Assessing Nonresponse Bias and Measurement Error Using Statistical Matching

John Dixon, Bureau of Labor Statistics, 2 Massachusetts Avenue, NE Washington, DC 20212
Dixon.John@bls.gov

Key Words: Nonresponse bias, Measurement error, Statistical matching

The estimation of nonresponse bias and measurement error share the problem of usually not having a criterion to assess the quality of the estimate. Nonresponse bias analysis often uses responders within the survey sample who are in some way similar to nonresponders to estimate the potential bias. This depends on the variables within the survey being related to both the likelihood of responding and also being related to the measure being estimated. An alternative method uses record linkage to get data about nonresponders from another source, often administrative data. The record linkage studies are limited by what data might be available and the quality of the linkage. Measurement error studies have similar problems with estimating error. Some methods include using the internal consistency of responses, re-interviews, or record linkage to provide a measure of the quality of response. Statistical matching uses surveys from different samples which have some similarities to compare estimates. The matching can be based on demographics and other sample characteristics which are thought to be related to the survey estimates. Propensity scores are one type of variable often used in matching.

1. Introduction

Three surveys were used to study the potential for using statistical matching to estimate nonresponse and measurement bias. The Consumer Expenditure Quarterly Interview Survey (CEQ) is a household survey which provides part of the “market basket” of consumer expenditures, which are the basis of the CPI as well as other indices. Sampled housing units in the Quarterly are interviewed for 5 consecutive quarters. These interviews are referred to as “time-in-sample” (TIS) 1 to 5. The National Health Interview Survey (NHIS) is a household survey to measure the population’s health. It is comprised of four components; household family, personal and child. It is administered to a household only once. The Current Population Survey (CPS) is a household survey which provides estimates of labor force participation, as well as many other measures. Households are interviewed for 8 months, in a 4-8-4 pattern (4 months of interviews, repeated for the same months 8 months later).

The Contact History Instrument (CHI) was added to the CEQ and the National Health Interview Survey (NHIS) in 2005 to collect detailed contact history data (Bates, 2004). The interviewer records times and outcomes of attempted contacts, problems or concerns reported by reluctant households, and strategies used to gain contact or overcome reluctance. This provides a very rich source for studying the interview process, which is only lightly used in this study. It was added to the CPS in 2009.

Dixon (2006) found that estimates of nonresponse bias weren't impacted much by the addition of call history variables. Those interviews which required a larger number of contacts where the interviewer changed mode of contact had lower expenditures (-39.4). This effect was partially offset by those interviews that required more contacts but where respondents who reported no concerns had higher expenditures (27.0). Those interviewers who reported "no strategy" for attempted contact ended up with lower expenditures (-66.2) and those who changed modes during the contact process ended up with higher expenditures.

2. Statistical Matching

To study nonresponse and measurement error, statistical matching compares responders or aggregate data from different sources to model differences in key survey estimates. The sources are typically other surveys or administrative records. The main problem with statistical matching studies is differences in sample composition and in question wording.

In this study I will use different sources, the CEQ, CPS, and the NHIS, to attempt to measure the differences in employment status. In statistical matching it is more common to study variables which aren't in both surveys. Samples will be matched on characteristics including sampling frame variables, similarities in process data, and responses to key survey questions. The estimation from the statistical matching will differ in response rate and in differences in the questions. These differences will be modeled to separate the nonresponse bias estimates from the measurement differences. One of the features of statistical matching is to include the variability in matching in the variance estimates. This can be done through multiple imputation, jackknife, or bootstrap estimation.

3. Data Sources

The CHI and CEQ data from 2006 through 2008 will be combined to provide the paradata for this study. 97,317 households were used for the analysis. Interviews which didn't provide CHI data were excluded. The National Health Interview Survey is an in-person cross-sectional survey. It also uses the CHI to collect call history information very similar to the CE.

The NHIS 2008 public use data file provided 44,358 cases with CHI data. Some data was linked from the family dataset. The Current Population Survey is currently implementing CHI; data from 2009 is used for this study (50,000 cases).

