Using Local Knowledge during Data Collection: Does It Make a Difference Who Applies it and When?

Rachael Walsh, James Christy, and Jack Marshall[1]
[1]U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 2023

Abstract
As part of a recent reorganization of its data collection activities, the Census Bureau created a new management structure. In addition to consolidating twelve Regional Offices to six, it created several new supervisory positions, like the Field Supervisor (FS) and the Survey Statistician – Field (SSF). The FS directly supervises interviewers within a geographic area, and the SSF manages multiple FSs within their collective area. At inception, only the SSF was granted rights to reassign cases to interviewers across FS areas. Initially, managers in some Regional Offices extended these rights directly to the FS. A uniform policy to reassign cases across FS areas was implemented in all Regional Offices in November 2014. This research assesses this change in management of case reassignment, specifically attempting to determine the effectiveness of the policy of allowing case reassignment by the FS early in the data collection process. After including covariates known to increase the level of effort necessary to resolve a case, as well as those affecting interview completion, multilevel models and logistic regression with random intercepts can use paradata to determine whether this modification of reassignment rights is changing data collection outcomes.

Key Words: reassignment, paradata, multilevel modeling, random intercepts

1. Introduction

As one of the largest data collection agencies in the U.S., the U.S. Census Bureau collects data on behalf of several federal agencies, generating nation-wide estimates. To do so, the Census Bureau data collection infrastructure has a centralized headquarters location in Washington, D.C., and six Regional Offices (ROs) across the Nation. The primary responsibility of the ROs is to oversee the day-to-day activities associated with field data collection operations. This includes managing the work assignments of interviewers. When fielding surveys, supervisors reassign cases from one interviewer to another to balance workload, cover vacant positions or to improve the likelihood of household response - an attempt at refusal conversion, for example. On average, almost one-quarter of cases from all major demographic surveys are reassigned at least once (Walsh and Coombs 2014).

As cases move from one interviewer to another, the cost of collecting data from that particular case increases. The number of contact attempts increases, and in turn, the hours and miles charged by interviewers associated with that case increases. Prior to the reorganization of data collection activities in 2012, statisticians in one of the twelve ROs managed interviewer workload, assignments and reassignments. In a decision linked to cost-savings, the Census Bureau closed half of the regional offices as part of its data collection reorganization. To accommodate the doubling of work in the remaining ROs,

Disclaimer: The views expressed on statistical, methodological, technical, or operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

the Bureau moved the management of interviewers to decentralized statisticians and supervisors. The purpose of this research is to evaluate the recent changes to the way in which we manage the reassignment process.

The realignment created two new, managerial positions in the field – Survey Statistician Field (SSF) and Field Supervisor (FS). Eight SSF areas were delineated in each of the ROs based on geography and population density. Each SSF within an RO is responsible for management of approximately eight to ten FSs. These FSs oversee the data collection for all current surveys in their role as direct supervisors of interviewers.

Although there was a general sense of the function of each position, the ROs were given flexibility in defining the specific roles and responsibilities for the SSF and FS positions. Work assignments were made initially by RO staff and provided to interviewers. Once original assignments were made, SSFs and FSs could make adjustments and reassignments for their assigned areas only. If an FS needed to be reassign a case to an FS area outside of their own area, this required the intervention of an SSF. The FS Area (FSA) boundaries are generally sufficient guidelines for case assignment by design. However, there are situations where it is more cost effective and efficient to cross FSA boundary lines. For example, when a case is located near the boundary line, an interviewer from the adjacent FSA may be closer to the sample unit.

Under the assumption that removing a degree of separation may make the process more efficient, the Census Bureau eliminated the restriction of requiring an SSF to intervene when moving cases across FSAs. In other words, an FS could reassign a case to an interviewer managed by another FS. The purpose of extending these reassignment rights to a lower, more localized level was to capitalize on the local knowledge of both the area and the caseloads of the interviewers. This localization push also provided the lower management level with more autonomy when determining caseload distribution.

