Relationship Between Household Nonresponse, Demographics, and Unemployment Rate in the Current Population Survey.

John Dixon, Bureau of Labor Statistics, Room 4915, 2 Massachusetts Ave., NE,
Washington, DC 20212-0001
Dixon_J@bls.gov
Keywords: "Survey nonresponse, Gross Flows, Unemployment"

Introduction

In the Current Population Survey, a household survey from which labor force estimates are produced, selected housing units remain in sample during a 16-month period. The households are interviewed during the first 4 and last 4 months of this period. These interview months are referred to as "month-in-sample" (MIS) 1 to 8.

Matching households between months allows an analysis of the relationship between nonresponse and estimates of the employment rate. Since change in employment may be related to the household's participation, the estimates of employment status may be affected. A recent study by Tucker and Kojetin (1997) showed that unemployment rates were related to nonresponse in the CPS. "Converts" (households that do not participate in the prior month) do not completely make up for the number of "Attriters" (households that do not participate in the following month), so their relative effect may not be offset. Moreover, they may differ on important characteristics, e.g.; race, ethnicity, or gender. The current study examines the nature of this relationship through an analysis of demographics and nonresponse and their resulting effect on labor force estimates.

Gross Flows

In this study "gross flows" uses the availability of information on one month to contrast the estimates from another month. For example, labor force estimates in month 1 are contrasted based on whether a household responded in month 2, and labor force estimates in month 2 are contrasted based on whether a household responded in month 1. For example; if the unemployment rate for month 1 is different for households who continued to respond in month 2 compared to those who did not respond, and this was not balanced by a difference in the other direction for those who responded in month 2 but did not respond in month 1, then some the estimates would be biased due to nonresponse.

Design

The CPS is a the monthly household labor force survey for the United States conducted by the U.S. Census Bureau for the U.S. Bureau of Labor Statistics. Approximately 48,000 eligible households are sampled each month in a two-stage clustered design. Households were matched for the years 1996 through 1999. Persons in the household who were not eligible for the labor force (e.g. under 16 years old) were excluded.

Analysis

The following tables are based on CPS adjacent months-in-sample data weighted by the base weight, which reflects the probability of selection, but does not adjust for non-response. Because of the differences in weighting, the labor force estimates will not be comparable to published estimates. The percentages reported are relative to the other categories, not the traditional unemployment rate, which is only relative to those in the labor force. The Mantel-Haenszel test provides a comparison of the availability of the data (non-response status for each month separately).

The Cochran-Mantel-Haenszel test provides a test of the comparability of tables contrasting months, and can be used as an indicator of the gross flow effect. None of the p-values are adjusted for multiple testing. The complex sampling used by the CPS was not accounted for in the p-values of the models.

Linear models provide a comparison of the means for unemployment for a number of demographic variables. The interaction of the "interview status" variable (response or nonresponse) and the "flow" variable (adjacent months) gives an estimate of the gross flow. Higher order interactions with the demographic variables show if they are related to any bias estimated by gross flows. Although none of the p-values are adjusted for multiple testing, the complex sampling is accounted for using the SAS™ procedure "surveyreg". The correlation between months was ignored in the design. Tables are provided for total nonresponse as well

Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

as for refusal and noncontact. The theory of nonresponse suggests that different causes may produce refusal and noncontact, but the combined effect is also of interest here, since that would produce the aggregate effect on estimates.

Results

An overall test of the impact of non-response on labor force estimates was examined in Table 1 by comparing two 3 by 2 tables (labor force by month). The 2nd month non-response was related to the 1st month labor force status (Mantel-Haenszel=172.009, p<0.001). Unemployment and employment were higher while those not in the labor force were lower for the non-response group. Similarly, the 1st month non-response was related to the 2nd month labor force status (Mantel-Haenszel-8.620, p<0.003). Employment was higher while unemployment and not-in-labor-force were lower. This difference between the two tables is reflected in the Cochran-Mantel-Haenszel test (623.421, df=2, p<0.0001) which contrasts the rows of the two tables. The gross flow of employment status from month to month is impacted by non-response, with unemployment reversing direction depending on whether the non-response occurred in the first or second month.

