EXAMINATION OF CENSUS OMISSION AND ERRONEOUS ENUMERATION BASED ON 1990 ETHNOGRAPHIC STUDIES OF CENSUS COVERAGE

Yukiko Ellis, U.S. Bureau of the Census
4401 Suitland Rd., Washington, D.C. 20233

Key Words: Census Coverage Errors, Ethnographic Evaluations

This paper reports results of an analysis of data from the Ethnographic Evaluation of the Behavioral Causes of Census Undercount for the 1990 Decennial Census (referred to as the 1990 Ethnographic Evaluation henceforth). The purpose of the analysis was to search for factors related to the two components of census coverage errors, omission and erroneous enumeration of persons, in the 1990 census.

An earlier paper (de la Puente, 1993a) on the 1990 Ethnographic Evaluation focused on census omissions and examined the effects of demographic variables on the outcome of census enumeration. This paper broadened the scope of the analysis on two fronts. First, not only census omissions but also erroneous enumerations were examined. Second, in addition to demographic variables, factors related to the social aspects of the sample areas were included in the analyses. The results of the analyses confirmed many of the results of the earlier studies on the census coverage errors but also shed some new light on the possible effects that social and demographic factors might have had on the outcome of the census enumeration.

BACKGROUND

The Census Bureau began a series of ethnographic evaluations of census coverage in 1986, culminating in the 1990 Ethnographic Evaluation. The history and the study design of the 1990 Ethnographic Evaluation have been documented by Brownrigg and Martin (1989). The 1990 Ethnographic Evaluation consisted of intensive studies of twenty-nine small areas conducted by ethnographers. Each principal ethnographer had a close tie with the community and previously worked in and resided near the study area. As part of the evaluation project, each ethnographer conducted an Alternative Enumeration (AE) which was an independent (from the census) listing of the Census Day residents in the sample area, using participant observation and ethnographic interviews. The AE person list was later linked to the census person list, and persons missed or erroneously counted by the census were identified in the Resolved Enumeration (RE). Each sample area included about 100 households in one or more census blocks. Twenty-eight sample areas were located in the continental U.S. and one in Puerto Rico. This paper will be concerned only with the twenty-eight sample areas in the continental U.S. The sample areas were selected, purposively, representing five groups (Blacks, Hispanics, Asians, American Indians, and recent immigrants) in which undercounts were known or suspected to be high. The sample areas were also selected from three settings: ethnically homogeneous urban areas, ethnically heterogeneous urban and suburban areas, and ethnically homogeneous rural areas. In all, there were a total of 110 census blocks, 3367 housing units and 8718 individuals in the RE.

One of the goals of the ethnographic evaluations was to understand and identify causes of differentially high undercount of minority males, especially of Black and Hispanic males. In the Ethnographic Coverage Reports (de la Puente, 1993b), the ethnographers reported that, in almost all sample areas, a group of factors, rather than one single factor, contributed to census omission and erroneous enumeration. The factors most frequently cited by the ethnographers were:

* Irregular and complex household arrangements;
* Language and illiteracy barriers;
* Concealment of information to protect resources; and
* Missed or erroneously enumerated housing units.

The ethnographers attempted to quantify the above factors and additional information such as residential mobility and presence of violence in the behavioral log that each ethnographer was asked to keep during the project. However, a review of the behavioral logs revealed variations in consistency and completeness, raising concerns about the reliability of cross-site comparisons.

This paper extended the study of de la Puente (1993a) by examining erroneous enumeration as well as census omission. The effect of sample areas on the census coverage was also investigated. Its importance was discussed but not quantified by de la Puente in his paper. In addition, the paper attempted to corroborate quantitatively some of the findings in the Ethnographic Coverage Reports through use of the census long-form questionnaires, in lieu of the behavioral logs, as proxies to summarize the social, economic, and educational backgrounds of the persons residing in and around the sample areas. Together with the demographic information from the AE and a variable that attempted to summarize the sample area effects, subsets of these factors that best predicted the outcome of either census

omission or erroneous enumeration were obtained.

LIMITATIONS

The AE data from the twenty-eight sample areas did not represent a probability sample. Hence, the results from this study should not be generalized to any population or group beyond the twenty-eight sample areas in the study.

