Selecting an Alternative Sample for the Occupational Employment and Wage Statistics Survey

David Piccone, Sertan Akinci and Ethan Ertel[1]
The U.S. Bureau of Labor Statistics, Washington D.C.

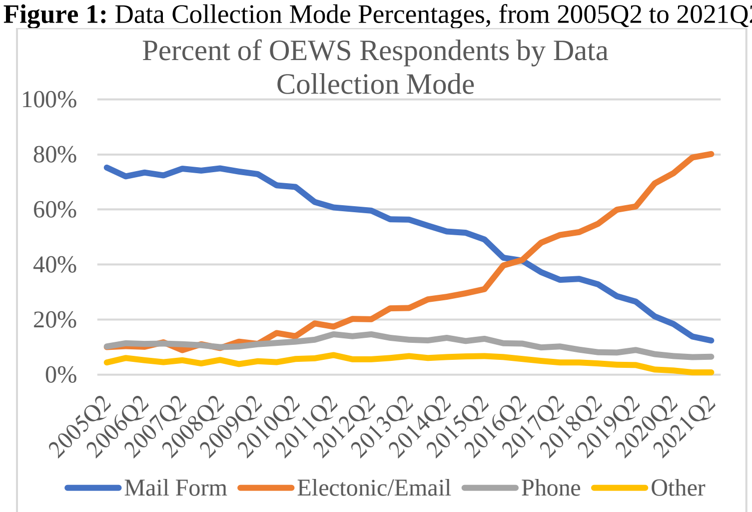
Abstract
In 2014 the Occupational Employment and Wage Statistics (OEWS) program at the Bureau of Labor Statistics (BLS) performed a feasibility study to determine if respondents to their survey were willing to provide additional data items. The results from this study were promising, leading to a follow-up test that began in 2022. For the follow-up test, the program wanted a sample that would allow the participating states to create annual estimates for their additional data items. This requires an alternative sample design since the current OEWS sample does not allow for annual estimates. A major requirement of the alternative sample is that it must maximally overlap with the official OEWS sample, to keep the cost of collecting additional units down. Another requirement is that the alternative sample must be selected in two bi-annual panels which correspond to the official OEWS sample collection periods. This results in about half of the sample being collected approximately six months after selection, requiring updating procedures for the alternative sample. In this paper, we describe the details of designing this alternative sample and its updating procedures for the OEWS program.

Keywords: sample design, sample overlap, power Neyman allocation, Neyman allocation, establishment survey, PPS sampling

1. Introduction

The Occupational Employment and Wage Statistics (OEWS) survey, conducted by the Bureau of Labor Statistics (BLS) provides comprehensive and reliable estimates on employment and wages across various occupations, industries, and detailed geographical areas. The survey employs a large sample that selects business establishments across all 50 United States, the District of Columbia, Puerto Rico, the US Virgin Islands, and Guam. The survey was initially designed to utilize a paper survey form mailed to sampled establishments asking for workers' occupation and wage. Over time, the OEWS program allowed respondents to submit their survey form by phone, electronically or by some other method such as fax or DVD/CD. Table 1 shows the percentage of OEWS data that is collected by the different modes from 2005Q2 to 2021Q2. Over these 16 years, the preferred data collection mode has shifted from the mailed form to electronic submission.

Figure 1: Data Collection Mode Percentages, from 2005Q2 to 2021Q2



[1] Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

Many of these electronically submitted responses come from reports produced by payroll processing software, which often includes information not solicited by OEWS, such as gender, hire date, Fair Labor Standards Act (FLSA) status, and hours worked. In 2015, the OEWS program conducted a multi-stage research project to explore what data items are commonly found within these payroll reports and if respondents are willing to provide them. The results from this research were promising, finding that OEWS could potentially collect five additional data items: hours paid, part- vs. full-time status, hire date, gender, and birth year of the worker. While it is feasible to collect additional data items, the research found that there will be cases that require extensive follow-up (Martinelli, 2015).

In March 2022, OEWS received funding to further test the collection of the additional data items. States could volunteer to have their OEWS sample augmented to create a yearly representative sample able to produce annual occupational estimates for the additional data items. In this paper, we will discuss how we designed the test sample for the states participating in the test.

2. Current OEWS Sample Design

2.1 Three-year Sample Rotation

The OEWS program produces occupational employment and wage estimates for very detailed industry and geography domains. To have adequate coverage, the OEWS survey requires a large sample of establishments. To collect such a large sample, the program implemented a three-year survey cycle, where approximately 187,000 establishments are sampled bi-annually, once in November and once in May, creating a yearly sample of about 374,000 establishments. Each bi-annual sample will be referred to as a panel sample. Any establishments selected in the previous two years are excluded from being selected into the current annual sample. For estimates, the current annual sample is combined with the previous two years of sample, to create a representative sample of 1.1 million establishments. While this achieves a large sample for estimates, it has the major drawback of preventing annual time series estimates. For any given year, about 748,000 establishments on the sampling frame are ineligible for selection due to being selected into the previous two years. This results in an annual sample that is not representative of the population and inadequate for calculating yearly time series estimates (US BLS, 2023).