Table 1

Labor Force Status by Interview Status			
1st Month Labor Force	**2nd Month interview**	**2nd Month nonresponse**	
Not in labor force	34.48%	30.08%	Mantel-Haenszel= 632.373, p< 0.0001
Employed	62.08%	65.95%	
Unemployed	3.43%	3.98%	
2nd Month Labor Force	**1st Month interview**	**1st Month nonresponse**	
Not in labor force	34.57%	31.46%	Mantel-Haenszel= 365.398, p< 0.0001
Employed	62.02%	65.00%	
Unemployed	3.41%	3.54%	
	Cochran-Mantel-Haenszel (row mean scores)= 2098.591 (df=2) p< 0.0001		

A simpler form of the gross flow matrix using just the unemployed relative to the employed would be:

Table 2 Unemployment ratio

	Interview Status		
Flow	I	N	All
Month 1	0.068	0.073	0.068
Month 2	0.070	0.069	0.070
All	0.069	0.071	0.069

In this table the unemployment ratio relative to employed was contrasted by whether they were interviewed in the adjacent month or not. This shows the higher unemployment rate of those who dropped out relative to those who stayed in. Those who converted the second month had a lower unemployment rate. Because more dropped out than were converted, the impact is almost entirely from the dropouts. This simpler table makes the display of effects relative to unemployment clearer for more complex gross flows. It also shows that the interviewed persons (I column) have the same rates as the aggregate column (ALL) which adds in the estimated effect for nonresponse (N column). This lack of effect on the estimates is due to the very small amount of nonresponse in the CPS. These numbers are weighted by the baseweight, which adjusts for the design, but doesn't adjust for nonresponse. The nonresponse adjustment would reduce the effect further. Models which estimate parameters for the tables presented here are in Appendix A (available in the long version of this paper).

Table 3: Type of nonresponse effect.

	Nonresponse Type			
Flow	I	N	R	All
Month 1	0.068	0.083	0.067	0.068
Month 2	0.070	0.076	0.065	0.070
All	0.069	0.080	0.066	0.069

Table 3 shows the flow relative to the type of nonresponse (I: interview, N: noncontact, R: refusal). Refusals show lower unemployment while noncontact shows higher unemployment. The effect would tend to cancel one another out, reducing the bias problem. Noncontact shows a stronger effect.

Table 4a: Gender effects.

	Interview		Nonresponse	
Flow	Male	Female	Male	Female
Month 1	0.063	0.072	0.076	0.083
Month 2	0.065	0.075	0.068	0.075

Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

The gross flows relative to gender shows higher unemployment for attrition, but a negligible effect for those who responded in the second month in sample. The effect appeared consistent for both genders.

Table 4b: Gender effects.

	Male			
Flow	I	N	R	All
Month 1	.064	.081	.065	.064
Month 2	.066	.076	.062	.066
All	.065	.079	.064	.065
	Female			
	I	N	R	All
Month 1	.072	.085	.069	.072
Month 2	.075	.076	.067	.075
All	.073	.081	.068	.073

Both genders showed a similar pattern as before; refusals show lower unemployment while noncontact shows higher unemployment. Males showed a stronger effect for refusal conversion (Month 2) than females.

Table 5a: Race effects.

	White		
Flow	I	N	All
Month1	.058	.064	.058
Month2	.060	.061	.060
All	.059	.063	.059
	Black		
	I	N	All
Month1	.138	.122	.137
Month2	.143	.112	.142
All	.141	.118	.140

This shows a lesser effect as before for Whites (.058 vs. .064 and .060 vs. .061), but a reverse effect for Blacks. The size of the effect might have an impact on the estimate for Blacks before adjustment for nonresponse (comparing the I column to the ALL column). This effect would be expected to disappear using the weights which compensate for nonresponse since race is one of the raking factors.

Table 5b: Race effects.

	White			
Flow	I	N	R	All
Month 1	0.058	0.071	0.060	0.058
Month 2	0.060	0.066	0.058	0.060
All	0.059	0.069	0.059	0.059
	Black			
	I	N	R	All
Month 1	0.138	0.139	0.108	0.137
Month 2	0.143	0.124	0.104	0.142
All	0.141	0.133	0.106	0.140

For Whites, the effect of noncontact attrition was strongest (Month 1) with higher unemployment. For Blacks refusal was strong for both attrition and conversion, but noncontact was more pronounced for conversion, with all effects related to lower unemployment.

