Veteran Employment Outcomes

Joyce Hahn, Henry Hyatt, Hubert Janicki,
Erika McEntarfer, Seth Murray, and Lee Tucker

April 24, 2020*

1 Overview

The challenges faced by military veterans transitioning into the civilian labor force are a subject of ongoing concern to policymakers. The Census Bureau's Veteran Employment Outcomes (VEO) are experimental statistics on Army veterans' labor market outcomes one, five, and 10 years after discharge, by military occupation, rank, demographics (age, sex, race, ethnicity, education), industry and geography of employment. These statistics are generated by linking veteran records provided by the U.S. Army to national administrative data on jobs at the U.S. Census Bureau. Coverage of the data is all enlisted soldiers in the Army who completed their initial term of service and were discharged between 2000 and 2015 (about 650,000 veterans). Although VEO currently cover only Army veterans, these statistics could potentially be expanded to other service branches.

These new data highlight the broad distribution of labor market outcomes for recent Army veterans, highlighting the role of industry and military occupation in post-military earnings outcomes. Some key findings from these new statistics include the following:

1. Private sector earnings upon exit are generally highest for veterans who were employed in electrical equipment repair and intelligence gathering operations while in military service.

2. The federal government is consistently one of the better paying employers across most military occupations, with veterans having a much higher propensity to work in the federal government than the civilian workforce at large.

*This paper documents a new Census Bureau experimental statistical product. The Veterans Employment Outcomes described here have been approved for release by the Census Bureau Disclosure Review Board (Release Number: CBDRB-FY20-088). Any opinions and conclusions expressed in this document are those of the authors and do not represent the views of the U.S. Census Bureau.

3. Industry of employment is an important predictor of earnings outcomes even within broad military occupation groups.

4. Employment rates fell for veterans exiting during and after the Great Recession, likely reflecting poor labor market conditions as well as increases in educational benefits for veterans.[1]

The Census Bureau is releasing Veteran Employment Outcomes together with an online data exploration tool so that veterans, military personnel, and policy-makers can easily explore these new statistics. As with other Census Bureau experimental statistical products, we seek feedback from users and stakeholders on the quality and usefulness of these data. The VEO may be regularly updated or expanded if there appears to be sufficient user demand and resources permit. The remainder of this document provides documentation on how the VEO data were constructed and how veteran confidentiality is protected in the released statistics.

2 Data Sources and Coverage

2.1 Veterans

The source of veteran information in the VEO is administrative record data from the Department of the Army, Office of Economic and Manpower Analysis. This personnel data contains fields on service member characteristics, such as service start and end dates, occupation, pay grade, characteristics at entry (e.g. education and test scores), and demographic characteristics (e.g. sex, race and ethnicity). Once the Army service member record data is transferred to the Census Bureau, personally-identifying information is stripped from the record and it is assigned a Protected Identification Key (PIK) that allows for service members to be matched with their employment outcomes in Census Bureau jobs data.

The VEO cover all Army service members who have completed their initial term of service - meaning they have served the time they signed up for when they enlisted and were not discharged early - and were discharged between 2000 and 2015. We additionally restrict our sample to service members who have completed active-duty service as enlisted soldiers (not as officers or warrant officers) and have final ranks E1-E9. Discharge rates for more senior military personnel are not sufficiently large to produce VEO statistics without significant statistical noise to protect those records.

[1] The Post-9/11 GI bill, passed in 2009, significantly increased the amount of educational benefits available to veterans who served after 2001.

2.2 Employment and Earnings

For data on veterans' employment and earnings once they are discharged from the Army, we link veteran records to a national database of jobs that is part of the Longitudinal Employer-Household Dynamics (LEHD) data at the U.S. Census Bureau. The LEHD data consists of quarterly earnings records for individual workers that are submitted by employers for the administration of state unemployment insurance (UI) benefit programs. These records are then linked to establishment-level data collected for the Quarterly Census of Employment and Wages (QCEW) program. Currently, 48 states and the District of Columbia share UI and QCEW data with the LEHD program as part of the Local Employment Dynamics (LED) federal-state partnership. Their UI data covers 95% of private sector, state and local government employment.[2]

We supplement the LEHD data with federal employment records sourced from the Office of Personnel Management (OPM). Currently, OPM data is available from 2000 to 2015. Since federal employment is a significant source of labor earnings among former service members, VEO data availability is restricted to include only years in which federal record data is available.

