raised in a state with these laws are must meet these requirements, but meat products sold within that state could be produced under either situation if it is imported from another state. So, only if meat is marketed as produced under these regulations are consumers able to make informed choices. This contributes to the ambiguous benefits of the animal welfare improvements. While policymakers should consider the social costs of such regulations, the private costs and welfare of consumers and producers should also be balanced to ensure the ability to maintain livelihoods. This can come by reducing costs to abide by such regulations or enhancing marketing efforts to couple the animal welfare regulations with proper promotion of state agricultural marketing campaigns (i.e Grown in California). Prior research has showed that redundant labels can be effective in regaining market share for products produced with multiple attributes that appeal to consumers (Ufer, Ortega, & Wolf, 2022). Other options include providing producer subsidies to help offset the initial decline in gross income experienced by producers (Bennett, 1997). Coupling policies for FAW regulation and marketing/branding campaigns and/or providing subsidies could create a win-win scenario for social and private welfare.

References

Bennett, R. M. (1997). Farm animal welfare and food policy. *Food policy*, *22*(4), 281–288. [27]

Bishop, R. C., & Heberlein, T. A. (2019). The contingent valuation method. In *Economic valuation of natural resources* (pp. 81–104). Routledge. [16]

Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*. [7, 8, 15, 39, 40, 42]

Boxall, P. C., & Adamowicz, W. L. (2002). Understanding heterogeneous preferences in random utility models: A latent class approach. *Environmental and Resource Economics*, *23*(4), 421–446. [20]

Caetano, C., Callaway, B., Payne, S., & Rodrigues, H. S. (2022). *Difference in differences with time-varying covariates.* [6]

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200-230. (Themed Issue: Treatment Effect 1) doi: https://doi.org/10.1016/j.jeconom.2020.12.001 [6, 7, 8, 9, 10, 11, 13, 39, 40, 41, 42, 43]

Carlier, A., & Treich, N. (2020, April). Directly Valuing Animal Welfare in (Environmental) Economics. *International Review of Environmental and Resource Economics*, *14*(1), 113-152. doi: 10.1561/101.00000115 [1]

Carlsson, F., Frykblom, P., & Lagerkvist, C. J. (2007). Farm animal welfare? Testing for market failure. *Journal of Agricultural and Applied Economics*, *39*(1), 61–73. [1]

Carson, R. T. (2012). Contingent valuation: A practical alternative when prices aren't available. *Journal of Economic Perspectives*, *26*(4), 27–42. [16]

Chen, X., Liu, Y., Jaenicke, E. C., & Rabinowitz, A. N. (2019). New concerns on caffeine consumption and the impact of potential regulations: The case of energy drinks. *Food Policy*, *87*, 101746. [2]

Clark, B., Stewart, G. B., Panzone, L. A., Kyriazakis, I., & Frewer, L. J. (2017). Citizens, consumers and farm animal welfare: A meta-analysis of willingness-to-pay studies. *Food Policy*, *68*, 112–127. [27]

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–96. [8]

Diener, A., O'Brien, B., & Gafni, A. (1998). Health care contingent valuation studies: A review and classification of the literature. *Health Economics*, *7*(4), 313–326. [16]

Dörnyei, K. R., & Gyulavári, T. (2016). Why do not you read the label?–an integrated framework of consumer label information search. *International Journal of Consumer Studies*, *40*(1), 92–100. [2]

Edenbrandt, A. K., & Lagerkvist, C.-J. (2021). Is food labelling effective in reducing climate impact by encouraging the substitution of protein sources? *Food Policy*, *101*, 102097. [2]

Gardner, J. (n.d.). Two-stage differences in differences. [7, 15, 39, 40, 42]

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*. [37, 38]

Hanemann, M., Loomis, J., & Kanninen, B. (1991). Statistical efficiency of double-bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics*, *73*(4), 1255–1263. [16]

Hausman, J. (2012). Contingent valuation: from dubious to hopeless. *Journal of Economic Perspectives*, *26*(4), 43–56. [16]

Howard, G., Roe, B. E., Interis, M. G., & Martin, J. (2020). Addressing attribute value substitution in discrete choice experiments to avoid unintended consequences. *Environmental and Resource Economics*, *77*(4), 813–838. [22]

Hudson, D. (2010). Discussion: The economics of animal welfare. *Journal of Agricultural and Applied Economics*, *42*(3), 453–455. [1, 2]

Imbens, G. W. (2021, August). Statistical significance, p-values, and the reporting of uncertainty. *Journal of Economic Perspectives*, *35*(3), 157-74. doi: 10.1257/jep.35.3.157 [8]

Kehlbacher, A., Bennett, R., & Balcombe, K. (2012). Measuring the consumer benefits of improving farm animal welfare to inform welfare labelling. *Food Policy*, *37*(6), 627–633. [26]

Lai, C.-Y., List, J. A., & Samek, A. (2020). Got milk? Using nudges to reduce consumption of added sugar. *American Journal of Agricultural Economics*, *102*(1), 154–168. [2]

Liljenstolpe, C. (2008). Evaluating animal welfare with choice experiments: An application to Swedish pig production. *Agribusiness*, *24*(1), 67-84. doi: https://doi.org/10.1002/agr.20147 [1]

Lin, W., & Nayga Jr, R. M. (2022). Green identity labeling, environmental information, and pro-environmental food choices. *Food Policy*, *106*, 102187. [2]

Lopez-Feldman, A. (2012). Introduction to contingent valuation using Stata. [20]

Lusk, J. L., Nilsson, T., & Foster, K. (2007). Public preferences and private choices: effect of altruism and free riding on demand for environmentally certified pork. *Environmental and Resource Economics*, *36*(4), 499–521. [1]

Lusk, J. L., Norwood, F. B., & Pruitt, J. R. (2006). Consumer demand for a ban on antibiotic drug use in pork production. *American Journal of Agricultural Economics*, *88*(4), 1015-1033. doi: https://doi.org/10.1111/j.1467-8276.2006.00913.x [1]