Table 6: Month-in-Sample effect.

	Month 1		Month 2	
MIS	I	N	I	N
1-2	0.054	0.063	0.065	0.051
2-3	0.064	0.072	0.070	0.068
3-4	0.069	0.081	0.073	0.075
4-5	0.073	0.076	0.067	0.072
5-6	0.067	0.069	0.070	0.067
6-7	0.070	0.072	0.073	0.067
7-8	0.073	0.074	0.075	0.070

This shows the higher unemployment rate for attrition (.065 vs. .054) compared to a reduction for conversion (.051 vs. .063) for the first two months in sample. The higher effect for attrition is consistent throughout the 8 months in sample, while the conversion effect reverses for several month-in-sample pairs, which probably contributes to the strength of the effect. MIS 4-5 is unique in that there is an 8 month interval between interviews, which may account for the reversal between months relative to the other MIS.

Table 7: Teenage Unemployment

	Nonresponse type			
Flow	I	N	R	All
Month1	0.209	0.237	0.184	0.209
Month2	0.211	0.219	0.178	0.211
All	0.210	0.229	0.181	0.210

Teenagers show the same effect as seen before, with higher unemployment for noncontact and lower for refusals, although the combined effect had no effect on the estimates.

Table 8: Hispanic effects

	Not Hispanic			Hispanic		
Flow	I	N	All	I	N	All
Mo1	.063	.070	.063	.113	.105	.113
Mo2	.066	.065	.066	.117	.104	.116
All	.064	.068	.064	.115	.105	.115

This shows Hispanics who don't respond have lower unemployment, but doesn't appear to have any effect on the overall estimate for Hispanics (0.115).

Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.

Linear Models

The following tables show the tests for gross flows using linear models. The "Flow" parameter shows the change between months, and the "Status" parameter shows the effect of nonresponse. The "Flow*Status" parameter indicates whether the flow is consistent between months relative to nonresponse. This is the gross flow indicator.

The variables added to the "flow" and "nonresponse" model were taken from the literature review and studies in Groves and Couper (1998), and a study by Tucker and Dixon (2000). They include; number of attempted contacts, presence of small children in the household, households in multilevel structures, household size, home ownership, relatives present, rural/urban, and population density.

In Table9a, the positive "Flow" parameter (.00286573) shows there is an increase in unemployment from month 1 to 2. The positive "Status" parameter (.0056345) shows a higher unemployment or nonresponse. The negative "Flow*Status" interaction parameter (-0.0070611) shows that the unemployment is higher for attrition (Month 1) than for conversion (Month 2).

Table 9a: Linear model of gross flow

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06760	0.00055412	<.0001
Flow	0.00285	0.00009748	<.0001
Status	0.00563	0.00183062	0.0021
Flow*Status	-0.00706	0.00081372	<.0001

Refusers have lower unemployment in month 2 relative to month 1 (-.0062967), and noncontacts have higher unemployment overall, but lower in month 2 relative to month 1.

Table 9b: Refusal and Noncontact

Parameter	Estimate	Std. Error	Pr>|t|
Intercept	0.06642	0.00055404	<.0001
Flow	0.00324	0.00010647	<.0001
Refuse	-0.00063	0.00240720	0.7924
Flow*Refuse	-0.00629	0.00087452	<.0001
Nocontact	0.01298	0.00232140	<.0001
Flow*Nocont	-0.00981	0.00166946	<.0001

The number of attempted contacts is a measure of how difficult a household was to contact. It was related to higher unemployment (CNT=0.0013294, adjusting for other variables), but had no detectable relationship to gross flow measures (cnt*flow; which tests the interaction between number of attempted contacts and month-to-month flow, cnt*status; which tests the interaction between number of attempted contacts and nonresponse, and cnt*status*fl; which test the interaction between month-to-month flow, nonresponse, and the number of attempted contacts).

Table 10a: Number of attempted contacts

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06542	.00080807	<.0001
Flow	0.00298	.00016807	<.0001
Status	0.00592	.00306062	0.0529
Flow*Status	-0.00837	.00135088	<.0001
CNT	0.00132	.00034298	0.0001
Cnt*flow	-0.00009	.00007551	0.2300
Cnt*status	-0.00039	.00102436	0.6990
Cnt*stat*fl	0.00061	.00045528	0.1784

The interaction between "number of attempted contacts", "month-to-month flow", and "noncontact" suggests the attrition and conversion effects have a counterbalancing effect, reducint the impact on estimates.