The Denver RO was the first to implement this, giving FSs this capability during realignment as part of the FS roles and responsibilities. The other five ROs restricted the ability to reassign cases across FSA to the SSF level during the realignment process. However, in April and December 2013, the New York and Chicago ROs, respectively, granted FSs permission to reassign cases across FSA. Implementation of an official policy occurred in November 2014, and now all FSs in all six offices have this capability.

2. Research Questions and Data Implications

The purpose of this research is to determine whether the application of this local geographic knowledge and case distribution was a cost-effective policy change. We fit multilevel repeated measures models with random effects to existing operations data from a production setting to evaluate the implementation of a non-experimental design. This research addresses the following research questions:

- What is the prevalence of reassignment across Field Supervisor Areas (FSAs)?
- Has the new, more localized policy of allowing the FS to transfer cases to another FSA reduced the cost of data collection?

To answer these research questions, we used paradata from three sources – the Contact History Instrument (CHI), payroll data, and geographic indicators. The CHI collects information from interviewers about the type, timing, and outcome of each contact attempt made to a sample unit. Interviewers are instructed to complete the CHI immediately following each contact attempt. The current interviewer payroll system is an internet-based application, recording the time spent on all data collection activities in 15-minute increments, by survey-specific project code. Interviewer training and current field procedures require interviewers to enter payroll information daily. We used two sources of geographic paradata – the sampling frame and the geographic management structure.

When using paradata to address the research questions, both the autonomy provided to ROs, as well as the paradata structure itself, have implications requiring mitigation. Providing the ROs with autonomy during the data collection process creates two major issues. First, the period for implementing the new policy was not consistent nationwide. Each RO implemented the new policy at a different time in an inconsistent manner that was not monitored. As this was not part of an experimental design, there exists a potential for unobserved bias introduced by other factors not accounted for in the models.

The second factor attributable to RO autonomy that requires mitigation is related to unavoidable differences in caseload distributions by survey by region, which is due to the requirements unique to each of many surveys. RO autonomy when managing data collection is necessary as the Census Bureau does not manage data collection for just one survey, but rather operates as a system of surveys, collecting data on behalf of other federal agencies for the generation of nationally representative key survey estimates.

The third issue requiring mitigation arises from the paradata systems themselves. The current data structure has several associated implications when attempting to make inferential statistical conclusions. First, the contact attempts for each case are collected by interviewers through the use of the CHI. While the CHI provides a rich source of paradata at the contact attempt level for each case, the CHI does not correspond with the current payroll system. In other words, there is no direct way to determine how much the interviewers are charging for the time spent working each case. As such, the contact attempt level paradata must be matched to the daily hours and miles reported by each interviewer. In certain circumstances, aggregates are an acceptable alternative. However, when looking specifically at the cost of reassigning a particular case, this issue requires special attention.

These issue result in limitations for the inferences made from the results of this analysis. Each of these issues were mitigated through a combination of data selection and model specification, which the following sections detail. Despite the application of compensatory measures, a shortcoming of this research is the inability to infer causality associated with the new policy.

3. Data and Covariate Selection

To assess the new reassignment policy, this research analyzed paradata from the major demographic surveys fielded by the U.S. Census Bureau. The interview period for the major demographic surveys included in the analysis are confined to a one month structure. To reduce seasonality effects as well as minimize the potential for unobserved bias resulting from data collection procedures that vary from month to month depending on the Regional Office and across-survey distribution, we selected one month that could be compared across regions based on the policy implementation dates. This also aids in the mitigation of the first two data issues associated with the data implications. Further mitigation with the application of these data was applied through model specifications presented in a subsequent section.

Table 1 shows the policy implementation dates and selected analytic dates by Regional Office. With implementation dates of April, November, and December, the only months available for comparison were January, February, and March. Considering the snowstorms in the northeast significantly impacted data collection efforts, March was selected as the best option.

