Projections of Sample Sizes for Offspring exam 11 / Omni1 exam 6

For brevity, throughout this document I use "exam 11" to denote the combined exam 11 / exam 6.

1. Projecting survivorship to start of exam cycle
   - Select people alive (01,01,2020)
   - Apply lifetables US whites 2009 by sex, single years of age
   - Estimate survival from last contact to start of exam, 10-01-2025
   - Probability is l(b) / l(a), where l(a) is age at last contact, l(b) is age at exam start
   - Both l(a) and l(b) are interpolated linearly between integer ages, e.g., floor(z) and ceiling(z)
   - Program is projection_survival11.sas
   - Estimated number surviving to start of exam 11 is 1817

     Sex breakdown is 731 men, 1086 women (rounded)

2. Projecting participation conditional on living to start of exam
   2.1 Participation at recent exams (9, 10)
   - First, tabulate actual participation among people who were alive at start of those exam cycles

   2.2 Model participation using cohort, sex, age at start of exam, residence (NE/NY vs other), participation at prior exams (back 1 and back 2 cycles)
   - Model 9 fits participation at exam 9, including variables above
     o Key variables were age at start of exam cycle (and age-70)^2, sex, cohort, residence in New England + New York vs other, indicators for participation at each of exams 7 and 8

   - Model 10 fits participation at exam 10, including variables above
     o Key variables were age at start of exam cycle, cohort and residence in New England + New York vs other, indicators for participation at each of exams 8 and 9
   - Another model fits in-person participation at exam 10
   - Program is attend9_attend10.sas

3. Projecting combined survival and participation (i.e., unconditional participation probability)
   - There are several options to consider regarding which participation patterns to apply, because exam 10 was disrupted by Covid-19 such that participation was low and perhaps does not represent what to expect in the future
   - In all participation projections, I used age at the start of exam 11
   - First, I used participation Model 9
     o Fit9a: completely disregard exam 10, plugging in participation at exams 8 and 9 as predictors of participation at exam 11
     o Fit9b: use participation at exams 9 and 10 as the predictors
     o Fit9x: do not penalize failure to participate at exam 10; use exams 9 and 10 if a person participated in exam 10, but use exams 8 and 9 when they did not attend exam 10
   -

Participation – fitted model	Variable name	Projected number

Fit9a (use ex 8,9 participation)	psurv11_att11_fit9a	1185	
Fit9b (use ex 9,10 participation)	psurv11_att11_fit9b	1058	
Fit9x (use 10 when beneficial)	psurv11_att11_fit9x	1148	

Participation – fitted model	Variable name	Projected number
Fit10a (use ex 8,9 participation)	psurv11_att11_fit10a	1062
Fit10b (use ex 9,10 participation)	psurv11_att11_fit10b	989
Fit10x (use 10 when beneficial)	psurv11_att11_fit10x	1041

Program is projection2019_combined_survatt11.sas

Exported SAS database has variables idtype, id, sex, age (integer) at start of exam 11 (10/01/2025), p_surv11 = probability of survival to start of exam 11, p_att11 = probability of attendance conditional on survival, and p_surv11_att11 = p_surv11 * p_att11 = probability of survival and attendance at exam 11.

Breakdowns of surviving attendees (weighted by p_surv11_att11) are shown below, rounded to integers.

Cohort	Count
1=Offsp	1011
7=Omni1	174

Female	Count
0	493
1	693

NE/NY	Count
0	169
1	1017

Age11_group	Sum
55	3
60	14
65	64
70	203
75	334
80	310
85	175
90	68
95	13
100	1

Age11_group	Men	Women
55	2	1
60	7	7
65	31	33
70	88	115
75	132	202
80	125	185
85	79	96
90	25	43
95	3	10
100	0	1
All Ages	492	693