QSS20 Social Impact Practicum: Investigating DOL Staffing and H2A Violations

Chad Lasseter, Bess Neiblum, Alexis Reaves, and Nicole Schwartz
Quantitative Social Science

chad.h.lasseter.22@dartmouth.edu
bess.m.neiblum.23@dartmouth.edu
alexis.d.reaves.20@dartmouth.edu
nicole.a.schwartz.22@dartmouth.edu

Abstract

Our primary object of study was the relationship between H2A violation counts, associated penalties, and trends in enforcement agency staffing at the state level. Using data from the Office of Personnel Management (OPM) as well as the Wage and Hourly Division (WHD) of the Department of Labor (DOL), we ran regressions on employment levels and variables such as back-wages per case, civil monetary penalties per case, proportion of cases resulting in back-wages, and proportion of cases resulting in civil monetary penalties. At the national level, we found that there was little correlation between number of employees and penalties per case. While there was greater correlation between number of employees and percentage of cases that result in a penalty, this trend is not consistent with past data and must be taken with a grain of salt. At the state level, agency employment is declining in all states of interest save for Alabama. Furthermore, this decline is associated with a lower percentage of cases that end in penalties as well as a lower amount of fines per case. Given that the existing literature posits that staffing has an impact of enforcement, our findings may be reason for concern. Hopefully the results of this study produces actionable insights that can help improve the lives of migrant workers.

1 Introduction

Foreign workers have historically been an extremely important part of the United States economy, as immigrants throughout the centuries have worked hard to contribute to the country they call home. Following this tradition, the H2A program was initiated by the United States DOL to fulfill the labor demands of United States farms who could not find enough agricultural workers within the country. The program allows employers to hire foreign workers on temporary visas if they can prove that they are facing a labor shortage and can also certify that they maintain health and safety standards. Some of the employer requirements include providing housing, transportation, payment at least twice a month, a copy of the work contract, and a pre-determined number of hours for their guest employees. Additionally, firms must provide workers' compensations for injuries taking place on the job and may not pass the cost of obtaining the H2A certification onto the employee (DOL.gov).

However, due to their foreign status and temporary visas issued by the program, migrant workers face numerous challenges. According to their DOL H2A contracts, they are bound to their employers and cannot change jobs like most workers can in a competitive job market. They also cannot bring their families with them and have no hope of obtaining citizenship at the end of the program. All of this puts guest workers at a great risk of abuse by their employers, who do not face a great incentive to abide by all of the DOL's laws regarding the H2A program. Even though there are extensive protections for migrant farm workers, they are enforced to such a small extent that the protections do not really hold much weight for employers. Abuses by agricultural employers include wage theft, unsafe housing and work conditions, and other violations of their rights, and in extreme cases, neglect by agricultural employers has led to the death of some migrant workers.

Government agencies like the DOL work to ensure the rights of migrant workers by investigating and penalizing employers who commit violations, yet the WHD of the DOL may not have the capacity to investigate all of these violations properly or at all. Staffing may represent an important factor in the launching and effectiveness of these investigations, especially given that the WHD is severely understaffed and underfunded, according to a report by the Economic Policy Institute or EPI (Costa, Martin, and Rutledge 2020).

In order to test if this were the case, our team decided to focus on WHD employment data and H2A violation case data to investigate possible correlations between agency employment levels and H2A violations across states. Our project may inform policy changes that will result in better enforcement of the rights of migrant workers, who represent a vulnerable population.

2 Related work

A report by the Southern Poverty Law Center (SPLC) detailed several of the abuses by agricultural employers that occur under the H2A guest worker program. Their work was based on legal cases and interviews. While the first version of their report in 2007 proposed a reforming of the program, the updated 2013 version concluded that this program is so full of corruption and abuse that it should not and cannot be reformed or expanded upon by Congress. The SPLC found that agricultural employers often do not reimburse their workers for travel costs, which can often be exorbitant. Furthermore, they exaggerate the benefits of working in the US to convince workers to enter the country, neglecting to mention the deep debt that can come from migrating. The report also discusses the severe imbalance of power between the employer and employee; because the employer has the power to decide whether the workers leaves, stays, has a job, or doesn't, the worker's fate is essentially always in the hands of the employer. This power imbalance encourages abuse by employers because they know that there is almost nothing guest workers can do about those abuses. In many cases, agricultural employers literally confiscate the legal documents of the workers and use them as a threat to prevent disobedience. Additionally, on the wage and hour side, many farmworkers make much less than the minimum wage they are promised. In short, they are often victims of wage theft. The SPLC report also touches on insufficient and unsafe housing conditions, sexual harassment, and extremely unsafe working conditions that can lead to injuries, as well as other forms of abuse by employers. As we mentioned earlier, the DOL simply does not have the resources to investigate all of these abuses and thus struggle to ensure fair treatment for farmworkers. ”The number of wage and hour investigators in the DOL declined by 14% between 1974 and 2004, and the number of completed compliance actions declined by 36%,” writes the SPLC in their report. They also found that ”In 2011, the DOL certified 7,000 employer applications for H-2A workers97 but conducted only 157 investigations into H-2A employers that same year” (Bauer 2013). These numbers clearly represent an issue, and the situation has since further deteriorated. This report provides a good basis, but the researchers use a more qualitative approach, conducting their research through legal cases and interviews with guestworkers. Our data-based approach will allow us to more clearly investigate the relationship between WHD staffing and investigation and enforcement of H2A violations.

