The Impact of Corporate Social Responsibility on Animal Welfare Standards: Evidence From the Cage-free Egg Industry

Samara Mendez* Jacob Peacock†

December 13, 2023

Abstract

Corporate social responsibility initiatives have become increasingly popular to address animal welfare issues. Over the last decade, hundreds of companies globally have committed to source cage-free eggs to improve welfare for laying hens. In many cases, these commitments were made following public pressure campaigns by advocates for farmed animals, representing a large portion of all spending on advocacy for farmed animals. The effect of these commitments to improve the issues that they target is not well known, and a better understanding of the impact may affect millions of dollars of animal advocacy funding. This study uses panel data from 44 countries over 13 years to estimate the effect of commitments on the proportion of hens in cage-free housing, while controlling for legislation and bird deaths from avian influenza. We use a correlated random effects model to account for unobserved heterogeneity between countries and unobserved correlation over time within individual countries. We test for potential endogeneity due to a feedback loop between commitments and housing using a control function test with pressure campaigns as an instrument for commitments. We estimate that one additional commitment in a country leads to an average 0.035 percentage point increase in cage-free housing in that country. Similarly, the average effect of an additional year under cage-free legislation is a 0.04 percentage point increase, although not statistically significant. We discuss implications of these results and comment on limitations and future extensions of this research.

*Rethink Priorities. ORCID: 0000-0002-8179-744X. Corresponding author: samara@rethinkpriorities.org

†Rethink Priorities. ORCID: 0000-0002-4834-8132.

Executive Summary

Corporate social responsibility initiatives have become popular to address animal welfare issues, like commitments to source cage-free eggs to improve layer hen welfare. This paper investigates **whether corporate cage-free egg commitments cause changes in the proportion of hens living in cage-free housing.**

We attempt to answer this question with observational data and statistical methods to reduce bias.

- We create **a new data set observing 44 countries over 13 years**, tracking four variables: hen housing, corporate cage-free commitments, cage-free legislation, and highly pathogenic avian influenza (see Section 3.1).
- We use a correlated random effects model to identify the effect of commitments on housing separately from other factors that may bias our estimate (see Section 3.2.1).
- We test for potential bias due to: a causal feedback loop between commitments and housing; missing observations; a potential connection between Germany's high imports and high commitments; and variable transformations (see Section 3.2.2).

We find that, on average within a given country, a one-commitment increase leads to a 0.035 (95% CI: 0.01 – 0.06) percentage point rise in the share of cage-free hen housing.

- Because of the large number of hens in the industry, the impact of a few new commitments per year is substantial. For example, the 16 new commitments made by Canadian food businesses in 2017 represent a 0.56 percentage point increase in Canadian cage-free housing, which in turn may affect **nearly 147,000 Canadian hens** (see Section 5).
- Our tests for potential bias suggest that our estimation method is robust to several different factors that typically affect observational studies (see Section 4). However, we acknowledge that some bias may still affect our results. **Therefore, we place moderate confidence in our findings.**

We recommend:

- Advocates should continue to pursue corporate commitments to increase the prevalence of cage-free housing for layer hens.
- Given the strength of evidence found here for a causal effect of cage-free commitments, similar commitments might be explored for other animal welfare improvements.
- Advocates should fund future research to determine whether the interaction between corporate commitments and legislation magnify or dampen their individual effects.
- Advocates and funders should continue tracking data on welfare commitments and support better data collection of welfare outcomes and other covariates for future welfare interventions.

Contents

1 Introduction

Consumers generally value animal welfare [19, 36, 10]. However, this valuation does not always manifest in purchase behavior, and some explanations for this gap include information asymmetries [56] and overlapping heuristics [77]. The misalignment between values and purchases represents a market failure both through the externalization of poor animal welfare and the loss of consumer surplus. Previous research has shown that caged egg consumers will vote to ban the eggs that they purchase, suggesting that legislation of certain welfare standards may be one solution to correcting this market failure [56]. Alternately, firms can use market-based corporate social responsibility (CSR) initiatives like public commitments and product labels to correct the information asymmetry, differentiating the credence qualities[1] of their products and reducing the externality of animal suffering.

In 2005, food corporations in the United States (US) began to signal their intention to improve animal welfare by committing to source eggs from layer hens in cage-free housing. Layer hens housed in battery cages, the predominant commercial housing system, are intensively confined in less than half a square foot of space per bird [76] and suffer from numerous welfare issues [54, pp. 12–17]. Cage-free housing for layer hens provides improved welfare, allowing hens space to move more freely and express natural behaviors. As of 2022, nearly 500 retailers, restaurants, and food service companies with operations in the US have made cage-free commitments [15], as have many food businesses globally. These commitments often include deadlines to complete the sourcing transition within several years. Commitments are sometimes facilitated by dialogue with animal advocacy groups, like The Humane Society of the United States, The Humane League,[2] and Mercy for Animals. In other cases, the commitments followed public campaigns by animal advocacy groups employing pressure tactics, including on-the-ground protests, newspaper ads, petitions, shareholder activism, and undercover investigations.

This study evaluates the impact of corporate commitments to source cage-free eggs on layer hen housing. We construct a novel longitudinal data panel tracking a group of 44 countries on an annual basis between 2006 and 2018, gathering data on hen housing conditions, corporate commitments, pressure campaigns, legislation to ban cages,[3] and the number of birds killed by highly pathogenic avian influenza (henceforth *avian influenza* or *flu* in variable names) outbreaks. Our main hypothesis is that corporate cage-free commitments increase the proportion of layer hens living in cage-free housing systems in a country's egg industry. Figure 1 motivates this hypothesis in the case of the US. Animal advocates and industry analysts have publicly expressed support for the hypothesis [67, 75] and further

[1] Qualities that cannot be observed by the consumer even after purchase, like labor, animal welfare, or environmental protection standards during production [22].

[2] See Section 6.3 "Conflicts of interest" for more information on the relationship between The Humane League and the authors' affiliations.

[3] Here we use *legislation* to refer to any law affecting hen housing, rather than the legal sense of a law passed specifically by a legislative body.

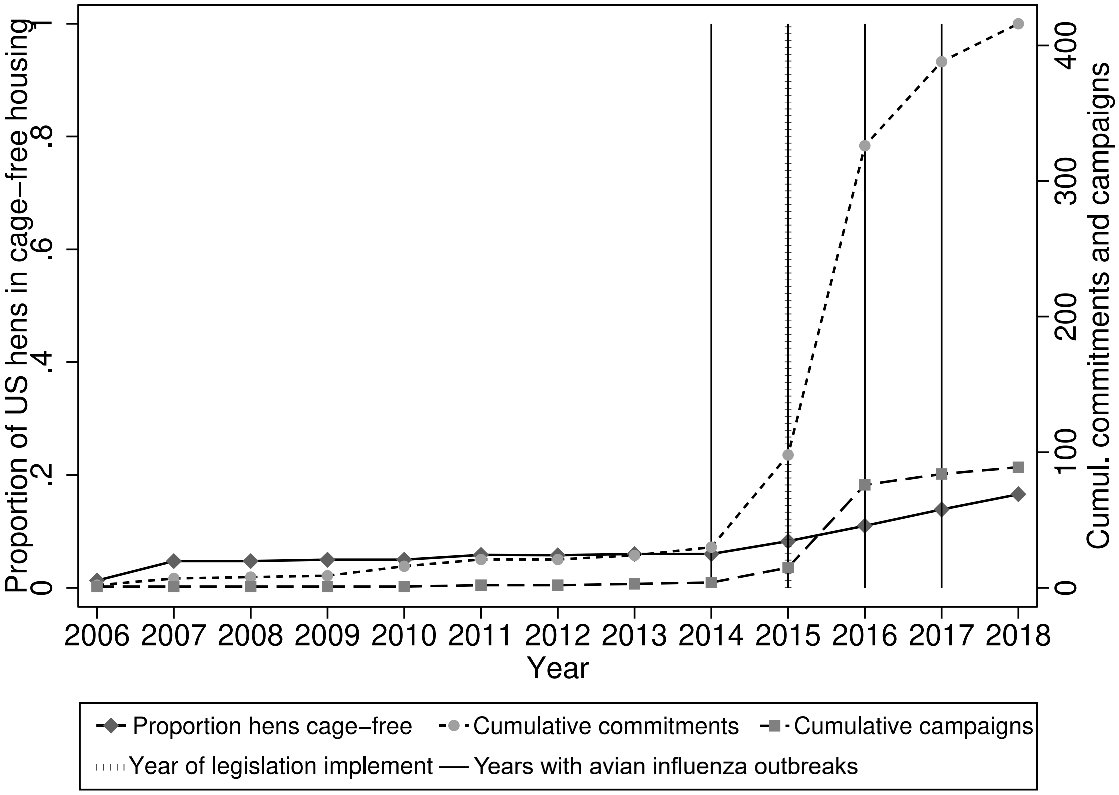


Figure 1: The cumulative number of corporate commitments in the US and the proportion of US hens in cage-free housing from 2006 to 2018. The two are closely correlated, with an increase in cage-free housing associated with the increase in commitments, preceded by years of only modest increases in cage-free housing. In 2015, California implemented Proposition 2 and AB 1437, which combined affect egg producers throughout the US. Campaigns appear to lag commitments partly because many commitments were pledged by corporate initiatives or through dialogue with animal advocates, without the need for advocates to escalate to pressure campaign tactics. Data for this article are available at `https://osf.io/VTE94/`.

support is found in previous studies of the US cage-free commitments discussed below.

We estimate the causal effect of commitments on the proportion of cage-free housing, with legislation and deaths from avian influenza as covariates, using an identification strategy that takes advantage of the longitudinal nature of the data to address the various sources of bias that arise in non-randomized studies. We are first concerned with endogeneity from unobserved variables that differ by country, such as consumer preferences or price premiums. Second, we consider a potential feedback loop between commitments and housing, as food retailers and their egg suppliers may make their decisions simultaneously. We use a correlated random effects (CRE) model, which can account for the endogeneity and accommodate the nonlinear (fractional) dependent variable. The model is a flexible alternative to the more common fixed effects (FE) model, as it allows for correlation between explanatory variables and unobserved heterogeneity within clusters and also decomposes the policy-relevant within effect (the impact of increasing commitments within countries over time) and the between effect (the effect of average commitments on housing in a country). We test for evidence of simultaneity using a control function instrumental variables method, with pressure campaigns as an instrument for commitments, following Wooldridge [82]. The results do not show evidence of such simultaneous endogeneity. We identify our preferred specification, a nonlinear CRE model, by examining diagnostic tests and residual plots.

Our results show that the average effect of a one-commitment increase is a 0.035 percentage point increase in the proportion of cage-free housing. Similarly, the average effect of implementing cage-free legislation is a 0.04 percentage point increase in cage-free housing (not statistically significant). We do not find a meaningful nor statistically significant effect on cage-free housing from avian influenza. The between effects show that a one-commitment increase in a country's average commitments leads to an average increase of 0.21 percentage points in the country's cage-free housing. For legislation, the between effect indicates the difference between a country with no legislation at any time during the sample period versus one with legislation throughout the entire period. We find a much larger 60.88 percentage point increase in a country's cage-free housing, although this effect should be interpreted with caution due to limited variation in the average years under legislation.

Our results are, to the best of our knowledge, the first global, quantitative estimates of the causal impact of commitments on cage-free housing, adding to a small literature that investigates the impact of corporate animal welfare commitments on the standards they attempt to improve. Several studies have previously examined the role of corporate commitments in increasing the percentage of hens in cage-free housing, as well as improving the welfare of chickens raised for meat, termed *broiler chickens*. Most directly relevant to our work is Sarek's [64] quantitative model of US cage-free egg production, which attributes an increase of between 2.1 and 10 percentage points in cage-free egg production from 2005 to 2018 to corporate commitments. To estimate the counterfactual percentage of cage-free housing, the analysis uses consumer willingness-to-pay and demand, retail prices and price premiums between caged and cage-free eggs, avian influenza rates, and historical trends. As the author acknowledges, data limitations restrict the explanatory power of the estimates. Šimčikas [68] examines the cost-effectiveness of advocacy to obtain corporate commitments in improving animal welfare, including both layer hens and broiler chickens. The calculation of the number of chickens affected by commitments includes a descriptive analysis of the counterfactual scenario without intervention by animal advocates. The study estimates between 9 and 120 years of chicken life (combing laying hens and broiler chickens) are improved per dollar spent by advocacy groups to obtain corporate commitments. Capriati [13] uses a case series to examine the role of animal advocacy in obtaining corporate commitments, including two cage-free egg and two broiler commitments, and estimates the cost-effectiveness of these efforts. The report positively links animal advocacy efforts to corporate commitments and estimates 10 years of layer hen life are improved per dollar spent [13, p. 58]. Collectively, these studies all suggest that corporate commitments, and advocacy efforts to obtain those commitments, caused increases in the percentage of cage-free housing. In addition, Šimčikas [68] and Capriati [13] suggest corporate commitments have caused improvements in the welfare of broiler chickens. However, these studies are limited by a lack of rigorous strategies for causal identification, a focus on the US, and the exclusion of the potential effects of legislation (though Šimčikas [68] qualitatively discusses how legislation could affect cost-effectiveness estimates).

Two additional studies examine broiler chicken welfare commitments. Saatkamp [63] examine factors that contributed to the Dutch broiler chicken industry's transition to higher welfare standards. Using expert elicitation, they find "initiating and triggering actions by non-governmental organizations" [63, p. 1] and a willingness on the part of the entire value chain to change as two of five main causal factors. Furthermore, a "rapid and complete change" [63, p. 2] in the industry toward higher welfare standards is attributed to a 2012 decision by several large retailers to discontinue carrying conventional broiler chicken meat in their stores. Reis and Molento [62] use content analysis of the annual reports of two major broiler chicken producers and five semi-structured interviews with industry decision makers to conclude that the European adoption of broiler chicken welfare standards may improve animal welfare. While these studies have the same limitations as the previously cited studies, the findings again support the role of corporate commitments in improving animal welfare and advocacy campaigns in achieving those commitments.

Campaigns by advocacy groups and the resulting corporate commitments have also been suggested to cause reductions in the number of rabbits harmed in cosmetics testing, animals killed for fur production, and dolphins harmed as bycatch in the tuna industry. Animal Rights International ran numerous corporate campaigns against the testing of cosmetics on animals and popularized the strategy in animal advocacy [67]. In a book-length case series, Singer [69, p. 113] partially attributes reductions in the number of animals harmed in cosmetics testing in the late 1970s and the eventual abandonment of animal testing by many large cosmetics companies in 1989 [69, p. 135] to these campaigns and commitments. In a case study of the global fur industry, Bollard [12] attributes some of the recent decline in the number of animals killed for fur to targeted campaigns that obtained 1,017 fur-free corporate commitments. Lastly, in a case study of the tuna industry, Mitchell [48] found corporate pledges to source dolphin-safe tuna were caused by consumer boycotts and environmental advocacy. These pledges may have then led to US and international government regulation, which ultimately succeeded in reducing the number of dolphins harmed through tuna fishing. These studies suggest a more general effect of corporate commitments, and campaigns by advocacy groups to obtain those commitments, for improving animal welfare across a variety of animal industries and theories of change.

Our study also informs the broader literature on CSR and market-driven governance, where a thread of research focuses on the effectiveness of corporate commitments, pledges, and voluntary labeling in causing change in the supply chain. Empirical research on the impact of corporate initiatives and commitments to limit deforestation shows mixed results, ranging from small reductions in deforestation [49, 35] to no statistically significant effect [50, 7]. The empirical literature on the effect of voluntary suppliers' codes of conduct on workers' rights generally finds a small but statistically significant effect [26], and the effect size tends to increase with improvements to study methodology such as more comparable control groups and finer-grained data [52, 53]. These results suggest positive effects of corporate commitments in non-animal industries as well. This study will extend the literature to consider the impact of CSR as it pertains to animal welfare.

We contribute to the existing literature on cage-free commitments in the following ways. First, we improve on attempts to directly identify their causal impact on cage-free housing, using a novel data set and panel data estimation models to account for confounding factors that might bias the results. Second, we improve the literature by examining global commitments and explicitly including legislation and avian influenza outbreaks as additional control variables. In particular, estimating of the impact of legislation adds value to the literature, since laws may also be important determinants of the percentage of cage-free housing in a country.

The rest of the paper proceeds as follows. In Section 2, we develop our hypotheses against the background of the egg industry and the historical setting of our study. Section 3 describes the data as well as the estimation and identification strategies. Section 4 presents the results, which suggest that an increase in commitments leads to an increase in cage-free housing in a country over time, and higher average commitments in a country are correlated with more cage-free housing in any year. Section 5 concludes with a discussion of the implications of these results, potential limitations, and possible future extensions.

