PLOS ONE

Enhancing detection of labor violations in the agricultural sector: A multilevel generalized linear regression model of H-2A violation counts

Arezoo Jafari[1]*, Priscila De Azevedo Drummond[1], Shawn Bhimani[2], Dominic Nishigaya[3], Aidong Adam Ding[4], Amy Farrell[3], Kayse Lee Maass[1]*

1 Department of Mechanical and Industrial Engineering, Northeastern University, Boston, Massachusetts, United States of America, 2 D'Amore-McKim School of Business, Northeastern University, Boston, Massachusetts, United States of America, 3 School of Criminology and Criminal Justice, Northeastern University, Boston, Massachusetts, United States of America, 4 Department of Mathematics, Northeastern University, Boston, Massachusetts, United States of America

* jafari.a@northeastern.edu (AJ); k.maass@northeastern.edu (KLM)

Abstract

Agricultural workers are essential to the supply chain for our daily food, and yet, many face harmful work conditions, including garnished wages, and other labor violations. Workers on H-2A visas are particularly vulnerable due to the precarity of their immigration status being tied to their employer. Although worksite inspections are one mechanism to detect such violations, many labor violations affecting agricultural workers go undetected due to limited inspection resources. In this study, we identify multiple state and industry level factors that correlate with H-2A violations identified by the U.S. Department of Labor's Wage and Hour Division using a multilevel zero-inflated negative binomial model. We find that three state-level factors (average farm acreage size, the number of agricultural establishments with less than 20 employees, and higher poverty rates) are correlated with H-2A violations. These findings offer valuable insights into where H-2A violations are being detected at the state and industry levels.

Introduction

The U.S. agricultural industry is crucial for ensuring food security and contributing to the national economy. This industry comprises various sectors, such as crop production, livestock farming, and food processing. However, it faces numerous challenges such as labor shortages [1], climate change [2], and a growing demand for sustainable practices [3]. As a result, there is an increasing need for migrant farm workers.

Migrant farm workers are essential for fulfilling the labor demands in the U.S. agricultural sector, particularly in seasonal and labor-intensive areas. International migrants occupy a significant portion of entry-level jobs in agriculture [4]. Employers often cite the lack of qualified local workers, the superior work ethic of migrants, and, occasionally, the lower labor costs

associated with migrant workers as reasons for their preference [4]. The latter is especially true when migrants are excluded from pension and other benefit programs, which can add 20 to 30 percent to labor costs [4]. However, migrant workers are susceptible to exploitation, as they face job-related risks, such as exposure to farm chemicals, substandard living conditions, and violations of work hours and wages.

To address labor shortages in the agriculture industry, the U.S. has implemented the H-2 (and, since 1986, the H-2A) visa programs since 1952. These programs enable farmers to fill seasonal farm jobs with guest workers from other countries when they anticipate a shortage of domestic workers, including U.S. citizens, other legally authorized workers, and undocumented workers. Employers seeking to recruit and hire H-2A workers must first obtain certification from the Office of Foreign Labor Certification (OFLC) of the U.S. Department of Labor (DOL) [1]. It is important to note that the H-2 programs are not new phenomena but rather part of a long history of the United States relying on low-paid workers to support the agriculture industry, beginning with slavery and progressing through sharecropping and the Bracero program, which brought workers from Mexico to fill the labor needs in agriculture. [5, 6]. It should be noted that not all H-2A workers face exploitation, yet the history of labor exploitation impacts the H-2A program, necessitating ongoing efforts to enhance labor protections.

Farm workers on H-2A visas are an essential component of the agricultural industry, as indicated by the rapid increase in H-2A workers employed in the U.S. agricultural industry over the past decade [7, 8]. Most agricultural industries have experienced significant growth in H-2A employment, particularly in industries with high labor requirements and seasonal employment [7]. From fiscal year 2010 to 2020, the number of jobs certified for H-2A worker employment increased from approximately 75,000 to 275,000. Six states comprised 55 percent of these H-2A certified jobs: Florida with 14 percent, Georgia and Washington each with 10 percent, California at 9 percent, North Carolina at 8 percent, and Louisiana at 4 percent [9]. Around 80 percent of job certifications in fiscal year 2020 lead to the issuance of visas for H-2A workers [9]. However, not all employers complete the hiring process for H-2A workers, and certain workers may occupy more than one certified role.

While H-2A policies state that H-2A workers must receive certain benefits not consistently afforded to non-H-2A agricultural workers, such as required workers' compensation coverage [10], these benefits don't always come to fruition. Additionally, H-2A workers are limited to working for the employer who applied for their services through the DOL. (However, during the COVID-19 pandemic, temporary changes were made to the Department of Homeland Security requirements for H-2A change of employer requests and H-2A maximum period of stay exceptions [11].) This restriction, coupled with language barriers, lack of knowledge about their rights, and limited access to legal assistance, leaves workers with few options if they encounter abusive work conditions or wage theft. In situations of abusive labor experiences, their only employment options are to return to their home countries, which they may be reluctant to choose due to losing their current job, fear of blacklisting, financial limitations, and other uncertainties. As a result, some workers may feel compelled to continue working under exploitative conditions [12].

Farm workers who rely on their employment to support themselves and their families often face economic vulnerability and hesitate to speak up for themselves, even in situations that may put their lives at risk [13]. Additionally, their limited financial resources make it difficult for them to access the legal services necessary to assert their rights fully [14]. One of the significant obstacles facing farm workers is their access to health services, as they are exposed to a range of occupational and environmental health risks, which result in high levels of physical injury and illness. The challenges to farm workers' use of health services include inadequate English proficiency, limited education, low income, and a lack of health insurance [15].

Several studies discuss strategies to improve and target efforts to protect farm workers [16–18]. Farm workers in the United States earn some of the lowest incomes in the labor market and experience a high rate of injuries [16, 19]. The U.S. DOL reveals that the federal and state governments lack the resources for even a one-time inspection of all labor camps in a state [20]. Costa et al. [16] also reported that there is a low probability (1.1%) that each farm employer will be investigated by the DOL Wage and Hour Division (WHD) per year. They recommended strategies to target violators, including increasing wage and hour staffing and enforcement funds, targeting the farm labor contractors who are the biggest violators, and ensuring sufficient penalties to stop future violations. Critical to these efforts is understanding factors that increase the risk of H-2A violations.

In this paper, we used WHD data that recorded the number of H-2A violations detected by DOL between 2010 and 2020. To examine the factors correlated with the number of H-2A violations, we assembled a dataset containing various factors from multiple sources. This allowed us to pinpoint the factors significantly associated with H-2A violations, thereby offering insights for optimizing resource allocation, such as budget, time, and personnel resources. This may allow inspectors to prioritize worksites based on the likelihood of labor violations, thereby improving the effectiveness of strategies to address such violations. We employed a multilevel zero-inflated negative binomial model to identify the correlation between factors and the count outcome. Our findings provide valuable insights that can be used by inspection agencies to optimize their investigative strategies, resulting in more efficient and effective enforcement of labor regulations.

Our study makes several contributions to the literature on H-2A violations in the agricultural industry. Firstly, we conducted a literature review to identify the key factors that may be correlated with the detection of H-2A violations. Secondly, while previous studies have used quantitative analyses on certain factors, to the best of our knowledge, our study is unique in its application of a multilevel zero-inflated negative binomial model. This approach allows for exploring correlations between significant factors and H-2A violations at both state and industry levels, providing a better understanding of the issue. Thirdly, based on the suggested model, we found that the average size of farms and number of small/medium establishments are correlated with the mean of H-2A violations. Additionally, labor intensity, number of task forces, duration of alleged violations, fatal injury rate, and minimum standard wage are associated with the likelihood of zero H-2A violation counts at both state and industry levels. The poverty rate is associated with both mean of H-2A violations and the likelihood of zero H-2A violation counts. By pinpointing the significant factors associated with H-2A violations, our study offers valuable insights that can guide policy decisions in the realm of labor regulation and enforcement.

The rest of this paper is structured as follows: The next section provides a literature review on factors that may be related to H-2A violations and discusses relevant multilevel and zero-inflated models employed in previous studies. The Data section details the factors considered for our regression analysis and outlines the data sources used to obtain these variables. In the Method section, we describe the regression models and multilevel generalized linear models applied to identify correlations between factors and H-2A violation counts. The Results section presents and interprets our analysis results, emphasizing key factors that significantly correlate with H-2A violations in the agricultural industry. Lastly, in the Discussion and Conclusion sections, we summarize our findings, discuss the practical implications of our study, and suggest strategies to improve the effectiveness of investigations. Overall, our research provides valuable insights that can inform policy decisions and support the development of more effective approaches for detecting labor violations in this industry.

Literature review

This section aims to briefly review relevant studies focused on factors correlated with the risk of labor violations, as well as regression models employed in the analysis of similar data structures.

Factors associated with labor violations among agricultural workers

Many scholars have studied the factors that increase the vulnerability of farm workers to exploitation. Agricultural work environments are notorious for increased risk of occupational injuries due to environmental heat stress, pesticide exposure, heavy workloads, and other workplace safety concerns [15, 21–26], including fatal injuries [27–30]. Social distancing was particularly challenging in agricultural workplaces during the COVID-19 pandemic [31]. Prior studies also document the impacts such work environments and stressors have on farm workers' mental health [32–35]. Additionally, wage violations [36–38] and violence [39] are widely documented in the agricultural industry. Prior research has explored how these issues are impacted by the contract type and recruitment strategy [7, 16, 40, 41], farm size [42–44], demographic characteristics of workers [45–47], migrant workers [7, 20] and labor inspections [13, 48]. We sought to explore the relationship of these factors to the number of H-2A violations detected by the DOL. However, due to the lack of available data on each of these factors during a common time period, our analysis focuses on the subset of factors for which data is available (as described further in the Data Section). Below, we summarize the literature on the aforementioned factors for which data is available and are included in our study. Specifically, we focus on farm size (according to multiple metrics: acreage, number of employees), labor intensity, average net income per farm, poverty rate, and fatal injury rate. We also explore the relationship between the presence of human trafficking taskforces and the state minimum standard wage on detected H-2A violations.

Farm size. Several studies have focused on farm size in relation to labor exploitation. Some have found that labor regulations are less strict for smaller firms (number of employees) [42, 43]. An article examining the impacts of firms' size on the risk of labor violations found a high probability of labor violation among employees in small (less than ten employees) and medium-sized firms (10 to 19 employees) rather than larger enterprises [42].

The literature highlights the importance of considering the correlation between the land size of farms (measured in acres) and labor violations, in addition to the impacts of labor laws on farms with fewer employees. Harrison and Getz [44] found that although larger acreage farms generally offer better job quality than smaller ones, these advantages are disproportionately available to white, U.S.-born workers. Migrant workers face challenges such as job insecurity, limited professional growth, and fear of losing their jobs due to their legal status. Their study highlights the need to examine the relationship between farm size and labor violations, considering the varying job quality and opportunities for different worker groups. Larger acreage farms may not always be more prone to labor violations, but specific subgroups, such as migrant workers, could be more vulnerable due to the mentioned factors.

