Getting US Agricultural Labor Markets Right: Implications for Data, Modeling, and Policy Analysis

Alexandra E Hill, University of California Berkeley, alihill@berkeley.edu

Selected Paper prepared for presentation at the 2024 Agricultural & Applied Economics Association Annual Meeting, New Orleans, LA; July 28-30, 2024

1. Introduction

Realistic estimates of the size, geographical dispersion, and other characteristics of the farm workforce are vital for targeting government resources and supporting empirical and predictive analyses of the agricultural sector (Hill, Ornelas, and Taylor 2021). National models conventionally assume a single farm labor market with homogeneous, country-wide labor demand and supply, whereas, in reality, the US farm labor market is an amalgamation of localized markets linked together by commuting of workers and follow-the-crop migration (Taylor and Charlton 2018). This is at least partially attributable to gaps in our understanding of farm labor markets and in data availability. As one example, aggregate estimates of the size of the US crop workforce vary widely by source: 2017 estimates range from nearly 1.7 million workers (US Census of Agriculture—CoA) to fewer than 0.5 million (USDA Farm Labor Survey—FLS). Not surprisingly, sub-national measures can exhibit even greater variation and, perhaps more importantly, are often not available in publicly accessible data sources.

This paper provides a framework for thinking about and modeling the US farm labor market and assembling useful farm labor data. We identify different types of hired crop workers, with relevance from either the labor demand or supply side. On the demand side, the worker typology is grounded in distinct job or task needs and associated worker types from the perspective of crop employers. On the supply side, the typology is based on where employees originate, how they engage with agricultural employers, and whether and where they settle. Distinguishing between these types of workers is important because changes in policies, technology availability and adoption, outside job opportunities, and other factors impacting either the demand or supply of farmworkers, can impact each of these groups differently.

Combining these supply and demand side considerations, we identify five distinct worker types, which we depict in a tree diagram (Figure 1). At the top level, we distinguish between local and migrant workers. Local workers are settled in a particular location in the US, directly hired by farms, and commute daily to work. They are the first worker type; migrant workers comprise the remaining four. Migrant workers temporarily relocate for work; thus, their labor market spans a much wider area. We differentiate between migrant workers hired directly by growers versus those hired indirectly, through farm labor contractors (FLCs), as well as migrant workers employed through the US temporary agricultural worker visa (H-2A) program versus non-H-2A workers.

We next propose geographic delineations of US crop labor markets for the two top-level worker types: local and migrant. For local workers, we assume labor market areas are reasonably proxied by US commuting zones, because these workers are defined by their daily commute to work. For migrant workers, we generate novel labor market areas, which we refer to as "agricultural labor sheds." They are akin to commuting zones, but with a much wider geographic extension, tailored to a migrant crop workforce willing to travel longer distances for employment than local farm workers.

The "agricultural labor shed" concept is crucial for modeling farm labor markets. National models that ignore labor market areas implicitly treat the country as a single labor shed and assume perfect labor mobility within it. Simulations using a modified Simple-G model show that this oversight can result in exaggerated estimates of the local responses to market shocks, while underplaying the distributional impacts of local policy interventions for the welfare of farmers and farm workers (Ray et al. 2023).

We then construct estimates of key employment characteristics at the worker-type and agricultural labor shed levels. To do so, we summarize and compare available data on US crop workers with information on employment counts, wages, and working hours. We identify and discuss the reasons for discrepancies across five distinct data sources (Census of Agriculture, Farm Labor Survey, National Agricultural Workers Survey, Quarterly Census of Earnings and Wages, and H-2A Visa programmatic disclosure data).

To the best of our knowledge, our study is the first to provide a strategic and comprehensive overview of available data on the US crop workforce, firmly grounded in a realistic theoretical framework of the US farm labor market, and to identify how differences in data collection and purpose drive discrepancies in agricultural labor estimates. We summarize all available estimates of farm employment at the national level and, based on this discussion, we recommend improved ways to estimate employment, wages, and hours worked on the farm. This study is the first to generate these disaggregated measures related to employment for the US crop workforce, which is critical to improve the accuracy and breadth of empirical research within agricultural labor economics and agricultural economics generally.

2. Conceptual Framework: Modeling Modern US Farm Labor Markets

Agricultural labor markets are unique for a multitude of reasons (Fisher and Knutson, 2012). On the demand side, there is substantial heterogeneity in the labor needs of farms in terms of number of workers, necessary skillsets, and timing of needs (within a year, season, and even day). Differences in labor demand can be impacted by a range of conditions, including spatial and temporal weather patterns, biological features of the crops they are producing, the technologies they use, and more. On the supply side, workers are heterogeneous in terms of their experience working in agriculture, which contribute to their skillsets, their unique preferences for types of working arrangements, geographic locations, willingness to relocate for work, and their opportunities for and desire to work in other industries. Differences in labor supply can similarly be impacted by a range of factors, largely related to individual preferences, but that might be correlated with worker demographics like age, family status, legal immigration status, and more.

The factors that influence both labor demand and supply have changed substantially over time, particularly in recent years, making modern US agricultural labor markets increasingly distinct from historic ones. Recent data suggest that non-H2A visa workers are aging, increasingly settled, and are much less likely to engage in follow-the-crop migration (Gold et al., 2022). On the demand side, shifting climate conditions and new technology adoption have changed both the types and timing of worker needs. Additionally, H-2A visa workers are becoming increasingly prominent in agricultural operations as a way of filling the labor gap between what is needed when, and what is available. To the best of our knowledge, no existing studies offer an up-to-date characterization of US farm labor markets that account for these factors.

In this section we outline distinct types of workers that are prominent in modern farm labor markets and explain from both demand and supply perspectives the economically important distinctions amongst these different worker types. We then geographically define the labor market areas for two high-level distinctions between these worker types – local and migrant workers.

2.1 Types of Workers

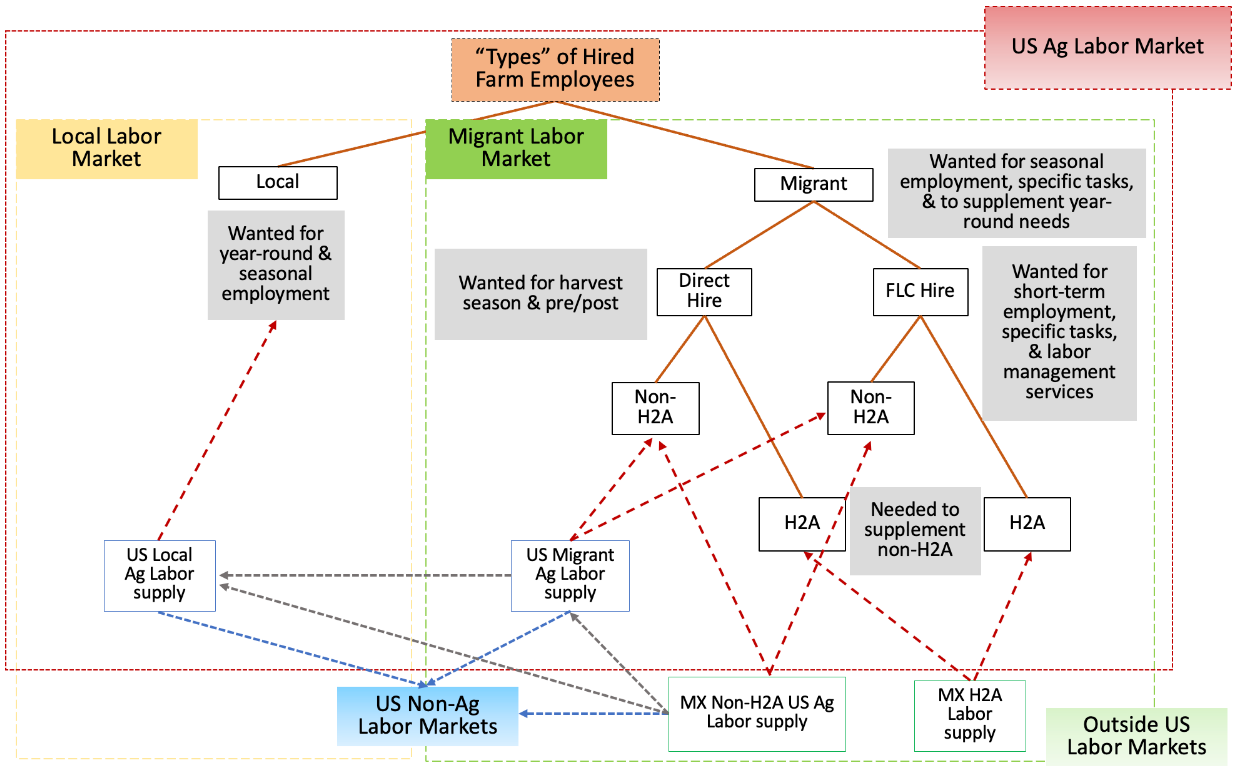


Figure 1. US Farm Labor Market Worker Typology

Notes: Illustration of US farm labor markets and the "types" of workers participating. At the top node of the tree we differentiate between two fundamentally distinct worker types: local and migrant workers. We further disaggregate migrant workers into those who are hired directly by farms and those hired indirectly, through a farm labor contractor (FLC). We define local workers such that they are always hired directly by farms. Both direct and FLC hired workers can be either H-2A visa holders or non-H-2A workers. All worker typologies consist of local agricultural workers, domestic migrant agricultural workers, Mexican non-H-2A agricultural workers, and Mexican H-2A agricultural workers. All non-H-2A workers might also exit the agricultural labor market for jobs in the US non-farm sector. In addition, domestic farm workers might transition to being local workers (though the opposite is uncommon) and Mexican non-H-2A workers might transition to being either local or migrant US workers.