2 Conceptual Framework and Hypothesis Development

To establish our main hypotheses, we provide institutional background about the modern egg industry and historical context around the study period to understand the most important features determining layer hen housing. As with many agricultural supply chains, the egg industry is highly integrated, so the egg producer often owns the hatchery, laying barns, and processing facilities [16]. This integration minimizes transaction costs, reduces the number of decision-makers, and gives producers control over production capabilities [16]. In this section, we focus on the aspects of egg production that inform how egg producers choose their flock sizes and housing systems year after year. We also confine the historical narrative to broad examples of how avian influenza outbreaks, commitments, and legislation affected the egg industry.

2.1 Productive Cycles and Short-run Production Variables

The most recognizable stage of egg production is *laying*, in which *table-egg-laying hens* produce eggs intended for human consumption. Egg producers raise egg-laying hens in flocks, which are groups of chickens of approximately the same age. Once a flock of young chickens, *pullets*, enters the egg-laying barn, the egg producer can add new pullets to incomplete flocks for up to three weeks [76]. These hens will remain together in the same barn for the length of the laying cycle, which depends on breed, season, consumer preferences, feed costs, and egg prices [65]. The laying cycle is typically 60 to 72 weeks but may be extended to 109 weeks through induced molting [30, 58, 65]. Molting is the annual process in which birds lose their feathers and temporarily slow or stop egg production, and the natural molting cycle can be induced by farmers through light or feed restriction, which can be harmful to hen welfare [6]. After the molting period, hens return to another, potentially shorter, laying cycle, in which they produce larger, higher-quality eggs at a faster rate [58]. Producers can vary the length of cycles, molting, timing of flock replacement, and reductions of stocking density to respond to short-term changes in market conditions [65, 85, 25].

2.2 Housing Types and Long-run Production Variables

Egg producers rarely under-utilize space in existing hen housing facilities, so the decision to increase cage-free flocks requires either building new housing structures or converting existing structures. Different housing systems have different implications for flock management and animal welfare, so we provide a high-level overview of the features of these systems to better understand the producers' investment decisions as well as the advocacy and legislative history discussed below. The most relevant housing systems for this study are battery cage and cage-free housing; however, we briefly examine features of several other relevant housing systems.

Most egg-laying hens are currently housed in *battery* cages (also called *non-enriched* or *barren* cages), which are stacked in rows, and are themselves inside a barn. These cages provide very little space per hen and provide no way for hens to express natural behaviors, resulting in poor hen welfare [23, 24]. Enriched cages, commonly used in European Union member countries, have more space per hen and contain enrichments like nests, perching space, litter, and feed troughs within the cages to allow hens to express natural behaviors [28]. Enriched cages are seen as an intermediate welfare improvement between battery cages and cage-free housing [9]. *Cage-free* housing includes a variety of systems without cages. Specifically, hens must be able to move in a way that promotes their welfare, be protected from predators, and have access to litter. Hens are provided enrichments such as perches, nests, and scratching areas to allow them to perform natural behaviors [2]. *Aviary* and *barn* systems are two examples of cage-free housing, in which the birds can freely roam inside the buildings. In barn systems, birds generally live on one level, while aviary systems have multiple levels for perching [5]. Other systems may exceed the housing space requirements of cage-free systems, including *pasture-raised*, as defined and certified by American Humane and Certified Humane, and USDA-certified *free-range* [2, 38, 4].

The transition from battery cages to cage-free housing in the US is estimated by industry analysts to cost between $30 and $40 per bird (depending on cage-free housing type, and including both long-term capital costs as well as variable operation costs) and require several years to complete [33]. From an empirical modeling perspective, we expect that the changes in percentage of hens in cage-free housing are influenced less by short-term variables like previous flock size and more by long-term variables that enter the housing investment decision like long-run price trends, demands of downstream corporate customers (e.g., cage-free commitments), retail consumer demand, and legislation.

2.3 Highly Pathogenic Avian Influenza

For six months starting in December 2014, an avian influenza virus swept through US chicken and turkey flocks. Over 50 million birds died from contracting avian influenza directly or were killed to quarantine the avian influenza, representing 12% of the US egg-laying flock [61]. The sudden supply shock increased US egg prices to record highs and reduced exports, and the stigma of the disease impacted both domestic and overseas consumer demand. Although observed outbreaks of the disease ended sharply in June 2015, the effects on the market and market prices were long-lasting. Historically often volatile, egg prices in 2016 were even more volatile than usual, and the attempt to rebuild flocks after the depopulation caused imbalances in the usually-seasonal supply of eggs [61]. As prices of conventional eggs (the vast majority of eggs produced at the time) increased and the price premium between cage-free and conventional eggs shrank, more consumers chose cage-free eggs in the grocery stores. Ramos et al [61] suggests that expanded demand for cage-free eggs alongside the need to replace flocks after the avian influenza outbreak may have prompted egg producers to invest in cage-free housing and increase their cage-free flock sizes relative to conventional flocks.

Avian influenza impacted countries in Europe less dramatically, and the birds affected were largely not egg-laying hens [3]. India and China both suffered from frequent outbreaks during our sample period, while other countries like Argentina, Brazil, Guatemala, and Ireland did not report any birds killed due to avian influenza. We do not explore the factors that affect avian influenza outbreaks, but there are no apparent systematic patterns in our data governing which countries experience outbreaks. Further, there does not appear to be a systematic relationship in our data between avian influenza outbreaks and cage-free housing, suggesting that the US experience may be unique. Nevertheless, we include avian influenza in our estimation model as a control variable.

2.4 Animal Advocacy and Corporate Commitments

In the early 2000s, US animal advocacy groups began to work on securing pledges from consumer-facing food businesses to source cage-free eggs, rather than demand changes directly from egg producers [67]. The first comprehensive national commitments to source all egg products exclusively from cage-free hens were made in 2005 by retailers Whole Foods, Wild Oats, and Earth Fare. Other early commitments to source cage-free eggs were either made by niche retailers and restaurants or covered only part of a company's egg usage. For example, major institutional food service companies Sodexo and Aramark committed to sourcing cage-free shell eggs in 2012 and 2013, respectively, but this commitment did not include liquid eggs [70, 78]. The brand identities of these early mover retailers and the long timelines of the commitments point to the long-term nature of the process to expand cage-free egg production and suggest that companies may have considered the existing cage-free egg supply in their decision to make commitments. Retailers such as Whole Foods likely had established sources for cage-free eggs, and other companies like Bon Appétit Management Company reported that they coordinated closely with suppliers to make changes to their hen housing [86]. The degree to which existing supply impacted a company's decision likely depended on the size and market position of the company: Sodexo and Aramark are two of the three largest companies in their industry, and therefore they may have influenced their suppliers' housing decisions as opposed to waiting for their suppliers to make housing changes [71].

In response to initial success, advocacy groups invested more time in corporate dialogue and pressure campaigns. In early 2015, Sodexo and Aramark updated their commitments to cover all egg products by 2020, and their competitor Compass Group North America released a commitment to cover all egg products in their US operations by 2019. Soon after, other companies began to make commitments in quick succession. The majority of cage-free commitments were made between 2015 and 2017, the period during which most advocacy efforts were conducted [15]. Many companies who committed during this time, like McDonald's, Costco, Walmart, Kroger, Denny's, and Nestle, use a large quantity of the eggs sold in the US and thus affect a large number of hens.

The case of warehouse grocer Costco illustrates the connection between campaigns and commitments. In 2015, a coalition of advocacy groups began a pressure campaign against Costco to secure a cage-free commitment after failed attempts at dialogue. Pressure tactics included on-the-ground protests, newspaper ads, petitions, shareholder activism, and publication of undercover footage of poor welfare standards in Costco's supply chain. The corporation made a cage-free commitment in December 2015 and now provides regular updates on its progress. The most recent update indicates that 89% of shell eggs sold in its US operations were cage-free as of September 2018, which impacts roughly 9.6 million hens each year [20, 1].

As most large US companies have now committed to sourcing cage-free eggs, the influx of cage-free commitments has moved from the US to other countries globally. Certain European countries like Belgium, Germany, and France gained commitments between 2005–2014, a few years earlier than the groundswell in the US. However, other countries like Italy, Latvia, and Poland have begun to see a large number of new commitments, starting in 2016. The expanded efforts of international advocacy groups like the Open Wing Alliance likely explain these recent commitments. To date, about 800 cage-free commitments affect European operations, and about 400 commitments affect countries outside of the US and Europe [15]. These observations provide invaluable variation over time and space that can be used to better estimate the causal effect of commitments on the percentage of hens in cage-free housing.

H1a: Commitments by food companies to source cage-free eggs from their suppliers will positively impact the proportion of cage-free housing, holding other factors constant.

H1b: Advocacy group pressure campaigns will positively impact the number of cage-free commitments, holding other factors constant.

2.5 Hen Housing Legislation

Among the European Union (EU) member states, legislation adopted in 1999 establishing minimum standards for egg-laying hens effectively banned the use of battery cages beginning in 2012 [27]. Germany, Austria, the Netherlands, and Sweden passed legislation of their own, with Germany and the Netherlands establishing stricter standards than the EU legislation [41]. Since the EU-wide legislation still allows for the use of enriched cages, advocacy groups have continued to work towards cage-free commitments for European food businesses. Outside of the EU, Switzerland banned all types of cages for egg-laying hens in 1992 [51, 41].

In 2008, California adopted a ballot initiative, Proposition 2, titled "Standards for Confining Farm Animals." Proposition 2 prohibited the confinement of certain farmed animals in a manner that does not allow them to lie down, stand up, fully extend their limbs, and turn around freely [73]. While the adoption of the law represented a victory for groups that had campaigned for the law since 2005, the vague wording of the bill allowed for enriched cages, and some producers switched from battery cages to enriched cages as a means of compliance with the new law [67]. In 2010, the California legislature passed AB 1437, requiring all eggs sold in California to meet the standards outlined in Proposition 2 by January 1, 2015 [37].

Following the lessons learned from Proposition 2 and AB 1437, initiatives in other states banned sales of *all* eggs from caged hens, regardless of whether they were sourced from outside the state or producers using enriched cages. In 2016, Massachusetts voters approved such a measure with 77% support, which will take effect in 2022 [18]. In California in 2018, Proposition 12, or the "Prevention of Cruelty to Farm Animals Act," banned cages for layer hens starting in December 31, 2021 and improved on Proposition 2 to require at least 144 in^2 (929 cm^2) per hen of floor space in barns without cages by January 1, 2020 [43, 42]. In 2019, legislatures in Washington, Oregon, and Michigan adopted cage-free laws, including sales bans, which come into effect in January 2024, January 2024, and December 2024 respectively [21, 11, 74].

H2: Legislation to ban the use of any kind of cages in hen housing will positively impact the proportion of cage-free housing, holding other factors constant.

3 Data and Methods

3.1 Data

We test our main hypotheses using observational country-level data on hen housing, corporate cage-free commitments, cage-free legislation, and avian influenza. Summary statistics for the full data panel are shown in Table 1, and Figure 2 visualizes the variables of interest in our study, plotted by country. All publicly available data and the code for cleaning and analysis of this data can be found in the Open Science Framework repository at `https://osf.io/vte94/`. All subsequent references to directories and files refer to this repository. The data cleaning code in the file `/code/wrangle.R` is written for the data analysis software `R` (version 3.6.1) [60]. The publicly available data are located in the files `/data/raw/production-public.csv` and `/data/raw/commitments-campaigns.csv`. Instructions for downloading the data and code for replication of the analysis are available in the file `README.md`. We preregistered our data collection and analysis plan [47], and our methods in this study follow the preregistration closely. Some deviations due to data availability and the results of diagnostic tests are noted throughout.

Table 1: Summary statistics for main estimation sample

	N	mean	sd	min	max
prop_hens_cage_free	504	0.270	0.274	0	1
cumul_commitments	624	10.54	35.05	0	416
cumul_campaigns	624	1.918	6.669	0	89
legislation	624	0.309	0.463	0	1
birds_killed_by_flu	624	147,794	896,727	0	1.601e+07

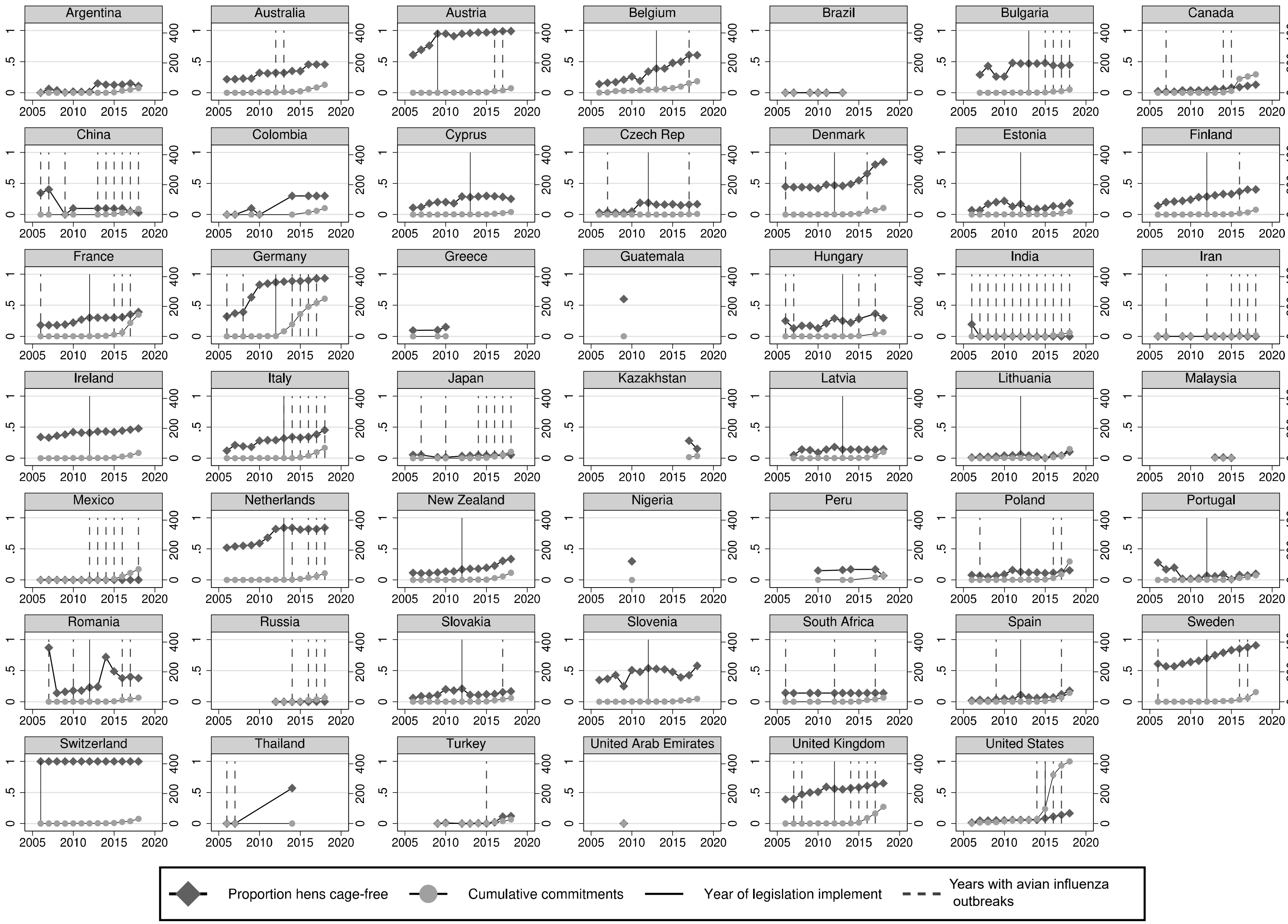


Figure 2: Trends in the data, by country and year. The connected black line denotes the dependent variable *prop_hens_cage_free*, while the connected gray line shows the explanatory variable of interest, *cumul_commitments*. The solid vertical line depicts the year in which a country's cage-free legislation was implemented, and dashed vertical lines indicate years in which at least one recorded death due to avian influenza occurred. Data for this figure are available at `https://osf.io/VTE94/`.

3.1.1 Proportion of Hens in Cage-free Housing

The percentage of hens living in cage-free housing by country and year is collected from annual publications of the International Egg Commission (IEC), an association of international egg industry professionals [39], or US and EU government reports [46, 17] and stored as a proportion in the variable *prop_hens_cage_free*. The IEC annual reports were accessed on December 1, 2019.[4] The IEC collects data from a representative in each country, who in turn collects data from producers within their country. Details of the data collection methods are not published and may be highly heterogeneous by country. For example, based on our own investigations, we know that all producers in the United States with flock sizes $\geq 30,000$ birds voluntarily report information about cage-free flock sizes to a government department (and an employee of the department acts as the representative for the US), while EU member states are required to inspect a representative sample of producers and report cage-free flock sizes to the European Commission as part of the 1999 battery cage ban. However, many countries reporting to the IEC do not provide details about how their data are collected.