Table 10b: Number of attempted contacts

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06424	.00081185	<.0001
Flow	0.00330	.00018289	<.0001
refuse	0.00063	.00426348	0.8816
Flow*Refuse	-0.00642	.00149341	<.0001
Nocontact	0.01190	.00366549	0.0012
Flow*Nocont	-0.01328	.00257388	<.0001
CNT	0.00133	.00034465	0.0001
Cnt*flow	-0.00003	.00008138	0.6839
Cnt*refuse	-0.00081	.00148391	0.5832
Cnt*rf*fl	0.00003	.00055000	0.9468
Cnt*noc	0.00010	.00121789	0.9339
Cnt*nc*fl	0.00146	.00073672	0.0461

The presence of small children in the household was related to higher unemployment, but had no detectable relationship to gross flow measures.

Table 11a: Small children in the household

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06747	.00055344	<.0001
Flow	0.00283	.00009720	<.0001
Nonresponse	0.00559	.00183182	0.0023
Flow*Status	-0.00699	.00081247	<.0001
KID	0.08847	.01367031	<.0001
Kid*flow	-0.00062	.00358565	0.8608
Kid*status	0.00490	.04937519	0.9209
Kid*stat*fl	-0.03353	.02214511	0.1299

There was also no relationship to refusal and noncontact.

Table 11b: Small children in the household

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06631	.00055359	<.0001
Flow	0.00324	.00010625	<.0001
refuse	-0.00054	.00240900	0.8207
Flow*Refuse	-0.00628	.00087108	<.0001
Nocontact	0.01281	.00232106	<.0001
Flow*Nocont	-0.00959	.00166676	<.0001
KID	0.08532	.01402731	<.0001
kidfl	-0.00201	.00462458	0.6630
kidref	-0.06282	.05646700	0.2659
kidrffl	-0.00613	.04931289	0.9009
kidnoc	0.05786	.06287321	0.3574
kidncfl	-0.08470	.06457218	0.1896

The household living in a multiunit structure was related to higher unemployment, but had no detectable relationship to gross flow measures (mul*status).

Table 12a: Multiunit structure

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06235	.00056706	<.0001
Flow	0.00277	.00010218	<.0001
Nonresponse	0.00473	.00204878	0.0207
Flow*Status	-0.00619	.00087062	<.0001
MUL	0.01938	.00107100	<.0001
Mul*flow	-0.00002	.00017590	0.8675
Mul*status	-0.00405	.00265066	0.1257
Mul*st*fl	-0.00064	.00123850	0.6052

The effect of nonresponse and "multiunit structure" above is probably due to noncontact with households living in multiunit structures and not contacted having lower unemployment (-0.0105104).

Table 12b: Multiunit structure

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.06143	.00056668	<.0001
Flow	0.00300	.00011052	<.0001
refuse	-0.00310	.00260585	0.2328
Flow*Refuse	-0.00546	.00090727	<.0001
Nocontact	0.01346	.00268656	<.0001
Flow*Nocont	-0.00895	.00191691	<.0001
MUL	0.01928	.00109528	<.0001
Mul*flow	0.00071	.00020294	0.0004
Mul*refuse	0.00415	.00398354	0.2975
Mul*rf*fl	-0.00138	.00159796	0.3874
Mul*nocont	-0.01051	.00302069	0.0005
Mul*nc*fl	-0.00108	.00218590	0.6196

The household size was related to higher unemployment, higher unemployment in the second month (num*fl), but lower unemployment for nonresponse (num*stat). The effect was consistent for attrition and conversion (num*st*fl).

Table 13a: Household size

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.02788	.00114703	<.0001
Flow	0.00125	.00023366	<.0001
Nonresponse	0.02096	.00423471	<.0001
Flow*Status	-0.00747	.00188917	<.0001
NUM	0.01181	.00036749	<.0001
Num*flow	0.00046	.00006961	<.0001
Num*stat	-0.00371	.00137290	0.0068
Num*st*fl	0.00011	.00061455	0.8508

The lower unemployment effect for nonresponse above is probably due to refusal (num*refuse).

