Do Monetary Incentives Increase Business Survey Response Rates? Results from a Large Scale Experiment

Paul P. Biemer[1], Christopher Ellis[2], Angela Pitts[2], Kimberly Aspinwall[2]
RTI International and University of North Carolina[1]
RTI International[2]

Abstract

This paper describe an experiment to evaluate the effects on response rates and survey costs of offering a $20 prepaid incentive to establishment points of contact (POCs) in a large-scale establishment survey. POCs have a critical role in this survey, in that they coordinate data-collection activities among the employees within their establishments who are selected to participate in the survey. A summary of findings from the experiment and a discussion of the analyses completed for the POC incentive experiment will be provided.

Keywords: O*NET, nonresponse, occupation survey, employee sampling, job analysis, use of money orders.

1. Introduction

This paper describes an experiment to evaluate the effects on response rates and survey costs of offering a $20 prepaid incentive to establishment points of contact (POCs) in a large-scale establishment survey. POCs have a critical role in this survey in that they coordinate data collection activities among the employees within their establishments who are selected to participate in the survey. The paper provides a summary of findings from the experiment and documents the analyses completed for the POC incentive experiment.

1.1 Background

The Occupational Information Network (O*NET) is a comprehensive system for collecting, organizing, describing, and disseminating information on occupational requirements and worker attributes. The O*NET database is designed to be the most comprehensive standard source of occupational information in the United States.

The O*NET Program is an ongoing effort to populate and maintain the O*NET database with valid, reliable, and current occupation and skills data. O*NET data are used by a wide range of audiences, including individuals making career decisions, the public workforce investment system and schools making training investment decisions, educational institutions preparing the future workforce, and employers making staffing, economic development, and training decisions. The O*NET program provides a common language and framework of occupational and skill requirements to meet the needs of various federal programs, including workforce investment and training programs of the Departments of Labor (DOL) and Education (ED). Further information about the O*NET program can be found at the following websites: www.doleta.gov/programs/onet and www.onetcenter.org.

The primary method for collecting these data is the *Establishment method.* This method incorporates a two-stage design that uses (1) a statistical sample of establishments expected to employ workers in each specific occupation and (2) a sample of workers in the occupations within each sampled establishment. The sampled workers are asked to complete the survey questionnaires.

Four domain questionnaires are used to collect data from sampled workers: Skills, Knowledge, Generalized Work Activities, and Work Context. Sampled workers are asked to complete one randomly assigned domain questionnaire, a basic demographic questionnaire, and a brief, occupation-specific task inventory. Workers may either complete the paper questionnaire and return it via mail or complete an online questionnaire at the project Web site.

Data collection operations for the main study began in June 2001 and are continuing. The sampling proceeds by *waves* which are defined by a set of occupations for which employee data are to be collected. Each wave is further divided into sub-waves which are subsamples of establishments selected for the occupations in the wave. Occupations which were not completed within a sub-wave are targeted in a subsequent sub-wave until the desired number of questionnaires are completed for each domain. Each sub-wave generally requires about 7 months to complete. The sub-waves are also interwoven across primary waves. This approach produces a longer time interval between related sub-waves, resulting in more efficient sampling, reduced respondent burden, and a higher yield of completed occupations.

1.2 Purpose

This document describes the design and implementation of the incentive experiment and provides results showing the effect of the incentive on establishment and employee response rates. The results presented here are based on data from 14 data collection waves, which included approximately 25,000 establishments.

Prior to conducting this experiment, the following incentives were offered to the POCs and selected employees in the O*NET sample establishments:

- For employers who agree to participate: the O*NET Toolkit for Business (a packet of information about the O*NET Program that managers can use for human resource planning, including a guide for writing job descriptions).

- For POCs: a desk clock with the introductory mailing, and a framed Certificate of Appreciation to those who agree to participate (combined cost of both is less than $10).
- For employees: a $10 prepaid cash incentive.

The purpose of the POC experiment was to examine the effects of offering POCs who agree to participate a prepaid $20 incentive, in addition to the clock and certificate. The survey methods literature (see, for example, Burnside, Bishop and Guiver, 2005, and Keesling, 2000) suggests that incentives have the potential to significantly and positively affect both the establishment and employee response rates in enterprise surveys. It was thought that the POC would find the incentive appealing given the burden on him/her, the need for sustained cooperation over an extended period of time (several weeks), and the absence of strong positive forces of direct benefit to the establishments and employees to participate in the study. Since employees are offered a $10 payment for completing their questionnaires, POCs might also expect some kind of remuneration in addition to the usual O*NET incentives for completing their tasks.

