Agricultural peak employment estimation

Employment Security Department
WASHINGTON STATE

Agenda

- Review
  - ESD requirements
  - Job order identification
  - USDOL determination thresholds
  - Employer estimation
- Employment estimation
  - Method overview
  - Example results
- Moving forward

Review: ESD requirements

- USDOL/ETA for 232 requires:
  - Total number of employers contacted during the survey
  - Total number of respondents
  - Total number of U.S. workers (employment) reported by employers
  - Estimated number of employers (business locations)
  - Estimated number of U.S. workers (employment)
    - Estimated number of crop variety activity workers (employment)

Review: job order identification

- Federal guidelines encourage surveys for any commodity activity to which one or more of the following conditions apply
  1. One hundred or more workers were employed in the previous season, or are expected to be employed in the current season
  2. The crop activity has an unusually complex wage structure
  3. The crop or crop activity has been designated by the national office as a major crop or crop activity
  4. Foreign workers were employed in the previous season, or employers have requested or may be expected to request foreign worker in the current season

Source: U.S. Department of Labor, Employment and Training Administration. Handbook No. 385: p. I-115

Review: USDOL determination threshold

USDOL threshold requirements

Number of estimated workers in crop activity area	Percent needed to make a determination
100 – 349	100%
350 – 499	60%
500 – 799	50%
800 – 999	40%
1,000 – 1,249	35%
1,250 – 1,599	30%
1,600 – 2,099	25%
2,100 – 2,999	20%
3,000 or more	15%

Examples:

Crop-variety	Activity	Reported workers	Estimated workers	Proportion (reported/ estimated)	Threshold	Wage determination
Apple, Gala	Harvest	3,500	20,000	18%	15%	Yes
Cherry, Red	Harvest	2,000	15,000	13%	15%	No
Pear, Bosc	Harvest	800	2,000	40%	25%	Yes
Berry, Strawberry	Harvest	240	500	48%	50%	No

Source: U.S. Department of Labor, Employment and Training Administration. Handbook No. 385: p. I-114

Review: employer estimation method

- Log-linear models for capture recapture:
  1) Determine the probability of a unit to experience a capture history
     - Example: Determine the likelihood of a crop-variety firm responding to the surveys
  2) From understanding the probability of capture, the expected number of units having a capture history can be determined
  3) The expected number of units having a capture history then is re-expressed as a log-linear model
     - Expression as a log-linear model aids in reducing inherent bias from the data and allows the fitting of a regression model to estimate abundance
  4) Fit a log-linear model
     - Poisson regression, deals with count data
     - Helps us identify bias, correct any bias found and produce a stable estimate
     - Enables the estimation of firms missed during the search occasions
  5) Abundance estimation
     - Produces final abundance estimate
     - Uses the number found at least once and the estimated number missed

Review: employer estimation analytical steps

Descriptive statistics

- Transform data to a usable format (matrix of capture histories)
  - Assign binary indicators for each capture occasion
- Produce descriptive statistics for capture-recapture data

Model fitting

- Fit various log-linear models for a closed population
- M_0, M_t, M_b

Model selection

- Produce fit statistics for the number of captures on each capture occasion and model performance.
  - AIC, BIC, standard error, etc.
- Using model fit statistics, select the model to be used for estimation

Abundance estimate

- Apply the selected model to compute the closed population abundance estimate and 95% confidence interval

Review: industry employer estimates

Industry (NAICS)	Adjusted 2017 QCEW firm count	Abundance estimate	Absolute error	Absolute percent error	Low 95	Hi 95
Other vegetable and melon farming	225	181	44	20%	128	284
Apple orchards	588	549	39	7%	483	633
Grape vineyards	156	149	7	4%	118	201
Berry (except strawberry) farming	176	180	4	2%	137	253
Fruit and tree nut combination farming	18	13	5	28%	8	>37.5
Other noncitrus fruit farming	713	695	18	3%	625	782
All other miscellaneous crop farming	209	217	8	4%	129	442

Source: Employment Security Department/LMEA, 2015 and 2017 Agricultural Peak Employment Wage and Practice Employer survey; U.S. Bureau of Labor Statistics, QCEW

