DRAFT – Not to be cited

Distr.: General
30 March 2016

English only

Economic Commission for Europe

Conference of European Statisticians

Joint Eurostat/UNECE Work Session on Demographic Projections
Geneva, 18-20 April 2016
Item 3 of the provisional agenda
Sub-national projections

Projecting the regional explicit socioeconomic heterogeneity in India

Prepared by Wittgenstein Centre (IIASA, VID/ÖAW, WU), Austria[1]

Summary

We developed a multi-regional multistate population projection model for India that can simultaneously simulate population heterogeneity in the demographic (age-sex), socio-economic (educational attainment), and spatial (States/UT and rural/urban) dimensions. This is a first model of its kind for India. We applied the model to study the impact of considering socio-economic and spatial heterogeneity at the sub-national level on the overall population prospects of India.

We populated the model with data and estimates and defined a baseline scenario by conducting in-depth analysis of data from Census and Sample Registration System (SRS) on five components namely: fertility by education, mortality, internal migration, education, and rural reclassification. We ran the projection and presented some preliminary results.

The projection resulted in slower rate of urbanization in India from 31% in 2011 to 41% in 2051 compared to the UN projections because of largely rural, less educated, large population in Bihar and Uttar Pradesh that are slowing down the momentum of urbanization in India. The same reason applies for explaining higher rate of population growth in India resulting from the bottom-up approach than when the projections are done for two regions (rural and urban) of India.

We plan to work further on issues identified in this iteration of the baseline scenario. In addition, we will conduct several sensitivity tests and define alternative scenarios with relevant narratives for India (for e.g. SDG scenario) as well as for the Shared Socio-economic pathways.

[1] [Prepared by Samir KC, Research Scholar, Wittgenstein Centre (IIASA, VID/ÖAW, WU), Austria and Professor, Asian Demographic Research Institute, Shanghai University, Shanghai, China; and Markus Speringer, Research Assistant, Wittgenstein Centre (IIASA, VID/ÖAW, WU), Austria and Professor]

I. Introduction

1. We are interested to explore how population size and structure will evolve within a region under different socio-economic scenario. Traditionally, past trends of demographic rates are extrapolated into the future, often along the path traversed by forerunners, and when needed experts inputs are included. While trend extrapolation is seen as a neutral way to speculate about the future, often determinants – proximate or distal- of the variable of interest are already known specially the ones with lagged effect and should not be ignored as the inclusion of such knowledge will enhance the quality of the scenario and the projection itself. One example of such a determinant is educational attainment of women that could determine the number of children they might have in the future, educational level of the population is also positively associated with healthy and longer life. Other example is a classic case of smoking, prevalence level in a cohort is linked to the remaining life expectancy. These determinants could be also related to new policies or new innovation (technological or social).

2. The two fundamental sources of heterogeneity in demographic rates are age and sex of individuals. More, recently, educational attainment has been added as the third most important source of population heterogeneity. There are many other individual level sources of heterogeneity (labor force participation, income level, etc.) related to 'who you are', in addition, the source of heterogeneity about 'where you live' is important. The place of residence, the community, the environment, the law of the land, etc. influences ones decisions either by choice or by force.

3. In this paper, we developed demographic projections of Indian population, disaggregated by age, sex, educational attainment, rural/urban residence, and by 35 states, as the first exploration study. Projection results show that accounting for spatial heterogeneity leads to faster population growth in India than when the projections are simply done for rural and urban India. The significant difference was attributed to a very high level of socio-demographic heterogeneity across states. It also reveals that what matters for the prospects of Indian population is not only where they reside but also who they are (particularly by educational attainment).

4. The UN estimates population of India to be more than 1.3 billion by 2015 i.e., 1 in 6 person in this World is living in India (United Nations 2015). About 31% of the population (as reported in 2011 Census) were rural residents and we want to know how these two populations living in rural and urban regions will evolve in the future.

5. Often employed method to estimate the future population by rural and urban is by extrapolating the urbanization rate. The UNPD's urbanization projection was based on extrapolating the current urbanization trend into the path of traversed by other forerunner countries in the past. The model to a large extent does not consider the population size or

structure of the country. The UNPD repeatedly overestimated the pace of growth in the urban population in India. Others, like NCAR, extended and modified the method used by the UN World Urbanization Prospects (United Nations 2014), allowing different speed of urbanization (low, medium and high) for countries with a population greater than 1 million and a land area greater than 1000 km^2. Often these projected urbanization rates are not translated into age and sex structure leaving room for users to make their own guesses and apply various proportional distribution methods to the national population projection.

6. India is the second largest country by population with 35 federal States/UT that has constitutional authority to make decision on various policies directly or indirectly affecting population. Our exploration of the Indian demographic data revealed a very high level of heterogeneity in population size, structure, rates of mortality, fertility, and migration (between States) as well as the levels of human capital varied greatly by States/UT and place of residence. Our preliminary conclusion was that by not considering spatial heterogeneity (by States/UT) we might lose important demographic evolutions occurring within India. Secondly, it is of great interest to demographers to study the population dynamics within States of India, the largest of the States, UP, has the population more than 200 million. Hence, we decided to project the population by States and the Union territories (State/UT) of India disaggregated by age, sex, educational attainment and place of residence to answer following research question:

7. *What would be the impact of considering socio-economic and spatial heterogeneity at the sub-national level on the overall population prospects of India?*

II. Data and Methods

8. As a first step, we acquired (downloaded) published data from the website of Office of Registrar General of India (ORGI). The data from two latest Censuses (2001 and 2011) and various reports from Sample Registration Survey (SRS) for the period since 1999 was used in this exercise.