It is recognized, however, that the two-stage sample design of this establishment survey is rather uncommon and no literature directly speaks to the effects of incentives in surveys with designs and target populations that are similar to the O*NET data collection program. Moreover, the literature on the use of incentives in establishment surveys is rather sparse and inconclusive regarding the effectiveness of incentives for general employee populations. Thus, it was decided to test the effectiveness of the $20 incentive by conducting an experiment. The experiment is described in some detail in ***Section 2***.

For this report, the results of the experiment were analyzed to examine the effect of the incentive on both the establishment and employee response rates. In addition, the relative costs of the two protocols were examined, since the monetary incentive has the potential to at least partially offset its inherent cost through greater efficiencies in the data collection process and higher response rates.

The following section provides details of the methodology employed in conducting the incentive experiment, including the research objectives, the experimental design, and the data collection protocol. ***Section 3*** presents the results of the analysis of establishment and employee response rates and costs. Finally, ***Section 4*** contains a discussion of these results and outcomes.

2. Summary of Sample Design and Data Collection Methods

2.1 Research Objectives

The primary objective of the POC incentive experiment was to determine the effect on response rates of offering the POC a monetary incentive of $20 at the recruitment stage of the O*NET data collection process. It was hypothesized that this incentive would significantly increase the establishment response rate relative to the control (non-incentive) treatment. It was further hypothesized that the incentive would slightly increase employee response rates due to its potential motivating effects on the POC during the follow-up stages of the process.

Exhibit 1. Number of Establishments and Employees in the Experimental Waves

	Incentive	No Incentive	Total
Establishments	7,874	2,624	10,498
Employees	22,309	7,694	30,003
Total	30,183	10,318	40,501

An additional anticipated effect of the incentive was the speed with which the POC distributed the O*NET questionnaires to the employees; it was hypothesized that POCs given the incentive would be more motivated to complete their work. It was believed that this higher level of motivation would translate into the POCs' completing their work at a faster pace. Further, it was thought that the speed with which employees returned their questionnaires might increase, potentially decreasing the number of follow-up calls to the POCs asking them to prompt employees to return their questionnaires. Thus, cost variables were examined to test the above hypotheses.

2.2 Sample Design

The design for testing the incentive was a split ballot design with two treatments: the $20 incentive and a control, $0 incentive treatment. The two treatment conditions were identical except for the incentive and a few changes in the survey procedures that were necessary to implement the incentive condition. There were two experimental units under study: the establishments and the employees within the establishments. Our experimental design randomly assigned establishments to the treatment and control groups, and thus employees were randomly assigned to each group in approximately the same proportions.

Note, however, that once an establishment was assigned to a condition, all employees within that establishment received the same treatment. This type of random assignment induces so-called clustering effects in the employee outcomes, which are taken into account in the subsequent analysis. In addition, the interactions of the RTI telephone interviewers (called Business Liaisons or BLs) with the POCs were carefully monitored to ensure equal levels of effort across both case types.

Another choice in the design was the proportion of sample establishments to allocate to each experimental condition. While an even split of half of the cases going to the incentive treatment and half to the control may be optimal for maximizing the power of significance tests, it is not optimal for maximizing the response rates for the data collection if the hypotheses regarding the incentive effects are supported. Therefore, expecting incentives to increase response rates, the use of incentives was maximized while achieving the objectives of the experiment by allocating approximately 75% of establishments to the incentive condition, with the remainder assigned to the control group.

As shown in ***Exhibit 1***, 7,874 establishments were assigned to the incentive treatment and 2,624 to the control group.

2.3 Data Collection Protocol

The data collection portion of the experiment closely followed the protocol of the main data collection program. The following modifications were made to support the experiment.

POC Incentive. The $20 monetary incentive was provided to the POC in the form of a money order which are preferable to cash for several reasons (see, for example, Dommeyer, 1988 and Church, 1993). First, because the POC mailing may be opened by persons other than the POC, a money order was thought to be less likely than cash to be misplaced before reaching the POC. Further, noting that some POCs might want to donate the money to charity, a money order facilitated the forwarding of the payment to a charity and provided a record of the transaction for the POC. Therefore, blank money orders were purchased in bulk, which was easier, quicker, and less expensive than using checks. Money orders are more secure than cash but offer all of the advantages of checks, including credibility, ease of transfer, and cost savings, compared to checks.

