Reconciling LFPR and income measures

Madison informal notes

June 22, 2023

Problem: Shigeru's computation from March showed the following labor force participation rates by income for groups defined by their **household** income. My computations showing LFPR by income for groups defined by **family** income have pretty different values for the $25,000< group.

Point 1: Familes are different from households. Households may be single-family or multi-family. Everyone with a household income under $25k also has a family income under $25k because in the CPS, families are defined as being a level below household. Household income is equivalent to the sum of the family incomes for each family in the household. This allows for the existence of cases where two families each have family incomes below 25k which sum to a household income greater than 25k.

Consider the universe of individuals with family income less than $25,000. Each person falls into one of four types:

1. In single-family household, family income < $25,000, household income < $25,000

2. In single-family household, family income < $25,000, household income $\geq$ $25,000

3. In multi-family household, family income < $25,000, household income < $25,000

4. In multi-family household, family income < $25,000, household income $\geq$ $25,000

Household income being the sum of all family incomes for each family in a household should mean that there are no instances of the second type. However, this is not true in our data.

Table 1: Family income < $25,000 Universe

Persons	Year		
	2020	2021	2022
Single-family, hh<25k	12,437	14,024	12,711
Multi-family, hh<25k	1,068	1,286	1,128
Single-family, hh$\geq$25k	293	185	236
Multi-family, hh$\geq$25k	3,365	3,671	3,240
Total	17,163	19,166	17,315

Table 2: Family income < $25,000

Persons	Year		
	2020	2021	2022
Single-family	12,730	14,209	12,947
Multi-family	4,433	4,957	4,368
Total	17,163	19,166	17,315

Notes: Sample is all participants 16+ from the March Supplements with family income (FTOTVAL) less than $25,000

Table 3: Household income < $25,000

Persons	Year		
	2020	2021	2022
Single-family	12,883	14,555	13,170
Multi-family	1,216	1,488	1,321
Total	14,099	16,043	14,491

Notes: Sample is all participants 16+ from the March Supplements with household income (HTOTVAL) less than $25,000

My computations of labor force participation were based on family income, so my below-25k group contains all three cases. Shigeru's computations of labor force participation were based on household income so his below-25k group contains people from only the first two cases. I think his group is on-average poorer than mine. I verified this by computing the average family income and average HH income for my below-25k group and his below-25k group, but my point is simply that since poorer people have lower labor force participation, it may make sense for his group to have a lower average labor force participation than mine.

Notice that the number of people in single family households (SFH) with HH income less than 25k **should** be less than the number of people in SFHs with family income less than 25k because one is a subset of the other. However, these counts show that the number of people in SFH with HH income less than 25k is greater than the number of people in SFH with family income less than 25k in each year. This is anomalous. I found that each year, there are small group of people (around 600 total, 5:1 ratio of SFH to MFHs) who have htotval less than 25k but a missing value for ftotval, causing the counts for people in SFH with family incomes less than 25k to be understated by about 500 people and MFH by about 100 people. If you add 500 to the counts for people in SFH with **family** income less than 25k, they are greater than the counts for people in SFH with **HH** income less than 25k, as we might expect.

How many people family incomes under $25k and HH incomes **above** $25k? And are they truly all in multi-family households as we thought? To answer this, we can count the people with FTOTVAL <$25k and HTOTVAL≥$25k and tabulate them by single/multi-family status:

Table 4: Frequency of people with FTOTVAL<$25k and HTOTVAL≥$25k

Persons	Year		
	2020	2021	2022
Single-family	293	185	236
Multi-family	3,365	3,671	3,240
Total	3,658	3,856	3,476

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k and HTOTVAL≥$25k

We see the majority of cases of scenario 3 are people living in multifamily households, as expected.

Table 5: Mean LFPR among people with FTOTVAL<$25k (Madison's)

Persons	Year		
	2020	2021	2022
Single-family	29.3	26.9	26.6
Multi-family	56.8	53.6	56.2
Total	36.6	33.8	34.3

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k. Weights used are marsupwt.

My LFPRs in the 'total' row are slightly different from the June 15 document I sent Shigeru. This is because I mistakenly did not use weights when I made those tabulations. All tabulations done in this document use the MARSUPWT set of weights. I did check that when I remove weights, these LFPRs for people with FTOTVAL<$25k are identical to those in my June 15 document.

Table 6: Mean LFPR among people with HTOTVAL<$25k (similar to Shigeru's calculation)

Persons	Year		
	2020	2021	2022
Single-family	28.4	26.3	25.9
Multi-family	38.3	37.3	43.7
Total	29.2	27.3	27.4

Notes: Sample is all participants 16+ from the March Supplements with HTOTVAL<$25k. Weights used are marsupwt.

Notice that the overall LFPR for the above group is not exactly as Shigeru had calculated in March. We will see why that is in a moment.

Table 7: Mean LFPR among people with FTOTVAL<$25k and HTOTVAL≥$25k

Persons	Year		
	2020	2021	2022
Total	63.4	59.9	61.3

Notice also how much higher the LFPR is for the group that is present in my calculations but not Shigeru's. It makes sense that my number was higher. I will confirm this at the end by showing a sum of weighted averages.

Point 2: Shigeru's household income measure was not simply the one given by CPS (HTOTVAL). Constructing variable h_tot_income_ASEC_calc, he showed exactly which income line items are encompassed in htotval (see memo from March). To that sum, he added in energy assistance, food stamps, and COVID economic impact payments to create variable h_tot_income_calc. He defines the income groups using *this* measure of household income. This measure takes larger values than HTOTVAL, especially in years 2021 and 2022 because that's when the EIPs are being paid out. This means that there will be more people in those years who fall into the case of family income < 25k and HH income > 25k when we use these numbers for HH income.

This is shown here:

Table 8: Frequency of people with FTOTVAL<$25k and h_tot_income_calc ≥$25k

	Year		
Persons	2020	2021	2022
Total	4,085	7,146	6,488

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k and h_tot_income_calc≥$25k

Table 9: Mean LFPR among people with h_tot_income_ASEC_calc <$25k (Shigeru's calculation)

	Year		
Persons	2020	2021	2022
Total	28.7	25.0	24.2

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k and h_tot_income_calc≥$25k

Now I will show what proportion each case constitutes of the overall subsample (family income<25k):

Table 10: Proportion of subsample

	Year		
Persons	2020	2021	2022
HH income < 25k	0.76	0.63	0.63
HH income ≥ 25k	0.24	0.37	0.37
Together	1.00	1.00	1.00

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k

And here are the LFPRs for those respective cases:

Table 11: Mean LFPR by group

Persons	Year		
	2020	2021	2022
HH income < 25k	28.7	25.0	24.2
HH income ≥ 25k	61.6	48.9	51.5
Together	36.6	33.8	34.3

Notes: Sample is all participants 16+ from the March Supplements with FTOTVAL<$25k

Multiplying each LFPR by its respective group's proportion and then summing yields the same values as in the 'Together' row. These are also the same as appeared in the total row in Table 4.

I am aware that the average LFPRs for the HH<25k group are not exactly the same as Shigeru's March calculations. They are off by tenths of a percentage point. This partly is because I condition on having family income nonmissing and less than 25k. When I include people with missing value for family income, the LFPRs for HH<25k group get even closer to Shigeru's original March calculations. I cannot figure out why they are not identical. I also don't have Shigeru's stata file for making that exact table, which would probably be helpful.