Version 1.0.0 | Research paper

Author: Vojtech Royce

Who does the work?

Sex composition in eight U.S. occupational categories, 2025

Working paper 1.0.0. Author review pending. This manuscript has not undergone external peer review.

Abstract

Public debates about contribution often move directly from an occupation's workforce composition to claims about the importance or replaceability of an entire population. This paper establishes a smaller, reproducible result. Eight selected categories in the U.S. Bureau of Labor Statistics Current Population Survey show pronounced differences in the share of employed people reported as women. In 2025, women represented 3.5% of electricians and 87.3% of registered nurses. The all-employment benchmark was 47.1%. These estimates describe who held jobs, not how much service a worker produced, why a person entered an occupation, or what would happen after a workforce disruption. We preserve the source rows, distinguish aggregate categories from detailed occupations, and report the unusual eleven-month observation window. The result supplies an empirical starting point for research on occupational concentration while identifying the additional evidence required for a model of adaptation and service continuity.

The question

How uneven is the recorded sex composition of selected occupations associated with construction, transport, emergency response, and care? The question concerns observable employment patterns. It does not ask whether a group deserves more rights or whether an average describes every individual. Those are different claims with different evidentiary requirements.

The selection contains recognizable jobs from the motivating research discussion. It was not drawn randomly, and the observed values were known before this manuscript was written. Accordingly, this is an exploratory descriptive analysis. It cannot establish how unusual these occupations are among all occupations, and it does not test a preregistered population-wide hypothesis. Publishing that selection rule matters because a compelling set of examples can otherwise be mistaken for comprehensive coverage.

The useful substantive contrast is visible without a composite score. Several selected construction, emergency, and driving roles have a low recorded female share. Several selected care roles have a high recorded female share. Preserving both sets in the same display allows readers to inspect concentration across kinds of work without first accepting an author's valuation of them.

Data and population

The numerical source is BLS annual CPS Table 11, titled Employed people by detailed occupation, sex, race, and Hispanic or Latino ethnicity. The table was last modified on February 20, 2026 and was accessed for this release on September 7, 2026 UTC. Employment totals are in thousands of people; the women column is a percentage of each row's employed population. These are rounded survey estimates, not a register of every worker. BLS Table 11

The CPS employment population is civilian and noninstitutional, age 16 and over. Its employment concept includes employees, people working in their own businesses, qualifying unpaid family workers, and people temporarily absent from jobs. For people with multiple jobs, occupation refers to the main job, normally the one with the greatest usual hours. Volunteer work and work around one's own household do not qualify as employment by themselves. Thus the table does not measure all socially useful labor or every person capable of performing a task. BLS concepts and definitions

The 2025 observation window requires prominent disclosure. October household survey data were not collected because of the federal government shutdown. BLS therefore calculated eleven-month annual averages and states that they are not strictly comparable with annual averages for other years. We retain the 2025 label together with that caveat in downloads and visualizations. We do not fill October with September, create an artificial twelfth month, or interpret a difference from 2024 as a clean year-over-year change. BLS Table 11 notes

Results

Exact BLS row	Total employed, thousands	Women, %	Men, % derived
Construction and extraction occupations	8,490	4.3	95.7
Electricians	1,063	3.5	96.5
Firefighters	352	5.1	94.9
Driver/sales workers and truck drivers	3,583	7.7	92.3
Registered nurses	3,528	87.3	12.7
Home health aides	726	86.1	13.9
Nursing assistants	1,404	86.9	13.1
Childcare workers	1,106	93.2	6.8

All original totals and female percentages come from BLS Table 11. The displayed male shares are calculated as 100 minus the published female percentage, using the table's sex categories. They are not an additional independently estimated source column.

The construction and extraction row is an aggregate that contains electricians. It must not be added to the electrician row to obtain a total workforce. The remaining named rows are separate occupational categories, but this selected collection still does not define an

exhaustive sector. In particular, the driver row combines driver/sales workers with truck drivers. Relabeling it as truck drivers alone would claim a specificity the source does not provide.