2.2 Frame Construction

The OEWS sampling frame is primarily created from administrative information found on unemployment insurance (UI) reports compiled by the state workforce agencies. Since employers are required by law to participate in their state's UI program, these reports contain information on nearly every establishment in the nation. Once a quarter, all the states' UI reports are combined and cleaned to create a national sampling frame called the Quarterly Census of Employment and Wages (QCEW). In May and November of each year, OEWS extracts a list of in-scope establishments from the QCEW for its sampling frame. This is supplemented using auxiliary lists of establishments in the railroad industry and Guam, which are not covered by the UI program. In 2022, there were approximately 8.3 million in-scope establishments on the OEWS sampling frame.

2.3 Certainty Units

There is a subset of frame establishments that the OEWS program deems important enough to their estimates that they are included with certainty in the full, three-year sample. These are the OEWS certainty units that have a probability of selection of one over a three-year period. The exact rules for identifying the certainty units are confidential. Once the certainty units are identified, they are subtracted from the frame, so that the non-certainty units can be allocated to the sampling strata.

2.4 Stratification and Allocation

The OEWS program uses stratified sampling, where strata are defined by state, area, industry, and ownership for educational units. Area is defined by metropolitan statistical areas (MSAs) or balance of state (BOS) areas, industry is defined by 4-digit North American Industry Classification System (NAICS) codes, and ownership is defined by whether an establishment is a private or government unit. Strata are defined so that establishments within strata have homogeneous occupational staffing patterns (i.e., the occupations found at the establishment) and wages. Stratified sampling allows for a more efficient sample to be drawn by minimizing the variance of responses within strata, thus reducing the number of sample units needed for each stratum. Another benefit of stratified sampling is that it allows for an increase in precision for the OEWS estimates when used in conjunction with an optimal sample allocation.

The OEWS program uses two allocations to assign a full three-year, non-certainty sample (i.e., 1.1 million sample units minus the number of certainty units) to strata. This is done every May and November. Once the full sample is allocated, it is divided by six to get the one panel sample allocation. The OEWS uses an optimal and minimum allocation.

The optimal allocation is the power Neyman allocation, which was first introduced by Bankier (1988). This allocation assigns sample proportionate to the product of the strata size raised to power p, and the occupational variability within the strata. It is important to note that if $p = 1$ the power Neyman allocation is equal to the Neyman allocation, which would provide the most precise national OEWS estimates. Sub-national estimates are also important to the OEWS program, so they opt to use a value of ½ for p, to spread sample from the largest strata to mid-sized and smaller strata. Below is the formula for the power Neyman allocation.

$$n_h^{eff} = n \frac{\sqrt{X_h} S_h}{\sum_{all\ h} \left(\sqrt{X_h} S_h \right)} \tag{1}$$

where,

n_h^{eff} = the amount of non-certainty sample allocated to stratum h (state by MSA/BOS by NAICS4 by educational ownership) using the efficient allocation
n = the national 3-year sample size minus the number of certainty units
X_h = the number of non-certainty employees in stratum h
S_h = the measure of occupational variability within stratum h

The minimum allocation ensures every stratum has a minimum amount of sample for estimates. This allocation is a function of the number of non-certainty frame units in the strata:

$$n_h^{min} = \begin{cases} N_h & \text{if } N_h \leq 3 \\ 3 & \text{if } 4 \leq N_h \leq 11 \\ 6 & \text{if } N_h \geq 12 \end{cases} \tag{2}$$

where,

n_h^{min} = the minimum allocation for stratum h (state by MSA/BOS by NAICS4 by educational ownership)
N_h = the number of non-certainty frame establishments in stratum h

The final amount of non-certainty sample allocated for each stratum, n_h, is the maximum of the minimum and power Neyman allocations. The national non-certainty sample size used in formula 1 is iteratively changed until the final amount of sample allocated, which after reconciling the two different allocations, is about 1.1 million minus the number of certainty units. The last step of the OEWS allocation is to divide each stratum allocation amount by six, to get the final non-certainty allocation for the bi-annual sample: $n_h^{pan} = n_h/6$

2.5 Sample Selection

Once the sample is allocated, a bi-annual non-certainty sample is selected using a probability proportionate to size (PPS) scheme. The measure of size assigned to each establishment on the frame is the mean employment level for each state by size class[2] cell that the establishment belongs to. This causes all establishments within the same stratum and size class to have equal probability of selection, allowing for simple random sampling (SRS) at that level. Under SRS, OEWS can minimize the sample overlap with other establishment surveys at the BLS by using permanent random numbers (PRNs). PRNs are random numbers assigned to all establishments on the QCEW that remain unchanged for the duration that the establishment is in business. Since the PRNs are randomly

[2] Size classes are categories that are used to group establishments with a similar number of employees together. The OEWS based their size class definitions on the OMB standard (OMB, 1982).

generated, the OEWS and other BLS establishment surveys can use them for selecting their samples. In OEWS, within each stratum and size class cell, establishments are sorted by PRNs and the first $n_{h,s}^{pan}$ establishments are selected from some starting point between 0 and 1, where $n_{h,s}^{pan}$ equals the number of non-certainty sample units allocated to stratum h, and size class s cell. To minimize the sample overlap, each establishment survey at BLS targets a different part of the PRN range.

It should be noted that for each bi-annual sample, all certainty units on the frame that are eligible for selection (i.e., was not selected in the previous five bi-annual samples), are selected.