2.2.1 Annual Earnings and Labor Market Attachment

A challenge when constructing earnings distributional measures from administrative data is the "fat left tail" i.e. the mass of individuals who mostly did not work during the year but held at least one short-duration job. These records include students and stay-at-home caregivers who have some earnings during the year but are not in the labor force full-time. So that our distributional measures reflect earnings outcomes for attached workers, we follow the typical approach of restricting earnings outcomes to those workers with evidence of strong work attachment during the year. Specifically, we only report earnings of veterans who i) work at least three quarters in the calendar year and ii) earn at least the equivalent of working full time at the federal minimum wage.[3]

Annual earnings are generated for workers who meet the criteria above by summing earnings for all jobs worked during the year. When reporting on employer characteristics, we use the characteristics of the employer the veteran received the highest earnings from in that year.

2.2.2 Employment Counts

Similarly, annual employment counts in VEO are counts of workers who met the labor force attachment criteria above in that year. We also release a count of the total number of workers who

[2] See Abowd et al. (2009) for details.

[3] This is the same labor force attachment restriction used in the Census Bureau's Post-Secondary Employment Outcomes, which uses LEHD data to track employment outcomes for college graduates.

did not meet that threshold. This group includes workers with zero earnings, very low earnings, and potentially workers whose earnings are not covered in the LEHD and OPM data.[4]

A more technical description of how the VEO measures are constructed is contained in the following section. Subsequent sections also provide more details on how veteran characteristics are defined including how we aggregated Military Occupation Speciality (MOS) codes for the VEO data. The last section provides details on the confidentiality protection system used to protect veteran privacy in the released statistics.

3 Concepts and Measures

3.1 Concepts

Our primary unit of analysis is an Army veteran denoted by subscript i in year t and quarter q.

Employment Status

We denote an earnings record for veteran i at employer j in year t and quarter q as w_{ijtq}. If a worker has positive earnings, we classify their employment status as employed; otherwise, they are considered non-employed. Employment status for worker i is therefore denoted by firm j in year t and quarter q as:

$$m_{ijtq} = \begin{cases} 1, & \text{if } w_{ijtq} > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Note that veterans may have earnings records with multiple employers in a given year and quarter. All earnings are adjusted for inflation using the CPI at a quarterly frequency.

Labor Market Attachment

Earnings measures can be sensitive to the inclusion of individuals who are loosely attached to the labor force. Since we do not have measures of hours worked that convey the degree of labor market attachment, we instead impose two sample restrictions when calculating moments of the earnings distribution. First, we require that individuals be employed at any combination of employers for at least three quarters in the calendar year $(m_{it}^{3+} = 1)$. Those employed for two or fewer quarters are denoted as $m_{it}^{3+} = 0$. Second, we require that individuals' total annual earnings, w_{it}, exceed a minimum earnings threshold. We set the threshold at the full-time earnings (35 hours times

[4]Notable coverage gaps in our data include: self-employed individuals, postal service employees, and many agricultural workers.

50 weeks, 12.5 weeks in each quarter) equivalent of the real federal minimum wage.[5] Since the minimum wage varies over our sample, we denote this threshold as $\underline{w_t}$ and require that $w_{it} \geq \underline{w_t}$.

Annual Earnings

Although earnings measures report veterans' real earnings across all employers in a given calendar year, we first generate worker i's annual earnings at firm j in year t by summing the worker's quarterly real earnings at the firm:

$$w_{ijt} = \sum_{q \in 1,2,3,4} w_{ijtq} \tag{2}$$

The worker's total annual earnings then aggregate the worker's real earnings across all employers in the calendar year:

$$w_{it} = \sum_{j \in J} w_{ijt} \tag{3}$$

where J denotes the set of all possible employers.

Dominant Job

When tabulating workers' earnings by employer characteristics, we focus on the characteristics of each worker's dominant job in the calendar year. The dominant job for worker i in year t is the job with the highest annual earnings:

$$d_{ijt} = \begin{cases} 1, \text{if } w_{ijt} > w_{ikt} \forall j \neq k \\ 0, \text{otherwise.} \end{cases} \tag{4}$$

Employer-reported industries can change within a given calendar year. To identify industry for the worker's dominant job in each year (employer j_{it}^{dom} where $d_{ijt} = 1$), we sum up the individual's total earnings at that employer by reported industry. Then, we assign the industry in which the individual earned the most in that job and year.

