Are There Substantive Differences between Sampling and Censusing Employees in Organizational Climate Surveys?

FedCASIC 2016

U.S. Census Bureau

Suitland, MD

5/3/16

Taylor Lewis[1] and Lorraine Latimore[1], U.S. Office of Personnel Management (OPM)

Nikki L. Graf, University of Mannheim

[1]The opinions, findings, and conclusions expressed in this presentation are those of the authors and do not necessarily reflect those of the U.S. Office of Personnel Management.

Outline

I. Background on the Federal Employee Viewpoint Survey (FEVS)

   I. History, Methodology, and Instrument

   II. Sample Design and the Sample/Census Determination

   III. The Call for All Agencies to be Censused

II. Retrospective Analysis – Exploiting the FEVS 2012 Natural Experiment

III. FEVS 2016 Wording Experiment

IV. Summary and Further Research Ideas

Background on the FEVS

- The FEVS is an annual organizational climate survey administered by the U.S. Office of Personnel Management (OPM) to roughly 900,000 federal employees from 80+ agencies (biennial from 2002 to 2010)

- Web-based instrument consists mainly of Likert-type attitudinal items (e.g., perceptions of leadership, job satisfaction) sent via personalized link embedded in an email message

- Thematically-linked groups of survey items are combined to form indices – one of the most highly visible is the Employee Engagement Index (EEI), which is comprised of 15 items covering three sub-factors of overall engagement

The Sample/Census Determination

- Described in more detail in the FEVS Technical Report (OPM, 2015)

- Prior to survey launch, agencies provide organization codes that are used to place each eligible employee into a hierarchical set of work units used for reporting purposes

- Marginal agency sample size is a function of the size of terminal work units in the organizational structure – employees in smaller work units sampled at a higher rate

- If the organization structure dictates that 75% or more of agency is to be sampled, then a census is conducted instead