The Race variable in this study was defined as Hispanic, Black, and Other. The group "Other" group included Asians, American Indians, and Whites. This definition of "Other" category hence limited what one could learn specifically about Asians and American Indians.

The data based on the census long-form questionnaires from each sample area and one ring of blocks surrounding the sample area were used in the analyses under an assumption that the socioeconomic landscape stays stable over an area covering several contiguous census blocks at a fixed point in time. (Ellis, 1995).

METHODOLOGY

Two logistic regression models were fit to data derived from the RE and census long-form questionnaires, using the SAS's LOGIST procedure with stepwise option: one for census omission with two response categories (missed/correctly enumerated in the census), the second one for erroneous enumeration (erroneously/correctly enumerated).

From the RE data, the following eight explanatory variables were defined: Age, Gender, Geography, Household (HH) size, Marital status, Race, Relation of an individual to the householder (in whose name the house was owned or rented), and Source. Source indicates whether the ethnographer enumerated the individual by direct observation, by information supplied by a household resident, or by other means such as information from neighbor, owner of building or administrative records. All but the Household size variable were treated as discrete variables.

The following eight continuous variables were defined based on the census long-form questionnaires collected within the sample areas and their surrounding blocks:

%Ereturn: % households enumerated by enumerators and not by mail;
%FemHH: % female householders with no spouse;
%Foreign: % persons born abroad;
%LowEd: % persons with less than high school education among persons 18+ years old;
MedInc: Median household income;
%OthLan: % persons who spoke a language other than English at home;
%Owner: % owner-occupied housing units;and
%Vac: % vacant housing units;

These eight variables were also used to group the twenty-eight sample areas into clusters, using the SAS's CLUSTER procedure. See Ellis (1995) for clustering of the sample areas. Figure 1 illustrates the five clusters of the twenty-eight sample areas in the framework of the sample design by race/ethnicity and type of setting. The five clusters are:

Cluster 1 (9 sample areas): Hispanic and Asian immigrants with low MedInc and high %LowEd;
Cluster 2 (9 sample areas): Blacks, high %FemHH and high %Vac;
Cluster 3 (5 sample areas): Rural homeowners, Hispanic and American Indian;
Cluster 4 (3 sample areas): Hispanic and Asian immigrants with high MedInc and low %LowEd;
Cluster 5 (2 sample areas): List/Enumerate sample areas.

For each of the discrete explanatory variables, a set of design variables was formed to represent the categories of the variable, using the reference cell coding method (p.48, Hosmer and Lemeshow, 1989). The category in which the persons had the lowest odds of being missed (or erroneously enumerated) in the census, given all other explanatory variables in the model, was chosen as the reference cell. For a continuous variable, the estimated odds ratio was computed for an increase of twenty percent in the variable. For the MedInc variable, the estimated odds ratio was computed for an increase of $1000 in the variable.

RESULTS

Census Omission

Table 1 shows the maximum likelihood estimates of the coefficients, standard errors, and odds ratios for a multiple logistic regression model in which the census omission was the binary response variable.

All eight discrete variables were found to contribute significantly in explaining the outcome of census omissions, after controlling for the variables in the model. The Cluster and Relation variables were found to have the strongest effects on the response variable. Persons in Cluster 4 (Hispanic and Asian immigrants with high median household income) were twenty-nine times more likely and persons in Cluster 1 (Hispanic and Asian immigrants with low median household income) were eleven times more likely to be missed by the census than persons in Cluster 3 (Rural homeowners, Hispanic and American Indian). Clusters 3 and 5, which included rural sample areas in the study, had relatively low odds of persons being missed in the census in comparison to other clusters.

Persons not related to householders in the RE were almost four times more likely to be missed by the census than householders. With respect to the race

variable, the odds ratios of 1.87 for Hispanic and 1.17 for Black in relation to Other were lower than what one might have expected, probably because the sample in this study included disproportionately large proportions of Asians (1/3) and American Indians (1/3) in the 'Other' race category compared to the general population.

Among the eight census long-form variables and the HHsize variable, all but the MedInc, %Othlan, and %Ereturn variables were found to contribute significantly to the model, given all other variables in the model. Among the six continuous variables, %Foreign had the largest effect on the response. The negative sign on its parameter estimate indicates that the higher the proportion of persons born abroad in a sample area the lower the likelihood of a resident of the sample area being missed in the census. This finding appears to contradict what one might have expected. This will be further discussed later. For %LowEd, an increase of 20 percent in the proportion of persons with less than high school education in a sample area would increase the chance of being missed by the census by almost twofold.