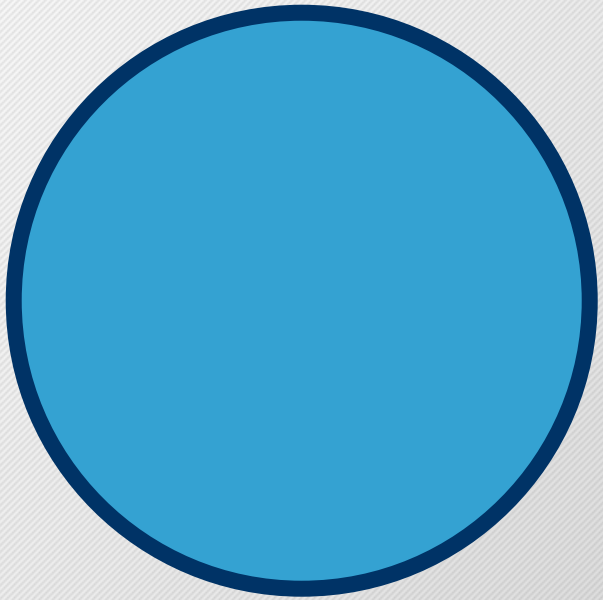
Employment estimation: method overview

- Iterative proportional fitting (raking algorithm)
  - Repeatedly adjust a set of data (survey responses) so that its marginal totals match specified marginal control totals (population totals)
    - Iterative algorithm for estimating cell values of a contingency table such that the marginal totals remain fixed and the estimated table decomposes to an outer product
    - Consists of two cycles that checks convergence criteria over the control variables
    - Control totals = Employer estimates (capture-recapture)
  - Procedure results in calibration weights
    - Calibrated weights adjust survey responses for survey non-response, bias and employer representation

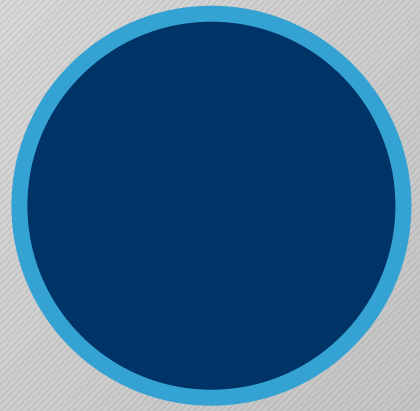
Employment estimation: method overview continued…

Raking algorithm example

Population margins



Sample margins



Employment estimation: method overview continued…

Base

Sample	A	B	C	Total
X	1	2	1	**4**
Y	3	5	5	**13**
Z	6	2	2	**10**
Total	**10**	**9**	**8**	**27**

Population	A	B	C	Total
X	?	?	?	**9**
Y	?	?	?	**19**
Z	?	?	?	**15**
Total	**14**	**15**	**14**	**43**

Iteration 1: row cycle

Raking (1)	A	B	C	Total
X	2.25	4.5	2.25	**9**
Y	4.384615	7.307692	7.307692	**19**
Z	9	3	3	**15**
Total	**15.63462**	**14.80769**	**12.55769**	**43**

Iteration 1: column cycle

Raking (1)	A	B	C	Total
X	2.0148	4.5584	2.5084	**9.081624**
Y	3.9262	7.4026	8.1470	**19.47581**
Z	8.0590	3.0390	3.3446	**14.44257**
Total	**14**	**15**	**14**	**43**

Employment estimation: method overview continued…

Iteration 1: results

Raking (1)	A	B	C	Total
X	2.0148	4.5584	2.5084	**9.081624**
Y	3.9262	7.4026	8.1470	**19.47581**
Z	8.0590	3.0390	3.3446	**14.44257**
Total	**14**	**15**	**14**	**43**

Population	A	B	C	Total
X	?	?	?	**9**
Y	?	?	?	**19**
Z	?	?	?	**15**
Total	**14**	**15**	**14**	**43**

Iteration 2: row cycle

Raking (2)	A	B	C	Total
X	1.9967	4.5175	2.4859	**9**
Y	3.8303	7.2217	7.9480	**19**
Z	8.3701	3.1563	3.4737	**15**
Total	**14.19702**	**14.89547**	**13.90751**	**43**