Figure 1 illustrates our depiction of the distinct types of hired employees in modern farm labor markets and offers some insights into why, from the labor demand perspective, differentiating between these types of workers is important. At the top of the tree, we distinguish between local and migrant workers. We define local workers as those directly hired by crop employers who do not travel the country for employment opportunities. These might be year-round workers or seasonal workers who find sufficient temporary employment opportunities within a local area.

Local workers are necessary for both year-round and seasonal employment in agriculture; without these workers employers would not be able to accomplish farm tasks in the off-peak seasons, including planting, applying fertilizers, herbicides and pesticides, soil and equipment maintenance, and early and late season harvest activities. However, if all workers were local and seeking these off-peak employment opportunities, agricultural operations would not have sufficient jobs to support them. Migrant workers, who typically appear in an area when labor demands peak, are thus key complements to the local workforce to accommodate the large seasonal changes in worker demand without mass unemployment in the off seasons. These workers constitute the remainder of the agricultural workforce.

Migrant workers can engage in follow-the-crop migration on their own – temporarily relocating from place to place with the harvest season and seeking employment opportunities on their own, thus being directly employed by employers. Migrant workers also constitute those employed by farm labor contractors, who manage this movement and negotiate terms of employment on behalf of their workers, sometimes piecing together year-round employment in agriculture, other times providing temporary employment opportunities. Among migrant workers who are directly hired by crop employers and those hired through farm labor contractors (FLCs), workers can be either H-2A workers (those who enter the US under the H-2A visa program for temporary agricultural employment) or non-H-2A workers. We treat H-2A workers as migrants regardless of employment type because they have effectively similar mobility as migrant workers: they can be contracted on a needs-basis across the country. Similarly, we treat all workers employed through FLCs as migrant workers because these contractors typically move workers from job to job, often covering large geographic areas, certainly outside the boundaries of commuting zones, based on employer needs.

This distinction between local and migrant workers is key for several reasons. First, migrant workers have historically played a pivotal role in US agricultural production. Those who historically engaged in follow-the-crop migration followed primarily south-to-north paths, based on crop and location specific seasons of production, to piece together year-round employment. These workers have long been acknowledged as vital to harvest activities which require a large influx of workers for a relatively short period of time, but these harvest activities occur at different times of the year in different locations. Second, one of the most notable trends in agricultural labor markets in recent years has been the decline in the migrant workforce (Ray et al., 2023; NAWS, 2022) Thus differentiating between these worker types is important in understanding how the farm labor market has evolved over time. Third, local and migrant workers serve different purposes from the demand-perspective. Local workers are available year-round and thus are available to perform necessary tasks in the off-season, such as maintenance and repairs. Migrant workers were historically desired for harvest activities alone, but in modern labor markets can also be used for year-round employment as needed to supplement the local labor force. The increasing applications for H-2A visa workers for the maximum one-year employment term speaks to this rising demand for workers with specific skillsets to stay with the farms year-round.

At the bottom of the tree we highlight the origins of these different worker types and illustrate how workers can transition types across time. Both local and migrant workers can originate from US farm labor markets or labor markets outside of the US. Those from US labor markets can be settled and contribute to the local labor supply, or migratory. Migratory workers might transition to local workers by settling, but the opposite is rare. Similarly, both types of workers might also transition out of agriculture, seeking employment in US non-farm industries, but workers from non-farm industries are unlikely to enter the farm labor market. Finally, in reality there are both H-2A and non-H-2A farmworkers in the US originating from other countries, but the majority are from Mexico and we model accordingly.

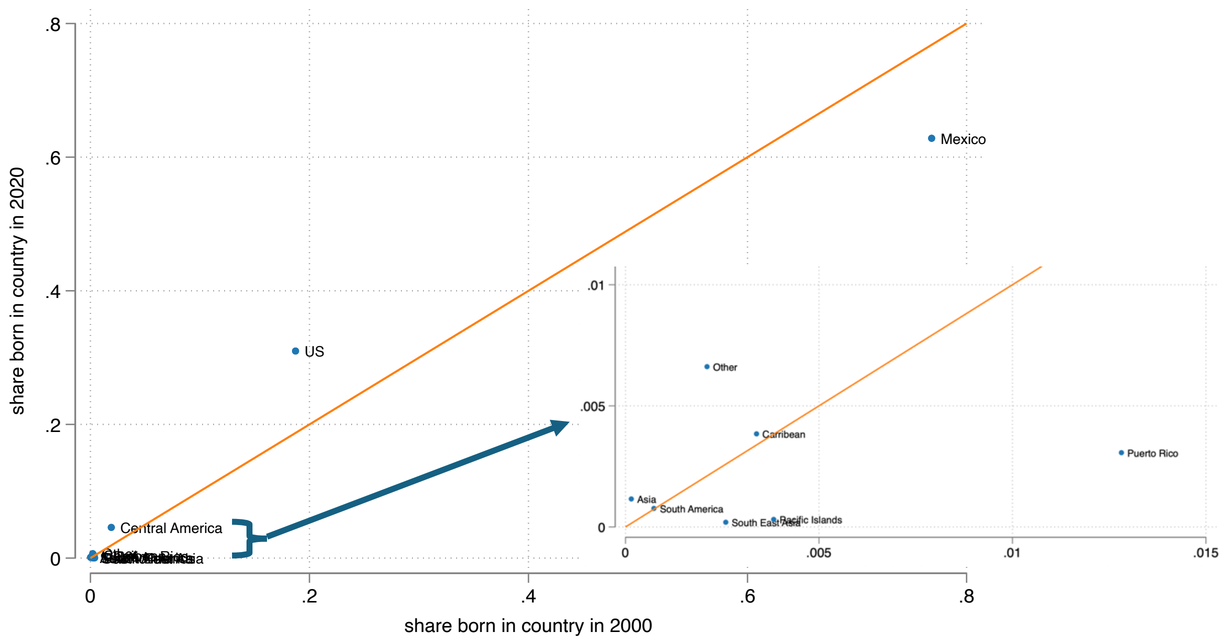


Figure 2. US Farmworker Countries of Origin, 2000 and 2020

Notes: Scatter plot of the share of US crop workers born in indicated regions in 2020 (y-axis) versus 2000 (x-axis). The orange line (y=x) clarifies changes over time in region of origin; points above the line indicate that the share of US crop workers born in that region was higher in 2020 than in 2000 and points below indicate the opposite. The overlayed figure on the right shows a zoomed in version of the regions of origin for less than 1 percent of the US farm workforce. Figure estimated by the authors using data from the National Agricultural Workers Survey (NAWS). Shares are estimated using NAWS survey weights and data from the 1999/2000 and 2019/2020 survey rounds.

Figure 2 supports this claim and illustrates that this has changed little in the past two decades. Using data from the National Agricultural Workers Survey (NAWS), the figure plots the share of non-H-2A US crop workers by their country of origin in 2000 (on the x-axis) and 2020 (on the y-axis). In both periods, most workers were born in Mexico, but the share decreased over time; 77% of workers in 2000 were born in Mexico versus 62% in 2020. US-born workers comprise the next largest share of the farm workforce, and this share increased over time; 19% of workers in 2000 were born in the US versus 32% in 2020. Mexico- and US-born workers comprised more than 94% of the farm workforce in these years, with the remainder coming from Central America, Puerto Rico, the Caribbean, and other areas.

2.2 Labor Market Areas

A key contribution of this paper is the construction of distinct delineations for labor markets (or labor sheds) for local and migrant agricultural workers. The extent to which agricultural workers are able and willing to move between locations is an important determinant of a variety of outcomes at the local, regional and global levels, including input choices, crop mix, agricultural production and prices. However, little is understood about within-country migratory patterns of US agricultural workers. Existing delineations of labor markets are created for a variety of purposes including the degree to which groups of counties are economically “connected” (i.e., wage similarities and simultaneous shifts in employment rates) or “contained” (i.e., workers live and work within these areas), using labor market information across all industries (Fowler and Jensen, 2020). The groups of counties most “connected” or “contained” with regard to agricultural labor markets likely differ from those for other sectors. We fill this gap and define the relevant labor sheds for local and migrant farm workers.

We define the markets for local workers as commuting zones, which are representative of daily commuting patterns for US workers employed across all sectors (cite CZ papers). However, to the best of our knowledge, no prior work has attempted to define labor markets for migrant agricultural workers. We considered a variety of approaches for constructing these migrant labor sheds before arriving at our preferred delineations, building on the USDA-ERS Resource Regions.

Figure 3 shows these preferred geographical delineations of US migrant labor sheds. They are constructed from aggregating and disaggregating USDA ERS Farm Resource Regions to better capture migratory patterns that are consistent with the timing of crop production within regions, that is, creating regions that capture south-to-north market sheds within which workers are likely to travel to secure year-round employment in agriculture.