Table 2 contains a table of the observed proportion of census omissions, conditional on Relation and Race. The proportions of census omissions for householders and spouses were found to be close in magnitude in each race/ethnicity category. The proportion was slightly higher for "other relative" than for householders and spouses. For persons not related to householders, the proportions increased greatly and ranged from 43 percent to 46 percent across the Race categories. Hence, these proportions in Table 2 indicated, first, that there appeared to be a two-way interaction between the Race and Relation variables; and secondly, given that a person was not related to the householder, the proportion of census omissions appeared to be statistically independent of the Race variable.

Table 3 shows a frequency table of enumeration status (correctly enumerated/partial HH miss/whole HH miss) by Race and Geography. A "partial HH miss" is a person who was missed in a household in which at least one other resident was correctly enumerated in the census. A person was a "whole HH miss" if nobody in the household was correctly enumerated by the census.

Table 3 shows that each race group had a distinct pattern of census omission by geography. Among persons in the "Other" race category, persons living in urban areas were most likely to be missed in the census compared to those living elsewhere. For each type of geography, the proportion of whole HH misses was almost twice as big as the proportion of partial HH misses. In the Black category, persons living in rural and urban areas were more likely to be missed than those living in urban/suburban areas. In rural areas, a Black person was five times more likely to be a whole HH miss than to be a partial HH miss. These whole household misses might reflect the difficulty in finding housing units, and consequently, missing everybody living in the units, in sparsely populated rural areas where, for example, addresses might not be marked clearly or units were hidden from public view down rural roads. The whole HH misses were also predominant among Blacks in urban areas as well. In the Harlem, NY, sample area where crimes were widespread, run-down buildings appeared abandoned but were not, and some brownstones seemed to be one family dwelling, but in fact, contained numerous housing units, all missed by the census. In the Hispanic race category, persons living in urban/suburban and urban areas were more likely to be missed than those living in rural areas. Partial HH misses were just as numerous as whole HH misses at each level of Geography among Hispanics. This could be partially explained by the ethnographers' observations that irregular housing went hand in hand with complex or irregular household arrangements. Because of a shortage of affordable housing, families and unrelated individuals doubled up in single housing units, some of which might have been illegally converted.

Erroneous Enumeration

Table 4 shows the maximum likelihood estimates of the coefficients, standard errors, and odds ratios for a multiple logistic regression model in which the erroneous enumeration was the binary response variable. Among the eight discrete variables, the Gender, Race, and Geography variables were found not to contribute significantly in predicting the erroneous enumeration, after statistically adjusting for all other variables in the model. Among the remaining four discrete variables, the Cluster variable was again found to play a major role in predicting the response. This time, however, persons in Cluster 1 (Hispanic and Asian immigrants with low median household income) were found to be seventeen times more likely to be erroneously enumerated than persons in Cluster 4 (Hispanic and Asian immigrants with high median household income). In the Miami, FL, sample area in Cluster 1, households were visited more than once by the census workers, resulting in duplicate enumeration. In the San Diego, CA, sample area, also in Cluster 1, irregular housing was responsible for multiple enumeration and other erroneous enumerations of households. More than half of the erroneous enumeration in Cluster 3 (Rural homeowners, Hispanic and American Indian) took place in the Marion County, OR, sample area where migrant workers in a migrant worker camp were erroneously enumerated by the

census.

Next to the Cluster variable, the Relation variable had a strong effect on the response variable. Persons not related to householders and "other relatives" had higher odds of being erroneously enumerated than householders and spouses. Residential mobility among persons not related to householders was mentioned in the Ethnograhic Coverage Reports as one of the contributing factors of erroneous enumeration in the sample areas with large immigrant populations. In another case, an erroneous inclusion of adult children was cited as one explanation of erroneous enumeration in the two sample areas with a sizable Chinese population.

Among the nine continuous variables, the %Foreign, %FemHH, %Vac, and HHsize variables did not contribute significantly to the model, after controlling for all other variables in the model. Among the remaining five continuous variables, the %OthLan variable had the strongest effect on the response. The %OthLan variable, as with the %Foreign variable for census omission, had a negative sign on its parameter estimate, indicating that the lower the proportion of persons speaking a language other than English at home the higher the likelihood of erroneous enumeration.