Labor intensity. The main characteristic of high-intensive jobs is the low use of machinery and intensive use of manual labor [49]. These increase workers' exposure to nature and its adversities. Articles have also documented that high-intensive work is highly dependent on temporary migrant workers [7, 50]. One such group of migrant workers that is particularly relevant to this research is workers on H-2A (agricultural) visas. Castillo et al. [7] described the trends in the H-2A program in the agriculture sector, finding H-2A employment growth in vegetables and melons, fruits, and tree nuts that reflect sectors with higher labor intensity. Since migrant workers are susceptible to labor exploitation because of their apprehension

about job loss, heightened exposure to hazardous work environments, and extended working hours [51], it is crucial to investigate H-2A violations in labor-intensive workplaces.

Income. The literature highlights the need to examine the correlation between farm income and labor exploitation in the agricultural sector. We could not find articles that specifically identify a relationship between farm income and labor violations in the agricultural industry. However, the literature discussed that businesses maximize labor efficiency to earn profits. This pressure to extract maximum profit from the work process often results in various labor standard issues, such as excessive working hours, low wages, and an over-reliance on contingent migrant workers [52].

Task forces. Barrick et al. [13] studied law enforcement's ability to identify, investigate and prosecute labor trafficking among farm workers. The study revealed a discrepancy between the perceptions of law enforcement, including task forces, regarding labor trafficking operations and the actual reports of labor trafficking by agricultural workers in parts of North Carolina. They recommended that law enforcement agencies in agricultural areas extend their missions to include the protection of laborers, farms, and farm camps in their routine investigations. Farrell et al. [48] explored strategies law enforcement, including multi-agency task forces, can use to improve detection, including developing non-traditional partnerships with labor inspections and local regulatory agencies.

Poverty rate. Schwarz et al. [38] presented major clusters of trafficking risk factors, including economic insecurity, house insecurity, education gaps, and migration. They expressed that the relationship between poverty and labor trafficking is strong since poverty restricts people's options and makes them vulnerable to labor exploitation. Also, several studies discussed that the poverty rate influences the chance of children becoming child laborers [53]. Marinescu et al. [36] identified the negative correlation between wages and violations enforced by the U.S. Occupational Safety and Health Administration (OSHA) and WHD. Nagurney [37] employed a network equilibrium model to demonstrate that wages play a crucial role in decisions to hire migrant workers. The study also revealed that engaging in illicit practices, such as misleading potential workers about their compensation, can financially benefit farmers, thereby emphasizing the need for stricter oversight and control.

Fatal injury rate. Some articles suggest that agriculture is one of the most dangerous occupational sectors regarding fatal injuries [27–29]. Rivara [30] studied fatal and nonfatal injuries among agricultural workers in the US and found that farm machinery, such as tractors, is the most common reason for fatal and nonfatal injuries.

State minimum standard wages. The Fair Labor Standards Act (FLSA) establishes the federal minimum wage and directly elicits the minimum wage provisions for employers to compensate all employees legally. Additionally, the Adverse Effect Wage Rate (AEWR) is a critical component of the H-2A visa program, designed to protect both U.S. and foreign farm workers from adverse impacts on wages due to the employment of H-2A visa holders. Set by the DOL, the AEWR is the minimum wage that employers must pay to H-2A workers and any domestic workers in similar employment to prevent the undercutting of local wage standards [54]. The literature suggests that when the minimum wage increases beyond what a firm can or is willing to pay, maintaining a job match may necessitate paying sub-minimum wages to some workers [55]. As a result, workers might stop seeking enforcement of the minimum wage, for instance, by choosing not to report violations of minimum wage regulations, when they fear their job is in danger [55, 56]. It is important to consider the role of minimum wage standards in labor violations, particularly within the agricultural sector.

Multilevel and zero-inflated models

Multilevel modeling is a robust statistical technique for analyzing data with group structures. This method has been employed across various disciplines to tackle diverse research questions, effectively demonstrating its worth in discerning both within-group and between-group effects and interactions [57]. Multilevel modeling enables researchers to accommodate intricate data structures and has found applications in various social science domains, such as education [58, 59], health [60, 61], environmental studies [62, 63], and violence-related research [64, 65].

Some articles have depicted the linear association between predictors and continuous outcomes by constructing multilevel linear models for detecting significant factors [66, 67]. On the other hand, the generalized linear multilevel models typically utilized for analyzing count, binary, or categorical data differ from the former. Leclerc et al. [68] employed a mixed-effect logistic regression analysis to study the effects of potential guardianship on the severity of child sexual abuse for nested data. Martinez-Schuldt et al. [69] studied the willingness of immigrant community members to notify law enforcement after being victimized. They used multilevel logistic regression analysis for a binary dependent variable.

Several studies particularly focused on count data when the data are clustered or grouped [70–76], using the Poisson distribution [71, 72] or negative binomial distribution [73]. Count data often exhibit excess zeros, meaning there are more zeros in the data than expected from a Poisson or negative binomial distribution. In such cases, two commonly used models are the zero-inflated Poisson (ZIP) regression model and the zero-inflated negative binomial (ZINB) regression model [74–79]. Seabright et al. [76] fitted zero-inflated generalized linear multilevel models to investigate the potential impacts of post-marital residence patterns on the size of women's social groups, as well as their access to alloparental childcare among Tsimane forager–farmers in lowland Bolivia. Rorie et al. (2022) [77] presented a ZINB mixed effect model to identify predictors' effects on adolescent victimization. Niedhammer et al. [78] examined the association between psychosocial work factors and sickness absence in 31 Europe countries. They used the multilevel negative binomial hurdle model to study the count of sickness absence with three hierarchical levels and excessive zeros. Using the zero-inflated negative binomial model with random effects, Forst et al. [79] modeled the association between factors, including the occupational category, demographic characteristics, and work-related injury counts for repeated measures within zip codes. This study provided evidence to support the potential benefits of community-based approaches for reducing the burden of workplace injuries and promoting occupational health and safety.

Our study employs multilevel and zero-inflated models as valuable tools for understanding the factors associated with H-2A labor violations. Leveraging a unique dataset, our analysis contributes to the existing literature on the factors that are correlated with H-2A labor violations. In addition, multilevel modeling is particularly advantageous for our research, as it accounts for varying predictor effects across categories, such as state or industry. This observed variation in our data allows for a more accurate assessment of the relationship between these factors and H-2A violation counts compared to previous methods. Furthermore, we incorporate a zero-inflated model in our study to address the presence of structural zeros in the data. This is essential for our analysis, as there are likely two distinct mechanisms generating detected H-2A violations, one of which does not produce detected violations greater than zero count. The zero-inflated model accurately accounts for these structural zeros, offering a more comprehensive understanding of the factors that are correlated with H-2A labor violations.

Data

The data used in this analysis comes from multiple sources, which are described in detail in this section. We also discuss the data cleaning process and feature selection to identify the most important factors used in the regression analysis.

Data collection

This research utilizes data from multiple sources to examine the correlation between various factors and H-2A visa program violations, which could help inform policymakers and inspection agencies in detecting violations in the agricultural industry. To analyze labor violations in the agricultural sector, we used data from the Wage and Hour Division (WHD) of the U.S. DOL covering the period from 2010 to 2020 [80]. Specifically, cases with a "findings start date" and "findings end date" between 2010 and 2020 were included. The dataset includes information on multiple types of violations, such as H-2A, H-1A, H-1B, OSHA, and FLSA violations, per case, investigated by DOL. Investigations are initiated in multiple ways. In addition to proactively targeting industries known for high violations or where vulnerable workers and low-wage challenges are prevalent, the WHD receives confidential complaints that initiate investigations [81]. Given this context, this study aims to explore the correlation between various factors and the H-2A violation count to provide insights that could enhance the precision of these investigations. The dataset includes 175,126 cases investigated by DOL in a wide range of industries. Given our research's focus on labor violations in the agricultural sector, we relied on a specific subset of data related to the count of H-2A violations per case. Specifically, the data was filtered to include only those cases with North American Industry Classification System (NAICS) codes that began with 11, corresponding to agriculture, forestry, fishing, and hunting industries. This filtering ensured that the analysis focused specifically on labor violations in the agricultural sector where H-2A violations may occur and resulted in 12,041 remaining cases. We next excluded H-2A violation counts for Puerto Rico and the District of Columbia from our analysis as our study focused on the 50 U.S. states. Additionally, finding data on relevant factors for these regions was challenging. This further reduced the number of cases from 12,041 to 11,976. Of the 11,976 cases within the 50 U.S. states with NAICS code starting with "11", 2,523 were found to have greater than zero H-2A violations reported.

However, it is important to note that for the cases with zero H-2A violations reported, it is unclear whether the DOL investigated for H-2A violations and found no violation occurred or whether DOL focused on investigating other types of labor violations and therefore was not looking for H-2A violations. This presented a challenge for the analyses, and we explain in detail how we addressed it in the methodology section. Additionally, some cases were found to have multiple H-2A violations. Thus, the 2,523 cases with non-zero H-2A violations had a total of 73,781 H-2A violations.

H-2A violations were found in all 50 states and 17 of the 19 agricultural industries over the 10 year period analyzed (see Figs 1 and 2, respectively). Florida, California, and Nevada clearly show more H-2A violations than other states. However, it is worth noting that factors such as the number of H-2A workers employed, number of farms, climate, and population of these states may contribute to the higher number of violations and that Fig 1 provides the count of H-2A violations identified without normalizing for these factors. Indeed, when accounting for the number of H-2A workers certified in each state (see Fig 3), Nevada continues to display more violations than other states, while Florida and California have a more moderate number of violations per H-2A worker certified. New Mexico, Kansas, and Alaska also have a substantially higher number of normalized H-2A violations than most states (Fig 3), although they have a low H-2A Violation Count (Fig 1). Note that the dataset for the number of H-2A
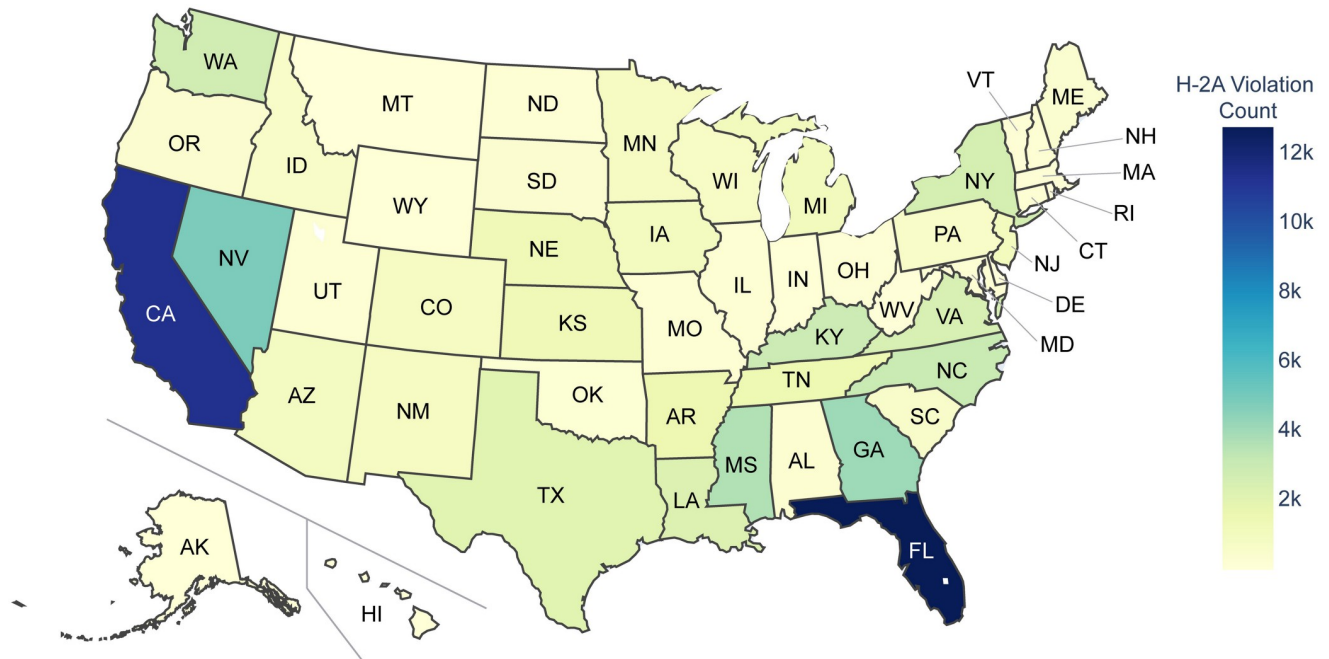
Fig 1. Geographical distribution of H-2A violations. The heatmap illustrates the total number of identified H-2A violations within the agricultural sectors per state from 2010–2020.

