The Use of Utility Bills as an Information Retrieval Aid in the American Housing Survey[1]

George R. Carter III, Ph.D.
U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

Abstract
Utility costs are an important component of housing costs. The American Housing Survey (AHS) is a key source of housing cost and utility data in the United States. To increase the accuracy of utility data reported, respondents to the AHS are encouraged to use actual bills to aid in information retrieval to the extent they are available. Using data from the 2007 AHS on occupied housing units, this paper presents research on respondents' use of bills in answering questions about electric and gas utilities. Demographic, economic, and regional variations in bill usage are examined. A majority of respondents were found to not have bills available to aid in answering questions about specific months. Respondents who are owners, older, and have higher levels of education are more likely to have complete billing records. Methods for improving data quality through record usage are discussed.

Key Words: record usage, questionnaire design, response process

1. Introduction

Utility costs are an important component of housing costs. The American Housing Survey (AHS) is a key source of housing cost and utility data in the United States. The AHS is collected by the U.S. Census Bureau for the Department of Housing and Urban Development (HUD) and is used by HUD to track changes in housing over time and to develop housing policy. Among the utility data the AHS collects are data on electric, gas, and oil bills, as well as bills for other fuels, trash collection, and water. To increase the accuracy of utility data reported, respondents to the AHS are asked to refer to their utility bills to aid in information retrieval to the extent that bills are available. Using data from the 2007 AHS on occupied housing units, this paper presents research on respondents' use of bills in answering questions about electric and gas utilities. Demographic, economic, and regional variations in bill usage are examined. In what follows, I first review research on advance letters and describe the ways in which electricity and gas costs are collected in the AHS. I then review prior research on the use of respondent records in surveys, discuss theoretical expectations for the research and methods used in the research. Results are then presented, including descriptive statistics and logistic regression models predicting record usage. I conclude with a summary of findings and a discussion of areas for future research.

1.1 AHS Advance Letter
Much of the research on advance letters has examined how advanced letters increase response rates (De Leeuw et al. 2007, Mann 2005, Link and Mokdad 2005, Hembroff, et al. 2005, Goldstein and Jennings 2002), but not on how advance letters prepare the

---

[1] This paper is intended to inform interested parties of ongoing research and encourage discussion of work in progress. The views expressed on methodological, technical and operational issues are those of the author and not necessarily those of the U.S. Census Bureau.

respondent for the interview. In the AHS advance letter, respondents are asked to fill out a worksheet on their utility costs and mortgage costs in preparation for responding to the survey. All housing units in the survey receive the advance letter or are provided the advance letter at the doorstep.

UTILITY COSTS FOR MOST RECENT MONTHS

Utility	January	April	August	December
Electricity	$	$	$	$
Gas	$	$	$	$

This worksheet is not a questionnaire but an aid to help you when our field representative contacts you. **Do not mail this back to the U.S. Census Bureau.** Keep it for your reference. Thank you again for your time and effort in completing this important survey!

Completing the survey will be easier if you will look up the following bills in your checkbook or other records before the interview. Record the amounts for the past year in the spaces provided, skipping any costs that do not apply to you. **Please keep this worksheet** for reference, so you can refer to it when our field representative contacts you.

Figure 1: Utilities Worksheet, American Housing Survey Advance Letter

1.2 How Electricity Costs are Collected in the AHS
Electricity costs are collected in the AHS by first asking the respondent if they use electricity in their home. If they are billed separately for electricity, they are asked how many months each electric bill covers. If the respondent is billed separately for electricity and gave the number of months each bill covers, they are asked: "Do you have any records available showing your costs for electricity (Can you please get them now?)" If they have records available, they are asked: "(From your records, what were the costs for electricity for the months of....) January, April, August, and December." To avoid misreporting, respondents who provide information on one or no monthly bills are asked to report their last bill and the month it was for.

1.3 How Gas Costs are Collected in the AHS
Similar questions are used to collect information on gas costs in the AHS. Respondents are first asked if they use gas in their home. If they say yes, they are asked if the gas is from underground pipes or bottled gas. If the gas is piped, they are asked if they have bills and are asked the monthly amounts of those bills. As with the questions for electricity, respondents who provide information on one or no bills are asked to report their last bill and the month it was for.