3. Alternative OEWS Sample Design

OEWS received funding for a single fiscal year to test the collection of the additional data items. Since the current sample design requires three-years' worth of sample to produce unbiased estimates, the OEWS managers decided to use an alternative sample for the test. The main requirement of the alternative sample design is that it is representative of the population using only one year's worth of sample. This will allow for unbiased estimates of the additional data items after one year of collection. We achieve this by simply allowing the previously sampled establishments to be eligible for selection in the current year. In other words, every establishment on the frame has a positive probability of selection. This means that some establishments selected into the alternative sample could have been selected in the previous two years by the official OEWS sample. The sample size for the alternative sample is equal to 374,000, which is the same as one year's worth of the official OEWS sample.

Another requirement for the alternative sample is that it must be designed to create the best occupational employment and wage estimates for the domains that OEWS currently use. While the goal of the alternative sample is to test collecting and estimating additional data items, we will not attempt to design the sample to create best estimates for any of the additional data items. There are two reasons for this. First, we currently do not have any information on how establishments will report on the additional data items. This information is key when designing a sample to optimally estimate a data item since it will directly affect how the sample is allocated and selected. For example, to improve the estimates of a particular data item, an increased number of sample units would be allocated to parts of the frame where that data item is more variable and harder to estimate. Since we currently have no information on how the data items are distributed across the frame, we could only guess how to allocate the sample to improve the estimates which may or may not actually improve them. Second, and more importantly, the main goal of the OEWS is to provide detailed information on occupational employment and wages; collection of the additional data items is secondary to this purpose.

The last requirement for the alternative sample is that it maximally overlaps with the official OEWS sample to keep the cost of collection low. To facilitate this, we must select the alternative sample in the May/November time periods of the official OEWS sample. The overlapping sample will be used for the official OEWS estimates, along with the test estimates for the additional data items. Later in this section, we will discuss the methods we used to increase the sample overlap between the official and alternative samples.

3.1 Certainty Units

Unlike the official sample, the alternative sample does not have any criteria for identifying certainty units. Instead, we allow the sample design to determine which establishments are to be selected into the sample with a probability of one. We will discuss this more in the section below on selecting the alternative sample.

3.2 Stratification

The official OEWS sample has about 145,000 strata, with the mean number of establishments within each stratum around 60. These are very detailed strata which require a large sample for proper coverage. This was the catalyst for the three-year sample rotation which allows OEWS to select 1.1 million establishments over three years. This works out to be around 7.5 sample units per stratum for the official sample. Since the alternative sample is one-third the size of the official sample, there would be too few sample units for proper coverage if we used the same strata definition as the official sample. For that reason, we decided to collapse the geography dimension of the strata definition, keeping the rest of the definition the same. Instead of stratifying by every MSA or BOS area, we collapse similar areas within states into what we call aggregate areas. Aggregate areas comprise of MSAs that are geographically close to one another and are formed so that each aggregate area accounts for roughly the same proportion of total employment within its respective state. All BOS areas are put into a single aggregate area. For most states, large city MSAs are not aggregated with any other areas.

By using aggregate areas, the number of strata falls to around 49,000, with a mean of about 174 establishments per stratum. Like before, this averages out to be around 7.5 sample units per stratum for the alternative sample. Table 1 below shows the difference in the stratum size distribution of employment and establishments between the current and alternative samples.

Table 1: Size Distribution of Official vs. Alternative Strata in 2022

	Official Stratification		**Alternative Stratification**	
Total Strata	144,973		49,383	
	Employment	**Establishments**	**Employment**	**Establishments**
Mean	927	59.1	2,722	173.5
10th Percentile	7	1	9	1
25th Percentile	32	3	157	10
Median	133	10	647	41
75th Percentile	495	33	2,120	134
90th Percentile	1,644	100	5,909	372

We can see from Table 1 that the size of the strata in terms of employment and number of establishments grows substantially under the alternative stratification definition.

The alternative stratification plan will not guarantee the same level of sample coverage for MSA and BOS areas as the current design. The alternative stratification allows for precise design-based estimates for national, state, aggregate area and detailed industry estimates but will not guarantee the same level of precision for MSA and BOS area estimates. The OEWS estimation methodology currently uses a modeling approach to predict occupational information for all establishments on the frame, which is then aggregated to calculate estimates. This approach will help with calculating MSA and BOS area estimates using the alternative sample. We decided to collapse the stratification on the MSA and BOS area dimension, but not on the industry dimension, since industry is a stronger predictor of what occupations are found in an establishment, and thus more important for the modeling approach used for estimates.

3.3 Allocation

Like the current design, we use a two-allocation approach for the alternative sample design that uses an efficient and a minimum allocation. We decided to test two efficient allocations, the power Neyman and Neyman allocation.

$$n_h^{pn} = n\frac{\sqrt{X_h}S_h}{\sum_{all\ h}\left(\sqrt{X_h}S_h\right)} \tag{3}$$

$$n_h^{ney} = n\frac{X_h S_h}{\sum_{all\ h}(X_h S_h)} \tag{4}$$

where,

n_h^{pn} = the amount of sample allocated to stratum h (state by aggregate area by NAICS4 by educational ownership) using the power Neyman allocation

n_h^{ney} = the amount of sample allocated to stratum h (state by aggregate area by NAICS4 by educational ownership) using the Neyman allocation

n = the national 1-year sample size

X_h = the number of non-certainty employees in stratum h

S_h = the measure of occupational employment variability within stratum h

The goal of testing both allocations is to quantify the tradeoff between the precision of the national and sub-national estimates when using each allocation.