Assignment of Cases to BLs. As a wave was introduced, cases were randomly assigned to BLs. This random assignment process ensured that BLs were working multiple Standard Industrial Classifications (SICs), that any given SIC had two or more BLs working its cases, and that each BL had a combination of incentive and control group cases. Approximately 75% of each BL's assignment for a wave consisted of incentive cases, and 25% consisted of control cases, although this was allowed to vary across BLs.

Data Collection Procedures. The standard data collection protocol for the O*NET survey is as follows:

- Step 1: Verification Call to Receptionist
- Step 2: Screening Call to Point of Contact (POC)
- Step 3: Send Information Packet
- Step 4: Recruiting Call to POC
- Step 5: Sampling Call to POC
- Step 6: Send Questionnaire Packet
- Step 7: Send Toolkit
- Step 8: 7-Day Follow-up Call to POC
- Step 9: Send Thank You/Reminder Postcards
- Step 10: 21-Day Follow-up Call to POC
- Step 11: 31-Day Follow-up Call to POC
- Step 12: Send Replacement Questionnaires
- Step 13: 45-Day Follow-up Call to POC

The experimental protocol differed from the standard protocol in the following ways:

- The Information Package, which is sent to the POC prior to the Recruitment Call, contained a newly developed brochure that described the program's various POC-, company-, and employee-level incentives. For the experiment, two versions of this brochure were used: one that referenced the $20 POC incentive (which was distributed to the POCs in the incentive group) and another (for the control group) that did not mention the $20 money order.
- The "Who, What and How" brochure, containing frequently asked questions regarding the O*NET Data Collection Program, was revised for the treatment group to reflect the additional POC incentive.
- During the Recruiting Call, the BL explained the various program incentives to the POC. This explanation was expanded for the cases in the incentive group to include the $20 money order.

Near the end of the Sampling Call, the BL informed the POC of the forthcoming shipment of questionnaires. For the incentive group, the BL reminded the POC that the shipment would include the previously mentioned $20 money order. The questionnaire shipment to the POC for cases in the incentive group contained the money order. It was enclosed in an envelope with the message "Your Special Gift Enclosed" printed on the outside. The payee line on the money order was left blank so that the POC could specify the desired payee. We also enclosed a one-page instructional sheet to help the POC complete the money order.

3. Analysis and Results

In this section, three key areas of analysis are addressed: establishment response rates, employee response rates, and costs—or, more specifically, the effect of the incentive on the nonresponse follow-up effort. All analyses were conducted using unweighted data to determine how the POC incentive affected the reported O*NET response rates, which are also unweighted.

Establishment Response Rate Analysis. For the analysis of establishment response rates, the hypothesis stated in ***Section 2.1*** was tested using the usual normal approximation to the t-test for two means. A two-tailed test was used to permit the unexpected, yet possible, outcome of a reduction in response rate due to the incentive. For this test, the absolute difference between the control group and incentive group response rates was computed as follows:

$$|d_{C\text{-}I}| = |p_C - p_I| \qquad (1)$$

where p_C is the estimated response rate for the control group and p_I is the estimated response rate for the incentive group. If the absolute difference was greater than $1.96 \times s.e.(d_{C\text{-}I})$, where $s.e.(d_{C\text{-}I})$ is the standard error of the difference, then the hypothesis that the two response rates are equal at the 5% level of significance (i.e., $\alpha = 0.05$) was rejected.

The response rate for this analysis was computed by dividing the total number of establishments that participated by all known eligible establishments. Thus, the numerator of the establishment response rate is the total number of establishments that successfully completed the sampling stage, and the denominator is the total number of establishments in the sample minus business- and SOC-ineligible establishments.

It is possible that establishment size and other characteristics of establishments interacted with the incentive treatment. For

Exhibit 3. Independent Variables in the Analysis of Establishment Response Rates

Variable	Definition
Group	1 = Control 2 = Treatment
Industry	1 = Agriculture, Mining, Construction, Manufacturing 2 = Transportation, Wholesale trade, Retail Trade, Finance, Services, Government
Size	1 = 1 to 24 2 = 25 to 249 3 = 250 or more
Urban/Rural	1 = Urban 2 = Rural
SOCs Listed on Selected Occupation List (SOL)	1 = 1 2 = 2 or 3 3 = 4 or more

example, POCs at smaller establishments may have had a different reaction to the monetary incentive than did POCs at larger establishments. The industry associated with an establishment could also play a role in the overall effectiveness of the incentive. For example, POCs in industries that mainly employ office workers may react differently to the incentive than those in industries where non–office workers predominate. Therefore, estimating the effect of the incentive on groups of establishments with common characteristics was of interest in our analysis. However, our ability to evaluate the effectiveness of the incentive on various types of establishments was limited to just a few variables that are available from the sampling frame. These variables are defined in ***Exhibit 3***.