3.2 Measures

The goal of this data product is to release employment and earnings measures.

[5] All VEO tabulations are calculated and presented in 2018 constant dollars.

Employment Counts

We construct two distinct employment counts by strength of labor market attachment. First, the count of strongly attached workers, i.e. the count of all individuals who meet the two sample restrictions in the calendar year:

$$\sum_{(w_{it} \geq \underline{w_t}, m_{it}^{3+}=1, i \in I(m), j_{it}^{dom} \in J(s))} d_{ijt} \tag{5}$$

Second, we produce a count of all other veterans in our data that do not meet our labor market attachment criteria. This count includes loosely attached workers, i.e. the count of all individuals with non-zero earnings in the calendar year who fail to meet one or both of the sample restriction criteria (($m_{it}^{3+} = 0$ and/or $w_{it} \geq w_t$) and $w_{it} > 0$), and those with no observed earnings for the calendar year in the LEHD infrastructure files ($w_{it} = 0$). All employment measures are summarized in Table 1.

Table 1: Summary of Measures for Release

Measure Name	**Individual-level Description**	**Aggregated Description**
Annual Earnings Percentiles (25th, 50th, 75th)	$\sum_{j \in J} \sum_{q \in 1,2,3,4} w_{ijtq}$	$F(w_{it} \mid w_{it} \geq \underline{w_t}, m_{it}^{3+} = 1, i \in I(m), j_{it}^{dom} \in J(s))$
"Above-Threshold" Count	$\begin{cases} 1, \text{if } d_{ijt} = 1 \wedge w_{it} > \underline{w_t} \\ 0, \text{otherwise} \end{cases}$	$\sum_{(w_{it} \geq \underline{w_t}, m_{it}^{3+}=1, i \in I(m), j_{it}^{dom} \in J(s))} d_{ijt}$
Unobserved and "Below-Threshold" Count	$\begin{cases} 1, \text{if } m_{it}^{3+} = 0 \vee (d_{ijt} = 1, w_{it} \leq \underline{w_t}) \\ 0, \text{otherwise} \end{cases}$	$\sum_{i \in I(m)} (1 - m_{it}^{3+}) + \sum_{(w_{it} \leq \underline{w_t}, m_{it}^{3+}=1, i \in I(m), j_{it}^{dom} \in J(s))} d_{ijt}$

Earnings Percentiles

We report the 25th, 50th, and 75th percentiles from the earnings distribution of veterans who meet the two sample restrictions. We denote the distribution from which these percentiles are calculated as:

$$F(w_{it} \mid w_{it} \geq \underline{w_t}, m_{it}^{3+} = 1, i \in I(m), j_{it}^{dom} \in J(s)) \tag{6}$$

where $I(m)$ denotes the set of individuals with characteristic of interest m (e.g. sex, race, pay grade, military occupation, etc.), and $J(s)$ denotes the set of dominant employers with characteristic of interest s (e.g. state or industry).

Table 2: VEO Tabulations

Table Name	Separation Cohort	Characteristics
veoo3	8yr cohort	3-digit DOD Occupation Code
veoo2ns	8yr cohort	2-digit DOD Occupation Code by Sector
veoo2gs	8yr cohort	2-digit DOD Occupation Code by State
veoo2p	4yr cohort	2-digit DOD Occupation Code by Pay Grade
veogs	2yr cohort	State
veos	2yr cohort	Sex
veoe	2yr cohort	Education
veorh	2yr cohort	Race by Ethnicity
veop	2yr cohort	Pay Grade
veons	2yr cohort	Sector
veot	2yr cohort	AFQT Score Tercile
veoa	2yr cohort	Age
veox	2yr cohort	Experience

Note: Each tabulation is available for one, five, and ten years post separation.