2. Related Research

Little research has examined the use of records by respondents in answering survey questions. Research by (Marquis, Moore, and Bogen 1993) explored ways to

increase record use in the Survey of Income and Program Participation (SIPP). The assumption of such studies is that record use improves response accuracy. This assumption, however, has not been tested. To test this with the AHS, one would need to compare administrative data on utility costs to self-reported survey responses. In the research presented in this paper, I examine which respondents have records and use them to answer utility questions.

3. Hypotheses and Methods

3.1 Who Should We Expect to Have Records?
Respondents who are in charge of paying the bills as well as persons with more organized filing habits are expected to be more likely to have bills available and use them in answering utility questions. For this reason, I expect the respondents who are the householder or spouse of householder to be more likely to have bills and respondents who are a child[2] and other relative of the householder to be less likely to have bills. Older respondents, owners, and those with higher levels of education and higher incomes are hypothesized to have more organized filing habits and thus more likely to use have bills. Finally, I consider the role of interviewer presence on reports of record usage. With in-person interviews, the field representative (interviewer) can visibly see if the respondent is using records. This could decrease the number of respondents saying that they have records when they are answering the questions.

3.2 Methods
Data for the research come from occupied housing units on the 2007 American Housing Survey (AHS) Public Use File and Internal Use File. Logistic regression models were estimated predicting the log odds of the respondent having using bills and answering utility questions for all four months for which utility data are collected (January, April, August, and December). Model 1 predicts that the respondent has all four electricity bills. Model 2 predicts that the respondent has all four gas bills. The logistic regression models take the common form:

$$\ln(\text{odds}) = \ln(p_i(1-p_i)) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots\ldots\beta_p X_{pi}$$

Independent variables in the models include respondent demographic and economic characterstics, housing characteristics, and the mode of interview (in-person vs telephone).

4. Results

4.1 Descriptive Statistics, Electricity Bills[3]
Overall, 39.5% of respondents had at least one electricity bill and 35.9% had all four. In no subgroup examined did all respondents have bills for all four months. Respondents

---

[2] Children of the householder must be 16 years old or older to be an eligible respondent for the American Housing Survey.
[3] All differences have been tested at the 90% level.

who were white, female, non-Hispanic, 65+, had more education, and higher incomes were more likely to have bills (Chart 1)[4].
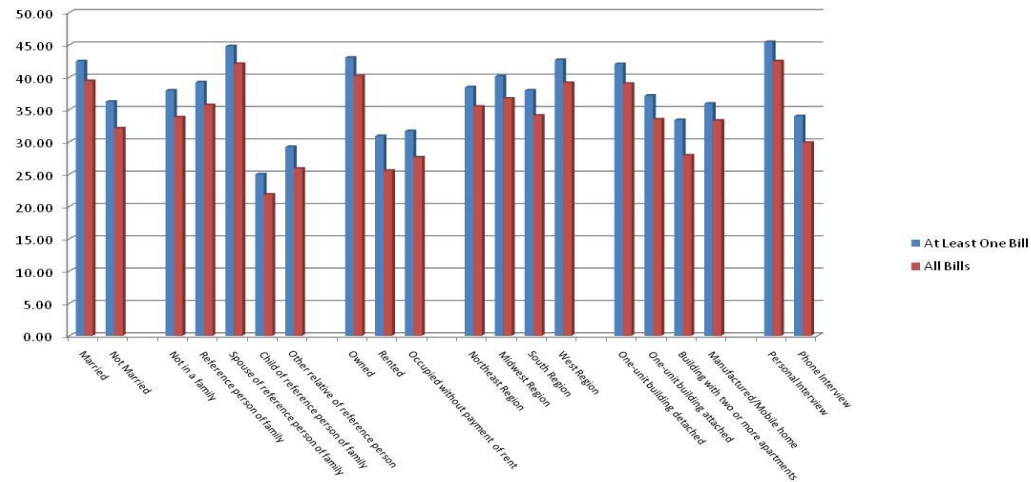
Married respondents, as well as owners, and those in one-unit detached buildings were more likely to have bills. Respondents who were either a child of the householder or an other relative of the householder were less likely to have bills. Personal interviews were more likely to have bills than telephone interviews. Little regional variation was shown (Chart 2).

