Inferential Methods to Identify Possible Interviewer Fraud Using Leading Digit Preference Patterns and Design Effect Matrices

Moon J. Cho, John L. Eltinge, David Swanson,
U.S. Bureau of Labor Statistics
Moon J. Cho, 2 Massachusetts Avenue NE, Washington, DC 20212
(Cho_M@bls.gov)

Key Words: Benford's Law, Cluster sample, Consumer Expenditure Survey, Curbstoning, Misspecification effect, Rao-Scott adjusted goodness-of-fit test .

1. Introduction

Interviewer fraud can damage the data quality severely. How can we detect it? Turner et al. (2002) used response patterns to detect falsification. They reported that suspected falsifiers could be noticeable by an unexpectedly high yield of interviews per assigned sample address, and/or unusual response rates for specific reported variables on behaviors. Turner et al. also discussed the systematic differences between suspected falsifiers and other interviewers in providing the verification means, such as telephone numbers of the respondents. Biemer & Stokes (1989) proposed a statistical model for describing dishonest interviewer behavior, which was applied to a general quality control sample design and several associated cost models. A 1982 U.S. Bureau of the Census study indicated a higher degree of cheating in urban areas (Biemer & Stokes). The study also shows a substantial and highly significant tendency for relatively inexperienced interviewers to cheat more frequently for the two largest demographic surveys, the Current Population Survey and the National Crime Survey (Biemer & Stokes).

We used the leading digits to detect curbstoning in this paper. The effect of the sampling design, such as stratification and clustering, on standard Pearson chi-squared test statistics for goodness of fit is investigated.

Statistical methods for analyzing cross-classified categorical data has been extensively developed under the assumption of multinomial sampling. However, most of the commonly used survey designs involve clustering and stratification and hence the multinomial assumption is violated (Rao & Scott, 1981). Literature has shown that clustering can have a substantial effect on the distribution of the standard Pearson chi-squared test statistic, χ^2 and that some adjustment to χ^2 may be necessary, without which one can get misleading results in practice (Rao & Scott, 1981). Rao & Scott developed a simple correction to χ^2 which requires only the knowledge of deffs (or variance estimates) for individual cells in the goodness of fit problem (Rao & Scott, 1981). The original Rao & Scott papers considered inference for one vector of proportions, based on (essentially) one sample.

In this paper, we are considering inference for a large number of proportion vectors $p_i, i = 1, ..., I$, where I is the total number of interviewers. Only a small portion of an interviewer's workload can be verified because of the limited resources. Therefore, we addressed the optimum allocation of resources such as re-interview time using the optimal decision rule.

2. Data Collection in Consumer Expenditure Survey

Consumer expenditure surveys are specialized studies in which the primary emphasis is on collecting data relating to family expenditures for goods and services used in day-to-day living (BLS Handbook, 1997, p.160). The current survey consists of two separate surveys, each with a different data collection technique and sample: In the Interview survey, each consumer unit (CU) in the sample is interviewed every 3 months over five calendar quarters. The sample for each quarter is divided into three panels, with CU's being interviewed every 3 months in the sample panel of every quarter. The interviewer uses a structured questionnaire to collect both demographic and expenditure data in the Interview survey. In the Diary (or recordkeeping) survey, demographic data is collected by the interviewer, whereas expenditure data is entered on the diary form by the respondent family for two consecutive 1-week periods (BLS Handbook, 1997, p.161).

The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the U.S. Bureau of Labor Statistics.

Both surveys are conducted by personal visits with telephone usage limited to appointment scheduling and follow-up calls for information missed at the time of the proposed interview. The current method for identification of possible interviewer fraud or gross reporting error in the Consumer Expenditure Survey is a reinterview program. The reinterview is conducted by a member of the supervisory staff. A subsample of approximately 6 percent of households in the Interview survey and 17 percent in the Diary survey are reinterviewed on an ongoing basis (BLS Handbook, 1997, p.161).