https://doi.org/10.1371/journal.pone.0302960.g001

workers certified in each state does not reliably report the sector of the business seeking certification, so Fig 3 focuses on H-2A violations and certifications for all industries. Since the H-2A visa is an agricultural visa, a vast majority of the identified H-2A violations are within the agricultural sector (i.e., NAICS code beginning with '11'). Therefore, it is not surprising that the
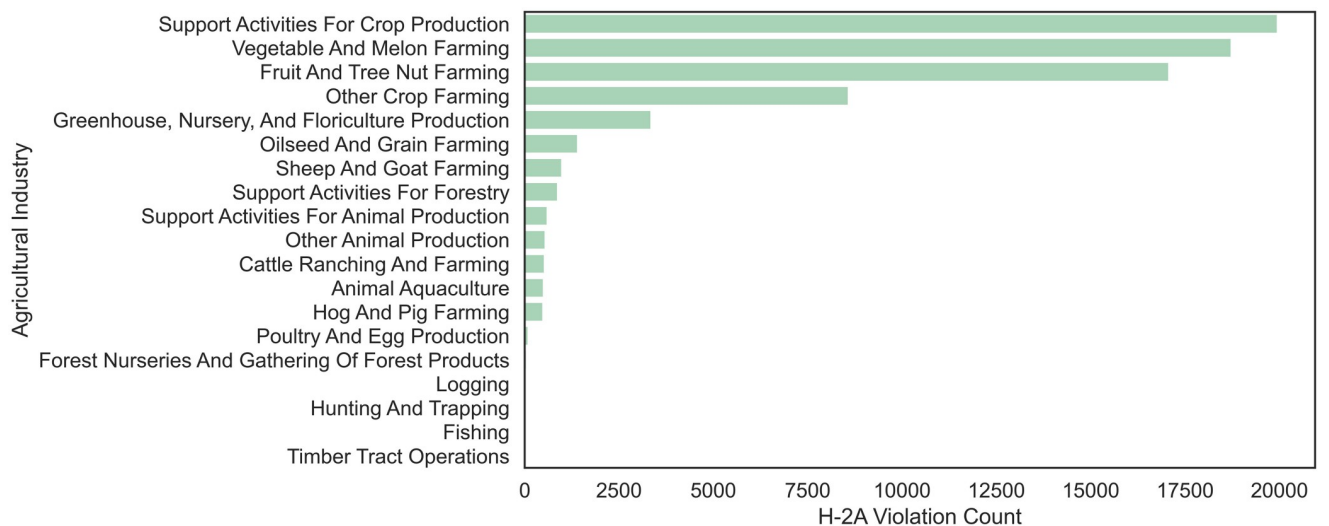

Fig 2. H-2A violations by industry. The bar plot shows the distribution of H-2A violation counts across agricultural industries, highlighting the industries with the most violations from 2010–2020.
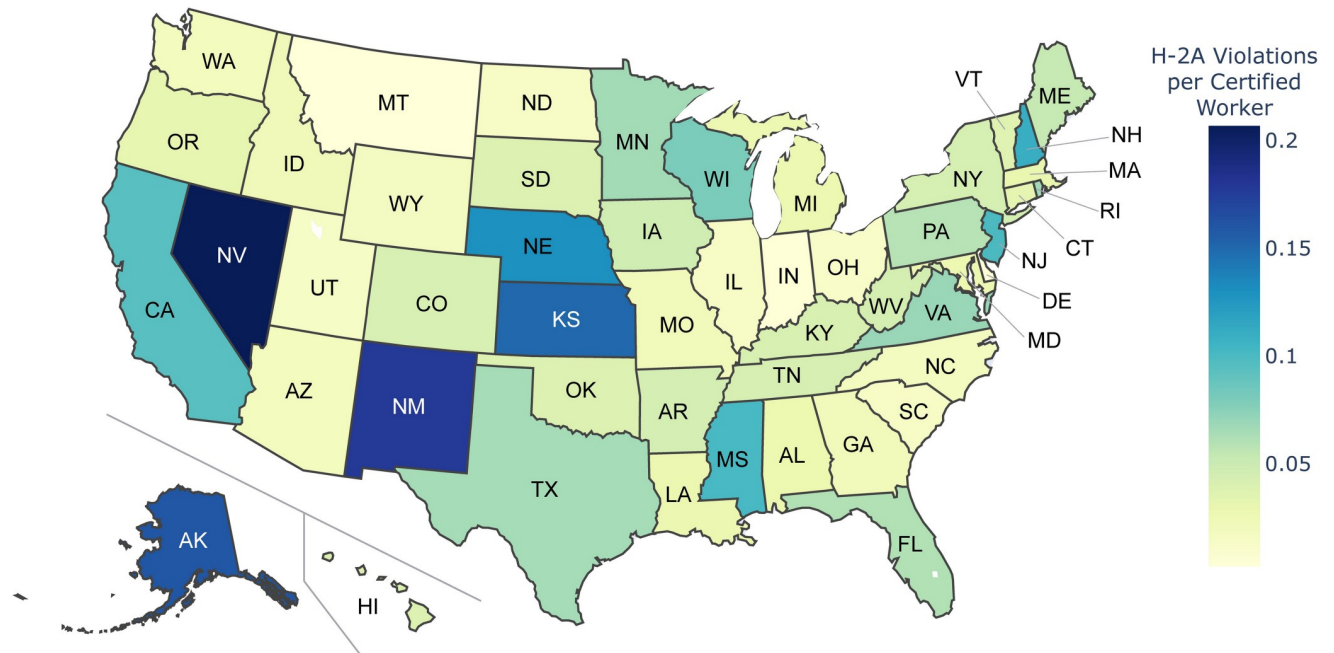
https://doi.org/10.1371/journal.pone.0302960.g002

Fig 3. H-2A violations per H-2A worker certified. The heatmap illustrates the total number of identified H-2A violations (across all industries) divided by the number of H-2A workers certified in each state from 2010–2020.

https://doi.org/10.1371/journal.pone.0302960.g003

distribution of H-2A violations identified within the agricultural sector (Fig 1) closely resembles the distribution of H-2A violations identified in all sectors (see S1 Fig).

The bar chart reveals that substantially more H-2A violations have been identified in the Support Activities for Crop Production (19,950), Vegetable and Melon Farming (18,727), and Fruit and Tree Nut Farming (17,073) industries than in other agricultural industries; some industries have had very few H-2A violations identified, such as Poultry and Egg Production (96), Forest Nurseries and Gathering of Forest Products (52), Logging (14), and Hunting and Trapping (2). No H-2A Violations were identified in the Fishing and Timer Tract Operations industries during the 10 year period. This is not particularly surprising, as the notable disparity in the incidence of H-2A violations among distinct agricultural sectors can be attributed, to some extent, to the level of H-2A employment within each industry. The combined analysis highlights geographical and industry-specific H-2A violation trends.

We combined WHD data with external data from multiple sources to form our consolidated dataset. The external data consists of various factors that we categorized into three groups: agriculture land, inspection, and population (see Table 1). Several studies in the Literature Review Section that discussed the potential factors related to the vulnerability of farm workers motivated us to study their impacts on the number of H-2A violations. In the following paragraphs, we describe the factors considered in the analysis. A summary of the factors is provided in Table 2.

Average size of farms (acres).   Each state's average acreage size of farms was collected from the United States Department of Agriculture (USDA) database from the National Agricultural Statistics Service (NASS) [82] using the `Average farm size (acres)` field. The data represent the size of any place from which a thousand dollars or more of agricultural products were produced, sold, or normally would have been sold during the fiscal year. We collected data from 2010 to 2020. This dataset contained information for the number of farms,

Table 1. Data frame of factors categorized into three groups.

Category	Factor
Agriculture Land	Average Size of Farms (acres)
	Number of Farms
	Number of Small/Medium Establishments
	Labor Intensity
	Average Net Income per Farm
Inspection	Number of Task Forces
	Length of Alleged Violations
Population	Poverty Rate
	Fatal Injury Rate
	Minimum Standard Wages

https://doi.org/10.1371/journal.pone.0302960.t001

land in farms (acres), and average farm size per state annually. The average farm size ranged from 55 (Rhode Island) to 2745 (Wyoming) acres.

Number of farms.   The number of farms in each state from 2010 to 2020 was obtained from the NASS USDA database `Number of farms` field [82]. The number of farms per state ranged from 680 farms (Alaska, 2010–2012) to 248,500 farms (Texas, 2013 & 2017), with an average number of 41,653 farms per state.

Number of small/medium establishments (by number of employees).   The number of small/medium establishments was collected from the Quarterly Census of Employment and Wages (QCEW) database from the U.S. Bureau of Labor Statistics [83]. The data represents the employment data from companies in the U.S. by type of industry (NAICS code) and state. It represents the number of establishments recorded quarterly in each year. The data available is only for the number of farms counted in the first quarter of each year. It might exclude H2A workers depending on the state labor regulation, which is a limitation. Although this data is expected to exclude around 20% of the agricultural workers, it is considered a representative dataset for operation in the agricultural sector from 2010 to 2020 used in USDA reports [7]. We define small/medium farms as those with less than 20 employees. The justification for this threshold is based on regulations on labor rights. For example, under California law, organizations with less than 26 employers abide by less strict labor regulations, being allowed to pay a lower minimum wage [84]. The U.S. Federal legislation also exempts the minimum wage for certain organizations. Specifically, organizations with an annual gross volume of sales or

Table 2. Description of factors.