4 Tabulations

Employment and earnings outcomes are tabulated by the following worker characteristics: separation cohort, sex, race and ethnicity, education at enlistment, AFQT score tercile, pay grade at separation, and military occupation. They are also tabulated by the state and industry of each worker's dominant employer. All outcomes are additionally reported one, five, and ten years post separation from the Army. All tabulations are detailed in Table 2.

Separation Cohort

In order for changes in veterans' labor market outcomes to be evaluated over time, we group veterans into cohorts based on their year of separation from active-duty service. For most tables, we use 2-year cohorts. However, for some tables, data quality constraints limit the extent to which we can provide detailed cohort groups. As a result, these tables are aggregated to 4-year and 8-year cohorts. Statistics for the following separation cohort bins are therefore generated:

- *Two year bins*: 2000-2001, 2002-2003, 2004-2005, 2006-2007, 2008-2009, 2010-2011, 2012-2013, and 2014-2015.
- *Four year bins*: 2000-2003, 2004-2007, 2008-2011, and 2012-2015.
- *Eight year bins*: 2000-2007 and 2008-2015.

Age, Sex, Race and Ethnicity

Worker demographic information allows for comparison of employment and earnings outcomes by age, sex, race, and ethnicity. Age is defined as age at separation from service and is taken from Army records. We then create two age groups: age <25 and age 25+. Sex, race, and ethnicity are taken from the LEHD data to allow for representation of race and ethnicity information in a manner that is consistent with the format used by the U.S. Census Bureau, presented in the LEHD schema available at `https://lehd.ces.census.gov/data/schema/latest/lehd_public_use_schema.html`. In a small number of cases when demographic characteristics are not available from the LEHD data, we use the variables from Army records and recode as necessary to maintain consistency. Sex has two bins: Male and Female. Race has six categories: White Alone, Black or African American alone, American Indian or Alaska Native Alone, Asian Alone, Native Hawaiian or Other Pacific Islander Alone, and Two or More Race Groups. Ethnicity has two bins: Hispanic and Not Hispanic.

Education at Enlistment

Eligibility for Army enlistment is dependent on meeting certain education thresholds. As a result, nearly all Army service-member records include their education level at time of enlistment. We use Army administrative data to generate three education-level categories: General Educational Development (GED) Test, High School Diploma, and Some College or Higher.

Armed Forces Qualification Test (AFQT) Score Tercile

All Army recruits take the Armed Forces Vocational Aptitude Battery (ASVAB), which is used to assess their skills at time-of-entry and their fit for any particular military occupation. A subset of the ASVAB is used to calculate the Armed Forces Qualification Test (AFQT) score. This score is reported as a percentile that is relative to a reference group and is comparable across time. We report employment and earnings outcomes by AFQT terciles, with 1 corresponding to scores 0-33, 2 to 34-66, and 3 to 67-100.

Pay Grade

We use pay grade at separation to capture each service member's performance during active-duty service. Due to sparsity of cells, some pay grade categories are aggregated into larger bins. Reported pay grade bins include: E1, E2, E3, E4, E5, E6, and E7-E9, with E1 being the pay grade for Privates and E7-E9 being the pay grades for senior non-commissioned officers (i.e. Sergeants

First Class, Master Sergeants or First Sergeants, Sergeant Majors, Command Sergeant Majors, or Sergeant Majors of the Army). Note that when labor market outcomes are reported by pay grade crossed with military occupation, we use a higher level of aggregation and report two grouped pay grade bins, E1-E5 and E6-E9.

Years of Service

We use three bins to capture the distribution of tenure for active-duty service at year of separation. They are: 0-5, 6-19, and 20+ years. Note that most enlisted service members serve less than five years and career personnel are eligible for retirement at 20 years of service.

Military Occupation

Occupation for enlisted personnel within the Army is defined by a Military Occupation Specialty (MOS) code. MOS code usage varies over time as new occupations are created and old ones are eliminated or reorganized. To account for these changes, we aggregate MOS occupation codes to the Department of Defense's Military Occupational Specialty Classification codes at the 2- and 3-digit level. The aggregation is done using the O*NET Military Crosswalk available at `www.onetcenter.org/dl_files/2010/military_crosswalk.zip`. A list of 2-digit DOD occupation codes is presented in Table 3. We do not track veterans who separated from service with an unassigned occupation group (i.e. 19X).