Lusk, J. L., Schroeder, T. C., & Tonsor, G. T. (2014). Distinguishing beliefs from preferences in food choice. *European Review of Agricultural Economics*, *41*(4), 627–655. [22]

McCluskey, J. J., Grimsrud, K. M., Ouchi, H., & Wahl, T. I. (2003). Consumer response to genetically modified food products in japan. *Agricultural and Resource Economics Review*, *32*(2), 222–231. [16]

McFadden, D., & Train, K. (2017). *Contingent valuation of environmental goods: A comprehensive critique*. Edward Elgar Publishing. [16]

Neill, C. L., & Holcomb, R. B. (2019). Does a food safety label matter? Consumer heterogeneity and fresh produce risk perceptions under the Food Safety Modernization Act. *Food Policy*, *85*, 7–14. [16]

Neill, C. L., & Williams, R. B. (2016). Consumer preference for alternative milk packaging: The case of an inferred environmental attribute. *Journal of Agricultural and Applied Economics*, *48*(3), 241–256. [16, 22]

Olynk, N. J., Tonsor, G. T., & Wolf, C. A. (2010). Verifying credence attributes in livestock production. *Journal of Agricultural and Applied Economics*, *42*(3), 439–452. [1]

Ortega, D. L., & Wolf, C. A. (2018). Demand for farm animal welfare and producer implications: Results from a field experiment in michigan. *Food Policy*, *74*, 74–81. [27]

Rimal, A. (2005). Meat labels: consumer attitude and meat consumption pattern. *International Journal of Consumer Studies*, *29*(1), 47–54. [2]

Roe, B. E., Interis, M. G., & Howard, G. E. (2018). Utilizing subjective beliefs in stated preference models: Issues and solutions. [22]

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *arXiv preprint arXiv:2201.01194*. [8, 14]

Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101-122. doi: https://doi.org/10.1016/j.jeconom.2020.06.003 [7, 39, 40, 42]

Sumner, D. A., Matthews, W. A., Mench, J. A., & Rosen-Molina, J. T. (2010). The economics of regulations on hen housing in california. *Journal of Agricultural and Applied Economics*, *42*(3), 429–438. [1]

Tonsor, G. T., Olynk, N., & Wolf, C. (2009). Consumer preferences for animal welfare attributes: The case of gestation crates. *Journal of Agricultural and Applied Economics*, *41*(3), 713–730. [1]

Tonsor, G. T., Wolf, C., & Olynk, N. (2009). Consumer voting and demand behavior regarding swine gestation crates. *Food Policy*, *34*(6), 492-498. doi: https://doi.org/10.1016/j.foodpol.2009.06.008 [1, 20, 26]

Tonsor, G. T., & Wolf, C. A. (2010). Drivers of resident support for animal care oriented ballot initiatives. *Journal of Agricultural and Applied Economics*, *42*(3), 419–428. [1]

Ufer, D., Ortega, D. L., & Wolf, C. A. (2022). Information and consumer demand for milk attributes: Are redundant labels an effective marketing strategy? *Applied Economic Perspectives and Policy*, *44*(2), 960–981. [27]

Van Der Merwe, D., Kempen, E. L., Breedt, S., & De Beer, H. (2010). Food choice: Student consumers' decision-making process regarding food products with limited label information. *International Journal of Consumer Studies*, *34*(1), 11–18. [2]

Venkatachalam, L. (2004). The contingent valuation method: A review. *Environmental Impact Assessment Review*, *24*(1), 89–124. [16]

Viske, D., Lagerkvist, C. J., & Carlsson, F. (2006). Swedish consumer preferences for animal welfare and biotech: A choice experiment. [1]

Weaver, D., & Finke, M. (2003). The relationship between the use of sugar content information on nutrition labels and the consumption of added sugars. *Food Policy*, *28*(3), 213–219. [2]

Zhu, C., Lopez, R. A., & Liu, X. (2016). Information cost and consumer choices of healthy foods. *American Journal of Agricultural Economics*, *98*(1), 41–53. [2]

Online Appendix

A Additional Tables and Figures

Table A.1: Summary of Gestating Sow Crate Laws

State	Year	Format of Provision	Details
Arizona	2006	Ballot Measure (Prop 204, enacted as Ariz. Rev. Stat. Ann. § 13-2910.07)	Prohibits confining veal calves and gestating sows where they cannot lie down, fully extend limbs, or turn around freely for the majority (or all) of the day.
California	2008	Ballot Measure (Prop 2, enacted as Cal. Health and Safety Code, Ch. 13.8 § 25990)	Prohibits confining veal calves, gestating sows, and hens where they cannot lie down, fully extend limbs, or turn around freely.
Colorado	2008	Legislation (Colo. Rev. Stat. Ann. § 35-50.5-102)	Prohibits confining veal calves and gestating sows where they cannot stand up, lie down, or turn around freely.
Florida	2002	Ballot Measure (Amendment 10, codified in Fla. Const. Art. X, § 21)	Prohibits gestation crates that prevent a pig from turning around freely; also includes a general statement that inhumane treatment of animals is a concern of Florida citizens.
Maine	2009	Legislation (Me. Rev. Stat. tit. 7, § 4020)	Prohibits confining veal calves and gestating sows where they cannot stand up, lie down, fully extend limbs, or turn around freely for the majority (or all) of the day.
Michigan	2009	Legislation (Mich. Comp. Laws. Ann. § 287.746)	Prohibits cruel confinement of veal calves, gestating sows, and hens where they cannot stand up, lie down, fully extend limbs, or turn around freely for the majority (or all) of the day.
Oregon	2007	Legislation (Or. Rev. Stat. § 600.150)	Prohibits confinement of gestating sows for more than twelve hours in a way that does not allow them to lie down, fully extend limbs, and turn around freely.
Rhode Island	2012	Legislation (4 R.I. Gen. Laws Ann. §4-1.1-3)	Prohibits confining veal calves and gestating sows where they cannot stand up, lie down, fully extend limbs, or turn around freely.
Ohio	2025	Regulations (Ohio Admin. Code §901:12-5-03 [veal]; §901:12-8 [swine])	Requires veal calves to be housed in group pens by 10 weeks of age where they can stand without impediment, rest using normal postures, groom, eat, turn around, and lie down. Limits sow gestation crates after 2025 to post-weaning for a period of time that seeks to maximize embryonic welfare and allows for confirmation of pregnancy.

