Chapter 5

Building Alternative Scenarios

Abstract New dimensions added in the projection models in Chap. 4, the labour force participation and the sector of activity, are modelled using regression parameters. In this chapter, we show how building alternative scenarios with this type of inputs. In the first example, we test an assumption in which women with young children at home have the same participation rate as other women. The second example, we test a scenario in which labour force participation rates of women gradually increase and reach those of men by 2060. The code to implement those alternative scenarios is explained, and some results are presented.

Keywords Microsimulation · Population projection · Demography · Method · SAS

5.1 Building Alternative Scenarios from Regression Parameters

At this point, we have replicated a reference scenario for a multistate/multiregional population projection for India and added two dimensions, the labour force participation and the sector of activity. For fertility, mortality, migration, and education, the input files already include rates for each projected year. Therefore, alternative scenarios can be built simply by changing the value of the inputs.

The parameters for the labour force participation and the sector of activity are however regression parameters. The sector of activity has a “cohort” parameter that allows change over time, but labour force participation does not. Assuming no change in parameters means that change in labour force participation would only result from changes in the composition of the population by age, sex, education, region and fertility (as a reminder, we have a parameter for women who had a kid in the past

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-79111-7_5) contains supplementary material, which is available to authorized users.

5 years). As the model is constructed at this point, changing a parameter in the input file would generate an immediate change starting from 2015, with no further evolution for the rest of the projection.

In this chapter, we will show two examples of alternative scenarios. In the first one, we test a scenario in which women with a young child at home would have the same participation rate and probability of working in the formal sector as other women, all other things being equal. In the second one, we test a scenario in which the labour force participation rates of women gradually increase and reach those of men in 2060. Obviously, none of these scenarios have a predictive purpose. They should be interpreted as “what if” scenarios that aim to measure how the model reacts to change in its different components and provide some policy-relevant information about the socioeconomic dynamics of the country.

5.2 Example 1: The Impact of Having a Young Child on Labour Force Participation and the Sector of Activity

For the first example of an alternative scenario, we will test one in which women with a young child at home have the same participation rate and the same chance to work in the formal sector as other women, all other things being equal. When creating an alternative scenario, the first thing to do is to copy and paste the folder where codes, parameters, and outputs are located. Otherwise, files will be replaced. In the example presented above, the folder was called “Chap. 4”. We copy and paste this folder and change the name to “In this chapter_YoungChild”.

This alternative scenario is easy to create and doesn’t require any major change in the codes. We can just switch to 0 the two parameters for young_kid in parameters files `lfp.csv` and `formal.csv` that are applied to women with a young kid, as highlighted in yellow in Fig. 5.1.

We then change the name of the scenario in the `%let` statement. This will automatically change the name of the scenario folder everywhere in the codes.

```
%let scenario_name= Chapter5_YoungChild;
```

We can then run the scenario and explore the results in the file `outputTotal.csv`. In Fig. 5.2, we compare the projected labour force size for women from this scenario to the Chap. 4 scenario. For the purpose of this analysis, we will consider Chap. 4 scenario as the “reference” scenario, since it is the business as usual scenario with constant parameters. In 2060, the scenario YoungChild manages to increase the number of female workers by 6%, by switching 7 million housewives into workers. This scenario also increases the number of women working in the formal sector by 6 M, an increase of 10% compared to the reference scenario.

Indeed, as shown in Fig. 5.3, the scenario YoungChild manages to improve the participation rate as well as the proportion of workers in the formal sector by several

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	sex	edu	young_kicregion	intercept	edu_p	agegr_p	agegr2_p	agegr_ed	agegr2_ec	young_kic	young_kicregion_p		
2		1		-4.4198							0		
3		1	0								0		
4		1	1								0		
5		1 e1			0.3929								
6		1 e2			0.9207								
7		1 e3			0								
8		1 e4			-1.5251								
9		1 e5			-2.3395								
10		1 e6			1.0876								
11		1				0.1847							
12		1					-0.00248						
13		1 e1						0.000397					
14		1 e2						-0.0266					
15		1 e3						0					
16		1 e4						0.0685					
17		1 e5						0.1092					
18		1 e6						-0.00883					
19		1 e1							7.27E-06				
20		1 e2							0.000251				
21		1 e3							0				
22		1 e4							-0.00087				
23		1 e5							-0.00133				
24		1 e6							7.98E-06				
25		1 e1	0									0	
26		1 e2	0									0	
27		1 e3	0									0	
28		1 e4	0									0	
29		1 e5	0									0	
30		1 e6	0									0	
31		1 e1	1									0	
32		1 e2	1									0	
33		1 e3	1									0	
34		1 e4	1									0	
35		1 e5	1									0	
36		1 e6	1									0	

Fig. 5.1 Screenshot of the parameter file lfp.csv for the YoungChild scenario (opened with Excel)

Fig. 5.2 Projected female labour force size in India, 2010–2060, Reference and YoungChild scenarios

Fig. 5.3 Projected labour force participation rate and proportion of workers in the formal sector by age in 2060, women, India, Reference and YoungChild scenarios

percent for the age groups 20 to 34, which are the main age groups that parameters for having a young kid are concerned with. For the labour force participation, this gain is significant, but rates are still way below those of men. This suggests that while having a young kid at home negatively affects the participation rates of women, it is only a minor factor for explaining their lower rates, and policies that seek to change this might focus on other issues.