More information comes from the EPI. A 2020 report found that "Farmworkers...are not being protected effectively by federal labor standards enforcement" and that documented violations are likely represent a mixture of under-reporting and lack of investigations (Costa et al.). Their conclusion was based on data from the DOL's WHD. Over the past two decades, the number of investigations into potential violators has steadily declined. At the same time, there has been an increase in the amount of back wages owed to farmers as well as civil monetary penalties (CMPs). Between the years of 2000 and 2019, employers were fined over $75 million dollars in employee back wages and over $60 million dollars in CMPs. Part of these penalties are directly in response to H2A violations; in 2019, an average of $485 dollars were owed per employee due to an H2A-associated fine. Overall, the amount of back wages owed to migrant farmworkers totaled $2.4 million dollars while the amount of CMPs totaled $2.8 million dollars in the same year. These fines resulted from violations such as employers inadequately paying and housing their workers. Relevant to our project, the report calculated only a 1.1% chance of a potential violator being investigated during any given year; a mere 1,200 employers are investigated out of a total of 107,000. The authors speculate that this is due to lack of WHD funding to carry out investigations, bluntly stating that "Farm employers can violate wage and hour laws and reasonably expect that those violations will never be detected" (Costa et al. 2020). By providing more research on the possible correlation between staffing trends and investigations, our study can add to the existing literature and may convince policymakers to increase funding to the WHD.

3 Data

Two main sources were used for our analysis. To acquire data about the number of employees in the WHD within the U.S. DOL, we used a collection of data sets from the U.S. OPM. This data came in the form of a series of raw text files corresponding to quarterly employee data as well as additional text files containing descriptions of codes and dictionaries for each variable in the main file. For example, we used the location file to map the location numbers in the main text file to two letter state abbreviations. Prior to 2011, the data was not reported every quarter so we focused on the data from 2011 to 2020 for consistency. The data set got down to the granularity of state data for each quarter. This was a major limitation for our analysis. We would have liked the data to include zip codes so that we could have differentiated between employees and offices on a more narrow level than just state. Given this limitation, we performed our analysis at the state-quarter level by aggregating the number of employees in each state at the point in time of each quarter from 2011 to 2020.

The second data source was Wage and Hour Compliance Action Data from the U.S. DOL. This data set includes all concluded WHD compliance actions since fiscal year 2005. Each row corresponds to a particular case ID and includes the company name, address, zip code, and NAICS code as well as information about how many violations were found, back wage amounts, number of employees due back wages, and civil money penalties assessed. For the purpose of this analysis, we focused on the violations, back wages, and CMPs related to H2A violations. We subsetted this data to cases that resulted in one or more H2A violations since investigations that result in no violations are not material. Since our OPM data ultimately contained the years from 2011 to 2020, we focused on those same years for this data as well. This will be discussed further in methods. Our unit of analysis for this data was each case. A limitation of the WHD data is the lack of a clear date for when an investigation begins.

4 Methods

The first step in the cleaning process for the OPM data was to download the quarterly DOL employment data files from the OPM website. To clean the data, we first defined a number of helper functions. One function subsetted a given dataset to only the WHD agency and to only the Wage and Hour Investigation Series occupation. There were other occupations within the WHD agency, but the prevalence with which the investigation series occupation occurred within this agency led us to conclude that this was the most important and relevant occupation and the correct one on which to focus. We posited that this occupation would be the most consequential in investigating violations. Another function added columns for the relevant year and quarter to a given dataset. And a final meta function ran both of the other functions on a given dataset. Since all of the csv files used the same naming conventions, we generated a list of these names and used them to read in the files. We fed the datasets to the final function and created a list of cleaned dataframes. We then combined those dataframes into one long dataset containing the employee data from all years and quarters since 2011.