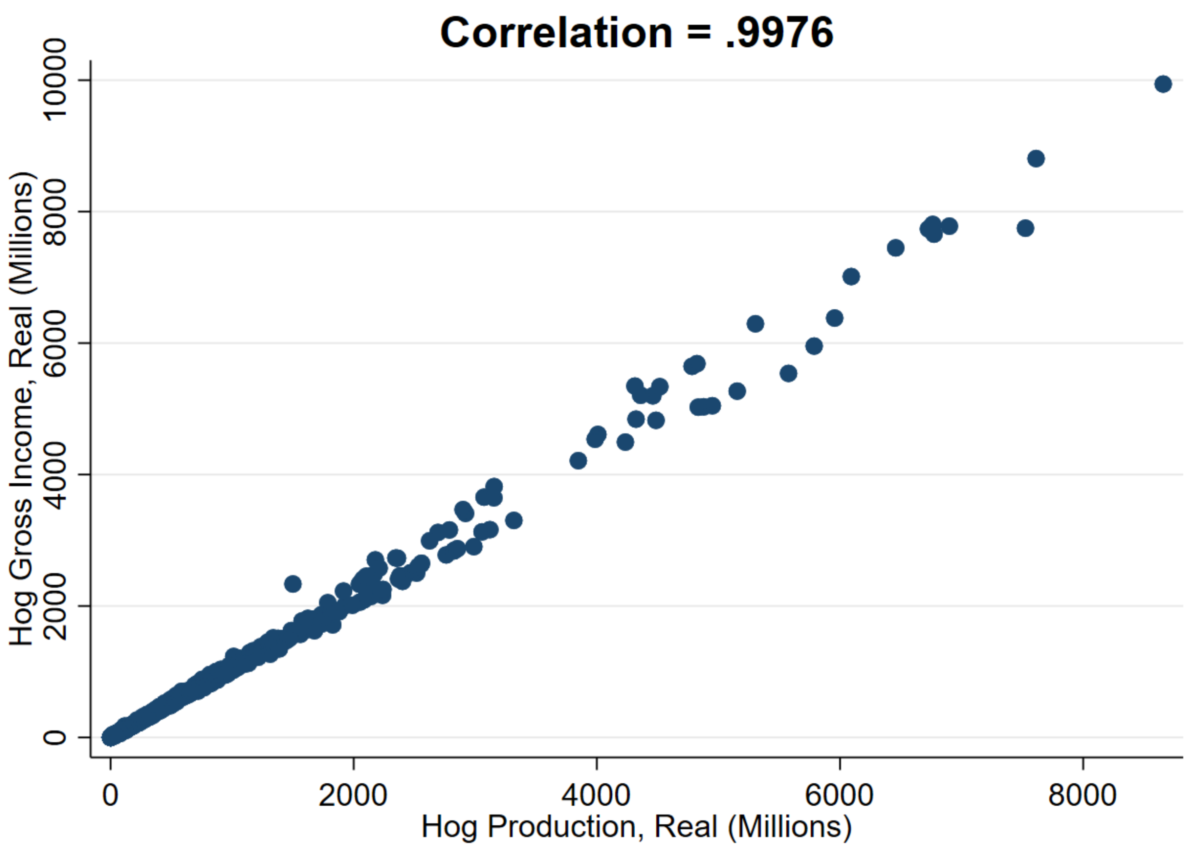
Table A.2: Summary Statistics

Panel A: Aggregate

	Mean	Median	SD	Min	Max	Count
Pounds of Hog Sold (millions)	596.28	83.31	1486.31	0	15077.2	1634
Hog Gross Income, Real (Millions)	402.02	58.86	969.11	0	9941.1	1634
Hog Production, Real (Millions)	373.64	53.69	879.48	0	8658.5	1634
Value (Thousands)	822.88	458.50	1065.80	0	8892.0	1634
Gest. Sows Law Effective	0.06	0.00	0.25	0	1.0	1650

Note: Value refers to home consumption, excluding inter-state trade, as defined by the USDA. Hog data comes from the USDA NASS Quick Stats portal. Gestating sow laws comes from the Animal Welfare Institute.

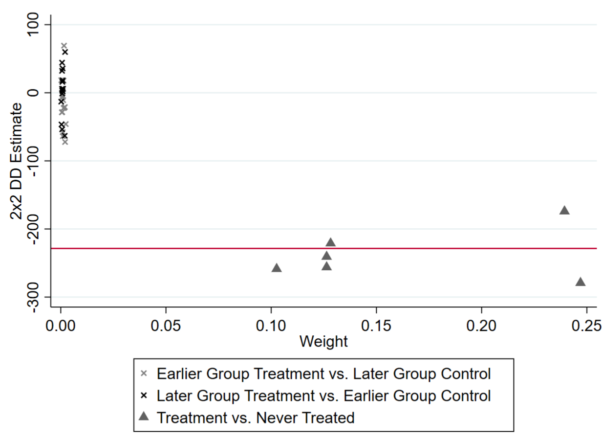
Figure A.1: Scatterplot of Producer Gross Income and Production Income



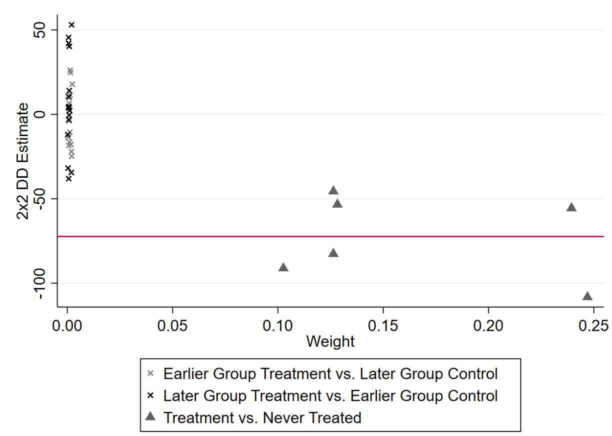
Note: This scatterplot shows the unconditional correlation between hog gross income and hog production. The data is at the state-year level. The unconditional correlation coefficient is shown at the top.