We tested three different minimum allocations while researching the alternative design. The first minimum allocation is like the official sample in that it is dependent on the number of frame establishments in each stratum:

$$n_h^{min1} = \begin{cases} 1 & if\ N_h < 12 \\ 2 & if\ N_h \geq 12 \end{cases} \tag{5}$$

In the tables below this minimum allocation will be referred to as the "OEWS-like" minimum allocation.

The second minimum allocation aims to ensure that there are a minimum number of observations for the most common occupational estimates in the industry and MSA-BOS area domains, since this is part of the criteria that determines if an OEWS estimate is published or suppressed. By coordinating the minimum allocation with the publishability rules, we hope to increase the number of publishable estimates. The minimum number of observations needed for an estimate to be published is confidential; we will denote this as x. We use past OEWS microdata to determine the most common occupations for each industry and area domain. These are the occupations that make up the top 90th percentile of employment within each domain. Next, we determine which occupations are found in the size classes within each industry and area. Since the OEWS is selected using a PPS sample, we can determine the expected number of sample units that would fall in each size class within each industry and area domain, given some sample size. By knowing the occupations found in each size class and the expected percentage of sample units that would fall in each size class, we determine the likelihood of collecting each of the common occupations, given a sample of only one establishment. These likelihood measures would be greater than zero but less than or equal to one. The inverse of this likelihood measure is the expected sample size required to collect one observation of the common occupation. We multiply this value by x to get the expected sample size needed to collect x observations for each common occupation. We use the maximum value of the expected sample sizes across all the common occupations to determine the final minimum sample size needed for each industry and area domain.

To reconcile the minimum allocation with the efficient sample, a further step is needed to distribute the industry and area minimums to each sampling stratum. Again, since we know that the sample is selected using PPS sampling, we can determine the expected sample sizes for the strata found within each industry and area domain. After this step, every state by aggregate area by NAICS4 by educational ownership strata will have two minimum allocations; one that ensures at least x observations for the most common occupational estimates within industry domains and one that ensures at least x observations for the most common occupational estimates within area domains. The final minimum allocation, n_h^{min2}, is the maximum value of the industry and area minimums. In the tables below this allocation will be referred to as the "targeted" minimum allocation.

The third allocation is simply the maximum value between the first and second minimum allocations, $n_h^{min3} = \max(n_h^{min1}, n_h^{min2})$. In the tables below this minimal allocation will be referred to as the "hybrid" minimum allocation.

We tested both efficient allocations with each of the three minimum allocations, resulting in six different allocations.

3.4 Sample Selection

Like the official sample, we use a PPS sampling scheme for the alternative sample which assigns the same size to all establishments within stratum and size class cells. The alternative sample also uses PRNs to select the sample. We target the same part of the PRN range as the official sample to increase the overlap between the alternative and official samples.

The main difference between how the official and alternative samples are selected is the timing of the sample selection. Since the official sample uses a three-year survey cycle, one-sixth of the full three-year sample is selected each bi-annual panel. For the alternative sample, the full one-year sample is selected in November, and then split into two equal parts to be collected in the November and May panels. The half of the sample collected in the November panel is fielded shortly after selection, whereas the half that is collected in May is held for roughly six months. During this period, some establishments go out of business or out of scope for OEWS. We will refer to these establishments as deaths. Conversely, some establishments come into business or in-scope during the six months, which we will refer to as births. For the alternative sample selected in May, we use updating procedures to handle the births and deaths occurring between the time the sample was selected and collected.

Handling the deaths is straightforward. We simply compare the frame from the November panel to the May panel and find all the establishments that no longer exist. These are the frame deaths. Any frame deaths in the May alternative sample are removed.

Updating the sample to include births is more involved. Like when dealing with the deaths, we compare the November and May frames to identify the establishment that exists on the May frame but not on the November frame. These are the frame births, which we sample to update the half of the alternative sample that is collected in May. We allocate the birth sample to strata in such a way that the birth sample is selected at the same rate as the full sample. For example, if a stratum has 100 employees and was allocated 5 sample units for the full sample, the sampling interval for the stratum is 100 divided by 5, or 20 employees per sample unit. If the birth frame for the same stratum has 25 employees, then the birth sample allocated is 25 divided by 20 or 1.25 sample units. We use a random rounding algorithm to either round the final birth sample up to 2, with probability of 0.25 or down to 1, with probability 0.75. This gives a birth sample unit roughly the same selection probability as a similarly sized full sample unit, which in turn gives the two units roughly the same sampling weight.

Once the deaths are removed from the May alternative sample, and the birth sample is added, the sample is then ready for collection. For both the November and May panels, the alternative sample is fielded at the same time as the official sample.

3.5 Swapping Algorithm

An important requirement of the alternative sample is that it maximally overlaps with the official sample to keep the cost of collection low. To achieve this, we used an algorithm that replaces nonoverlapping sample units in the alternative sample with sample units found in the official sample. We will refer to this as the swapping algorithm. We first find the nonoverlapping sample in the alternative and official sample, within each stratum. Then for each sample unit in the nonoverlapping alternative sample, we look for a similar sized sample unit from the nonoverlapping part of the official sample. If there is only one, we replace the nonoverlapping alternative sample unit with that unit. If there are more than one, we give preference to the nonoverlapping official sample unit that is closest in terms of employment and within the same MSA area as the nonoverlapping alternative sample unit. If there are ties, then we randomly select the official sample unit to swap into the alternative sample. We allow a nonoverlapping official sample unit to be swapped into the alternative sample up to two times, effectively reducing the sample size of the alternative sample.