3. CES Data Overview

Swanson et al.(2003) analyzed overall proportions of leading digit, and the related covariance matrices. We studied the overall proportions in detail in this paper. Our first question was whether we see any distinct patterns in the overall proportions with respect to quarters. Table 1 shows ratio estimates from Quarter 2 to Quarter 5. The table also shows the values of the standard errors from the balanced repeated replication method (SE_{BRR}) with 44 replicate weights, and the values of the standard errors that one can obtain from multinomial assumption on the underlying observations (SE_{χ^2}). Note that the values of SE_{χ^2} is consistently smaller than the ones of SE_{BRR}. The ratio of these two values, $\frac{SE_{BRR}}{SE_{\chi^2}}$, can be considered as a deft (Kish, 1995), and is the range of this ratio is between 0.996 and 1.601. We did not observe any major change in the values of proportions with respect to the quarters. When we formally tested the quarterly effect on the leading digit 2, the value of test statistic is 5.20 with the critical value 9.01 at $\alpha = 0.05$.

In the following sections, we will examine the test statistics for the proportions collected by each interviewer.

4. Options for Quadratic-Form Test Statistics

In our example data of the U.S. Consumer Expenditure Survey (CES), individual field representatives are closely tied to a sample design. We will consider a household as a cluster within an interviewer. Define $x_{icj} = (x_{1icj}, \cdots, x_{9icj})'$ a 9×1 vector such that $x_{dicj} = 1$ if the leading digit is d, and 0 otherwise, for a given interviewer i, an interview c and an expenditure item j. Let J_{ic} be the number of non-zero reports obtained for an interview c by an interviewer i, and $x_{ic} = \sum_{j=1}^{J_{ic}} x_{icj} = (x_{1ic}, \cdots, x_{dic}, \cdots, x_{9ic})'$ be a 9×1 vector where x_{dic} is the number of non-zero reports with a leading digit d for an interview c by an interviewer i. Let $\hat{p}_i$ be a mean vector of leading digits for interviewer i.

$$\begin{aligned} \hat{p}_i &= (\hat{p}_{1i}, \cdots, \hat{p}_{9i})' \\ &= \frac{1}{J_i} \sum_{c=1}^{n_i} \sum_{j=1}^{J_{ic}} x_{icj} \\ &= \frac{1}{J_i} \sum_{c=1}^{n_i} x_{ic} \end{aligned}$$

where n_i is the number of interviews conducted by an interviewer i, and $J_i = \sum_{c=1}^{n_i} J_{ic}$ is the total number of non-zero reports obtained by an interviewer i.

The general test statistic is:

$$T_i = (\hat{p}_i - p_0)' V_i^{-1} (\hat{p}_i - p_0) \ .$$

where p_0 is a 9×1 reference proportion vector. p_0 can be a vector of proportions predicted by Bedford's Law, or a vector of empirical CES proportions.

The following are possible estimators of V_i.

4.1 Design-based variance estimator

A design-based variance estimator of $\hat{p}_i$, $\hat{V}(\hat{p}_i)$, is

$$\begin{aligned} \hat{V}(\hat{p}_i) &= \hat{V}\left(\frac{1}{J_i}\sum_{c=1}^{n_i} x_{ic}\right) \\ &= \hat{V}\left\{\frac{1}{n_i}\sum_{c=1}^{n_i}\left(\frac{n_i}{J_i}x_{ic}\right)\right\} \\ &= \hat{V}\left(\frac{1}{n_i}\sum_{c=1}^{n_i} y_{ic}\right) \\ &= \frac{1}{n_i}\frac{1}{(n_i-1)}\sum_{c=1}^{n_i}\left\{(y_{ic}-\hat{p}_i)(y_{ic}-\hat{p}_i)'\right\} \end{aligned}$$

where $y_{ic} = \frac{n_i}{J_i} x_{ic}$. A test statistic is:

$$T_{i\,1} = \left(\hat{p}_{i(8)} - p_{0(8)}\right)' \left\{\hat{V}(\hat{p}_{i(8)})\right\}^{-1} \left(\hat{p}_{i(8)} - p_{0(8)}\right)$$

where $\hat{p}_{i(8)}$ is the first eight elements of $\hat{p}_i$. Since $\hat{V}(\hat{p}_i)_{9\times 9}$ will be always singular because of the constraint on $\hat{p}_i$ in our case, we use $\hat{p}_{i(8)}$ to compute $T_{i\,1}$.