Factor	Source	Type	State/Industry
Average Size of Farms	USDA	integer	state
Number of Farms	USDA	integer	state
Number of Small/Med Establishments	QCEW	integer	state
Labor Intensity	USDA, [7]	binary	industry
Average Net Income per Farm	USDA	continuous	state
Number of Task Forces	OJP	integer	state
Length of Alleged Violations	WHD	integer	state and industry
Poverty Rate	USCB	continuous	state
Fatal Injury Rate	BLSIIF	continuous	state
Minimum Standard Wages	DOL	continuous	state

https://doi.org/10.1371/journal.pone.0302960.t002

business done of at least $500,000 and specific organizations with fewer than nine employees are exempt from this requirement [85]. Additionally, employees who work for companies with fewer than ten employees do not have the right to receive unemployment insurance [86]. The dataset includes different categories for firm sizes, such as less than 5 employees, 5 to 9 employees, 10 to 19 employees, 20 to 49 employees, etc. As a result, we combined the groups with fewer than 20 employees to represent the small/medium establishment variable, as small/medium-sized firms with less than 20 employees have been defined in a previous study [42].

Labor intensity.   To study the effects of labor intensity on labor violation counts, we classified each type of crop in the NAICS code as either high-intensity (1) or low-intensity manual labor (0). A labor-intensive crop is a relative measure of the number of hours of human labor required to produce the same yield in dollars in comparison to other crops [7]. It was difficult to establish a definitive threshold to determine which crops require high-intensity labor due to factors such as technological advancements and changes in farming practices over time [87]. Therefore, we used the same classification Castillo et al. [7] used in a 2021 USDA report. According to their findings, high-intensive crops usually employ more H-2A workers than non-high-intensive ones [7]. To help improve the classification, we also used other papers to corroborate the classification of which NAICS code is associated with high-intensity labor [16, 87].

Average net income.   Net farm income was extracted from the U.S. Department of Agriculture (USDA), Economic Research Service, Farm Income and Wealth Statistics [88]. This dataset includes the `net farm income` per state by year for 2010—2020. To find the average net income per farm, the net farm income is divided by the number of farms per state each year. The number of farms is gathered by USDA's reports of farms and land in farms [82]. The average yearly net income varied between -$21,369.12 to $263,372.17 across states.

Number of task forces.   Due to the variety of human trafficking task forces throughout the U.S. and a lack of data related to how many task forces exist in each state, we focused on the number of Enhanced Collaborative Model task forces funded through the Office of Justice Programs Office for Victims of Crime (OVC) that aim to combat human trafficking [89]. These task forces include victims, law enforcement, social service providers, and other governmental and non-governmental partners to help trafficking victims and provide them with appropriate services. Enhanced Collaborative Model task forces have specific provisions focused on labor trafficking, which may provide an indicator of law enforcement awareness and readiness to identify labor trafficking and associated exploitation. This data consists of information about the funded task forces, award amount, award status (e.g., open, closed, etc.), etc. To transform the data for regression analysis, we counted the unique number of active OVC Enchanced Collaborative Model task forces per by year. The number of task forces ranged from 0 to 15 (California, 2019), with an average of 1.002 task forces in a state each year.

Length of alleged violations.   We calculated the length of the alleged violations per case investigated in the WHD dataset by taking the difference of the `findings_end_date` and the `findings_start_date` of the alleged violations [80]. This time period indicates the estimated duration of violations that occurred per case. Only cases related to agricultural businesses (filtered by NAICS codes that begin with 11) during 2010—2020 were included in our analysis. To determine the state associated with the business in each case, we used the `st_cd` field in the WHD dataset.

Poverty rate.   We obtained poverty statistics by state from The U.S. Census Bureau (USCB) historical poverty tables [90]. Specifically, we used data from the `Percent in Poverty` field in *Table 19. The number of Poor and Poverty Rate by State: 1980 to 2021* for the years 2010 to 2020. Although we study labor violations in agriculture, due to the lack of available data related to poverty rates specifically for people employed in the agricultural

sector, the poverty rate data we used includes households working in agricultural and non-agricultural jobs. The data ranges from a minimum poverty rate of 0.9% (New Hampshire, 2009–2010) to 25.8% (North Dakota, 2012).

Fatal injury rate. As the lack of tracking systems makes detecting and understanding trafficked workers' health and risk patterns difficult, we extract data from the U.S. Bureau of Labor Statistics, Injuries, Illness, and Fatalities (BLSIIF) dataset [19], which compares risk among worker groups with varying employment levels. Due to incomplete data on nonfatal injuries during our specified timeframe, we relied on the fatal injury rate from 2010—2020 for our analysis. The data ranges from a minimum fatal injury rate of 0.9% (New Hampshire, 2010) to 17.7% (North Dakota, 2012).

Minimum standard wages. We used the DOL wage data for changes in basic minimum wages under state law from 2010 to 2020 obtained by the U.S. DOL Division of Fair Labor Standards Act and Child Labor Wage and Hour Division [91]. DOL wage data has some exceptions related to minimum wage values for specific states. To maintain consistency and uniformity in data analysis, we addressed issues by taking the following steps: First, data includes minimum wage values for some states that are lower than the federal minimum wage due to some exceptions, such as a low number of employees. We decided to use the federal minimum wage for states rather than the lowest wage a state can legally pay if all exceptions are met. Second, five states do not have a state minimum wage, which leaves them subject to federal law. We used the federal minimum wage of $7.25 for these states for our analysis [92]. Third, some states have a range of numbers for the minimum wage. In these cases, we decided to conservatively use the higher value since, in some states, this may indicate the state's initiative to increase wages within their legal framework.

Data pre-processing and feature selection

In this section, we first applied feature scaling to standardize all regressors, as they were initially on different scales. This process ensured that no particular predictor would dominate the analysis due to its scale. To explore the dataset further and identify correlations between regressors or factors, we utilized a correlation matrix, which provided valuable insights into the relationships among the variables.

To check for multicollinearity, we employed the variance inflation factor (VIF) to assess the degree of independence among the factors. We discovered that some factors, such as the number of farms, number of small/medium establishments, and average net income per farm, exhibited high VIF values. To select the most relevant features, we employed Pearson correlation analysis and found that the number of small/medium establishments had the strongest correlation with the target variable, outperforming the other two factors. Additionally, we leveraged Feature Importance from the random forest regression model and used recursive feature elimination with k-fold cross-validation to determine the most influential factors for our model. These methods, including the correlation matrix and random forest feature importance, indicated that the number of small/medium establishments should be considered as a factor in our model.

Lastly, we had to select between the number of farms and the average net income per farm based on the VIF results. Since the average net income per farm was derived from the number of farms, we decided to retain the average net income variable, as it implicitly represents the number of farms. This choice allowed us to avoid potential multicollinearity issues while still incorporating valuable information in our analysis.

While no single solution perfectly addresses collinearity, the model selection procedures we used aimed to alleviate the effects of collinearity. This approach focuses on including only

variables that have the strongest association with the response variable, selected from a group of correlated predictors. All data cleaning and pre-processing tasks were performed using Python 3.7, ensuring a consistent and robust approach to preparing the data for subsequent analysis. This streamlined process allowed us to efficiently identify and incorporate the most relevant factors into our model, ultimately yielding more accurate and meaningful results.

Method

In this section, we will explain the modeling approach used to establish a relationship between the predictors and the H-2A violation count, which is the target variable. We begin by exploring various regression models suitable for count data, where many observations have zero counts. Subsequently, we will implement a multilevel model for the grouped data with two levels, state and industry.

Regression models for count data

The dependent variable is the number of H-2A violations detected by DOL for each distinct WHD investigation, which we refer to as a "case". This count denotes the number of workers identified within a single investigation who faced violations. It can only take non-negative integer values and experiences excess zeros. These types of count data are modeled with General linear models (GLMs) and General linear mixed models (GLMMs) using either negative binomial (NB) or Poisson distributions [93, 94].

We first investigated whether the Poisson distribution would be an appropriate fit for our data by conducting a dispersion test. Upon examining the results, we found that the p-value was statistically significant, indicating the presence of overdispersion in the count data. This overdispersion suggests that the variance is greater than the mean of the outcome variable, which is not ideal for a Poisson distribution. We decided to consider the NB regression model since this model is more adept at handling overdispersed count data, making it a more suitable choice compared to the Poisson regression model.

It is common to use a zero-inflated model when there are more zeros in count data than a simple model predicts. To check whether such zero-inflated models are needed, we conduct a model comparison between the simple models and zero-inflated models. Zero-inflated general linear models are a mixture of distributions, one that has degenerate zero counts (logit)—a zero-inflated model—and another that has degenerate integer counts following Poisson and Negative binomial (NB) distributions—a conditional model [95]. We compared Zero-inflated Negative Binomial (ZINB), and NB models fitted to the dataset using the Akaike information criterion (AIC), Bayesian information criterion (BIC), and likelihood ratio test. AIC and BIC serve as measures for assessing the goodness of fit of statistical models while considering their complexity. Both criteria aim to find the optimal balance between fitting the data well and avoiding overfitting due to unnecessary complexity. Lower values of AIC and BIC signify a better fit of the model to the data, as they indicate a more favorable balance between data fit and model complexity. The formula used to compute AIC is $AIC = -2log(L) + k$, where L is the maximum value of the likelihood function for the model and k is the number of estimated parameters in the model [96]. The BIC penalizes the model complexity more strongly than the AIC, and both are useful for model selection, depending on the trade-off between goodness of fit and model complexity. As shown in Table 3, the ZINB model has a lower AIC and BIC value compared to the NB model, indicating the zero-inflated component of the model effectively captures the remaining overdispersion that the NB model does not entirely explain. In addition to AIC and BIC, we used the likelihood ratio test to compare the two models. A likelihood ratio test is a validation approach that compares a simple model, such as the NB model,

Table 3. Comparison of NB and ZINB Models.

Model	AIC	BIC
NB Model	30543.4	30624.7
ZINB Model	28804.4	28959.6

with a special case of that model that includes the zero-inflated part, such as the ZINB model. According to the test, additional parameters in the ZINB model improve the model's fit significantly compared to the NB model, confirming the conclusion of the AIC and BIC comparison. Therefore, we used ZINB regression for the number of H-2A violations that exhibit overdispersion and excess zeros.