As in the case of census omission, the higher the proportion of persons with less than high school education and the lower the proportion of owner-occupied housing units in a sample area the higher the likelihood of erroneous enumeration.

The MedInc variable, with a positive sign on its parameter estimate, indicated that the higher the median household income of a sample area the more likely it was for persons living in the sample area to be erroneously enumerated. Note that the median household income at site level in the study ranged from $5,000 to $30,000.

DISCUSSION

This paper has attempted to describe patterns of undercount and overcount within the selected sample areas, and to examine possible behavioral causes through indirect measurements of the social, economic, and educational backgrounds of the study areas.

The comparison of the multiple logistic regression models for census omission (Table 1) and erroneous enumeration (Table 4) indicates that persons who had high odds of being missed in the census had both similarities and differences from persons who had high odds of being erroneously enumerated. A person who was either in the age group of 18-29 or who was not related to the householder had high odds of not only being missed, but also being erroneously enumerated in the census, controlling for all other variables in the model. Also, the higher the proportions of persons with less than high school education or in renter-occupied units in a sample area the higher the odds of census omission and erroneous enumeration in the sample area. The result on the educational variable was consistent with the ethnographers' observations that illiteracy among recent immigrants was one of the contributing factors to census coverage errors. The result on the tenure variable might be a reflection of irregular and complex household arrangements being more prevalent among renter-occupied units where the tenants might be unwilling to reveal their living arrangements to an outsider, as often reported by the ethnographers. Also, one would expect persons in renter-occupied units to be more mobile than persons in owner-occupied units, leading to more coverage errors.

For both census omission and erroneous enumeration, the Cluster variable played a major role. Albeit the sample in the study was not a probability sample, the result of the analyses poses a question about the validity of the assumption of geographic homogeneity with respect to census coverage. In 1990 the mechanism of undercount was assumed to be different by region. Hence, post-strata were defined within the Census Division. Under this assumption, one might have expected large variability in coverage errors within clusters in this study since each cluster consisted of the sample areas from different Census Regions. Yet, the estimated standard errors for the clusters were found to be stable and small in comparison to their estimated coefficients both for census omission and erroneous enumeration.

Another question of interest is: Does a stratification scheme that incorporates variables on socioeconomic/ educational backgrounds of the sample areas help fine-tune the scheme that only utilizes the race/Hispanic origin and urban/rural variables?

Both multiple logistic regression models included a few parameter estimates whose signs were contrary to the conventional wisdom. The %Foreign and %OthLan variables were good indicators of language barriers that might have existed among the residents in a sample area. The ethnographers often cited a language barrier to be one of the important contributing factors to census coverage errors, especially in the sample areas heavily populated by Hispanic and Asian persons. One possible explanation for the negative signs on their parameter estimates is that it was not the lack of knowledge of English per se that caused a person to be missed or erroneously enumerated in the census in this study group. The sample areas with large proportions of recent immigrants included larger proportions of persons with the characteristics that influenced census coverage errors (such as being 18-29 years old, unrelated to the householder, and living in a renter-

occupied housing unit) than the remaining sample areas. When we statistically adjusted for these characteristics, we found that the %Foreign and %OthLan variables had an inverse relationship with census coverage errors.

In order to estimate numbers and characteristics of people missed and erroneously enumerated in the 1990 Census, the Census Bureau conducted the Post-Enumeration Survey (PES), an independent coverage survey with a probability sample, a few months after the census. Refer to Hogan (1993) for a complete background discussion of the 1990 PES. The 1990 Ethnographic Evaluation results agreed with most of the findings from the 1990 PES and de la Puente (1993a). One factor that surfaced as having an especially important role in predicting within-household census coverage errors in this study as well as in Moriarity (1993) and Ellis (1994) was the Relation variable. If a person was not related to the householder, then the person was found to have a very high risk of being either missed or erroneously enumerated in the census. For example, these are:

* Unrelated individuals living together for the sole purpose of sharing the rent;
* Individuals in households that contained two or more "nuclear" families; and
* Mobile or ambiguous household members.

