Background
The paper is embedded in an interdisciplinary case-study at the International Institute for Applied Systems Analysis (IIASA) that investigates the impact of Socioeconomic Heterogeneity in Model Applications (SCHEMA) and on the environment and well-being in India.

This study is motivated by two research questions: (1) How does the accounting of socioeconomic heterogeneity, measured by educational attainment, improve population projections for India?, and (2) How will changing patterns in urbanization affect the population projection, depending on the spatial scale (national vs. subnational) considered in the projections?

Introduction
Both of these research questions represent fundamental questions in the field of spatial demography and population research as social heterogeneity is strongly discussed, especially for highly populated urban areas. Lewin (2014)\(^1\) postulated a correlation or interdependence between socio-economic, demographic and other characteristics that produce social heterogeneity. Since these socio-demographic factors can show deviant peculiarities depending on the spatial patterns, for e.g. between urban and rural place of residence, the social and spatial heterogeneity in population dynamics has to be considered.

Fig.1) Differentials in Total Fertility Rate in India by education, state, residence and region, 2013
(Source: Sample Registration System 2013, India)

Figure 1 illustrates the variation in the level of Total Fertility Rate (y-axis) in India in 2013 by level of education (x-axis) and by States and Union Territories (State/UT) in India. TFR for rural and urban place of residence are presented separately in two panels. The average TFR for India is presented by the black lines. The colors of the points represent 6 different regions of India.

Three observations can be made, first, a steep negative gradient by education is visible reflecting the negative association between education and actual fertility. A slight positive slope for women with university education shows that those women possibly desire more children and are in a better economic position to afford them. This U-shape can be observed in some European countries.

---

Secondly, the same gradient is visible for urban and rural areas in India, but on a different level. This implicates a spatial dimension of fertility related to place of residence. Thirdly, we can see a large deviation within and between regions and states, as for instance states in Central India shows significant higher fertility levels among all educational groups, compared to the Indian average. The spatial dimension is not only visible in the case of fertility, but also of mortality and migration, so that an inclusion of spatial and social heterogeneity in the projections would better inform the prospective population development of India.

Methods
In this project we developed a multi-dimensional population projection model that projects the population of India by five dimensions: three cover personal characteristics (age, sex, and educational attainment) and two spatial characteristics (35 states/union territories, and with rural and urban, 2 residences). In total 70 sets of subnational populations are projected in 5 yearly steps from 2010 up to 2050.

This projection of the population is achieved using a multi-state model by levels of education. Education transitions are considered between the six education categories until age group 30-34 years, after which we assume that the achieved educational attainment will not change.

So far, we have prepared a base-line scenario to show the impact of different spatial layers, (A) national, (B) national by residence, and (C) by State/UT and by residence, on the projection outcome. Our assumptions for the base-line scenario are based on the analysis of trends and differentials in demographic rates and education transitions from various sources (e.g. 2001 and 2011 census, SRS - Sample Registration System, DHS).

SRS data show that among women with higher education levels the fertility has levelled off at below replacement level in almost all populations, while among women with lower education levels the TFR has been steadily declining. Based on this, we assumed that for women with higher levels of education the current fertility levels will remain the same, while for women with lower education levels the TFR will further decline to below replacement level of 2.1 by 2015. Empirical data from DHS on education, residence and State/UT were used to convert TFR into ASFR. As mortality data by education were not available we apply standard life tables from the base period by residence and by States/UT (SRS data), currently, we hold this constant. We plan to extrapolate the trend in sex-specific life expectancy at birth following the UN assumptions for India. The final projection will be presented at the PAA.

The internal migration rates between and within states/union territories by residence were estimated from 2001 census (see Figure 2) as more recent migration data are not published yet. As soon as they are available (hopefully by the end of 2015) the migration numbers will be updated. A similar approach was used for international migration, whereby the numbers here are quite low compared to other demographic events.

Lastly, the transitions between educational groups were analyzed for all spatial units, retrieved from 2011 census data. Based on this analysis, education attainment progression ratios (EAPRs) were extrapolated into the future. In case a region has lower levels of educational attainment compared to national average, a convergence to the sex and residence specific average Indian pattern by 2050 was assumed.