The fitted ZINB model reflects two possible outcomes for H-2A violations reported in this dataset: zero H-2A violations or a positive number of H-2A violations. Some cases reporting zero H-2A violations reflect structural zeros, while others reflect non-structural (random) zeros. Structural zeros encompass cases in which a zero is reported in the dataset because either (a) DOL investigated other (non-H-2A) types of violations; since they were not looking for H-2A violations, no violations were found or reported, or (b) the organization's structure is such that H-2A violations are not possible, such as the organization does not hire H-2A workers. In investigation cases where DOL is inspecting for H-2A violations and the organization's structure is such that a H-2A violation could potentially occur, the reported number of H-2A violations can be zero or greater than zero and are generated using the NB distribution portion of the model. Zero counts in these cases are non-structural (random) zeros and refer to situations in which the organization could potentially have had violations, but no violation was found because even though DOL investigated for H-2A violations, (a) the organization did not have a violation during the period in which they were investigated (i.e., true negative) or (b) DOL could not detect the H-2A violations that were occurring during the inspection (low detectability) (i.e., false negative).

We assumed that structural zeros occur with probability Φ. Therefore, the probability mass function (pmf) for a ZINB can be formulated as follows [97]:

$$\Pr(Y = y) = \begin{cases} \Phi + (1 - \Phi)g(Y = 0), & \text{if } y = 0 \\ (1 - \Phi)g(Y = y), & \text{if } y = \mathbb{N}^+, \end{cases} \tag{1}$$

where the random variable $Y \in \mathbb{N}$ is the count data, y is the realization of Y, and $g(Y)$ is the pmf of the negative binomial distribution. The distribution is formulated in terms of the mean (μ) and dispersion parameter ($\alpha = \frac{1}{r}$), where r is the predefined number of successes that occur according to the definition of the NB distribution.

$$g(Y = y) = Pr(y|\mu, \alpha) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(y + 1)\Gamma(\alpha^{-1})} \left(\frac{1}{1 + \alpha\mu}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu}{1 + \alpha\mu}\right)^{y} \tag{2}$$

The negative binomial components are used to estimate the intercept and coefficients of the regression model fitted to the dataset, thereby allowing us to accurately capture the relationships between the variables and make meaningful inferences based on the fitted model.

Multilevel generalized linear models

We explored using ZINB linear models to analyze H-2A violation counts to study the effects of different factors. These factors were explained in detail in the Data section. Our data
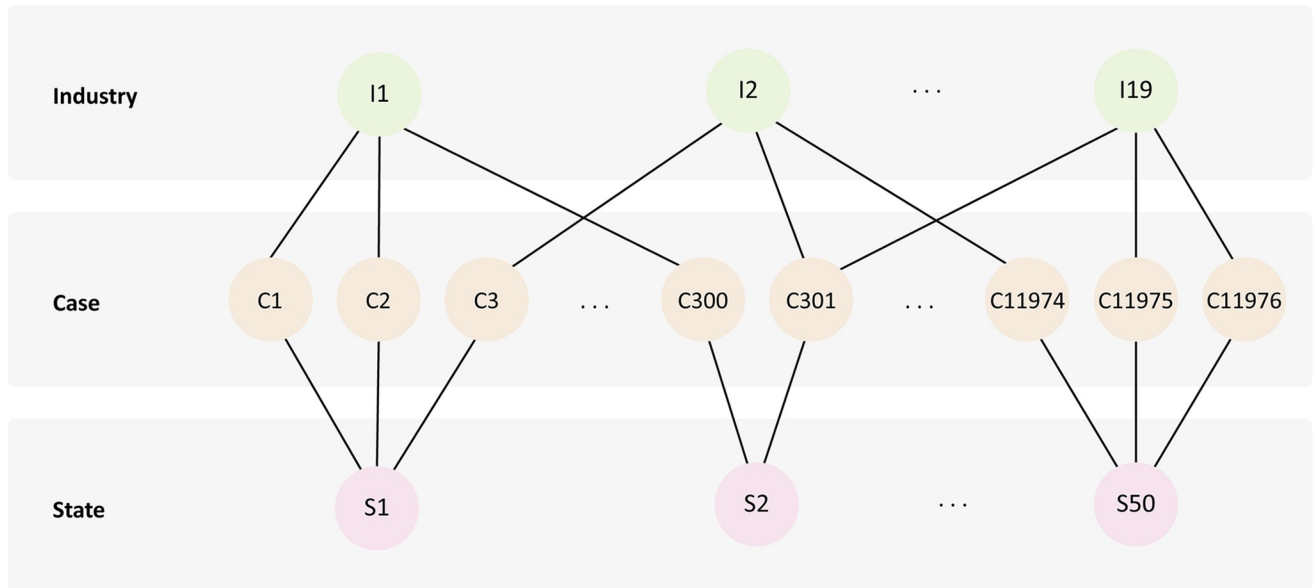
Fig 4. Cross-classified data. The multilevel diagram illustrates the cross-classification of two levels of study.

https://doi.org/10.1371/journal.pone.0302960.g004

exploration revealed the presence of multiple groups in the dataset, including state and industry. However, we also found that these groups are not hierarchically nested within each other, but instead, they cross [98]. This can be presented schematically for different cases contained within the cross-classification of 50 states by 19 industries, as in Fig 4. Cases are at the individual level (level 1), and the two higher levels are industry and state (level 2).

Based on the structure of our data, we employed a multilevel zero-inflated negative binomial linear model (MZINBLM) with two crossed random effects—state and industry—in both the conditional and ZI parts of the model. We compared the performance of this model with other alternatives, such as the zero-inflated negative binomial linear model (ZINBLM) without accounting for state effect or industry effect, MZINBLM with only one level (accounting for either state effect or industry effect but not both), and MZINBLM with two levels (accounting both state and industry) in the conditional model. We evaluated the models by comparing AIC, BIC, and likelihood ratio tests. The findings reveal that the MZINBLM with two crossed random effects in both parts of the model (e.g., conditional and ZI) provided a superior fit to the data as compared to the other models considered, see Table 4.

Let $Y_{i(jk)}$ represent the count outcome for the i-th case, belonging to the j-th industry group and k-th state group, $\mu_{i(jk)}$ be the conditional mean of $Y_{i(jk)}$, and $X_{i1}, \ldots, X_{ip}$ be the p predictor variables at the case level. The MZINBLM formulation is as follows [99]:

Table 4. Model selection comparative analysis using AIC, BIC, and log-likelihood measures.

Model	Random Effects (Cond Model)	Random Effects (ZI Model)	AIC	BIC	logLik
NB	-	-	30543.4	30624.7	-15260.7
ZINB	-	-	28804.4	28959.6	-14381.2
MZINB1	state	-	28545.2	28707.8	-14250.6
MZINB2	state & industry	-	28423.6	28593.6	-14188.8
MZINB3	state & industry	state	28031.6	28209	-13991.8
MZINB4	state & industry	state & industry	27812.5	27997.3	-13881.3

https://doi.org/10.1371/journal.pone.0302960.t004

Level One:

$$\log\left[\frac{\Phi_{i(jk)}}{1 - \Phi_{i(jk)}}\right] = a_{i(jk)} + a_1 X_{i1} + \cdots + a_p X_{ip} \tag{3}$$

$$\log(\mu_{i(jk)}) = \beta_{i(jk)} + \beta_1 X_{i1} + \cdots + \beta_p X_{ip} + \epsilon_i \tag{4}$$

$$\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2) \tag{5}$$

Level Two:

$$a_{i(jk)} = a_0 + w_j^{(1)} + w_k^{(2)} \text{for} \quad i = 1, 2, \ldots, 11976 \tag{6}$$

$$\beta_{i(jk)} = \beta_0 + u_j^{(1)} + u_k^{(2)} \quad \text{for} \quad i = 1, 2, \ldots, 11976 \tag{7}$$

$$u_j^{(1)} \sim \mathcal{N}(0, \sigma_{u^{(1)}}^2) \tag{8}$$

$$u_k^{(2)} \sim \mathcal{N}(0, \sigma_{u^{(2)}}^2) \tag{9}$$

$$w_j^{(1)} \sim \mathcal{N}(0, \sigma_{w^{(1)}}^2) \tag{10}$$

$$w_k^{(2)} \sim \mathcal{N}(0, \sigma_{w^{(2)}}^2) \tag{11}$$

Where $a_0, \ldots, a_p$ and $\beta_0, \ldots, \beta_p$ are the corresponding regression coefficients for the logit model and conditional model, respectively; $w_j^{(1)}$, $u_j^{(1)}$ and $w_k^{(2)}$, $u_k^{(2)}$ are the random effects of intercept for state and industry respectively, and ϵ_i is the residual error. The variance of the random effects are denoted by $\sigma_{u^{(1)}}^2, \sigma_{u^{(2)}}^2, \sigma_{w^{(1)}}^2$, and $\sigma_{w^{(2)}}^2$.

The results of this study were obtained using R 4.2.2 with glmmTMB package for multilevel zero-inflated generalized linear models. It includes the conditional model, which is the negative binomial distribution, and the zero-inflated model [100].

Results

As per the methodology, several models were fit, and their performance was evaluated based on AIC, BIC, and likelihood tests. The best model was selected by comparing the models using these criteria, considering the balance between the goodness of fit and model complexity. We used a multilevel zero-inflated generalized linear model to analyze the count of H-2A violations, taking into account fixed effects such as the average size of farms, number of small/medium establishments, labor intensity, average net income per farm, number of task forces, length of alleged violations, poverty rate, fatal injury rate, and minimum standard wages, as well as crossed random effects from state and industry.

The selected model identified the significant factors associated with the count outcome. Table 5 presents the results of the selected MZINBLM, where the first part pertains to the conditional model. This part describes the correlation between the independent variables and the logarithm of the count data, which can be interpreted as the change in the expected H-2A violation counts associated with a unit increase in the corresponding variable while holding all other variables constant. The second part corresponds to the zero-inflation model. This

Table 5. The multilevel zero-inflated negative binomial linear model.

Factor	Estimate	Pr($>$ |z|)
Conditional Model		
Intercept	2.407	$<$ 2e-16***
Average size of farms	0.122	0.006**
Number of small/med establishments	1.369	0.000***
Labor intensity	-0.030	0.921
Average net income per farm	-0.129	0.185
Number of task forces	-0.017	0.559
Length of alleged violations	0.123	0.307
Poverty rate	0.253	0.008**
Fatal injury rate	-0.109	0.169
Minimum standard wages	-0.082	0.159
Zero-inflated Model		
Intercept	2.829	0.000***
Average size of farms	-0.125	0.247
Number of small/med establishments	-0.215	0.821
Labor intensity	-2.386	0.039*
Average net income per farm	0.025	0.838
Number of task forces	-0.066	0.026*
Length of alleged violations	-1.585	$<$ 2e-16***
Poverty rate	1.065	$<$ 2e-16***
Fatal injury rate	-0.725	2.74e-07***
Minimum standard wages	-0.441	1.43e-09***

Significance levels: 0 ***, 0.001 **, 0.01 *

includes log odds of count outcome (estimate) for predicting excess zeros and their corresponding p-values.

The results indicate that the intercept, average size of farms, the number of small/medium establishments, and the poverty rate have a significant correlation with the mean H-2A violation counts in the conditional model. Additionally, the zero-inflated model suggests several factors are significantly associated with the outcome, including labor intensity, number of task forces, length of alleged violations, poverty rate, fatal injury rate, and minimum standard wages. Notably, the average net income per farm is not a significant factor, indicating no significant relationship between this variable and the count outcome.