9. These data were first processed to produce distribution and estimates consistent to our age and education categories. We defined six levels of educational attainment namely: “no education”, “some primary”, “completed primary”, “completed lower secondary”, “completed upper secondary” and “completed post-secondary”. The population was disaggregated in five yearly age groups with ‘100+’ as the last age group.

10. The definition urban type of residence is according to the Census 2011 and was based on population size (>5000 inhabitants), population density (400 inhabitants per km2), and proportion of males working in non-agricultural sector as main occupation (>75%). The administrative regions that are officially a “village” but fulfill the above criteria was reclassified as “Census Town” (CT). While administrative urban regions are known as “Statutory Town” (ST). (See details in section D) (Census India 2011)

a) Fertility

11. Demographic transition theory explains how fertility declines from a very high level (~7 children per women) to a low value (~ 2.1). India is currently towards the end of the demographic transition, with TFR of 2.3 children per woman in 2013 (The Registrar General & Census Commissioner, India 2014) reported for India.

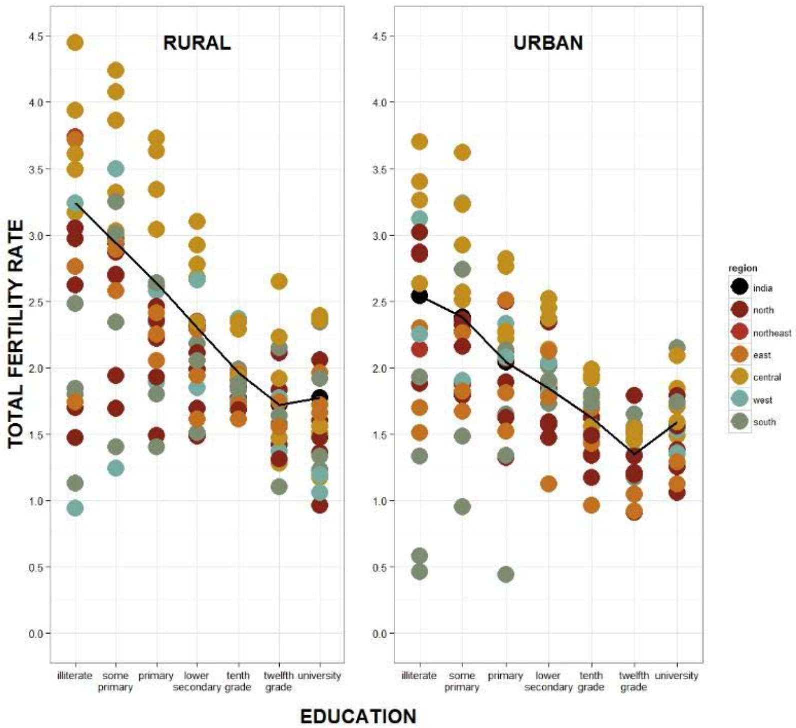
12. The trend extrapolation of Indian TFR by different group of researchers (RGI, UN and WIC) unanimously predicts further decline in the fertility level. However, UN and WIC projections assumes slower decline in the TFR compared to projections done by locals (Office of the Registrar General of India 2006). The expectation by RGI (done in 2006) predicts quite well the reported TFR value in 2013 by SRS. RGI predicted that the fertility level of 2.0 will be reached during the period of 2021-25 almost 15 years earlier than what UN and WIC have in their projection.

13. Due to high variability in the level of TFR between the Indian States (spatial heterogeneity), instead of assuming overall Indian fertility and then deriving the State's fertility, a bottom up approach could have led to the expectation of faster decline. Local experts might be aware of this and hence were predicting faster decline. Most of the spatial heterogeneity, however, can be explained by different composition of the population such as educational level, place of residence, overall development level, ethnicity, cultural practices, religion etc.

14. A consistent linear education differential exists in the States/UT of India (Figure 1), except for the highest education group, where the TFR is higher than among the women with upper secondary education. This is based on the data during 1999-2013. However in some States/UT, the differential is diminishing, that the less educated woman is following the path of fertility experienced earlier by women with higher education. This is true for woman with no education or some primary education, and happening in greater speed.

15. A consistent linear education differential exists in the States/UT of India (Figure 1), except for the highest education group, where the TFR is higher than among the women with upper secondary education. This is based on the data during 1999-2013. However in some States/UT, the differential is diminishing, that the less educated woman is following the path of fertility experienced earlier by women with higher education. This is true for woman with no education or some primary education, and happening in greater speed.