Table 1. Policy Implementation and Analytic Dates by Regional Office

Regional Office	Implementation Date	Analytic Date
Denver	During Realignment	Excluded
New York	April 2013	March 2013/2014
Chicago	December 2013	March 2013/2014
Atlanta, Philadelphia, Los Angeles	November 2014	March 2014/ 2015

Because reassigned cases by design require additional effort, the analytic sample was restricted to cases that were reassigned from one interviewer to another at least one time. Appendix A provides the overall sample sizes by RO to illustrate the prevalence. Note that the data chosen supports an assessment of the reassignment process overall. The goal of this research is to determine if the localization policy had a positive or negative impact on the operational efficiency associated with collecting data from more difficult cases. In addition to the prevalence of reassignments across FSAs, five level of effort and cost metrics were evaluated: number of contact attempts – sub setting out personal visits as a separate indicator – number of times a case was reassigned from one interviewer to another, and hours and miles charged per case.

3.1 Identifying Across-FSA Case Reassignment

While each case has a geographic indicator identifying the Census tract and block group in which the address is located, our interviewers do not. Instead, when an interviewer works a case, his/her code is attached to the contact attempt record. As interviewers report directly to one FS, the FS area is matched to the case based on the interviewer's assignment. Monitoring the change in interviewer codes and associated FS codes can be used to identify case reassignment.

Each CHI record captures the code of the interviewer recording the entry. *Within*-FSA reassignment occurs when the interviewer code changes from one contact attempt to another but the interviewers report to the same FS, and therefore the FSA code remains the same. *Across*-FSA reassignment occurs when both the interviewer and the FS code change from one contact attempt to the next. This introduces a limitation to this analysis in that it assumes each interviewer assigned the case records a contact attempt in CHI. Despite this limitation, this is the only way to identify case reassignment within the current production data.

3.2 Generating Cost Metrics

The five level of effort and cost metrics were generated using a combination of the CHI and payroll paradata. The number of contact attempts per case is a summation of all telephone attempts and personal visits made to the sample unit during the interview period. For the purposes of this research, personal visit attempts were a separate dependent variable. Personal visits incur more expenses than telephone attempts, as hours and miles are necessary to make a personal visit. One of the goals of the new policy was to reduce the hours and miles charged per case by capitalizing on local knowledge. The FS may not be privy to personnel information of interviewers not under his/her direct supervision – for example an interviewer may be on vacation, or have an unusually high caseload during that interview period. Without corresponding personnel information pertaining to schedules and workload, across-FSA reassignment triggered by the policy may result in additional reassignments. As such, the number of times a case is reassigned is also included as a dependent variable.

Because the payroll system is designed to charge survey-specific project codes, we do not have a direct measure of the hours and miles charged per case. The interviewer code and the date, however, can link the contact history data and the payroll data. When matching by interviewer code and calendar day, six percent of the contact attempt records did not have a corresponding payroll entry for that day.[1] These records contributed to the total number of contact attempts as well as personal visits, but were not included in the models that generated the cost per case metrics.

The same interviewer makes multiple contact attempts each day. We therefore fit a multilevel repeated measures model such that contact attempts were nested within interviewers each day, as seen in Equation 1. Equation 1 is the combined multilevel equation containing both the interviewer- and day-level models (Hox 2010).

$$y_{ti} = \beta_{0i} + \beta_1 x_{1ti} + u_{0i} + u_{1i} x_{1ti} + e_{ti} \qquad [Eq.\ 1]$$

Hours charged (y_{ti}) by the interviewer (i) each day (t) was regressed on the number of personal visits, telephone contact attempts, and miles driven, by survey. The length of time expended on each personal visit is dependent on the outcome of that attempt. As such, personal visits were parsed out as noncontacts, contact but unable to interview, and completed interviews. As hours is a bounded, count variable, we applied a negative binomial regression (Agresti and Finlay 2009). x_{1ti} is the time indicator, while u_{0i} represents the residual error terms at the interviewer level and e_{ti} represents the residual error term at the day level.