5.3 Example 2: Gender Equality in Labour Force Participation

In this second example of an alternative scenario (GenderEquality), we test a scenario in which labour force participation rates of women gradually increase and reach those of men (with the same age/education/region) by 2060. This scenario will show the impact on the labour force of an efficient policy reducing gender equalities and empowering women. For this kind of scenario, where changes occur gradually over time, we cannot directly change the parameter for women in the input parameters file `lfp.csv` because this would generate an immediate change in the labour force participation.

We will show in this section how to adapt the model to make changes in parameters gradually over a period of years. We will build a scenario in which parameters for women reach those of men by 2060 (though keeping the negative parameters for having a young child). As there are no big gaps between men and women for the proportion of workers in the formal sector, we will not touch this module.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	sex	edu	agegr	young_kid	region	intercept	edu_p	agegr_p	agegr_edu_p	young_kid_p	young_kid_edu_p	region_p	intercept2060	edu_p2060	agegr_p2060	agegr_edu_p2060	young_kid_p2060	young_kid_edu_p2060	region_p2060
2	1					-0.9037							3.9222						
3	1									0							0		
4	1			1						-0.4282							-0.4282		
5	1	e1				0.3907									-0.5112				
6	1	e2				0.1768									-0.2577				
7	1	e3				0				0					0				
8	1	e4				-0.443									0.3855				
9	1	e5				-0.3102									0.9575				
10	1	e6				0.5389									0.3696				
11	1	e1		0						0								0	
12	1	e2		0						0								0	
13	1	e3		0						0								0	
14	1	e4		0						0								0	
15	1	e5		0						0								0	
16	1	e6		0						0								0	
17	1	e1		1						0								0	
18	1	e2		1						0								0	
19	1	e3		1						0								0	
20	1	e4		1						0								0	
21	1	e5		1						0								0	
22	1	e6		1						-0.2921								-0.2921	
23	1		15							-1.2557					-3.9989				
24	1		20							-1.0963					-1.1446				
25	1		25							-0.2771					-0.3098				
26	1		30							-0.0384					0.4394				
27	1		35							0					0				
28	1		40							-0.1096					-0.0042				
29	1		45							-0.303					-0.4817				
30	1		50							-0.7808					-1.3789				
31	1		55							-0.8383					-1.5168				
32	1		60							-1.2301					-2.9482				
33	1		65							-1.6653					-4.0251				
34	1		70							-2.1789					-4.5075				
35	1	e1	15							0.0035						1.2222			

Fig. 5.4 Screenshot of the parameter file lfp.csv for the scenario GenderEquality (opened with Excel)

Since this is a new scenario, we again need to copy and paste the model and its assumptions for reference (Chap. 4) into a new folder (we label it “In this chapter_GenderEquality”). In the code, we then change the name of the scenario accordingly:

```
%let scenario_name= Chapter5_GenderEquality;
```

We then need to reshape the input file for labour force participation. For women, we have initial parameters for 2015, and we want those parameters to gradually evolve until they reach those of men in 2060. The parameter file thus needs to have two columns for each set of parameters: one for the initial value, one for the value of 2060. In Fig. 5.4, we show how the file should look. Columns A to E are variable names and their categories. Columns F to L are their corresponding initial parameters. Columns M to S (ending in “2060”) are their corresponding parameters in 2060, which are a copy of the men’s parameters (plus the parameters for the variable young_kid and its interaction with education).