It should be noted that the swapping algorithm comes at a cost in terms of both bias and variance to the estimates calculated from the alternative sample. The added bias comes from using establishments that were not directly sampled, and the added variance comes from decreasing the effective sample size.

4. Results

While researching the alternative sample design we first evaluated how well the six different allocations performed in terms of natural overlap with the official sample. We picked 2021 to check this since the official sample was already selected for that period. Using the official 2021 frame, we selected alternative samples using each allocation. We then selected the sample using a true PPS sampling scheme, and not the approximate PPS sampling that OEWS officially uses, where PRNs are used to select the sample. We chose to select the alternative sample this way so that we can have an independent natural overlap baseline to evaluate how much the PRN coordination and swapping algorithms increase the overlap. Table 2 below shows the natural unit overlap between the alternative and official samples, by allocation type. Table 3 shows the natural employment overlap.

Table 2: Unit overlap between the official and alternative samples in 2021, by allocation

Efficient Allocation	Minimum Allocation	Official Sample Units	Alternative Sample Units	Overlap	%
Neyman	OEWS-like	374,126	373,933	58,033	15.5%
Neyman	Targeted	374,126	373,942	58,219	15.6%
Neyman	Hybrid	374,126	373,853	58,816	15.7%
Power Neyman	OEWS-like	374,126	373,923	63,063	16.9%
Power Neyman	Targeted	374,126	373,947	61,450	16.4%
Power Neyman	Hybrid	374,126	374,092	61,472	16.4%

Table 3: Employment overlap between the official and alternative samples in 2021, by allocation

Efficient Allocation	Minimum Allocation	Official Sample Employment	Alternative Sample Employment	Overlap	%
Neyman	OEWS-like	28,204,283	58,135,381	9,304,875	33.0%
Neyman	Targeted	28,204,283	57,925,488	9,312,189	33.0%
Neyman	Hybrid	28,204,283	57,525,744	9,331,182	33.1%
Power Neyman	OEWS-like	28,204,283	54,292,875	9,517,968	33.7%
Power Neyman	Targeted	28,204,283	53,111,172	9,448,683	33.5%
Power Neyman	Hybrid	28,204,283	52,816,390	9,450,893	33.5%

The unit overlap is around 15.6 percent for the Neyman allocations and 16.5 percent for the power Neyman allocations. The employment overlap is around 33 percent for the Neyman allocations and 33.6 percent for the power Neyman allocations. The highest unit and employment overlap occurs for the power Neyman allocation with OEWS-like minimums. This is not surprising since this allocation is the most like the official allocation.

It is interesting to note that while the official and alternative sample have roughly the same number of units, the alternative sample has more than double the amount of employment. This is because the official sample selects a PPS sample spread across three years, and the comparison in the table uses only a single year of the official sample. Approximately two-thirds of the largest establishments were ineligible for selection in 2021 since they were sampled in the previous two years. It is also interesting that the Neyman allocations have between 7 and 9 percent more employment than the power Neyman allocations. This is because the Neyman allocation assigns more sample units to the largest strata, compared to the power Neyman allocations. The larger strata tend to skew more towards larger establishments resulting in the Neyman allocation to be more skewed towards larger establishments.

Next, we used a simulation study to compare the bias and variance of the estimates using each of the six allocations. For this, we utilized an output from another project where every establishment on the 2019Q2 QCEW is assigned predicted occupational employment and wage data. For each allocation we drew 150 samples from the predicted QCEW using Poisson PPS sampling. For the simulation study we ignored non-response. For each simulation sample we calculated design-based estimates for national, NAICS4 industry and MSA/BOS areas. By treating the predicted values on the QCEW as truth, we were able to calculate the relative root mean square error (RRMSE) for each estimate, which captures both the bias and variance of the estimate.

$$RRMSE_{D,o} = \frac{\sqrt{\sum_{s=1}^{150}\left(\hat{x}_{D,o,s} - X_{D,o}\right)^2/150}}{X_{D,o}} \qquad (6)$$

where,

$RRMSE_{D,o}$ = relative root mean square error measure for occupation o, in domain D. Domains can be the nation, or individual industries or areas.

$\hat{x}_{D,o,s}$ = design-based estimates for occupation o, in domain D, using simulation sample s. This could be an estimated employment level or a mean wage estimate.

$X_{D,o}$ = true value for occupation o, domain D. This could be a true employment level or mean wage estimate.

Tables 4, 5 and 6 have the RRMSE distributional statistics for the national, industry and MSA/BOS occupational employment estimates. We put the RRMSE results for the occupational mean hourly wage estimates in Appendix A.