Figure A.2: Goodman-Bacon TWFE Diagnoses

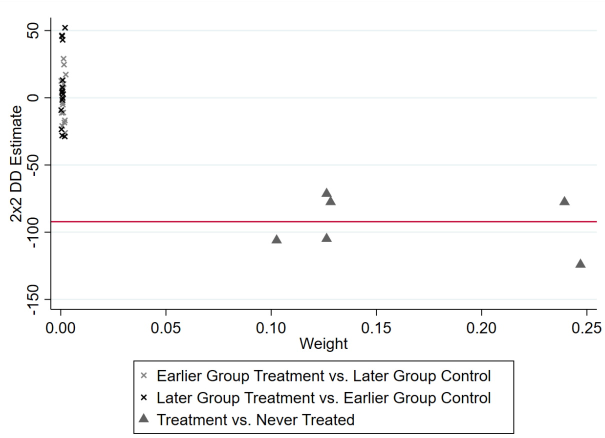
(a) Pounds Sold



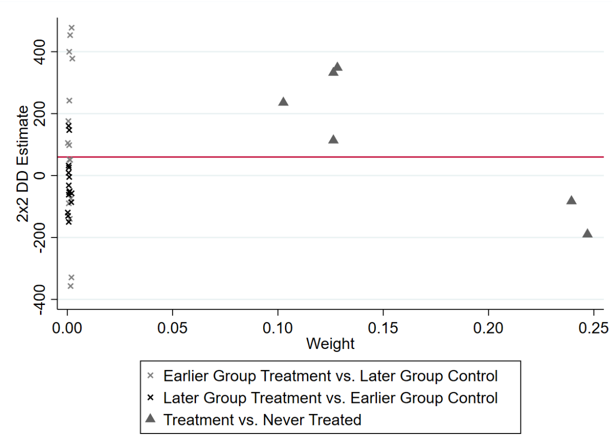
(b) Production



(c) Gross Income



(d) Value



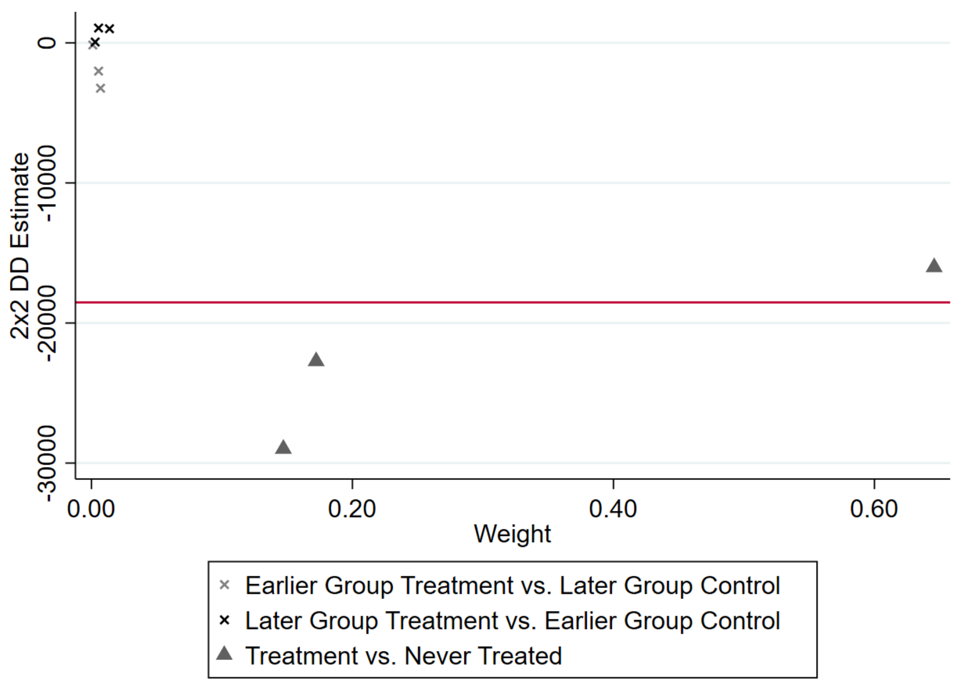
Note: Scales differ. Figures show the decomposition of the TWFE static estimate into its respective 2X2 comparisons (Goodman-Bacon, 2021). The x-axis is the weight placed on the comparison by TWFE. The y-axis is the effect size of the "2X2" used in the aggregate effect (Goodman-Bacon, 2021). Triangles are the unproblematic comparisons. The red line shows the TWFE estimate. In (a), the dependent variable is pounds of hog sold. In (b), the dependent variable is producer production income. In (c), the dependent variable is producer gross income. In (d), the dependent variable is home consumption value.

Table A.3: Summary Statistics, County-Level

	Mean	Median	SD	Min	Max	Count
Hog Sold (Head, Naive Missing)	74455.45	967.50	315549.00	3	9251890.0	11552
Hog Sold (Head, Naive Zero)	54693.46	281.00	272437.51	0	9251890.0	15726
Hog Sold (Head, Merge Zero)	67311.73	651.00	300828.60	0	9251890.0	12778

Note: Data comes from the USDA NASS Quick Stats Portal. Hog Sold (Head, Naive Missing) is the raw data in which missing values in county-year NASS data and county-years with no merge are both treated as missing. Hog Sold (Head, Naive Zero) replaces all missing values with zero regardless of source of missingness. Hog Sold (Head, Merge Zero) replaces all unmerged counties with zero, but leaves NASS data missings as such.