Figure 1 Total Fertility Rate in India by level of education and States/UT separately for Rural and Urban place of residence (Source: SRS 2010-2013)



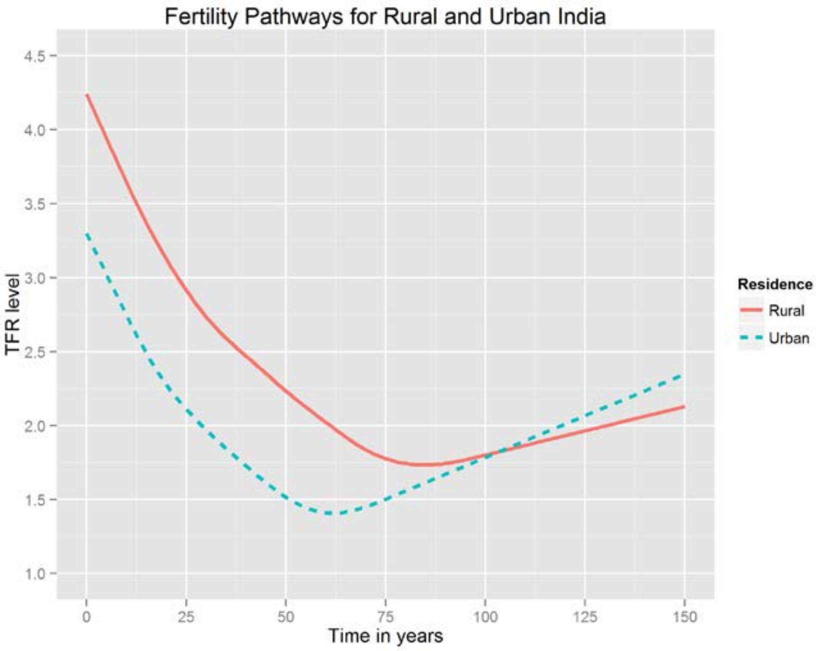
16. We defined two fertility pathways for rural and urban type of residence in India using education specific TFR from the period 1999-2013. Trend shows that fertility is declining rapidly among none or some primary educated women. The fertility rate among primary educated and lower secondary educated is declining slowly and seems levelling off. The fertility rate among upper secondary educated women is the lowest in many cases. And finally, the fertility rate among tertiary educated women is often slightly higher than among the upper secondary educated women.

17. As explained earlier, after various iterations final national pathways separately for rural and urban region were chosen by first aligning the fertility trend for each education and then fitting a smooth spline separately.

18. The smooth spline showed (Figure 2) that the fertility declines to a level (1.73 for rural and 1.40 for urban) and then increases (this is a phenomenon observed in many Western countries). Very low fertility is observed in many Northern States/UT of India mostly among the urban dwellers and could follow the Southeast Asian pattern. However, in many Southern States/UT as well as among the most educated women, the TFR is not so low and is around 1.8 child (e.g. in Tamil Nadu). Further both UN (mostly in the range of 1.6-2.0) and WIC (1.75) assumes higher level for ultimate fertility. Based on these arguments for the baseline scenario we assumed TFR of 1.75 as the ultimate value that fertility among all groups in the urban region will converge. For the rural region, we expect this ultimate level

to be higher than 1.75. The gap in the ultimate fertility level between the rural and urban region was set to the same gap between the minima of the two fertility pathways (0.33, see above).

Figure 2 ***Fertility Pathways for Rural and Urban India, based on SRS 2007, 2010-2013 TFR (Data Source SRS)***



19. The smooth spline showed (Figure 2) that the fertility declines to a level (1.73 for rural and 1.40 for urban) and then increases (this is a phenomenon observed in many Western countries). Very low fertility is observed in many Northern States/UT of India mostly among the urban dwellers and could follow the Southeast Asian pattern. However, in many Southern States/UT as well as among the most educated women, the TFR is not so low and is around 1.8 child (e.g. in Tamil Nadu). Further both UN (mostly in the range of 1.6-2.0) and WIC (1.75) assumes higher level for ultimate fertility. Based on these arguments for the baseline scenario we assumed TFR of 1.75 as the ultimate value that fertility among all groups in the urban region will converge. For the rural region, we expect this ultimate level to be higher than 1.75. The gap in the ultimate fertility level between the rural and urban region was set to the same gap between the minima of the two fertility pathways (0.33, see above).

20. The two national fertility pathways were then used to project the education specific fertility in 70 sub-regions of India. In cases where the fertility level is already below respective pathways, the gap was allowed to remain to carry forward the low-fertility behaviour of women in the region. Fertility for women with up to completed primary education will level off at the ultimate values assumed for rural (2.08) and urban (1.75) regions. Fertility for

women with at least lower secondary will follow the path. If the current value is already less than the minima, we let the difference be maintained.