Table 4: RRMSE distribution statistics for national occupational employment estimates, by allocation

			National Emp RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	790	4.6%	0.8%	1.5%	3.0%	5.7%	9.9%
Neyman	Targeted	790	4.8%	0.9%	1.6%	3.2%	6.0%	9.9%
Neyman	Hybrid	790	4.7%	0.9%	1.6%	3.2%	5.9%	9.8%
Power Neyman	OEWS-like	790	4.6%	1.1%	1.8%	3.2%	5.5%	8.8%
Power Neyman	Targeted	790	5.3%	1.3%	2.2%	3.8%	6.5%	10.1%
Power Neyman	Hybrid	790	5.3%	1.4%	2.2%	3.8%	6.6%	10.3%

Table 5: RRMSE distribution statistics for NAICS4 industry occupational employment estimates, by allocation

			Industry Emp RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	46,543	36.3%	3.2%	9.8%	23.0%	48.4%	86.2%
Neyman	Targeted	46,543	37.3%	3.2%	9.7%	22.6%	49.3%	89.6%
Neyman	Hybrid	46,543	36.6%	3.2%	9.5%	22.1%	48.7%	88.0%
Power Neyman	OEWS-like	46,543	31.0%	2.9%	8.4%	19.3%	41.4%	74.7%
Power Neyman	Targeted	46,543	36.5%	3.9%	10.0%	22.7%	48.7%	87.1%
Power Neyman	Hybrid	46,543	36.7%	3.9%	10.0%	22.8%	49.1%	87.8%

Table 6: RRMSE distribution statistics for MSA/BOS area occupational employment estimates, by allocation

			Area Emp RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	221,028	67.3%	9.3%	23.8%	50.5%	93.7%	147.3%
Neyman	Targeted	221,028	66.0%	8.0%	21.4%	46.5%	90.7%	150.2%
Neyman	Hybrid	221,028	64.5%	8.0%	21.3%	46.2%	89.1%	145.7%
Power Neyman	OEWS-like	221,028	56.4%	8.8%	20.9%	42.9%	77.6%	120.5%
Power Neyman	Targeted	221,028	52.7%	6.5%	17.1%	36.9%	72.0%	118.6%
Power Neyman	Hybrid	221,028	52.9%	6.4%	17.2%	37.0%	72.3%	119.1%

We should make clear that the results in Tables 4 through 6 are meant for comparing the six allocations to each other, and not to give an estimate of the precision of the estimates if implementing these allocations. For example, we cannot say that using the Neyman allocation with OEWS-like minimums for the alternative sample will give us national occupational employment estimates with a median RRMSE of 4.6 percent. We cannot say this because what we are treating as truth in our simulation study is not the actual truth, but rather predictions that have their own biases and variances associated with them. The goal of these tables is to show the relative standings of each allocation in terms of precision.

Table 4 shows that the Neyman allocations produce slightly more precise (i.e., lower RRMSE measures) national employment estimates than the power Neyman allocations. This is true if looking at the average or median RRMSE. One interesting finding is that the power Neyman allocation using OEWS minimums have estimates that are almost as precise as the Neyman allocation using the OEWS minimums, and it has better results than the Neyman allocations with the Targeted and Hybrid minimums. Said another way, there is not much precision being lost for the National employment estimates when using the power Neyman allocation with OEWS-like minimums over the Neyman allocations. All differences between the six allocations when considering the national employment estimates are quite small.

Table 5 and 6 shows that the relative standings switch from the Neyman to the power Neyman allocations when considering the sub-national employment estimates. For both the industry and area estimates, the power Neyman allocations produce more precise estimates when focusing on the average and median RRMSE. The allocation that produces the most precise industry estimates is the power Neyman allocation with OEWS-like minimums, where the most precise area estimates come from the power Neyman allocation with Targeted minimums. It is important to point out that the targeted minimum allocation sets minimums at the MSA/BOS areas, whereas both efficient allocations and the OEWS-like minimum allocation use aggregate areas. Not surprisingly, setting minimums at the MSA/BOS area level helps the precision of those estimates.

It should be noted that many of the sub-national estimates that feed into Tables 5 and 6 would be suppressed due to lack of precision. Part of the rules that OEWS uses to determine if an estimate is publishable is how precise it is in terms of estimated standard-error of the estimate, which would be correlated with the RRMSE. The lack of precision is most stark when looking at the MSA/BOS area estimates. This is not too surprising, since the alternative sample does not use MSA/BOS areas for their strata definition. The official estimation methodology uses modeling to help with the precision for these area estimates. Still, using a single year's worth of OEWS sample for the MSA/BOS area estimates will result in many of these estimates to be suppressed due to lack of precision.

We also looked at the differential response rates when using the six different allocations. To do this, we used a response propensity model trained on the 2019Q2 OEWS data. We use a logistic regression model with size class, 3-digit NAICS industry and state for the predictors. For this analysis, we excluded state and federal government data since their response rate is nearly 100 percent. We used the response propensity model to predict a response propensity score for each unit selected in each simulation sample. From there, we were able to calculate expected response rates for each of the 150 simulation samples selected for each allocation. Table 7 below shows the average, minimum, median, and maximum response rates for the six allocations. For comparison's sake, the 2019Q2 response rate for OEWS, when excluding state and federal government, is 70.6 percent.

Table 7: Estimated response rates, by allocation

		Estimated Response Rates			
Efficient Allocation	**Minimum Allocation**	**Avg.**	**Min**	**Median**	**Max**
Neyman	OEWS-like	67.8%	67.8%	67.8%	67.9%
Neyman	Targeted	68.0%	67.9%	68.0%	68.0%
Neyman	Hybrid	68.1%	68.1%	68.1%	68.2%
Power Neyman	OEWS-like	69.0%	68.9%	69.0%	69.0%
Power Neyman	Targeted	69.3%	69.3%	69.3%	69.4%
Power Neyman	Hybrid	69.4%	69.4%	69.4%	69.5%

The estimated response rates across the 150 simulated samples are quite stable for each allocation. All six proposed allocations have lower response rates than the official 2019Q2 OEWS sample. The Neyman allocations and power Neyman allocations have response rates that are about 2.5 and 1.3 percent lower, respectively, than the official response rate. This is mainly due to the alternative allocations being more skewed towards larger establishments which tend to have a lower response propensity than smaller establishments.