Next, we turned to the WHD data for cleaning. After downloading the dataset from the DOL's website, we first isolated the columns of interest; namely, those involving H2A violations and associated penalties. Following this we subsetted the data to only include cases that had non-zero H2A violation counts. We eliminated cases with no violations because these investigations were not consequential since no violations were found. A WHD office that performed more zero-violation investigations is not better or more efficient than an office that does not investigate these cases. If anything, it could represent additional wasted resources. This data was then filtered to investigations occurring between the years 2012 and 2020. We did not include year 2011 because as will be discussed shortly, we incorporated a one year lag between the WHD investigation date and the employee data to account for the length of time it takes to run an investigation. Since the OPM employee data begins in 2011, the earliest investigation case should be in 2012. Finally, we added columns for the year and quarter of the findings_start_date, findings_end_date, and ld_dt columns. We ultimately selected the ld_dt column to use as our starting date of investigation for each case. As discussed above, this was a limitation of the data so this variable was our best option.

After preparing the WHD dataset, we went back to the OPM data to continue cleaning and preparing for a merge now that we had a better understanding of the WHD violations data. We planned to merge the two datasets based on state, year, and quarter, but the OPM data had number-codes for its location column, rather than two-digit state abbreviations. To recode this column, we used the "DTloc" file from the OPM as well as the state crosswalk that had been provided to us earlier in the course. This crosswalk is a dataframe containing state and country names and their abbreviations. We created a dictionary linking the location number code in the main OPM dataframe to the full state name in the DTloc file. Then the crosswalk was used to convert the full name into the 2-digit abbreviation so that it would match the st_cd column in the WHD data. The result became the state_abbrev column in our data.

Next, we prepared the OPM data to be combined with the WHD data. We needed to collapse the OPM dataset so that the number of employees was summed for each state-year-quarter combination. Since the WHD violations data was at the investigation-case level, the number of employees corresponding to the state, year, and quarter of the investigation would be added to the row. First, a "shell" dataframe of all the year, quarter, and state combinations was created. Then, the OPM data was grouped by year, quarter, and state abbreviation so that the number of employees could be summed for each state-year-quarter combination. This result was then concatenated with the shell dataframe. Any state-year-quarter combination for which there was no employee data was given a value of zero.

Finally, WHD employee information from the OPM data had to be combined with the WHD violations data. It was determined that the number of employees in the four quarters prior to the investigation start date in the WHD data would most likely have the greatest effect on the outcome of the investigation. Therefore, for every case in the WHD violations data, the number of employees (total_employment variable) was determined by taking the average of the number of employees in that state in the prior four quarters. The result was an investigation-case level dataset with the columns from the WHD data, plus a total_employment column for the average number of WHD investigators in the state during the prior four quarters. The resulting dataframe was used for our analysis.

To create the visualizations, the cleaned data was grouped by year and state (sometimes by total employment as well to fix data glitch). From here, .agg was used extensively to create summary statistics such as count, sum and mean. These stats were then displayed against each other. Additionally, a "pursued" variable was also created to demonstrate the proportion of cases that resulted in backwages or civil monetary penalties being assigned.

Lastly, a fixed effects regression was performed to regress back wages per violation and civil monetary penalties per violation on the number of employees while accounting for time effects and unobserved differences across states that are constant over time. The linearmodels package was used for this analysis.

A table of means summarizing our data is below. Observations were grouped by year.

Table 1: **Table 1: Table of Means Summary**

lddtyear	caseviolcnt	h2aviolcnt	h2abwatpamt	totalemployment
2015	38.73	35.87	7236.65	36.24
2016	23.11	20.20	6564.58	37.45
2017	26.80	23.09	7206.50	41.57
2018	28.01	23.23	5402.28	39.95
2019	32.70	27.06	4901.26	42.25
2020	26.91	22.88	7248.99	44.79

5 Results

Our initial findings were that while there was virtually no correlation between number of employees and penalties per case, there is a possible correlation between number of employees and the proportion of cases resulting in a penalty. As we continued our exploration and narrowed in on select target states, this did not always prove to be the case. In fact, some states showed some strong correlations. These are detailed in the results tables below.

Nationwide Results

Below is a general scatterplot of reported H2A violations in relation to enforcement capacity over the years. Each dot represents a state in a given year and a trend line is applied to illustrate the lack of correlation.