Exhibit 4 shows the results of the comparison of response rates for establishments defined by the characteristics in ***Exhibit 3***. The first column of the table shows the variable being tested, and the second and third columns report the response rates for the control and treatment groups, respectively. The column labeled "Diff" is the difference between the response rates; i.e., $d_{C\text{-}I}$ in (1). The next column reports the standard error of $d_{C\text{-}I}$, and the p-value for the hypothesis test of no difference (i.e., $|d_{C\text{-}I}| = 0$) is reported in the last column. A p-value less than 0.05 indicates the control and treatment response rates are significantly different at the $\alpha = 0.05$ level.

Overall, the control group had a response rate that is about 1.7 percentage points greater than the incentive group—62.0% for the control group compared with 60.3% for the incentive group. The difference is not significant, although it portends what is seen in the subgroup analysis: for 9 of the 10 subgroups compared, the control group has a higher response rate than the incentive group. Note further that none of the differences in ***Exhibit 4*** are statistically significant at the $\alpha = 0.05$ level, indicating that there is insufficient evidence to conclude that treatment and control group response rates differ apart from sampling variation.

Exhibit 4. Comparison of Establishment Response Rates (in Percent) by Treatment for the Independent Variables

Category	Control	Treatment	Diff	SE of Diff	P-Value
Total	61.97	60.27	1.69	1.10	0.12
Industry					
Agriculture, Mining, Construction, Manufacturing	48.90	48.42	0.48	2.70	0.86
Transportation, Wholesale/Retail Trade, Finance, Services, Government	64.71	62.82	1.89	1.19	0.11
Establishment size					
1-24 employees	57.86	56.93	0.92	1.79	0.61
25-249 employees	68.85	66.19	2.67	1.51	0.08
250+ employees	47.69	48.79	-1.10	3.18	0.73
Urban status					
Urban	59.40	57.94	1.46	1.24	0.24
Rural	71.64	69.75	1.89	2.25	0.41
Number SOCs on SOL					
1 SOC	61.31	59.46	1.84	1.48	0.21
2-3 SOCs	57.69	56.74	0.95	2.32	0.68
4+ SOCs	68.26	66.23	2.03	2.28	0.37

These data suggest that the incentive had no effect on response rates for groups of establishments defined by the variables in ***Exhibit 3***. Although the main effects in the exhibit are not significant, there still may be significant interaction effects. That is, response rates for the control and incentive may differ for various combinations of the independent variables. To explore this possibility, a logistic regression model was fitted to the establishment data and all pairwise combinations of the independent variables were simultaneously entered into the model in the form of a three-way interaction with the grouping variable. To further explain the model, consider the model for a single pair of variables, say establishment size (*S*) and industry (*I*). The basic model for these two variables is as follows:

$$\log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = u + u_i^G + u_j^I + u_k^S + u_{ij}^{GI} + u_{ik}^{GS} + u_{ijk}^{GIS} \qquad (2)$$

where p_{ijk} denotes the response rate for the subgroup defined by the i^{th} treatment condition (G), j^{th} industry category (I), and k^{th} size category (S) and the *u*-variables denote the model effects associated with the variables in the superscript labels. Of particular interest is the three-way interaction term in this model, u_{ijk}^{GIS}. This term is used to determine whether the effect of the incentive condition varies by the six combinations of industry and size. If the interaction term GIS differs significantly from 0, then the response rates differ between control and incentive for at least one combination of the categories of S and I. Otherwise, there is no evidence of a difference for any combination of these two variables.

