

# Characteristics of Falsified Interviews<sup>1</sup>

Chrishelle Lawrence, US Census Bureau  
Elizabeth Love, US Census Bureau

## Abstract

Data falsification occurs when an interviewer intentionally deviates from established survey interviewing procedures. Falsification may be influenced by several factors, such as interviewer experience level and workload, survey subject matter, location, and length of survey. While past research projects have focused on the impact of these factors individually, few studies have examined the joint association between these characteristics and data falsification. The current research is an exploratory analysis that focuses on the relationship between the above-mentioned factors and data falsification. Results from formal investigations of data falsification on the Census Bureau's major demographic surveys from 2005 to the present will be analyzed along with respondent traits. Findings from this analysis will be used to generate a profile of falsified interviews.

**Key Words:** Data falsification, data quality

## 1. Introduction

Data quality continues to be a major concern for survey researchers. Shrinking response rates and coverage errors are common factors that negatively impact data quality. Yet, few other determinants directly undermine quality more than falsification. Deliberately misrepresenting sample units (e.g. recording a sample unit as a vacant lot when it is actually a habitable home), intentionally removing names from a household roster, and fabricating an entire interview are all examples of falsification that damage data quality.

Survey researchers and analysts focus their attention on deterring, detecting, and mitigating falsification and its damaging affect on data quality. Data analysts have employed several techniques to uncover falsified data and interviewers that falsify. Many researchers have tried a wide array of statistical methods to detect falsified data. These analytical approaches include identifying outliers on influential variables and paradata such as time stamps, production over time, and survey items that control large differences in skip patterns (Bushery et al., 1999). Other quantitative techniques include comparing the frequency of numeric responses to well-known distributions or mathematical theories and statistical quality control techniques (Porrás et al., 2004, Schäfer et al., 2004, and Biemer et al., 1989).

Survey managers who work closely with interviewers attempt to profile problem interviewers through observation and reinterview. Although supervisors silently monitor interview-respondent interactions primarily to check that interviewers adhere to survey procedures, these non-intrusive reviews can also provide hard evidence against

---

<sup>1</sup> Disclaimer: Any views expressed are those of the authors and not necessarily those of the U. S. Census Bureau.

interviewers that cheat. Reinterview, a second contact with a survey participant to verify information obtained in the original interview, is a frequently used quality control method. Based on the findings from reinterview, survey managers can confirm high-level details such as household roster accuracy, sample unit status (i.e. vacant or demolished), and that an interview actually took place (“Interviewer Falsification in Survey Research”, 2004). Major statistical organizations search for trends amongst falsifying interviewers. Research from interviewer management studies identified survey characteristics that might be associated with inclinations to falsify data. Interviewers with more experience were hypothesized to falsify less noticeably so quality assurance programs were designed to have higher sampling rates for more experienced interviewers (Wetzel and Hood et al., 1993).

Several studies employed statistical strategies to characterize falsified data, yet few try to profile falsified interviews by characterizing interviewers, respondents, and other paradata. Data extracted from a quality assurance form that details the findings from formal investigations of data falsification were examined. These investigations were performed on interviews from major demographic surveys conducted by the US Census Bureau from 2005 to 2009. The current research focuses on profiling falsified interviews by examining the relationship between several factors, such as survey subject matter and interviewer experience level, and the likelihood to falsify.

## 2. Methodology

Whenever a case in an interviewer’s assignment is suspected of falsification, an investigation must be conducted to determine whether the case contains fabricated information. Investigators must also find out if there was a failure to follow standard survey procedures. All details of the investigation must be documented in the *Quality Assurance Form* (QAF). The QAF allows the regional offices (ROs) to fully document the details of their investigations. At the conclusion of the investigation, a final judgment about the suspicion of falsification is made as well as what actions will be taken against the interviewer (if any).