For each interviewer, the multilevel model generates subject-specific coefficients, which controlled for the effect of the interviewers' daily planning and strategies. The coefficients were then applied to the data, solving the equations for hours per type of contact attempt by each interviewer on each survey worked.[2] Based on each contact

[1] The missing records were analyzed separately to determine if there were any patterns to the non-matching records. After an extensive review, it was determined that these records were missing at random, therefore exclusion of these records did not negatively affect the models.
[2] Almost three-quarters of Census Bureau interviewers work on more than one survey.

attempt to the case by the interviewer associated with the attempt, the data were then aggregated to the case level to determine the hours per case.

The model mitigated the third data issue – an indirect match of paradata systems to generate cost metrics – but was only used for determining hours per case. The sum of miles reported each day by the interviewer was divided by the total number of cases with personal visit contact attempts by that interviewer on the same day. The miles per case were also aggregated up to the case level, generating the final cost metric.

3.3 Case Characteristics

This research focused on the use of paradata for survey analytics, and therefore did not use any of the survey response data for two main reasons. First, during the data collection process, it is not feasible at this time to incorporate survey response data that have not been processed or edited in real-time evaluations. Second, the use of paradata reduces the potential for nonresponse bias in that paradata are available for both responding and nonresponding households. Using survey response data limits analyses to responding cases, and the focus of this research is operational efficiency pertaining to all cases.

Without using survey response data, the case level information was restricted to the information provided from the sampling frame, like the block group identifier. The block group identifier matched the case to the American Community Survey (ACS) 5-year estimates, which include a tertiary strata indicator of the block group social and demographic characteristics indicative of survey response (Durrant, D'Arrigo, and Steele 2011; Erdman, Adams, and O'Hare 2015; O'Muircheartaigh and Campanelli 1998; Steele and Durrant 2011; West and Olson 2010). For the purposes of this research, this strata indicator was recoded into a dichotomous indicator, flagging cases that were in the two most difficult strata. These case characteristics were included in the models with the aforementioned dependent variables.

The data selected for analysis served as compensatory measures for the three issues previously mentioned. Restricting the data to the month of March mitigated the non-experimental design and inconsistent implementation across ROs. Selecting one interview period for analysis also minimizes the confounding effects resulting from the system of surveys by which the Census Bureau operates. These effects are often seasonal so the restriction minimizes the potential for bias. Generating the cost metrics from models compensates for both the system of surveys as well as the non-integrated paradata systems. Fitting models further mitigates the data issues arising from RO autonomy. The next section details the model fit and mitigation for data implications.

4. Methods

To address both research questions – the prevalence and potential cost savings of across-FSA reassignments – descriptive analytics compared the raw estimates, then models were fit to control for case, interviewer, and regional characteristics. The majority of Census Bureau interviewers work multiple surveys, and each of the surveys differs in length, therefore it was important to acknowledge this in the models for both prevalence and cost savings. To control for the differing survey characteristics, dummy indicators were included for each survey in the model. Including dummy survey indicators also served as a control for variance in caseload distribution regionally, mitigating the system of surveys issue. Further mitigation was applied in the analytic methods for both the prevalence and effectiveness outcomes of interest.