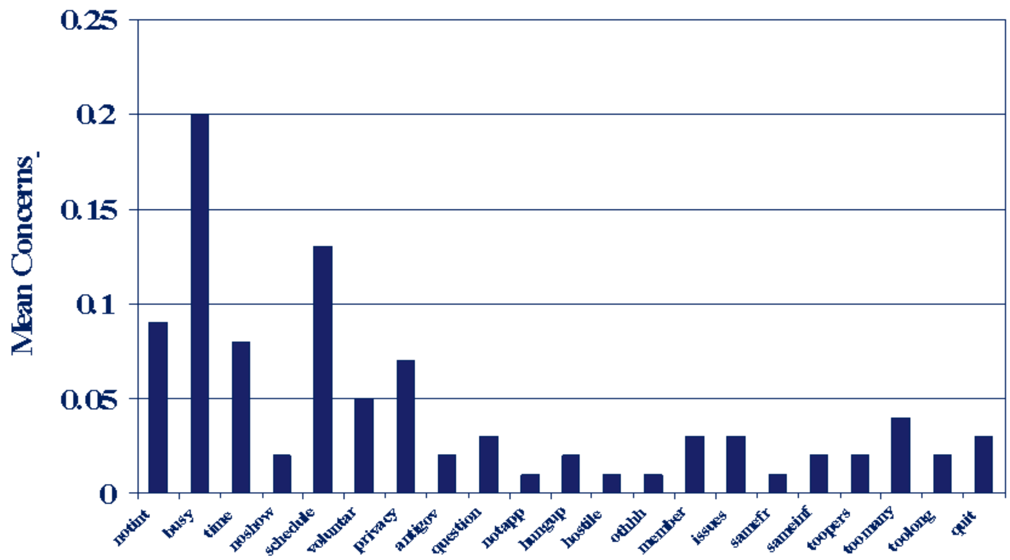
4. Methods

Logistic models (SAS version 9.1) were used to predict nonresponse in the CEQ using variables from CHI. A factor analysis (using MPlus) was used to examine inter-relationships among the respondents concerns expressed in CHI for the CEQ, CPS, and the NHIS separately. Demographic characteristics (e.g.: family structure) and similarities on the CHI were used to create a matching variable in the surveys. Estimates of nonresponse bias and measurement differences were compared using differences in those estimates.

5. Results

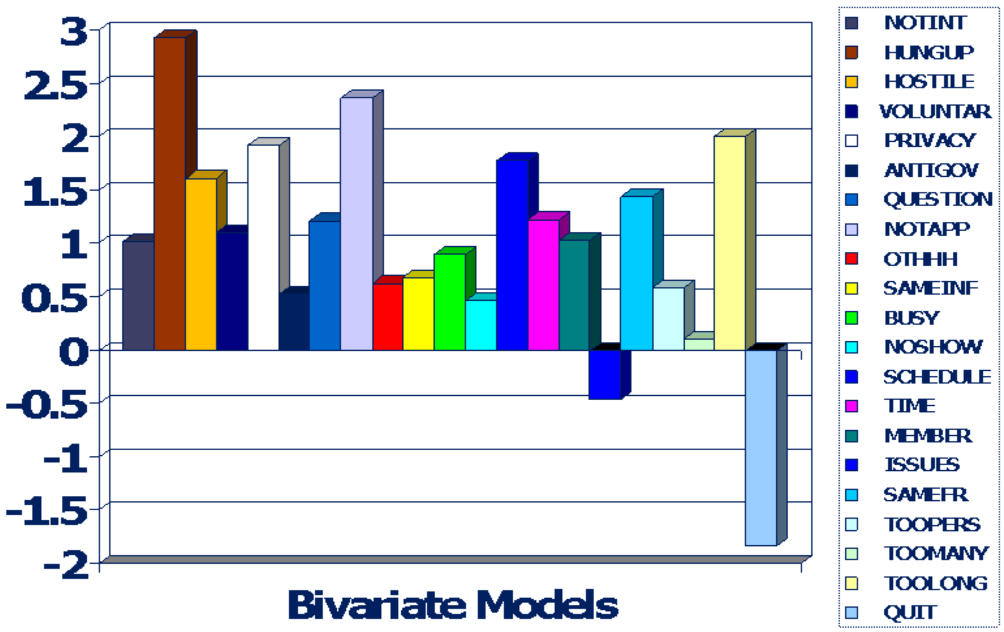
The most common concern expressed by respondents was “busy” (Figure 1), followed by “schedule difficulties”, and “not interested”, which were also most predictive of a refusal outcome. Other notable concerns were “time the interview takes” and “privacy concerns.”