When first completing the form, the investigator notes information about the interviewer, survey, time of suspected falsification, and the interviewer’s assignment size. Next, details on what initially caused the suspicion of falsification are documented. Usual causes of suspicion include reinterview results, contact from respondents, and suspicious data. For each case suspected of falsification, the investigator must make a decision on both falsification and failure to follow survey procedures. Reasons for confirmation of falsification include misclassifying sample units to avoid interviewing and entering fictitious survey data. Reasons for confirmation of failure to follow survey procedures include not using a laptop for interviewing and not following respondent self-response rules. Investigators must also specify the techniques used during the course of the investigation, such as contacting the original respondents or asking the suspected interviewer for further explanation of their work. In addition, details must be provided on the characteristics of cases that are confirmed of falsification (discrepancies, timing, patterns, and extent of falsification) as well as those confirmed for failure to follow established survey procedures.

After the investigation is completed, final action will be taken against the interviewer who is suspected of falsification. An interviewer can be cleared, confirmed, or still suspected of falsification. Possible actions include termination, additional training,

observations, or no action (if the interviewer is cleared of data falsification). The RO has 60 days to complete the investigation and notify Census headquarters of their findings.

All data on the QAFs are then manually entered into an electronic database. The data used in this analysis are taken from the 2005 version of the QAF. Although there have been some slight formatting updates to the form since then, the questions have remained the same. Therefore, the data are consistent within this five-year period of analysis. An independent reviewer validated the reliability of the data. For the purposes of this analysis, the final outcome was categorized as a binary variable by collapsing cases that were either cleared or still suspected into the not-confirmed category.

Data collected from the QAF are characteristics of one of the following four categories: interviewers, respondents, surveys, and location. Features of interviewers included experience level, workload, prior offenses (single or repeated), and whether they followed procedures. Experience level was measured in years and modeled as binary and multilevel in separate models to determine which would work best. Workload was described in two ways: the total number of cases the interviewer had at the time of the investigation and the number of surveys the interviewer worked concurrently. Both of these workload dimensions were given at the start of the investigation. Respondent characteristics were age (elderly or not), language barrier (yes or no), and hard to reach (yes or no). Respondents can be hard to reach for interviewing or physically hard to reach because they live in buildings or homes that are difficult to access (e.g., gated communities). It is important to note that survey data were not used to determine respondent characteristics. These characteristics are based on the opinion of the investigator who completed the QAF. The only survey attribute was periodicity (one-time or panel). Originally, surveys were examined to see if they could be grouped according to subject matter, however, none of the surveys were similar enough to be grouped together. A list of the surveys is appended in Attachment A. Lastly, the environment category included region (West, Northeast, Midwest, and South), seasonality (summer, spring, winter, and fall), urbanicity (urban or rural), and income level (low income or not). Urbanicity and income level were determined by the investigator.

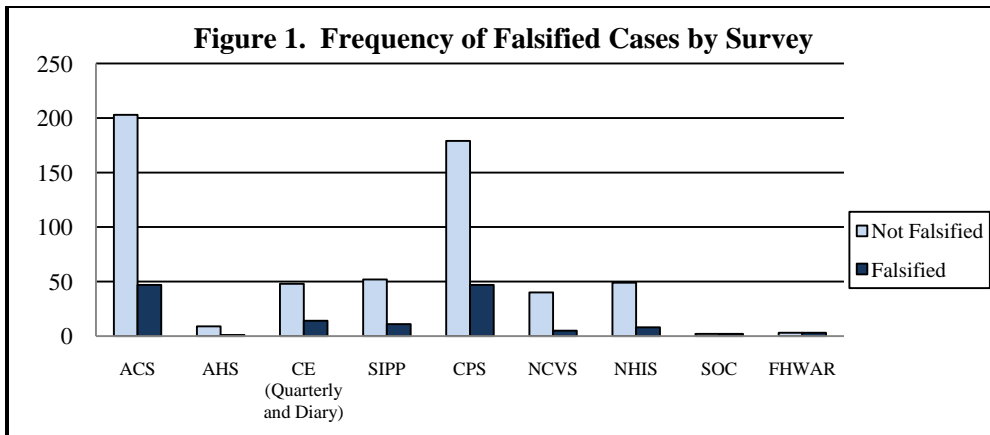
### **3. Statistical Analysis**

Several predictors were considered for model selection. All predictors were tested to gauge how strongly each explanatory variable was associated with the propensity to falsify. Results from this preliminary analysis are not shown but aided model selection. An unconditional logistic regression model was built using a backwards elimination process. This model selection procedure forced a first order term to be in the model if it was a factor in an interaction term. Only second order interaction terms were considered. The binary response variable was confirmed of falsification (yes or no). The significance level was set at 0.10, the Census Bureau standard. Rather than using a more complicated model to account for dependent observations from interviewers that were investigated more than once (up to eight times), one case was randomly selected to account for the lack of independence.