b) Mortality

21. In India, SRS estimates for life expectancy at birth for females and males were 69.3 years 65.8 years respectively for the period 2009-2013 (midyear as 2011). In the past the mortality situation was worse for females. For the first time in 1980-85, the life expectancy at birth among females (55.1 years) became higher than that for males (54.8 years). The sex difference widened as the increase in the life expectancy at birth for females increased faster than that for males, 2.05 vs 1.9 years (between 2000-2005 and 2005-2010) and further widened with a gain of 2.47 vs 1.6 years for males and females respectively between the period 2005-2010 and 2010-2015.
22. At the state level, the life expectancy at birth varies between States/UT levels. The life expectancy at birth is always higher in the urban areas compared to the rural region with an exception of Kerala in the most recent data from SRS. Through time, the life expectancy seems to be converging, rapidly in the rural areas. The convergence is happening faster among females than males.
23. We generated sex-specific average pathways for the future gain in life expectancy at birth by regressing the gain between two periods on the value from initial period. We fitted simple linear regression and extrapolated the life expectancy into the future using the regression results. For each States/Sex, we started with recently observed average rate of change and converged to the predicted value by 2030.
24. We set lower limits for the predicted value, 0.75 year per five year for males and 1 year for females. This leads to widening of the life expectancy between males and females, which we think will happen in the future. Few rules and limitations were imposed. The five-yearly change in life expectancy at birth was limited to a maximum of 3 years. The gain in life expectancy at birth will converge to the general predicted average gain by 2030. Within each state, the life expectancy in rural area was restricted to remain lower or equal to that in the urban region. And, the gap between the rural and urban region was limited to the most recent observed values.
25. Once the life expectancy were ready, we applied Gompertz transformation method as implemented by KC et al. 2010 to calculate life tables for the calculated life expectancy at birth. We used the life tables for India from the UN Medium Variant in the World Population Prospect, 2015, as standard life tables. (KC et al. 2010)

c) Migration (Internal)

26. The internal migration between rural and urban regions within and between the States/UT, altogether 70 spatial units, is one of the main determinants of the population dynamics in

India. The data for the estimating the flows between the rural and urban regions by States/UT was not readily available and had to be estimated from different available tables from the Census 2001. (2011 is not yet published). Following are the steps and list of data used:

i. In the first step, we used the data on the total number of migrants by sex in the current place of residence (by States/UT and by rural/urban) and last place of residence.

ii. Five yearly age and sex distribution of migrants, who has been living in the current region (destination) since up to 9 years, is available by origin State/UT and by rural/urban. Age distribution of those who moved during last five years is ideal for our projection. Since, the only available data is for the duration 0-9 years, we used the data to distribute the total out-migration to each of the 70 regions by age. We first calculated the destination specific age distribution from each origin. We then divided the age-sex-origin-destination specific number of migrants by the total population at the origin, to roughly estimate the migration rate[2]. Inspecting the distribution revealed some anomalies that needed some further treatment and smoothing. The problem mostly occurs for ages with small population. Also, the problem may be due to reporting error. Also, the age patterns of the migration rates from origin were similar. Hence, we decided to use same age-pattern to all destinations. Further analysis of the migration data should be done for the future update.[3]

27. We assumed that the age and sex specific migration rates will remain constant between the 70 regions of India. With a single set of data, it is difficult to know the trend. Once the data migration data from the 2011 Census will be released, we will conduct further analysis.

d) Urbanization

28. The change in population size and structure of in urban regions could occur due to: i) natural increase (births minus deaths); ii) migration (in minus out), and; iii), reclassification of a rural region to urban. The first two are the inherent part of the projection model. However, urbanization through reclassification needs a separate analysis, firstly to understand what is happening and secondly on how to make future assumptions.

Urbanization through Reclassification

29. The rate of reclassification is difficult to predict. A recent paper by Pradhan (2013) has estimated the number of villages that were classified as Census town (CT) in the census 2011. The classification of villages into Census town was based on three criteria, namely,

[2] A proper way to deal with this is to divided the number of migrants (Age-Sex-Origin-Destination or ASOD) by the number of population at origin = current population + those who left the region – those who came to the region. We will update this in the next run.

[3] We will update this in the next run after correcting for abnormal behavior of the rates.

population size, population density, and proportion of males working in non-agriculture as main occupation. The paper estimates that almost 29.5% of the growth in the urban population (91m) is due to the new CTs (Table 2 in Pradhan (2013)). We used the data estimated from the paper by Pradhan (2013) to estimate the proportion of population reclassified to CTs from Villages. (Pradhan 2013)

30. In order to make assumptions on the transition ratio from Village to CTs/STs for each state, we explored the relationship between various factors (rural population size, proportion of rural residence). The following figure shows a negative relationship between the transition ratios and the proportion urban in each state. The seven outliers are in two groups, the first with very high proportion of urban and smaller States/UT where more than 50% of the rural population make transitions to CT population, the second group consists of states like Kerala and Goa, also with hihgher proportion of urban population.