Iteration 2: column cycle

Raking (2)	A	B	C	Total
X	1.9689	4.5492	2.5024	**9.020525**
Y	3.7771	7.2724	8.0008	**19.05038**
Z	8.2539	3.1784	3.4968	**14.92909**
Total	**14**	**15**	**14**	**43**

Employment estimation: method overview continued…

- Post-stratification adjustment
  - Classifying sample units into groups after data collection using information collected in the survey and auxiliary information to adjust weights to population control totals or for nonresponse adjustment
  - Adjusting the weights within each cell so that the weights sum to the known population marginal totals
    - *Example:*
      - 10 total employers in a specific population, all are surveyed, each have an initial weight of 1
      - 2 employers respond, initial weights of 1 get adjusted to 5, summing to 10
  - Post-stratified weight: $w_{2j} = w_{1j} \frac{\sum_{i \in \mathcal{U}} \mathbb{I}[i \in \mathcal{C}_k]}{\sum_{l \in \mathcal{S}} w_{1l} \mathbb{I}[l \in \mathcal{C}_k]}$
    - $W_{1j} =$ base sample probability weight
    - $\mathbb{I}[.] =$ indicator function taking the value of 1 when its argument is true and 0 otherwise
    - $\mathcal{C}_k =$ post-stratification cells
    - $\mathcal{U} =$ finite population
    - $\mathcal{S} =$ sample of the finite population

Employment estimation: method overview continued…

Raking algorithm procedure

1. Initialize
   - Use the base weights to initialize the raked weight
     - Initialize the iteration counter $k \longleftarrow 0$ and weights as $w_j^{0,p} \longleftarrow w_{1j}$
2. Increment
   - Use the end result of the previous outer cycle iteration to initialize the weights for the current outer cycle iteration
     - Increment the iteration counter $k \longleftarrow k+1$, update the weights $w_j^{k,0} \longleftarrow w_j^{k-1,p}$
3. Inner cycle (post-stratify)
   - Post-stratify with respect to the given control variable
     - Go over the control variables $v = 1, \ldots, p$ and update the weights

$$w_j^{k,v} = \begin{cases} w_j^{k,v-1}, & x_{vj \neq 0} \\ w_j^{k,v-1} \dfrac{T[X_v]}{\sum_{l \in S} w_l^{k,v-1} x_{vl}}, & x_{vj=0} \end{cases}$$

4. Return the weights ($w_j^{k,p}$) at the final stage as the calibrated weights
5. Multiply the reported employment by the calibrated weights to determine total employment

Employment estimation: analytical steps

Employer estimation

- Descriptive statistics
- Model fitting
- Model selection
- Abundance estimate (employer estimate)

Identify estimation cells

- Identify employment estimation cells (job order identification)
  - What crop-variety-activities must we estimate for and analyze
- Attach finite population margins to sample data (employer estimate)