Upon analyzing the regression results, we noted a positive correlation between the significant factors and the target variable in the conditional model. However, it is essential to note that correlation does not necessarily indicate causation. This relationship may not imply causality due to potential confounding factors, reverse causality, or random variation in the data. Therefore, interpreting these results requires caution, and further research with rigorous study designs and control for confounding factors is needed to establish a causal relationship between these variables. Specifically, a one-unit increase in average farm size correlates with $e^{0.122}$ = 1.13 increase in the mean H-2A violation counts. Furthermore, the positive coefficient for the number of small/medium establishments suggests that DOL identified more H-2A violations in states or industries with higher numbers of small/medium establishments (less than 20 employees). The poverty rate is significant in conditional and ZI models. It illustrates that states with higher poverty rates have higher mean H-2A violation counts. According to the ZI

model, an increase in the poverty rate is associated with a higher likelihood of structural zero, where no violations are discovered in organizations that are not at risk of H-2A violations. Therefore, it is more likely that zero counts for states with high poverty rates happen due to one of two reasons. First, DOL investigated other violations (not H-2A). Second, the organization does not hire H-2A workers, making it impossible for such organizations to have H-2A violations.

The ZI model also highlights other significant factors such as labor intensity, the number of task forces, the length of alleged violations, the fatal injury rate, and minimum standard wages. The ZI model presents a negative correlation between these predictors and the log of odds ($log\left(\frac{\Phi}{1-\Phi}\right)$), indicating that increasing one of these predictors is associated with less likelihood of excessive zero (structural zero). The states with higher labor intensity, task forces, length of alleged violations, fatal injury rates, and minimum standard wages are more likely to have non-structural zeros. Zero H-2A violation counts in these states could happen for two reasons. Although they are at risk of violations, DOL could not detect them, or no H-2A violations occurred during the investigation period.

Our analysis reveals that while specific predictors are consistently correlated with H-2A violations across all areas, the actual levels of these violations vary significantly from one state or industry to another. Each state and industry starts from a different baseline regarding violation frequency. The differing intercepts are significant; they indicate that factors unique to each state and industry also play a crucial role in H-2A violations beyond the predictors we have analyzed. Understanding these unique baselines is essential for enhancing our understanding of the issue. It is important for inspectors to consider these findings within a broader context, combining them with other relevant data to ensure equitable and well-rounded decision-making processes in resource allocation and inspection strategies.

Discussion

This paper presented a multilevel zero-inflated negative binomial regression model to handle cross-classified data with excessive zeros, particularly regarding labor violations in agriculture. We employed a multilevel regression model to investigate the association between different factors and the frequency of H-2A labor violations. The study emphasizes the need for a flexible labor inspection strategy that considers the location and type of agricultural industry.

The ZI model revealed several essential factors correlating with the probability of excessive zeros. Our analysis of DOL's investigations showed that the reasons for zero H-2A violation counts vary depending on the location and the type of agricultural industry. The ZI model's insights suggest that zero counts in industries with higher labor intensity are more likely to occur either because the DOL could not detect the violations or no violation occurred during the investigation. Similar relationships can be observed for other factors, such as the number of task forces, length of alleged violations, fatal injury rate, and minimum standard wages. This study offers valuable insights for labor inspectors and agencies involved in H-2A labor inspections. These findings can guide them in making better-informed decisions regarding resource allocation, encompassing aspects such as budget, the number of inspectors, and inspection time. By doing so, they can enhance the effectiveness and efficiency of resource distribution, particularly in areas with a higher risk of H-2A violations.

This study has several limitations that should be considered. While the regression analysis identified correlations between certain factors and H-2A violation counts, it did not establish causal relationships. Even though establishing causative relationships would be more valuable from a policy perspective, this study is intended as a foundation for such future research. Therefore, future studies should focus on investigating these causal effects, drawing from the

insights of this study. Additionally, the absence of comprehensive data related to farm workers, including health issues, access to healthcare, income, housing conditions, and demographic information, limited our dataset construction. As a result, we relied on available data sources. This constraint suggests that our findings should be interpreted carefully, and additional, more extensive research is necessary to gain a deeper understanding of H-2A violations in the agricultural sector.

It is also important to acknowledge that while the DOL plays a role in identifying violations, there is still room for improvement in effectively uncovering abuses. There are numerous reasons why H-2A workers might not report abuses or concerns, and previous research has consistently highlighted the need for more proactive efforts to uncover such abuses. Furthermore, the WHD data we obtained did not clarify whether the DOL conducted H-2A violation inspections in cases with zero counts. This leaves the possibility of undetected or unreported cases, which may have led to an underestimation of H-2A violations in specific states and industries. We attempted to address this issue by using a zero-inflated model to account for excessive zeros. It is also important to note that there are multiple types of rules and regulations H-2A employers are expected to adhere to (including rules related to housing standards, work hours, wage rates, etc.) [101]. However, it is not clear whether non-adherence to all of these rules and regulations results in the type of H-2A violations documented in the WHD dataset. Furthermore, the dataset lacks detailed categorization that would distinguish between the severity of violations, ranging from minor to severe violations. This limitation highlights the need for more detailed data to inform future research and the development of more effective investigation strategies by identifying areas where compliance efforts could be most impactful.

Furthermore, the role of labor contractors is significant within the H-2A program and warrants further exploration. Our current analysis could not delve into this aspect due to the limited detail in the WHD's violation data. This limitation highlights an important avenue for future research, dependent on the availability of more detailed and specific data. However, this highlights the need for improved data collection strategies to better track labor inspections and accurately assess the prevalence of H-2A labor violations.

In addition to the insights that can be gained in this study, the use of H-2A violation data may also provide critical insights into labor trafficking, which is significantly underidentified. Though the data is not perfect, it shows the potential for administrative data related to workplace safety and to help inform our understanding of places and situations where workers may be vulnerable to abuses and are in need of protection.

Conclusion

In conclusion, our study provides essential insights for labor enforcement agencies in shaping specific H-2A violation inspection strategies for each state and agricultural sector. By employing a zero-inflated negative binomial model, this study offers unique insights into the distribution of detected H-2A violations across different states and agricultural sectors. This approach allows us to identify not only states and sectors with high numbers of reported violations but also those with unexpectedly low or zero counts. It is essential to recognize that the absence of H-2A violations in certain states or industries does not necessarily imply that no violations have occurred. For instance, in regions or industries with specific characteristics (e.g., higher labor intensity, number of task forces, length of alleged violations, fatal injury rate, and minimum standard wages), there is a potential for violations to either remain undetected or not occur during official investigations. Additionally, our results highlight the importance of considering factors such as the average size of farms, the number of small/medium establishments, and the poverty rate in planning inspections. Based on our research, states or industries falling

within the upper range of these categories have more detected H-2A violations. By providing insights into the state and agricultural sector features correlated to where H-2A violations are currently being detected, our findings may inform future WHD labor inspection strategies regarding where to allocate limited inspection resources.

Supporting information

S1 Fig. Geographical overview of H-2A violation counts in all sectors. The heatmap illustrates the total number of identified H-2A violations per state from 2010—2020. Note that these violations span various industries, not just those with NAICS codes beginning with "11". (TIF)

S1 File.
(DOCX)

Acknowledgments

We thank Aubrey Sneesby, Chase Childress, Cao Jainfei, Elaine Klatt, Joseph Pak, and Margaret Clark for their thoughtful comments and assistance related to gathering data, data cleaning, and methodology.

Author Contributions

Conceptualization: Arezoo Jafari, Shawn Bhimani, Amy Farrell, Kayse Lee Maass.

Data curation: Arezoo Jafari, Priscila De Azevedo Drummond, Dominic Nishigaya.

Formal analysis: Arezoo Jafari.

Funding acquisition: Shawn Bhimani, Amy Farrell, Kayse Lee Maass.

Investigation: Arezoo Jafari, Priscila De Azevedo Drummond, Dominic Nishigaya.

Methodology: Arezoo Jafari, Shawn Bhimani, Aidong Adam Ding, Amy Farrell, Kayse Lee Maass.

Project administration: Kayse Lee Maass.

Resources: Kayse Lee Maass.

Software: Arezoo Jafari.

Supervision: Shawn Bhimani, Amy Farrell, Kayse Lee Maass.

Validation: Arezoo Jafari, Amy Farrell, Kayse Lee Maass.

Visualization: Arezoo Jafari.

Writing – original draft: Arezoo Jafari.

Writing – review & editing: Arezoo Jafari, Priscila De Azevedo Drummond, Shawn Bhimani, Dominic Nishigaya, Aidong Adam Ding, Amy Farrell, Kayse Lee Maass.

References

1. Martin P, Rutledge Z. Proposed changes to the H-2A program would affect labor costs in the United States and California. California Agriculture. 2022; 75(3):135–141. https://doi.org/10.3733/ca.2021a0020

2. Parajuli R, Thoma G, Matlock MD. Environmental sustainability of fruit and vegetable production supply chains in the face of climate change: A review. Science of The Total Environment. 2019; 650:2863–2879. https://doi.org/10.1016/j.scitotenv.2018.10.019 PMID: 30373063

3. Gollin D, Parente S, Rogerson R. The Role of Agriculture in Development. American Economic Review. 2002; 92(2):160–164. https://doi.org/10.1257/000282802320189177

4. Martin PL. Migrant Workers in Commercial Agriculture; 2016. Available from: http://www.ilo.org/global/topics/labour-migration/publications/WCMS_538710/lang--en/index.htm.

5. Martin PL, Mamer J. The Farm Labor Market. Proceedings of the Academy of Political Science. 1982; 34(3):223–234. https://doi.org/10.2307/1173742

6. Mandeel EW. The Bracero Program 1942-1964. American International Journal of Contemporary Research. 2014; 4:171–184.

7. Castillo M, Simnitt S, Astill G, Minor T. Examining the Growth in Seasonal Agricultural H-2A Labor. U. S. Department of Agriculture, Economic Research Service; 2021. 226.

8. Bruno A. H-2A and H-2B Temporary Worker Visas: Policy and Related Issues. United States of America: TRAC: Transactional Records Access Clearinghouse; 2020.

9. Castillo M, Martin P, Rutledge Z. The H-2A Temporary Agricultural Worker Program in 2020. U.S. Department of Agriculture, Economic Research Service; 2022.

10. H-2A Guest Workers Fact Sheet; 2020. Available from: https://www.ncfh.org/h-2a-guest-workers-fact-sheet.html.

11. Mizelle CR. Temporary Changes to Requirements Affecting H-2A Nonimmigrants due to the COVID-19 National Emergency: Extension of Certain Flexibilities. U.S. Department of Homeland Security; 2020. Available from: https://www.federalregister.gov/documents/2020/12/18/2020-27661/temporary-changes-to-requirements-affecting-h-2a-nonimmigrants-due-to-the-covid-19-national.