Based on the results of the natural overlap and simulation study, we decided that the best allocation for the alternative sample is the power Neyman with the OEWS-like minimums. The main reason for this was the increased precision that the power Neyman allocation had for the sub-national estimates, which is an important product of the OEWS. More research is needed to understand if the targeted minimum allocation will increase the number of publishable estimates, as intended. Ideally, we would use a field test where the targeted minimum allocation is implemented after using the OEWS-like minimum allocation, keeping all other aspects of the sample design fixed, to see if the number of publishable estimates increases.

For the last step of our research, we added in the PRN coordination and swapping algorithm to the alternative sample. Both help with increasing the sample overlap between the alternative and official samples. The PRN coordination has the added benefit of decreasing the overlap between the alternative sample and other BLS establishment survey samples. Tables 8 and 9 show the unit and employment overlap when adding in the PRN coordination and swapping algorithm.

Table 8: Unit overlap when including the PRN coordination and swapping algorithm

Overlap	Official Sample Units	Alternative Sample Units	Overlap	%	Supple-ment needed
Natural	374,126	373,923	63,063	16.9%	310,860
PRN Coordination	374,126	373,981	101,759	27.2%	272,222
PRN Coordination + Swapping	374,126	344,049	207,281	60.2%	136,768

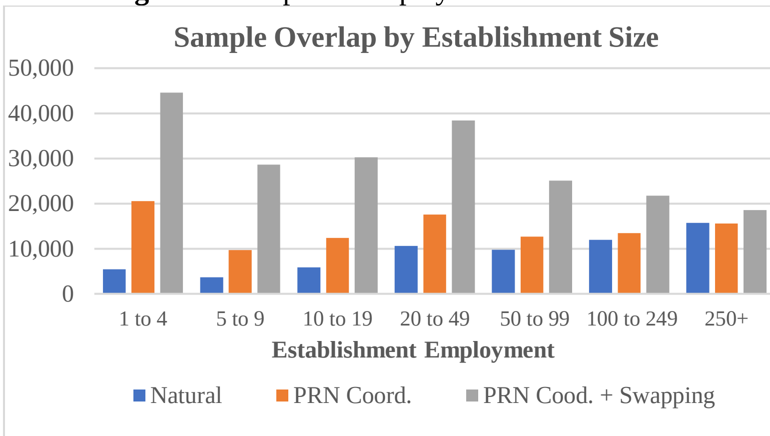
Table 9: Employment overlap when including the PRN coordination and swapping algorithm

Overlap	Official Sample Employment	Alternative Sample Employment	Overlap	%
Natural	28,204,283	54,292,875	9,500,226	33.7%
PRN Coordination	28,204,283	54,151,310	9,908,537	35.1%
PRN Coordination + Swapping	28,204,283	51,941,917	12,693,016	45.0%

By changing the selection mechanism from true PPS sampling to the approximate PPS using PRNs, we increase the unit overlap between the alternative and official sample by 10.3 percent, from 16.9 to 27.2 percent. The employment overlap increased by only 1.4 percent, from 33.7 to 35.1 percent. The average employment for each additional overlapping sample units added was 10.6. When adding in the swapping algorithm, the unit overlap increased by 33.0 percent, to 60.2 percent, and the employment overlap increased by 9.9 percent, to 45.0 percent. The average employment for a unit swapped into the alternative sample was 26.4.

In Figure 2 below we show the number of overlapping sample units by establishment size categories for the natural overlap, the overlap when we add PRN targeting and the overlap when we add the swapping algorithm. The natural overlap increases as the size of the establishment gets larger. This pattern goes away when we add in the PRN targeting. The overlap shifts from being skewed towards larger establishments, to having more overlap found in the smallest and mid-sized establishments. Most of the added overlap from the swapping algorithm occurs in the bottom four size classes. There is a ceiling to the amount of overlap that can occur in the largest size classes, due to the three-year sample cycle of the official sample. As mentioned earlier, two-thirds of the largest establishments are ineligible for selection into be in the 2021 official sample, due to being selected in the previous two years. Since this is not a requirement for the alternative sample, these large establishments have a high probability of being in the alternative sample and thus will never have a chance of overlapping between the two samples. This is why the PRN targeting and swapping algorithm mostly adds small establishments to the overlap.

Figure 2: Sample overlap by establishment size



The swapping algorithm causes the sample size for the alternative sample to decrease by 8.0 percent. The sample employment decreases by 4.1 percent. While not ideal, this was necessary to keep the cost of collecting the alternative sample manageable. Table 8 shows that the swapping algorithm causes a 49.8 percent drop in the amount of additional sample units (i.e., supplemental sample) needed for the alternative sample. The overlapping and supplemental sample will be asked questions about the addition data items along with the normal OEWS questions on occupational employment and wages. The part of the official sample that does not overlap is asked only about occupational employment and wages. A nice feature of the alternative sample is that not only will it be able to provide annual estimates for the additional data items, but it will also be able to calculate annual occupational employment and wage estimates.