For each country, we choose between IEC and government reporting of cage-free data depending on which source contains the most observations across our 13-year sample. IEC and government data are highly correlated (Pearson correlation $\rho = 0.969$). For countries with only one missing observation of cage-free housing, we search for additional data from other sources such as individual countries' egg industry association reports or poultry news organizations. All supplementary sources can be found in the `data_notes` column of the data file `/data/raw/production-public.csv`. Despite efforts to complete our panel, we were unable to obtain data on the percentage of hens housed cage-free for every country in every year. This missing data results in an unbalanced panel, which requires additional consideration in the estimation strategy to protect against information bias. To reduce the potential for bias, we examined possible patterns among demographic variables in countries either missing observations or completely missing from our sample. We do not find any evidence of relationships between missingness and any observed factors.

International trade between countries with high numbers of corporate commitments and low domestic production of eggs may affect our hypothesis that corporate commitments from food businesses with operations in a particular country increase the percentage of cage-free housing in only that country. However, nearly all the countries in our sample are largely self-sufficient in egg production [31]. Germany is the only exception, with both a low self-sufficiency rate and a high number of corporate commitments.[5] As such, we will omit Germany from our main analysis; however, the results of a

[4]Two reports were not available to the public. We privately obtained data from these reports through communication with IEC members; as such, we omit data from these two reports in our public data set. We provide the public data panel and estimation results from this article replicated using these data in the directory `/data/raw/`.

[5]See [80] for a detailed analysis of the spatial shifts in EU egg production, with a special focus on Germany, during

robustness check regression that includes Germany (see Table 9 Column 2) are qualitatively similar.

3.1.2 Corporate Commitments and Pressure Campaigns

Corporate commitment (*cumul_commitments*) and pressure campaign (*cumul_campaigns*) data are obtained from ChickenWatch.org [15], a website that tracks food business commitments to source higher welfare animal products, including cage-free eggs. The data are input by animal advocacy groups involved in securing each commitment, with a dedicated administrator reviewing the data. Each observation records details about a single publicly stated commitment, including corporation name, sector (retailer, restaurant, producer, etc.), date on which the commitment was published, deadline for compliance, countries covered, and whether the commitment was obtained through dialogue or a pressure campaign. Note that commitments may not be one-to-one with corporations or legal entities: for example, a commitment might represent only a single brand or subsidiary. While this might potentially produce some inconsistency in the unit of analysis, there is no obvious source of systematic bias. We aggregate commitments and campaigns at the country level. In each case, we apply a cumulative sum over time, counting the number of commitments or campaigns applying to a given country in the current and all previous years of our sample, which accounts for the dynamic nature of the producers' decision to invest in cage-free housing. Table 1 shows that the distribution of cumulative commitments is skewed with the mean much closer to the minimum than the maximum value, which Figure 2 shows to be driven by the high numbers of commitments in a small number of countries (the US, Canada, Germany, Poland, France, and the United Kingdom). Similarly, the US, Canada, France, and the United Kingdom drive a long right tail in the campaigns data. Campaigns follow similar country-level trends as commitments, so the data are omitted from the plots in Figure 2 for visual clarity.

In Section 4.4, we discuss transformations of the covariates that deviate from the preregistered covariates described above to provide alternative interpretations of the results. For commitments, ideally we would perform this analysis at the food business level and account for the market share of each company when estimating the effect of each commitment. However, egg market share data are generally not available for the food companies within our data. A second-best strategy might be to calculate each country's proportion of food retailers committed to sourcing cage-free eggs. This proportion would normalize the commitments variable to remove the influence of factors like national population as it pertains to company size. However, data on the total number of food retailers operating in a country are limited, given the wide number of countries and years in our panel. As an approximation for this alternate interpretation, we normalize the yearly cumulative number of

the period following the EU's ban on battery cages.

commitments by the maximum cumulative commitments for each country, yielding a proportion of total commitments.

3.1.3 Legislation

Information on legislation is collected from public news sources and communication with experts on the subject. The indicator variable *legislation* equals one in the year that a country implements legislation banning the use of any kind of cage in layer hen housing and in all years following, and zero in all previous years. No countries in our sample implement and then later overturn legislation. We choose an indicator variable rather than, for example, a composite index to simplify data collection and model specification in this analysis, where legislation is not the primary effect of interest. Similarly, we aggregate legislation banning battery cages with legislation banning all cages because very few countries have implemented total cage bans.

3.1.4 Avian influenza

Avian influenza data are collected from the World Organization for Animal Health's World Animal Health Information System dashboard [84]. The variable $birds\_killed\_by\_flu$ represents the number of domestic commercial poultry birds (including turkeys, quails, guinea fowl, and chickens) killed during an official avian influenza outbreak. Table 1 shows the wide variability in the number of deaths due to disease: in most years, most countries experience no deaths due to avian influenza, but in years with outbreaks, like the US in 2014–2015, many birds died. For visual clarity, Figure 2 simply indicates years in which at least one recorded death due to avian influenza occurred as opposed to the number of birds killed.

For an alternative interpretation of the avian influenza variable, we normalize the number of birds killed by the total population of domestic chickens. This normalization removes the influence of relative differences between countries in population and flock sizes. Data on total population of domestic chickens are collected from the Food and Agriculture Organization of the United Nations' FAOSTAT Production database [32].

3.2 Statistical Methods

3.2.1 Estimation Strategy

We are primarily interested in the causal effect of corporate cage-free commitments on the proportion of cage-free hen housing in a given country. We hypothesize that simultaneity bias, or a feedback loop, may exist between the two variables, as food retailers and their egg suppliers may make their decisions simultaneously. There may also be unobserved factors that influence how egg producers choose to house hens. To minimize bias in identifying the causal relationship between commitments and cage-free housing, we must account for the observed and unobserved variables, correlations between variables, and the potential simultaneity bias. We do so by exploiting the longitudinal nature of the data to analyze the impacts of corporate commitments at the country level over time in our estimation procedure.

The estimation equation, a model of the proportion of hens living in cage-free housing in country i and year t, is as follows:

$$\begin{aligned} prop\_hens\_cage\_free_{it} = \beta_0 &+ \beta_1\, cumul\_commitments_{it} + \beta_2\, legislation_{it} \\ &+ \beta_3\, birds\_killed\_flu_{it} + \boldsymbol{year\_t}\,\boldsymbol{\Gamma} \\ &+ \nu_i + u_{it} \end{aligned} \tag{1}$$

where

- $prop\_hens\_cage\_free_{it}$ is the proportion of all commercial egg-laying hens living in cage-free housing in country i and year t;
- $cumul\_commitments_{it}$ is the cumulative number of corporate cage-free commitments in country i and year t;
- $legislation_{it}$ is an indicator variable for whether country i has implemented legislation banning either battery cages or all cages in year t;
- $birds\_killed\_flu_{it}$ is the number of chickens killed by avian influenza outbreaks in country i and year t;
- $\boldsymbol{year\_t}$ ($t = 2006, ..., 2018$) is a vector of indicator variables equal to one in year t and zero in other years;
- $\beta_0, \beta_1, \beta_2, \beta_3$, and the elements of vector $\boldsymbol{\Gamma}$ are coefficients to be fit during the estimation;
- ν_i and u_{it} are the unobserved country and country-year effects, respectively, in country i and year t

We estimate this equation using correlated random effects (CRE) estimation methods with regressions that account for our main sources of bias [55, 83, 40]. We choose our preferred specification from among several according to information from diagnostic plots and hypothesis tests. We further compare the performance of the CRE method against the fixed effects (FE) method, a more commonly used method for estimating effects within clusters, and show that CRE is a flexible alternative to FE. CRE allows us to estimate the effects of explanatory variables within individual countries over time ("within effect") as well as the average effect of differences in time averages of explanatory variables between countries ("between effect"). The within effect, which asks "how do corporate commitments affect the proportion of cage-free housing over time in a given country?" is the most policy-relevant effect, and the between effect helps us understand how a change in a country's average level of an explanatory variable affects the value of the dependent variable. All diagnostics and estimation were performed using `STATA` analysis software version 13 [72], and the estimation code is available in the file `/code/analysis.do`.

To estimate equation 1 using CRE, we employ a Mundlak transformation to model the unobserved country effect ν_i as conditional on the observed explanatory variables [55, p. 122]. Specifically, the transformation assumes that the country-level error ν_i can be decomposed into the sum of the cluster means of the observed explanatory variables and an unobserved country-level error term μ_i with $\mu_i \sim N(0, \sigma^2_{\mu_i})$ [55, 66]. This assumption allows for the explanatory variables to be correlated with the country-level heterogeneity through the cluster means, an important feature that allows us to account for potential unobserved variable bias and to causally interpret the estimated effects of our explanatory variables. To implement the Mundlak transformation, we construct the cluster means, or averages over time within a country, of all explanatory variables and add them to equation 1 [55, 83]. We compute these time averages using only complete cases, or years in which we observe the values of all dependent and explanatory variables for a country; therefore, averaging the year indicators does *not* remove all variation. Thus, we arrive at the equation to be estimated in all CRE regressions:

$$\begin{aligned} \mathit{prop\_hens\_cage\_free}_{it} = {} & \beta_0 + \beta_1\, \mathit{cumul\_commitments}_{it} + \beta_2\, \mathit{legislation}_{it} \\ & + \beta_3\, \mathit{birds\_killed\_flu}_{it} + \boldsymbol{year\_t}\, \boldsymbol{\Gamma} \\ & + \xi_1\, \mathit{avg\_commitments}_i + \xi_2\, \mathit{avg\_years\_legislation}_i \\ & + \xi_3\, \mathit{avg\_birds\_killed\_flu}_i + \boldsymbol{avg\_year\_t_i}\, \boldsymbol{\Pi} \\ & + \mu_i + u_{it} \end{aligned} \tag{2}$$

where, in addition to the other variables described in equation 1,

- $\mathit{avg\_commitments}_i$ is the average cumulative commitments in country i across all of i's complete-case years;
- $\mathit{avg\_years\_legislation}_i$ is the years with legislation in country i across all of i's complete-case

years;

- $avg\_birds\_killed\_flu_i$ is the number of birds killed by avian influenza in country i across all of i's complete-case years;

- $\boldsymbol{avg\_year\_t_i}$ is a vector of averaged year indicators, computed for each country i across all of i's complete-case years;

- ξ_1, ξ_2, ξ_3, and the elements of vector $\mathbf{\Pi}$ are coefficients to be fit during the estimation

Here, ξ_1, ξ_2, ξ_3 are interpreted as the difference between the within effect and the between effect of each explanatory variable k. Therefore, we can recover the between effects by adding the estimated coefficient on each explanatory variable in our regression results (the within effect β_k) to the estimated coefficient on the time average of that variable (ξ_k).

Our first regression to estimate equation 2 follows a baseline nonlinear CRE method, which assumes the panel is balanced and therefore does not account for potential heteroskedasticity induced by an unbalanced panel [55]. We refer to this as the *homoskedastic nonlinear CRE regression*, and we estimate it by fitting a generalized linear model (GLM) with a probit link function to equation 2. Data are clustered at the country level. The probit link function accounts for the proportional dependent variable and constrains the estimated response to the range $[0, 1]$.

We estimate a second nonlinear CRE regression, modified from the homoskedastic nonlinear CRE regression. This allows us to account for the unbalanced nature of the panel data, which are clustered at the country level. We refer to this as the *heteroskedastic nonlinear CRE regression*. Following Wooldridge [83, p. 139] and Joshi and Wooldridge [40], we create "unbalancedness" fixed effects (FEs) to group countries which appear in the same number of years by defining indicator variables such that $T_{ir} = 1$ if country i has complete cases for $r = 2, \ldots, 13$ years of our 13-year data sample and zero otherwise. Countries with data in only one year are dropped from this analysis [40]. We estimate the regression using a heteroskedastic GLM with a probit link function that includes the "unbalancedness" FEs in both the expectation equation and the variance equation of the procedure. The heteroskedastic GLM is characterized by explicitly estimating a variance equation that depends on certain explanatory variables. By specifying the variance to depend on the "unbalancedness" FEs, this method allows us to account for heteroskedasticity that may occur due to the number of years in which a country appears in our data [79]. We include the "unbalancedness" FEs in all CRE regressions for consistency in order to compare estimates across regression. However, only the heteroskedastic nonlinear CRE regression estimates the variance equation that includes the "unbalancedness" FEs.

The coefficients from the probit function in the two nonlinear CRE regressions are not directly interpretable, so we calculate the average partial effects (APEs) of each estimate, which takes the average of the partial effects (the partial derivative of the outcome variable with respect to the explanatory variable, with other variables held fixed) of each variable evaluated at all values of the other explanatory variables [55]. This calculation provides an average effect of a change in one explanatory variable for any values of the other variables, which is equivalent to the interpretation of the linear coefficients.

Lastly, we estimate two additional regressions, a linear CRE regression and a FE regression, as robustness checks for our nonlinear regressions and to illustrate the benefits of using CRE versus the more common FE in our empirical setting. The CRE method in general is a flexible alternative to the FE method (which is commonly used to handle unobserved heterogeneity in panel data) as the CRE computes both the within and between effects, and the within estimates of the linear CRE are equivalent to the FE estimates. Linear CRE can handle the unbalanced panel and endogenous explanatory variables if needed, and the Mundlak transformation ensures that the between effect of time-invariant variables can be recovered from the regression results. For the linear CRE regression, we estimate equation 2 by fitting a GLM with a Gaussian identity link function using maximum likelihood estimation methods. Data are clustered at the country level. However, this regression is misspecified because our fractional dependent variable violates the assumption of the Gaussian identity link function that the error term is normally distributed error.

The FE regression is estimated by fitting FE ordinary least squares (OLS) to equation 1, with standard errors clustered at the country level. As with linear CRE, FE can conveniently handle endogenous explanatory variables and account for the unbalanced panel without modification, but it is also misspecified because of the fractional dependent variable. Further, the FE method does not allow us to observe the between effect, since all time-invariant variables are dropped by construction. Comparing the results of the two linear methods confirms that the linear CRE within estimates are equivalent to the FE estimates, and linear CRE provides estimates of the between effects that are unavailable from FE.

To choose a preferred specification, we perform post-estimation statistical tests to check assumptions about the distribution and the correlation structures of the data. Along with tests for endogeneity described below, visual inspection of the response residuals and the Akaike information criterion (AIC) will inform our preferred specification selection.[6] We use Pearson residuals, which standardize the response residuals of the nonlinear CRE regressions to be approximately normally distributed, to better compare diagnostics across our four regressions.[7] Additionally, we test whether the random effects

[6] We calculate AIC scores using the formula AIC = $\frac{-2\ln\hat{L}+2k}{N}$ where $\ln\hat{L}$ is the log likelihood (or pseudo log likelihood), k is the number of parameters, and N is the number of observations for each estimation procedure.

[7] Pearson residuals are calculated by dividing the response residuals $y_i - \hat{\mu_i}$ by a function of the fitted value that

coefficients are statistically significantly different from the fixed effect coefficients using the null hypothesis that $\boldsymbol{\xi}$ jointly equal zero. We test the significance of the between effect, calculated as $\beta_i + \xi_i$, using a χ^2 test of $H_0: \ \beta_i + \xi_i = 0$. Finally, we test the main hypothesis of interest with the null hypothesis $\beta_1 = 0$; that is, we test whether the within-country impact of corporate commitments on the proportion of hens living in cage-free housing is statistically significant.

3.2.2 Identification Strategy

Identification of the causal relationship between corporate commitments and the proportion of hens in cage-free housing using observational data is at high risk of statistical bias. We discuss the issues that potentially affect our study—namely endogeneity due to simultaneity and unobserved variables—and the modeling choices we make to account for those issues.

Simultaneity may occur in our empirical setting when corporate leaders make decisions to adopt cage-free commitments in the same time period as producers decide to adopt cage-free housing systems, and both look to the others' choices to inform their decisions. We test for potential simultaneity using a two-stage control function instrumental variable (IV) method with *cumul_campaigns* as our instrument for *cumul_commitments*. Should the test indicate the presence of simultaneity, we can then use the control function IV method as part of our main estimation strategy. The choice of instruments requires close scrutiny to ensure that they fit the assumptions of IVs [8, p. 79]. We argue that *cumul_campaigns* is contemporaneously exogenous to the outcome and fits the relevance and exclusion criteria necessary for strong instruments. Conceptually, we have shown in Section 1 that pressure campaigns have played a role in obtaining corporate commitments, which has been acknowledged by both animal advocates and agricultural industry figures. Since public pressure campaigns are almost exclusively not waged against egg producers, who are not well known among the public, campaigns are not directly correlated with changes in hen housing. Further, anecdotal evidence suggests that egg producers were initially quite resistant to making changes based on the demands of the advocacy groups running cage-free campaigns [29]. We support our choice of instrument with diagnostics and statistical tests, namely an F-statistic hypothesis test on the coefficients from the first-stage control equation regression to determine whether the instrument *cumul_campaigns* and other exogenous variables jointly affect the endogenous variable *cumul_commitments*. While there may be concern that this instrument faces the same bias from omitted variables discussed below, our panel data methods to handle omitted variable

depends on the link function family. In the case of the Bernoulli and binomial families used for our nonlinear CRE regressions, the Pearson residuals are calculated as $r_i = \frac{y_i - \hat{\mu_i}}{\sqrt{\hat{\mu_i}(1-\hat{\mu_i})}}$ where $\hat{\mu_i}$ is the fitted value of observation i [34, p. 365]. For linear regressions, including our linear FE regression and our linear CRE regression (estimated using the identity link function), the standardization function equals 1 so the response residuals $y_i - \hat{\mu_i}$ and Pearson residuals $\frac{y_i - \hat{\mu_i}}{1}$ are equivalent.

bias in the variables of interest also cover the instrument.