Figure 1: Mean Rates of CHI concerns.



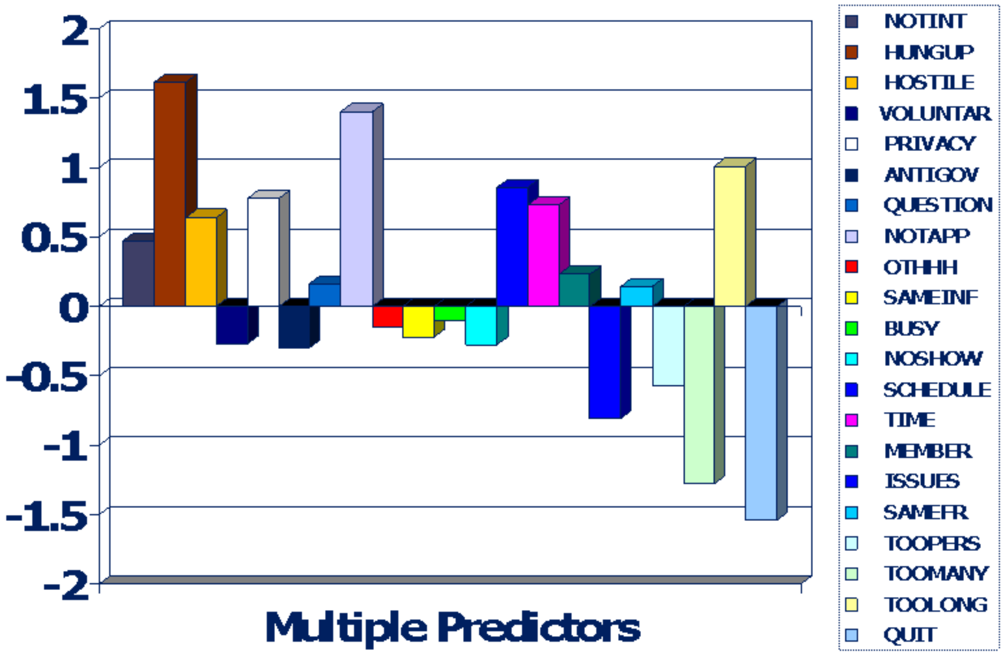
Logistic models for the relationships between concerns and refusal (Figure 2), shows some strong effects. The univariate logistic models showed positive relationships between most of the concerns and refusal during some of the interviews. “Family issues” (issues, which was not significantly related to refusal) and “intends to quit” are the two related to not refusing.

Figure 2: Predicting nonresponse with bivariate logistic models.