31. We excluded the seven outliers and fitted a curve (general linear model – normally distributed error with log link function) ***log(y) = A+Bx***, (where, y is transition ratio and x is proportion of rural population). We then let the proportion of rural population at the end of each projection period predict the transition ratio to CTs from Villages. We let the seven outliers to be constant in the future. The predicted proportion was then used to reclassify rural population to urban population.

e) Educational Attainment

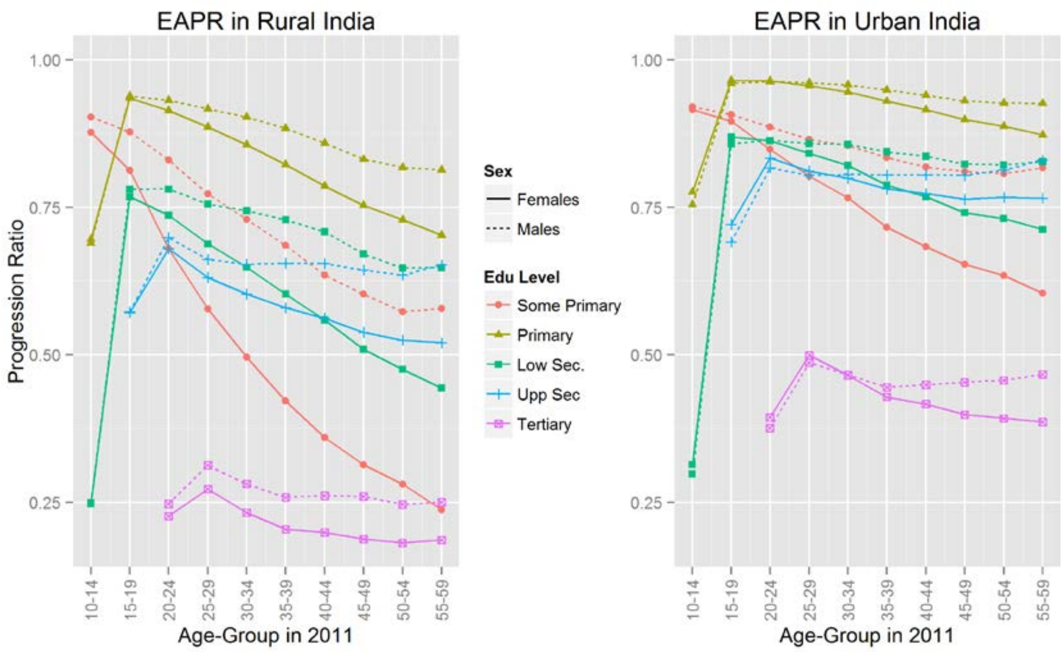
32. We defined six levels of educational attainment namely: “no education”, “some primary”, “completed primary”, “completed lower secondary”, “completed upper secondary” and “completed post-secondary”. The education distribution was available by more than six categories in the Census (2011). We aggregated for the six categories to match the ISCED definition (ref) and studied the education transition between these six levels of education.

33. For a given educational attainment level, we defined education attainment progression ratio (EAPR) to the next educational level as the proportion who completed the next level of educational attainment among those in the current level. For e.g., if in a cohort, 90% have completed at least primary education and 45% have completed at least lower secondary, then the EAPR to lower secondary completion is the ratio of the proportion who completed lower secondary to those with at last primary completed (i.e. 45%/90%. = 0.5).

34. To study the trend, the education distribution in older cohorts provide us information about cohort specific education transitions in the past. Distribution and transitions from consecutive cohorts can be used to analyse the trends in different education categories (see Figure 3). Figure 3 shows the EAPR in five yearly cohorts for rural (left panel) and urban (right panel) regions of India for five educational attainment categories (five colours) for males (dashed lines) and female (solid lines) reported in Census 2011.

35. We analysed each of the trend drawn from several cohorts and defined future education scenarios essentially by extrapolating the trend and in some cases by applying some 'expert' opinion. For e.g. while all other transitions were allowed to become universal, the transition to tertiary from upper secondary was limited to 70% in the urban and 50% in the rural areas. Also, for those regions with slower speed of change than the national one (by States/UT, residence and sex), we allowed the speed (slope) to converge to the national one.

Figure 3 Educational Attainment Progression Ratios in India for five educational levels among males and females by place of residence (Rural and Urban) in 2011 (Source: Census 2011, and own calculation)



Results

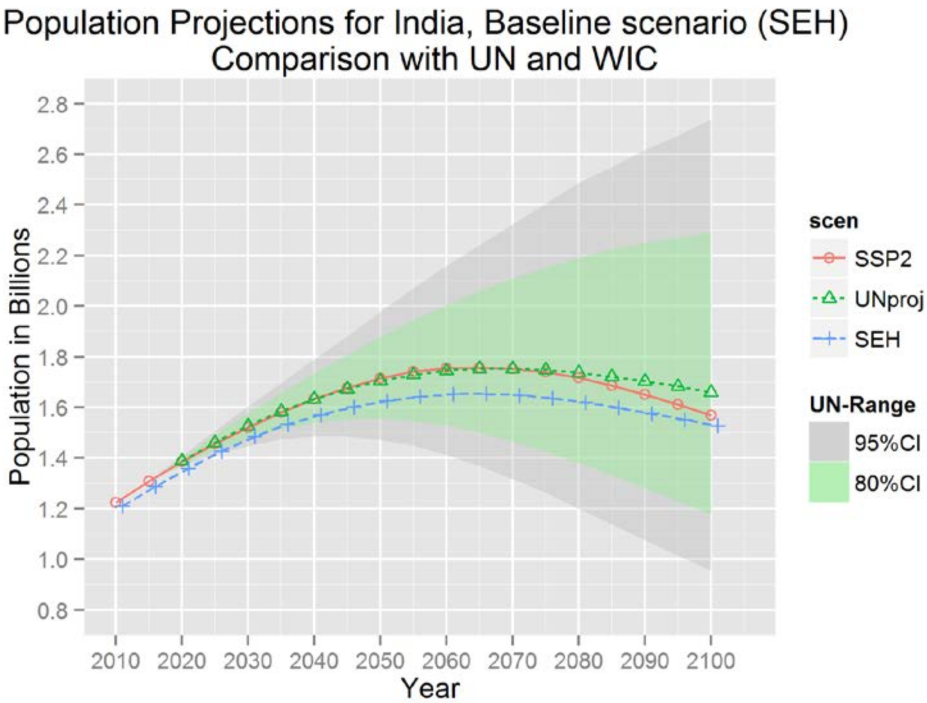
36. We defined a baseline scenario for India and projected the population by age, sex, and educational attainment in each State/UT of India by rural and urban place of residence.