Table 3: Occupations

DOD Occupation Code	**Group Title**
10X	Infantry, Gun Crews, and Seamanship Specialists
11X	Electronic Equipment Repairers
12X	Communications and Intelligence Specialists
13X	Health Care Specialists
14X	Other Allied Professions
15X	Functional Support and Administration
16X	Electrical/Mechanical Equipment Repairers
17X	Craftsworkers
18X	Service and Supply Handlers

The corresponding 3-digit DOD occupation codes are listed in Table 4.

Table 4: Occupations

DOD Occupation Code	Title
101	Infantry, General
102	Armor and Amphibious, General
103	Combat Engineering, General
104	Missile Artillery, Operating Crew
106	Small Boat Operators
108	Unmanned Vehicle System (UVS) Operators, General
110	Navigation, Communication, and Countermeasure, N.E.C.
111	Shipboard and Other Fire Control
112	Missile Guidance and Control
115	ADP Computers, General
116	Teletype and Cryptographic Equipment, General
119	Electronic Instruments, N.E.C.
120	Radio Operators
122	Air Traffic Control
123	Intercept Operators (Code and Non-Code)
124	Operational Intelligence
125	Combat Operations Control, General
126	Communications Center Operations,General
127	Cyberspace Operations, General
130	Medical Care and Treatment, General
131	Radiology
132	Diet Therapy
133	Dental Care, General
134	Medical Administration
140	Photography, General
141	Surveying
142	Weather, General
143	EOD/UDT
145	Musicians, General
149	Firefighting and Damage Control
150	Recruiting and Counseling
151	Legal

Continued on next page

Table 4: Occupations

DOD Occupation Code	Title
152	Combined Personnel and Administration, General
153	Operators/Analysts/Programmers
154	Disbursing
155	Supply Administration
156	Chaplain's Assistants
157	Information and Education, General
160	Aircraft, General
161	Automotive, General
162	Lineman/Central Office
164	Aviation Ordinance
165	Auxiliaries
166	Electric Power
169	Other Mechanical and Electrical Equipment, General
170	Metal Body Repair
171	Construction Equipment Operation
172	Utilities, General
174	Lithography, General
176	Fabric, Leather, and Rubber, General
180	Food Service, General
181	Motor Vehicle Operators
182	Missile Fuel and Petroleum
183	Law Enforcement, General
184	Laundry and Personal Service, General
186	Forward Area Equipment Support, General

Employer Geography

Employment and earnings outcomes are available for each of the 50 states and the District of Columbia. A worker is assigned to a given state if their dominant employer for the calendar year paid UI compensation for that worker in that state. For federal employees, we use the location of the government agency to establish state. States are identified by their Federal Information Processing Standard (FIPS) state code.

Employer Industry

We provide statistics by the industry of the dominant employer, at the North American Industrial Classification System (NAICS) Sector level with the federal government added as a separate category. This results in the 21 categories presented in Table 5.

Table 5: Industry

2-digit NAICS	NAICS Sector
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support of Waste Management and Remediation
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other
92	Public Administration
99	Federal Government

5 Confidentiality Protection Procedures

We implement a set of confidentiality protection procedures that follow those used for the Post-Secondary Employment Outcomes (PSEO) data product.[6] All employment counts are protected by geometric noise infusion while all earnings percentiles use geometric noise infusion in the histogram approach. We summarize implementation of these protection procedures below.

[6] PSEO protections are described in Foote, Machanavajjhala, and McKinney (2019).

5.1 Disclosure Protection of Employment Counts

Employment counts are a set of count queries that we protect using a geometric mechanism for noise.

Consider a count ($n_{m,s,t,\tau}$) by worker characteristics (e.g. cohort, sex, occupation, etc) m, in state and industry s, in table t, τ years after separation. To protect the count, we use the geometric noise mechanism to add noise such that the protected count is:

$$\tilde{n}_{m,s,t,\tau} = n_{m,s,t,\tau} + \eta \tag{7}$$

where $\eta \sim X - Y$, where $X,Y \sim Geometric(p)$ where $p = 1 - \frac{1}{e_t^{\varepsilon}}$. Therefore, $\tilde{n}_{m,s,t,\tau}$ is ε-differentially private.