The test for potential simultaneity follows Wooldridge [82]. We use the control function IV method to determine whether the endogeneity is due to unobserved heterogeneity at the country level (and is therefore handled by the CRE method alone) or due to contemporaneous correlation with the idiosyncratic error term (and would therefore be more consistently estimated by combining the control function method with the CRE method). This method estimates a first-stage control equation by regressing *cumul_commitments* on *cumul_campaigns*, the other exogenous variables and their respective averages, then includes the residuals in the second-stage regression of equation 2. We choose the control function method over the more commonly used two-stage least squares (2SLS) IV method to test for endogeneity for two reasons. First, the control function method is more efficient than 2SLS when either the dependent variable or the endogenous explanatory variables are nonlinear (the two methods are equivalent in the linear case) [82, 59]. Second, the control function method provides a more robust test of endogeneity than the 2SLS test, which alleviates some of the general concerns around overreliance on the 2SLS tests of endogeneity [82].

We estimate the first-stage control regression using pooled OLS; then we include the residuals from the control regression in both the linear and nonlinear CRE second-stage regressions of equations 1 and 2, respectively [82]. For comparison, we perform the same procedure for the FE regression, estimating the control regression with FE OLS. The test for endogeneity is a t test of the null hypothesis that the coefficient on the residuals is zero—in other words, that *cumul_commitments* is contemporaneously exogenous [82, p. 435]. If we were to reject the null of exogeneity, we would then use a similar control function IV method throughout our analysis. However, the results of the test for endogeneity (Table 2) indicate that we cannot reject the null hypothesis that *cumul_commitments* is uncorrelated with the country-year error. These results imply that simultaneity does not bias our results, therefore we omit the control function IV method from the remainder of our analysis. We support this decision with a robustness check comparing a regression estimated using the control function IV method to one without (described in Section 4). This check shows that the results are qualitatively similar and confirms that simultaneity does not bias our results. We focus the remainder of this discussion on approaches that do not include the IVs.

Another threat to identification in our study comes from omitted variable bias, which occurs when unobserved factors that directly impact the proportion of cage-free hens are also correlated with one of the explanatory or IVs. For example, unobserved consumer preferences may influence corporations' willingness to make commitments as well as producers' willingness to adopt cage-free housing systems, and perhaps even advocates' decision of which campaigns to pursue. Similarly, the unobserved price premium between cage-free and conventional eggs may influence both producers' and corporations' decisions. We address potential omitted variable bias by using the CRE method, which takes advantage of variation in the data across time and between individual countries to account for bias by either removing time-invariant unobserved effects or allowing for correlation between countries.

Panel data methods are powerful tools to address many types of endogeneity; however, unbalanced panels (or panels in which not every individual appears in every time period) may introduce additional bias if selection into a time period is systematically related to the observed variables or to the variable value that is missing. This situation may induce heteroskedasticity, or variance that is not constant over all observations, in the unbalanced panel [83]. In our context, we may worry that countries which fail to report on proportion of cage-free housing in a given year may be related to one another in terms of population size, national income, political institutions, or other factors; and that may also influence the number of commitments made, legislation implemented, or avian influenza control methods. As discussed in Section 3.1, we do not find any evidence of relationships between missingness and any observed factors when looking for potential correlations. Additionally, our estimation methods account for issues arising from unbalanced panels. The fixed effects and linear correlated random effects estimates are consistent using unbalanced panels without modification, and we estimate one of the nonlinear CRE regressions using a heteroskedasticity-robust GLM estimator to account for potential induced heteroskedasticity in the nonlinear model [83].

3.3 Deviations From the Preregistration

The estimation procedures described above deviate from our preanalysis plan in a few ways, primarily related to the unbalanced panel. Upon compiling all of the data, we discovered that our panel was unbalanced, so we estimated a nonlinear CRE regression that accounts for unbalanced panels alongside our preregistered regressions. The control function IV test for endogeneity was added after learning more about the procedure through the unbalanced panel literature. Based on the results of this test, we have not used the IV estimation method that we preregistered for our main estimation analysis. To further support this decision, we include the results of our preferred specification (the homoskedastic nonlinear CRE regression) with and without IVs in the appendix to illustrate that the results are qualitatively similar using both methods. We planned to use `R` for all estimation, but found `STATA` provided better support for CRE estimation of unbalanced panels. We calculate a more-standard AIC instead of the preregistered quasi-AIC due to the underlying estimation methods used by `STATA` commands compared to `R` commands.

4 Results

4.1 Pre-estimation Diagnostics

Table 2 reports the results of the nonlinear CRE control function IV regressions, which provide the pre-estimation tests of instrument strength and endogeneity. Results for the linear CRE and FE regressions are qualitatively similar and can be found in Table 5. Column 1 shows the first-stage control function regression, with the dependent variable *cumul_commitments*, estimated with OLS. Column 2 shows the second-stage regression of equation 1, with the dependent variable *prop_hens_cage_free*, estimated using the nonlinear CRE. This table provides two key diagnostic results that allow us to refine the main regression procedure. First, the F test of the control regression rejects the null that the instruments jointly have no effect on *cumul_commitments*; that is, *cumul_campaigns* is a suitably strong instrument for *cumul_commitments*. Second, the lack of statistical significance of the coefficient on *residuals* in Column 2 indicates that we cannot reject the null that *cumul_commitments* is exogenous with respect to idiosyncratic error [40, 44]; that is, we conclude that *cumul_commitments* does not have a meaningful within-country unobserved effect on hen housing in our data. To further support our decision to omit IVs from the main estimation, we present the results of our preferred specification with and without IVs in Table 6. The comparison is estimated on a fully balanced subsample of our data, with countries that appear in all 13 time periods, as the full estimation analysis using the control function procedure requires a balanced panel in order to adjust the standard errors [55, 59]. As our interest is primarily the comparison of estimates across regressions, rather than inference, the balanced subsample is sufficient. The results with and without IVs are qualitatively similar, so we conclude that residual bias from endogeneity of *cumul_commitments* does not meaningfully affect our main estimation results.

4.2 Main Results

The above diagnostic tests indicate that we can proceed with the main estimation analysis using regressions that account for within and between effects, a fractional dependent variable, and an unbalanced data panel but do not account for an endogenous explanatory variable. Table 3 presents the results of our preregistered regressions (Columns 1, 3, and 4) as well as the additional regressions that account for the unbalanced panel (Column 2). The pre- and post-estimation diagnostics and residual plots discussed in the following section indicate that the homoskedastic nonlinear CRE regression (Column 1) is the preferred specification because it best accounts for our empirical setting and the structure of our data.

Columns 1 and 2 present the APEs of the nonlinear regressions, which are required for casual

Table 2: Control function IV regressions

Model	(1) OLS 1st stage	(2) NL CRE 2nd stage
Dependent variable	*cumul_commitments*	*prop_hens_cage_free*
cumul_commitments		0.0082
		(0.0086)
		-0.0086 - 0.0249
cumul_campaigns	4.5126***	
	(0.0516)	
	4.4113 - 4.6139	
residuals		-0.0074
		(0.0092)
		-0.0254 - 0.0105
birds_killed_flu	9.11e-07***	1.31e-08
	(2.98e-07)	(1.58e-08)
	3.26e-07 - 1.50e-06	-1.79e-08 - 4.41e-08
legislation	1.2380	0.0109
	(1.1071)	(0.1069)
	-0.9376 - 3.4137	-0.1986 - 0.2204
avg_commitments	0.1332***	0.0062
	(0.0213)	(0.0054)
	0.0913 - 0.1751	-0.0044 - 0.0167
avg_birds_killed_flu	-3.38e-06***	-9.63e-07**
	(9.93e-07)	(4.24e-07)
	-5.33e-06 - -1.42e-06	-1.79e-06 - -1.33e-07
avg_years_legislation	-1.9771	2.2136**
	(1.8678)	(0.9061)
	-5.6476 - 1.6934	0.4376 - 3.9896
intercept	20.9039*	-0.2965
	(11.1892)	(3.2335)
	-1.0845 - 42.8923	-6.6340 - 6.0410
Observations	488	488
F test of instrument strength (p)	553 (0)	
Year FE	YES	YES

This table reports the first and second stages of the control function instrumental variables (IV) regression, estimated with ordinary least squares (OLS) and nonlinear correlated random effects (NL CRE) respectively. The second-stage regression includes the results of the test for simultaneity: that is, statistical significance of the *residuals* coefficient. The t test of simultaneity uses the null hypothesis that the *residuals* coefficient is equal to zero. F tests of instrument strength in Column 1 indicate that *cumul_campaigns* is a suitably strong instrument for *cumul_commitments*.
Standard errors (in parentheses) are clustered by country. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are below the standard errors.

interpretation and for comparison to the estimates from the linear methods. (The nonlinear coefficients are reported in Table 7.) Column 1 presents the homoskedastic nonlinear CRE regression. The within effect of commitments suggests an average increase of 0.00035 in the proportion of cage-free

Table 3: Main regression results

Model	(1) NL CRE APEs Preferred specification	(2) NL CRE Het APEs Accounts for Heteroskedasticity[a]	(3) L CRE	(4) FE
Dependent variable	*prop_hens_cage_free*	*prop_hens_cage_free*	*prop_hens_cage_free*	*prop_hens_cage_free*
cumul_commitments	0.00035***	0.00028***	0.00014	0.00014
	(0.0001)	(0.0001)	(0.0002)	(0.0001)
	0.0001 - 0.0006	0.0001 - 0.0005	-0.0002 - 0.0005	-0.0001 - 0.0004
birds_killed_flu	$4.56\text{x}10^{-09}$	$2.24\text{x}10^{-09}$	$-1.14\text{x}10^{-10}$	$-1.14\text{x}10^{-10}$
	$(4.17\text{x}10^{-09})$	$(1.09\text{x}10^{-08})$	$(2.30\text{x}10^{-09})$	$(3.72\text{x}10^{-09})$
	$-3.62\text{x}10^{-09}$ - $1.27\text{x}10^{-08}$	$-1.90\text{x}10^{-08}$ - $2.35\text{x}10^{-08}$	$-4.61\text{x}10^{-09}$ - $4.38\text{x}10^{-09}$	$-7.43\text{x}10^{-09}$ - $7.20\text{x}10^{-09}$
legislation	0.00041	0.00142	0.05240**	0.05240***
	(0.0294)	(0.0195)	(0.0243)	(0.0139)
	-0.0572 - 0.0580	-0.0368 - 0.0396	0.0048 - 0.1000	0.0251 - 0.0797
avg_commitments	0.00179	0.00220	-0.00075	
	(0.0015)	(0.0017)	(0.0007)	
	-0.0012 - 0.0048	-0.0011 - 0.0055	-0.0021 - 0.0006	
avg_birds_killed_flu	$-2.88\text{x}10^{-07}$**	$-3.11\text{x}10^{-07}$**	$-2.99\text{x}10^{-08}$	
	$(1.19\text{x}10^{-07})$	$(1.45\text{x}10^{-07})$	$(6.22\text{x}10^{-08})$	
	$-5.20\text{x}10^{-07}$ - $-5.52\text{x}10^{-08}$	$-5.96\text{x}10^{-07}$ - $-2.68\text{x}10^{-08}$	$-1.52\text{x}10^{-07}$ - $9.21\text{x}10^{-08}$	
avg_years_legislation	0.60837***	0.57875***	0.58310***	
	(0.2262)	(0.2203)	(0.2055)	
	0.1650 - 1.0517	0.1469 - 1.0106	0.1804 - 0.9858	
intercept			0.74455	0.19305***
			(1.0423)	(0.0127)
			-1.2983 - 2.7874	0.1681 - 0.2180
Observations	488	432	488	488
Year FEs	Y	Y	Y	Y
Year means	Y	Y	Y	N
Estimated parameters	28	22	22	16
Variance equation	N	Y	N	N
Number of "unbalanced" FEs	10	2	10	0
Number of countries	44	34	44	44
Pseudolikelihood	-177.2503	-220.7618	90.6427	601.8038
AIC	0.841	1.124	-0.281	-2.401
χ^2 test of CRE model (p)	158463 (0)	161027 (0)	14927 (0)	
χ^2 test, between effect of *cumul_commitments* (p)	2.016 (0.156)	2.287 (0.130)	0.562 (0.454)	
χ^2 test, between effect of *legislation* (p)	7.179 (0.00738)	6.837 (0.00893)	8.738 (0.00312)	

This table reports the four main regressions: the nonlinear correlated random effects (NL CRE); the NL CRE accounting for heteroskedasticity (Het); the linear correlated random effects (L CRE); and the fixed effects (FE). The average partial effects (APEs) of nonlinear models are reported here for interpretation, while the coefficients are reported in Table 7. Akaike information criterion (AIC) = $\frac{-2\ln\hat{L}+2k}{N}$. The Durbin-Wu-Hausman χ^2 statistics to test the CRE estimators uses the null hypothesis that the CRE estimators are not jointly significantly different from the FE estimators. The χ^2 test of heteroskedasticity uses the null hypothesis that the variables included in the variance during the estimation procedure are not jointly significant. The χ^2 test of the between effect uses the null hypothesis that the between effect $(\beta_i + \xi_i)$ equals zero, or $H_0 : \beta_i = -\xi_i$.
Standard errors (in parentheses) are clustered by country and are obtained by the delta method for the APEs. *** p <0.01, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are shown below standard errors.
[a] Column 2 estimates the heteroskedastic nonlinear CRE regression, which includes a variance equation that depends on the "unbalancedness" FEs, on the subsample of countries that appear in at least 10 of the 13 years in our sample to ensure convergence. Because of the restricted sample, this regression has fewer "unbalancedness" FEs than the other CRE regressions.

housing from a one-commitment increase, or a 0.035 percentage point increase. The between effect, 0.213 (= 0.035 + 0.178), shows an average increase of 0.21 percentage points associated with a one-commitment increase in the country's commitments; however, this effect is not statistically significant. We do not find a meaningful effect on cage-free housing from the number of birds killed by avian influenza. The average effect of a country implementing legislation is a 0.04 percentage point increase in cage-free housing above the level of cage-free housing in the absence of legislation. The between effect indicates the difference between a country with no legislation at any time during the sample period versus a country with legislation throughout the entire period: a 60.88 (= 60.84 + 0.04) percentage point increase in cage-free housing. The large magnitude of the effect reflects the magnitude of the intervention; however, variation in *avg_years_legislation* is limited and potentially driven by a handful of countries, warranting caution. Most countries in the EU implemented legislation in the same year, and only a few pieces of legislation have been implemented outside of these countries.

The estimated APEs of the heteroskedastic nonlinear CRE regression, which accounts for the unbalanced panel, are shown in Column 2. The heteroskedastic probit estimation is computationally difficult when the number of independent variables included in the variance equation is large; to ensure convergence, we restrict the estimation sample such that the "unbalancedness" indicators in the variance equation are minimized while the number of countries and number of observations in the sample are maximized. This optimal subsample occurs when we keep countries that appear in at least 11 out of the 13 time periods (34 countries, 432 observations, 3 complete-case indicator variables in the variance) so that the number of "unbalancedness" FEs is 2, down from 10 in the unrestricted sample. We could imagine that this method of creating the subsample may introduce bias if there is a relationship between a country reporting in almost all periods in our sample and the country's prevalence of cage-free housing. However, our data do not display systematic evidence of such bias, as shown by comparing the summary statistics of the full sample versus the restricted (Table 8) and by visually examining trends in the explanatory variables among the included countries (Figure 3). For robustness, we estimate the homoskedastic nonlinear CRE regression on the same restricted subsample to compare the APEs. The subsample results (Table 9, Column 1) are similar in statistical significance and sign to homoskedastic nonlinear CRE results, apart from the effect of *legislation* (which is not statistically significant). We conclude that the differences in results are driven by the estimation procedures and not the restricted sample. The results show an average 0.028 percentage point increase in cage-free housing from one additional commitment. The between effect is similar to the results from the preferred specification, with a 0.25 ($= 0.028 + 0.22$) average percentage point increase in cage-free housing from an additional commitment. The within effect of implementing legislation (0.14 percentage points) is larger in magnitude than the preferred specification, although this result is also imprecisely estimated. The between effect of legislation in this regression is similarly large and imprecisely estimated as the preferred specification: comparing a country with no legislation at any time during the sample period to a country with legislation throughout the entire period shows a 58.14 ($= 0.14 + 58$) percentage point increase in cage-free housing.