Figure A.3: Goodman Bacon Decomposition, County-Level



Note: Figures show the decomposition of the TWFE static estimate into its respective 2X2 comparisons (Goodman-Bacon, 2021). The x-axis is the weight placed on the comparison by TWFE. The y-axis is the effect size of the "2X2" used in the aggregate effect (Goodman-Bacon, 2021). Triangles are the unproblematic comparisons. The red line shows the TWFE estimate. In (a), the dependent variable is pounds of hog sold. In (b), the dependent variable is producer production income. In (c), the dependent variable is producer gross income. In (d), the dependent variable is home consumption value.

Table A.4: Estimates of Effect of Gestating Sows Law on Production, County Level, CS Only

	(1) Naive Miss	(2) Naive Zero	(3) Merge Zero
Gestating Sows Law	-24906.66	-57895.57	20001.51
	(26311.10)	(52385.81)	(20020.22)
Observations	15726	15726	15726

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Cluster-robust standard errors, by state, in parentheses. All regressions include county and year fixed effects. Gestation Crate Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. The dependent variable is number of hogs sold (in head). All columns use the group-time effect aggregation from Callaway and Sant'Anna (2021). It uses never treated as comparison units, the doubly robust differences-in-differences estimator based on inverse probability of tilting and weighted least squares (Sant'Anna & Zhao, 2020), and shows the simple aggregation.

Table A.5: Estimates of Effect of Gestation Crate Law on Production, County Level and Cluster

Panel A: Naïve Missing		
	(1)	(2)
Gestation Crate Law	-18780.83***	-24414.04***
	(3047.50)	(3545.26)
	[0.00]	
Observations	11341	11552
TWFE	X	-
Imputation Method^{a}	-	X
Panel B: Naïve Zero		
	(1)	(2)
Gestation Crate Law	-14403.45***	-14880.09***
	(3267.10)	(2656.09)
	[0.00]	
Observations	15726	15726
TWFE	X	-
Imputation Method^{a}	-	X
Panel C: Merge Zero		
	(1)	(2)
Gestation Crate Law	-16602.70***	-21468.43***
	(2739.21)	(3263.78)
	[0.00]	
Observations	12656	12778
TWFE	X	-
Imputation Method^{a}	-	X

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Cluster-robust standard errors, by county, in parentheses. All regressions include county and year fixed effects. The dependent variable is hogs sold, measured in head. Gestation Crate Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. In Panel A, all missing outcomes are left missing. In Panel B, all missing outcomes are replaced with zero. In Panel C, only missing values that arise due to no county-year in the USDA Quick Stats data are counted as zeroes. Column 1 estimates the law's effect using two-way fixed effects OLS. ^{a}Column 2 uses the imputation estimator of Borusyak et al. (2021) and Gardner (n.d.).

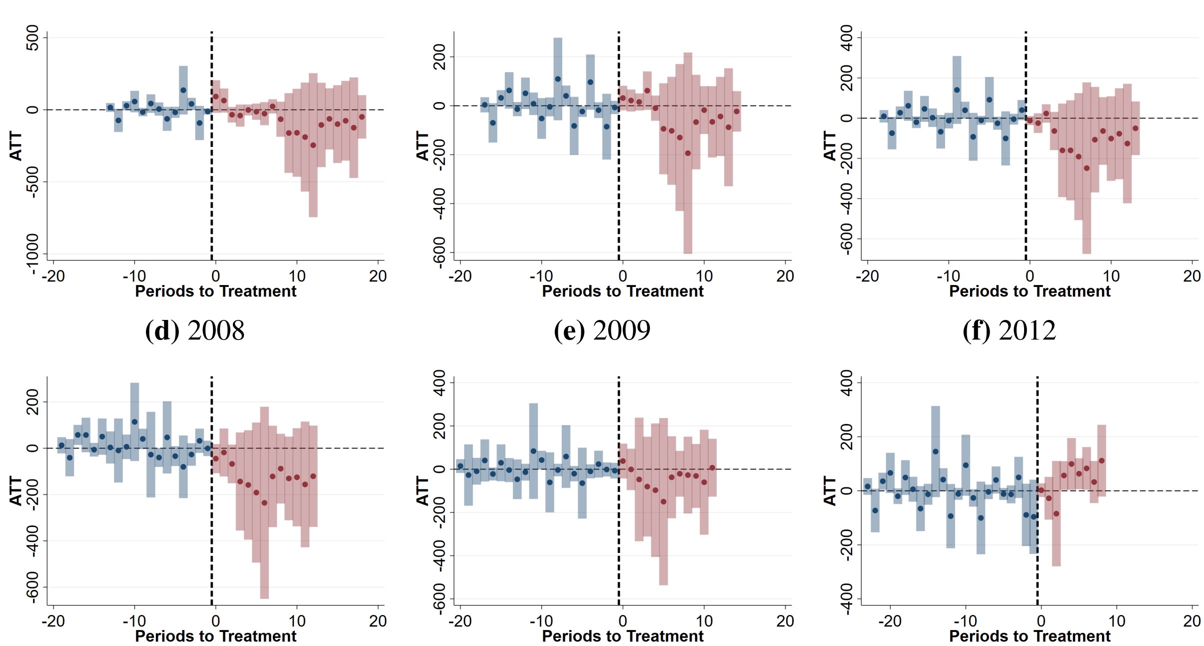
A.1 Production Income Results

Table A.6: Estimates of Effect of Gestation Crate Laws on Production

Panel C: Production	(1)	(2)	(3)
Gestation Crate Law	-72.27	-73.07	-67.58
	(56.63)	(54.38)	(43.81)
	[0.21]		
Observations	1634	1634	1650
TWFE	X	-	-
Imputation Method[a]	-	X	-
CS Method[b]	-	-	X

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Cluster-robust standard errors, by state, in parentheses. All regressions include county and year fixed effects. Gestation Crate Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. The dependent variable is production income. Column 1 estimates the law's effect using two-way fixed effects OLS. [a]Column 2 uses the imputation estimator of Borusyak et al. (2021) and Gardner (n.d.).[b]Column 3 uses group-time effect aggregation from Callaway and Sant'Anna (2021). It uses never treated as comparison units, the doubly robust differences-in-differences estimator based on inverse probability of tilting and weighted least squares (Sant'Anna & Zhao, 2020), and shows the simple aggregation.