The models that were actually fitted were more complex than (2) since all pairwise combinations of the independent variables interacting with the treatment group, *G*, were simultaneously entered into the model. In addition, since only hierarchical models were considered (for ease of interpretability), all second-order interactions and main effects made up of variables

Exhibit 5. Comparison of Employee Response Rates (in Percent) by Treatment for the Independent Variables

Category	Control	Treatment	Difference	SE of Diff	P-Value
Total	73.55	72.72	0.83	1.0914	0.45
Industry					
Agriculture, Mining, Const, Manufacturing	64.58	69.57	-4.99	4.53	0.27
Transportation, Wholesale/Retail Trade, Finance, Services, Government	74.13	72.95	1.18	1.12	0.29
Establishment size					
1-24 employees	71.46	71.38	0.07	2.55	0.98
25-249 employees	74.29	73.52	0.78	1.24	0.53
250+ employees	68.03	65.92	2.12	4.49	0.64
Urban Status					
Urban	72.17	71.19	0.98	1.30	0.45
Rural	77.22	77.31	-0.09	1.98	0.96
Number SOCs on SOL					
1 SOC	72.02	69.86	2.16	1.68	0.20
2-3 SOCs	72.31	71.75	0.56	2.54	0.82
4+ SOCs	74.92	74.78	0.14	1.65	0.93
Occupation Group					
Management, Business, Mathematics, Engineer	79.83	80.83	-1.00	2.41	0.68
Social Services, Food, Maintenance, etc	70.17	71.3	-1.13	2.22	0.61
Healthcare, Protective Services	59.81	59.49	0.32	4.12	0.94
Social Science, Legal, Education, Arts	74.58	73.01	1.58	1.31	0.23

contained in the three-way interactions were also entered into the model.

Such a large model is over-specified and contains many terms that are not statistically significant. A more parsimonious model is required to strengthen the relationships between the variables in the model and to improve the precision of the statistical tests. To obtain an optimum model, a stepwise elimination model selection process was implemented that deleted the highest-order interaction term in the model whose p-value most exceeded 0.10. After deleting this term, the model was rerun and the elimination process was repeated in a stepwise fashion until either all terms in the model were significant at $\alpha = 0.10$ or only main effect terms remained in the model.

This model selection process produced a model with only main effect terms and a few interaction terms that did not involve *G*. This analysis found that no pairwise combinations of the independent variables produced a difference between the treatment and control response rates that were significant at the 0.10 level or lower. Thus, there is no evidence of any incentive effect for any pairwise combination of the variables in ***Exhibit 3***. To confirm this result, the selection process was reversed. A forward stepwise selection process was implemented which added each three-way interaction involving *G* to the model, including all lower order terms derived from this interaction. Again, only those terms that were significant at $\alpha = 0.10$ were retained. The process continued until all three-way interactions containing *G* had been considered. As in the backward stepwise elimination approach, no interaction terms involving *G* were retained in the model, confirming the earlier finding of no incentive effects for combinations of explanatory variables.

Employee Response Rate Analysis. The analysis of employee response rates mirrored the approach taken for the establishment

Exhibit 6. Final Model for Estimating Incentive Effects for Three-way Interactions

Term[1]	DF	F	P-Value
Industry × Size	2	2.75	0.0637
Industry × Occupation	1	14.06	0.0002
Size × Occupation	6	4.49	0.0002
Group × Size × Urban	2	3.02	0.0490
Group × Urban × Occupation	3	2.96	0.0311

[1] Note: All lower order terms were included in the model, but fit statistics are reported only for the highest-order term involving the variable.

response rates. The employee response rate, for this analysis, was defined as the ratio of the number of returned questionnaires to the number of questionnaires sent out. As for the establishment survey analysis, the incentive effects at the main effect level (marginals for each dependent variable) were first examined, and then logistic regression was employed to assess the interaction effects.

Independent variables used in the employee analysis were: Group, Industry, Establishment Size, Urbanicity, Number of SOCs on Selected Occupation List, and Occupation Group. Note that, except for one variable, Occupation, the variables correspond to those in ***Exhibit 3*** defined for establishments. The categories of Occupation conform to SIC codes that share the same leading digits. Note that POCs may work with employees in several occupations within an establishment, which could attenuate the differences of the incentive effect across occupations. Our analysis made no attempt to account for this form of clustering, however.

The effects of the incentive for each independent variable are shown in ***Exhibit 5***. As for the establishment-level analysis, there is no evidence that the incentive improved response rates for any subgroup defined by a single independent variable. The overall difference, which is not significant, is less than one percentage point in favor of the control group. Of the 13 subgroup comparisons in the exhibit, 9 favor the control group and only 4 are in the direction of higher response rates for the incentive group.

To evaluate the effects of the incentive on employees with characteristics defined by combinations of the independent variables, a logistic regression model like that in (2) was fitted consisting of three-way interaction effects defined by the two independent variables and the treatment variable, *G*. As described for the establishment-level analysis, both forward and backward elimination processes were conducted using the same model fitting rules.