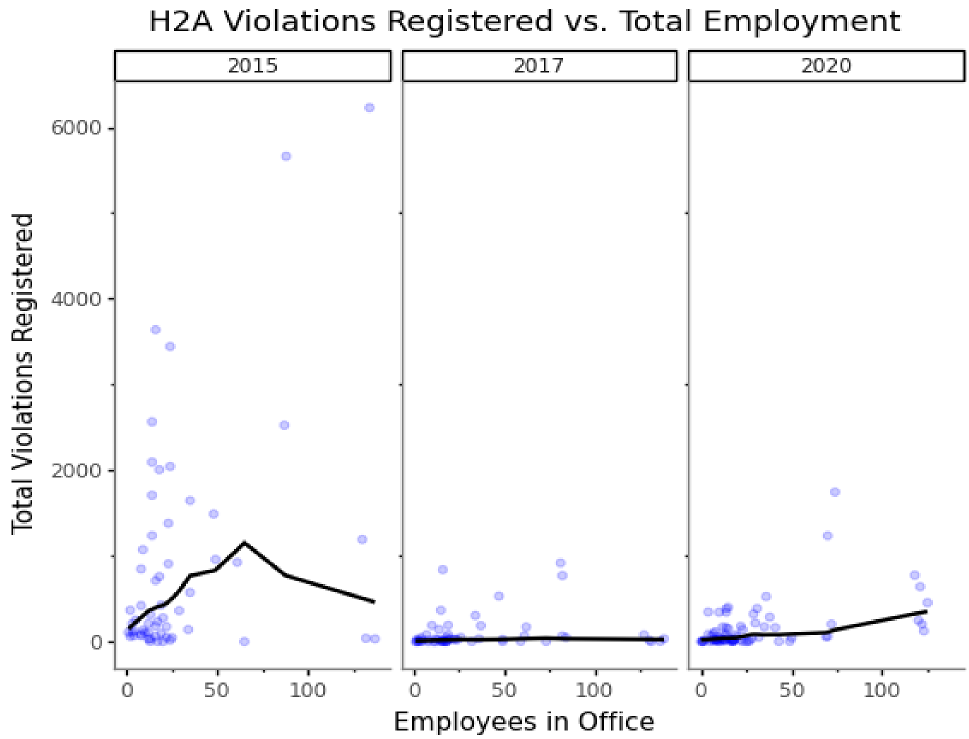


Figure 1: Total H2A Violations vs. Total Employment Above is a scatterplot of total H2A violations reported in the given year. Each dot represents a state in a given year. Total violation counts include all violations regardless of whether they resulted in a fine or penalty occurring.

The next area of focus in our analysis focused on the penalties associated with these general case counts. How much bang-for-the-buck were states getting for their enforcement personnel? To quantify this we focused on Civil Monetary Penalties and Back Wages assessed to respective violations. Below is a selection of graphs detailing these inquiries. As above, each dot represents a specific state in that select year and trend lines are applied to illustrate correlation (or in some cases, the lack thereof).