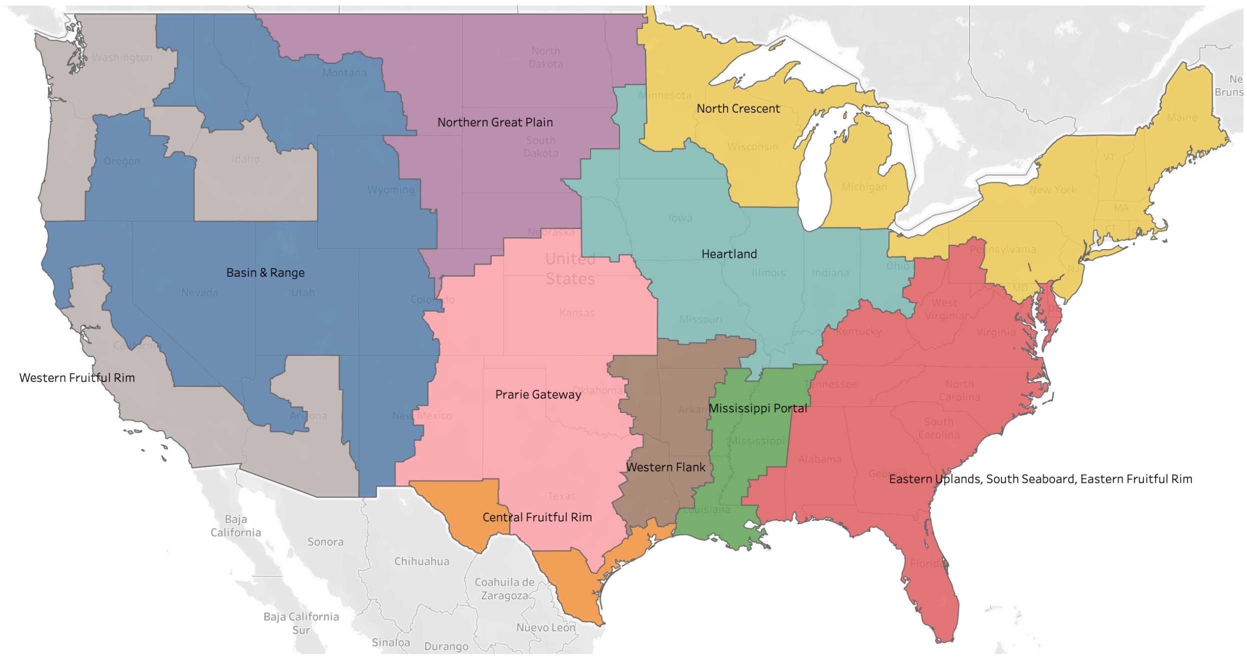


Figure 3. Migrant Labor Sheds

Notes: Geographic delineations of the "agricultural labor sheds" for migrant US farmworkers. Each shed is comprised of connected US counties and based on aggregating and disaggregating USDA farm resource regions to better capture North-South commuting patterns of migratory farmworkers.

The USDA ERS Farm Resource Regions divide the U.S. into 9 distinct regions that reflect "geographical specialization in production of U.S. farm commodities" (Heimlich, 2000). These resource regions group together counties with similarities in farm commodities – for example, drawing distinctions between areas that primarily produce cattle and wheat and those producing fruits and vegetables – and farm types – for example, family versus nonfamily farms, smaller versus larger farms, and relative farm production value. While the primary intent of these resource regions is to reflect farm supply-side similarities across the U.S., these regions are also likely to capture labor supply and demand similarities. Different farm commodities require specific labor skillsets and experience. As one example, row crops like corn and wheat require relatively few workers to operate machinery for harvest, whereas many fresh fruits still require workers who are able and willing to hand harvest. Even within these broad commodity categories and for specific tasks, workers often have preferences for the working conditions associated with specific commodities. For example, strawberry harvesters rarely seek harvesting opportunities in tree fruits, preferring harvest activities without ladders and perhaps enjoying working in closer proximity to their coworkers than is typical in orchards. Because of these factors, ERS resource delineations serve as a good starting point for delineating migratory flows for crop work.

However, one limitation of the ERS resource regions is that they do not capture seasonal variation in production in a manner useful for considering the behavior of agricultural workers. In particular, we would expect migrant, or follow-the-crop, workers to move primarily in south-to-north patterns over the course of the year, as this enables them to capitalize on the differences in harvesting times across the US (whereas moving in East-West patterns would not make sense because harvest times would be similar). To adjust for this, we use a data-driven approach to determine whether any of the ERS-delineations should be disaggregated or aggregated. These disaggregated and aggregated ERS farm resource regions are our first considered (and preferred) labor shed delineation. For completeness, however, we compare these delineations with three alternatives including the "raw" or original farm resource regions, in Appendix A, along with alternative a detailed data-driven justification for our preferred delineations discussed here.

Our first consideration in deciding whether to disaggregate any farm resource regions is whether the region could capture vertical, rather than horizontal, movement. The most obvious candidate for disaggregation was the fruitful rim, which includes counties in Washington, Oregon, Idaho, California, Arizona, Texas, and Florida. We separate the fruitful rim into three regions to be more consistent with capturing vertical migrant flows: the western fruitful rim, consisting of all fruitful rim counties in Washington, Oregon, Idaho, California, and Arizona; the central fruitful rim, consisting of all fruitful rim counties in Texas, and the eastern fruitful rim, consisting of all fruitful rim counties in Florida. Next, we assert that labor market areas should be geographically connected and accordingly disaggregate the Eastern Uplands and Southern Seaboard regions into separate segments east and west of the Mississippi Portal region. Because the western portions of these two disaggregated regions are small, adjacent, and form a North-South area, we then combine these two areas west of the Mississippi Portal into the Western Flank region (the Western Flank of the Eastern Uplands and Southern Seaboard).

To evaluate how well these regions capture farm labor mobility, we consider whether and to what extent migrant workers reported working across and within each region. To do this, we use data from the restricted-access version of the National Agricultural Workers Survey (NAWS) which includes information on farm and non-farm employment, as well as periods of unemployment for each interviewed worker for the prior year. Using these data, we define migrant workers as those who reported doing farm work in more than one county in the prior year. We restrict our sample to workers who meet this definition of migrant in all applications of the NAWS for constructing or evaluating migrant labor shed delineations, and later when using the NAWS to estimate the share of migrant workers in the crop workforce. Thus our classification of "migrants" are workers who perform farmwork in multiple counties within the past year.

Using these data we estimate key performance statistics for four alternative delineations of migrant worker labor market areas: (1) our preferred regions described above, (2) dividing the US into 4 North-South regions along state boundaries, (3) ERS Farm Resource Regions, and (4) Commuting Zones. We detail the construction of each alternative labor market area in Appendix A, and summarize multiple statistics related to their performance in terms of accurately capturing migrant farmworker flows in Table 1. We produce these performance indicators by cleaning and summarizing the NAWS data as follows: we remove duplicate observations of worker-county employment to capture the unique number of workers with jobs in each county; next, for each region, we reduce the sample to workers who reported performing farmwork within that region within the past year (including their current farm job that is the place of the interview); we then calculate the number of unique workers in the sample for the region, the proportion of those workers who worked in multiple counties *within* the potential labor shed delineation, and the proportion of those workers who worked in counties in other regions. Table 1 summarizes the average number of jobs held by each worker who reported working within a region, the percent of workers working multiple jobs in the same region (across all regions), and the percent of workers working jobs in multiple regions (across all regions). In general, these summary statistics illustrate that our USDA Farm Resource Region-based approach to constructing migrant labor markets perform well; 83% of migrant workers worked more than one job within the same region, and 33% worked at least one job in a different region. This represents an improvement over both the standard ERS Resource regions and commuting zone delineations, but performs somewhat worse in terms of these measures than the North-South regions. However, we believe that our preferred delineations represent an improvement over the North-South regions, because they divide the US into a larger number of labor markets (10 versus 4), and more accurately divide the US in terms of agricultural production characteristics.

Table 1. Performance of Alternative Migrant Farmworker Labor Market Areas

	Region			
	1 (Preferred Regions)	2 (North-South)	3 (ERS Resource)	4 (CZs)
Avg. # of jobs per worker within regions	1.852	2.013	1.728	1.211
% working multiple jobs in same region	83.04	89.19	77.69	55.71
% working in multiple regions w/in last year	33.47	22.35	43.76	82.91

Next, we turn to the data sources available on US farmworker employment and outline our approach to estimating key employment variables at the labor shed and worker-typology levels.

3. Data on Agricultural Workers

There are several sources of data with information on the number of farmworkers and other employment characteristics. We provide an overview of what we believe are the five most accurate and comprehensive data sources with such information. We provide an overview of each data source, highlight their relative strengths and weaknesses, and discuss the reasons for differences in the estimates they produce. We then justify our preferred data sources for different metrics and outline our approach to aggregating and disaggregating available data to generate desired employment statistics at the worker type-labor market area level. The analysis consists of iteratively performing these manipulations for the full breadth of available data from 2002 through 2022. Recognizing that there is a logic for existing data, we discuss reasons for discrepancies across data sources to motivate our preferred employment estimates, but we also present the full range of estimates one can reasonably draw.