$\hat{V}(\hat{p}_{i(8)})$ can also be singular when n_i is small. A range of n_i, the number of interviews covered by an interviewer i, is 1 to 185 in our example, a mean of 25, a first quartile of 2, a median of 9, and a third quartile of 37. Among total number of 1235 interviewers, there are 234 interviewers who conducted

Table 1: Quarterly Ratio Estimates and Associated Standard Errors

Quarter 2		Quarter 3		Quarter 4		Quarter 5	
$\hat{R}_{Q2}$	SE_{BRR} SE_{χ^2}	$\hat{R}_{Q3}$	SE_{BRR} SE_{χ^2}	$\hat{R}_{Q4}$	SE_{BRR} SE_{χ^2}	$\hat{R}_{Q5}$	SE_{BRR} SE_{χ^2}
0.3059	**0.0012** *0.0010*	0.3053	**0.0014** *0.0011*	0.3010	**0.0013** *0.0011*	0.3018	**0.0015** *0.0011*
0.1918	**0.0011** *0.0009*	0.1931	**0.0014** *0.0009*	0.1953	**0.0012** *0.0010*	0.1941	**0.0013** *0.0009*
0.1208	**0.0010** *0.0007*	0.1231	**0.0008** *0.0008*	0.1246	**0.0009** *0.0008*	0.1245	**0.0010** *0.0008*
0.0897	**0.0008** *0.0006*	0.0894	**0.0007** *0.0007*	0.0911	**0.0010** *0.0007*	0.0908	**0.0008** *0.0007*
0.1038	**0.0007** *0.0007*	0.1051	**0.0008** *0.0007*	0.1050	**0.0011** *0.0007*	0.1051	**0.0011** *0.0007*
0.0679	**0.0006** *0.0006*	0.0696	**0.0007** *0.0006*	0.0693	**0.0009** *0.0006*	0.0671	**0.0006** *0.0006*
0.0485	**0.0007** *0.0005*	0.0477	**0.0006** *0.0005*	0.0463	**0.0005** *0.0005*	0.0480	**0.0006** *0.0005*
0.0451	**0.0006** *0.0005*	0.0431	**0.0006** *0.0005*	0.0434	**0.0007** *0.0005*	0.0439	**0.0005** *0.0005*
0.0265	**0.0005** *0.0004*	0.0235	**0.0006** *0.0004*	0.0241	**0.0006** *0.0004*	0.0247	**0.0004** *0.0004*

only one interview. This indicates that we have quite a few interviewers whose n_i is very small. We found that $\hat{V}(\hat{p}_{i(8)})$ is too unstable.

4.2 Variance estimator for ratio estimator

Consider the number of non-zero reports obtained by an interviewer i, J_i, as a random variable. The ratio estimator is

$$\begin{aligned}\hat{R}_i &= \left(\sum_{c=1}^{n_i}\sum_{j=1}^{J_{ic}} x_{icj}\right) / \left(\sum_{c=1}^{n_i}\sum_{d=1}^{9} x_{dic}\right) \\ &= \left(\sum_{c=1}^{n_i} x_{ic}\right) / J_i\end{aligned}$$

.

Recall that J_i is the number of non-zero reports obtained by an interviewer i . Then the variance estimator for the ratio estimator is:

$$\begin{aligned}&\hat{V}(\hat{R}_i) \\ &= \{n_i(n_i-1)\}^{-1} \times \\ &\quad \left[(J_i/n_i)^{-2}\sum_{c=1}^{n_i}\left\{(x_{ic}-\hat{R}_i\,J_{ic})(x_{ic}-\hat{R}_i\,J_{ic})'\right\}\right].\end{aligned}$$

A test statistic is:

$$T_{i\,2} = \left(\hat{R}_i - p_{0(8)}\right)' \left\{\hat{V}(\hat{R}_{i(8)})\right\}^{-1} \left(\hat{R}_i - p_{0(8)}\right) .$$

Note that $\hat{V}(\hat{R}_{i(8)})$ shares the same problems that $\hat{V}(\hat{p}_{i(8)})$ has with small values of n_i.