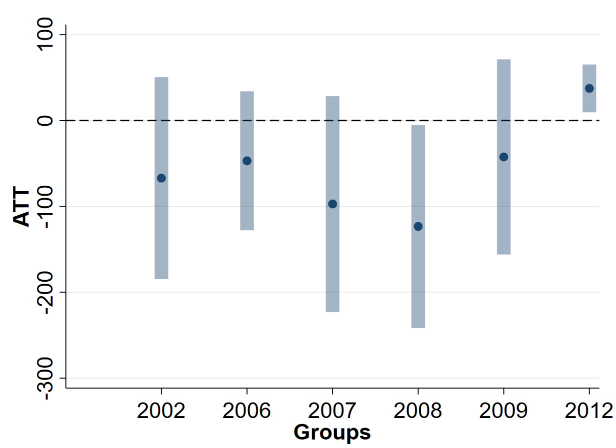
Figure A.4: Event Studies by Group for Production



(a) 2002 **(b)** 2006 **(c)** 2007 **(d)** 2008 **(e)** 2009 **(f)** 2012

Note: The y-scale differs across figures. These figures report coefficients from Equation 2 for the estimation of dynamic effects of the law change on producer production income. Each figure represents a different treatment cohort. The coefficients are obtained from the event-study aggregation of Callaway and Sant'Anna (2021). Coefficients represent the change in pounds of hog sold for states that adopt the gestation crate size law relative to non-adopting state in the years before and after the law passes, as compared with the year immediately prior to the law passing. Gestation crate law adoption dates comes from the Animal Welfare Institute and consumer value comes from the USDA NASS Quick Stats. The whiskers represent 95% confidence intervals. The standard errors are clustered at the state-level and are calculated using the wild bootstrap.

Figure A.5: Production Income Group Level Averages



Note: Estimates are group-level aggregates from Callaway and Sant'Anna (2021). Wild bootstrapped standard errors. 95% confidence intervals shown.

A.2 Observational Value Results

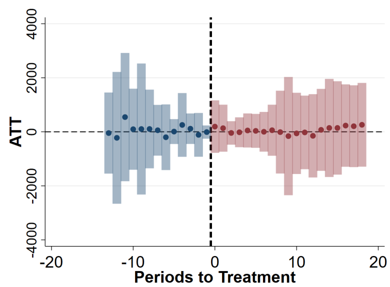
Table A.7: Estimates of Effect of Gestation Crate Laws on Pounds Sold, Producer Gross Income, and Value

Panel D: Value	(1)	(2)	(3)
Gestation Crate Law	60.14	75.76	14.13
	(139.19)	(134.41)	(64.39)
	[0.67]		
Observations	1634	1634	1650
TWFE	X	-	-
Imputation Method[a]	-	X	-
CS Method[b]	-	-	X

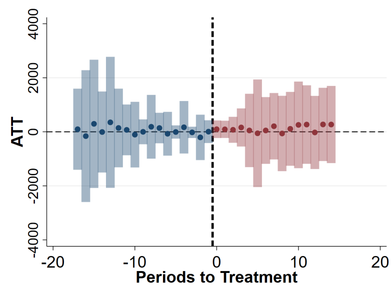
Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Cluster-robust standard errors, by state, in parentheses. All regressions include county and year fixed effects. Gestation Crate Law is a binary variable equaling 1 if it is in post-adoption period in a state that passes a law dictating minimum gestation crate sizes. In Panel A, the dependent variable is pounds of hog sold. In Panel B, the dependent variable is producer gross income. In Panel C, the dependent variable is the value of farm home consumption, excluding inter-state trade. Column 1 estimates the law's effect using two-way fixed effects OLS. [a]Column 2 uses the imputation estimator of Borusyak et al. (2021) and Gardner (n.d.).[b]Column 3 uses group-time effect aggregation from Callaway and Sant'Anna (2021). It uses never treated as comparison units, the doubly robust differences-in-differences estimator based on inverse probability of tilting and weighted least squares (Sant'Anna & Zhao, 2020), and shows the simple aggregation.