Table 13b: Household size

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.02618	.00116168	<.0001
Flow	0.00184	.00025413	<.0001
refuse	0.02068	.00560046	0.0002
Flow*Refuse	-0.00597	.00194701	0.0022
Nocontact	0.02266	.00513422	<.0001
Flow*Nocont	-0.00659	.00376001	0.0796
NUM	0.01192	.00037209	<.0001
Num*flow	0.00039	.00007398	<.0001
Num*refuse	-0.00583	.00171707	0.0007
Num*rf*fl	-0.00008	.00062914	0.8875
num*nocont	-0.00097	.00172212	0.5731
num*nc*fl	-0.00107	.00132671	0.4162

Household ownership was related to lower unemployment and lower unemployment the second month. There was a nonsignificant trend toward higher unemployment (adjusting for the other variables) in the interaction of the gross flow (own*st*fl), suggesting ownership may obscure a small amount of higher unemployment, although the attrition and conversion effects would reduce the impact on the estimates.

Table 14a: Home ownership

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.10491	.00128236	<.0001
Flow	0.00303	.00022895	<.0001
Nonresponse	-0.00055	.00342244	0.8721
Flow*Status	-0.00730	.00172868	<.0001
OWN	-0.05164	.00135133	<.0001
Own*flow	-0.00057	.00024934	0.0223
Own*status	0.00210	.00404016	0.6018
Own*st*fl	0.00344	.00192443	0.0738

The gross flow interaction above may come predominantly from refusal (own*rf*fl), although noncontact (own*nc*fl) contributes in the same direction.

Table 14b: Home ownership

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.10342	.00131558	<.0001
Flow	0.00491	.00026686	<.0001
refuse	-0.00331	.00492352	0.5009
Flow*Refuse	-0.00842	.00204610	<.0001
Nocontact	0.00311	.00407393	0.4441
Flow*Nocont	-0.01043	.00306007	0.0007
OWN	-0.05021	.00138428	<.0001
Own*flow	-0.00248	.00028544	<.0001
Own*refuse	0.00080	.00565275	0.8861
Own*rf*fl	0.00440	.00221184	0.0465
Own*nocont	0.00084	.00485451	0.8615
Own*nc*fl	0.00508	.00355227	0.1526

Relatives present in the household was related to higher unemployment, higher unemployment in the second month, but lower unemployment for nonresponse. No gross flow interaction effect was found.

Table 15a: Relatives present

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.03266	.00076069	<.0001
Flow	0.00160	.00016850	<.0001
Nonresponse	0.02017	.00283321	<.0001
Flow*Status	-0.00739	.00126520	<.0001
REL	0.02381	.00051130	<.0001
Rel*flow	0.00074	.00010957	<.0001
Rel*status	-0.00832	.00197185	<.0001
Rel*st*fl	0.00003	.00082692	0.9671

The lower unemployment effect above is probably due to both refusal and noncontact.

Table 15b: Relatives present

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.03038	.00075925	<.0001
Flow	0.00185	.00018875	<.0001
refuse	0.01345	.00393177	0.0006
Flow*Refuse	-0.00500	.00139899	0.0004
Nocontact	0.02903	.00347319	<.0001
Flow*Nocont	-0.00891	.00241710	0.0002
REL	0.02437	.00051706	<.0001
Rel*flow	0.00078	.00012004	<.0001
Rel*refuse	-0.00864	.00266942	0.0012
Rel*rf*fl	-0.00092	.00096519	0.3402
Rel*nocont	-0.00722	.00245754	0.0033
Rel*nc*fl	-0.00086	.00175947	0.6229

Rural location was related to lower unemployment and lower unemployment the second month.

Table 16a: Rural location

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.07120	.00068501	<.0001
Flow	0.00298	.00011661	<.0001
Nonresponse	0.00350	.00206421	0.0900
Flow*Status	-0.00670	.00093049	<.0001
RUR	-0.01369	.00116284	<.0001
Rur*flow	-0.00057	.00021117	0.0068
Rur*status	0.00641	.00447343	0.1516
Rur*st*fl	-0.00152	.00191099	0.4239

There is an interaction between “rural location”, “month-to-month flow”, and noncontact (rur*nc*fl). Since the interaction involving refusal is in the opposite direction, this may explain why the interaction above was not significant. The interaction would reduce the impact of nonresponse on the estimation of unemployment.