5. Misspecification Effect Matrices

A misspecification effect is one measure of the effect of a complex sample design. It measures the bias of a variance estimator which is computed under misspecified design or modeling assumptions. The eigen structure of a misspecification effect matrix allows comparison of some sensitivity properties of competing quadratic-form test methods. Define Δ_i a 8×8 multivariate design effect matrix for an interviewer, i. Then

$$\Delta_i = \Sigma_i^{-\frac{1}{2}} V_i \Sigma_i^{-\frac{1}{2}}$$

where $\Sigma_i^{-\frac{1}{2}}$ is an inverse of a square root symmetric matrix of Σ_i

$$\Sigma_i = \frac{1}{J_i}\left\{diag(p_{0(8)}) - (p_{0(8)})(p_{0(8)})'\right\} ,$$

and V_i is a corresponding covariance matrix of our choice. The misspecification matrix for $\hat{V}(\hat{p}_i)$ is:

$$\Delta_{i\,1} = \Sigma_i^{-\frac{1}{2}} \hat{V}(\hat{p}_i) \Sigma_i^{-\frac{1}{2}}$$

The misspecification effect matrix for $\hat{V}(\hat{R}_i)$ is:

$$\Delta_{i\,2} = \Sigma_i^{-\frac{1}{2}} \hat{V}(\hat{R}_i) \Sigma_i^{-\frac{1}{2}}$$

The eigenvalues of the misspecification effect matrix are called generalized design effects. Rao & Scott used the mean and coefficient of variation of these eigenvalues to develop modifications of certain chi-square goodness-of-fit test statistics.

5.1 Modified Chi-Squared Tests

Consider a naive χ^2 statistic to test how a distribution of the leading digit proportions from non-zero reports obtained by an interviewer differs from the distribution of the reference proportions.

$$\chi_i^2 = J_i \sum_{d=1}^{D} \left\{ (\hat{p}_{di} - p_{d0})^2 / p_{d0} \right\}$$

where d is $1, \cdots, D = 9$.

Define λ_i the eigenvalues and $\bar{\lambda}$ the average eigenvalue from the estimated misspecification effect matrix. The first order Rao-Scott adjusted test is:

$$\begin{aligned} \chi_{iM1}^2 &= \chi_i^2 \,/\, \bar{\lambda} \\ &= (D-1)\, \chi_i^2 \,/ \left(\sum_{d=1}^{D} \{(1-\hat{p}_{di})\ deff_{di}\} \right) \end{aligned}$$

where $deff_{di} = \{\hat{p}_{di}(1-\hat{p}_{di})/J_i\}^{-1} \hat{V}(\hat{p}_{di})$, and $\hat{V}(\hat{p}_{di})$ is the d-th diagonal element of $\hat{V}(\hat{p}_i)$. The performance of χ_{iM1}^2 depends on the eigenstructure of the misspecification effect matrix. When $\hat{\lambda}_i$ are approximately equal to their average, $\bar{\lambda}$, χ_{iM1}^2 is approximately distributed as a χ^2 random variable with $D-1$ degrees of freedom (Lee & Eltinge).

The second order approximation which is the better approximation can be used if there is more information available on the $\hat{\lambda}_i$'s (Rao & Scott, 1981).

$$\chi_{iM2}^2 = \chi_{iM1}^2 \,/\, (1+\hat{a}^2)$$

where $\hat{a}^2 = \left\{(D-1)\bar{\lambda}^2\right\}^{-1} \sum_{d=1}^{(D-1)} (\hat{\lambda}_{di} - \bar{\lambda})^2$.

For example, we have an interviewer who has conducted 45 interviews with total of 725 non-zero reports. The following are the computed statistics for the interviewer: $T_{i\,2} = 37.01$,$\chi_i^2 = 26.47$, $\chi_{iM1}^2 = 14.74$ with $\bar{\lambda} = 1.80$, $\chi_{iM2}^2 = 14.45$ with $\hat{a}^2 = 0.02$.

5.2 Other Related Tests

The examination of the expenditure data collected by the CEQ survey in the year 2000 shows that the leading digits of those expenditures follow Benford's Law quite closely (Swanson, Cho, and Eltinge, 2003). However, digits such as 5's or 9's do not follow Benford's Law as closely as others. When comparing the leading digits collected by an unusual field representative with overall CEQ's leading digits distribution, the leading digit of 5's or 9's of an unusual field representative do not follow the overall CEQ's leading digits distribution (Swanson, Cho, and Eltinge, 2003). We may test deviations from null proportion only on those leading digits using an univariate t-test or Bonferroni test statistics:

$$t = \frac{\hat{p}_{di} - p_{d0}}{\hat{V}(\hat{R}_{di})}$$

where $\hat{V}(\hat{R}_{di})$ is the diagonal element of $\hat{V}(\hat{R}_i)$ for the leading digit d.