- One notable exception was the FEVS 2012 Census

FEVS Sample Size and Response Rate Trends
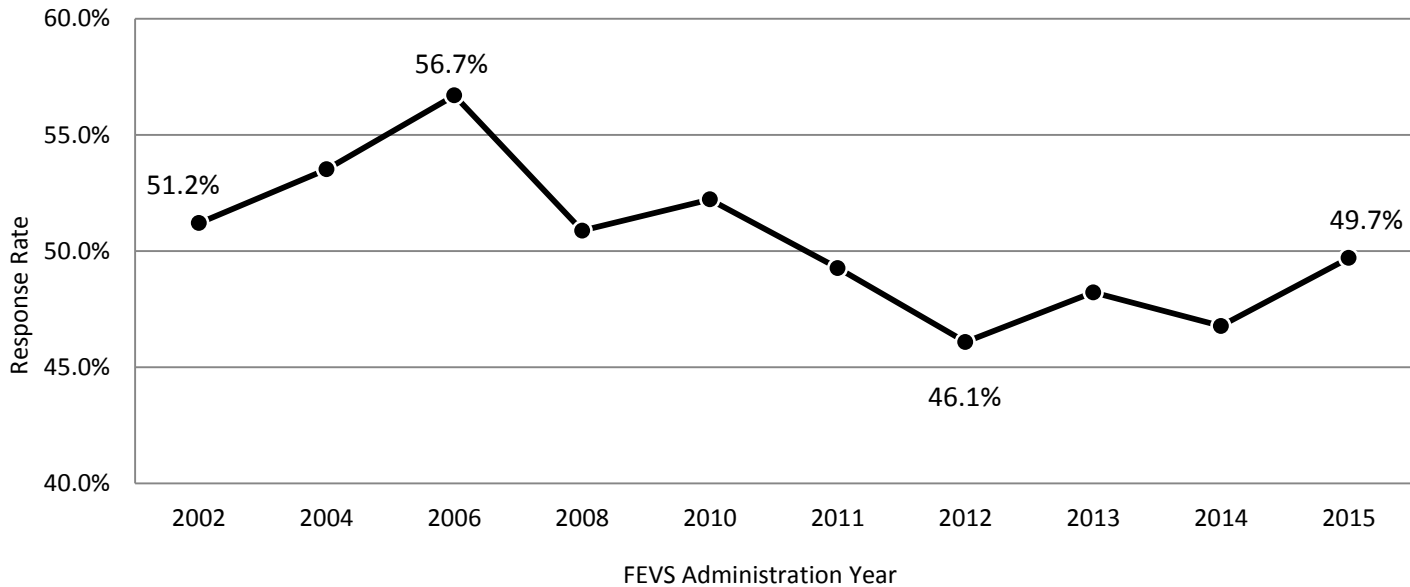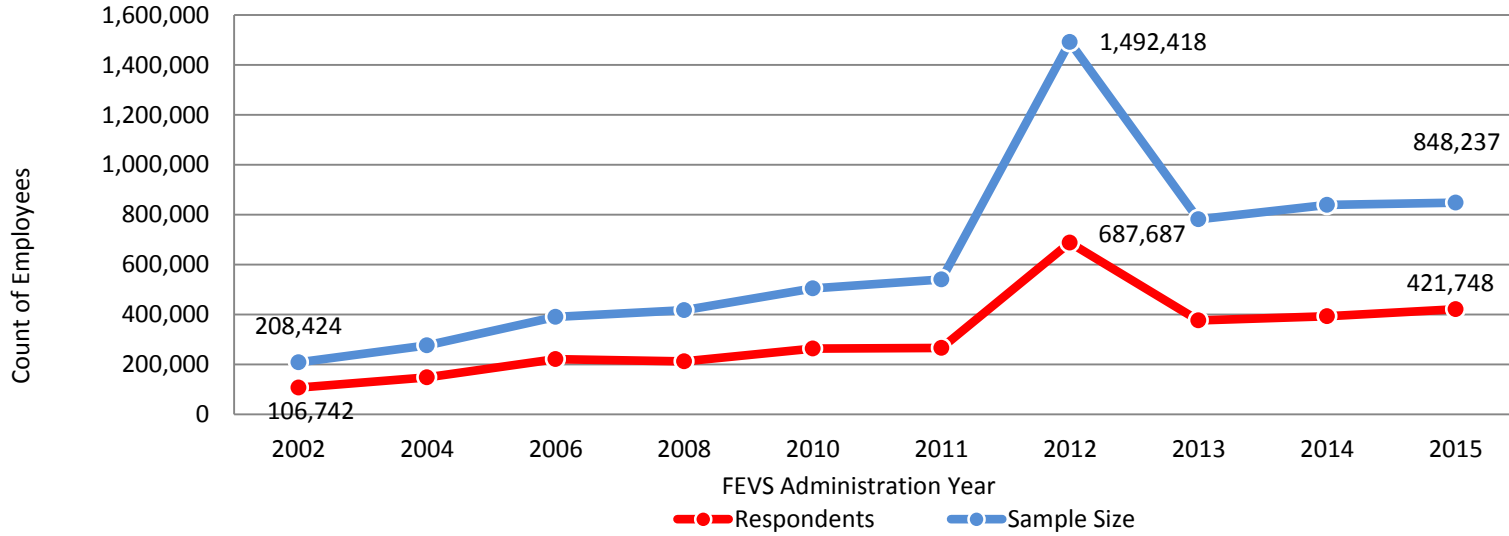
The Push for a Census

- The impetus for the FEVS 2012 census was desire for deepest possible level of analysis within the participating agencies (Berry, 2012)

- Sampling methodology developed for FEVS 2013 and beyond maintains reporting breadth of FEVS 2012 census, yet the FEVS Team continues to hear calls to transition to a perennial census

- Some reasons are logistical or for messaging purposes:
  - Marginal cost of surveying non-sampled employees is negligible given data collection mechanisms already in place
  - Senior leaders understand merits of sampling, but want all employees to have the opportunity to participate and give their "say"

- Recently, we have heard a new reason: a belief that response rates and EEI scores will increase if the agency conducts a census

Theorizing on a Census Effect

- If such an effect can be proven to exist, the FEVS Team would genuinely consider refining the FEVS sampling methodology to allow an agency to conduct a census regardless of its organizational structure

- Literature review turned up plenty of discussion regarding pros/cons of conducting a census versus a sample, but no directly relevant (quasi)experimental research into the phenomenon

- One potentially applicable theory is that of *diffusion of responsibility*, summarized in Barron and Yechiam (2002):
  - Darley and Latané (1968) – motivation to help lessened if others are perceived to be able to help (e.g., Kitty Genovese murder)
  - Diekmann (1985) – individuals in a theoretical game setting less likely to volunteer to help for the greater good of group if they knew someone else already had volunteered

- If true in FEVS, conducting a census could lead to response rate *decrease*

Retrospective Analysis

- Grounds for a natural experiment: 13 agencies went from a sample to a census in FEVS 2012, and 13 agencies (many of the same) went from a census to a sample in FEVS 2013

- For each of the two administration thresholds, we formulated a *first-differenced* estimator (Wooldridge, 2012)

$$\Delta_i = \beta_0 + \beta_1 d_i + \varepsilon_i$$

where Δ_i is change in the (base-weighted) RR or EEI for the ith agency, and d_i is a 0/1 indicator variable for changing from a sample to census (or vice versa)