Ellis (1994) reported that, nationwide, Blacks and Hispanics had larger proportions of persons not related to householders (48 percent and 49 percent, respectively) compared to White and Other (33 percent). The proportion of being missed in the census among persons not related to householders was uniform across tenure groups and household sizes. In this ethnographic study, it was again found that the proportion of being missed in the census was uniform across race groups among persons not related to householders. In other words, given that a person was not related to the householder, the probability of being missed in the census seemed to be independent of other factors. Hence, the differences in household composition may be the main contributing factor of the differential undercount within households. Fay (1989) also suggested this possibility in his analysis of the Current Population Survey data.

SUGGESTIONS FOR FUTURE RESEARCH

The ethnographic studies have proved useful in elucidating various causes of undercount. More research is needed to see whether a post-census coverage survey can incorporate the ethnographic evaluation to improve census coverage, especially in the hard-to-enumerate areas.

More research is needed on improving enumeration and estimation of persons not related to householders. The research could be done at different stages: at the census questionnaire design stage (revise wording and add probes, for example), at the sample selection stage (target areas with an overcrowding problem), or at the estimation stage (post-stratify by the Relation variable). These issues address the within-household coverage errors. As for the issues addressing the whole-household errors, a better method of address listing needs to be devised, especially for irregular and hidden housing units. Once again, a possible solution is the utilization of participant observations in ethnographic studies in which somebody familiar with each sample area would canvass and list housing units.

REFERENCE

Brownrigg, L. and Martin, E. (1989). "Proposed Study Plan for Ethnographic Evaluation of the Behavioral Causes of Undercount." Paper presented to the Census Advisory Committees of the American Statistical Association and on Population Statistics at the Joint Advisory Committee Meeting, Alexandria, VA, April 13-14, 1989.

de la Puente, M. (1993a). "A Multivariate Analysis of the Census Omission of Hispanics and Non-Hispanic Whites, Blacks, Asians and American Indians: Evidence from Small Area Ethnographic Studies." Proceedings of the Survey Research Methods Section, American Statistical Association, 641-646.

de la Puente, M. (1993b). "Why Are People Missed or Erroneously Included by the Census: A Summary of Findings from Ethnographic Coverage Reports." Paper prepared for the Advisory Committee for the Design of the Year 2000 Census Meeting, Suitland, MD, March 3, 1993.

Ellis, Y. (1994). "Categorical Data Analysis of Census Omissions." DSSD 1990 REX Memorandum Series #PP-10. Washington, DC: Bureau of the Census.

Ellis, Y. (1995). "Clustering of Twenty-Eight Sample Areas in the Ethnographic Evaluation of the Behavioral Causes of Census Undercount for the 1990 Decennial Census." Unpublished document. Washington, DC: Bureau of the Census.

Fay, R. (1989). "An Analysis of Within-household Undercoverage in the Current Population Survey." Proceedings of the Bureau of the Census Fifth Annual Research Conference, pp.156-175. Washington, DC: Bureau of the Census.

Hogan, H. (1993). "The 1990 Post-Enumeration Survey: Operations and Results." Journal of the American Statistical Association, 88, 1047-1060.

Hosmer, D. and Lemeshow, S. (1989). Applied Logistic Regression. New York: Wiley.

Moriarity, C. (1993). "Characteristics of Census Error-Additional Results." DSSD 1990 REX Memorandum Series #PP-7. Washington, DC: Bureau of the Census.

Figure 1. Five Clusters of 28 Sample Areas in the Framework of Sample Design By Race/Ethnicity and Type of Setting

Race/ Ethnicity	Type of Setting		
	Urban	Urb/Sub	Rural
Asian	● ● ○	● ● X X	
Immigrant	● ●	● ○ ○	
Hispanic	●	● X	■ ■
Black	X X X X	X ■	X □
Am.Indian			■ ■ □

Legend: ●: A sample area in Cluster 1;
X: A sample area in Cluster 2;
■: A sample area in Cluster 3;
○: A sample area in Cluster 4;
□: A sample area in Cluster 5.