### 4. Results

The falsification database contained the findings of 1,095 investigations from 2005 to 2009. The work of 791 interviewers was covered in this time period and resulted in 242 confirmed cases. Most interviewers had a single allegation of misconduct, but 10% (73 out of 791) of them had four to eight prior suspicions. After accounting for missingness and removing additional observations at random for interviewers with multiple investigations, there were 735 cases left for analysis.

Of the 735 observations, 143 cases ended in confirming suspicions of falsification. Frequency of confirmed cases ranged from 16 to 30% of all suspected cases each year with the highest rate of occurrence in 2005 and the lowest in 2007 (see Table 1). Aggregated counts revealed that the American Community Survey and the Current Population Survey had the most cases suspected and confirmed cases with 47 out of 250 and 47 out of 226 respectively (see Figure 1). It is also important to note that the American Community Survey and Current Population Survey conduct the most interviews of all the demographic surveys. As a result, we would expect their number of suspected and confirmed cases to be much higher than the other surveys.



**Table 1. Frequency of Falsified Cases by Year**

	2005	2006	2007	2008	2009	Total
Confirmed Falsification	14	41	29	40	19	143
Total Cases	46	223	174	173	119	735

#### 4.1 Preliminary Results

Interviewers were more likely to falsify when households were in areas that were urban, low income, or physically hard to reach. Additionally, interviewers’ tendency to cheat increased when respondents were elderly or had language barriers. Interviewers were less likely to falsify in the fall and most likely to falsify in the spring.

The distribution of falsifying interviews for several other predictors seemed evenly spread. One-time surveys were not more likely to have more interviewers that falsified data than panel or longitudinal surveys. Interviewers who had larger workloads or worked on more surveys concurrently did not falsify more than those who had smaller workloads and worked on less surveys.

Of the cases where interviewers were confirmed of falsification, a quarter (36 out of 143) identified any extenuating circumstances that contributed to their delinquency. Interviewers most often claimed personal or family circumstances (e.g. illness, death) and deadline pressures as mitigating factors that caused them to falsify (21 and 13 respectively). Very few interviewers blamed their misbehavior on cumbersome workloads or time conflicts with another job (4 and 5 respectively).

#### 4.2 Model Selection

The backwards elimination procedure yielded a model with the following categorical variables: failure to follow procedures (yes or no), prior offenses (yes or no), experience level, and region. Additional testing was done to determine how experience level should be categorized.

##### *Experience Level*

Experience level was modeled as binary and multilevel. Several models were tested to identify which group of interviewers was most likely to falsify. In all three treatments, prior offenses, failure to follow survey procedures, and region were included in the model, and all of these variables were significant, including experience level. Initially, experience level was treated as a binary variable (experience less than or greater than or equal to 5 years). In this model, interviewers with less than 5 years of experience were more likely to falsify. The major disadvantage of this model is that with only two groups, specificity is lost.

**Table 2. Treatment of Experience Level**

Treatment	Model Results
<b>Binary</b> < 5yrs ≥ 5yrs	<ul style="list-style-type: none"> <li>Less experienced interviewers were more likely to falsify</li> <li>No significant interaction terms</li> </ul>
<b>Multilevel (A)</b> < 2 yrs 2-5 yrs 5-10 yrs ≥ 10 yrs	<ul style="list-style-type: none"> <li>Less experienced interviewers were more likely to falsify</li> <li>Significant interaction between experience and following procedures</li> <li>Parameter estimates become less interpretable</li> </ul>
<b>Multilevel (B)</b> < 2 yrs 2-10 yrs ≥ 10 yrs	<ul style="list-style-type: none"> <li>Less experienced interviewers were more likely to falsify</li> <li>Significant interaction between experience and following procedures</li> </ul>

Next, experience level was treated as a multilevel variable with four levels: less than 2 years, 2-5 years, 5-10 years, and greater than 10 years. Similar to the binary treatment, less experienced interviewers were more likely to falsify. There was also a significant

interaction between experience level and failure to follow survey procedures. However, the parameter estimates became less interpretable. Rather than having an inverse linear relationship between experience level and the likelihood to falsify, there was a quadratic one. It was expected that as experience level increased, the likelihood to falsify would decrease, but in this model, the likelihood to falsify was more volatile.