4.1 Prevalence Methods

Descriptively, evaluating the prevalence of across-FSA reassignment was simply looking at the differences in the percent of all reassigned cases that were reassigned across-FSA. To test for statistically significant differences, we applied a logistic regression model with random intercepts, as seen in Equations 2 and 3.

$$log\left\{\frac{\pi_{iA}}{1-\pi_{iA}}\right\} = \eta_{iA} = \beta_0 + \beta_A + \gamma_i \qquad [Eq.\ 2]$$

$$log\left\{\frac{\pi_{iB}}{1-\pi_{iB}}\right\} = \eta_{iB} = \beta_0 + \beta_B + \gamma_i \qquad [Eq.\ 3]$$

where Equation 2 calculated the probability of across-FSA reassignment prior to policy implementation and Equation 3 calculates the same probability post-implementation. The logistic regression models use the random intercepts to account for fixed group effects within the RO (Agresti and Finlay 2009). Accounting for the fixed group effects mitigates RO autonomy and inconsistent implementation. Including the survey dummy indicators controls for the effect of operating as a system of surveys.

4.2 Effectiveness Methods

Descriptively, paired difference tests compared the five metrics pre- and post-policy implementation. To mitigate the issues resulting from a non-experimental implementation design and RO autonomy, a multilevel repeated measures model with random intercepts was fit to the data (refer to Equation 1). The log-link function linked the mean of the response variable to the explanatory variable while confining the response to be nonnegative (Madsen and Thyregod 2011). Separate models were run for each of the five dependent variables, taking into consideration the distribution of each. The first four – contact attempts, personal visits, number of times reassigned, and hours charged – were bounded, count variables and therefore require a negative binomial regression (Agresti and Finlay 2009; Hilbe 2011). The miles per case had a Poisson distribution, and thus the model was fit in the applicable manner.

The repeated measures option in the model made the within-RO comparison, controlling for the regional variation in active surveys and RO autonomy. The random intercepts measured across-RO effects, controlling for the environmental characteristics not included in the model under the assumption that the model is generating the RO effects in comparison to the overall slopes and intercepts, which would account for the potential for bias introduced by excluding environmental controls (Hox 2010). Some of these issues include things like declining survey response rates, increased costs of data collection, political climate, etc.

5. Results

Fitting models to existing operations data provided analytic results that had been mitigated to control for the lack of a randomized, experimental design when implementing the new localization policy. Looking at the descriptive results then fitting models provided a robust interpretation of the potential implications of the new localization policy on the relationship between case reassignment procedures and both the prevalence and cost metrics. While the results are not able to infer causality, we can see the differences in the metrics associated with the reassignment policy in the results from the analyses. Also remember that the results displayed show five of the six ROs since the Denver RO had no base of comparison.

5.1 Prevalence Results

Table 2 shows the percentage of reassigned cases that were reassigned across-FSA relative to the new policy implementation. The prevalence did increase; however, when tested with the random intercepts model, the only RO where the difference was statistically significant was Atlanta, which was also the only RO to see a decrease in the prevalence of across-FSA case reassignment (Refer to Appendix A for applicable sample sizes by RO).

Table 2. Percent of Reassigned Cases with Across-FSA Reassignment, by RO

Regional Office	Pre-Policy	Post-Policy	Percentage Point Difference
New York[a,b]	32.82%	36.47%	3.65
Chicago[a,b]	43.61%	47.93%	4.32
Atlanta[b,c]	33.50%	30.72%	-2.78*
Philadelphia[b,c]	38.67%	42.86%	4.19
Los Angeles[b,c]	40.99%	42.14%	1.15

Note: *indicates statistically significant difference at the $p \leq 0.05$ level.
Source: U.S. Census Bureau Current Surveys Paradata, March 2013[a], 2014[b], 2015[c].

Overall, where increases were seen, it was less than five percentage points in any given RO. Though looking at the Chicago RO, we see that almost half of all reassignments in that region are across-FSA reassignments.

5.2 Effectiveness Results

Turning our attention now to the reassignment process as a whole, this section details the evaluation of level of effort and cost metrics associated with the reassignment of a case from one interviewer to another, not taking into consideration whether the reassignment was across- or within-FSA. Table 3 displays the results from the paired difference test.