Columns 3 and 4 present the results of the two linear regressions, the linear CRE and FE regressions respectively, both of which can accommodate unbalanced panels without modification [40]. We note that the estimated coefficients are the same but the standard errors are different, as expected [40].

The within effect of the linear CRE regression in Column 3 is a 0.014 percentage point increase in the fraction of cage-free housing; this effect is equivalent to the FE result. The between effect is -0.06 (= 0.014 + $[-0.074]$), indicating that a one-commitment increase in average commitments is associated with a 0.06 percentage point decrease in cage-free housing. Notably, the within effect of commitments is not statistically significant in either linear regression. On the other hand, the within effect of legislation shows that implementing cage-free legislation leads to a statistically significant 5.2 percentage point increase in cage-free housing. The between effect indicates that a country with legislation throughout the entire period experiences an average of 63.55 percentage point difference in cage-free housing relative to a country with no legislation at any time during the sample period.

4.3 Post-estimation Diagnostics

We examine the hypothesis tests and AIC scores reported in Table 3 as well as residual plots (Figure 4-6) to compare the fit of our four main regressions. The residual plots, tests of CRE parameters, and conceptual arguments for using a nonlinear model for the proportional dependent variable point us away from linear models. The test of heteroskedasticity indicates the presence of bias induced by the unbalanced panel; but the lower AIC score, fewer estimated parameters, and similar residual plots of the heteroskedastic regression compared to the homoskedastic indicate that accounting for heteroskedasticity does not provide enough improvement in fit to justify the reduced sample size. These diagnostics collectively inform our choice of the homoskedastic nonlinear CRE regression (Column 1) as our preferred specification.

The test of the CRE parameters (jointly χ^2 testing $H_0 : [\xi_1, \xi_2, \xi_3] = \mathbf{0}$) indicates that CRE methods (Columns 1, 2, and 3) are preferred over FE (Column 4) [83]. The test for heteroskedasticity (the χ^2 test of the null hypothesis that the unbalanced panel does not produce heteroskedasticity, or that the variables included in the variance are jointly equal to zero) suggests that we should reject the null of homoskedasticity.

The AIC scores of the FE and linear CRE regressions are the lowest and second lowest respectively, with the low scores reflecting the fact that AIC scores favor regressions with fewer coefficients as well as regressions with greater-magnitude likelihood scores. The FE method excludes the year and covariate means and "unbalancedness" fixed effects by construction, while the linear CRE regression omits some of these variables due to collinearity. The homoskedastic nonlinear CRE regression (Column 1) has a greater AIC than either linear regression but estimates the greatest number of parameters. Further, the homoskedastic nonlinear CRE has a lower AIC than the heteroskedastic nonlinear CRE regression (Column 2), implying that the nonlinear CRE heteroskedastic regression provides a relatively worse fit for our data despite estimating fewer parameters and accounting for the unbalanced panel.

Residual plots provide further evidence in favor of the nonlinear CRE method. Plots of the residuals versus fitted values (Figure 4) suggest that there are some unaccounted-for patterns in the data, but the nonlinear CRE regressions both account for these patterns better than the linear regressions. The residuals plotted versus time (Figure 5) display the desired homoskedasticity, with the nonlinear CRE regressions (Figures 5a and 5b) looking more homoskedastic than the FE regression in (Figure 5d). On the other hand, the residuals plotted against commitments (Figure 6) suggest that there may be some heteroskedasticity in another of the explanatory variables that has not been accounted for: the plots show funneling or a change in variance of the residuals as the number of commitments reach the higher end of our data. The source of this heteroskedasticity is not clear, although the small number of countries with large numbers of commitments (US, Canada, Germany, Poland, France, and the United Kingdom) may be one cause of estimation issues.

Taken collectively, these diagnostics point away from the linear methods and toward the homoskedastic nonlinear CRE regression that can handle the most important sources of bias without a major loss of information from the reduced sample required by the heteroskedastic nonlinear CRE's estimation procedure. Thus, the homoskedastic nonlinear CRE is our preferred specification.

4.4 Covariate Transformations

The results of models estimated using our preregistered variables illustrate the impact of a single additional commitment, implementing legislation, or a single bird killed due to avian influenza on the proportion of cage-free housing. For additional perspective and exploratory analysis beyond our confirmatory preanalysis plan, we transform two of the key explanatory variables as described in Section 3.1 and re-estimate the homoskedastic nonlinear CRE regression using these transformed variables. Table 4 compares the results of the homoskedastic nonlinear CRE regression with *cumul_commitments* in Column 1 (reprinted from Table 3, Column 1) against a regression using the proportion of total commitments (*prop_cumul_commitments*) in Column 2 and a regression using the proportion of the national flock killed by avian influenza (*prop_birds_killed_flu*) in Column 3.

The results for *prop_cumul_commitments* indicate that a 1 percentage point increase in the proportion of commitments leads to a 0.12 percentage point increase in the proportion of cage-free housing, on average. We interpret this result carefully, noting that the proportion variable is constructed as the fraction of 2018 commitments as opposed to all possible commitments, which would be the total number of food retail businesses in each country. Therefore, this result cannot predict the impact of an increase in the proportion of commitments after 2018. The results for *prop_birds_killed_flu* did not

Table 4: Regressions with covariate transformations

Model	(1) NL CRE Preferred specification	(2) NL CRE Using proportional commitments	(3) NL CRE Using proportion of flock killed by avian influenza
Dependent variable	*prop_hens_cage_free*	*prop_hens_cage_free*	*prop_hens_cage_free*
cumul_commitments	0.00035***		0.00027***
	(0.0001)		(0.0001)
	0.0001 - 0.0006		0.0001 - 0.0004
birds_killed_flu	$4.56\text{x}10^{-09}$	$2.95\text{x}10^{-09}$	
	$(4.17\text{x}10^{-09})$	$(3.77\text{x}10^{-09})$	
	$-3.62\text{x}10^{-09}$ - $1.27\text{x}10^{-08}$	$-4.44\text{x}10^{-09}$ - $1.03\text{x}10^{-08}$	
legislation	0.00041	0.01720	-0.00344
	(0.0294)	(0.0277)	(0.0305)
	-0.0572 - 0.0580	-0.0372 - 0.0716	-0.0632 - 0.0564
avg_commitments	0.00179		-0.00081
	(0.0015)		(0.0010)
	-0.0012 - 0.0048		-0.0029 - 0.0012
avg_birds_killed_flu	$-2.88\text{x}10^{-07}$**	$-1.75\text{x}10^{-07}$**	
	$(1.19\text{x}10^{-07})$	$(8.05\text{x}10^{-08})$	
	$-5.20\text{x}10^{-07}$ - $-5.52\text{x}10^{-08}$	$-3.33\text{x}10^{-07}$ - $-1.77\text{x}10^{-08}$	
avg_years_legislation	0.60837***	0.64127***	0.61444***
	(0.2262)	(0.2139)	(0.2200)
	0.1650 - 1.0517	0.2221 - 1.0604	0.1832 - 1.0456
prop_commitments		0.12795***	
		(0.0404)	
		0.0488 - 0.2071	
avg_prop_commitments		1.07114	
		(0.7847)	
		-0.4669 - 2.6092	
prop_birds_killed_flu			-0.43467
			(0.9708)
			-2.3374 - 1.4680
avg_prop_birds_killed_flu			28.00082
			(39.2564)
			-48.9402 - 104.9419
Observations	488	479	404
Year FEs	Y	Y	Y
Year means	Y	Y	Y
Estimated parameters	28	22	24
Number of "unbalanced" FEs	10	10	10
Number of countries	44	42	37
Pseudolikelihood	-177.2503	-173.2638	-139.7234
AIC	0.841	0.815	0.811
χ^2 test CRE model (p)	158463 (0)	180278 (0)	478352 (0)
χ^2 test, between effect of *cumul_commitments* (p)	2.016 (0.156)	2.413 (0.120)	0.254 (0.614)
χ^2 test, between effect of *legislation* (p)	7.179 (0.00738)	9.369 (0.00221)	7.849 (0.00508)

This table reports the average partial effects (APEs) of the homoskedastic nonlinear correlated random effects (NL CRE) regression, using alternate explanatory variables. Column 1 presents the preferred specification as reported in Table 3, Column 1; Column 2 reports the results using the fraction of 2018 commitments; and Column 3 reports the results using the fraction of a country's flock killed by highly pathogenic avian influenza. Akaike information criterion (AIC) $= \frac{-2\ln \hat{L}+2k}{N}$. The Durbin-Wu-Hausman χ^2 statistics to test the CRE estimators uses the null hypothesis that the CRE estimators are not jointly significantly different from the FE estimators. The χ^2 test of heteroskedasticity uses the null hypothesis that the variables included in the variance during the estimation procedure are not jointly significant. The χ^2 test of the between effect uses the null hypothesis that the between effect $(\beta_i + \xi_i)$ equals zero, or $H_0 : \beta_i = -\xi_i$. Standard errors (in parentheses) are clustered by country and are obtained by the delta method for the APEs. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are shown below standard errors.

reach statistical significance, which may be due to the fact that apart from the US experience with avian influenza in 2014–2015, few countries in our sample experienced significant outbreaks of the disease. Thus the vast majority of the proportional outbreak observations are at or near 0. We note that the AIC scores of these regressions are lower than our preferred specification, but the sample size and number of countries retained in these regressions are smaller. These alternative interpretations may provide some additional insight into our main results, but ultimately illustrate that our original interpretations of the explanatory variables are the most useful and the best fit for the largest sample of our data.

5 Discussion and Conclusions

Our study tests the impact of corporate commitments to source cage-free eggs on the proportion of hens in cage-free housing at a country level. The answer to this question may help activists allocate their scarce resources and help corporations determine the expected impacts of their corporate social responsibility initiatives. We find that one additional cage-free commitment has a small positive average effect (0.035 percentage point increase) on the proportion of hens living in cage-free housing. In the context of the number of commitments made each year and the number of hens in the industry, the impact of commitments is substantial. For example, 16 food businesses made new commitments affecting their Canadian operations in 2017, which according to our estimates may represent a 0.56 $(= 16 * 0.035)$ percentage point increase in the country's cage-free housing. To visualize the impact on egg-laying hens, such a percentage point increase in cage-free housing might affect nearly 147,000 out of the 26,189,873 total hens in Canada in 2017 $(146,663 = 0.0056 * 26,189,873)$. We also find that implementing cage-free legislation has a similar effect (0.04 percentage points) on a country's cage-free housing, although this result is not statistically significant. Avian influenza does not appear to have a meaningful effect on housing. These results confirm our preregistered main hypotheses and provide more insight into the relative effect sizes of each factor.

This study has some limitations. Our correlated random effects methodology attempts to address endogeneity bias, namely simultaneity and unobserved variable bias, but residual bias may remain due to limited, non-experimental data. Data limitations may affect our study in the following ways. First, our estimation does not include the effects of prices because most countries in our sample do not provide public data on price premiums. Large price premiums on cage-free eggs may encourage egg producers to convert to cage-free housing, and by not measuring these premiums, the effects we estimate may be biased upward by the effect of premiums on encouraging cage-free housing. On other hand, the bias in the results may not be very pronounced due to the minimal short-run variation in price premiums. In the US, price premiums on cage-free eggs have historically been fairly stable apart from the years affected by avian influenza [14, 57, 45]. Second, our study estimates the effect of increasing the *number* of commitments without regard to the market share of the food businesses that are making the commitments. Little public data are available on the number of eggs used by each food retailer, so we cannot calculate company-level market shares. For this reason, the effect of an additional commitment may be biased toward zero, as our commitments variable may mask the heterogeneous effect on housing of certain companies with very large market shares. Finally, our data sample ends in the year 2018, which may be too soon to see some of the effects of commitments on housing. Because egg suppliers may need to build cage-free housing before they can fulfill their customers' demand for cage-free eggs, food businesses generally include a later deadline by which to meet their cage-free pledge. Many of the commitments made in the mid-2010s stated deadlines of 2020, 2022, or 2025, for example. While our data do show a clear upturn in cage-free production, the full impacts of the commitments may not be known until later.

Information bias, which occurs where intervention status or outcome data are biased, may impact our study through measurement error in the process of collecting data on commitments. Animal advocacy groups who submit commitment data to ChickenWatch.org are not incentivized to omit or falsify commitments, but duplication of commitments may be a concern. We addressed this concern during the data cleaning process by removing explicit duplicates. As noted in Section 3.1, because commitments can represent a single subsidiary or brand of a larger corporation, there is also likely some heterogeneity in the unit of analysis tracked by commitments. However, we find no reason to think this minor heterogeneity might systematically bias our results. We do not have detailed information on the industry's collection methods of the *prop_hens_cage_free* variable, but we verify industry-collected data against government statistics. We find that industry and government measures of housing are highly correlated ($\rho = 0.969$). Since government statistics usually rely on self-reports from producers, fraud is possible. However, reports of fraud in cage-free housing are not widespread: our own examination of the US government cage-free data found that reporting is voluntary and anonymous, and therefore producers have little incentive to report fraudulent data [46].

Selection bias results from either the systematic exclusion of certain subjects from the study due to factors related to both the explanatory and dependent variables. Our use of egg industry data for the dependent variable mitigates this concern, due to the consistency of reporting. In particular, annual reports have tracked hen housing since 2007, before most of the corporate commitments occurred, and few countries have been added or removed since. The reporting countries account for the vast majority of cage-free commitments in the ChickenWatch.org data set and over half of countries with any commitments. The egg industry reports also include countries in which no cage-free commitments have been made. Nonetheless, there remains a risk of bias from the industry's initial selection of countries, although we believe this risk to be small.

Future research could extend our current analysis to estimate the relative impact of details of legislation—like wholesale bans on the sales of non-cage-free eggs and bans on enriched or battery cages—by constructing a more detailed data set of legislative features. Additionally, the effects of legislation might vary with time, so future research using analysis methods that account for differences in legislation implementation timing across countries might provide insight into the long-term effects of legislation. Our analysis might also be extended to consider questions of enforcement efficiency as more compliance and enforcement mechanism data become available.

Further research might move beyond our current static model to a dynamic CRE model if more data become available [81]. A dynamic model of the relationship between cage-free commitments and hen housing might provide a richer understanding of egg producers' decisions. Our cumulative definition of key variables preserves the long-term effects of commitments on hen housing; however, the short-term effects of commitments on housing may interest egg producers and legislators. Finally, spatial-dynamic models might also account for shifts in egg production and trade, which may enable the analysis of cross-border impacts of commitments in countries with very high or very low self-sufficiency rates, like Germany.

RETHINK PRIORITIES

This research is a project of Rethink Priorities. It was written by Samara Mendez and Jacob Peacock, and it was conceptualized and planned when both authors were affiliated with The Humane League Labs (see Section 6.3). If you like our work, please consider subscribing to our newsletter. You can explore our completed public work here.

6 Declarations

6.1 Author Information

Samara Mendez✉ samara@rethinkpriorities.org, https://orcid.org/0000-0002-8179-744X.

Jacob R Peacock, jacob@rethinkpriorities.org, https://orcid.org/0000-0002-4834-8132

6.2 Acknowledgements

This draft benefited from extensive consultation with Matthew Butner; research feedback by Emiliano Huet-Vaughn, Romain Espinosa, Zach Freitas-Groff, Aakanksha Melkani, and participants at the NAREA Scholars' Circle and RECAP research groups; and helpful discussions with Josh Balk, Chris

Holbein, Julie Janovsky, Janosch Linkersdörfer, Paul Petersan, Giovana Vieira and Wendy Watts. All remaining errors are our own.

6.3 Conflicts of Interest and Funding

Samara Mendez and Jacob Peacock were affiliated with The Humane League Labs (THLL) when this research was conceptualized and planned. THLL performs scientific research to inform animal advocacy strategy. THLL is a program of, and currently fully funded through, The Humane League (THL), a 501(c)(3) nonprofit organization that "exists to end the abuse of animals raised for food." THLL is editorially independent from THL, and any other potential funders, in reporting research results. The design, execution, analysis, interpretation, and reporting of THLL research is performed entirely by THLL staff, without oversight by other THL staff or leadership. To further mitigate potential conflicts of interest, THLL demonstrates commitment to transparency by adhering to open science practices, including public preregistration of studies and analysis plans as well as publication of supporting data, computer code, and materials for all THLL research.

Neither The Humane League Labs nor its parent organization, The Humane League (`https://thehumaneleague.org/`), have reviewed or endorsed the final version of this paper.