In this multilevel treatment of experience level, interviewers with 2-5 years of experience and 5-10 years were not significantly different. Consequently, a third treatment of experience level was considered with three levels: less than 2 years, 2-10 years, and greater than 10 years. Again, less experienced interviewers were more likely to falsify, and there was a significant interaction between experience level and failure to follow survey procedures. These variables, including the interaction term determined the final model used.

### 4.3 Model Results

Interviewers with prior offenses were more likely to falsify data (log-odds=0.54, CI: [0.17, 0.90]). Interviewers who followed procedures were also significantly linked to higher log- odds of falsifying of data (log-odds =0.60, CI: [0.06, 1.15]). Interviewers in the Northeastern and Southern regions of the country had similar tendencies to falsify. For those areas, interviewers were more likely to fabricate than those in the Western and Midwestern parts of the nation. Less experienced interviewers were more likely to falsify. The group most likely to falsify was less experienced interviewers who failed to follow procedures (See Table 3).

**Table 3. Model Estimates**

<b>Factor</b>	<b>Log Odds</b>	<b>90% CI</b>
Prior Offenses++ (None is the reference category)	0.54	[0.17, 0.90]
Failure to Follow Procedures++ (No is the reference category)	0.60	[0.06, 1.15]
Region (West is the reference category)		
Midwest	-0.34	[-0.83, 0.14]
Northeast++	0.62	[0.17, 1.08]
South++	0.56	[0.10, 1.02]
Years of Experience (< 2 yrs is the reference category)		
2 ≤ Experience < 10 yrs++	-1.08	[-1.72, -0.45]
Experience ≥ 10 yrs++	-2.09	[-3.60, -0.98]
Procedures*Experience		
Procedures * 2 ≤ Experience < 10 yrs++	0.96	[0.21, 1.72]
Procedures* Experience ≥ 10 yrs	1.11	[-0.27, 2.77]

++ Indicates significant association

## 5. Discussion

In generating a profile of falsified interviews, there were several variables to consider. Interviewer characteristics were considered first. Heavy workloads could increase the likelihood of falsification because interviewers may feel pressure to complete their interviewing assignments and get high response rates. Workload was analyzed in two ways: the number of surveys worked concurrently and the number of cases in the interviewer's assignment. Neither of these factors was significant.

Next, survey characteristics were reviewed. The type of survey could increase the likelihood of data falsification if the survey contains sensitive questions or if the interviewer has experienced respondent reluctance or refusals. Surveys used in this analysis covered the following topics: census, housing, economics, employment, crime, health, construction, and recreation. The periodicity of the survey was also considered. Surveys were either one time interviews or panels (multiple interviews). Neither the subject matter of the survey nor its periodicity was significant. Environmental variables to consider were seasonality, urbanicity, and low income area (yes or no). The summer and winter months have the most holidays, and in addition, many people take vacations during these months. Interviewers may be inclined to falsify or shortcut the interviewing process in order to allow more time for family or special events. Seasonality was insignificant. Urbanicity and whether or not the respondents live in a low income area were also insignificant.

Lastly, respondent characteristics were analyzed. For every interviewer confirmed for data falsification, the investigator must note any patterns or characteristics of the respondents in the falsified units on the QAF. The following respondent characteristics were considered for the model: age (elderly or not), language (language barrier or not), and physical barriers to the residence. None of these respondent characteristics were significant.

No demographic or administrative data were used in this analysis. This lack of information limited the depth of research. For example, the number of languages spoken by an interviewer may be correlated with falsifiers that experienced language barriers. Also, older interviewers may be less likely to falsify data for elderly respondents. Neither of these conjectures could be validated with the current data. These and other interviewer effects cannot be fully explored without descriptive data for the interviewers. Additionally, this research only considered falsification investigations within the past five years. Expanding the research with more years of data may uncover additional factors related to falsification. Lastly, the data from the QAF can be enriched by consulting other sources to get more information on the nature of the surveys. For example, the average time to complete the survey could possibly be a better indicator when considering survey characteristics.