While the relationship between the policy implementation and the cost and level of effort metrics was not consistent in all areas, each RO seemed to have consistent results. In the New York RO, all metrics (with the exception of the number of reassignments) saw a decrease with the implementation of the new policy. In Philadelphia, the level of effort was decreased, but none of the cost metrics were statistically significant. In Chicago, the impact of the policy was negative in that the number of times a case was reassigned, as well as the hours and miles charged by interviewers increased. In Atlanta, the implementation of the new policy was associated with an increase in the hours and miles charged by interviewers. Los Angeles had the most negative reaction to the new policy in that all but one measure showed a statistically significant increase – the number of personal visit attempts.

Table 3. Cost Metric Means and Paired Differences, by RO

	New York[a,b]	Chicago[a,b]	Atlanta[b,c]	Philadelphia[b,c]	Los Angeles[b,c]
Contact Attempts					
Pre	6.65	8.15	6.73	7.05	6.41
Post	6.17	8.04	6.58	6.77	6.76
Difference	-0.47***	-0.12	-0.15	-0.28**	0.35**
Personal Visits					
Pre	4.09	4.47	4.16	4.24	4.14
Post	3.66	4.31	4.11	3.74	4.03
Difference	-0.44***	-0.16	-0.05	-0.50***	-0.11
Reassignments					
Pre	2.13	2.26	2.16	2.17	2.14
Post	2.16	2.33	2.15	2.17	2.19
Difference	0.03**	0.07***	0.00	-0.01	0.05***
Hours					
Pre	5.04	3.56	3.64	3.88	3.42
Post	4.01	3.85	3.98	3.76	3.96
Difference	-1.03***	0.29**	0.34***	-0.12	0.54***
Miles					
Pre	59.63	32.19	40.74	37.48	34.20
Post	42.49	39.15	43.32	36.74	38.39
Difference	-17.14***	6.96***	2.58*	-0.74	4.19***
N Cases	2,572	3,142	2,712	3,318	3,436

Note: * $p \leq 0.05$ level, ** $p \leq 0.01$, ***$p \leq 0.001$.

Source: U.S. Census Bureau Current Surveys Paradata, March 2013[a], 2014[b], 2015[c].

Overall, when looking at the descriptive analysis and the paired difference tests for statistically significant differences, only the New York and Philadelphia ROs showed significant decreases in the level of effort and cost in association with the localization of reassignment procedures. In Chicago, Atlanta, and Los Angeles, the implementation of the new localization policy was associated with increases in the level of effort and cost metrics. However, when looking at the data descriptively, we were unable to apply control indicators that previous research found to be predictive of survey response.

Table 4 shows the results from the multilevel repeated measures model, which included a control indicator for the sociodemographic characteristics associated with survey response. When interpreting the results from Table 4, keep in mind that the second level of the multilevel model merely enhances the relationship of the level-one predictors with the outcome variable (Hox 2010). The direction and magnitude of the RO vector tells you how it affects the relationship between time and the dependent variable. Given we are interested in the effects of the policy implementation as denoted by the time indicator, we focus the discussion of the results on the statistically significant RO level effects.

Only six RO level effects were statistically significant. Chicago saw positive, significant effects with respect to the relationship between the time indicator and contact attempts, and, to a lesser extent, personal visits. In other words, the negative relationship between time and level of effort was enhanced in Chicago – the policy implementation may have contributed to an increase in the number of contact attempts made in Chicago – a finding consistent with the paired difference test results.