We merge this parameter file to the population file in the labour force module in a way similar to what we did previously, with SQL codes. From the code for Chapter 4, we simply add for each variable of the model the name of parameters in 2060 from the input file, as highlighted in yellow in the code below.

```

/*Labor force participation*/
/*Implementing parameters*/
proc sql;
  create table pop_lfp1 as
  select
    p.*,
    t1.intercept, t1.intercept2060, t2.edu_p, t2.edu_p2060,
    t3.agegr_p, t3.agegr_p2060, t4.agegr2_p, t4.agegr2_p2060, t5.agegr_edu_p,
    t5.agegr_edu_p2060, t6.agegr2_edu_p, t6.agegr2_edu_p2060, t7.young_kid_p,
    t7.young_kid_p2060, t8.young_kid_edu_p, t8.young_kid_edu_p2060, t9.region_p
    , t9.region_p2060
  from
    pop2 p
  left join
    ( select sex, intercept, intercept2060
      from param.lfp
      where not missing(intercept)
    ) t1
  on p.sex=t1.sex
  left join
    ( select sex, edu, edu_p, edu_p2060
      from param.lfp
      where not missing(edu_p)
    ) t2
  on p.sex=t2.sex and p.edu=t2.edu
  left join
    ( select sex, agegr_p, agegr_p2060
      from param.lfp
      where not missing(agegr_p)
    ) t3
  on p.sex=t3.sex
  left join
    ( select sex, agegr2_p, agegr2_p2060
      from param.lfp
      where not missing(agegr2_p)
    ) t4
  on p.sex=t4.sex
  left join
    ( select sex, edu, agegr_edu_p, agegr_edu_p2060
      from param.lfp
      where not missing(agegr_edu_p)
    ) t5
  on p.sex=t5.sex and p.edu=t5.edu
  left join
    ( select sex, edu, agegr2_edu_p, agegr2_edu_p2060
      from param.lfp
      where not missing(agegr2_edu_p)
    ) t6
  on p.sex=t6.sex and p.edu=t6.edu
  left join
    ( select sex, young_kid, young_kid_p, young_kid_p2060
      from param.lfp
      where not missing(young_kid_p)
    ) t7
  on p.sex=t7.sex and p.young_kid=t7.young_kid
  left join
    ( select sex, edu, young_kid, young_kid_edu_p, young_kid_edu_p2060
      from param.lfp
      where not missing(young_kid_edu_p)
    ) t8
  on p.sex=t8.sex and p.edu=t8.edu and p.young_kid=t8.young_kid
  left join
    ( select sex, region, region_p, region_p2060
      from param.lfp
      where not missing(region_p)
    ) t9
  on p.sex=t9.sex and p.region=t9.region;
quit;

```

In the resulting population file `pop_lfp1`, each individual has logit parameters for the year 2010 and 2060. We now need to adjust the calculation of the labour force participation rate in the section simulating the labour force participation event in the temporary population file `pop_lfp2` for the population aged 15–74.

```
/*Labour force participation event*/
data work.pop_lfp2;
set work.pop_lfp1;
labour=0;
if 15<=agegr<74 then do;
  (...)
```

In the temporary variables `lab2010` and `lab2060`, we sum up those parameters for the initial (2010) and final (2060) years, respectively.

```
(...)
lab2010 = intercept + agegr_p*agegr + agegr2_p*agegr*agegr + edu_p +
agegr_edu_p*agegr + agegr2_edu_p*agegr*agegr
+ region_p + young_kid_p + young_kid_edu_p;
lab2060 = intercept2060 + agegr_p2060*agegr + agegr2_p2060*agegr*agegr +
edu_p2060 + agegr_edu_p2060*agegr + agegr2_edu_p2060*agegr*agegr
+ region_p2060 + young_kid_p2060 + young_kid_edu_p2060;
(...)
```

Then, in another temporary variable `lab`, we interpolate the sum of logit parameters for the current period (`endyr`). From this, we calculate the corresponding probability of participating in the labour force (`prob_lab`) and simulate the event with a random experiment.

```
(...)
lab = lab2010 + (%endyr-2010)*(lab2060 - lab2010) / (2060-2010);
prob_lab=exp(lab)/(1+exp(lab));
if rand('uniform')<prob_lab then labour=1;
end;
(...)
```

Finally, we drop parameters and temporary variables.

```
(...)
drop intercept agegr_p agegr2_p edu_p agegr_edu_p agegr2_edu_p region_p young_kid_p
young_kid_edu_p lab prob_lab lab2010 lab2060
intercept2060 agegr_p2060 agegr2_p2060 edu_p2060 agegr_edu_p2060 agegr2_edu_p2060
region_p2060 young_kid_p2060 young_kid_edu_p2060;
run;
```

No further changes are required to run this scenario. The interpolation implemented directly in the code of the model allows for a smooth and gradual increase of the labour force participation rates of women, as shown in Fig. 5.5 which has been built from the file `outputTotal.csv`.

Results of this scenario would give approximatively the same population size by age, sex, education and region as the reference scenario in Chap. 4, but very different outcomes in terms of labour force size, as shown in Fig. 5.6. The size of the workforce is thus about 60% higher in the GenderEquality scenario than in the baseline scenario (1.0G vs. 0.6G), and therefore, the country reaps a much greater benefit from the demographic dividend, as suggested by the labour force dependency ratio which, by 2060 in the GenderEquality scenario, is less than half of what it was in the reference scenario (0.79 versus 1.86). This outcome highlights the high stakes of including labour force participation and its sources of heterogeneity in population projections.

Fig. 5.5 Projected labour force participation rates by age for women, India, 2010–2060, GenderEquality scenario

Fig. 5.6 Projected population according to the labour force status, India, 2010–2060, Reference and GenderEquality scenarios

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.