One difference in the employee level analysis is that the clustering of employees within establishments was explicitly accounted for in the model estimation process by treating the establishment as a primary sampling stage. SUDAAN® software was used to appropriately account for the sample clustering effects. As described for the establishment analysis, the model obtained by the backward stepwise elimination approach was confirmed using a forward stepwise selection process. The final

Exhibit 7. Comparison of Model-Based and Design-Based Estimates of Employee Response Rates

Category	Model-Based Estimates			Design-Based Estimates			% of Pop.
	Control	Trt	Diff[1]	Control	Trt	Diff[1]	
Size by Urbanicity							
Size (1-24)							
Urban	75.1	74.3	0.8	72.5	70.0	2.6	22.0
Rural	79.2	77.4	1.8	67.5	75.7	-8.3	7.4
Size (25-249)							
Urban	72.7	72.5	0.2	72.9	72.0	0.9	29.8
Rural	76.0	79.3	-3.3	77.8	78.0	-0.2	12.1
Size (250+)							
Urban	64.0	68.6	-4.6	62.5	64.1	-1.6	23.1
Rural	91.2	74.4	16.7*	88.0	72.2	15.8*	5.7
Urbanicity by Occupation							
Urban							
Management, Business, Mathematics, Engineer	80.4	80.2	0.2	80.8	79.8	1.1	11.4
Social Services., Food, Maintenance, etc.	73.1	74.6	-1.5	68.6	69.6	-1.0	47.1
Healthcare, Protective Services	56.1	55.5	0.6	62.1	58.8	3.3	8.2
Social Science, Legal, Education, Arts	73.4	72.9	0.4	72.5	71.6	0.9	8.1
Rural							
Management, Business, Mathematics, Engineer	77.2	85.7	-8.5*	76.7	84.0	-7.3~	3.3
Social Services., Food, Maintenance, etc.	77.2	80.9	-3.7	73.7	75.2	-1.5	19.2
Healthcare, Protective Services	46.4	60.1	-13.7	47.8	64.5	-16.7~	1.2
Social Science, Legal, Education, Arts	81.5	79.3	2.2	80.2	77.5	2.7	1.6

[1] * - significant at the alpha=0.05 level; ~ - significant at the alpha=0.10 level

model selected for the subsequent interaction effects analysis, shown in ***Exhibit 6***, is the best model in terms of fit and parsimony obtained by the forward and backward selection processes.

As shown in ***Exhibit 6***, two 3-way interactions are significant—Group × Size × Urban and Group × Urban × Occupation. These results indicate that differences between response rates for incentive and control groups were detected for several groups defined by combinations of Size and Urbanicity and Urbanicity and Occupation. To determine which combinations are significant and the directions of the differences, response rates predicted by the final model were estimated, as shown in ***Exhibit 7***.

Exhibit 7 contains three sections that are best viewed simultaneously. The first section of the table, labeled Model-Based Estimates, provides the predicted response rates from the model in ***Exhibit 6*** for control and incentive groups defined by combinations of variables contained in the two significant three-way interactions. The differences in these response rates, the standard error of the differences, and the p-values associated with the test of "no difference" are also included in the table.

Also contained in the table are the corresponding estimates produced from a purely design-based inference (i.e., no explicit model was used to compute the estimates). These estimates are included in the table for comparison with the model-based estimates in order to help interpret and verify the differences in response rates estimated by the model. The designed-based estimates have the advantage that they are not subject to any bias associated with model misspecification since they are not based upon a model. However, the standard errors of the design-based estimates usually exceed those of the model-based estimates, which tend to be more efficient. By comparing both sets of estimates, we can take advantage of the strengths of both estimation approaches.

Finally, the last column of the table contains our best estimate of the proportion of the total population of establishments represented by the row characteristics in the table. This is used to gauge the importance of an observed incentive effect. For example, an effect on response rates for a population subgroup that represents 20% of all employees may be considered more important than an effect on a 5% population subgroup.

First consider the top half of the table, which summarizes the effects for establishments defined by Urbanicity and Size. Both the model-based and design-based estimates indicate that one subgroup has a significant difference between incentive and control—establishments with 250 or more employees in rural areas of the country. Surprisingly, the difference is 16 to17

percentage points in favor of the *control* group, which is consistent for both the model-based and design-based estimates. As indicated in the last column, this is a relatively small group of employees, constituting about 6% of the total population. Still, the result is surprising since it suggests that the incentive had a negative effect on response rates for the employees in these establishments.