Bang for the Buck

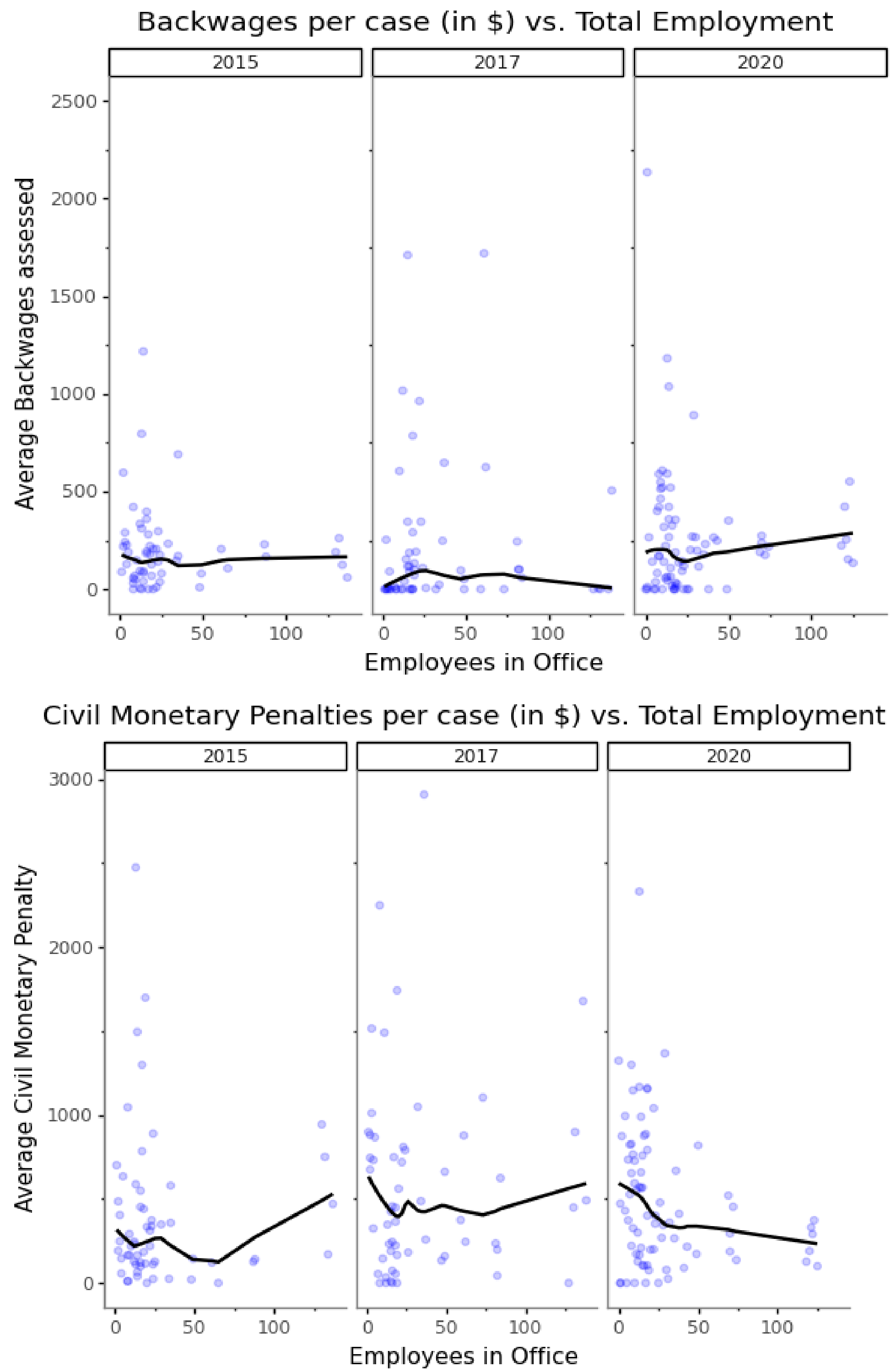
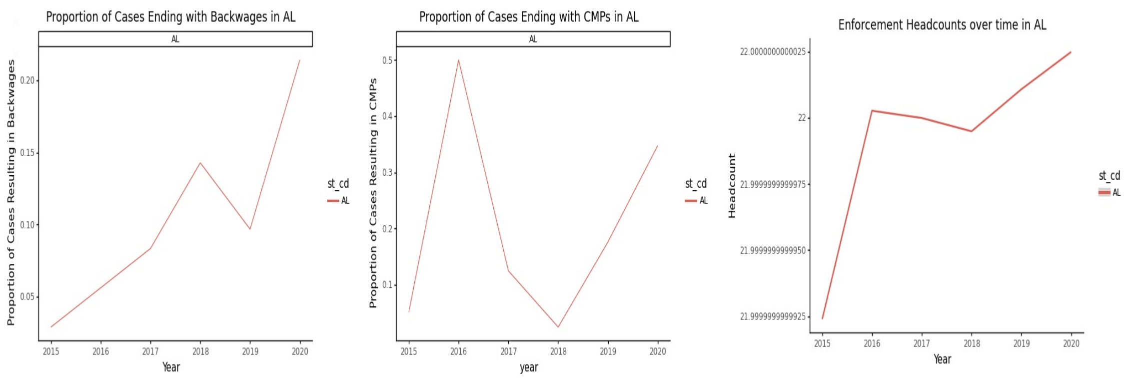


Figure 2: Proportion of Cases Resulting in CMPs vs. Total Employment and Civil Monetary Penalties per case (in $) vs. Total Employment The first set of graphs are scatterplots representing the proportion of DOL's opened cases that resulted in a civil monetary. The Y axis represents this proportion, which is total cases resulting in CMPs divided by total H2A violations reported. Each

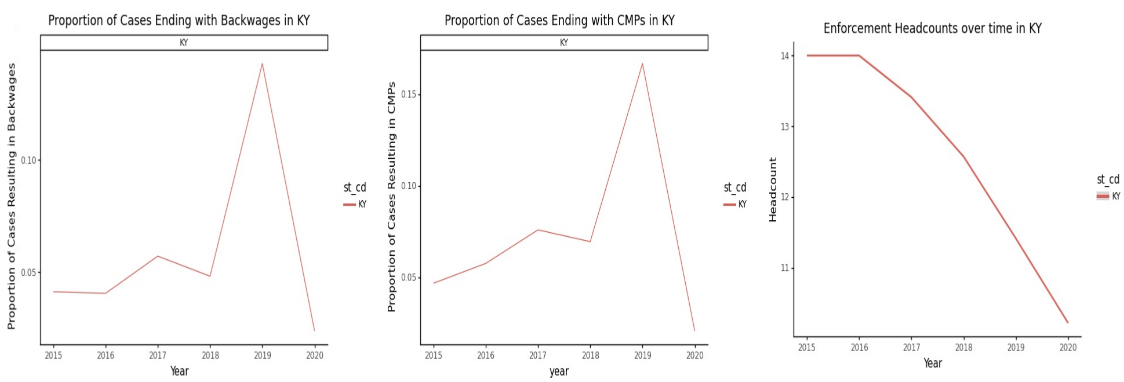
dot represents the proportion for one state in the given year. The second set of graphs are scatterplots representing the DOL's success in converting cases into CMPs. The Y axis represents the average amount of CMPs, which is calculated by diving the sum of CMPs collected by a given state by the total number of H2A cases opened. Each dot represents this average for the given year.