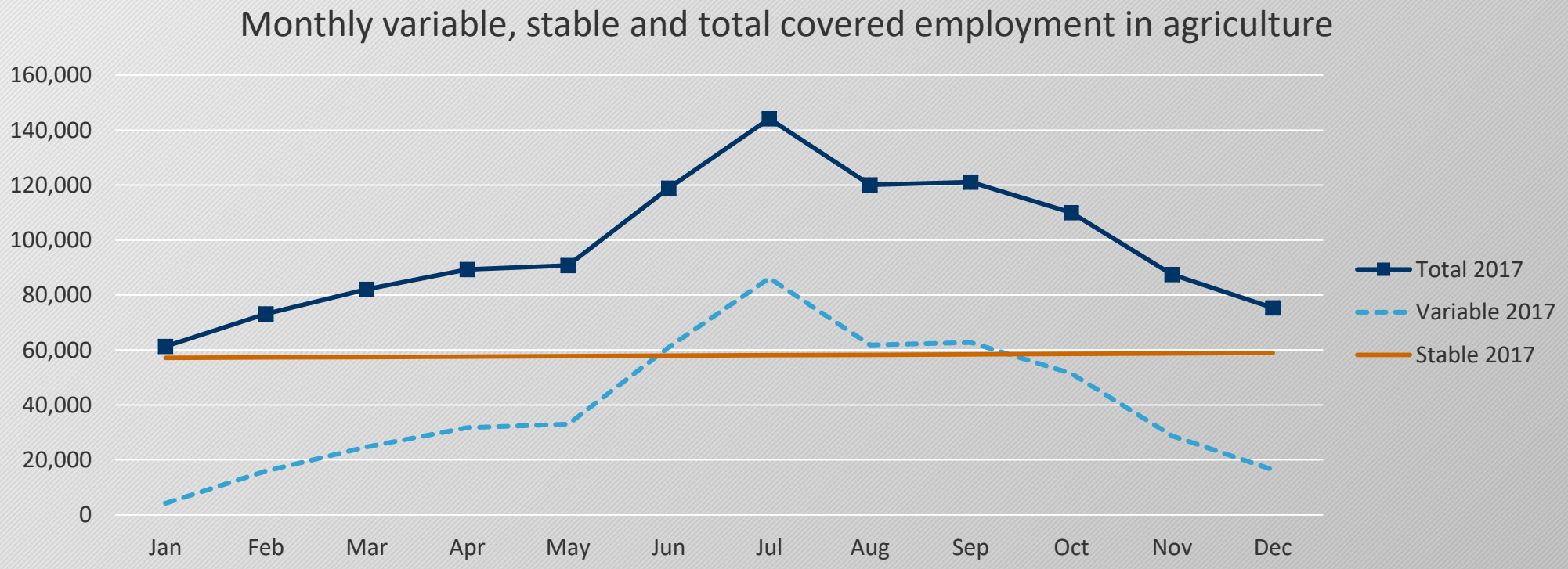
Employment estimation

- Initialize
- Increment (outer cycle iteration)
- Inner cycle iteration (post-stratify)
- Return calibrated weights
- Multiply reported employment by calibrated weights (employment estimate)

Employment estimation: method overview continued…

- General requirements:
  - Summed marginal cell values must be non-zero
  - Marginal column and row values must sum to the same value
  - Sample cell values should be smaller than population cell values
- Assumptions:
  - Population in question is finite
  - Each employer has the same initial probability (base weight) of responding

Example results: application to estimate industry peak employment



Source: Employment Security Department/LMEA; U.S. Bureau of Labor Statistics, QCEW

Example results: application to estimate industry peak employment continued…

Industry (NAICS)	Estimated peak employment	Reported peak employment (2017)	2017 peak variable industry employment (QCEW)	Absolute error	Absolute percent error
Other vegetable and melon farming	4,328	874	2,584	1,744	67%
Apple orchards	23,821	5,284	23,603	218	1%
Grape vineyards	3,785	1,777	3,574	211	6%
Berry (except strawberry farming)	6,432	1,274	6,125	307	5%
Fruit and tree nut combination farming	1,603	1,313	1,402	201	14%
Other noncitrus fruit farming	26,926	8,839	25,042	1,884	8%
All other miscellaneous crop farming	6,353	841	6,000	353	6%
Aggregate	**73,248**	**20,202**	**68,331**	**4,917**	**7%**

Source: Employment Security Department/LMEA, 2017 Agricultural Peak Employment Wage and Practice Employer survey; U.S. Bureau of Labor Statistics, QCEW

Example results: application to estimate peak crop activity employment

Crop	Activity	Estimated employment	Reported employment	Proportion reported	Threshold	Determination
Apples	Harvesting	33,946	5,899	17%	15%	Yes
Berries	Harvesting	6,826	1,533	22%	15%	Yes
Cherries	Harvesting	30,604	10,604	35%	15%	Yes
Pears	Harvesting	12,325	2,265	18%	15%	Yes
Grapes	Harvesting	5,529	621	11%	15%	No

Source: Employment Security Department/LMEA, 2017 Agricultural Peak Employment Wage and Practice Employer survey