In this paper, we use data from the US Census of Agriculture, the Farm Labor Survey, the National Agricultural Workers Survey, the Quarterly Census of Earnings and Wages, the American Community Survey, and H-2A Disclosure Data. Table 2 summarizes each of these data sources. The table highlights the level of granularity each data source offers, including industry (in particular, whether the data are available separately for crop and animal agriculture), worker typology (ability to distinguish between direct and contract hire, local or migrant workers, or H-2A and non-H2A workers), and geography (the smallest geographic area at which the data can be aggregated). The table also summarizes the relevant employment-related variables available in each data source at the indicated level of aggregation, along with the years they are available. To the best of our knowledge, these data sources comprise the universe of data that are national in scope and provide reliable, longitudinal information on US farm employment.[1]

The Census of Agriculture (CoA) is conducted every five years by the US Department of Agriculture's National Agricultural Statistics Service (NASS). The CoA includes all US farms and ranches with at least $1,000 in annual sales of agricultural products. It provides the most comprehensive data on US agricultural production and the individuals involved. This project uses the publicly available summary data from the CoA available through NASS's searchable Quick Stats database. These public data summarize information collected in the CoA at the county, state, or national level, depending on the requested statistics. In this study we use information on the number of workers and farm payroll costs provided in the CoA. The CoA asks farmers and ranchers to report their total number of direct hire farm employees, their total payroll costs for direct hire and contract hire workers, and the number of migrant workers on their operation (that were either hired directly or through a contractor). Limitations to the CoA include that respondents are not asked for the number of contract-hired workers they employ, and that at the county-level all employment statistics are only available for crop and animal operations combined, whereas for this study we aim to estimate employment on crop farm operations only. The CoA does provide the information separately for crop and animal operations at the state-level, but because state boundaries do not align with our farm labor sheds, this has limited utility for our analysis.

We also access data from the Farm Labor Survey (FLS) through NASS's Quick Stats database. Most information from the FLS is available at the NASS Farm Labor Region level which divides the country into 15 multi-state regions and 3 single-state regions for California, Hawaii, and Florida.[2] Because these regions do not align with our farmworker

---

[1] The Current Population Survey and American Community Survey also allow researchers to identify individuals employed in US agriculture, but both have limited utility because they are household-based surveys, and because farmworkers often live in non-traditional housing, vastly underestimate the number of farmworkers in the US.

[2] See page 26 in USDA Farm Labor report (2023) for information on the states included in each of these regions: https://downloads.usda.library.cornell.edu/usda-esmis/files/x920fw89s/v405tw18s/dn39zk84n/fmla1123.pdf

labor market areas, the FLS has limited utility for producing estimates of employment at the level of granularity we are seeking. However, for completeness and to highlight differences across available data sources, we include the FLS in our national-level summary of employment information. The publicly available FLS data summarize information for crop and animal workers separately for the number of jobs, the total contract and direct hire payroll costs, and the direct-hire hourly wage rate. It additionally summarizes average weekly workhours for crop and animal workers combined.

We use the restricted access version of the National Agricultural Workers Survey (NAWS) in our analysis, which includes the county of employment, enabling us to aggregate statistics at the agricultural worker labor shed level.[3] The NAWS is the only dataset we use that collects data from farm employees, rather than employers, and the NAWS is restricted to individuals currently employed in US crop production. The NAWS is an employment-based, random sample of currently employed farmworkers conducted by the US Bureau of Labor Statistics (BLS). The survey sampling methodology consists of randomly sampling workers, based on their place of employment, using information on farm locations from the Farm Labor Survey. As such, while the NAWS offers comprehensive and regionally representative information on farmworkers, their families, employment, and much more, the survey cannot be used to estimate the total number of farmworkers, or the total payroll cost to employers. However, the survey does provide granular and useful information on key employment attributes omitted in employer-based surveys, including hourly wages and annual earnings, weekly workhours, and annual workweeks. In addition, because the NAWS collects detailed demographic information on interviewees, it also allows for estimating these statistics separately for migrant and local workers, and direct hire and contract hire workers.

The Quarterly Census of Earnings and Wages (QCEW) provides a quarterly count of employment and earnings reported by employers. The QCEW is primarily informed by administrated data from state unemployment insurance systems and is supplemented by data from two BLS surveys. The QCEW covers more than 95 percent of U.S. jobs and provides annual reports on nonresponse and undercounting. The publicly available QCEW summarize this information at the quarter-county-industry code level. Data are available at up to the 6-digit NAICS code level, but here use the three-digit NAICS codes that identify establishments in the crop (111) and animal (112) sectors. The QCEW is understood to undercount employment in several sectors including agriculture because certain workers and employers are not covered by the unemployment insurance system (BLS, 2023). The BLS provides annual reports on the estimated number of included and excluded wage and salary agricultural workers, which suggest that 7 to 19 percent of these workers are excluded from the QCEW (BLS, 2004; BLS, 2022).

[3] Relative to other nationally representative surveys, the NAWS has a small sample size; it consists of 1,200-3,600 in-person interviews with farmworkers each fiscal year. Because of this, the NAWS data provide an accurate representation of US farmworkers when data are appropriately aggregated. Typical NAWS analyses aggregate data at NAWS region level (which divides the US into 5 multi-state regions and California on its own) using two fiscal year groups (which comprise a single interview cycle). For this study, we aggregate the data at either the national farm labor market area levels and present summary statistics in two fiscal year groups.

The final data source we use in this paper come from the US Department of Labor Foreign Labor Certification Annual Performance Data for the H-2A Visa program. For the purposes of this project, we have cleaned and harmonized these datasets for years 2006 to 2022 and have provided these data for public use in our project GitHub repository. The H-2A disclosure data consist of selected information extracted from H-2A visa worker requests, including the number of workers requested, the state, and in later years zipcode, of the proposed worksite, dates of requested and granted employment, the promised hourly wage rate, the and the planned weekly hours of work. In this project, we treat the number of certified H-2A worker positions as the number of H-2A workers in the US farm labor market. Ideally, we would prefer a measure of the number of H-2A visa workers who enter the US to work, but such data, along with geographic information on their worksites, are not currently available (cite Phil).

Table 2. Data Sources, Levels of Aggregation, and Years Available

Data Source	Industry Level	Worker Type	Geographic Level	Employment Information Provided					Years
				# Jobs	Payroll Cost $	Hourly Wage $	# Weekly Work Hours	# Annual Work Weeks	
CoA	Crop + Animal	Direct	County	Yes	Yes	No	No	No	2002, 2007, 2012, 2017, 2022
		Contract	County	No	Yes	No	No	No	
		Migrant	County	Yes	No	No	No	No	
	Crop, Animal	Direct	State	Yes	Yes	No	No	No	2012, 2017, 2022
		Contract	State	No	Yes	No	No	No	
FLS	Crop + Animal	Direct	Region	Yes	No	Yes	Yes	No	2000-2022
	Crop, Animal	Direct	Region	Yes	No	Yes	No	No	
	Crop, Animal	Direct + Contract	Region	No	Yes	No	No	No	
NAWS	Crop	Direct, Contract, Migrant	Region	No	No	Yes	Yes	Yes	2000-2022
QCEW	Crop, Animal	Direct, Contract	County	Yes	Yes	No	No	No	2000-2022
H-2A Program Data	Crop, Animal	Direct, Contract	Zip code (2015-22); City (2008-14); State (2006-07)	Yes	No	Yes	Yes	Yes	2006-2022

4. Estimates of Agricultural Employment: Number of Workers, Total Wage Payments, Wage Rates, Weekly Working Hours, and Annual Workweeks

Estimates of farm employment are useful to and desired by a variety of stakeholders, including academics, policymakers, and agribusiness owners and operators. However, there are multiple ways to define farm employment and the relevant or desired definition varies by application. For example, policymakers and agribusiness owners might be most interested in knowing the total number of individuals available to do farm work, the number of hours they are willing to work, or their outside employment opportunities. On the other hand, researchers modeling agricultural production might be more interested in the total number of jobs, work hours per employer, total expenditures on wages, or worker earnings.

In an ideal world, we would be able to produce all of these estimates. This would require data that link individual workers to each of their employers, along with details on their employment, including working hours, wages, and benefits. Unfortunately, this is infeasible with currently available data, although it is possible that with a combination of employee and employer tax records, such a dataset could be constructed.

Here we focus on what is possible to estimate.
Here we outline our approach to estimating US farm employment and associated terms of employment, using available data.

We estimate the number of jobs, rather than number of workers, because such estimates are more widely available and comprehensive. These estimates come from aggregating employer-level information on their reported number of employees. Because this information comes from employers, rather than employees, it reflects the number of farm jobs and should be larger than the number of distinct farm workers, since each farm employee could work for multiple employers within a given year.