TABLE 2. Observed Proportion (and Number) of Census Omission, Conditional on RELATION and RACE

RELATION	RACE		
	Other	Black	Hispanic
Householder	10% (111)	19% (124)	12% (71)
Spouse	11% (64)	20% (40)	13% (44)
Other Relative	14% (197)	20% (178)	20% (239)
Not Related	43% (31)	46% (42)	45% (80)

TABLE 3. Type of Census Omission by RACE and GEOGRAPHY

RACE AND GEOGRAPHY	TYPE OF CENSUS OMISSION		
	Correct Enumeration	Partial HH Omission	Whole HH Omission
Other:Urban	696 (83%)	54 (6%)	88 (11%)
Urb/Sub	1082 (89%)	45 (4%)	92 (7%)
Rural	970 (90%)	37 (3%)	74 (7%)
Black:Urban	682 (78%)	57 (7%)	133 (15%)
Urb/Sub	473 (83%)	38 (7%)	59 (10%)
Rural	288 (75%)	16 (4%)	81 (21%)
Hispanic:Urban	532 (77%)	74 (11%)	82 (12%)
Urb/Sub	621 (75%)	80 (10%)	126 (15%)
Rural	722 (91%)	42 (5%)	30 (4%)

Maximum Likelihood Estimates of the Multiple Logistic Regression Model of the Probabilities of Being:

TABLE 1. Missed in the Census

Variable	Est. Coeff	Est. Std.Err	Est.Odds Ratio
RACE: Black vs Other	0.1541	0.1053	1.17
Hisp vs Other	0.6270	0.1008	1.87
AGE: 0-17 vs 50+	0.2208	0.1378	1.25
18-29 vs 50+	0.5810	0.1223	1.79
30-49 vs 50+	0.2969	0.1117	1.35
GENDER: Male vs Female	0.2021	0.0704	1.22
RELATION: Spouse vs Hsehldr	0.1444	0.1224	1.16
Oth Rel vs Hsehldr	0.2770	0.1157	1.32
Non-Rel vs Hsehldr	1.3323	0.1467	3.79
MARITAL STATUS:			
Marrd vs Sp/Dv/Wd	0.1352	0.1336	1.14
Single vs Sp/Dv/Wd	0.3428	0.1248	1.41
CLUSTER: Clstr1 vs Clstr3	2.3731	0.2542	10.73
Clstr2 vs Clstr3	1.7942	0.2654	6.01
Clstr4 vs Clstr3	3.3662	0.2745	28.97
Clstr5 vs Clstr3	0.3006	0.2479	1.35
GEOGRAPHY: Urban vs Urb/Sub	0.2625	0.1078	1.30
Rural vs Urb/Sub	0.9570	0.1852	2.60
SOURCE: Obser vs HHmem	0.5072	0.0831	1.66
Other vs HHmem	0.6137	0.1385	1.85
HHsize	0.0546	0.0163	1.06
%LowEd	2.5600	0.4783	1.67*
%Foreign	-5.3612	0.4341	0.34*
%FemHH	-3.6582	0.5198	0.48*
%Owner	-1.2160	0.2932	0.78*
%Vac	-2.2777	0.5227	0.63*
Constant	-2.7399	0.3872	

TABLE 4. Erroneously Enumerated in the Census

Variable	Est. Coeff	Est. Std.Err	Est.Odds Ratio
AGE: 18-29 vs 0-17	0.6325	0.1178	1.88
30-49 vs 0-17	0.5402	0.1473	1.72
50+ vs 0-17	0.3355	0.1534	1.40
RELATION: Hsehldr vs Spouse	0.1158	0.1311	1.12
Oth Rel vs Spouse	0.6175	0.1572	1.85
Non-Rel vs Spouse	0.6245	0.2121	1.87
MARITAL STATUS:			
Marrd vs Sp/Dv/Wd	0.3570	0.1450	1.43
Single vs Sp/Dv/Wd	0.2002	0.1516	1.22
CLUSTER: Clstr1 vs Clstr4	2.8222	0.3159	16.81
Clstr2 vs Clstr4	1.7959	0.3091	6.02
Clstr3 vs Clstr4	2.3183	0.2546	10.12
Clstr5 vs Clstr4	0.4241	0.4691	1.52
%LowEd	1.2864	0.4442	1.29*
MedInc	1.98E-4	0.17E-4	1.22**
%OthLan	-2.9151	0.2989	0.56*
%Owner	-1.9124	0.2927	0.68*
%Ereturn	2.5909	0.4976	1.68*
Constant	-8.0719	0.5293	

* The estimated odds ratio was computed for an increase of 20% in this variable.

** The estimated odds ratio was computed for an increase of $1000 in the median household income.