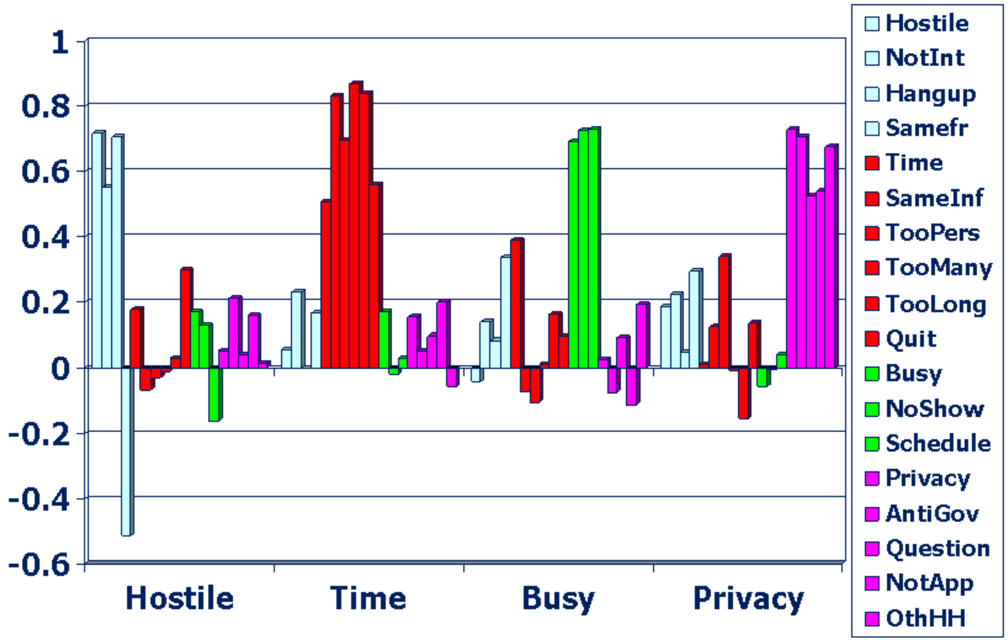
The multivariate model (Figure 3) showed some coefficients which reversed direction or became non-significant when adjusting for the other variables. The univariate estimates could be interpreted as the relationship between respondent concerns and refusals, while the multivariate estimates could be interpreted as the unique relationship of those concerns beyond the other concerns. The multivariate model gives a more complete picture of the complex relationship between concerns and refusal. The most common concern, "busy", showed a strong relationship with refusals, but didn't contribute anything beyond the other variables. "Not interested", another frequent concern had a strong relationship with refusal even after adjusting for the other variables. Counter to expectations, the concern "planning to quit" showed a strong negative relationship to refusal, even after controlling for other concerns. This result might have been related to increased efforts by the interviewer to persuade respondents planning to quit to stay with the survey.

Figure 3: Predicting nonresponse with a multivariate logistic model.



To examine the inter-relationships between concerns a factor analysis was used (Figure 4).

Figure 4: Factor pattern for CHI concerns.



Four factors were identified based on concerns expressed in the contact history instrument. The "Hostile" factor included hostile behavior, hangups, "not interested", and in a negative relationship, wanting the same FR from the previous interview. The "Privacy" factor included concerns about privacy, expression of anti-government sentiment, not understanding the survey, not thinking the survey was applicable to them, and other household members advising the respondent not to participate. Although the "Time" and "Busy" factors may seem similar, the "Time" factor was negatively related to refusal and seems more related to being overwhelmed or over burdened while the “Busy” one seems to relate to the respondent not being able to find the time. The prediction of refusal based on the 4 scales gave a pseudo R-square of .19 compared to a model using all the items which went into the scales pseudo r-square of .20.

By comparing those most like refusers on their contact history characteristics we can make inferences about potential nonresponse bias. Other studies support this approach, noting a relationship between CHI data and refusal on both the NHIS and the CEQ (Bates 2004, Bates et. al. 2008, Maitland et. al. 2009). This study extended that research to examine differences in the interview experience and subsequent refusals based on respondent concerns in the CPS, NHIS and the CEQ.

The studies by Bates (2004), and Henley and Bates (2006) found that in the NHIS, the *number* of concerns was a more important predictor of refusal than *specific* concerns. They found privacy concerns, the voluntary nature of the survey, “not interested”, and “Survey takes too long” to be the primary concerns for refusers. This study found a similar overall pattern, but added “schedule difficulties” to the list. Some unexpected effects were found with negative relationships to refusal for respondents which had “family issues” (issues) or “intends to quit survey” (quit). The family issues were thought to make the difficult process of reporting expenditures more difficult, but although non-significant, that didn’t seem to be a problem for respondents.

The present study found a similar factor structure as the Maitland et. al. 2009 study. The similar factor structure allows the creation of rough indices of the factors described. The matching was based on sample demographics (family size and age), and response patterns on the CHI (using scales based on the factors common the two surveys). Random matches between respondents to the three surveys were selected, and treated as an imputation in a multiple imputation framework.