State Level Analysis After noting minimal correlations at the nationwide level, we decided to narrow our scope and focus on a few select states that the Social Impact Practicum (SIP) was interested in. These states were Alabama, Kentucky, Louisiana, Mississippi, Tennessee and Texas. Below are the graphs pertaining to these state specific analysis. From left to the right, the graphs show: 1. The change over time in the proportion of opened H2A cases that end in Backwages being assessed. 2. The change over time in the proportion of opened H2A cases that end in Civil Monetary Penalties (CMPs) being assessed. 3. The change over time in DOL staffing in that specific state.

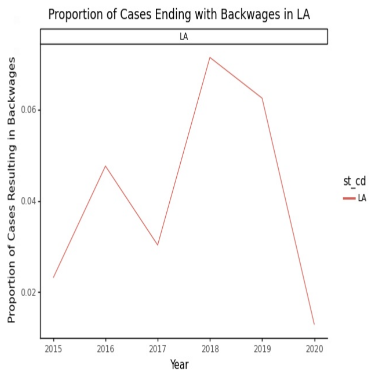
Alabama

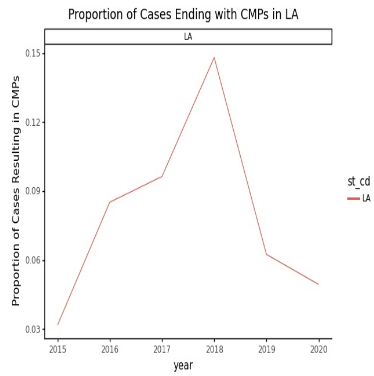


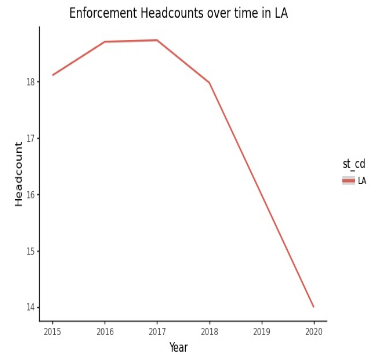
Kentucky



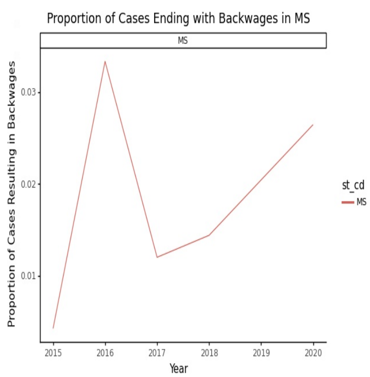
Louisiana

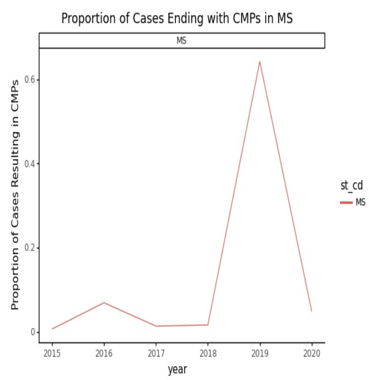


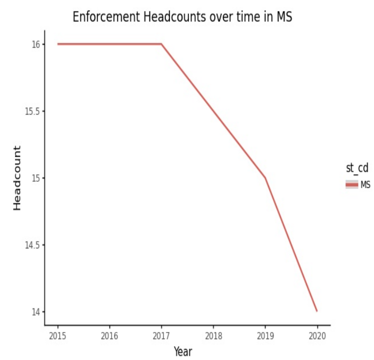




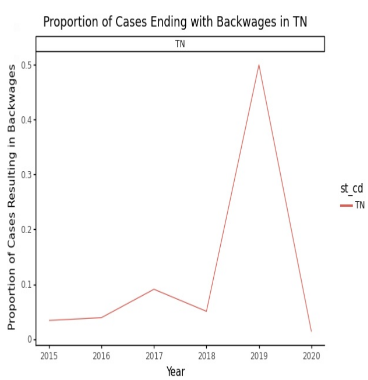
Mississippi

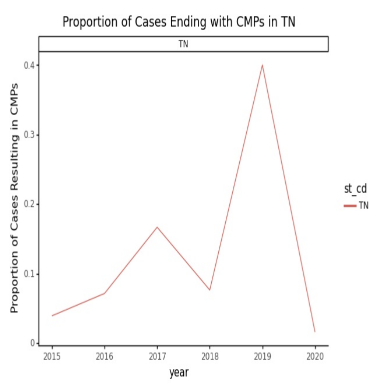


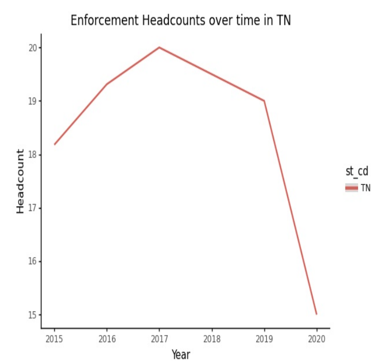




Tennessee







Texas

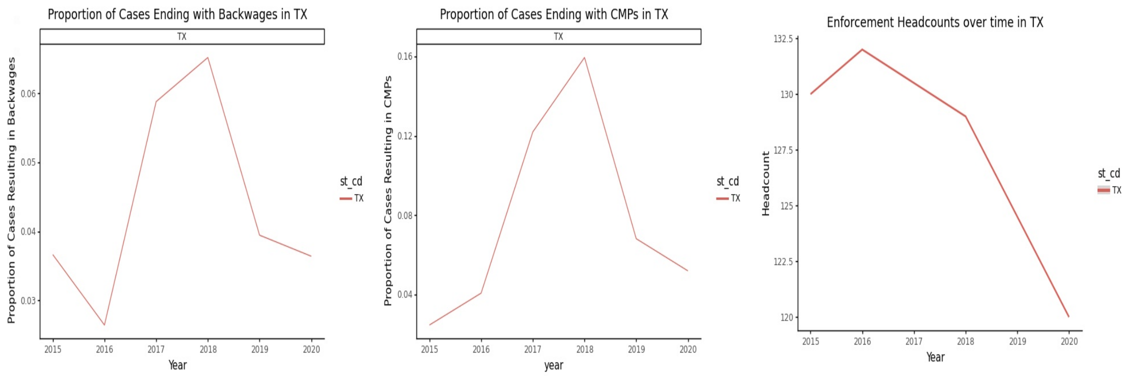


Figure 3: Change over time in (1) the proportion of H2A cases that result in back wages, (2) the proportion of H2A cases that result in CMPs, and (3) DOL staffing for the six states of interest Above is a set of line plots detailing the success of the DOL in the particular state in prosecuting H2A cases. The total number of cases are defined as those having a non-zero H2A violation count in the WHD dataset, and is used to calculate the proportion. The leftmost graph represents the proportion of opened H2A cases in the given state converted into assessed backwages. The middle graph represents the proportion of opened H2A cases in the given state converted into assessed CMPs. The rightmost graph is a smoothed line graph representing the total employee headcount in the given states DOL division. The comparison between conversion and headcount can be used to evaluate the effectiveness of the given state's DOL in effectively prosecuting H2A violations.

Fixed Effects Regression

To support our claims above, we performed a fixed effects regression of the amount of back wages per violation and the amount of civil monetary penalties per violation on the number of employees. We used data containing the average amount of back wages and civil monetary penalties as well as the number of employees for each state-year combination. Using a fixed effects approach allowed us to control for time effects and account for the unobserved differences across states that are constant over time.

In the first regression of back wages per violation on the number of employees, we obtain a coefficient of 0.0119 and a T-statistic of 0.4199. Holding constant time and state, an additional investigation employee is association with a 0.01 increase back wages per violation. This coefficient is not statistically significant. An extra one cent in back wages is not contextually significant either.

In the regression of civil monetary penalties per violation on the number of employees, we obtain a coefficient of 0.0273 and a T-statistic of 1.2817. Holding constant time and state fixed effects, an additional investigation employee is associated with a 0.03 increase in civil monetary penalties per violation. With a T-statistic less than two, this coefficient is also not significant. The regression tables are shown below.