Table 4. Repeated Measures Models Regressing Cost Metrics on Cases, by RO

	Attempts	Personal Visits	Reassignments	Hours	Miles
Intercept	1.679***	0.447**	0.680***	1.103*	6.337
	(0.068)	(0.085)	(0.063)	(0.341)	(5.256)
Difficult Area	0.113***	0.219***	0.029***	1.370***	9.063***
	(0.008)	(0.010)	(0.009)	(0.046)	(0.637)
Time	-0.003	-0.052***	0.013	-0.019	-0.541
	(0.008)	(0.009)	(0.009)	(0.044)	(0.613)
New York	-0.045	-0.040	-0.022	0.492***	11.717***
	(0.041)	(0.024)	(0.015)	(0.160)	(3.163)
Philadelphia	0.019	-0.013	-0.007	-0.035	-2.681
	(0.040)	(0.023)	(0.015)	(0.159)	(3.155)
Chicago	0.147***	0.078***	0.047***	-0.083	-3.743
	(0.040)	(0.023)	(0.015)	(0.159)	(3.152)
Atlanta	-0.062	0.006	-0.006	0.076	0.303
	(0.041)	(0.024)	(0.016)	(0.161)	(3.174)
Los Angeles	-0.058	-0.032	-0.012	-0.450**	-5.596
	(0.041)	(0.023)	(0.015)	(0.159)	(3.157)

N=24,708 cases with a maximum of 5,731 per group (RO).
Note: *indicates statistically significant difference at the $p \leq 0.05$ level.
Source: U.S. Census Bureau Current Surveys Paradata, March 2013[a], 2014[b], 2015[c].

In the New York RO, the negative effects of the time indicator show that the relationship between the time indicator and the number of hours and miles charged by interviewers was significantly enhanced. In the New York RO, interviewers charged fewer hours and miles after the policy was implemented. Again, the model supports the finding from the descriptive analysis. Los Angeles, unfortunately, saw the opposite relationship between policy implementation and hours charged by interviewers. The regional effects in Los Angeles were negative, which decreases the relationship between the number of hours charged by interviewers and the time indicator, which was negative.

The repeated measures models support the findings from the descriptive analysis and paired difference tests. Overall, the localization of the reassignment policy was associated with decreased level of effort and cost in the New York and Chicago ROs, but a increases in the Los Angeles RO. The policy and efficiency associations were not statistically significant in either the Philadelphia or Atlanta ROs.

6. Discussion

The combined findings regarding the nationwide implementation of the policy were inconclusive. At most, 25 percent of cases are reassigned in any given interview period. Between one-third and one-half of that subset are then reassigned across-FSA. Even when using all current surveys fielded by the U.S. Census Bureau nationwide, because the regional effects need to be considered, the sample size with which to test the new policy is never greater than 4,000 cases in any area. While this could be further reduced if restricting the analysis to comparisons between within- and across-FSA reassignments, we do not have a way to empirically test the need to reassign cases across FSA boundaries. Working within a system of surveys, with many moving parts, as well as unobservable factors, this research was only able to assess the implications of policy implementation on the reassignment process as a whole.

Despite this limitation, we were able to show some relationships between the policy implementation and both level of effort and data collection cost metrics using existing data from a production environment without an experimental design. The results at the RO level were consistent across measures, suggesting that the benefits of localizing the reassignment policy were dependent on the regional policy prior to implementation. One specific example pertains to the differences in the data collection management styles between the New York and Los Angeles ROs.

In the New York RO, initial case assignment is based solely on the FSA boundary lines. In the Los Angeles RO, however, RO staff has both the time and the capacity to perform geo-spatial analyses prior to making the initial case assignment. In other words, in the Los Angeles RO, case assignment is made irrespective of FSA boundary lines and instead is based on the geographic distance between the closest interviewer and the sample unit.

The localization policy was designed to increase consistency across ROs in making initial case assignments based on proximity without the restriction of FSA boundaries. This policy is beneficial only under the assumption that the local knowledge and efficacy exceeds that of the RO level. Unfortunately, given the workload, knowledge base, and available software and data, the FS may not have the same capabilities as the RO staff for such a policy implementation to be effective. In situations like New York, the decrease in level of effort and cost of data collection following the policy implementation may be explained, at least in part, by the substantial application of local knowledge. Presumably, when the localization of the case reassignment removes an analytic component designed to reduce costs, as in situations like Los Angeles, the effect of the policy on cost and level of effort is diluted. The results of this analysis support these assumptions, but causal inferential statements regarding policy are not possible without a randomized experimental design.