5.2 Disclosure Protection of Earnings Percentiles

To protect earnings percentiles, we use noise infusion in the histogram approach. Consider the conditional earnings distribution described in Equation (6). We can approximate it by a histogram defined over H bins, such that all observed earnings fall into one of H intervals. These can be denoted $[b_1, b_2), ..., [b_h, b_{h+1}), ..., [b_{H-1} b_H), [b_H, \infty)$. We choose to have an earnings histogram with 21 bins, where the minimum and maximum values are set to \$10,000 and the 99.9th percentile of the earnings distribution, respectively. Each bin represents a 5th percentile increment of a publicly-available lognormal earnings distribution. We base the mean and variance parameters on the earnings distribution of veterans found in the ACS.[7] The corresponding bins are presented in Table 6.

We also include a bin for workers with zero earnings, or positive earnings but less than the earnings threshold. The count of observations in each bin h is protected by adding noise,

$$\tilde{n}_{h,m,s,t,\tau} = n_{h,m,s,t,\tau} + \eta \tag{8}$$

where $n_{h,m,s,t,\tau}$ is a count of workers in bin h with characteristics (e.g. cohort, sex, occupation, etc) m, in state and industry s, in table t, τ years after separation and $\eta \sim X - Y$, where $X,Y \sim Geometric(p)$ where $p = 1 - \frac{1}{e_t^{\varepsilon}}$. Therefore, $\tilde{n}_{h,m,s,t,\tau}$ is ε-differentially private.

[7] These parameters are estimated using the 2010-2014 5-year ACS Public-Use Microsample for veterans who were on active duty in 1990 onward and reported at least \$10,000 in real earnings. The mean is 10.7814 and standard deviation is 0.71134.

Table 6: ACS Veteran Earnings Histogram

Bin	Lower Bound	Upper Bound
1	10000	14933
2	14933	19337
3	19337	23021
4	23021	26442
5	26442	29780
6	29780	33136
7	33136	36582
8	36582	40182
9	40182	44003
10	44003	48117
11	48117	52617
12	52617	57619
13	57619	63291
14	63291	69872
15	69872	77745
16	77745	87560
17	87560	100575
18	100575	119733
19	119733	155042
20	155042	193998
21	193998	433482

Due to the composition properties of differential privacy established in Hay et al. (2010), the entire histogram is also ε-differentially private. To evaluate the extent to which a change in ε_t impacts the accuracy of the counts associated with the observed empirical distribution approximated by our histogram, we create the following accuracy measure:

$$CA_t = 1 - \frac{1}{100}\sum_{\eta}\frac{1}{2N_t}\sum_{h,m,s,\tau}|\tilde{n}_{h,m,s,t,\tau} - n_{h,m,s,t,\tau}| \tag{9}$$

We define CA_t as the average accuracy of all cells*bins combinations of histograms underlying table t, where the average is defined over 100 draws of η. $N_t = \sum_{h,m,s,\tau} n_{h,m,s,t,\tau}$ is the number of workers contributing to the table. These workers are defined over all possible combinations of (h,m,s,τ) for a table t.

From Equation (8), we can construct an ε-differentially private empirical CDF, from which we construct the percentiles of interest. The empirical CDF is defined as:

$$\tilde{F}(h|m,s,t,\tau) = \frac{\sum_{l=1}^{h} \tilde{n}_{l,m,s,t,\tau}}{\sum_{l=1}^{H} \tilde{n}_{l,m,s,t,\tau}} \tag{10}$$

To construct the percentile of interest from $\tilde{F}(h|m,s,t,\tau)$, we assume earnings are uniformly distributed within a bin. We then use linear interpolation to calculate earnings at any percentile between two estimated bins. The counts associated with the earnings percentiles are then defined over all of the bins, i.e. $\sum_{h=1}^{H} \tilde{n}_{h,m,s,t,\tau}$.