a) Total Population and Fertility rate

37. Population of India will increase rapidly to 1.6 billion by 2046 and will start to decelerate reaching a peak at above 1.65 billion by 2066 (*Figure 4*). The population will then decline below 1.53 billion by the end of the century. The rate of increase is moderate and is driven by the increase in life expectancy. A moderate decline in fertility is expected due to increasing level of educational attainment among women.

38. Compared to the UN projection (Figure 4), peak value in our projection for India is less by 100 million than UN's Medium Variant of 1.75 billion in 2068. By the end of the century, the gap between the two projections, UN and SEH, widens to 130 million with the UN projection reaching 1.66 billion in 2100.

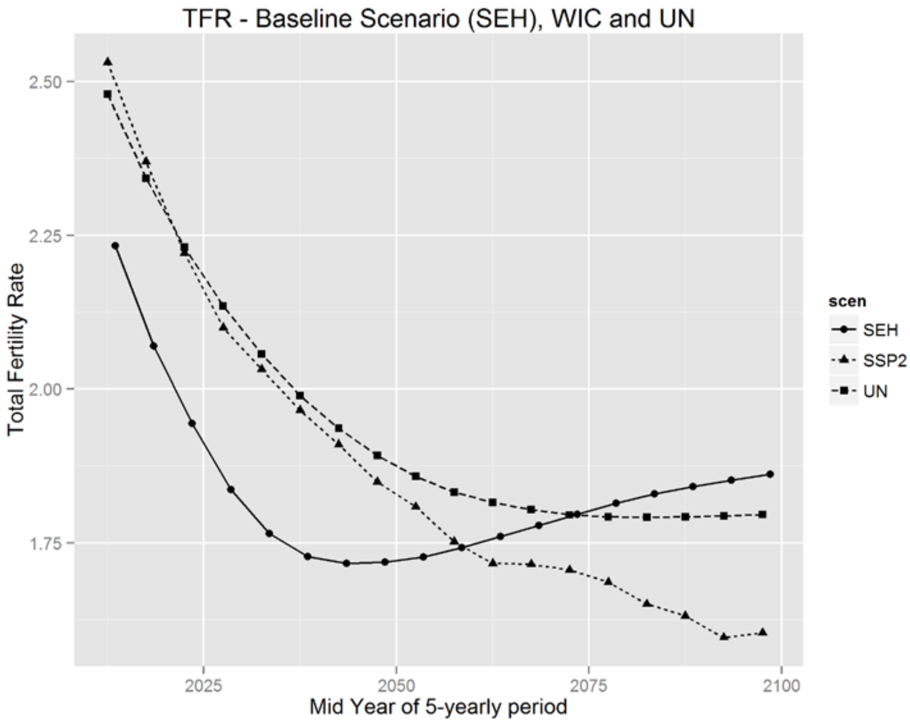
Figure 4 Population Projections for India, baseline scenarios (SEH) along with UN Medium Variant and probabilistic projection, and Wittgenstein Center's projection (WIC)



39. The WIC projection (Figure 4) is almost identical to UN up to 2070s after which the decline is much steeper than the UN and projected to meet our projection (SEH) shortly after the end of the century. The differences between SEH with the WIC and UN are already visible at the first step of projection (2015/2016) mainly due to the difference in fertility assumptions and differential spatial level considered in the projection. While UN and WIC has assumed total fertility for the whole of India, we calculated it by taking weighted averages of ASFRs by States/UT, types of residence, and level of educational attainment among women. The UN assumed a TFR value of 2.48 children per woman for the period 2010-2015 and WIC assumed it to be 2.53 (Census India 2011). Our projection shows the value to be 2.23 children per woman for the period 2011-2016 which is significantly lower than both the UN and WIC. We have discussed this in earlier section about the higher estimates of fertility by the UNPD and the WIC (see section 0)

40. In our projection, the baseline TFR for women in each education group by type of residence and State/UT is derived averaging four annual TFRs reported in SRS during 2010-2013. The rural and urban TFR for the period 2011-2016 in our projection are 2.5 and 1.69 respectively. One reason for lower level of overall fertility in our projection could me the assumed proxy (national TFR by type of residence) for the 15 (out of 35) States/UT for which SRS did not estimate.