Results
Referring back to our two research questions, whether the accounting for socioeconomic (educational attainment) and spatial (place of residence and subnational) heterogeneity affect our projections for India. And the answer is YES.
The explanation lies mostly in the fertility differentials and the active domestic migration network as shown on Figure 2. The largest flows are hereby from the rural areas of Uttar Pradesh (UP) to urban areas in the same and other states. Large migration flows were affecting in 2001 the urban areas in Maharashtra (MH) with 10 cities with more than 1 million inhabitants, including megacities like Mumbai, Pune or Nagpur.

Therefore considering in population projections spatial regions beyond the national level, such as by residence or subnational units, will automatically require the consideration not only of international but also domestic migration that may affect the inner-country origin and target regions.

Fig.2) Internal Migration in India by States & Residence, 2001

Projecting the Indian population by residence separately shows us the expected increase in both urban and rural regions, peaking in the mid of the 21st century (see Figure 3) before declining again. By considering rural to urban differences the growth of the rural population cushions the growth of the share of urban population as this share just increases from 31 percent in 2010 to 34 percent in 2050. As a result the total Indian population would increase from 1.2 billion inhabitants in 2010 to 1.6 billion by 2050 (see Figure 4), before declining to 1.3 billion in 2100 instead of .... If projections are done for the whole of India.

Considering not only urban and rural differences, but also different administrative levels like states, we get a quite different picture. Here we can see a higher increase for both, rural and urban regions in India, which aggregates to a higher total population that peaks later and declines thereafter at a slower pace. The gains in population size by adding the state-level is mainly influenced by the increase of rural population. For instance, in 2050 the total population difference between the national and the state-level projection would be 70 million, from which 54 million (77 percent) are contributed by the rural population. (see Figure 3 & 4)
In terms of social heterogeneity, here shown with education, the total population (males & females) of India will increase its share of population aged 25 years plus and no education will decrease in the baseline scenario from 39.3 in 2010 to 12.4 percent in 2050, while vice versa the proportion of those with upper secondary and post-secondary education would increase from 28.4 to 60.2 percent. (see Figure 5)

**But how does the educational composition differ by residence and gender?** – Figure 6, shows in a scatter plot the female to male ratio of population aged 25 years plus with upper secondary and higher education by states (points), region (point color), with the ratio for the urban population on the y-axis and for the rural population on the x-axis, between 2010 and 2050. The demarcation line at level 1.0 on both axis represents a gender balance, while the diagonal line shows the urban to rural difference in the respective year.

Here we can see in 2010 that women in urban areas are more educated than those living in rural areas. But women in both areas are lagging behind men, except in Kerala (KL) where massive investments in the education system in the past have brought massive improvements to female and general education. Kerala is also in a “leading position” in 2050, but the other states/UTs are catching up fast and converging
to gender balance. Also the urban and rural differences get in 2050 narrower in most states, except some lower populated Union Territories in the Southern and Western Regions.

Conclusion
This convergence in education by sex, residence and state is an implicit part of the projection assumptions that lead in the long run to higher equality within India based on the recent demographic and socioeconomic trends. More importantly, the projections show the importance of taking into consideration the spatial scale.

When not considering the state-level in the projection, we are implicitly assuming that each state has the same population weight throughout the projection. However, due to differences in the demographic and socioeconomic structure among the rural and urban populations between states, the overall composition will change in the future. For instance, the states of Uttar Pradesh (UP) and Bihar (BR) are inhabiting 30 percent of India’s rural population in 2010. Due to the high fertility levels among women living in rural areas of UP and BR, this share will most likely increase to 36 percent by 2050 and 44 percent by 2100. Their simple population weight lowers the pace of the fertility decline in rural India. Additionally, national projections ignore domestic migration flows that in general happen from higher fertility rural regions to lower fertility urban regions.

The base-line scenario will serve as basis for the development of further scenarios along different narratives. Additionally we will present sensitivity analysis to investigate the elasticity of different demographic determinants (fertility, mortality, and migration). We aim to present the here described work and this further model extensions at the PAA.