Figure A.6: Event Studies by Group for Value

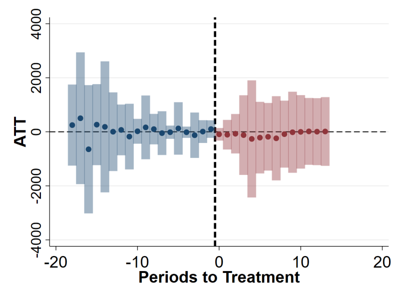
(a) 2002



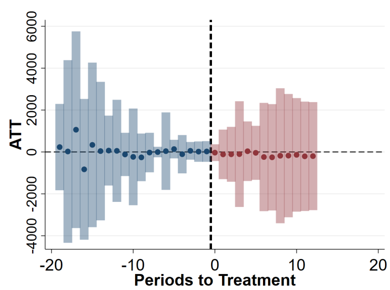
(b) 2006



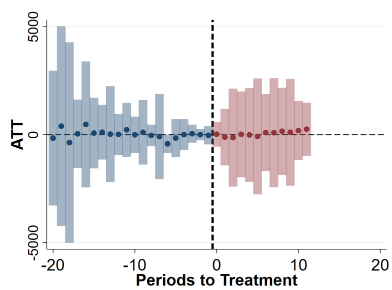
(c) 2007



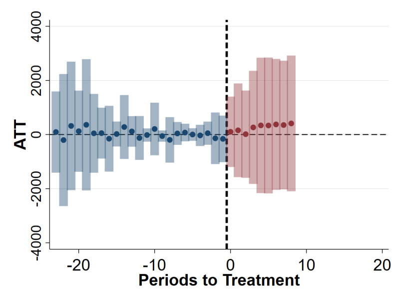
(d) 2008



(e) 2009

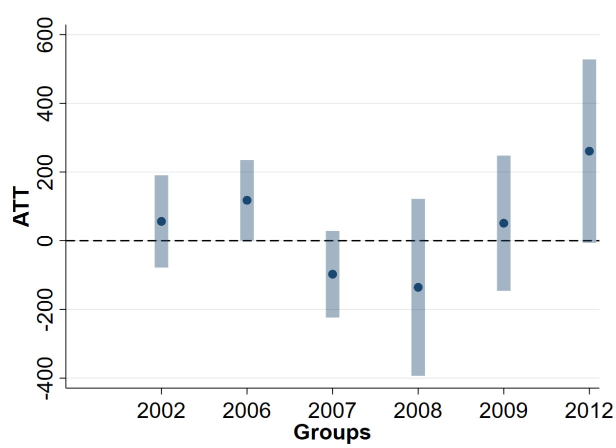


(f) 2012



Note: The y-scale differs across figures. These figures report coefficients from Equation 2 for the estimation of dynamic effects of the law change on consumer value. Each figure represents a different treatment cohort. The coefficients are obtained from the event-study aggregation of Callaway and Sant'Anna (2021). Coefficients represent the change in pounds of hog sold for states that adopt the gestation crate size law relative to non-adopting state in the years before and after the law passes, as compared with the year immediately prior to the law passing. Gestation crate law adoption dates comes from the Animal Welfare Institute and consumer value comes from the USDA NASS Quick Stats. The whiskers represent 95% confidence intervals. The standard errors are clustered at the state-level and are calculated using the wild bootstrap.

Figure A.7: Value Group Level Averages



Note: Estimates are group-level aggregates from Callaway and Sant'Anna (2021). Wild bootstrapped standard errors. 95% confidence intervals shown.

B Experimental instructions.

B.1 First stage of the contingent valuation.

a. Baseline

Before you proceed with the survey, please carefully read the following information. It will be repeated in the next few questions, but we ask that you take the time to read it carefully now.

Gestation crates (sometimes also called gestation stalls) refer to metal crates (approximately 7 feet long and 2 feet wide) that house female breeding stock (hogs/sows) in individually confined areas during an animal's four-month pregnancy. As of 2022, eight passed the laws against using gestation crates. Specifically, the law prohibits the use of crates in gestating hogs/sows so that they can stand up, lie down, or turn around freely.

Please answer the following questions as if you were actually making shopping decisions in the grocery stores.

Pork chops at $(4.0/4.5/5.0/5.5) per pound. The pork is produced under the Gestating Sow Law (i.e., no use of gestation crates.)

Pork chops at $4.50 per pound. The pork is not produced under the Gestating Sow Law (i.e., use of gestation crates.)

b. Quality & welfare treatment

Before you proceed with the survey, please carefully read the following information. It will be repeated in the next few questions, but we ask that you take the time to read it carefully now.

Gestation crates (sometimes also called gestation stalls) refer to metal crates (approximately 7 feet long and 2 feet wide) that house female breeding stock (hogs/sows) in individually confined areas during an animal's four-month pregnancy. As of 2022, eight passed the laws against using gestation crates. Specifically, the law prohibits the use of crates in gestating hogs/sows so that they can stand up, lie down, or turn around freely.

Research has found that the such Gestation Crate Laws increase the quality of pork products, and increases the welfare of hogs/sows.

Please answer the following questions as if you were actually making shopping decisions in the grocery stores.

Pork chops at $(4.0/4.5/5.0/5.5) per pound. The pork is produced under the Gestating Sow Law (i.e., no use of gestation crates.)

Pork chops at $4.50 per pound. The pork is not produced under the Gestating Sow Law (i.e., use of gestation crates.)

c. Quality treatment

Before you proceed with the survey, please carefully read the following information. It will be repeated in the next few questions, but we ask that you take the time to read it carefully now.

Gestation crates (sometimes also called gestation stalls) refer to metal crates (approximately 7 feet long and 2 feet wide) that house female breeding stock (hogs/sows) in individually confined areas during an animal's four-month pregnancy. As of 2022, eight passed the laws against using gestation crates. Specifically, the law prohibits the use of crates in gestating hogs/sows so that they can stand up, lie down, or turn around freely.

Research has found that the such Gestation Crate Laws increase the quality of pork products.

Please answer the following questions as if you were actually making shopping decisions in the grocery stores.

Pork chops at $(4.0/4.5/5.0/5.5) per pound. The pork is produced under the Gestating Sow Law (i.e., no use of gestation crates.)

Pork chops at $4.50 per pound. The pork is not produced under the Gestating Sow Law (i.e., use of gestation crates.)

d. Welfare treatment

Before you proceed with the survey, please carefully read the following information. It will be repeated in the next few questions, but we ask that you take the time to read it carefully now.

Gestation crates (sometimes also called gestation stalls) refer to metal crates (approximately 7 feet long and 2 feet wide) that house female breeding stock (hogs/sows) in individually confined areas during an animal's four-month pregnancy. As of 2022, eight passed the laws against using gestation crates. Specifically, the law prohibits the use of crates in gestating hogs/sows so that they can stand up, lie down, or turn around freely.

Research has found that the such Gestation Crate Laws increase the welfare of hogs/sows.

Please answer the following questions as if you were actually making shopping decisions in the grocery stores.

Pork chops at $(4.0/4.5/5.0/5.5) per pound. The pork is produced under the Gestating Sow Law (i.e., no use of gestation crates.)