While the purpose of this research is to profile falsified data, surveyors cannot avoid wondering about the more salient issue: why do interviewers falsify? The obvious response is to avoid doing work, but this simple explanation is unsatisfying and warrants closer examination. In an effort to address that issue, the QAF includes a section where the investigator can list any extenuating circumstances that placed the interviewer under duress and caused him/her to cheat. Suggested responses include deadline pressures (needing more time to start and/or follow-up cases), performance pressures, personal/family issues, time conflicts with other job(s), and cumbersome workload (too

many cases or very difficult cases). Only 36 out of 143 confirmed cases reported any mitigating circumstances. Moreover, investigators for 29 confirmed cases stated that they did not know if there were any special reasons for the falsification, and another 20 investigations did not have a response recorded. Other researchers have attempted to answer the same question. The consensus amongst other researchers that have attempted to explain why interviewers falsify is that performance ratings are partially based on response rates (Crespi and Keicker). Interviewers who want the benefits and privileges awarded to those with high performance ratings may cheat because it is easier and unlikely they will be caught. Others may have very difficult respondents and find it unfair to be penalized for their refusal to participate. There may be other valid reasons why interviewers falsify. More research, particularly cognitive interviews of interviewers, need to be conducted to make that determination.

### References

- Biemer, Paul P. and S. Lynne Stokes. "The Optimal Design of Quality Control Samples to Detect Interviewer Cheating." Journal of Official Statistics 5.1 (1989): 23-29.
- Bushery, J. M., Reichert, J. W., Albright, K. A., and Rossiter, J. C. (1999). Using date and time stamps to detect interviewer falsification. In *Proceedings of the Survey Research Methods Section* (pp. 316-20). Alexandria, VA: American Statistical Association.
- Crespi, Leo P. "The Cheater Problem in Polling." Public Opinion Quarterly. Winter (1945): 431-445.
- Hood, C. C. and Bushery, J. M. (1997). Getting more bang from the reinterview buck: Identifying at risk interviewers. In *Proceedings of the Survey Research Methods Section* (pp. 820-24). Alexandria, VA: American Statistical Association.
- "Interviewer Falsification in Survey Research: Current Best Methods for Prevention, Detection, and Repair of Its Effects." Survey Research: Newsletter from the Survey Research Laboratory 35.1 (2004): 1-5.
- Kiecker, P. L. and J. E. Nelson. "Marketing Research Interviewers and Their Perceived Necessity of Moral Compromise." Journal of Business Ethics 15 (1996): 1107-1117.
- Porras, J., English, N. 2004. Data-Driven Approaches to Identifying Interviewer Data Falsification: The Case of Health Surveys. In *JSM Proceedings, Survey Research Methods Section*. Toronto, Canada: American Statistical Association. 4223-4228.
- Schäfer, C., Schräpler, J., Müller, K., and Wagner, G. 2004. Automatic Identification of Faked and Fraudulent Interviews in Surveys by Two Different Methods. In *JSM Proceedings, Survey Research Methods Section*. Toronto, Canada: American Statistical Association. 4318-4325.
- Schreiner, I., Pennie, K., and Newbrough, J. 1988. Interviewer falsification in Census Bureau Surveys. In *JSM Proceedings, Survey Research Methods Section*. Washington, DC: American Statistical Association. 491-496.
- Wetzel, A.J. "Falsification by Field Representatives 1982-1992." Census Bureau Memorandum for Paula Schneider: 1993.



### **Acknowledgements**

The authors wish to acknowledge Candice A. Barnes and Aref N. Dajani for their assistance in this research.

**Attachment A**  
**Demographic Surveys Used in this Analysis**

**One-Time**

American Community Survey (ACS)

American Housing Survey (AHS)

National Survey of Fishing, Hunting, and Wildlife-Associated Recreation (FHWAR)

National Health Interview Survey (NHIS)

Survey of Construction (SOC)

**Panel**

Consumer Expenditure Survey (CE)

Current Population Survey (CPS)

National Crime Victimization Survey (NCVS)

Survey of Income and Program Participation (SIPP)