Figure 6: Fixed Effects Regression of Back Wages per Violation on Number of Employees

Dep. Variable:	backwages	**R-squared:**	0.0009
Estimator:	PanelOLS	**R-squared (Between):**	0.0045
No. Observations:	256	**R-squared (Within):**	-0.0027
Date:	Sun, Jun 06 2021	**R-squared (Overall):**	-0.0006
Time:	17:04:15	**Log-likelihood**	-1848.3
Cov. Estimator:	Unadjusted		
		F-statistic:	0.1764
Entities:	50	**P-value**	0.6750
Avg Obs:	5.1200	**Distribution:**	F(1,200)
Min Obs:	1.0000		
Max Obs:	6.0000	**F-statistic (robust):**	0.1764
		P-value	0.6750
Time periods:	6	**Distribution:**	F(1,200)
Avg Obs:	42.667		
Min Obs:	36.000		
Max Obs:	48.000		

	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
total_employment	0.0119	0.0282	0.4199	0.6750	-0.0438	0.0675

F-test for Poolability: 1.2142
P-value: 0.1708
Distribution: F(54,200)

Included effects: Entity, Time

Figure 7: Fixed Effects Regression of Civil Monetary Penalties per Violation on Number of Employees

Dep. Variable:	backwages	**R-squared:**	0.0009
Estimator:	PanelOLS	**R-squared (Between):**	0.0045
No. Observations:	256	**R-squared (Within):**	-0.0027
Date:	Sun, Jun 06 2021	**R-squared (Overall):**	-0.0006
Time:	17:04:15	**Log-likelihood**	-1848.3
Cov. Estimator:	Unadjusted		
		F-statistic:	0.1764
Entities:	50	**P-value**	0.6750
Avg Obs:	5.1200	**Distribution:**	F(1,200)
Min Obs:	1.0000		
Max Obs:	6.0000	**F-statistic (robust):**	0.1764
		P-value	0.6750
Time periods:	6	**Distribution:**	F(1,200)
Avg Obs:	42.667		
Min Obs:	36.000		
Max Obs:	48.000		

	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
total_employment	0.0119	0.0282	0.4199	0.6750	-0.0438	0.0675

F-test for Poolability: 1.2142
P-value: 0.1708
Distribution: F(54,200)

Included effects: Entity, Time

6 Discussion and Conclusion

Nationwide At the nationwide level, trends are limited. Trend lines from respective graphs are either in direct opposition to comparable lines in other plots or flat / non-significant. There is a slight trend in the CMP 2020 graphs, where total employment actually decreases both the proportion of cases that end in CMPS and the dollars of CMP per case. This trend is not supported by the historical data in 2015 or 2017 however, and thus cannot be taken as a fact or treated as significant.

State-Level Results become more interesting at the state level. Most notably, employee headcounts in all target states asides from Alabama have seen steady declines since 2015. This is alarming on the surface, but also seems to causing some issues at an enforcement level. In all states asides from Alabama (which has had almost no employee turnover), there has been an overall decline in cases resulting in backwages and CMPs assessed since 2015. Additionally, there have been sharp declines in these areas since 2018, which corresponds to when states saw the most significant employee reductions.

Although these are just 6 of 50 states, it seems there is a reason that the SIP is concerned with them in particular. While there is little evidence to suggest that the real occurrence of violations is decreasing, there is evidence to suggest that the degree to which they are enforced and prosecuted.

Regression In our fixed-effect regression analysis, it is possible that there could be time-variant confounding variables affecting our estimates.

Limitations There were a few limitations that limited our analysis. The first is that the zip codes where DOL agents were stationed were not available, and thus we could not perform these analysis at a zip code level. An additional limitation to the data was that individual case id's did not include information about the nature of the H2A violations, so no sub-setting to specific violation types was possible. Finally, we did not have an official start date for each case in the investigations data and we did not have specific information about how long investigation preparations and proceedings take, so the lag we calculated between employees and the investigation date was estimated. With more information, our results could have been more refined.

Works Cited

Bauer, Mary. 2013. “Close to Slavery: Guestworker Programs in the United States.” *Southern Poverty Law Center*. Retrieved June 6, 2021 (https://www.splcenter.org/20130218/close-slavery-guestworker-programs-united-states).

Costa, Daniel, Philip Martin, and Zachariah Rutledge. 2020. “Federal Labor Standards Enforcement in Agriculture.” *Economic Policy Institute*. Retrieved June 6, 2021 (https://www.epi.org/publication/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/).

U.S. Department of Labor Wage and Hour Division. “H-2A: Temporary Agricultural Employment of Foreign Workers.” *U.S. Department of Labor*. Retrieved June 6, 2021 (https://www.dol.gov/agencies/whd/agriculture/h2a).

U.S. Department of Labor Wage and Hour Division. ”Wage and Hour Compliance Action Data.” *U.S. Department of Labor*. (https://enforcedata.dol.gov/views/data_summary.php)

U.S. Office of Personnel Management. ”Data, Analysis Documentation.” *U.S. Office of Personnel Management*. (https://www.opm.gov/data/)