Chart 1: Percentage of Respondents with Electricity Bills
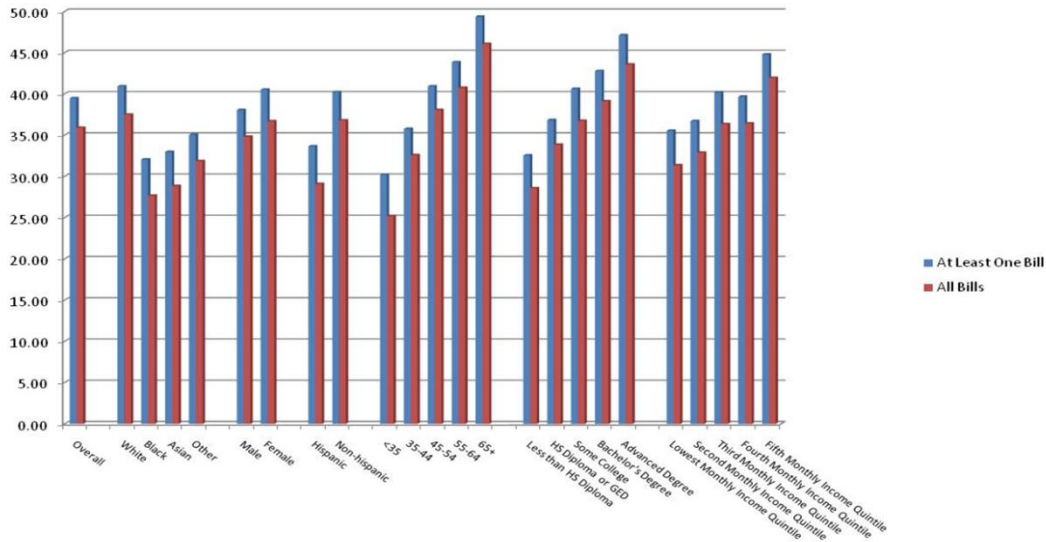
Source: 2007 American Housing Survey

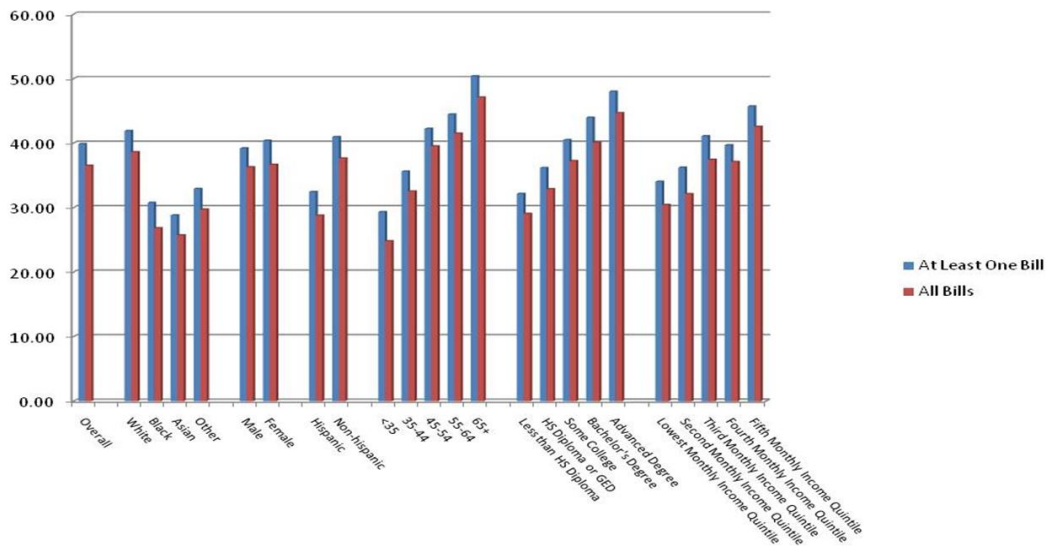**Chart 2**: Percentage of Respondents with Electricity Bills (cont.)