The only other group within the Urbanicity × Size interaction that approaches significance is employees in establishments with 25 to 249 employees in rural areas. However, for this group the model-based estimate shows a difference of about 3 percentage points in favor of the incentive (p-value of 0.12). Note, however, that this difference disappears in the design-based table. Note also that employees in establishments with 25 to 249 employees in rural areas constitute about 12% of the employee population.

Next, consider the subgroups defined by Urbanicity and Occupation at the lower half of ***Exhibit 7***. Here the results of the model- and design-based analyses are fairly consistent. The model-based analysis clearly indicates that employees in rural areas in Management, Business and Financial Operations, Computer and Mathematical Occupations, and Architecture and Engineering Occupations were positively affected by the incentive, responding almost 8.5 percentage points higher than their counterparts in the control group. This result is consistent with the design-based analysis, although there the p-value is larger due to the inefficiency of the design-based approach. This group of employees represents only about 3.3% of the O*NET employee population. The design-based analysis also suggests that employees in Healthcare Practitioners and Technical Occupations, Healthcare Support Occupations, and Protective Service Occupations may also have responded at a higher rate incentive, responding almost 8.5 percentage points higher than their counterparts in the control group. This result is consistent with the design-based analysis, although there the p-value is larger due to the inefficiency of the design-based approach. This group of employees represents only about 3.3% of the O*NET employee population. The design-based analysis also suggests that employees in Healthcare Practitioners and Technical Occupations, Healthcare Support Occupations, and Protective Service Occupations may also have responded at a higher rate with the incentive than without (significant at $\alpha = 0.10$). Note that this effect is not supported by the model-based analysis, which indicates that the difference, although considerable at approximately 13.7 percentage points, is not significant. Although a relatively small part of the employee population (about 1%), healthcare professionals are often surveyed and have historically responded at low rates. Thus, the fact that incentives may improve have general survey methodological importance.

Cost Analysis. In other studies (for example, the National Survey on Drug Use and Health, or NSDUH), the use of incentives has reduced costs by reducing the number of follow-up attempts needed to obtain an interview. It is conceivable that the same phenomenon could operate for the O*NET data collection; i.e., the number of follow-up attempts required per completed employee questionnaire could be less for the incentive group than for the non-incentive group.

However, the O*NET data collection protocol design makes it highly unlikely that there could be any real savings of effort even if response rates were substantially improved under the incentive condition. This is because every establishment receives a minimum of four follow-up calls as long as the number of nonresponding employees is one or more. The only scenario wherein these calls could be truncated early is if all sampled employees in the establishment respond prior to completing the fourth follow-up call. Given the very small percentage of establishments that actually achieve that level of participation, it is not surprising that there was essentially no meaningful difference in the number of calls per POC for the two treatments.

Two additional cost measures that were examined include the rate of increase within the cumulative employee response rate and the number of replacement questionnaire packages ordered. If, for example, the incentive group's cumulative employee response rate climbed at a more rapid pace than the control group's, then it follows that BLs spent less time on the phone discussing pending employee responses. Additionally, fewer orders for replacement questionnaires would represent a considerable cost savings in the categories of support labor, printing, materials, and postage.

Completion rates for the treatment and control groups were also compared as a function of the week of data collection. The rate of completion for the two experimental conditions were essentially the same. This suggests that the incentive payment did not increase the speed with which a wave was completed.
A comparison of the rates of ordering replacement questionnaires among the treatment and control groups was also completed, examining all sample waves. For a few waves, the control group required considerably more replacements than the incentive group. However, these were balanced out by the remaining waves, which generally show an opposite effect. Overall, however, the two groups performed similarly, and any differences between experimental groups in the exhibit can be explained by sampling variation. Thus, the incentive offered no savings in reducing the number of questionnaires that needed to be replaced in the process.

4. Conclusions and Outcomes

This experiment considered the effects on establishment and employee response rates of offering the POC a $20 incentive, in addition to the other incentives that the POC receives for O*NET participation. It was hypothesized that this monetary incentive would add to the benefits perceived by the POC for participating in the O*NET program. Since POCs may not fully understand all the requirements of O*NET participation when they are recruited, their commitment might decline as data collection progresses. If they are given a $20 incentive, they might be more committed to the O*NET program and be more motivated to follow up employee nonresponse. It also seemed logical that since employees are offered a $10 payment for completing their

questionnaires, POCs might also expect some kind of payment in addition to the usual O*NET incentives for completing their tasks.