4.1 Methods for producing disaggregated estimates of farm employment

Compare estimates of number of crop + animal direct hire farm jobs:

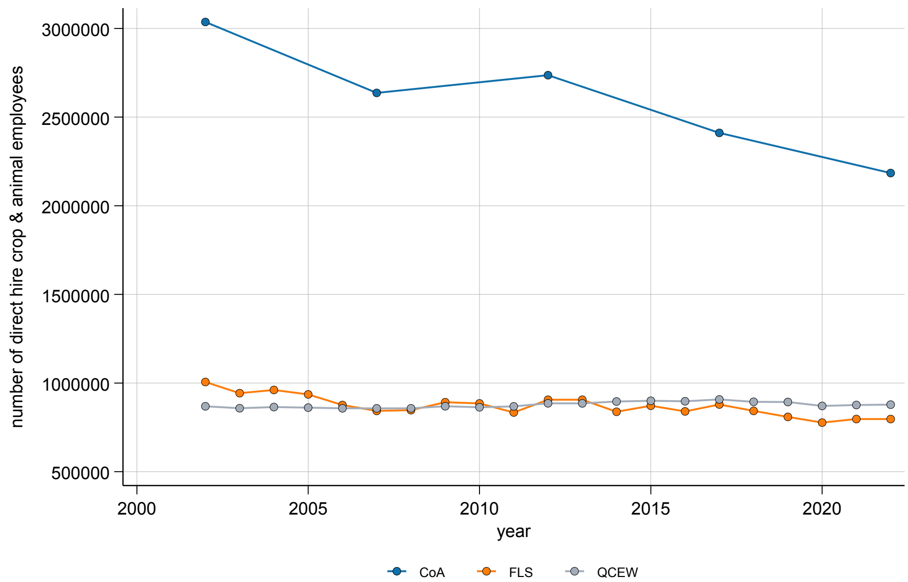


Figure 4. Estimates of Crop and Animal Employment

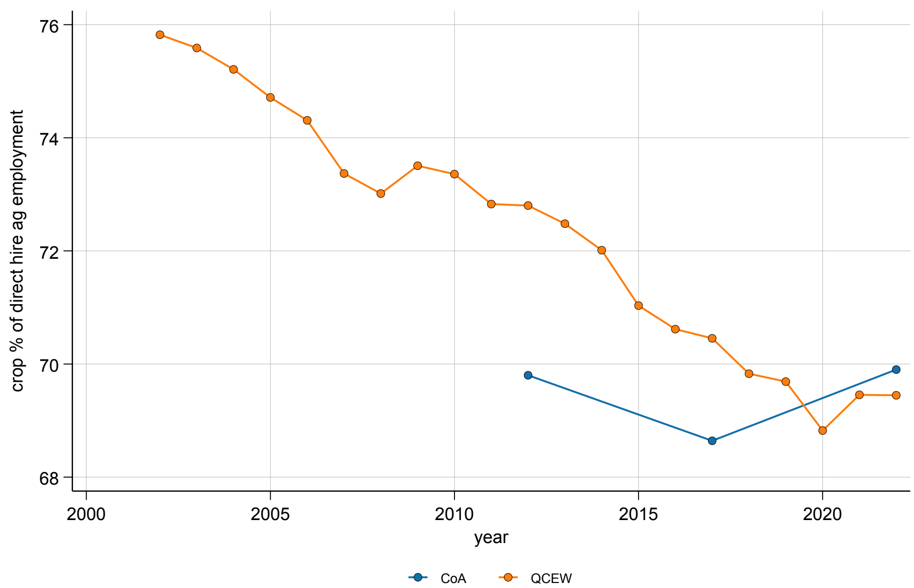
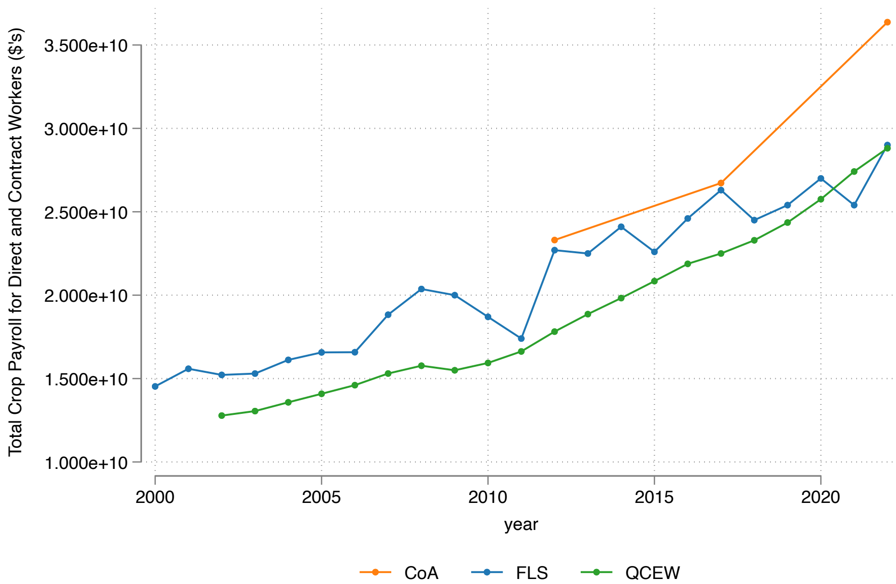


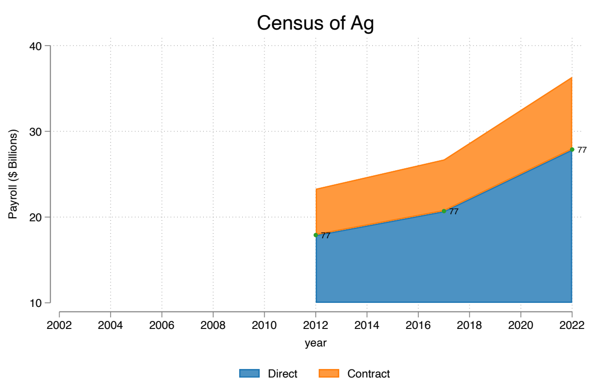
Figure 5. Estimates of Percent of Ag Employment in Crop Production

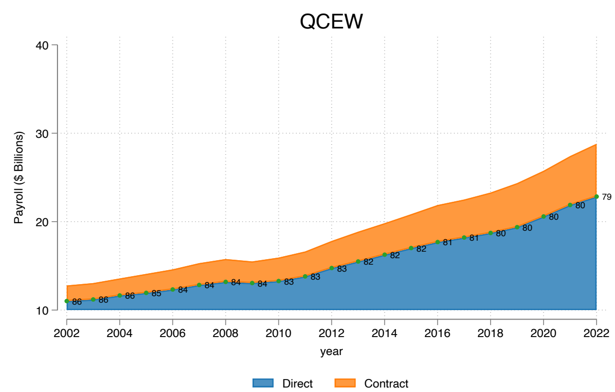
Notes: We identify direct hire employment in crop and animal production using NAICS 111 for crop production and NAICS 112 for animal production in the QCEW. The public-use data from the CoA include total direct hire number of workers by 4 and 5-digit NAICS codes beginning in 2012, so we identify crop employment using NAICS 1111, 1112, 1113, 1114, and 1119 and animal employment using NAICS 112111, 112112, 11212, 1122, 1123, 1124, 1125, and 1129.

Compare total payroll costs for crop employers by survey



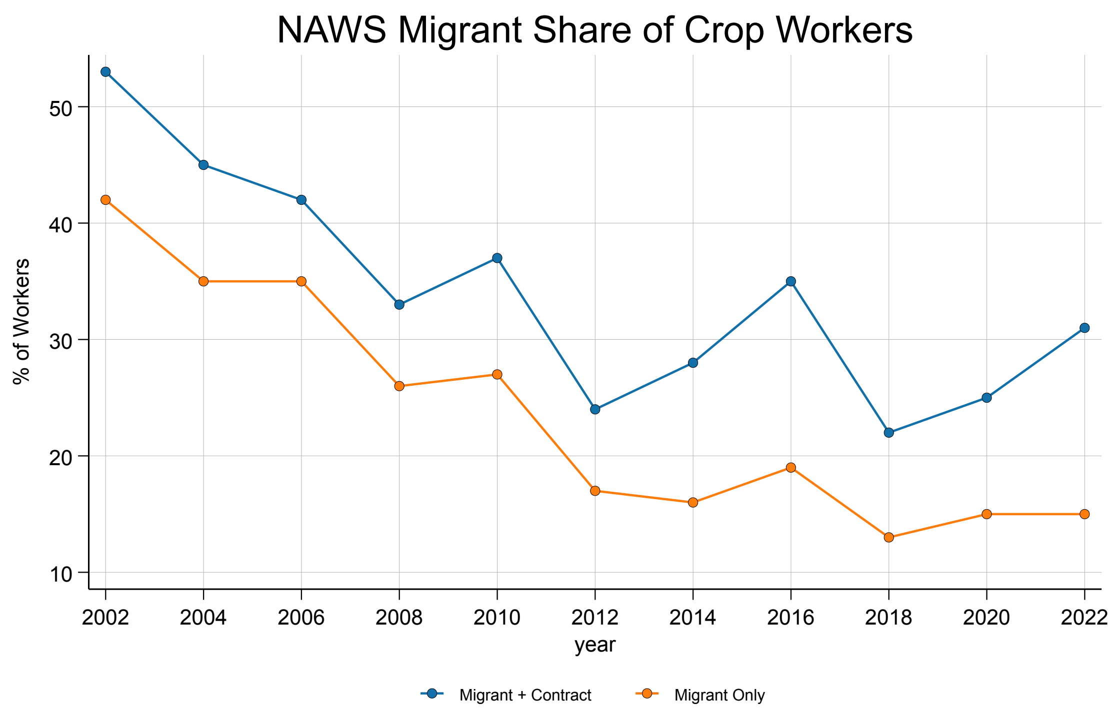
Separating Direct and Contract Hire Expenses:

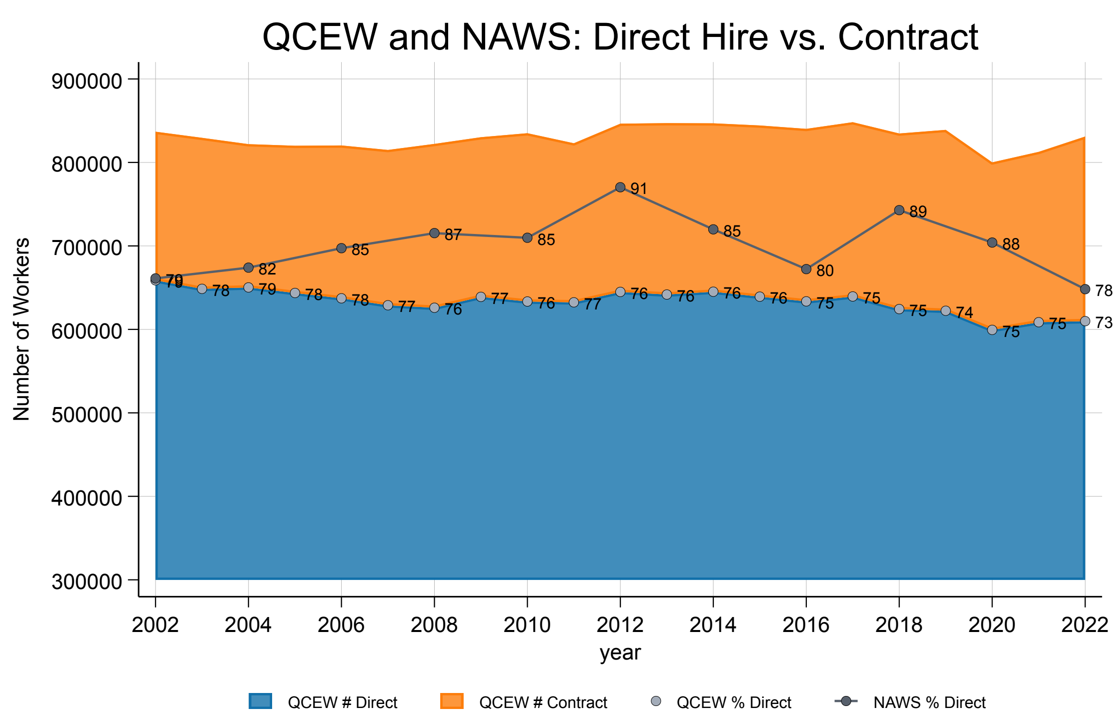




Notes: numbers indicate the direct hire percentage of total payroll (direct + contract payrolls). Census of Ag numbers are farm expenditures on contract workers, whereas QCEW are payroll expenses of farm labor contractors. Expenses for contract workers in CoA should be higher than QCEW because these will include FLC overheads. In addition, CoA asks employers to include "employer's cost for social security, worker's compensation, health and life insurance premiums, pension plans, etc." whereas the QCEW expenses include "bonuses, stock options, severance pay, the cash value of meals and lodging, tips and other gratuities. In some states, wages also include employer contributions to certain deferred compensation plans, such as 401(k) plans." "Covered employer contributions to old-age, survivors, and disability insurance; health insurance; UI; workers' compensation; and private pension and welfare funds are not reported as wages. Employee contributions for the same purposes, however, as well as money withheld for income taxes, union dues, and so forth, are reported, even though they are deducted from the worker's gross pay." For these reasons, it is intuitive that expenses in the CoA are higher than in the QCEW. The QCEW also excludes employers with fewer than 11 employees in some states (based on state UI policies). The QCEW estimates that this resulted in excluding roughly 18% of hired agricultural workers in 2022 (https://www.bls.gov/cew/publications/employment-and-wages-annual-averages/2022/home.htm#exclusions) and 8% in 2002 (https://www.bls.gov/cew/publications/employment-and-wages-annual-averages/2002/home.htm#Employment). Note that their estimates of the percent of agricultural workers excluded have risen over time.

Other Measures of Numbers of Jobs/Workers: # H-2A, % Migrant (NAWS), # Contract (QCEW), % Contract (NAWS)





Notes: QCEW estimates for direct hire crop employment use peak annual employment for NAICS 111 and peak annual employment for NAICS 115115 for Contract employment. NAWS estimates are weighted using survey-provided sampling weights, and are estimated in two-fiscal-year increments in accordance with the sampling methodology.