7. Conclusions

This research demonstrated the applicability of existing production paradata for analytic use of policy implementation when experimental design is or was not an option at the time of implementation. With sufficient inputs, and nationally representative data, assumptions can be made regarding unobserved characteristics that render the potential for bias minimal. In addition to providing an example demonstrating this feasibility of fitting models to existing data, this research also contributes to the literature with respect to the importance of regional effects when implementing data collection policies for nationally representative survey designs. In our example, the regional effects were significant contributing factors to the analysis of the localization policy due, impart to the existing policies and procedures in the area. Our findings suggest that the utilization of local knowledge may have enhanced operational efficiency during data collection. However, it may be possible to achieve similar results through geo-spatial analyses at higher levels within the organizational structure, preventing the addition of burden to lower managerial positions. Despite the inability of our research to infer causality, it demonstrates the importance of including regional effects when designing national policy as well as fitting models to nationally representative data.

References

Agresti, A. and Finlay, B. (2009). Statistical Methods for the Social Sciences, 4th Ed. Prentice Hall, Inc. Upper Saddle River, NY.

Durrant, G.B., D'Arrigo, J. and F. Steele (2011). Using Field Process Data to Predict Best Times of Contact Conditioning on Household and Interviewer Influences. *Journal of the Royal Statistical Society*, Series A 174 (4): 1029-1049.

Erdman, C., Adams, T. and O'Hare, B. (2015). Development of Interviewer Response Rate Standards for National Surveys. *Field Methods*.

Hilbe, J. (2011). Negative Binomial Regression, 2nd Edition. Cambridge, UK: Cambridge University Press.

Hox, J. (2010). Multilevel Analysis: Techniques and Applications Quantitative Methodology Series, 2nd Edition. Routledge. New York, NY.

Madsen, H., and Thyregod, P. (2011). Introduction to General and Generalized Linear Models. Chapman & Hall/CRC.

O'Muircheartaigh, C., & Campanelli, P. (1998). The relative impact of interviewer effects and sample design effects on survey precision. *Journal of the Royal Statistical Society Series A (Statistics in Society), 161*(1), pp. 63-77.

Steele, F. and Durrant, G. (2011). Alternative Approaches to Multilevel Modelling of Survey Non-Contact and Refusal. *International Statistical Review*, 79(1), 70-91.

West, B. and Olson, K. (2010). How Much of Interviewer Variance is Really Nonresponse Error Variance? *Public Opinion Quarterly*, 74(5), 1004-1026.

Walsh, R. and Coombs, J. (2014). Case Reassignment: When making contact is a two-person job. In *JSM Proceedings,* Survey Research Methods Section. Alexandria, VA: American Statistical Association, 3690-3700.

Appendix A. Sample Size by RO and Reassignment Status

	Pre-Policy				Post-Policy			
Regional Office	All Cases	Eligible Cases	All Reassignments	Across-FSA Reassignment	All Cases	Eligible Cases	All Reassignments	Across-FSA Reassignments
New York[a,b]	25,784	19,116	1,531	748	24,529	18,002	1,634	938
Chicago[a,b]	25,941	18,584	1,721	1,331	24,861	17,417	1,636	1,506
Atlanta[b,c]	24,823	17,250	1,330	670	27,575	19,584	1,879	833
Philadelphia[b,c]	24,483	18,084	1,613	1,017	25,156	17,963	1,896	1,422
Los Angeles[b,c]	24,461	19,045	1,578	1,096	24,510	19,107	1,988	1,448

Note: *indicates statistically significant difference at the $p \leq 0.05$ level.
Source: U.S. Census Bureau Current Surveys Paradata, March 2013[a], 2014[b], 2015[c].