Source: 2007 American Housing Survey

---

[4] Effects reported in the charts are singular, not cumulative.

4.2 Descriptive Statistics, Gas Bills[5]

Overall, almost 40% (39.89%) of respondents had at least one gas bill with 36.51% having all four. Relationships between the demographic and socioeconomic variables mirrored what was found with the electricity bills. In no group examined did all respondents have bills for all four months. Respondents who were white, non-Hispanic, 65+, had more education, and had higher income were more likely to have bills (Chart 3)[6]. Married respondents, as well as owners, and those in one-unit detached buildings were more likely to have bills[7]. Respondents who were either a child of the householder or an other relative of the householder were less likely to have bills. Personal interviews were more likely to have bills than telephone interviews. Respondents in the Northeast were less likely to have gas bills than respondents in other regions (Chart 4).

Chart 3: Percentage of Respondents with Gas Bills
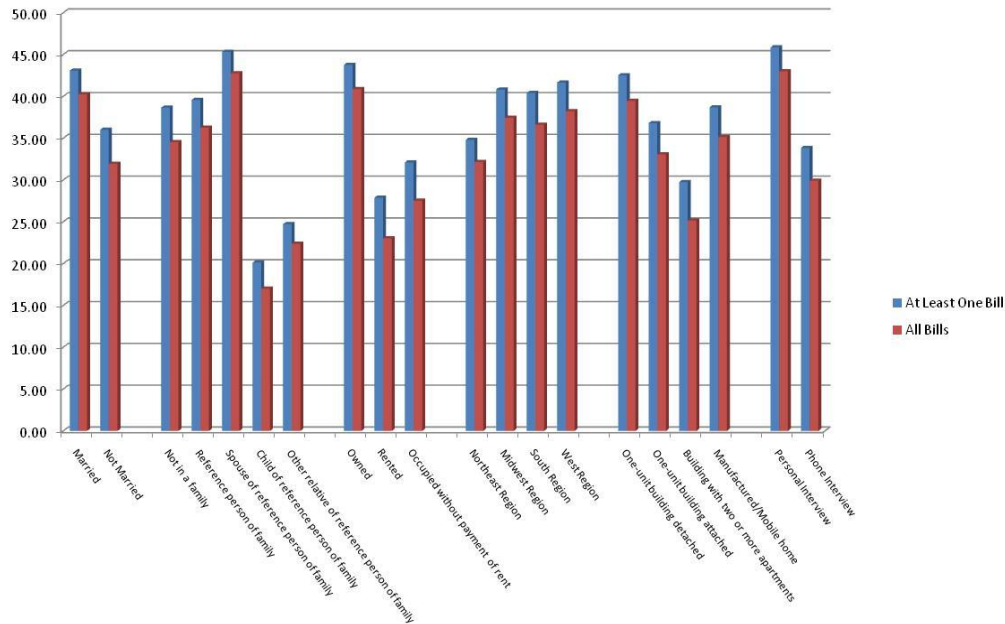

Source: 2007 American Housing Survey

---

[5] All differences have been tested at the 90% level.

[6] Effects reported in the charts are singular, not cumulative.

[7] One unit detached buildings are significantly different from all structure types except for manufactured/mobile homes.

Chart 4: Percentage of Respondents with Gas Bills (cont.)