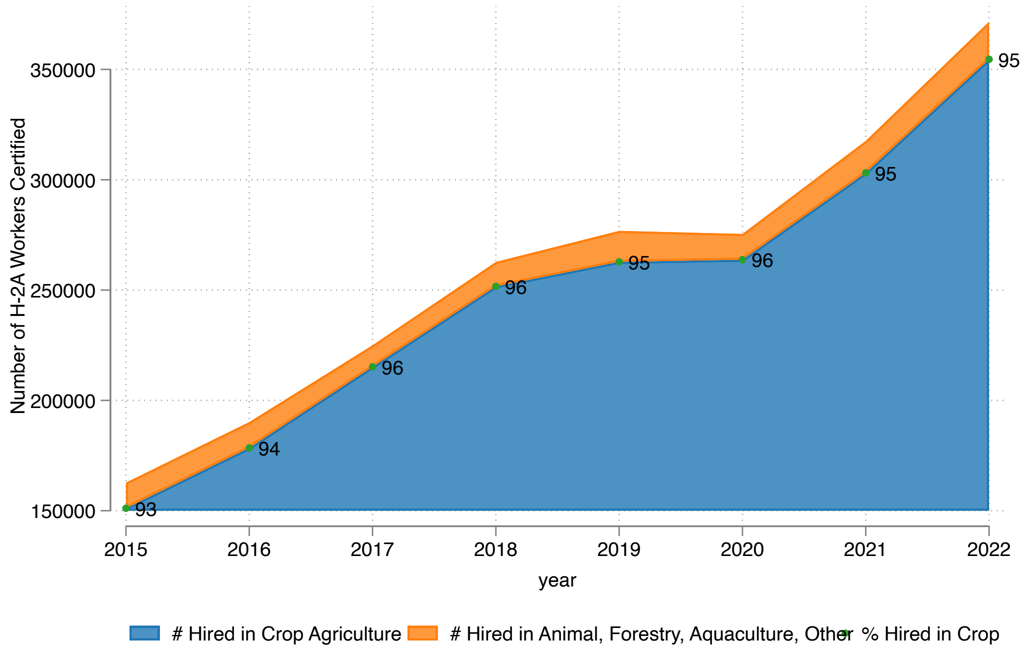


Figure 6. H-2A Data: Number of Workers Certified by Year, Crop versus Animal

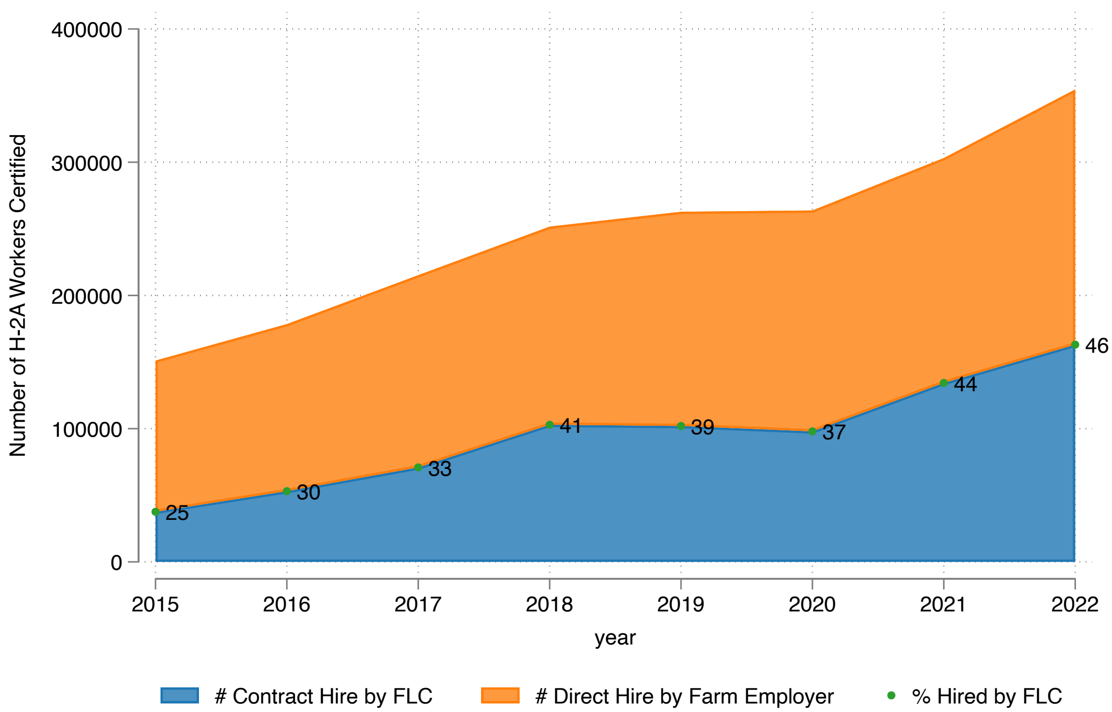
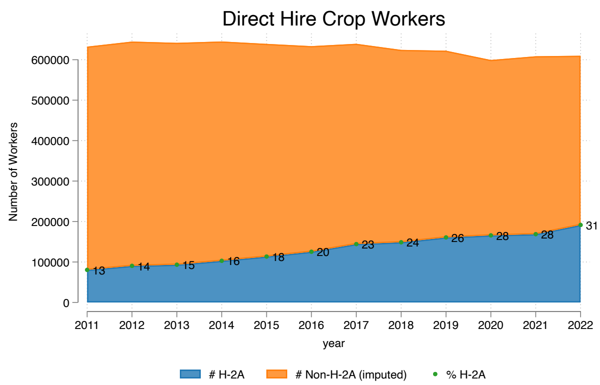
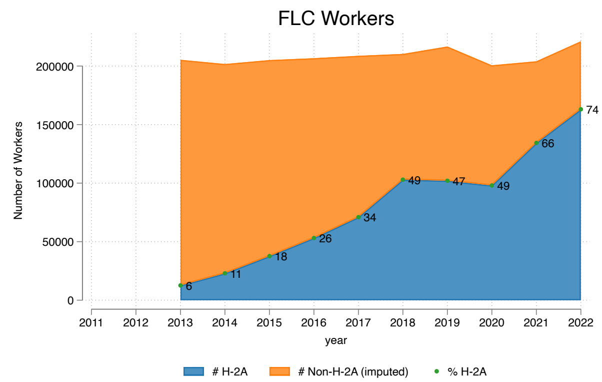


Figure 7. H-2A Data: Contract versus Direct Hire Crop H-2A Workers

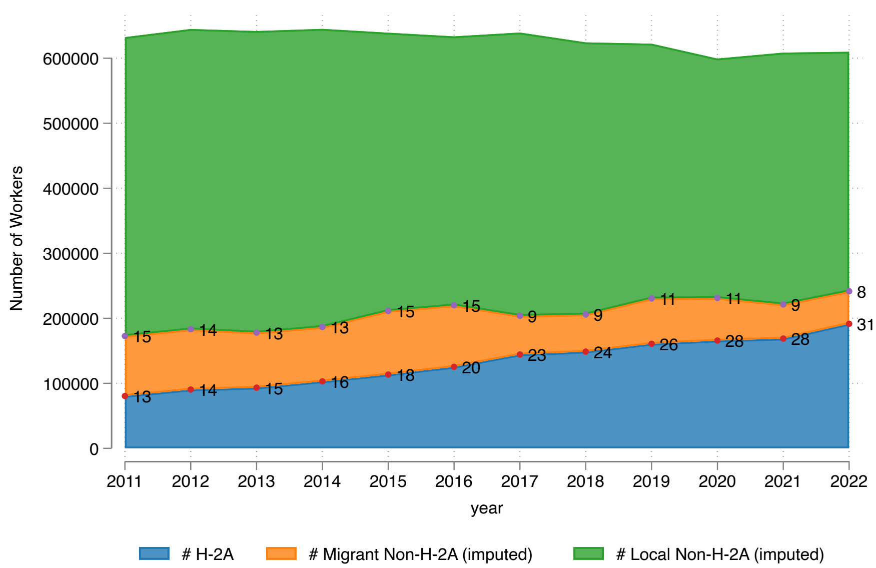
4.2 National estimates by worker type over time

Starting from the bottom of the tree, we have direct hire and contract hire H-2A workers. Subtracting these workers from the QCEW estimated employment counts (maybe adjust for known undercounts), we can impute the number of non-H-2A hired directly and hired by FLCs:





From here, we need to further disaggregate the non-H-2A direct hire workers into migrant and local workers. To do this, we apply the percentage of direct hire crop workers who identify as migrant from the NAWS



References

BLS (2023). QCEW Overview: Coverage. Washington, DC: Bureau of Labor Statistics. Retrieved from https://www.bls.gov/cew/overview.htm#coverage

BLS. (2022). QCEW Table A: Coverage exclusions in 2022 for selected workers. Washington, DC: Bureau of Labor Statistics. Retrieved from https://www.bls.gov/cew/publications/employment-and-wages-annual-averages/current/home.htm#exclusions

BLS. (2002). QCEW Table A: Unemployment Insurance Laws and Coverage. Washington, DC: Bureau of Labor Statistics. Retrieved from https://www.bls.gov/cew/publications/employment-and-wages-annual-averages/2002/home.htm#Coverage

Fisher, D. U., and Knutson, R. D. (2013). "Uniqueness of Agricultural Labor Markets." *American Journal of Agricultural Economics*, *95*(2), 463–469. http://www.jstor.org/stable/23358418

Fowler, C.S., and Jensen, L. (2020). "Bridging the Gap between Geographic Concept and the Data We Have: The Case of Labor Markets in the USA." *Environment and Planning A: Economy and Space 52(7)*: 1395–1414.

Fowler, C.S., Danielle C.R., and Jensen, L. (2016). "Reassessing and Revising Commuting Zones for 2010: History, Assessment, and Updates for U.S. 'Labor-Sheds' 1990–2010." *Population Research and Policy Review 35(2)*: 263–86. https://doi.org/10.1007/s11113-016-9386-0.

Gold, A., Fung, W., Gabbard, S., and Carroll, D. (2022). "Findings from the National Agricultural Workers Survey (NAWS) 2019-2020: A Demographic and Employment Profile of United States Farmworkers." JBS International, Research Report No. 16. Retrieved from: https://www.dol.gov/agencies/eta/national-agricultural-workers-survey/research

Heimlich, R. (2000). "USDA Farm Resource Regions." USDA Economic Research Service Agricultural Information Bulletin No. 760. Retrieved from: https://www.ers.usda.gov/publications/pub-details/?pubid=42299

Hill, A.E., Ornelas, I. and Taylor, E. (2021). "Agricultural Labor Supply." *The Annual Review of Resource Economics, 13*. DOI: 10.1146/annurev-resource-101620-080426

Ray, S., Hertel, T.W., Taylor, E., Hill, A.E., and Haqiqi, I. (2023). "Labor Markets: A Critical Link between Global-local Shocks and its Impact on Agriculture." *Environmental Research Letters, 18*: 035007. DOI: 10.1088/1748-9326/acb1c

Taylor, J.E. and Charlton, D. (2018). The Farm Labor Problem: A Global Perspective. Amsterdam: Elsevier Academic Press.

Online Appendix

Appendix A

The markets for 'local' workers are straightforward from the literature – we define these as commuting zones, which are representative of daily commuting patterns for U.S. workers employed across all sectors (cite CZ papers).[4] However, to the best of our knowledge, no prior work has attempted to define labor markets for migrant agricultural workers. We considered a variety of approaches for constructing these migrant labor sheds before arriving at the ERS Resource Region-based sheds used in the manuscript. This appendix documents our approaches and presents our justification for using the (slightly adapted) ERS Resource Regions instead of other possible delineations.

To begin, we considered four potential methods for constructing migrant labor sheds. For simplicity throughout this appendix we will refer to the methods as follows:

METHOD NUMBER	SHORT-HAND DESCRIPTION	FULL DESCRIPTION
1	Adapted ERS Resource Regions	10 Regions based on aggregating and disaggregating ERS Resource Regions.
2	North-South Delineations	4 regions consisting of states in North-South lines that we believe roughly approximate historical migration patterns.
3	Original ERS Resource Regions	The 9 ERS Resource Regions (link)
4	Commuting Zones (CZ)	We use the 2010 CZ delineations developed in Fowler, Rhubart, and Jensen (2016)

Detailed Descriptions of Labor Shed Delineations

ERS Resource Region-Based Labor Sheds

The USDA ERS Farm Resource Regions divide the U.S. into 9 distinct regions that reflect "geographical specialization in production of U.S. farm commodities" (https://www.ers.usda.gov/publications/pub-details/?pubid=42299). These resource regions group together counties with similarities in farm commodities (e.g., drawing distinctions between areas that primarily produce cattle and wheat and those producing

[4] In choosing these existing delineations as our definitions of local labor markets for agricultural workers, we did also attempt to re-implement the commuting zone methodology as described in XX, applied to a restricted subset of the ACS (the data used in constructing the original commuting zones) including only respondents who report working in agriculture. However, in the publicly available version of these data (accessed from iPUMS), the key variable for constructing commuting zones (the county FIPS code) is missing for the majority of respondents working in agriculture. ADD % of OBS MISSING FIPS. CHECK IF HOME OR WORK FIPS MISSING. COMPARISON TO MISSING FOR OVERALL SAMPLE?

fruits and vegetables) and farm types (e.g., family versus nonfamily farms, smaller versus larger farms, and relative farm production value). While the primary intent of these resource regions is to reflect farm supply-side similarities across the U.S., we hypothesized that these regions might also capture labor supply and demand similarities. In particular, because different farm commodities require different labor skillsets and experience (for example, row crops like corn and wheat require relatively few workers to operate machinery for harvest, whereas many fresh fruits still require workers who are able and willing to hand harvest), the ERS resource delineations should serve as a good starting point for delineating migratory flows for crop work.