Figure 5 shows the estimates of the employment rate for the different surveys. The first line is the estimate from each survey for someone in a household being employed (the data needed to be collapsed for the CPS, since employment is only reported at the household level for NHIS and CE)
The second line shows the estimates of employment for CE. The first is based on CE data adjusting for the other surveys (but the CE data dominate, with only a slight increase in standard error due to the matching). The estimate based on NHIS and CPS (not using the CE estimates of employment) show poorer estimates and larger standard errors.
This patter is similar for the third and fourth lines.

Figure 5: Employment Estimate Differences Between Surveys

	CE	**NHIS**	**CPS**
	.7453 (.0016)	.7447 (.0026)	.7394 (.00224)
CE	.7453 (.001602)	.8616 (.0044)	.8055 (.0073)
NHIS	.7008 (.0078)	.7447 (.002607)	.7474 (.0074)
CPS	.6502 (.0136)	.6078 (.0062)	.7394 (.002244)

Figure 6 shows the estimates of potential nonresponse bias. We saw that the CE has higher estimates of household employment than the CPS or NHIS. The first group of bars represents the combined 1st though 4th quintiles of the refusal propensity for each of the 3 surveys. Compare that to the next group, which represents those most like the refusers as the 5th quintile. The CPS and NHIS have such low nonresponse rates that the effect on the final estimates would be very low, with refusers being more likely to be employed. This is the opposite of the CE.

The next two groups of bars represent the 1st through 4th quintiles of the noncontact propensity to compare with the 5th quintile which represents the potential for noncontact bias. The CE and CPS have only a small amount of noncontact, so the impact on estimates would be very small. The noncontact bias for the CE is in the opposite direction from the refusal bias, so the overall bias would be reduced.