Figure 5 Total Fertility Rate in the baseline line scenario (SEH), Wittgenstein Center (WIC) and United Nations (UN) Medium projection



41. In our projection, the baseline TFR for women in each education group by type of residence and State/UT is derived averaging four annual TFRs reported in SRS during 2010-2013. The rural and urban TFR for the period 2011-2016 in our projection are 2.5 and 1.69 respectively. One reason for lower level of overall fertility in our projection could me the assumed proxy (national TFR by type of residence) for the 15 (out of 35) States/UT for which SRS did not estimate.

42. The other sources of difference in the population (size and structure) are mortality and migration. Finally, the assumptions of assuming internal migration rates constant if higher than actual values (in Census 2011), will result in more number of women migrating from high fertility areas to low fertility areas and implies less number of children born. This source needs to be investigated as well.

43. Appendix table A1 (Census India 2011) provides the result of population projection, urbanization level and proportion with at least lower secondary completed in States/UT of India.

b) Internal Migration and Urbanization

44. We find that the rate of urbanization in the baseline scenario seems too slow. It will take another forty years for India to increase by 10 percentage points, from 31% in 2011 to 41% in 2051. And will take another 50 years to become a majority urban country. UN on the other hand projects a very high urbanization level for India that is double than ours, i.e. 50% by 2050. This large difference between ours and the UN's can be explained. While, UN's method extrapolates the urbanization trend at the highest level of aggregation, we apply the bottom-up approach making assumptions at much granular level. UN's approach is to apply the experience from other countries, which might not be appropriate for India with such a huge and diverse population.

45. In our projection, the urbanization is affected by the rural-urban migration assumptions and the reclassification rule. We have kept both rural-to-urban and urban-to-rural migration constant. With growing urban population, the number of population moving from urban to rural area will also grow. A five yearly net flow from rural to urban will peak to about 10 million in the next decade and will decline to around 2.5 million by the end of the century.

46. By allowing the same rate of urban-to-rural migration, we are not implicitly speeding the process of sub-urbanization – i.e. people migration to the rural region but essentially commuting to work in non-agricultural jobs in the urban areas. However, we do not capture this process properly and large portion of this sub-urbanization are becoming 'rural' population in our model.

47. The future rural population will be quite different than what we see now. While many urban-to-rural population are contributing in the process of sub-urbanization, many will be contributing to the formation of new Census Towns. In some areas of India, around big cities for e.g. Delhi, Bengaluru, etc., we could already see this process unfolding. In our projection, we need to include this process of sub-urbanization that is not captured by the definition of "Census Town" employed in the Census 2011. Reclassification rates might increase in the future than what we have assumed in this paper based on migration date from Census 2001 (the migration result from the Census 2011, once available, will provide us more clues).

48. In addition to sub-urbanization and reclassification issues, the third reason for the low urbanization rate in our baseline scenario is due to population weights carried by larger States of India, such as Uttar Pradesh and Bihar with very large proportion of rural less educated population. On top of it, the rural-to-urban migration rate among women is lesser than among men, which essentially means that while men from rural area go to work in other States, women (wives) are left behind to bear and rear the children as well as to take

care of the household, farms and the elderly. The sensitivity of these two States is such that, when we calculate the speed of urbanization in rest of the States/UT, the speed is much faster from 35% in 2011 to 51% by 2061.

III. Conclusion

49. We developed a multi-regional multistate population projection model for India that can simultaneously simulate population heterogeneity in the demographic (age-sex), socio-economic (educational attainment), and spatial (States/UT and rural/urban) dimensions. This is a first model of its kind for India. We populated the model with data and estimates and defined a baseline scenario based on the analysis of data on five components (fertility by education, mortality, internal migration, education, and reclassification) from the Census and SRS. We ran the projection and presented some preliminary results.

50. The results show that the overall fertility for India is lower than estimated/assumed by UN and WIC. The projection resulted in slower rate of urbanization in India from 31% in 2011 to 41% in 2051 and we presented several explanations. The most important of all the reasons could be the largely rural, less educated, large population in Bihar and Uttar Pradesh that are slowing down the momentum of urbanization in India.

51. The same reason can explain higher rate of population growth in India resulting from the bottom-up approach than when the projections are done for two regions (rural and urban) of India. We plan to work further on issues in the baseline scenario that are identified in this iteration. In addition, we will conduct several sensitivity tests and define alternative scenarios with relevant narratives for India (for e.g. SDG scenario). We also plan to compare the results with the SSP2 scenario and project other SSP narratives relative to our baseline scenario.

References

Census India. 2011. "Census of India 2011. Provisional Population Totals. Urban Agglomerations and Cities." Delhi, India: Office of the Registrar General & Census Commissioner, India(RGI). http://censusindia.gov.in/2011-prov-results/paper2/data_files/India2/1.%20Data%20Highlight.pdf.

KC, Samir, Bilal Barakat, Anne Goujon, Vegard Skirbekk, Warren C. Sanderson, and Wolfgang Lutz. 2010. "Projection of Populations by Level of Educational Attainment, Age, and Sex for 120 Countries for 2005-2050." *Demographic Research* 22 (Article 15): 383–472. doi:10.4054/DemRes.2010.22.15.