However, one limitation of the ERS resource regions is that they do not capture seasonal variation in production in a manner useful for considering the behavior of agricultural workers. In particular, we would expect migrant, or follow-the-crop, workers to move primarily in south-to-north patterns over the course of the year, as this enables them to capitalize on the differences in harvesting times across the U.S. (whereas moving in East-West patterns would not make sense because harvest times would be similar). In light of this, we decided to use a data-driven approach to determine whether any of the ERS-delineations should be disaggregated or aggregated. These disaggregated and aggregated ERS farm resource regions are our first considered (and preferred) labor shed delineation, which we compare with three alterative delineations, including the “raw” or original farm resource regions.

Our primary consideration for determining how to perform these aggregations and disaggregations was whether and to what extent migrant workers reported working across and within each region. To do this, we used data from the restricted-access version of the National Agricultural Workers Survey (NAWS). In particular, we used information from the NAWS work-grid which includes information on farm and non-farm employment, as well as periods of unemployment for each interviewed worker for the prior year. Using these data, we define migrant workers as those who reported doing farm work in more than one county in the prior year. We restrict our sample to workers who meet this definition of migrant in all applications of the NAWS for constructing or evaluating migrant labor shed delineations. Our first consideration in deciding whether to disaggregate any farm resource regions was whether the region could capture vertical, rather than horizontal, movement. The most obvious candidate for disaggregation was the fruitful rim, which includes counties in Washington, Oregon, Idaho, California, Arizona, Texas, and Florida. In particular, we considered separating the fruitful rim into three regions to be more consistent with capturing vertical migrant flows. We separated this region into the western fruitful rim, consisting of all fruitful rim counties in Washington, Oregon, Idaho, California, and Arizona; the central fruitful rim, consisting of all fruitful rim counties in Texas, and the eastern fruitful rim, consisting of all fruitful rim counties in Florida.

We then construct heatmaps for each farm resource region and our disaggregated versions that show the density, or prevalence, of migrant workers working across these potential labor shed delineations. We construct these heatmaps using data from the NAWS workgrid. For each worker in the sample (the NAWS interviews between 1,500 to 3,600 workers each year), we observe the work history over the year prior to the interview, which includes information on the type (farm, non-farm, or a period of unemployment) and the location (county, if in the U.S.) of the job. To focus on migrant workers alone, we first reduce the observations in the workgrid to include only farm jobs (i.e., our heatmaps are specific to agricultural employment patterns and do not capture workers moving for nonfarm work, vacation, or other reasons). We then restrict the sample to include only workers who performed farmwork in more than one U.S. county within the past year. Thus our classification of "migrants" are workers who perform farmwork in multiple counties within the past year. We construct heatmaps for each potential region as follows: we remove duplicate observations of worker-county employment to capture the unique number of workers with jobs in each county; next, for each region, we reduce the sample to workers who reported performing farmwork within that region within the past year (including their current farm job that is the place of the interview); we then calculate the number of unique workers in the sample for the region, the proportion of those workers who worked in multiple counties *within* the potential labor shed delineation, and the proportion of those workers who worked in counties in other regions.

NAWS-Based Labor Sheds

Our third considered labor shed delineation comes from adapting the commuting zone (CZ) delineation approach detailed in Fowler, Rhubart, and Jensen (2016) to the NAWS data. In particular, we use a bottom-up, data-driven approach to define geographical labor shed areas suited for analysis of agricultural labor markets with finer resolution than heretofore possible. Fowler, Rhubart, and Jensen (2016) use the 5-year summary data from the American Community Survey (ACS) for 2006-2010 to replicate and update the USDA ERS CZ delineations last constructed in 2000. However, there are multiple limitations to relying on the ACS and these CZ delineations to define market areas for migrant agricultural workers. In particular, the CZ approach relies on information on respondents' place (county?) of residence and place of work. Thus, CZs capture within-day movement, or commuting, patterns. While these zones are useful in establishing the relevant labor markets for local workers, who transport daily from their place of residence to their place of work, they are less useful for capturing longer-run movement patters. For example, the CZ methodology would not capture follow-the-crop migration, wherein workers temporarily relocate for chunks of the year in accordance with agricultural employment opportunities. To address these and other shortcomings of the ACS, we modify the CZ methodology for use with the work history information provided in the NAWS workgrid.

Our modified measure of proportional flow can be written as $\frac{C_{ij}}{min(W_i,W_j)}$, where C_{ij} is the count of migrnat workers who worked in both commuting zone i and commuting zone j, and W_i and W_j are the number of (distinct) migrant workers who worked in CZs i and j, respectively, at any time in the previous year. This measure of proportional flow between two CZs is similar in construction to the measure from Fowler, Rhubart, and Jensen (2016) but differs in several key ways. Most importantly, rather than delineating labor market areas based on commuting distance from an individual's home to their work, we use the observed counties in which these distinct worker types have worked in the previous year. Secondly, while we innitially constructed measures of proportional flow across counties (i.e., where

$C_{ij}, W_i,$ and W_j are counts of workers within *counties* i and j), this level of granularity appeared to be asking too much of the NAWS data. While the NAWS is the only nationally representative survey of U.S. farmworkers, the sample size is small in general, but particularly in comparison to the ACS. Given the large data-needs of such a data-driven approach, we chose to use less granular geographic units than the CZ methodology. Finally, we also use data over a longer time span (the previous 20 years) than the CZ methodology (using 5 years of data) to increase the sample size. We then apply the clustering technique used in Fowler, Rhubart, and Jensen (2016) to identify the groups of counties with highest proportional flow according to each measure. The specific clustering technique we employ is called hierarchical agglomerative cluster analysis (Kaufman and Rousseeuw 2009), a widely used technique for grouping observations based on the linkage between them. This technique begins by assuming that each county belongs to its own cluster, then the two clusters (counties) with the strongest linkage (measured by proportional flow) are paired, the cluster linkages recalculated, and the next two clusters with the strongest linkage paired. The process continues until a stoppage point defined by the researcher is reached. Unlike the CZ methodology, which relies on the measures of proportional flow, alone, to determine clusters, we also include a matrix of indicator variables for CZ adjacency and CZs within the same broad geographic region (west, central, midwest, and east – as described in the following section). We included these additional criteria in the clustering to capture the stylized fact that migrant workers typically travel via ground transportation and are expected to follow tractable paths, rather than transporting accross the U.S.

At the outset, we hypothesized that this approach would yield more granular and accurate delineations of the areas of migrant labor mobility, but as we will demonstrate in our evaluation below, this approach did not perform well in comparison to our alternative (and simpler) approaches. We believe that this is largely due to the relatively small sample size in the NAWS. After restricting the sample to our definition of migrant workers (those perfmorming farmwork in more than one county in the prior year), our 20-year sample includes just XX workers, whereas the 5-year ACS that is used in constructing CZs includes more than XX respondents. In addition to adapting the CZ methodology to construct these NAWS-based alternative migrant labor shed delineations, we also consider the CZ delineations alone as a potential migrant labor shed. Despite their (logical) pitfalls for defining these zones, we were interested to document their performance relative to the alternatives. Not surprisingly, given their granularity, the CZ delineations performed poorly in terms of the density of workers working across delineations.

Annecdotal-Based Labor Sheds

Our final considered labor shed delineation comes from annecdotal evidence on historical migrant flows. Historically, U.S. migrant workers followed circular north-south paths across fairly large geographic regions in accordance with the seasonality of crop production activities, most often to be in a particular location at harvest time. These paths historically began in the south early in the year, moved northward during spring and summer, then headed south again, following crops and weather patterns that dictate harvest times (Arnedo, Rose, and Borges 2011; Taylor 1937). As an example, one well-documented historic route of agricultural laborers on the West coast was to begin the year harvesting in California's Imperial Valley, next to the Mexican border, move north to the Los Angeles or San Bernardino county areas in the spring, then move north to spend most of the summer and fall in central California near Bakersfield and Fresno, followed by returning to the south at the tail of the year (Taylor and Rowell 1938).

Our aim with these labor sheds was to be as encompassing as possible while still representing an improvement over treating the entire U.S. as a single labor market for migrant workers. As such, we chose to divide the U.S. into four north-south regions that should logically capture the bulk of migrant routes. Figure XX shows these delineations, which we refer to as the west, central, midwest, and eastern regions.

Among workers who reported working for 2 or more employers in the past year…

	Region			
	1 (Our Regions)	2 (Anecdotal North-South)	3 (ERS Resource)	4 (CZs)
Avg. # of jobs per worker within regions	1.852	2.013	1.728	1.211
% working multiple jobs in same region	83.04	89.19	77.69	55.71
% working in multiple regions w/in last year	33.47	22.35	43.76	82.91