Figure 6: Potential Nonresponse Bias

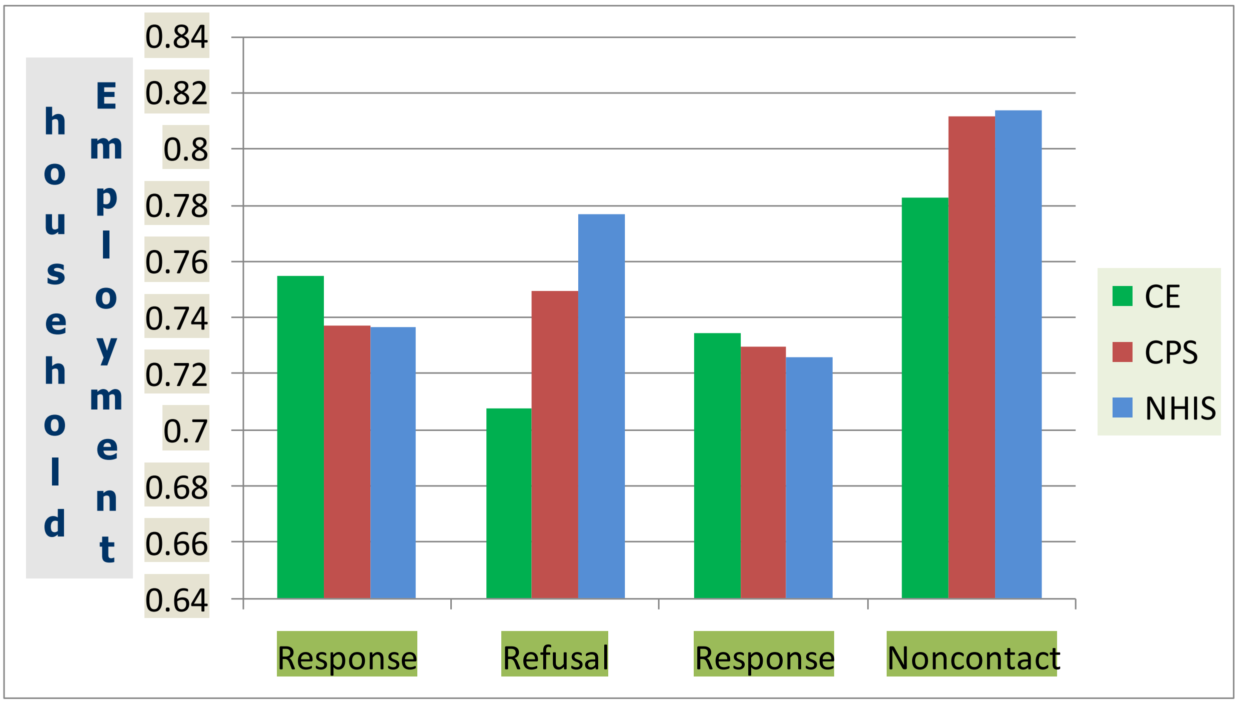
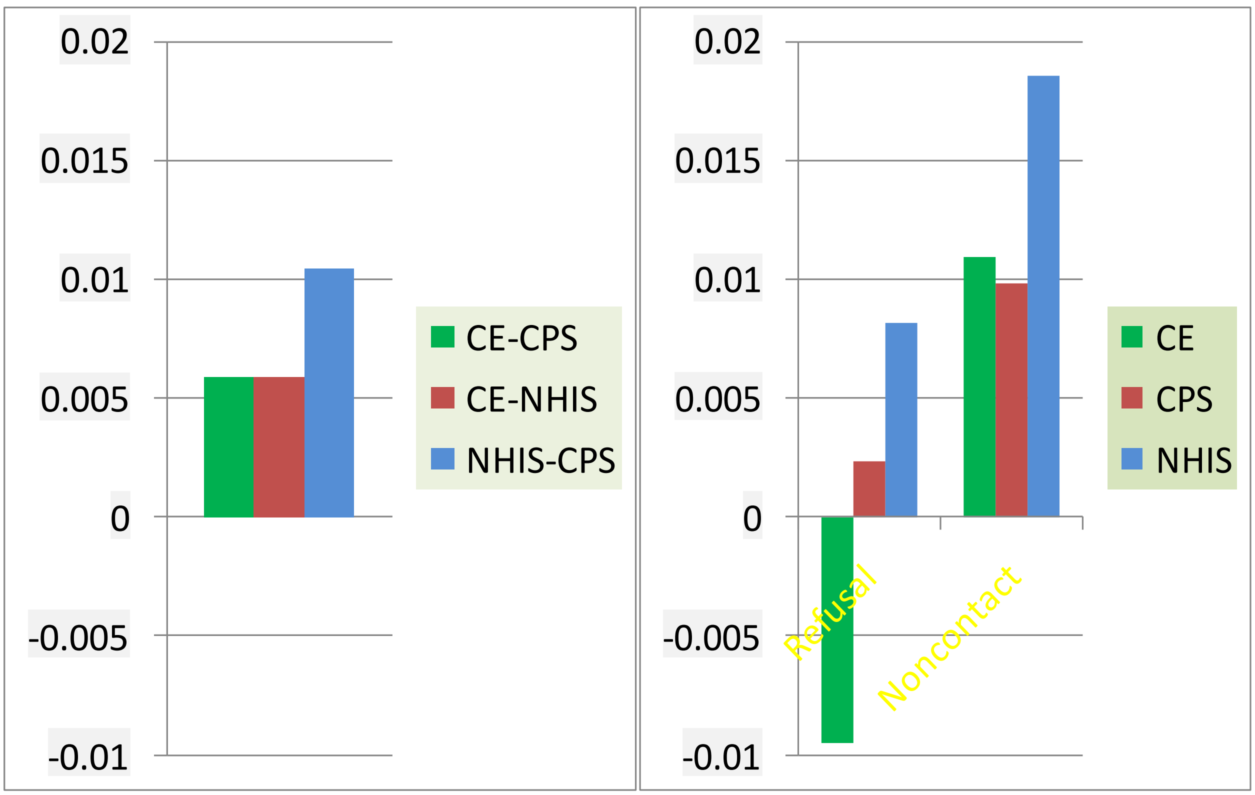


Figure 7 shows estimates of measurement differences and estimates of nonresponse bias. The left chart shows the differences in estimates between the surveys (used as an indicator of measurement error). The CPS and the NHIS differ the most, but the differences are small. The NHIS estimate is between the CE and CPS.