- We can interpret β_1 as the expected effect at agency level → testing H_0: $\beta_1 = 0$ versus H_1: $\beta_1 \neq 0$ provides insight into whether any observed effect is statistically significant

Agency-Level Results

- Table below reports estimated values of β_1 (i.e., expected percentage point changes) for both the RR and EEI models

Transition Year	Transition Type	Base-Weighted Response Rate	Employee Engagement Index
		Coefficient (p-value)	*Coefficient (p-value)*
2012	Sample → Census	4.03% (p = 0.1579)	-0.73% (p = 0.8101)
2013	Census → Sample	-5.55% (p = 0.0509)	1.12% (p = 0.6948)

- Changes are RR actually in the direction speculated by external stakeholders, but results are not statistically significant at the α = 0.05 level; nothing noteworthy about EEI changes

Limitations of Agency-Level Analysis

- Small sample size (n = 83 agencies) – data could be too aggregated to detect an effect that actually exists

- We had originally planned to replicate the first-differenced estimator analysis for work units below the agency level, but historic response rate information (i.e., prior to widespread use of organizational codes) is not as reliably trended

- Because the "treatment" of a sample/census was not experimentally controlled, we considered pursuing a class of propensity score adjustment/matching techniques (Rosenbaum and Rubin, 1983; Imbens and Rubin, 2015)

- Ultimately decided against that avenue, largely because we felt we lacked sufficiently predictive work-unit covariates – could explore covariates related to level of publicity regarding the survey within the agency

Individual-Level Analysis

- There is an implication that the individuals not afforded the opportunity to participate would participate at a higher rate if given that opportunity

- Using individual-level identifiers, we teased apart response rates for the FEVS 2012 sampled employees based on whether or not they were sampled in FEVS 2011

- We found that these "newly reached" employees responded at a notably lower rate than those who were in the prior FEVS sample: 43.7% versus 50.3%

- We then followed up on individuals who were part of a censused agency in FEVS 2012 but a sampled agency in FEVS 2013, and found that they were less likely to respond: 44.9% versus 54.8%

- Hence, results appear to be mixed

Email Wording Experiment for FEVS 2016

- To have better control over isolating and quantifying any potential effect of informing the employee that his/her agency was conducting a sample/census, we designated four agencies for an FEVS 2016 email reminder wording experiment

- Two of the agencies are conducting a sample, two are conducting a census

- One-half of the agency receives traditional wording, while the other half receives alternative wording to emphasize merits of census/sample

Pro-Sample Wording Example	Pro-Census Wording Example
"…You are one of those randomly selected to participate. Your responses represent not only your perceptions and sentiments, but those of fellow employees not selected…"	"…To obtain the most valuable and useful information possible, we are striving to hear the voice of every employee in <agency name>…"

Brief Summary

- In this talk we presented findings from a retrospective analysis into the impact of the FEVS 2012 natural experiment – both at agency and individual levels

- Results were inclusive

- An ideal experimental design would involve randomly assigning the census/sample "treatment" on work units – unfortunately, such a design is not feasible at this time

- Next best option, in our view, is an email wording experiment systematically manipulating whether an employee weighs in—assuming the message is read in its entirety—on the census/sample aspect when deciding whether or not to participate in the FEVS

References

Barron, G., and Yechiam, E. (2002). "Private E-Mail Requests and the Diffusion of Responsibility," *Computers in Human Behavior*, **18**, pp. 507–520

Berry, J. (2012). "Guide for Interpreting and Acting on Federal Employee Viewpoint Survey Results," November 23, 2012 Memorandum to the Chief Human Capital Officers (CHCO) Council. Available at: https://www.chcoc.gov/content/guide-interpreting-and-acting-federal-employee-viewpoint-survey-results.

Darley, J., and Latané, B. (1968). "When Will People Help in a Crisis?" *Psychology Today*, **2**, pp. 54–57.

Diekmann (1985). "Volunteer's Dilemma," *Journal of Conflict Resolution*, **29**, pp. 605–610.

Imbens, G., and Rubin, D. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, NY: Cambridge University Press.

Rosenbaum, P., and Rubin, D. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, **70**, pp. 41 – 55.

U.S. Office of Personnel Management. (2015). *Federal Employee Viewpoint Survey Technical Report*. Available at: https://www.fedview.opm.gov/2015/Published/.

Wooldridge, J. (2012). *Introductory Econometrics: A Modern Approach*. *Fifth Edition*. Mason, OH: Cengage Learning.

Thanks!

Questions/Comments?

Taylor.Lewis@opm.gov