5. Conclusion

This paper outlined the empirical work we did to design an alternative sample suitable for testing the collection and estimation of additional data items for the OEWS program. There were several requirements that dictated how we designed the alternative sample. First, it must be able to produce unbiased annual estimates using only one years' worth of sample. To do this, we had to deviate from the three-year survey cycle that the official sample uses and drop the rule that no establishment can be selected more than once every three years. For the alternative sample, all establishments were eligible for selection on the sampling frame. Since the alternative sample was one-third the size of the official sample, we broadened the strata definition by using aggregate areas instead of detailed MSA and BOS areas.

The second requirement was that the alternative sample must be designed to produce the best occupational employment and wage estimates possible. This caused us to use similar allocation and selection procedures as the official sample. We tested six different allocations, but ultimately chose the one most like the official sample allocation, based mainly on the results of our simulation study.

The last requirement was that the alternative and official sample must maximally overlap. To achieve this, we used PRNs to coordinate the selection between the official and alternative sample, and we also implemented a swapping algorithm. We required that the alternative sample unit could only be replaced by an official sample unit that was very similar to it in terms of geography, industry, and size. Even still, we acknowledge that the swapping algorithm comes at a cost by adding bias and variance to the estimates. However, the swapping algorithm was necessary to reduce the number of additional sample units needed for the alternative sample to bring costs down.

An exciting byproduct of this research is that it can help the OEWS program test methods that could be useful for creating occupational time series estimates. A promising feature of the alternative sample is that it can produce OEWS time series estimates, if used in consecutive years. As we mentioned in section 2.1, a major drawback of the official OEWS sample design is that it cannot provide unbiased time series estimates. This research can help test different changes that can be made to the OEWS sample design to make it suitable for time series estimation.

6. Acknowledgements

The authors would like to acknowledge and thank all those who contributed to this research project: Dave Byun and Edwin Robison from the BLS Statistical Methods Staff (SMS), Steve Miller from the BLS Office of Survey Methods Research (OSMR), Matt Dey from the BLS Office of Employment Research and Program Development (OERPD), and Laurie Salmon from the BLS Division of Occupational Employment and Wage Statistics (DOEWS).

7. References

Bankier, Michael D. (1988). Power Allocations: Determining Sample Sizes for Subnational Areas. *American Statistician*, Vol. 42, pp. 174-177.

Martinelli, Cori and Jones, Carrie K. (2015). Field-Testing the Collection of New Data Elements in the Occupational Employment Statistics Survey. In JSM Proceedings, Government Statistics Section. Alexandria, VA: American Statistical Association. 1553-1567.

Office of Management and Budget [OMB]. (1982). "New Statistical Standard on Comparability of Statistics on Business Size." Federal Register 47 (96), 21362-21363.

U.S. Bureau of Labor Statistics [US BLS], (2023). *Occupational Employment and Wages Technical Note*. Retrieved September 4, 2023, from http://www.bls.gov/news.release/ocwage.tn.htm

Appendix A: RRMSE results for mean hourly wage estimates

National Estimates

			National Wage RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	790	1.6%	0.3%	0.5%	1.0%	2.0%	3.5%
Neyman	Targeted	790	1.7%	0.3%	0.6%	1.1%	2.0%	3.5%
Neyman	Hybrid	790	1.6%	0.3%	0.6%	1.1%	2.0%	3.5%
Power Neyman	OEWS-like	790	1.6%	0.4%	0.6%	1.1%	1.9%	3.1%
Power Neyman	Targeted	790	1.9%	0.4%	0.7%	1.3%	2.3%	3.7%
Power Neyman	Hybrid	790	1.9%	0.4%	0.7%	1.3%	2.3%	3.7%

NAICS4 Industry Estimates

			Industry Wage RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	46,543	8.5%	0.7%	2.4%	5.3%	10.8%	19.4%
Neyman	Targeted	46,543	8.6%	0.7%	2.4%	5.3%	10.9%	19.8%
Neyman	Hybrid	46,543	8.4%	0.6%	2.4%	5.2%	10.7%	19.5%
Power Neyman	OEWS-like	46,543	7.4%	0.6%	2.1%	4.6%	9.4%	17.2%
Power Neyman	Targeted	46,543	8.4%	0.8%	2.5%	5.4%	10.8%	19.2%
Power Neyman	Hybrid	46,543	8.4%	0.8%	2.5%	5.4%	10.8%	19.3%

MSA/BOS Area Estimate

			Area Wage RRMSE					
Efficient Allocation	**Minimum Allocation**	**Number of Ests**	**Avg.**	**10th Pct**	**25th Pct**	**Median**	**75th Pct**	**90th Pct**
Neyman	OEWS-like	221,028	13.4%	1.8%	4.8%	9.9%	17.8%	28.2%
Neyman	Targeted	221,028	12.9%	1.5%	4.4%	9.3%	17.2%	27.7%
Neyman	Hybrid	221,028	12.9%	1.5%	4.4%	9.3%	17.1%	27.6%
Power Neyman	OEWS-like	221,028	12.3%	1.8%	4.5%	8.9%	16.1%	25.9%
Power Neyman	Targeted	221,028	11.5%	1.2%	3.7%	7.9%	15.3%	25.3%
Power Neyman	Hybrid	221,028	11.5%	1.2%	3.7%	8.0%	15.3%	25.3%