The right chart shows the refusal and noncontact bias estimates based on the differences between the total estimates and the estimates based on the 1st 4 quintiles of the nonresponse propensities. Because noncontact is so small for the CE and CPS, the difference in direction for the CE will reduce the refusal bias slightly, and the noncontact bias for the CPS will increase the overall nonresponse bias only slightly. Also, the CPS bias is estimated for the 1st Month in sample, in other interview months the direction changes, partially cancelling out. The NHIS may overestimate employment more than the other surveys.

Figure 7: Measurement and Nonresponse Bias



Summary

The CHI data showed factor patterns that could describe broad areas of concern. They related well in predicting refusal and noncontact. The nonresponse bias looked similar for the CPS and NHIS in household employment for both refusal and noncontact. The CE noncontact was reversed. The magnitude of measurement differences was similar to the magnitude of the nonresponse bias estimates. The regression approach used for statistical matching was more robust and had better diagnostics than other methods I tried (a binning method not reported here).

Limitations and future research

Additional methods of matching survey data need to be explored, both with the NHIS and CEQ, as well as other (the CPS and MEPS are likely candidates). The matching is probably stronger for these surveys than in other matching analysis, since the key survey estimates, employment estimates, are common to the surveys. Measurement issues with variables not in common would be interesting to explore. The variance increase due to matching was small but non-negligible. Other matching methods may give smaller variances. The regression approach of Scheuren and Moriarity looks particularly interesting.

References

Bates, N., Contact Histories: A Tool for Understanding Attrition in Panel Surveys., A paper presented at the 59th Annual AAPOR Conference, May 11-13, 2004

Bates, N., Dahlhamer, J., and Singer, E., "Privacy Concerns, Too Busy, or Just Not Interested? Exploring the Relationship Between Doorstep Concerns and Nonresponse Bias", Journal of Official Statistics, Vol. 24.4, December, 2008.

Chopova, B., Edgar, J., Gonzalez, J., King, S., McGrath, D., and Tan, L., "Assessing nonresponse bias in the CE Interview Survey: A summary of four studies", presented at the Federal Committee for Statistical Methodology Workshop on How to Do Nonresponse Bias Analyses in Household and Establishment Surveys, June 10, 2009.

Dahlhamer, J., Simile, C., and Taylor, B., "Exploring the Impact of Participant Reluctance on Data Quality in the National Health Interview Survey (NHIS)", paper presented at the Statistics Canada Symposium; Methodological Issues in Measuring Health, 2006.

Dixon, J., "The Effects of Item and Unit Nonresponse on Estimates of Labor Force Participation", Paper presented at the Joint Statistical Meetings, N.Y., N.Y., 2002.

Dixon, J., "Models of Nonresponse Bias, Survey Quality, Paradata, and Household Characteristics in the Consumer Expenditure Survey.", A paper presented at the International Survey Nonresponse Workshop, Omaha, Nebraska, 2006.

Henley, M., and Bates, N., "Using Call Records to Understand Response in Longitudinal Surveys", paper presented at the American Association for Public Opinion Research, 2006.

Mailtand, A., Casas-Cordero, C., and Kreuter, F., "An Evaluation of Nonresponse Bias Using Paradata from a Health Survey", paper presented at the Joint Statistical Meetings, Washington, DC., 2009.

Moriarity, C., and Scheuren, F., "A Note on Rubin's Statistical Matching Using File Concatenation With Adjusted Weights and Multiple Imputations", Journal of Business and Economic Statistcs, V. 21.1, 2003, pp. 65-73.

Rassler, S. Statistical Matching: A Frequentis6t Theory, Practical Applications, and Alternative Bayesian Approaches, Lecture Notes in Statistics, 168, 2002, New York, Springer.

Tucker, C., and Dixon, J., "Predicting Interviewer Nonresponse Rates from Household and Regional Characterstics", Paper presented at AAPOR, 2000.