6.4 Funding

This paper was initially funded as a project of The Humane League Labs. Neither The Humane League Labs nor its parent organization, The Humane League (`https://thehumaneleague.org/`), have reviewed or endorsed the final version of this paper.

6.5 Open Science

All data, materials, and code for this project are available at `https://osf.io/VTE94/`.

The analysis plan for this project was preregistered in accordance with open science principles. Before preregistration, some data for this project were obtained and data cleaning for data on corporate commitments and campaigns commenced. However, besides the US data on cage-free housing and corporate commitments (Figure 1), no data for this project were substantively analyzed prior to registration. During a public presentation, one author incidentally saw a graph of increasing cage-free egg production in a Scandinavian country that, to their recollection, supported the main hypothesis of this study. During searches to identify data availability, we examined the cage-free housing and corporate commitment data to assess which countries and years were available in the data. Examining this incomplete data informed plans for data cleaning and searches for further data, but did not substantively test or inform any of our preregistered hypotheses.

7 References

[1] Agricultural Analytics Division. *Cage-Free Statistics from USDA*. Tech. rep. United States Department of Agriculture, June 2018.

[2] Agricultural Marketing Service. *Shell Egg Grading Service FAQs*. Oct. 2015. URL: https://www.ams.usda.gov/publications/qa-shell-eggs (visited on 09/27/2018).

[3] Pablo Alarcon et al. "Comparison of 2016–17 and Previous Epizootics of Highly Pathogenic Avian Influenza H5 Guangdong Lineage in Europe". In: *Emerging Infectious Diseases Journal* 24.12 (Dec. 2018). DOI: 10.3201/eid2412.171860. URL: https://wwwnc.cdc.gov/eid/article/24/12/17-1860_article (visited on 09/01/2019).

[4] American Humane Farm Program. *Our Standards*. Tech. rep. American Humane Certified, June 2017, pp. 1–142. URL: http://humaneheartland.org/index.php?option=com_content&view=article&id=3&Itemid=106&jsmallfib=1&dir=JSROOT/Animal+Welfare+Full+Standards+%2B+Supplements (visited on 06/20/2019).

[5] American Veterinary Medicine Association. *A Comparison of Cage and Non-Cage Systems for Housing Laying Hens*. Undated. URL: https://www.avma.org/KB/Resources/Reference/AnimalWelfare/Pages/AVMA-issues-A-Comparison-of-Cage-and-Non-Cage-Systems-for-Housing-Laying-Hens.aspx (visited on 02/19/2019).

[6] American Veterinary Medicine Association. *Literature Review on the Welfare Implications of Induced Molting of Layer Chickens*. Tech. rep. American Veterinary Medical Association, Feb. 2010. URL: https://www.avma.org/KB/Resources/LiteratureReviews/Pages/Welfare-Implications-of-Induced-Molting-of-Layer-Chickens.aspx (visited on 08/30/2019).

[7] Christa M. Anderson, Gregory P. Asner, and Eric F. Lambin. "Lack of Association between Deforestation and Either Sustainability Commitments or Fines in Private Concessions in the Peruvian Amazon". In: *Forest Policy and Economics* 104 (July 2019), pp. 1–8. ISSN: 1389-9341. DOI: 10.1016/j.forpol.2019.03.010. URL: http://www.sciencedirect.com/science/article/pii/S1389934118303629 (visited on 09/13/2019).

[8] Joshua D Angrist and Alan B Krueger. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments". In: *Journal of Economic Perspectives* 15.4 (2001), pp. 69–85.

[9] Animal Welfare Division. *Welfare Implications of Laying Hen Housing–Literature Review*. Tech. rep. American Veterinary Medicine Association, Jan. 2012. URL: https://www.avma.org/resources-tools/literature-reviews/welfare-implications-laying-hen-housing (visited on 01/28/2020).

[10] Animal Welfare Institute. *Consumer Perceptions of Farm Animal Welfare*. Tech. rep. Animal Welfare Institute, May 2015, pp. 1–14. URL: https://awionline.org/sites/default/files/uploads/documents/fa-consumer_perceptionsoffarmwelfare_-112511.pdf (visited on 12/12/2018).

[11] Brian Blake. *Commercial Egg Layer Operations–Various Provisions*. July 2019. URL: http://lawfilesext.leg.wa.gov/biennium/2019-20/Pdf/Bills/Session%20Laws/House/2049-S.SL.pdf?cite=2019%20c%20276%20%C2%A7%207; (visited on 01/07/2020).

[12] Lewis Bollard. *Research Note: What Can We Learn From the Fur-Free Fight?* Aug. 2019. URL: https://us14.campaign-archive.com/?u=66df320da8400b581cbc1b539&id=edb84be486 (visited on 12/19/2019).

[13] Marinella Capriati. *Cause Area Report: Corporate Campaigns for Animal Welfare*. Tech. rep. Founders Pledge, Nov. 2018. URL: https://founderspledge.com/research/fp-animal-welfare (visited on 07/31/2019).

[14] Jae Bong Chang, Jayson L. Lusk, and F. Bailey Norwood. "The Price of Happy Hens: A Hedonic Analysis of Retail Egg Prices". In: *Journal of Agricultural and Resource Economics* 35.3 (2010), pp. 406–423. ISSN: 1068-5502. URL: https://www.jstor.org/stable/23243063 (visited on 08/28/2019).

[15] ChickenWatch. *ChickenWatch*. 2022. URL: https://www.ChickenWatch.org (visited on 01/26/2019).

[16] Phillip Clauer. *Commerical Poultry Industry*. Dec. 2011. URL: https://extension.psu.edu/commerical-poultry-industry (visited on 08/27/2019).

[17] European Commission. *EU Market Situation for Eggs: Forecast Working Group on Poultry & Eggs*. Tech. rep. European Commission, June 2019, pp. 1–22. URL: https://ec.europa.eu/info/food-farming-fisheries/animals-and-animal-products/animal-products/eggs_en.

[18] Abby Elizabeth Conway. *Mass. Voters Approve Question 3, Banning Certain Farm Animal Confinement Practices*. Nov. 2016. URL: https://www.wbur.org/politicker/2016/11/08/question-three-animal-confinement-results (visited on 08/27/2019).

[19] Amelia Cornish, David Raubenheimer, and Paul McGreevy. "What We Know about the Public's Level of Concern for Farm Animal Welfare in Food Production in Developed Countries". In: *Animals* 6.11 (Nov. 2016). ISSN: 2076-2615. DOI: 10.3390/ani6110074. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5126776/ (visited on 08/21/2019).

[20] Costco. *Animal Welfare*. Dec. 2018. URL: https://www.costco.com/sustainability-animal-welfare.html (visited on 08/23/2019).

[21] Kevin Daley. *Animal Industry Act*. Nov. 2019. URL: https://www.legislature.mi.gov/documents/2019-2020/publicact/pdf/2019-PA-0132.pdf (visited on 01/07/2020).

[22] Michael R. Darby and Edi Karni. "Free Competition and the Optimal Amount of Fraud". In: *The Journal of Law & Economics* 16.1 (1973), pp. 67–88. ISSN: 0022-2186. URL: https://www.jstor.org/stable/724826 (visited on 03/15/2022).

[23] Marian Stamp Dawkins. "Behavioural Deprivation: A Central Problem in Animal Welfare". In: *Applied Animal Behaviour Science* 20.3 (Aug. 1988), pp. 209–225. ISSN: 0168-1591. DOI: 10.1016/0168-1591(88)90047-0. URL: https://www.sciencedirect.com/science/article/pii/0168159188900470 (visited on 04/29/2022).

[24] Marian Stamp Dawkins and Sylvia Hardie. "Space Needs of Laying Hens". In: *British Poultry Science* 30.2 (June 1989), pp. 413–416. ISSN: 0007-1668. DOI: 10.1080/00071668908417163. URL: https://doi.org/10.1080/00071668908417163 (visited on 04/29/2022).

[25] Rich Duprey. *Cal-Maine Foods Scrambled by Low Egg Prices*. July 2017. URL: https://www.fool.com/investing/2017/07/20/cal-maine-foods-scrambled-by-low-egg-prices.aspx (visited on 08/26/2019).

[26] Niklas Egels-Zandén and Henrik Lindholm. "Do Codes of Conduct Improve Worker Rights in Supply Chains? A Study of Fair Wear Foundation". In: *Journal of Cleaner Production* 107 (Nov. 2015), pp. 31–40. ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2014.08.096. URL: http://www.sciencedirect.com/science/article/pii/S0959652614009196 (visited on 09/12/2019).

[27] European Commission. *Council Directive 1999/74/EC of 19 July 1999 Laying down Minimum Standards for the Protection of Laying Hens*. Aug. 1999. URL: http://data.europa.eu/eli/dir/1999/74/oj/eng (visited on 09/01/2019).

[28] European Commission. *Laying Hens*. Text. July 1999. URL: https://ec.europa.eu/food/animals/welfare/practice/farm/laying_hens_en (visited on 01/28/2020).

[29] Scott Fetridge. *Cage-Free Chicken Campaign Gains Unexpected Traction.* Feb. 2022. URL: https://www.foodmanufacturing.com/supply-chain/news/22044097/cagefree-chicken-campaign-gains-unexpected-traction (visited on 06/24/2022).

[30] Food and Agriculture Organization of the United Nations. *Chapter 1 - Egg Production.* Services Bulletin 150. Food and Agriculture Organization of the United Nations, 2003. URL: http://www.fao.org/3/Y4628E/y4628e03.htm (visited on 08/23/2019).

[31] Food and Agriculture Organization of the United Nations. *FAO Statistical Pocketbook 2012: World Food and Agriculture.* Rome: Food and Agriculture Organization of the United Nations, 2012. ISBN: 978-92-5-107083-3.

[32] Food and Agriculture Organization of the United Nations. *FAOSTAT Database.* 2022. URL: https://www.fao.org/faostat/en/#data (visited on 04/19/2022).

[33] Chad Gregory. *Cage-Free Transition Creates Uncertainty.* Dec. 2016. URL: http://www.poultrytimes.com/poultry_today/article_27f6af08-cc75-11e6-8cda-6f95559fd8e3.html (visited on 01/28/2020).

[34] James W. Hardin et al. *Generalized Linear Models and Extensions, Second Edition.* Stata Press, 2007. ISBN: 978-1-59718-014-6.

[35] Robert Heilmayr and Eric F. Lambin. "Impacts of Nonstate, Market-Driven Governance on Chilean Forests". In: *Proceedings of the National Academy of Sciences* 113.11 (Mar. 2016), pp. 2910–2915. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1600394113. URL: https://www.pnas.org/content/113/11/2910 (visited on 09/12/2019).

[36] Heinke Heise and Ludwig Theuvsen. "What Do Consumers Think about Farm Animal Welfare in Modern Agriculture? Attitudes and Shopping Behaviour". In: *International Food and Agribusiness Management Review* 20.3 (Mar. 2017), pp. 379–399. DOI: 10.22434/IFAMR2016.0115. URL: https://www.wageningenacademic.com/doi/abs/10.22434/IFAMR2016.0115 (visited on 08/22/2019).

[37] Jared Huffman. *Shelled Eggs: Sale for Human Consumption: Compliance with Animal Care Standards.* July 2010. URL: https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=200920100AB1437 (visited on 01/21/2020).

[38] Humane Farm Animal Care. *Animal Care Standards: Egg Laying Hens.* Tech. rep. Humane Farm Animal Care, Feb. 2018, pp. 1–45. URL: https://certifiedhumane.org/how-we-work/our-standards/ (visited on 06/20/2019).

[39] International Egg Commission. *Annual Review 2007*. Tech. rep. International Egg Commission, 2007. URL: https://www.internationalegg.com/wp-content/uploads/2015/11/Annual_Review_2007.pdf (visited on 08/29/2019).

[40] Riju Joshi and Jeffrey M. Wooldridge. "Correlated Random Effects Models with Endogenous Explanatory Variables and Unbalanced Panels". In: *Annals of Economics and Statistics* 134 (2019), pp. 243–268. ISSN: 2115-4430. DOI: 10.15609/annaeconstat2009.134.0243. URL: www.jstor.org/stable/10.15609/annaeconstat2009.134.0243 (visited on 12/05/2019).

[41] Deven King. *Legislation Changing Laying Hen Welfare Rules Worldwide*. Jan. 2019. URL: https://www.wattagnet.com/articles/36335-legislation-changing-laying-hen-welfare-rules-worldwide (visited on 09/01/2019).

[42] Malinda Larkin. *Another Ballot Initiative Increases Housing Size for Farm Animals*. Dec. 2018. URL: https://www.avma.org/News/JAVMANews/Pages/190101h.aspx (visited on 07/01/2019).

[43] Legislative Analyst's Office. *Proposition 12*. Nov. 2018. URL: https://lao.ca.gov/BallotAnalysis/Proposition?number=12&year=2018 (visited on 07/10/2019).

[44] Wei Lin and Jeffrey M. Wooldridge. "Chapter 2 - Testing and Correcting for Endogeneity in Nonlinear Unobserved Effects Models". In: *Panel Data Econometrics*. Ed. by Mike Tsionas. Academic Press, Jan. 2019, pp. 21–43. ISBN: 978-0-12-814367-4. URL: https://www.sciencedirect.com/science/article/pii/B9780128143674000022 (visited on 02/18/2022).

[45] Amy Mayer. *Cage-Free Egg Farms Take over California Market; Consumers Pay More*. Feb. 2022. URL: https://www.agri-pulse.com/articles/17171-californias-cage-free-egg-mandate-expands-market-for-producers-who-invested-in-new-systems (visited on 06/24/2022).

[46] Samara Mendez. *US Egg Production Data Set*. Tech. rep. The Humane League Labs, Aug. 2019. URL: https://osf.io/wta3q/ (visited on 08/18/2019).

[47] Samara Mendez, Jacob Peacock, and Matt Butner. *Impact of Corporate Commitments to Source Cage-Free Eggs on Layer Hen Housing*. Tech. rep. OSF Preprints, Aug. 2020. URL: https://osf.io/hkrsm/ (visited on 11/10/2021).

[48] Lorraine Mitchell. *Dolphin-Safe Tuna Labeling*. Economic Research Service Report AER-793. United States Department of Agriculture, Jan. 2001, pp. 22–25. URL: https://www.ers.usda.gov/webdocs/publications/41203/18892_aer793f.pdf?v=0 (visited on 11/07/2019).

[49] Daniela A. Miteva, Colby J. Loucks, and Subhrendu K. Pattanayak. "Social and Environmental Impacts of Forest Management Certification in Indonesia". In: *PLOS ONE* 10.7 (July 2015), e0129675. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0129675. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0129675 (visited on 09/12/2019).

[50] Anna Nordén, Jessica Coria, and Laura Villalobos. "Evaluation of the Impact of Forest Certification on Environmental Outcomes in Sweden". In: *University of Gothenburg Working Papers in Economics* No 657 (May 2016). ISSN: 1403-2465. URL: http://hdl.handle.net/2077/44417 (visited on 09/12/2019).

[51] Terrence O'Keefe. *What Can US Egg Producers Learn from Swiss Cage Ban?* Oct. 2011. URL: https://www.wattagnet.com/articles/10677-what-can-us-egg-producers-learn-from-swiss-cage-ban (visited on 09/01/2019).

[52] Chikako Oka. "Accounting for the Gaps in Labour Standard Compliance: The Role of Reputation-Conscious Buyers in the Cambodian Garment Industry". In: *The European Journal of Development Research* 22.1 (Feb. 2010), pp. 59–78. ISSN: 1743-9728. DOI: 10.1057/ejdr.2009.38. URL: https://doi.org/10.1057/ejdr.2009.38 (visited on 09/13/2019).

[53] Chikako Oka. "Channels of Buyer Influence and Labor Standard Compliance: The Case of Cambodia's Garment Sector". In: *Advances in Industrial and Labor Relations*. Ed. by David Lewin, Bruce E. Kaufman, and Paul J. Gollan. Vol. 17. Emerald Group Publishing Limited, Jan. 2010, pp. 153–183. ISBN: 978-1-84950-932-9. URL: https://www.emerald.com/insight/content/doi/10.1108/S0742-6186(2010)0000017008/full/html (visited on 09/13/2019).

[54] Open Wing Alliance. *The Case for Cage-Free Legislation*. Tech. rep. Open Wing Alliance, 2019, pp. 1–23.

[55] Leslie E. Papke and Jeffrey M. Wooldridge. "Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates". In: *Journal of Econometrics* 145.1-2 (July 2008), pp. 121–133. ISSN: 03044076. DOI: 10.1016/j.jeconom.2008.05.009. URL: https://linkinghub.elsevier.com/retrieve/pii/S030440760800050X (visited on 09/27/2019).

[56] Andrew S. Paul et al. "An Experiment on the Vote-Buy Gap with Application to Cage-Free Eggs". In: *Journal of Behavioral and Experimental Economics* 79 (Apr. 2019), pp. 102–109. ISSN: 2214-8043. DOI: 10.1016/j.socec.2019.02.005. URL: https://www.sciencedirect.com/science/article/pii/S2214804318300508 (visited on 04/18/2022).