Source: 2007 American Housing Survey

4.3 Electricity Bill Model[8]

Model 1 predicts the log odds of having all four electricity bills (Table 1). Exp(B)'s are presented in the table. They represent the odds that a predictor variable will have all four bills when compared to the reference group. Age, education, ownership, relation to the householder, and tenure were found to have strong effects. Respondents who are 65 and older are 2.23 times more likely to have all four bills than respondents who are under 35. Respondents who have an advanced degree are 1.62 times more likely to have all four bills compared to respondents with less than a high school education. Respondents with household incomes in the fifth quintile are 1.22 times more likely to have all four bills compared with respondents in the first income quintile. Respondents in owned units are 1.34 times more likely than respondents in rented units to have all four bills. Respondents who were interviewed in-person were 1.41 times more likely to have all four bills compared with respondents who completed the survey over the phone. Respondents who are children of the householder were .75 as likely as those who are the householder to have all four bills.

---

[8] Reference groups in the model include: White only; non-Hispanic; <35 years old; less than high school education; first monthly income quintile; not married; reference person; renter; living in the Northeast; unit in a building with 2+ apartments; and telephone interview.

Table 1: Electricity Bill Model

Ind Vars	Exp(β)	Standard Error	Significance
Sex			
Male	0.93	0.027	**
Race			
Black Only	0.84	0.044	***
Asian Only	0.71	0.076	***
Other	0.93	0.090	
Hispanic Origin			
Hispanic	0.94	0.044	
Age			
35-44	1.19	0.042	***
45-54	1.44	0.042	***
55-64	1.60	0.044	***
65+	2.23	0.044	***
Marital Status			
Married	1.08	0.044	*
Relation to Ref. Person			
Spouse	1.07	0.037	*
Child	0.75	0.112	***
Other Relative	0.74	0.157	*
Not in a family	0.98	0.043	
Tenure			
Owned	1.34	0.043	***
No Rent	0.97	0.110	
Education			
HS	1.19	0.044	***
Some College	1.38	0.045	***
College	1.49	0.050	***
Advanced Degree	1.62	0.056	***
Household Income Quintile			
Second	1.03	0.030	
Third	1.17	0.045	***
Fourth	1.10	0.048	*
Fifth	1.22	0.051	***
Structure Type			
One Unit Detached	0.97	0.044	
One Unit Attached	0.91	0.065	
Manu./Mobile Home	0.94	0.072	
Region			
Midwest	1.07	0.037	*
South	1.04	0.035	
West	1.29	0.040	***
Personal Visit	1.41	0.026	***
Intercept	0.18	0.071	***
-2 Log Likelihood	121,260,365		
Model Chi Square	4,902,299.32		
DF	31		
Total Cases	32,723		

Source: 2007 American Housing Survey
*=p<.10, **=p<.05, ***=p<.01

4.4 Gas Bill Model[9]

Model 2 predicts the log odds of having all four gas bills (Table 2). Exp(B)'s are presented in the table. They represent the odds that a predictor variable will have all four bills when compared to the reference group. Age, education, ownership, relation to the householder, and tenure were found to have strong effects. Respondents who are 65 and older are 2.37 times more likely than respondents under 35 years old to have all four bills. Respondents living in owned units are 1.43 times more like likely to have all four bills compared with respondents in rented units. Respondents with an advanced degree are 1.58 times more likely to have all four bills than respondents with less than a high school education. Respondents with household incomes in the fifth income quintile are 1.29 times more likely than respondents in the first income quintile to have all four bills. Respondents who had an in-person interview were 1.41 more likely than respondents who had a telephone interview to have all four bills. Respondents who are the child of the householder were .59 times as likely to have all four bills compared with respondents who are the householder.

5. Summary and Future Research

5.1 Summary

Most findings matched theoretical expectations. Regarding the effect of bill paying responsibility, the householder or spouse of the householder was more likely than the child of the householder to have all four bills. Regarding who is more likely to have bills readily available, older respondents and those with higher levels of education, higher incomes, and owners are likely to have all four bills. Owners were more likely than renters to have all four bills available. Personal interviews were expected to negatively affect bill usage. Results for mode of interview did not match expectations with personal interviews being more likely than telephone interviews to have all four bills. In personal interviews, respondents may have more social pressure to retrieve records when the interviewer is standing in front of them.

---

[9] Reference groups in the model include: White only; non-Hispanic; <35 years old; less than high school education; first monthly income quintile; not married; reference person; renter; living in the Northeast; unit in a building with 2+ apartments; and telephone interview.