Table 16b: Rural location

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.07001	.00068480	<.0001
Flow	0.00340	.00012747	<.0001
refuse	-0.00235	.00272432	0.3878
Flow*Refuse	-0.00644	.00101308	<.0001
Nocontact	0.01035	.00262089	<.0001
Flow*Nocont	-0.00798	.00190799	<.0001
RUR	-0.01355	.00116566	<.0001
Rur*flow	-0.00064	.00023126	0.0053
Rur*refuse	0.00461	.00583998	0.4294
Rur*rf*fl	0.00075	.00197577	0.7035
Rur*nocont	0.00868	.00562605	0.1226
Rur*nc*fl	-0.00877	.00393821	0.0259

Population density (size) was related to higher unemployment and higher unemployment the second month.

Table 17a: Population density

Parameter	Estimate	Std.Error	Pr>|t|
Intercept	0.0518856	.00082010	<.0001
Flow	0.0020194	.00015469	<.0001
Nonresponse	0.0061225	.00321204	0.0567
Flow*Status	-0.0054161	.00135693	<.0001
SIZ	0.0013857	.00007002	<.0001
Siz*flow	0.0000684	.00001204	<.0001
Siz*status	-0.0001813	.00022699	0.4245
Siz*st*fl	-0.0001193	.00009750	0.2212

Refusal and noncontact had nonsignificant impact on the relationship between density and unemployment.

Table 16b: Population density

Parameter	Estimate	STD.Error	Pr>|t|
Intercept	0.05095	.00082075	<.0001
Flow	0.00233	.00017025	<.0001
refuse	0.00158	.00416182	0.7032
Flow*Refuse	-0.00398	.00144238	0.0058
Nocontact	0.01106	.00390885	0.0047
Flow*Nocont	-0.01246	.00290364	<.0001
SIZ	0.00136	.00007021	<.0001
Siz*flow	0.00007	.00001329	<.0001
Siz*refuse	-0.00029	.00029763	0.3250
Siz*rf*fl	-0.00018	.00010684	0.0840
Siz*nocont	-0.00003	.00027736	0.8983
Siz*nc*fl	0.00022	.00020021	0.2671

Discussion

Similar to the Tucker and Kojetin study, this study found small differences in the flow of labor force estimates depending on nonresponse. The impact of nonresponse on the final estimates is likely to be negligible. The opposite effects of conversion and attrition as well as a moderating effect for refusal and noncontact for some of the demographic groups would minimize the impact on estimates. This study replicated the small differences in labor force estimates related to nonresponse found earlier, with higher unemployment rates for attrition. This effect was moderated by race and month-in-sample, since attrition and conversion effects differed for the groups. Blacks showed a strong effect for refusal, and an effect for noncontact conversions, with all the effects showing lower unemployment. Tucker and Dixon (2000) found higher nonresponse for Blacks relative to Whites, so this might be related to the degree of the effect. Groves and Couper found a different relationship beween Black households and refusals using a census match study. Caution should be exercised in drawing inferences about the gross flows for the impact of nonresponse for Blacks in this study since those who never responded may be different from those who attrited or converted. The gross flows depend on the occasional responders to estimate the impact on estimates, so differences between studies using other methods are useful in gauging the generality of the findings. A census match study would be much more definitive. Month-in-sample showed a mixed effect, with attrition having its' largest effect in the 3rd and 4th months. Conversion effects were very small in the first to second months, probably because those who didn't respond the first month were more like those who responded the second month. Gender showed an effect for refusal versus noncontact, with a stronger effect for males for refusal. Teenage unemployment showed higher rates for noncontact and lower rates for refusal similar to the overall sample. The effects would tend to cancel one another, with little impact on the estimates.

The linear models showed no detectable gross flow effect for "number of attempted contacts", "Small children in the household", or "population size". The "number of attempted contacts" was related to the combination of flow and noncontact, suggesting that attrition and conversion balanced one another to produce little impact on the estimates. The households living in "multiunit structures" had higher unemployment, but it wasn't related to overall nonresponse. It was related to noncontact, which was similar to studies of nonresponse (Groves and Couper (1998), Tucker and Dixon (2000)). The noncontact effect suggested lower unemployment for those not contacted, but the nonsignificant effect for refusal and higher unemployment attenuates the impact.