[57] Janet G. Peckham and Emily Galloway. "What's in a Name? Using IRI Scanner Data to Evaluate Retail Food Labeling for Shell Eggs". In: *Agricultural & Applied Economics Association Annual Meeting*. Chicago, Illinois, July 2017. DOI: 10.22004/ag.econ.258267.

[58] Penn State Extension. *Modern Egg Industry*. July 2012. URL: https://extension.psu.edu/modern-egg-industry (visited on 08/24/2019).

[59] Amil Petrin and Kenneth Train. "A Control Function Approach to Endogeneity in Consumer Choice Models". In: *Journal of Marketing Research* 47.1 (2010), pp. 3–13. ISSN: 0022-2437. URL: https://www.jstor.org/stable/20618950 (visited on 09/27/2019).

[60] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: http://www.R-project.org/.

[61] Sean Ramos, Matthew MacLachlan, and Alex Melton. *Impacts of the 2014-2015 Highly Pathogenic Avian Influenza Outbreak on the U.S. Poultry Sector*. Economic Research Service Report LDPM-282-02. United States Department of Agriculture, Dec. 2017, p. 22. URL: https://www.ers.usda.gov/webdocs/publications/86282/ldpm-282-02.pdf?v=43089.

[62] Germano Glufke Reis and Carla Forte Maiolino Molento. "Emerging Market Multinationals and International Corporate Social Responsibility Standards: Bringing Animals to the Fore". In: *Journal of Business Ethics* (Mar. 2019). ISSN: 0167-4544, 1573-0697. DOI: 10.1007/s10551-019-04144-5. URL: http://link.springer.com/10.1007/s10551-019-04144-5 (visited on 01/02/2020).

[63] Helmut W. Saatkamp et al. "Transition from Conventional Broiler Meat to Meat from Production Concepts with Higher Animal Welfare: Experiences from The Netherlands". In: *Animals* 9.8 (Aug. 2019), p. 483. DOI: 10.3390/ani9080483. URL: https://www.mdpi.com/2076-2615/9/8/483 (visited on 08/27/2019).

[64] Karolina Sarek. *Did Corporate Campaigns in the US Have Any Counterfactual Impact? A Quantitative Model*. May 2019. URL: http://www.charityentrepreneurship.com/1/post/2019/05/corporate-campaigns-in-the-us.html (visited on 08/27/2019).

[65] R. Schulte-Drüggelte. "Determining the Optimum Replacement Schedule for Commercial Layers: Does Molting Pay Off?" In: *Lohmann Information* 1 (Apr. 2013), p. 47. URL: https://d-nb.info/1035538687/34.

[66] Reinhard Schunck. "Within and between Estimates in Random-Effects Models: Advantages and Drawbacks of Correlated Random Effects and Hybrid Models". In: *The Stata Journal* 13.1 (Mar.

2013), pp. 65–76. ISSN: 1536-867X, 1536-8734. DOI: 10.1177/1536867X1301300105. URL: http://journals.sagepub.com/doi/10.1177/1536867X1301300105 (visited on 12/05/2019).

[67] Sara Shields, Paul Shapiro, and Andrew Rowan. "A Decade of Progress toward Ending the Intensive Confinement of Farm Animals in the United States". In: *Animals* 7.12 (May 2017), p. 40. ISSN: 2076-2615. DOI: 10.3390/ani7050040. URL: http://www.mdpi.com/2076-2615/7/5/40 (visited on 05/14/2018).

[68] Saulius Šimčikas. *Corporate Campaigns Affect 9 to 120 Years of Chicken Life per Dollar Spent*. Aug. 2019. URL: https://www.rethinkpriorities.org/blog/2019/8/7/corporate-campaigns-affect-9-to-120-years-of-chicken-life-per-dollar-spent (visited on 08/27/2019).

[69] Peter Singer. *Ethics into Action*. First. 1998. ISBN: 0-8476-9073-3.

[70] Sodexo. *Sodexo Announces Shift to Eggs from Cage-free Hens*. Aug. 2012. URL: https://web.archive.org/web/20150131201717/http://www.prnewswire.com/news-releases/sodexo-announces-shift-to-eggs-from-cage-free-hens-167996086.html (visited on 01/23/2020).

[71] Sodexo. *Sodexo Leverages Scope of Its Supply Chain to Make Positive Economic Impact in the Local Communities Where It Operates*. Oct. 2016. URL: https://www.prnewswire.com/news-releases/sodexo-leverages-scope-of-its-supply-chain-to-make-positive-economic-impact-in-the-local-communities-where-it-operates-300340694.html (visited on 04/18/2022).

[72] StataCorp. *Stata Statistical Software*. StataCorp LLC. College Station, TX, 2013. URL: https://www.stata.com/.

[73] State of California. *STANDARDS FOR CONFINING FARM ANIMALS Proposition 2*. 2008. URL: https://repository.uchastings.edu/ca_ballot_props/1282.

[74] State of Oregon. *Chapter 686*. 2019. URL: https://www.oregonlegislature.gov/bills_laws/lawsstatutes/2019orlaw0686.pdf (visited on 01/06/2020).

[75] Ron Sterk. *Egg Prices Fall on Slow Sales and Ample Supplies*. Feb. 2019. URL: https://www.foodbusinessnews.net/articles/13263-egg-prices-fall-on-slow-sales-and-ample-supplies?v=preview (visited on 08/27/2019).

[76] United Egg Producers. *Complete Guidelines for Cage and Cage-Free Housing*. Tech. rep. United Egg Producers, 2017. URL: https://unitedegg.com/wp-content/uploads/2017/11/2017-UEP-Animal-Welfare-Complete-Guidelines-11.01.17-FINAL.pdf (visited on 08/24/2019).

[77] Belinda Vigors. "Reducing the Consumer Attitude–Behaviour Gap in Animal Welfare: The Potential Role of 'Nudges'". In: *Animals* 8.12 (Dec. 2018), p. 232. ISSN: 2076-2615. DOI: 10.3390/ani8120232. URL: http://www.mdpi.com/2076-2615/8/12/232 (visited on 09/18/2019).

[78] Leon Walker. *Aramark Makes Cage-Free Egg Pledge*. Jan. 2013. URL: https://www.environmentalleader.com/2013/01/aramark-makes-cage-free-egg-pledge/ (visited on 01/23/2020).

[79] Richard Williams. "Fitting Heterogeneous Choice Models with Oglm". In: *The Stata Journal* 10.4 (Dec. 2010), pp. 540–567. ISSN: 1536-867X. DOI: 10.1177/1536867X1101000402. URL: https://doi.org/10.1177/1536867X1101000402 (visited on 01/26/2022).

[80] Hans-Wilhelm Windhorst. "Dynamics and Patterns of the EU Egg Industry". In: *Lohmann Information* 51.2 (2017), p. 7. URL: https://lohmann-breeders.com/media/2020/08/Windhorst.pdf.

[81] Jeffrey M. Wooldridge. "A Simple Method for Estimating Unconditional Heterogeneity Distributions in Correlated Random Effects Models". In: *Economics Letters* 113.1 (Oct. 2011), pp. 12–15. ISSN: 0165-1765. DOI: 10.1016/j.econlet.2011.05.019. URL: https://www.sciencedirect.com/science/article/pii/S0165176511001960 (visited on 01/12/2022).

[82] Jeffrey M. Wooldridge. "Control Function Methods in Applied Econometrics". In: *Journal of Human Resources* 50.2 (Mar. 2015), pp. 420–445. ISSN: 0022-166X, 1548-8004. DOI: 10.3368/jhr.50.2.420. URL: http://jhr.uwpress.org/content/50/2/420 (visited on 02/22/2022).

[83] Jeffrey M. Wooldridge. "Correlated Random Effects Models with Unbalanced Panels". In: *Journal of Econometrics*. Annals Issue in Honor of Jerry A. Hausman 211.1 (July 2019), pp. 137–150. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2018.12.010. URL: https://www.sciencedirect.com/science/article/pii/S0304407618302392 (visited on 08/20/2021).

[84] World Organisation for Animal Health (OIE). *OIE-WAHIS Data: Animal Disease Events*. 2022. URL: https://wahis.oie.int/#/events (visited on 12/14/2021).

[85] Hurria Yassin et al. "A Model for an Economically Optimal Replacement of a Breeder Flock". In: *Poultry Science* 91.12 (Dec. 2012), pp. 3271–3279. ISSN: 0032-5791. DOI: 10.3382/ps.2012-02290. URL: https://academic.oup.com/ps/article/91/12/3271/1569933 (visited on 08/26/2019).

[86] Helene York. *Inside One Corporation's Decision to "Go Humane"*. Feb. 2012. URL: https://web.archive.org/web/20220418203534/https://civileats.com/2012/02/21/inside-one-corporation%E2%80%99s-decision-to-%E2%80%9Cgo-humane%E2%80%9D/ (visited on 04/18/2022).

A Additional Figures and Tables

Table 5: Linear CRE and FE control function IV regressions

Model	(1) OLS 1st stage	(2) L CRE 2nd stage	(3) FE 1st stage	(4) FE 2nd stage
Dependent variable	*cumul_commitments*	*prop_hens_cage_free*	*cumul_commitments*	*prop_hens_cage_free*
cumul_commitments		0.0033		0.0001
		(0.0030)		(0.0001)
		-0.0027 - 0.0092		-0.0001 - 0.0002
cumul_campaigns	4.5126***		4.5501***	
	(0.0516)		(0.0976)	
	4.4113 - 4.6139		4.3532 - 4.7470	
L CRE residuals		-0.0033		
		(0.0031)		
		-0.0095 - 0.0028		
FE residuals				0.0018
				(0.0010)
				-0.0004 - 0.0039
birds_killed_flu	$9.11\text{x}10^{-07}$***	$-2.03\text{x}10^{-10}$	$9.22\text{x}10^{-07}$***	$-1.55\text{x}10^{-10}$
	$(2.98\text{x}10^{-07})$	$(2.69\text{x}10^{-09})$	$(2.64\text{x}10^{-07})$	$(2.52\text{x}10^{-09})$
	$3.26\text{x}10^{-07}$ - $1.50\text{x}10^{-06}$	$-5.48\text{x}10^{-09}$ - $5.08\text{x}10^{-09}$	$3.89\text{x}10^{-07}$ - $1.46\text{x}10^{-06}$	$-5.24\text{x}10^{-09}$ - $4.93\text{x}10^{-09}$
legislation	1.2380	0.0546**	1.1511	0.0534**
	(1.1071)	(0.0238)	(1.3802)	(0.0244)
	-0.9376 - 3.4137	0.0080 - 0.1012	-1.6322 - 3.9345	0.0042 - 0.1026
avg_commitments	0.1332***	-0.0006		
	(0.0213)	(0.0008)		
	0.0913 - 0.1751	-0.0021 - 0.0010		
avg_birds_killed_flu	$-3.38\text{x}10^{-06}$***	$-2.95\text{x}10^{-08}$		
	$(9.93\text{x}10^{-07})$	$(6.15\text{x}10^{-08})$		
	$-5.33\text{x}10^{-06}$ - $-1.42\text{x}10^{-06}$	$-1.50\text{x}10^{-07}$ - $9.09\text{x}10^{-08}$		
avg_years_legislation	-1.9771	0.5810***		
	(1.8678)	(0.2080)		
	-5.6476 - 1.6934	0.1734 - 0.9886		
intercept	20.9039*	0.7402	-0.4169	0.1930***
	(11.1892)	(1.0507)	(0.5673)	(0.0174)
	-1.0845 - 42.8923	-1.3191 - 2.7994	-1.5610 - 0.7272	0.1580 - 0.2280
Observations	488	488	488	488
F test of instrument strength (p)	553 (0)		1352 (0)	
Year FE	YES	YES	YES	YES
Number of countries	44	44	44	44

This table reports the first and second stages of the control function instrumental variables (IV) regressions for the linear correlated random effects (L CRE) and fixed effects (FE) regressions. The first stage regression is estimated using pooled ordinary least squares (OLS) and FE for the L CRE and FE models, respectively. The second-stage regression includes the results of the test for simultaneity: that is, statistical significance of the coefficient on *residuals*. The t test of simultaneity uses the null hypothesis that the coefficient on *residuals* is equal to zero. F tests of instrument strength in Column 1 indicate that *cumul_campaigns* is a suitably strong instrument for *cumul_commitments*.
Standard errors (in parentheses) are clustered by country. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are below the standard errors.

Table 6: Control function IV robustness check regressions

	(1)	(2)
Model	NL CRE Preferred specification, balanced subsample	NL CRE IV With control function IVs
Dependent variable	*prop_hens_cage_free*	*prop_hens_cage_free*
cumul_commitments	0.00033***	0.00027
	(0.0001)	(0.0004)
	0.0001 - 0.0005	-0.0006 - 0.0011
residuals		0.00080
		(0.0013)
		-0.0017 - 0.0033
birds_killed_flu	2.77x10^{-09}	2.30x10^{-09}
	(2.41x10^{-09})	(1.65x10^{-08})
	-1.96x10^{-09} - 7.49x10^{-09}	-3.01x10^{-08} - 3.47x10^{-08}
legislation	0.00323	0.00463
	(0.0228)	(0.0216)
	-0.0414 - 0.0479	-0.0378 - 0.0470
avg_commitments	0.00277	0.00282
	(0.0020)	(0.0040)
	-0.0011 - 0.0067	-0.0050 - 0.0107
avg_birds_killed_flu	-3.81x10^{-07}**	-3.79x10^{-07}
	(1.52x10^{-07})	(4.30x10^{-07})
	-6.79x10^{-07} - -8.37x10^{-08}	-1.22x10^{-06} - 4.64x10^{-07}
avg_years_legislation	0.66452***	0.663**
	(0.2452)	(0.3272)
	0.1840 - 1.1451	0.0217 - 1.3043
Observations	351	351
Year FEs	Y	Y
Year means	Y	Y

This table reports the average partial effects (APEs) of regressions comparing the robustness of the control function instrumental variables (IV) method. Results are estimated on the fully balanced subsample of the panel described in Section 4. Column 1 reports the results of the homoskedastic nonlinear correlated random effects (NL CRE) regression, while Column 2 reports the results of the same regression using the control function IV method.
Robust standard errors (in parentheses) are clustered by country and obtained from 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1. 95% confidence intervals are shown below standard errors.

Table 7: Coefficients of nonlinear CRE regressions

	(1) NL CRE Preferred specification	(2) NL CRE Accounts for heteroskedasticity
Dependent variable	*prop_hens_cage_free*	*prop_hens_cage_free*
cumul_commitments	0.00129***	0.000005***
	(0.0004)	(0.0000)
	0.0005 - 0.0021	0.0000 - 0.0000
birds_killed_flu	1.66x10^{-08}	4.03x10^{-11}
	(1.52x10^{-08})	(3.86x10^{-11})
	-1.31x10^{-08} - 4.64x10^{-08}	-3.53x10^{-11} - 1.16x10^{-10}
legislation	0.00149	0.00003
	(0.1071)	(0.0004)
	-0.2084 - 0.2114	-0.0007 - 0.0007
avg_commitments	0.00652	0.00004
	(0.0057)	(0.0000)
	-0.0046 - 0.0177	-0.0000 - 0.0001
avg_birds_killed_flu	-1.05x10^{-06}**	-5.60x10^{-09}**
	(4.16x10^{-07})	(2.36x10^{-09})
	-1.87x10^{-06} - -2.33x10^{-07}	-1.02x10^{-08} - -9.82x10^{-10}
avg_years_legislation	2.21776**	0.01041***
	(0.8984)	(0.0037)
	0.4569 - 3.9786	0.0031 - 0.0177
intercept	-0.30344	0.03601
	(3.2350)	(0.0561)
	-6.6439 - 6.0370	-0.0740 - 0.1460
Observations	488	432
Number of countries	44	34
Pseudolikelihood	-177.2503	-220.7618
AIC	0.841	1.124
χ^2 test of CRE model (p)	158463 (0)	161027 (0)
χ^2 test of heteroskedasticity (p)		991.9 (0)

This table reports the raw coefficient results of the nonlinear correlated random effects (NL CRE) regressions, from which the average partial effects (APEs) in Table 3 Columns 1 and 2, respectively, are derived. Column 2 accounts for heteroskedasticity (Het) and estimates the regression on a subsample of countries that appear in at least 10 of the 13 time periods in our sample. Akaike information criterion (AIC) = $\frac{-2\ln\hat{L}+2k}{N}$.
Standard errors (in parentheses) are clustered by country and are obtained by the delta method for the APEs. *** p <0.01, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are shown below standard errors.