Table 2: Gas Bill Model

Ind Vars	Exp(β)	Standard Error	Significance
Sex			
Male	0.99	0.038	
Race			
Black Only	0.75	0.061	***
Asian Only	0.58	0.099	***
Other	0.86	0.126	
Hispanic Origin			
Hispanic	0.91	0.060	
Age			
35-44	1.18	0.058	***
45-54	1.51	0.058	***
55-64	1.62	0.061	***
65+	2.37	0.061	***
Marital Status			
Married	1.09	0.060	
Relation to Ref. Person			
Spouse	1.09	0.050	*
Child	0.59	0.161	***
Other Relative	0.64	0.217	**
Not in a family	1.00	0.059	
Tenure			
Owned	1.43	0.059	***
No Rent	1.06	0.166	
Education			
HS	1.08	0.065	
Some College	1.33	0.065	***
College	1.45	0.071	***
Advanced Degree	1.58	0.079	***
Household Income Quintile			
Second	1.04	0.063	
Third	1.30	0.064	***
Fourth	1.19	0.171	**
Fifth	1.29	0.256	***
Structure Type			
One Unit Detached	1.03	0.065	
One Unit Attached	0.96	0.089	
Manu./Mobile Home	1.04	0.139	
Region			
Midwest	1.22	0.052	***
South	1.25	0.054	***
West	1.39	0.055	***
Personal Visit	1.41	0.035	***
Intercept	0.136	0.106	***
-2 Log Likelihood	63,391,072		
Model Chi Square	3,164,410.58		
DF	31		
Total Cases	17,162		

Source: 2007 American Housing Survey
*=p<.10, **=p<.05, ***=p<.01

5.2 Future Research

This study examines the question of who has records and uses them to answer survey questions, but does not assess whether record use improves response accuracy. To assess accuracy, one could compare actual records from utility companies to the responses reported by respondents. One could also ask for a copy of the respondent's own records to identify any possible problems encountered as the respondent reads the value off of the record to the interviewer. Research could also be conducted on whether advance letters improve record usage. A split panel study could be designed to test the advance letter effect and the effects of advance letter content on respondent record usage. Alternatively, other methods can be developed to collect utility data. As is done in the Residential Energy Consumption Survey (RECS), respondents can sign waivers so that their utility data can be requested from utility companies. Additionally, one could explore collecting tract level utility information from utility companies and model utility costs based upon location, housing, and household characteristics. These methods would reduce respondent burden by taking away the need to ask as many questions about utilities on surveys.

Acknowledgements

References

De Leeuw, Edith, Mario Callegaro, Joop Hox, Elly Korendijk, and Gerty Lensvelt-Mulders. 2007. "The Influence of Advance Letters on Response in Telephone Surveys: A Meta-Analysis." *The Public Opinion Quarterly*, Vol. 71, No. 3, pp. 413-443.

Goldstein, Kenneth M. and M. Kent Jennings. 2002. "The Effect of Advance Letters on Cooperation in a List Sample Telephone Survey." *The Public Opinion Quarterly*, Vol. 66, No. 4, pp. 608-617.

Hembroff, Larry A., Debra Rusz, Ann Rafferty, Harry McGee, and Nathaniel Ehrlich. 2005. "The Cost-Effectiveness of Alternative Advance Mailings in a Telephone Survey." *The Public Opinion Quarterly*, Vol. 69, No. 2, pp. 232-245.

Link, Michael W. and Ali Mokdad. 2005. "Advance Letters as a Means of Improving Respondent Cooperation in Random Digit Dial Studies: A Multistate Experiment." *The Public Opinion Quarterly*, Vol. 69, No. 4, pp. 572-587.

Mann, Christopher, B. (2005). "Do Advance Letters Improve Preelection Forecast Accuracy?" *The Public Opinion Quarterly*, Vol. 69, No. 4, pp. 561-571.

Marquis, K., J. Moore, and K. Bogen. 1993. "Effects of a Cognitive Interviewing Approach on Response Quality in a Pretest for the SIPP." *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 318-323.