Korn & Graubard demonstrated that Wald statistics behave poorly when simultaneously testing a large number of regression coefficients. They suggested that for some applications, the use of Bonferroni inequality on the individually tested regression coefficients can have greater power than the Wald procedure. The Bonferroni procedure is preferable to the Wald procedure when the number of variables is large compared with the number of sampled PSU's (Korn & Graubard).

6. Optimal Decision

Define π_i to be the probability that a given item recorded by interviewer i is erroneous. For simplicity, we will assume that a given interview records data that can be classified entirely as erroneous or not erroneous. In an interview that is not erroneous, all expenditure amounts are reported accurately. This ideal case would arise, for example, if the interview subject had maintained comprehensive and accurate records of all relevant expenditures by the consumer unit, and based all interview responses on these records. At the other extreme, we will define an interview to be erroneous if either: (a) the interviewer has fabricated all responses, i.e., the interview was "curbstoned" or (b) the interviewer accurately recorded all responses given by the interview subject, but the interview subject's responses were based on uninformed guesses. Many practical cases would fall between these two extremes. For example, an interviewer might fabricate responses in one section for which the interview subject refused to provide responses; or an interview subject might provide uninformed guesses about expenditures within groups for which another family member made the purchase decisions. However, these more complex cases are beyond the scope of the current work. In general, an ideal procedure would be the one that has a relatively high possibility of identifying interviewers who have digit-reporting patterns that deviate substantially from the overall pattern, while

also having a relatively low Type I error rate. For the Consumer Expenditure Survey data, this issue is complicated by the fact that the number of interviews covered by a given interviewer varies a great deal as we mentioned in the Section 4.1.

A is the cost of conducting the one reinterview, B is the cost of including an erroneous report, R is the number of reinterviews, and U is the number of undetected erroneous reports. Let C_i be the optimal cut off point which minimizes $E(cost) = A\,E(R) + B\,E(U)$. If $\hat{p}_i \le C_i$, then we reinterview all of the households which had previously been interviewed by interviewer i. Note that C_i depends on the number of households (n_i), the number of total items (J_i) reported by interviewer i, and the design effects.

$$
\begin{aligned}
&E(R, \textit{interviewer } i) \\
&= n_i\, P\left(\hat{p}_i \le C_i\right) \\
&= n_i P\left(\hat{p}_i \le C_i | Err\right)\ \pi_i \\
&\quad + n_i P(\hat{p}_i \le C_i | NotErr)\ (1-\pi_i) \\
&= n_i \pi_i\ P\left(\frac{\hat{p}_i - E_{GE}(\hat{p}_i)}{\sqrt{V_{GE}(\hat{p}_i)}} \le \frac{C_i - E_{GE}(\hat{p}_i)}{\sqrt{V_{GE}(\hat{p}_i)}} | Err\right) \\
&\quad + n_i(1-\pi_i) \\
&\quad \times P\left(\frac{\hat{p}_i - E_{GT}(\hat{p}_i)}{\sqrt{V_{GT}(\hat{p}_i)}} \le \frac{C_i - E_{GT}(\hat{p}_i)}{\sqrt{V_{GT}(\hat{p}_i)}} | Non\right) \\
&\approx n_i\ \pi_i\ \Phi\left(\frac{C_i - E_{GE}(\hat{p}_i)}{\sqrt{V_{GE}(\hat{p}_i)}}\right) \\
&\quad + n_i\ (1-\pi_i)\ \Phi\left(\frac{C_i - E_{GT}(\hat{p}_i)}{\sqrt{V_{GT}(\hat{p}_i)}}\right) \\
&= n_i\, \pi_i\, \gamma_E(C_i) + n_i\, (1-\pi_i)\, \gamma_T(C_i)
\end{aligned}
$$

where GE and GT are the distributions of the erroneous reports and the truthful reports respectively, $V_{GE}(\hat{p}_i) = \frac{p_{i,GE}(1-p_{i,GE})}{J_i}\, deff$, $p_{i,GE} = E_{GE}(\hat{p}_i)$, $V_{GT}(\hat{p}_i) = \frac{p_{i,GT}(1-p_{i,GT})}{J_i}\, deff$, $p_{i,GT} = E_{GT}(\hat{p}_i)$, $\gamma_E(C_i) = \Phi\left(\frac{C_i - E_{GE}(\hat{p}_i)}{\sqrt{V_{GE}(\hat{p}_i)}}\right)$, $\gamma_T(C_i) = \Phi\left(\frac{C_i - E_{GT}(\hat{p}_i)}{\sqrt{V_{GT}(\hat{p}_i)}}\right)$.