Table 8: Subsample summary statistics

	Full sample summary stats				Subsample summary stats			
	mean	sd	min	max	mean	sd	min	max
cumul_commitments	10.54	34.38	0	416	11.13	36.31	0	416
birds_killed_flu	165,620	999,176	0	16.01M	156,545	10.40M	0	16.01M
legislation	0.369	0.483	0	1	0.417	0.494	0	1
prop_hens_cage_free	0.258	0.264	0	1	0.280	0.269	0	1

$N = 488$ observations for all variables in the full sample, $N = 432$ observations in the subsample.

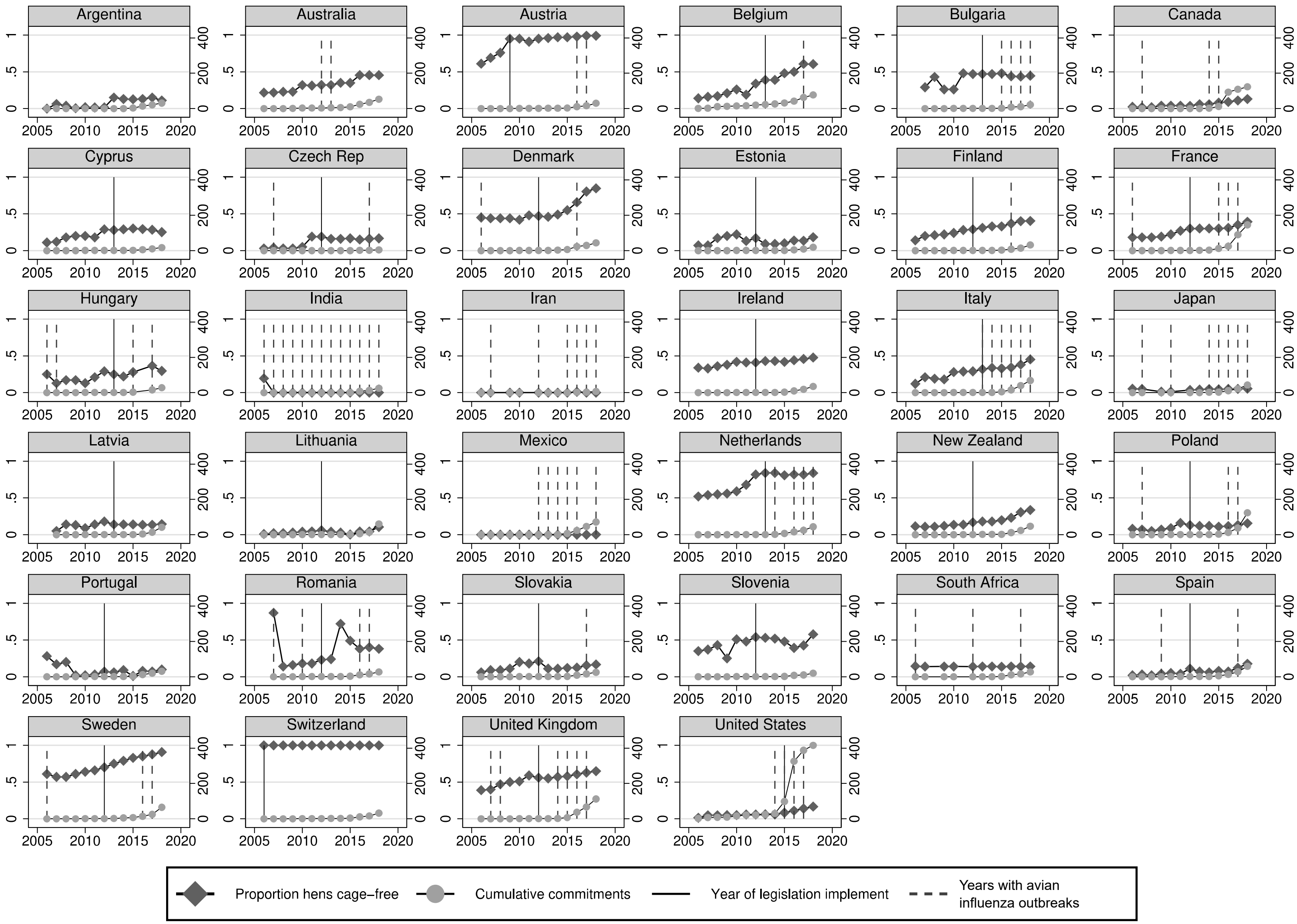
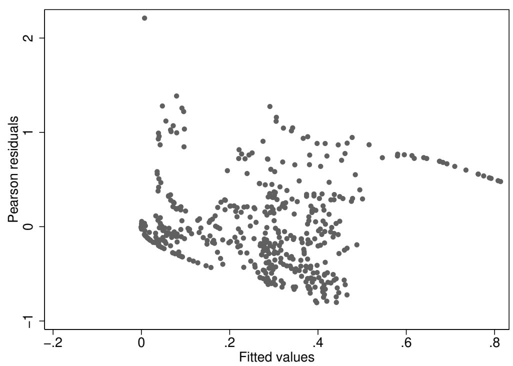


Figure 3: Trends in cage-free proportions, for countries appearing in more than 11 years of data. The connected black line denotes the dependent variable *prop_hens_cage_free*, while the connected gray line shows the explanatory variable of interest, *cumul_commitments*. The solid vertical line depicts the year in which a country's cage-free legislation was implemented, and dashed vertical lines indicate years in which at least one recorded death due to highly pathogenic avian influenza occurred.

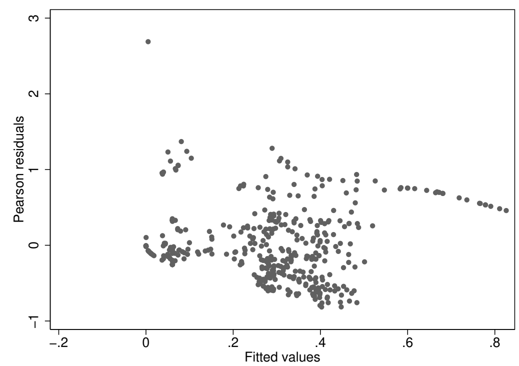
Table 9: Robustness check regressions

Model	(1) NL CRE Restricted subsample	(2) NL CRE Including Germany
Dependent variable	*prop_hens_cage_free*	*prop_hens_cage_free*
cumul_commitments	0.00036***	0.00046**
	(0.0001)	(0.0002)
	0.0001 - 0.0006	0.0001 - 0.0008
birds_killed_flu	6.50x10^{-09}	4.76x10^{-09}
	(7.42x10^{-09})	(4.10x10^{-09})
	-8.04x10^{-09} - 2.10x10^{-08}	-3.27x10^{-09} - 1.28x10^{-08}
legislation	0.00886	0.00380
	(0.0200)	(0.0295)
	-0.0304 - 0.0481	-0.0541 - 0.0617
avg_commitments	0.00193	0.00373***
	(0.0016)	(0.0009)
	-0.0013 - 0.0051	0.0019 - 0.0056
avg_birds_killed_flu	-1.04x10^{-06}**	-4.30x10^{-07}***
	(4.18x10^{-07})	(1.09x10^{-07})
	-1.86x10^{-06} - -2.19x10^{-07}	-6.44x10^{-07} - -2.15x10^{-07}
avg_years_legislation	0.63779***	0.62202***
	(0.2376)	(0.2275)
	0.1721 - 1.1035	0.1762 - 1.0678
Observations	432	501
Year FEs	Y	Y
Year means	Y	Y
Estimated parameters	21	28
Number of FEs	2	10
Number of countries	34	45
Pseudolikelihood	-166.6497	-182.49017
AIC	0.869	0.840
χ^2 test of CRE model (p)	48.65(8.74e-09)	229549 (0)
χ^2 test, between effect of *cumul_commitments* (p)	2.018 (0.155)	17.02 (3.70e-05)
χ^2 test, between effect of *legislation*(p)	7.199 (0.00729)	7.529 (0.00607)

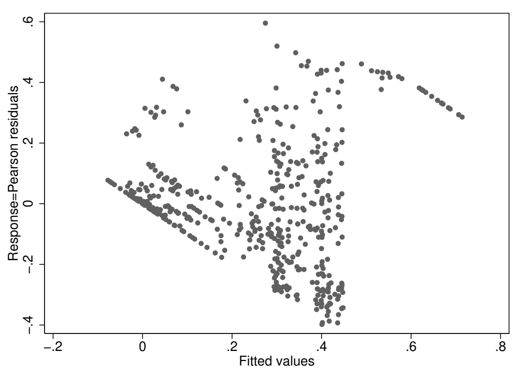
This table reports the average partial effects (APEs) of various robustness check regressions. Column 1 presents the APEs of the homoskedastic nonlinear correlated random effects (NL CRE) regression estimated on the restricted subsample used in Column 2 of Table 3. Column 2 presents APEs of the preferred specification with Germany included. Akaike information criterion (AIC) = $\frac{-2\ln\hat{L}+2k}{N}$. The Durbin-Wu-Hausman χ^2 statistics to test the CRE estimators uses the null hypothesis that the CRE estimators are not jointly significantly different from the FE estimators. The χ^2 test of heteroskedasticity uses the null hypothesis that the variables included in the variance during the estimation procedure are not jointly significant. The χ^2 test of the between effect uses the null hypothesis that the between effect $(\beta_i + \xi_i)$ equals zero, or $H_0 : \beta_i = -\xi_i$.
Standard errors (in parentheses) are clustered by country and are obtained by the delta method for the APEs. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 95% confidence intervals are shown below standard errors.



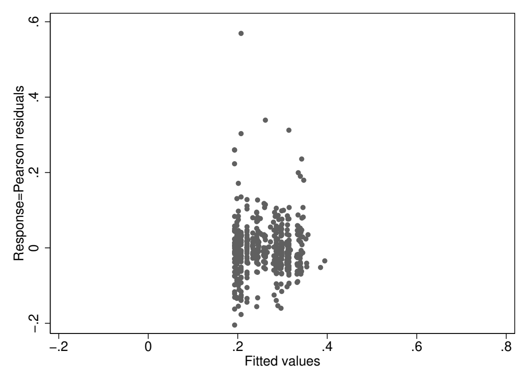
(a) NL CRE
(Preferred specification)



(b) NL CRE Het

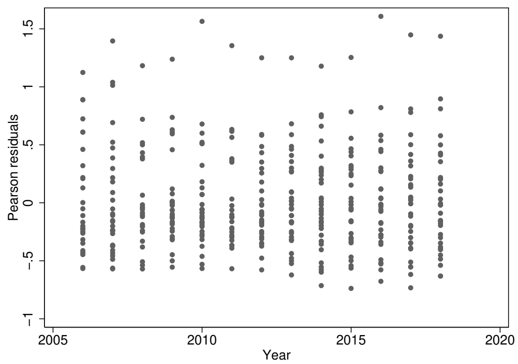


(c) L CRE

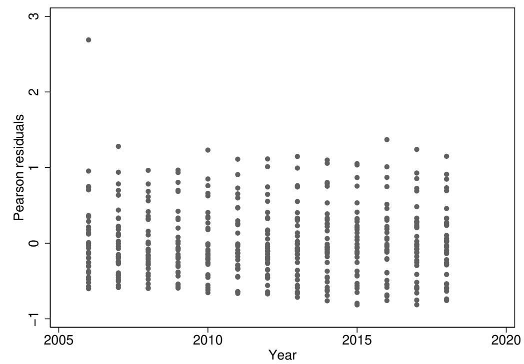


(d) FE

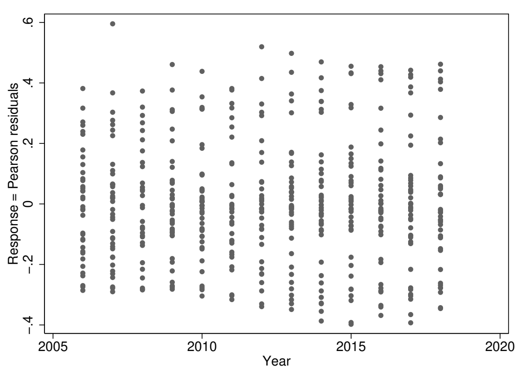
Figure 4: Pearson residuals versus fitted values plotted for the four main regressions shown in Table 3. Figure 6a plots residuals for the preferred specification, the homoskedastic nonlinear correlated random effects (NL CRE) model. Figure 6b shows the NL CRE model that accounts for heteroskedasticity. Figure 6c shows the linear CRE regression. Figure 6d shows the fixed effects (FE) regression. Notable artifacts of the data are apparent in these plots: the diagonal lower and upper bounds in plots 4c, 4a, and to a lesser degree 4b are the result of the dependent variable bounded between 0 and 1. The vertical outlier in plots 4d, 4a, 4b is India in 2006 (which can be seen in Figures 2 and 3) and is likely the result of a change in data collection methods. The extreme values in the right-hand tail of these plots belong to the same few countries that have relatively high proportions of cage-free housing.



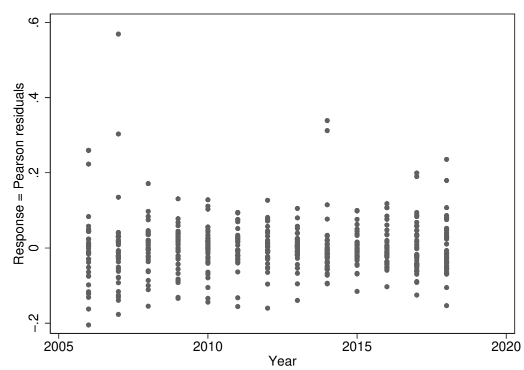
(a) NL CRE
(Preferred specification)



(b) NL CRE Het

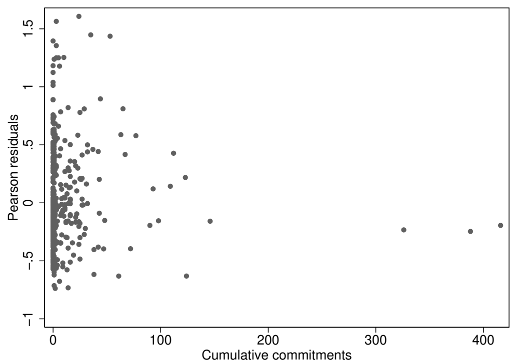


(c) L CRE

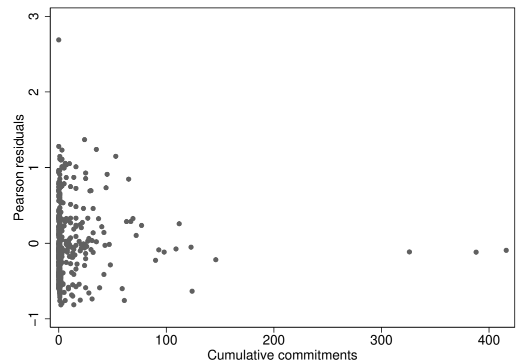


(d) FE

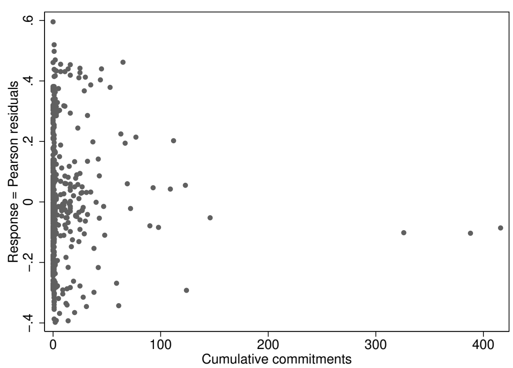
Figure 5: Pearson residual versus year plotted for the four main regressions shown in Table 3. Figure 6a plots residuals for the preferred specification, the homoskedastic nonlinear correlated random effects (NL CRE) model. Figure 6b shows the NL CRE model that accounts for heteroskedasticity. Figure 6c shows the linear CRE regression. Figure 6d shows the fixed effects (FE) regression.



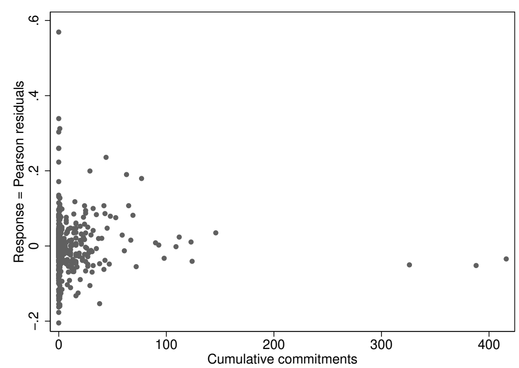
(a) NL CRE
(Preferred specification)



(b) NL CRE Het



(c) L CRE



(d) FE

Figure 6: Pearson residual versus cumulative commitments plotted for the four main regressions shown in Table 3. Figure 6a plots residuals for the preferred specification, the homoskedastic nonlinear correlated random effects (NL CRE) model. Figure 6b shows the NL CRE model that accounts for heteroskedasticity. Figure 6c shows the linear CRE regression. Figure 6d shows the fixed effects (FE) regression.