Example results: application to estimate peak crop variety activity employment

Crop	Variety	Activity	Estimated employment	Reported employment	Proportion reported	Threshold	Determination
Apple	Braeburn	Harvesting	1,942	385	20%	25%	No
Apple	Cripps pink	Harvesting	2,239	875	39%	20%	Yes
Apple	Fuji	Harvesting	14,069	2,365	17%	15%	Yes
Apple	Gala	Harvesting	23,763	3,627	15%	15%	Yes
Apple	Golden delicious	Harvesting	8,602	1,602	19%	15%	Yes
Apple	Granny smith	Harvesting	16,409	1,978	12%	15%	No
Apple	Honeycrisp	Harvesting	17,442	2,988	17%	15%	Yes
Apple	Red delicious	Harvesting	13,646	2,722	20%	15%	Yes

Source: Employment Security Department/LMEA, 2017 Agricultural Peak Employment Wage and Practice Employer survey

Example results: application to estimate peak crop variety activity employment continued…

Crop	Variety	Activity	Estimated employment	Reported employment	Proportion reported	Threshold	Determination
Berry	Blueberry	Harvesting	5,377	1,389	26%	15%	Yes
Berry	Raspberry	Harvesting	2,086	745	36%	25%	Yes
Berry	Strawberry	Harvesting	459	364	79%	60%	Yes
Cherry	Dark red	Harvesting	16,849	7,079	42%	15%	Yes
Cherry	Red	Harvesting	28,320	5,064	18%	15%	Yes
Cherry	Yellow	Harvesting	15,190	3,695	24%	15%	Yes
Pear	Bartlett	Harvesting	9,443	1,753	19%	15%	Yes
Pear	Bosc	Harvesting	9,872	730	7%	15%	No
Pear	D'Anjou	Harvesting	8,820	1,164	13%	15%	No

Source: Employment Security Department/LMEA, 2017 Agricultural Peak Employment Wage and Practice Employer survey

Moving forward

- March, 2019
  - Final employer and worker survey analysis and estimation
- April, 2019
  - Conference call with all stakeholders presenting final results
    - Feedback period of approximately one week
  - Submission of final results to USDOL
    - Publication of final results is contingent upon USDOL
  - Begin administrative planning for 2019 survey iteration

References

- Deming, W.E, Stephan, F. (1940). “On a least squares adjustment of a sampled frequency table when the expected marginal totals are known”. *The Annals of Mathematical Statistics, **11**(4), 427-444.*
- Deming, W.E. (1943). “Statistical adjustment of data”. *Bureau of the Census and Bureau of the Budget, Washington*. New York: Wiley.
- Holt, D., Smith, T.M.F. (1979). “Post stratification”. *Journal of the Royal Statistical Society. Series A (General), **142**(1), 33-46.*
- Ireland, C.T, Kullback, S. (1968). “Contingency Tables with Given Marginals”. *Biometrika, **55**(1), 179-188.*

References

- Lumley, T. (2004). “Analysis of complex survey samples”. *Journal of Statistical Software, **9**(1), 1-19.*
- Lumley, T. (2017). “survey: analysis of complex survey samples”. *R package version 3.32*. https://cran.r-project.org/web/packages/survey/survey.pdf
- Rao, J.N.K, Yung, W., Hidiroglou, M.A. (2002). “Estimating Equations for the Analysis of Survey Data using Poststratification Information”. *The Indian Journal of Statistics, **64**(A2), 364-378.*
- Valliant, R. (1993). “Poststratification and Conditional Variance Estimation”. *Journal of the American Statistical Association, **88**(421), 89-96.*

Contact information

Steven Ross, Director
Employment Security Department
Labor Market Information
Labor Market and Economic Analysis
(360) 507-9615
sross@esd.wa.gov

Gustavo Avilés, Manager
Employment Security Department
Program Evaluation, Research & Analysis
Labor Market and Economic Analysis
(360) 507-9552
gaviles@esd.wa.gov

Joshua Moll, Research Economist
Employment Security Department
Program Evaluation, Research & Analysis
Labor Market and Economic Analysis
(360) 507-9554
jmoll@esd.wa.gov

Toby Paterson, Research Economist
Employment Security Department
Program Evaluation, Research & Analysis
Labor Market and Economic Analysis
(360) 742-2491
mpaterson@esd.wa.gov