12. Owen C, Dank M, Breaux J, Banuelos I, Farrell A, Pfeffer R, et al. Understanding the Organization, Operation, and Victimization Process of Labor Trafficking in the United States. Trends in Organized Crime. 2015; 18(4):348–354. https://doi.org/10.1007/s12117-015-9257-9

13. Barrick K, Lattimore PK, Pitts WJ, Zhang SX. When farmworkers and advocates see trafficking but law enforcement does not: challenges in identifying labor trafficking in North Carolina. Crime, Law and Social Change. 2014; 61(2):205–214. https://doi.org/10.1007/s10611-013-9509-z

14. Guild A, Figueroa I. The Neighbors Who Feed Us: Farmworkers and Government Policy—Challenges and Solutions. Harvard Law & Policy Review. 2018; 13:157.

15. Robinson E, Nguyen HT, Isom S, Quandt SA, Grzywacz JG, Chen H, et al. Wages, Wage Violations, and Pesticide Safety Experienced by Migrant Farmworkers in North Carolina. NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy. 2011; 21(2):251–268. https://doi.org/10.2190/NS.21.2.h PMID: 21733804

16. Costa D, Martin P, Rutledge Z. Federal Labor Standards Enforcement in Agriculture: Data Reveal the Biggest Violators and Raise New Questions about How to Improve and Target Efforts to Protect Farmworkers. Bulgaria: Economic Policy Institute; 2020.

17. Liebman AK, Augustave W. Agricultural Health and Safety: Incorporating the Worker Perspective. Journal of Agromedicine. 2010; 15(3):192–199. https://doi.org/10.1080/1059924X.2010.486333 PMID: 20665305

18. Martin P. The Prosperity Paradox: Fewer and More Vulnerable Farm Workers. Critical Frontiers of Theory, Research, and Policy in International Development Studies. Oxford University Press; 2020.

19. Census of Fatal Occupational Injuries. Fatal Injury Rates by State, All Ownerships, 2007-21; 2022. https://www.bls.gov/iif/state-data/fatal-injury-rates-by-state-2007-2021.htm.

20. CENTRO DE LOS DERECHOS DEL MIGRANTE, I. Ripe for Reform: Abuse of Agricultural Workers in the H-2A Visa Program. (2020), https://cdmigrante.org/ripe-for-reform/

21. Zhang K, Arauz RF, Chen TH, Cooper SP. Heat Effects among Migrant and Seasonal Farmworkers: A Case Study in Colorado. Occupational and Environmental Medicine. 2016; 73(5):324–328. https://doi.org/10.1136/oemed-2015-103332 PMID: 26872487

22. Mac VVT, McCauley LA. Farmworker Vulnerability to Heat Hazards: A Conceptual Framework. Journal of Nursing Scholarship. 2017; 49(6):617–624. https://doi.org/10.1111/jnu.12327 PMID: 28806486

23. Vega-Arroyo AJ, Mitchell DC, Castro JR, Armitage TL, Tancredi DJ, Bennett DH, et al. Impacts of Weather, Work Rate, Hydration, and Clothing in Heat-related Illness in California Farmworkers. American Journal of Industrial Medicine. 2019; 62(12):1038–1046. https://doi.org/10.1002/ajim.22973 PMID: 30964208

24. Stoklosa H, Kunzler N, Ma ZB, Luna JCJ, de Vedia GM, Erickson TB. Pesticide Exposure and Heat Exhaustion in a Migrant Agricultural Worker: A Case of Labor Trafficking. Annals of Emergency Medicine. 2020; 76(2):215–218. https://doi.org/10.1016/j.annemergmed.2020.03.007 PMID: 32362432

25. Arcury TA, Summers P, Talton JW, Nguyen HT, Chen H, Quandt SA. Job Characteristics and Work Safety Climate Among North Carolina Farmworkers With H-2A Visas. Journal of Agromedicine. 2015; 20(1):64–76. https://doi.org/10.1080/1059924X.2014.976732 PMID: 25635744

26. Ronda-Pérez E, Moen BE. Labour Trafficking: Challenges and Opportunities from an Occupational Health Perspective. PLOS Medicine. 2017; 14(11):e1002440. https://doi.org/10.1371/journal.pmed.1002440 PMID: 29166395

27. Pratt SG, Hard DL. Injury Risk Factors Associated with AgriculturalWorkplace Fatalities. Journal of agricultural safety and health. 1998;(1):29–38. https://doi.org/10.13031/2013.15372

28. Pickett W, Hartling L, Brison RJ, Guernsey JR, Surveillance PCAI. Fatal Work-Related Farm Injuries in Canada, 1991-1995. CMAJ. 1999; 160(13):1843–1848. PMID: 10405669

29. Horsburgh S, Feyer AM, Langley JD. Fatal Work Related Injuries in Agricultural Production and Services to Agriculture Sectors of New Zealand, 1985–94. Occupational and Environmental Medicine. 2001; 58(8):489–495. https://doi.org/10.1136/oem.58.8.489 PMID: 11452042

30. Rivara FP. Fatal and Nonfatal Farm Injuries to Children and Adolescents in the United States. Pediatrics. 1985; 76(4):567–573. https://doi.org/10.1542/peds.76.4.567 PMID: 4047800

31. Luckstead J, Nayga RM Jr, Snell HA. Labor Issues in the Food Supply Chain Amid the COVID-19 Pandemic. Applied Economic Perspectives and Policy. 2021; 43(1):382–400. https://doi.org/10.1002/aepp.13090 PMID: 33173571

32. Hiott AE, Grzywacz JG, Davis SW, Quandt SA, Arcury TA. Migrant Farmworker Stress: Mental Health Implications. The Journal of Rural Health. 2008; 24(1):32–39. https://doi.org/10.1111/j.1748-0361.2008.00134.x PMID: 18257868

33. Ramos AK, Carlo G, Grant K, Trinidad N, Correa A. Stress, Depression, and Occupational Injury among Migrant Farmworkers in Nebraska. Safety. 2016; 2(4):23. https://doi.org/10.3390/safety2040023 PMID: 29644237

34. Earle-Richardson G, Jenkins PL, Slingerland DT, Mason C, Miles M, May JJ. Occupational Injury and Illness among Migrant and Seasonal Farmworkers in New York State and Pennsylvania, 1997–1999: Pilot Study of a New Surveillance Method. American Journal of Industrial Medicine. 2003; 44(1):37–45. https://doi.org/10.1002/ajim.10233 PMID: 12822134

35. Lindert J, von Ehrenstein OS, Priebe S, Mielck A, Brähler E. Depression and Anxiety in Labor Migrants and Refugees – A Systematic Review and Meta-Analysis. Social Science & Medicine. 2009; 69 (2):246–257. https://doi.org/10.1016/j.socscimed.2009.04.032 PMID: 19539414

36. Marinescu I, Qiu Y, Sojourner A. Wage Inequality and Labor Rights Violations. National Bureau of Economic Research; 2021. 28475.

37. Nagurney A. Attracting International Migrant Labor: Investment Optimization to Alleviate Supply Chain Labor Shortages. Operations Research Perspectives. 2022; 9:100233. https://doi.org/10.1016/j.orp.2022.100233

38. Schwarz C, Alvord D, Daley D, Ramaswamy M, Rauscher E, Britton H. The Trafficking Continuum: Service Providers' Perspectives on Vulnerability, Exploitation, and Trafficking. Affilia. 2019; 34 (1):116–132. https://doi.org/10.1177/0886109918803648

39. Norwood JS. Labor Exploitation of Migrant Farmworkers: Risks for Human Trafficking. Journal of Human Trafficking. 2020; 6(2):209–220. https://doi.org/10.1080/23322705.2020.1690111

40. Silverman SJ, Hari A. Troubling the Fields: Choice, Consent, and Coercion of Canada's Seasonal Agricultural Workers. International Migration. 2016; 54(5):91–104. https://doi.org/10.1111/imig.12266

41. Siqueira CE, Gaydos M, Monforton C, Slatin C, Borkowski L, Dooley P, et al. Effects of Social, Economic, and Labor Policies on Occupational Health Disparities. American Journal of Industrial Medicine. 2014; 57(5):557–572. https://doi.org/10.1002/ajim.22186 PMID: 23606055

42. Bhorat H, Kanbur R, Mayet N. Estimating the Causal Effect of Enforcement on Minimum Wage Compliance: The Case of South Africa. Review of Development Economics. 2012; 16(4):608–623. https://doi.org/10.1111/rode.12007

43. Pierre G, Scarpetta S. Do Firms Make Greater Use of Training and Temporary Employment When Labor Adjustment Costs Are High? IZA Journal of Labor Policy. 2013; 2(1):15. https://doi.org/10.1186/2193-9004-2-15

44. Harrison JL, Getz C. Farm size and job quality: mixed-methods studies of hired farm work in California and Wisconsin. Agriculture and Human Values. 2015; 32(4):617–634. https://doi.org/10.1007/s10460-014-9575-6

45. Barrick K, Lattimore PK, Pitts WJ, Zhang SX. Labor Trafficking Victimization among Farmworkers in North Carolina: Role of Demographic Characteristics and Acculturation. International Journal of Rural Criminology. 2014; p. 225–243. https://doi.org/10.18061/1811/61593

46. Zhang SX, Spiller MW, Finch BK, Qin Y. Estimating Labor Trafficking among Unauthorized Migrant Workers in San Diego. The Annals of the American Academy of Political and Social Science. 2014; 653(1):65–86. https://doi.org/10.1177/0002716213519237

47. Richards K. The Trafficking of Migrant Workers: What Are the Links Between Labour Trafficking and Corruption? International Migration. 2004; 42(5):147–168. https://doi.org/10.1111/j.0020-7985.2004.00305.x

48. Farrell A, Bright K, de Vries I, Pfeffer R, Dank M. Policing Labor Trafficking in the United States. Trends in Organized Crime. 2020; 23(1):36–56.

49. Zahniser S, Taylor JE, Hertz T, Charlton D. Farm Labor Markets in the United States and Mexico Pose Challenges for U.S. Agriculture. U.S. Department of Agriculture, Economic Research Service; 2018. EIB-201.

50. Moseley C. Working Conditions in Labor-Intensive Forestry Jobs in Oregon. Ecosystem Workforce Program, Institute for a Sustainable Environment, University of Oregon; 2006.

51. Preibisch K, Otero G. Does Citizenship Status Matter in Canadian Agriculture? Workplace Health and Safety for Migrant and Immigrant Laborers. Rural Sociology. 2014; 79(2):174–199. https://doi.org/10.1111/ruso.12043

52. Locke RM, Samel H. Beyond the Workplace: "Upstream" Business Practices and Labor Standards in the Global Electronics Industry. Studies in Comparative International Development. 2018; 53(1):1–24. https://doi.org/10.1007/s12116-017-9258-6

53. Sari LK, Wardana LO. Modeling of Child Labour Exploitation Status in Indonesia Using Multilevel Binary Logistic Regression. ITM Web of Conferences. 2021; 36:01008. https://doi.org/10.1051/itmconf/20213601008

54. H-2A Adverse Effect Wage Rates (AEWRs) | Flag.dol.gov;. Available from: https://flag.dol.gov/wage-data/adverse-effect-wage-rates.