Pork chops at $4.50 per pound. The pork is not produced under the Gestating Sow Law (i.e., use of gestation crates.)

B.2 Belief questions.

1. I believe a ban of gestation crates increases the quality of pork products.
2. I believe pork products generated using gestation crates are subject to more food safety risks.
3. Animals should not reside in confined areas like gestation crates.
4. I believe a ban on gestation crates increases the welfare of hogs.
5. I believe that natural pork is from animals not raised in gestation crates.
6. I rarely think about the use of gestation crates when purchasing hog products.
7. I would be less likely to purchase pork products which are generated using gestation crates.
8. I would likely pay more if pork products were generated under the ban of gestation crates.
9. The use of gestation crates by farmers reduces pork prices.

10. The use of gestation crates increases productivity of hog operations.

11. All things considered, the use of gestation crates by farmers is good for consumers.

12. Please mark the answer to this question as “Somewhat disagree”.

B.3 Demographic survey.

1. Please enter your age in years. (Please do not enter birth year)

2. Please indicate your gender.

3. Please indicate the highest level of education you have complicated.

4. What is your approximate annual household income before taxes?

5. Are there children under the age of 18 living in your household? If so, how many

6. Including yourself, how many people are currently living in your household?

7. What percentage of your household’s purchasing decisions (e.g. groceries) do you make?

8. What is the number of times per week you or someone in your household consumes pork products?

9. Please indicate your race.

10. In which state do you currently reside?

C Double-bounded model estimation process.

We define y_i^1 and y_i^2 as the dichotomous variables that capture the response to the choice questions in the first and second stage, then the probability that an individual chooses GSL products in the first stage and non-GSL products in the second stage (case a) can be expressed as:

$$Pr(y_i^1 = 1; y_i^2 = 0|z_i) = Pr(s,n) \tag{4}$$

$$\begin{aligned} Pr(s,n) &= Pr(p_1 \leq WTP \leq p_2) \\ &= Pr(p_1 \leq z_i^{'}\beta + u_i \leq p_2) \\ &= Pr\left(\frac{p_1 - z_i^{'}\beta}{\sigma} \leq \frac{u_i}{\sigma} < \frac{p_2 - z_i^{'}\beta}{\sigma}\right) \\ &= \Phi\left(\frac{p_2 - z_i^{'}\beta}{\sigma}\right) - \Phi\left(\frac{p_1 - z_i^{'}\beta}{\sigma}\right) \end{aligned} \tag{5}$$

Using symmetry of the normal distribution, we have the probability as:

$$Pr(s,n) = \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_1}{\sigma}\right) - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right) \tag{6}$$

In case (b) where individuals choose GSL products in both stages ($p_2 > p_1$), i.e., $y_i^1 = 1$ and $y_i^2 = 1$.

$$\begin{aligned} Pr(s,s) &= Pr(WTP > p_1, WTP > p_2) \\ &= Pr(z_i^{'}\beta + u_i > p_1, z_i^{'}\beta + u_i > p_2) \end{aligned} \tag{7}$$

Since $p_2 > p_1$, $Pr(z_i^{'}\beta + u_i > p_1 | z_i^{'}\beta + u_i > p_2) = 1$, and hence we have:

(8)
$$
\begin{aligned}
Pr(s,s) &= Pr(u_i \geq p_2 - z_i^{'}\beta) \\
&= 1 - \Phi\left(\frac{p_2 - z_i^{'}\beta}{\sigma}\right)
\end{aligned}
$$

By symmetry, $Pr(s,s) = \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right)$.

In case (c) where individuals choose non-GSL products in the first stage, and choose GSL products in the second stage ($p_2 < p_1$), i.e., $y_i^1 = 0$ and $y_i^2 = 1$.

(9)
$$
\begin{aligned}
Pr(n,s) &= Pr(p_2 \leq WTP \leq p_1) \\
&= Pr(p_2 \leq z_i^{'}\beta + u_i \leq p_1) \\
&= Pr\left(\frac{p_2 - z_i^{'}\beta}{\sigma} \leq \frac{u_i}{\sigma} < \frac{p_1 - z_i^{'}\beta}{\sigma}\right) \\
&= \Phi\left(\frac{p_1 - z_i^{'}\beta}{\sigma}\right) - \Phi\left(\frac{p_2 - z_i^{'}\beta}{\sigma}\right)
\end{aligned}
$$

(10)
$$
Pr(n,s) = \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right) - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_1}{\sigma}\right)
$$

In case (d) where individuals choose non-GSL products in both stages ($p_2 < p_1$), i.e., $y_i^1 = 0$ and $y_i^2 = 0$.

$$
\begin{aligned}
Pr(n,n) &= Pr(WTP < p_1, WTP < p_2) \\
&= Pr(z_i^{'}\beta + u_i < p_1, z_i^{'}\beta + u_i < p_2) \\
&= Pr(z_i^{'}\beta + u_i < p_2) \\
&= \Phi\left(\frac{p_2 - z_i^{'}\beta}{\sigma}\right)
\end{aligned} \tag{11}
$$

So, $Pr(n,n) = 1 - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right)$.

Finally, based on the four possible cases, the following likelihood function is used to estimate β and σ:

$$
\begin{aligned}
\sum_{i=1}^{N} \Bigg\{ & d_i^{sn} ln \left(\Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_1}{\sigma}\right) - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right)\right) \\
& + d_i^{ss} ln \left(\Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right)\right) \\
& + d_i^{ns} ln \left(\Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right) - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_1}{\sigma}\right)\right) \\
& + d_i^{nn} ln \left(1 - \Phi\left(\frac{z_i^{'}\beta}{\sigma} - \frac{p_2}{\sigma}\right)\right) \Bigg\}
\end{aligned} \tag{12}
$$

where $d_i^{sn}, d_i^{ss}, d_i^{ns}, d_i^{nn}$ are the indicator variables of the four cases, taking the value of 1 or 0.