Office of the Registrar General of India. 2006. "Population Projections for India and States, 2001-2026." REPORT OF THE TECHNICAL GROUP ON POPULATION PROJECTIO NS CONSTITUTED BY THE NATIONAL COMMISSION ON POPULATION. New Delhi: Office of the Registrar General and Census Commissioner of India. https://nrhm-mis.nic.in/Part%20B%20Demographic%20and%20Vital%20Indicators/Population%20Projection%20Report%202006%20by%20RGI.pdf.

Pradhan, Kanhu Charan. 2013. "Unacknowledged Urbanisation: New Census Towns of India." *Economic and Political Weekly* xlviii(36): p.43–51.

The Registrar General & Census Commissioner, India. 2014. "Sample Registration System (SRS)." *Www.censusindia.gov.in*. http://www.censusindia.gov.in/2011-Common/Sample_Registration_System.html.

United Nations. 2014. "World Urbanization Prospects: The 2014 Revision." ST/ESA/SER.A/352. New York, NY: Department of Economic and Social Affairs, Population Division, United Nations. http://esa.un.org/unpd/wup/Highlights/WUP2014-Highlights.pdf.

———. 2015. "World Population Prospects: The 2015 Revision, Key Findings and Advance Tables." Working Paper ESA/P/WP.241. New York, NY: Department of Economic and Social Affairs, Population Division. http://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf.

Appendix Table A1 Results of Baseline population projection

Region	Total Population (in millions)			Proportion Urban			Proportion with at least secondary education (15-64y)		
	2011	2031	2051	2011	2031	2051	2011	2031	2051
Andaman and Nicobar Is.	381	531	636	0.38	0.41	0.43	0.59	0.73	0.85
Andhra Pradesh	84573	94555	96181	0.33	0.40	0.47	0.36	0.59	0.80
Arunachal Pradesh	1384	2031	2499	0.23	0.29	0.34	0.40	0.60	0.76
Assam	31206	38074	40613	0.14	0.17	0.21	0.41	0.57	0.73
Bihar	104098	133197	150118	0.11	0.12	0.13	0.32	0.52	0.71
Chandigarh	1055	1767	2359	0.97	0.96	0.97	0.71	0.74	0.82
Chhattisgarh	25545	31121	33408	0.23	0.27	0.31	0.37	0.60	0.78
Dadra and Nagar Haveli	344	616	847	0.47	0.61	0.70	0.47	0.64	0.78
Daman and Diu	243	472	677	0.75	0.77	0.83	0.59	0.67	0.77
Goa	1459	2008	2442	0.62	0.75	0.81	0.63	0.75	0.85
Gujarat	60439	74858	82900	0.43	0.50	0.56	0.42	0.60	0.76
Haryana	25351	33926	40630	0.35	0.42	0.47	0.53	0.69	0.81
Himachal Pradesh	6865	8242	9170	0.10	0.13	0.15	0.60	0.78	0.90
Jammu and Kashmir	12541	14838	16142	0.27	0.31	0.36	0.51	0.72	0.85
Jharkhand	32988	41998	47065	0.24	0.31	0.37	0.36	0.56	0.74
Karnataka	61095	71010	75441	0.39	0.45	0.51	0.44	0.65	0.82
Kerala	33406	36386	35719	0.48	0.70	0.81	0.68	0.85	0.94
Lakshadweep	64	95	120	0.78	0.81	0.82	0.56	0.78	0.91
Madhya Pradesh	72627	91997	102161	0.28	0.31	0.34	0.36	0.58	0.76
Maharashtra	112373	140243	159372	0.45	0.54	0.61	0.53	0.71	0.84
Manipur	2856	3367	3517	0.29	0.37	0.44	0.60	0.74	0.84
Meghalaya	2967	3963	4615	0.20	0.23	0.25	0.35	0.54	0.71
Mizoram	1097	1276	1360	0.52	0.68	0.79	0.53	0.68	0.80
Nagaland	1979	2470	2677	0.29	0.33	0.37	0.50	0.63	0.76
Delhi	16788	25932	33994	0.98	0.98	0.99	0.68	0.73	0.81
Odisha	41974	48478	51031	0.17	0.21	0.26	0.37	0.59	0.78
Puducherry	1248	1640	1893	0.68	0.72	0.76	0.66	0.81	0.93
Punjab	27743	33610	37674	0.37	0.44	0.50	0.55	0.70	0.82
Rajasthan	68547	87947	98262	0.25	0.28	0.32	0.35	0.56	0.74
Sikkim	611	851	1014	0.25	0.22	0.20	0.43	0.57	0.72
Tamil Nadu	72147	79756	80999	0.48	0.54	0.60	0.52	0.73	0.90
Tripura	3674	4524	4894	0.26	0.37	0.46	0.44	0.59	0.75
Uttar Pradesh	199799	256006	282959	0.22	0.24	0.27	0.42	0.61	0.78
Uttarakhand	10086	12729	14486	0.30	0.38	0.44	0.56	0.74	0.85
West Bengal	91276	104878	107772	0.32	0.40	0.47	0.39	0.53	0.70
India	***1210827***	***1485391***	***1625646***	***0.31***	***0.36***	***0.41***	***0.43***	***0.62***	***0.78***