The experimental results provide no evidence that the incentive had any effects on establishment cooperation rates. The POC appeared just as likely to initially agree to participate in the O*NET data collection with the $20 incentive as without it. There are several possible explanations for this. POCs are initially presented with a fairly extensive array of motivating materials and gifts in the early stages of the recruitment process. It is conceivable that the $20 incentive seems a small incremental benefit compared with all the other benefits that are part of participating in the survey. Also, since most POCs conduct their O*NET work with the approval of their supervisors and, presumably, on company time, any additional monetary gift may be viewed as unnecessary or even unwanted by the POCs and their employers. Further, O*NET establishment response rates are already high compared to other establishment surveys, which indicates that the O*NET data collection protocol without incentives may be adequate for maximizing response rates.

Although the POC monetary incentive may not affect cooperation at the establishment level, it could still have an effect on the employee response rate. During the recruitment stage, POCs may not be fully aware of what O*NET participation involves. But as data collection progresses, they may be chagrin to learn of the time commitment required for generating the sample lists of employees, distributing questionnaires, recontacting nonresponding employees, and so on. If, by accepting the $20, the POC feels more obliged to reciprocate by carrying out his/her duties in the later stages of the process, particularly during the nonresponse follow-up stage, employee response rates could be positively affected even though initial response rates are not.

Evidence of any benefit for employee response rates is weak. In general, subgroups that showed a tendency toward a positive incentive effect (such as a few occupations in rural areas) were relatively small compared to groups showing no effect. An inexplicable and pronounced negative effect was also found for employees in large rural establishments—about 6% of all employees. However, in debriefing sessions, BL reports of negative reactions by the POC to the incentive offer were rare and BLs were unaware of any systematically negative effects of the POC monetary incentive. Thus the negative effect is inexplicable and may be regarded as spurious.

Taken as a whole, the employee analysis results suggest weak evidence at best of any possible effect of the monetary incentive on employee response rates. This finding, combined with the lack of evidence of any cost advantage using the incentive, leads to the conclusion that the $20 incentive, as implemented, offered no important benefits to the O*NET data collection.

In interpreting these findings, it is important to note the limitations of the experiment. One limitation is the randomization process used in the study. As previously noted, BL assignments comprised both control and incentive cases, with the latter type making up the majority of a BL's assignment. Such a design is not ideal since it introduces the potential for BL-induced cross-treatment contamination of effects. As an example, if the POC monetary incentive tended to motivate the BLs to improve response rates, then response rates for both the incentive and control groups could improve thereby attenuating the estimated effect of the incentive. This possibility was considered during the design phase of the experiment, but the solution—to randomize the assignment of BLs rather than establishments (i.e., POCs) to treatment and control groups—was deemed operationally infeasible.

A second consideration is the fact that the SOCs in the analysis represent a non-random sample of approximately 15% of all the SOCs that will ultimately be surveyed. Thus, while these results reflect the performance of the incentive on response rates to date, they may not predict the performance of the POC incentive on the more than 600 other SOCs that are not represented in the waves analyzed.

These limitations suggest there could be a small risk that the experiment results do not accurately predict the performance of the incentive in a non-experimental situation. That risk must be weighed against the cost of providing the POC incentive in the remaining waves of the O*NET data collection at a substantial cost. In discussing these findings with the BLs, no concerns were identified with regard to discontinuing the incentive.

Given the considerable cost of providing monetary incentives to the POC and the experiment results that suggest no significant increase in response rates or cost savings, it was decided that the experiment should be discontinued and that all newly recruited POCs should not be offered the monetary incentive.

References

Burnside, R., Bishop, G., and Guiver, T. (2005). "The Effect of an Incentive on Response Rates and Timing in an Economic Survey," Proceedings of the International Statistics Institute, Sydney, Australia.

Church, Allen H. (1993). Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. Public Opinion Quarterly. 57 (1) 62-79.

Dommeyer, Curt J. (1988) How Form of the Monetary Incentive Affects Mail Survey Response. Journal of the Market Research Society. 30(3):379-85.

Keesling, R. (2001). "The O*NET Data Collection Program. Pretest Results of Methodological Experiments on Employee Respondent Incentives," Paper presented at the Annual Meeting of the AAPOR.