Labor force participation by imputed family income

Madison informal notes

June 1, 2023

1 LFPR

Previously the LFPRs for predicted lowest income group were about 10 pps higher than for the actual lowest income group, when we estimate the imputation model over the whole sample. We observed that the model was throwing some people with actual family incomes under 25k (many of them NILF) into much higher predicted income categories, giving us a higher LFPR for the lowest predicted category. But that was when I was running the model over the **whole** ASEC sample by mistake and not merely the 16+ people. That's my bad. Here are the correct LFPRs by income category. The sample is all ASEC participants age 16+ with no other restrictions.

Table 1: LFPRs by actual family income, ASEC (% as of March)

	Year		
Income group ($)	2020	2021	2022
− 24,999	35.9	34.1	33.8
25,000 − 49,999	53.6	53.3	52.7
50,000 − 99,999	65.5	65.4	65.5
100,000 − 149,999	73.3	72.8	72.8
150,000 −	76.2	76.9	77.2
Overall (ASEC, 16+)	62.7	62.0	62.5
Overall (Official NSA)	62.6	61.5	62.4

Notes: Sample is all participants 16+ from the March Supplements

Table 2: LFPRs by IMPUTED family income, ASEC (% as of March)

	Year		
Income group ($)	2020	2021	2022
− 24,999	38.6	40.9	39.9
25,000 − 49,999	56.0	55.7	56.1
50,000 − 99,999	63.0	64.8	64.5
100,000 − 149,999	72.6	74.8	74.7
150,000 −	75.6	78.5	79.1
Overall (ASEC, 16+)	62.7	62.0	62.4
Overall (Official NSA)	62.6	61.5	62.4

Notes: Sample is all participants 16+ from the March Supplements. Income groups are defined using the predicted values of ftotval for the same observations that appear in the ASEC.

As you can see, the gap is a good deal narrower when I actually run the imputation model using the correct sample. Now, let's see if the gap between the two sets of rates closes at all if we do some sample selection. We want to exclude people in households that have little attachment to the labor market anyway. This just also has the added benefit of potentially getting rid of some of the people that the model was mis-assigning.

1.1 Exclude 2-person retiree households

Table 3: LFPR by actual family income category

	Year		
Income group ($)	2020	2021	2022
− 24,999	50.7	49.5	48.7
25,000 − 49,999	68.8	68.9	67.7
50,000 − 99,999	76.5	76.6	76.9
100,000 − 149,999	81.2	80.9	81.1
150,000 −	82.5	82.7	82.9
Overall	73.8	73.3	73.6

Notes: Sample is ASEC participants 16+, excluding 2-person households in which both people are retired.

Table 4: LFPR by predicted family income category

	Year		
Income group ($)	2020	2021	2022
− 24,999	51.9	52.2	51.3
25,000 − 49,999	69.7	69.2	69.7
50,000 − 99,999	76.4	76.1	76.0
100,000 − 149,999	80.6	81.3	81.1
150,000 −	79.8	79.5	80.1
Overall	73.8	73.3	73.6

Notes: Sample is ASEC participants 16+, excluding 2-person households in which both people are retired.

1.2 Exclude 1 and 2 person households wherein all members are retired

	year			
	2020	2021	2022	Total
Actual				
1	59.8	58.0	57.7	58.5
2	73.0	73.2	72.6	73.0
3	77.8	78.0	78.4	78.0
4	81.6	81.4	81.7	81.6
5	82.7	82.9	83.2	82.9
Total	76.7	76.3	76.7	76.5

	year			
	2020	2021	2022	Total
Predicted				
1	64.7	63.5	63.0	63.7
2	74.0	73.7	74.7	74.1
3	77.2	77.2	77.2	77.2
4	80.7	81.4	81.4	81.2
5	79.7	79.4	80.1	79.7
Total	76.7	76.3	76.7	76.5

1.3 Exclude people 22 and younger who are childless and have NILF activity listed as "In School"

	year			
	2020	2021	2022	Total
Actual				
1	47.6	46.5	45.7	46.6
2	62.5	62.2	61.2	62.0
3	71.8	71.8	71.9	71.9
4	78.0	77.9	77.7	77.9
5	80.3	80.9	81.0	80.7
Total	69.7	69.3	69.5	69.5

	year			
	2020	2021	2022	Total
Predicted				
1	50.0	49.9	49.0	49.7
2	64.0	63.3	63.7	63.6
3	70.4	70.4	70.1	70.3
4	77.8	78.6	77.7	78.0
5	78.9	78.8	79.3	79.0
Total	69.7	69.3	69.5	69.5

1.4 Exclude 2-person retiree households and young student non-parents

	year			
	2020	2021	2022	Total
Actual				
1	50.9	49.7	48.9	49.8
2	68.9	69.0	67.8	68.6
3	76.5	76.6	76.9	76.6
4	81.2	80.9	81.1	81.1
5	82.5	82.7	82.9	82.7
Total	73.9	73.4	73.7	73.6

	year			
	2020	2021	2022	Total
Predicted				
1	52.2	52.4	51.5	52.0
2	69.7	69.2	69.6	69.5
3	76.4	76.2	76.0	76.2
4	80.6	81.3	81.1	81.0
5	79.8	79.5	80.1	79.8
Total	73.9	73.4	73.7	73.6

2 RMSE

Since we improved the imputation model slightly by adding additional predictor variables, I re-computed the root mean squared error for the model.

Madison question: Does RMSE change when you estimate the model over a selected sample?

Table 5: RMSEs for updated log-linear imputation model, by variable

	Year		
Variable	2019	2020	2021, 2022
FTOTVAL	68,125	69,927	71,593
FAM_AGI	67,603	72,434	76,062
FEARNVAL	62,087	61,905	63,979
HTOTVAL	69,367	71,665	72,869