Women were 47.1% of all employed people in the source table. Relative to that descriptive benchmark, the electrician share was 43.6 percentage points lower and the registered nurse share was 40.2 points higher. Relative to a 50% parity reference, those distances are 46.5 and 37.3 points. These are two different comparisons. Choosing 50% is a reference convention; choosing 47.1% compares the occupation with observed overall employment. Neither reference establishes how a fair labor market ought to allocate workers. BLS Table 11

The scale information also changes the interpretation. A large percentage in a small occupation is different from a similar percentage in an occupation employing millions. At the same time, a larger headcount is not automatically a larger contribution or greater operational criticality. A single task can be a bottleneck even when few people perform it, while a large workforce can supply services with varying urgency and substitutability. The table reports the headcount side of that distinction; it contains no measure of the bottleneck side.

What the arithmetic establishes

Our reproduction script reads the preserved source extraction and matches each exact row label. It checks published employment totals and female percentages, then derives male shares and benchmark differences. It also calculates approximate sex-specific headcounts by multiplying the rounded row total by the rounded percentage. Those derived counts are marked as approximations and are not presented as separately published BLS estimates.

No significance test is attached to this display. An occupational population estimate such as 1,063 thousand is not the survey sample size. Treating it as a binomial denominator would produce unjustifiably narrow uncertainty intervals. Appropriate inference would require the survey's variance methodology or suitable microdata and design information. This release reports point estimates and arithmetic differences, allowing a reader to reproduce those claims exactly without implying a precision that has not been calculated.

From jobs to infrastructure

CISA identifies sixteen critical infrastructure sectors and describes dependencies between their systems. That taxonomy motivates attention to continuity of essential services. It does not supply an official crosswalk from each CPS occupation to a critical facility, and it does not assign a universal social-value weight to an occupation. Our broad topic labels are editorial navigation categories. They are not CISA certifications. CISA critical infrastructure overview

An electrician may work on new residential construction, maintenance, or another setting; the occupation label alone does not reveal the urgency of the tasks performed during a particular disruption. Registered nurses also work across settings and specialties. Knowing the number of people employed in each category does not identify which shifts, facilities, or services would lose qualified staff. A continuity model would need an occupation-to-task-to-service mapping, time-specific staffing constraints, and explicit coverage of the systems it claims to represent.

Care deserves consistent measurement within that model. The fact that unpaid household care is outside this employment table is a boundary of the data, not evidence that its contribution is zero. If household care affects whether other workers can remain available, omitting it could alter a modeled response. We do not assign that effect a number here. A future study should measure it with an appropriate source and avoid counting the same activity twice across market and household categories.

Alternative explanations and next evidence

Observed concentration is compatible with multiple interacting mechanisms, including training histories, job requirements, preferences, recruitment, constraints, discrimination, and career transitions. Table 11 does not distinguish those mechanisms. It has no randomized intervention, comparable applicant pool, direct skill assessment, or counterfactual occupational choice. The same descriptive percentage should therefore remain unchanged regardless of which explanation a reader favors.

The next empirical step is a complete and auditable occupational inventory, with a documented inclusion rule established before comparing concentrations. A separate task-level dataset should record required credentials, demonstrated competencies, training durations, reserves, and service outputs. Analysis should distinguish replacing tasks already mastered from learning tasks under emergency conditions. A person's recorded sex should not substitute for their measured qualification or capacity.

Research on workforce shocks should then state who is absent, whether the population's demand changes, how long the shock lasts, and what assistance remains available. Losing labor while all residents remain is a different counterfactual from removing people and their demand. Those scenarios cannot be merged into one curve without changing the estimand.

Limits and reproducibility

This paper concerns the United States and one unusual annual window. It covers eight purposefully selected categories, includes an overlapping aggregate row, excludes much unpaid labor, and uses rounded survey estimates without calculated confidence intervals. It supplies no causal decomposition, innate-ability estimate, prediction of collapse, or ranking of human value. These limits define exactly which claims a replication can examine.

The accompanying package contains the exact web-retrieval extraction used to verify the rows, its checksum, structured observations, and a deterministic calculation script. Direct local downloads of the original BLS HTML and PDF returned access errors during preparation, so the preserved object is explicitly a web extraction rather than a claimed original file. The public BLS URL remains available for independent comparison. The analysis can be checked from this snapshot without network access, and any later source revision should create a new dataset version rather than silently replacing these observations.

The defensible empirical finding is substantial occupational concentration in different directions across the selected jobs. The more ambitious questions about substitution, service failure, and recovery remain separate research tasks. Their answers require additional data and an explicit model; employment composition is the beginning of that investigation.