55. Clemens J, Strain MR. Understanding "Wage Theft": Evasion and avoidance responses to minimum wage increases. Labour Economics. 2022; 79:102285. https://doi.org/10.1016/j.labeco.2022.102285

56. Yaniv G. Minimum Wage Noncompliance and the Employment Decision. Journal of Labor Economics. 2001; 19(3):596–603. https://doi.org/10.1086/322074

57. Garson GD. Multilevel Modeling: Applications in STATA®, IBM® SPSS®, SAS®, R, & HLMTM. SAGE Publications; 2019.

58. Khine MS. Methodology for Multilevel Modeling in Educational Research: Concepts and Applications. Springer Nature; 2022.

59. Shirilla P, Solid C, Graham SE. The Benefits of Longitudinal Data and Multilevel Modeling to Measure Change in Adventure Education Research. Journal of Experiential Education. 2022; 45(1):88–109. https://doi.org/10.1177/10538259211027595

60. Leyland AH, Groenewegen PP. Multilevel Modelling for Public Health and Health Services Research: Health in Context. Springer Nature; 2020.

61. Seidu AA, Darteh EKM, Agbaglo E, Dadzie LK, Ahinkorah BO, Ameyaw EK, et al. Barriers to Accessing Healthcare among Women in Ghana: A Multilevel Modelling. BMC Public Health. 2020; 20(1):1916. https://doi.org/10.1186/s12889-020-10017-8 PMID: 33334326

62. Paul A, Lang JWB, Baumgartner RJ. A Multilevel Approach for Assessing Business Strategies on Climate Change. Journal of Cleaner Production. 2017; 160:50–70. https://doi.org/10.1016/j.jclepro.2017.04.030

63. Echavarren JM, Balžekienė A, Telešienė A. Multilevel Analysis of Climate Change Risk Perception in Europe: Natural Hazards, Political Contexts and Mediating Individual Effects. Safety Science. 2019; 120:813–823. https://doi.org/10.1016/j.ssci.2019.08.024

64. Uthman OA, Moradi T, Lawoko S. The Independent Contribution of Individual-, Neighbourhood-, and Country-Level Socioeconomic Position on Attitudes towards Intimate Partner Violence against Women in Sub-Saharan Africa: A Multilevel Model of Direct and Moderating Effects. Social Science & Medicine. 2009; 68(10):1801–1809. https://doi.org/10.1016/j.socscimed.2009.02.045 PMID: 19303687

65. Kiss L, Schraiber LB, Heise L, Zimmerman C, Gouveia N, Watts C. Gender-Based Violence and Socioeconomic Inequalities: Does Living in More Deprived Neighbourhoods Increase Women's Risk of Intimate Partner Violence? Social Science & Medicine. 2012; 74(8):1172–1179. https://doi.org/10.1016/j.socscimed.2011.11.033 PMID: 22361088

66. Jason LA, Salina D, Ram D. Oxford Recovery Housing: Length of Stay Correlated with Improved Outcomes for Women Previously Involved with the Criminal Justice System. Substance Abuse. 2016; 37(1):248–254. https://doi.org/10.1080/08897077.2015.1037946 PMID: 25962090

67. Augustyn MB, Fulco CJ, Agkebe D, Henry KL. Maternal Substance Abuse and the Mother-Child Relationship in Late Childhood. Substance Use & Misuse. 2021; 56(14):2106–2114. https://doi.org/10.1080/10826084.2021.1972313 PMID: 34514950

68. Leclerc B, Smallbone S, Wortley R. Prevention Nearby: The Influence of the Presence of a Potential Guardian on the Severity of Child Sexual Abuse. Sexual Abuse. 2015; 27(2):189–204. https://doi.org/10.1177/1079063213504594 PMID: 24145397

69. Martínez-Schuldt RD, Martínez DE. Immigrant Sanctuary Policies and Crime-Reporting Behavior: A Multilevel Analysis of Reports of Crime Victimization to Law Enforcement, 1980 to 2004. American Sociological Review. 2021; 86(1):154–185. https://doi.org/10.1177/0003122420978406

70. MacDonald JM, Lattimore PK. Count Models in Criminology. In: Piquero AR, Weisburd D, editors. Handbook of Quantitative Criminology. New York, NY: Springer; 2010. p. 683–698.

71. Huebner BM. Administrative Determinants of Inmate Violence: A Multilevel Analysis. Journal of Criminal Justice. 2003; 31(2):107–117. https://doi.org/10.1016/S0047-2352(02)00218-0

72. Marchment Z, Frith MJ, Morrison J, Gill P. A Multi-Level Analysis of Risky Streets and Neighbourhoods for Dissident Republican Violence in Belfast. ISPRS International Journal of Geo-Information. 2021; 10(11):765. https://doi.org/10.3390/ijgi10110765

73. Shaaban AN, Peleteiro B, Martins MRO. Statistical Models for Analyzing Count Data: Predictors of Length of Stay among HIV Patients in Portugal Using a Multilevel Model. BMC Health Services Research. 2021; 21(1):372. https://doi.org/10.1186/s12913-021-06389-1 PMID: 33882911

74. Almasi A, Eshraghian MR, Moghimbeigi A, Rahimi A, Mohammad K, Fallahigilan S. Multilevel Zero-Inflated Generalized Poisson Regression Modeling for Dispersed Correlated Count Data. Statistical Methodology. 2016; 30:1–14. https://doi.org/10.1016/j.stamet.2015.11.001

75. Hidano A, Gates MC. Why Sold, Not Culled? Analysing Farm and Animal Characteristics Associated with Livestock Selling Practices. Preventive Veterinary Medicine. 2019; 166:65–77. https://doi.org/10.1016/j.prevetmed.2019.03.005 PMID: 30935507

76. Seabright E, Alami S, Kraft TS, Davis H, Caldwell AE, Hooper P, et al. Repercussions of Patrilocal Residence on Mothers' Social Support Networks among Tsimane Forager–Farmers. Philosophical Transactions of the Royal Society B: Biological Sciences. 2022; 378(1868):20210442. https://doi.org/10.1098/rstb.2021.0442 PMID: 36440570

77. Rorie M, Park Sm, West MP. Do Structural Choice Theory and the 'Risky Lifestyles' Perspectives Predict Immunity as Well as Victimization? A Test Using Zero-Inflated Mixed-Effect SEM Analyses of Adolescent Victimizations in South Korea. International Review of Victimology. 2022; 28(3):345–366. https://doi.org/10.1177/02697580211057330

78. Niedhammer I, Chastang JF, Sultan-Taïeb H, Vermeylen G, Parent-Thirion A. Psychosocial Work Factors and Sickness Absence in 31 Countries in Europe. European Journal of Public Health. 2013; 23(4):622–629. https://doi.org/10.1093/eurpub/cks124 PMID: 23002241

79. Forst L, Friedman L, Chin B, Madigan D. Spatial Clustering of Occupational Injuries in Communities. American Journal of Public Health. 2015; 105(Suppl 3):S526–S533. https://doi.org/10.2105/AJPH.2015.302595 PMID: 25905838

80. Wage and Hour Compliance Action Data;. Available from: https://enforcedata.dol.gov/views/data_catalogs.php.

81. Fact Sheet #44: Visits to Employers;. Available from: http://www.dol.gov/agencies/whd/fact-sheets/44-flsa-visits-to-employers.

82. Farms and Land in Farms | ID: 5712m6524 | USDA Economics, Statistics and Market Information System;. https://usda.library.cornell.edu/concern/publications/5712m6524.

83. Jobe S. Quarterly Census of Employment and Wages U.S. Bureau of Labor Statistics;. https://www.bls.gov/cew/downloadable-data-files.htm.

84. Minimum Wage Frequently Asked Questions;. Available from: https://www.dir.ca.gov/dlse/faq_minimumwage.htm.

85. U S Department of Labor. Fair Labor Standards Act Advisor;. https://webapps.dol.gov/elaws/whd/flsa/screen75.asp.

86. Unemployment Insurance Tax Topic, Employment & Training Administration (ETA)—U.S. Department of Labor;. https://oui.doleta.gov/unemploy/uitaxtopic.asp.

87. Huffman WE. Trends, Adjustments, and Demographics, and Income of Agricultural Workers. Review of Agricultural Economics. 2005; 27(3):351–360. https://doi.org/10.1111/j.1467-9353.2005.00229.x

88. Farm Sector Financial Indicators, State Rankings;. Available from: https://data.ers.usda.gov/reports.aspx?ID=17839#P10674dbd9d2b433997310b0cf3cb570e_2_185iT0R0x16.

89. Task Force Funding;. https://data.ojp.usdoj.gov/Funding-API/Funding/a7v8-ei2f.

90. Bureau UC. Historical Poverty Tables: People and Families—1959 to 2021;. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html.

91. Changes in Basic Minimum Wages in Non-Farm Employment Under State Law: Selected Years 1968 to 2022;. http://www.dol.gov/agencies/whd/state/minimum-wage/history.

92. State Minimum Wages; 2022. https://www.ncsl.org/labor-and-employment/state-minimum-wages.

93. Mwalili SM, Lesaffre E, Declerck D. The Zero-Inflated Negative Binomial Regression Model with Correction for Misclassification: An Example in Caries Research. Statistical Methods in Medical Research. 2008; 17(2):123–139. https://doi.org/10.1177/0962280206071840 PMID: 17698937

94. Perumean-Chaney SE, Morgan C, McDowall D, Aban I. Zero-Inflated and Overdispersed: What's One to Do? Journal of Statistical Computation and Simulation. 2013; 83(9):1671–1683. https://doi.org/10.1080/00949655.2012.668550

95. Lambert D. Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing. Technometrics. 1992; 34(1):1–14. https://doi.org/10.2307/1269547

96. Cavanaugh JE, Neath AA. The Akaike Information Criterion: Background, Derivation, Properties, Application, Interpretation, and Refinements. WIREs Computational Statistics. 2019; 11(3):e1460. https://doi.org/10.1002/wics.1460

97. Young DS, Roemmele ES, Yeh P. Zero-Inflated Modeling Part I: Traditional Zero-Inflated Count Regression Models, Their Applications, and Computational Tools. WIREs Computational Statistics. 2022; 14(1):e1541. https://doi.org/10.1002/wics.1541

98. de Leeuw J, Meijer E, editors. Handbook of Multilevel Analysis. 1st ed. New York, NY: Springer; 2008.

99. Fielding A, Goldstein H. Cross-classified and Multiple Membership Structures in Multilevel Models:. vol. 69. DfES; 2006.

100. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB Balances Speed and Flexibility among Packages for Zero-Inflated Generalized Linear Mixed Modeling. The R journal. 2017; 9(2):378–400. https://doi.org/10.32614/RJ-2017-066

101. WHD Fact Sheets;. Available from: http://www.dol.gov/agencies/whd/fact-sheets.