Ownership was related to lower refusals and noncontact by Tucker and Dixon (2000). The present study found a slight trend toward higher unemployment for nonresponse adjusting for the flow effect. The overall impact collapsing across flow was not detectable. Attrition and conversion would cancel one another. The impact comes from refusal rather than noncontact. None of the other variables investigated in the linear models showed as strong an effect for canceling of effects. The reasons behind the effect of conversion of owners from refusal having higher unemployment might be of theoretical interest.

"Relatives present" and "Household size" showed more unemployment for the second month and lower unemployment for nonresponse. This suggests attrition may be a problem for these types of households, with refusal contributing for both, and noncontact contributing for "relatives present". Households involving family members would be less likely to participate, and may be placing barriers to contact, such as caller id and answering machines. In contrast, nonresponse involving households with unrelated members would be due to refusal.

"Rural location" had lower unemployment the second month, but no other effect. The finding was understandably reversed for "Population density". The attrition and conversion effects for noncontact cancelled out the effect for "rural location", but no effect for "population density". The direction of the effects due to refusal and noncontact were reverse, with higher unemployment for nonresponse in rural areas, but lower in higher density areas. While the effects weren't significant, the trends suggest examining interactions between other variables and measures of density or possible curvilinear effects would be useful. The effect of attrition due to noncontact for rural households related to lower unemployment estimates might be of interest.

The limitations of the study would include the assumption that the occasional responders represent the same relationship between response and labor force estimates as those households which never responded. Since the "occasional responders" are about the same size as the "never responders" the impact of the gross flows would on the unemployment ratio would be underestimated. There are also known problems with gross flows because they doesn't account for population growth or attrition. This study focussed on the later problem. The linear models may not have had enough power to detect small effects in some of the higher order interactions. Some of the coefficients were large enough to be of interest but with such high standard errors they proved statistically nonsignificant. The models also need to be examined for nonlinear effects. Other types of models, such as logistic regression, might find effects not possible with the linear models used here. More predictors and more complex models may also be needed. Tucker and Dixon (2000) found some interactions between predictors of nonresponse, so similar models my be useful in studying the impact of nonresponse on estimates.

These results suggest that strategies which attempt to reduce non-response bias might best be aimed at attrition, but the methods may need to vary by target group.

Additional methods may be useful in studying the relationships examined here. Census/CPS match data would provide a more complete picture of nonresponse. Examining other characteristics of nonrespondents based on other questions in the CPS (particularly supplements) could help find more useful segmentation stratagies for reducing bias due to nonresponse. Since separating nonresponse into refusals and noncontact showed different effects from the aggregate nonresponse, other characteristics of nonresponse might be helpful in understanding its' potential impact on estimates. The type of noncontact; phone machine, no one home, or other barriers to contact, may be useful to study. The employment status may also be modeled by the characteristics of the refusal, since the motivations and fears of the respondents which produce refusal may be related to their employment status.

References:

- Agresti, A., Categorical Data Analysis, Wiley, New York, 1990.
- Barkume, A.J., and Horvath, F.W., "Using gross flows to explore movements in the labor force", Monthly Labor Review, April, 1995.
- Dolton, P., Lindeboom, M., and Van den Berg, G.J., "Survey attrition: A taxonomy and the search for valid instruments to correct for biases", in Statistical Policy Working Paper 30, 1999 Federal Committee on Statistical Methodology Research Conference.
- Fletcher, J., and Schmidt, D., "Measuring Response Bias in Survey Research: An Analysis of Age Characteristics of Early Respondents and Resistors", Paper presented at AAPOR, 2001.
- Groves, R., and Couper, M., Nonresponse in Household Interview Surveys, Wiley, New York, 1998.
- Tucker, C, and Dixon, J., "Predicting Interviewer Nonresponse Rates from Household and Regional Characterstics", Paper presented at AAPOR, 2000.
- Tucker, C., and Kojetin, B., "The Impact of Nonresponse on the Unemployment Rate in the Current Population Survey", Paper presented at the International Workshop on Household Survey Nonresponse, 1997.