$$
\begin{aligned}
&E(U, \textit{interviewer } i) \\
&= n_i\, P(\textit{undetected erroneous report} \mid Err) \\
&= n_i \pi_i\, P\left(\hat{p}_i \ge C_i |\ Err\right) \\
&= n_i \pi_i P\left\{\frac{\hat{p}_i - E_{GE}(\hat{p}_i)}{\sqrt{V_{GE}(\hat{p}_i)}} \ge \frac{C_i - E_{GE}(p_i)}{\sqrt{V_{GE}(\hat{p}_i)}} | Err\right\} \\
&\approx n_i \pi_i \left\{1 - \Phi\left(\frac{C_i - E_{GE}(p_i)}{\sqrt{V_{GE}(\hat{p}_i)}}\right)\right\} \\
&= n_i \pi_i \left\{1 - \gamma_E(C_i)\right\}
\end{aligned}
$$

Option 2: Find C_i to minimize $E(R)$ subject to $E(U) \le R_i = n_i R_0$. From the definitions of E(R) and E(U), note that $E(R)$ is monotone increasing with respect to C_i, and $E(U)$ is monotone decreasing with respect to C_i. Hence the minimum of $E(R)$ can be obtained when $E(U) = n_i R_0$, i.e. when $C_i = \gamma_E^{-1}\left(1 - R_0/\pi_i\right)$.

Option 3: Find C_i to minimize $E(U)$ subject to $E(R) \le U_i = n_i U_0$. From the definitions of E(U) and E(R), note that $E(U)$ is monotone decreasing with respect to C_i, and $E(R)$ is monotone increasing with respect to C_i.

Define $\gamma(C_i) = n_i \pi_i\, \gamma_E(C_i) + n_i(1-\pi_i)\gamma_T(C_i)$ Then the minimum of $E(U)$ can be obtained when $E(R) = n_i U_0$, i.e. when $C_i = \gamma^{-1}\left(n_i U_0\right)$.

7. Discussion

We tested in the section 3 whether the quarter in which the data was collected matters. The same test can be applied for the level of experience of interviewers, or other relevant factors.

We considered the univariate case in the Section 6, but this can be extended to the more complex cases such as the multivariate one. One can also apply the mixture modeling method for parameters of GE and GT. A sensitivity analysis for the optimal decision would be of interest.

8. References

Biemer, P.P. and Stokes, S. L.(1989). The optimal design of quality control samples to detect interviewer cheating. *Journal of Official Statistics* 5, 23-39.

Bureau Of Labor Statistics (1997). *BLS Handbook of Methods*. U.S. Department of Labor.

Hill, Theodore (1995). A Statistical Derivation of the Significant-Digit Law. *Statistical Science* 10, 354-363.

Korn, E. and Graubard, B.(1990). Simultaneous testing of regression coefficients with complex survey data: use of Bonferroni t statistics. *The American Statistician* 44, 270-276.

Lee, S. and Eltinge, J.( ). Exploratory analysis of estimated design effect matrices computed from complex survey data.

McLachlan, G.J. and Basford, K.E.(1988). *Mixture Models.* Marcel Dekker, INC.

Rao, J.N.K. and Scott, A.J.(1981). The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence of two-way table. *J. Amer. Statist. Assoc.* 76, 221-230.

Rao, J.N.K. and Scott, A.J.(1984). On chi-squared tests for multi-way contingency tables with cell proportions estimated from survey data. *Ann. Statist.* 12, 46-60.

Swanson, D., Cho, M., Eltinge, J(2003). Detecting possibly fraudulent or error-prone survey data using Benford's Law. *Proceedings of the Section on Survey Research Methods, American Statistical Association* to appear.

Turner, C., Gribble, J., Al-Tayyib, A., and Chromy, J. (2002). Falsification in epidemiologic surveys: detection and remediation. *Technical Paper on Health and Behavior Measurement.* 53. Washington DC: Research Triangle Institute.