For an approved privacy budget B, we allocate it across tables following the social choice framework detailed in Abowd and Schmutte (2019). Consider the optimal disclosure protection ε-assignment problem for a set of T tables, each with K_t cells, that contains three percentiles (25th, 50th, and 75th). Each table represents a query on the same private database. To evaluate the accuracy of our B-differentially private data product, we consider the average accuracy of I draws of each cell, summed over all tables and percentiles of interest. Formally, this problem can be expressed as:

$$v(B) = \max_{\varepsilon_t} \left\{ \sum_{q \in \{25,50,75\}} \sum_{t=1}^{T} \frac{1}{K_t} \sum_{k=1}^{K_t} \frac{1}{I} \sum_{i=1}^{I} \left[1 - \frac{|m_{t,i,k,q}^{p}(\varepsilon_t) - m_{t,k,q}^{d}|}{m_{t,k,q}^{d}} \right] \right\} \tag{11}$$
$$\text{subject to:} \sum_{t=1}^{T} \varepsilon_t \leq B$$

where $v(B)$ represents the total average accuracy across all tables as a function of the underlying privacy budget B. Let $m_{t,i,k,q}^{p}$ and $m_{t,k,q}^{d}$ represent the protected and true data moments, respectively. For any approved privacy budget B, we can recover the optimal ε_t for each table t that maximizes our accuracy function.

The degree of suppression of small (protected) cells increases overall accuracy for any value of the privacy budget. Our baseline suppression value is 50, which is slightly higher than the value used in the PSEO data product. The ε values for each individual table have been identified by numerically solving Equation (11).

6 Output

The final data product is composed of 13 files, one for each of the 13 tables described in Table 2. These files contain the following:

- Relevant identifiers

- Protected employment counts of veterans who are above the threshold in one, five, and ten calendar years after separation from active-duty (denoted *y1_emp*, *y5_emp*, and *y10_emp*)
- Protected employment counts of veterans who are below the threshold or unobserved in one, five, and ten calendar years after separation from active-duty (denoted *y1_nonemp*, *y5_nonemp*, and *y10_nonemp*).[8]
- Protected employment counts of veterans who are above the threshold are accompanied by three earnings percentiles in one, five, and ten calendar years after separation from active-duty:
  - P25 (denoted *y1_p25_earn*, *y5_p25_earn*, and *y10_p25_earn*)
  - P50 (denoted *y1_p50_earn*, *y5_p50_earn*, and *y10_p50_earn*)
  - P75 (denoted *y1_p75_earn*, *y5_p75_earn*, and *y10_p75_earn*)
- Nine status flags for employment and earnings measures noting suppressions or other data limitations in one, five, and ten calendar years after separation from active-duty (denoted *status_y1_emp*, *status_y5_emp*, and *status_y10_emp*; *status_y1_nonemp*, *status_y5_nonemp*, and *status_y10_nonemp*; and *status_y1_earn*, *status_y5_earn*, and *status_y10_earn*). For example, some states enter the LED partnership after our earliest cohorts separate from the Army. As a result, there are state-specific restrictions in the reporting of employment counts and earnings percentiles. Possible values are detailed in Table 7.

Table 7: Status Flags

Status Flag	**Label**
-1	Data not available to compute this estimate. For VEO, this is typically due to lack of earnings data for later separation cohorts or lack of state data for earlier separation cohorts
1	OK
5	Value suppressed because it does not meet U.S. Census Bureau publication standards. For VEO, this is typically due to small protected counts.

References

[1] Abowd, John and Stephens, Bryce and Vilhuber, Lars and Anderson, Fredrik, and McKinney, Kevin and Woodock, Simon (2009) "The LEHD Infastructure Files and the Creation of the

[8] These counts are omitted from files that stratify workers by employer characteristics: state and industry.

Quarterly Workforce Indicators," *NBER Chapters in: Producer Dynamics: New Evidence from Micro Data*, pp. 149-230.

[2] Abowd, John M., and Ian M. Schmutte (2019) "An Economic Analysis of Privacy Protection and Statistical Accuracy as Social Choices." *American Economic Review*, 109 (1): pp. 171-202.

[3] Foote, Andrew David, Ashwin Machanavajjhala, and Kevin McKinney. 2019. "Releasing Earnings Distributions Using Differential Privacy: Disclosure Avoidance System For Post-Secondary Employment Outcomes (PSEO)". *Journal of Privacy and Confidentiality* 9 (2).

[4] Hay, Michael and Rastogi, Vibhor and Miklau, Gerome and Suciu, Dan (2010) "Boosting the Accuracy of Differentially Private Histograms Through Consistency," *Proc. VLDB Endow.*, 3(1-2), pp. 1021–1032