Distr.: General
1 April 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 7 of the provisional agenda
Beyond projections by age and sex

Education-specific labor force projections: painting the global picture

Note by Elke Loichinger[1]

Summary

National and international organizations regularly perform labor force projections by age and sex. Adding another dimension to this already three dimensional endeavor brings along challenges but still warrants serious consideration, given significant education-differentials in labor force participation and constant changes in populations' education composition.
In this overview paper, I discuss the possibilities and constraints of performing education-specific labor force projections for a wider range of countries. The biggest constraint is data availability; readily available data from national statistical offices and international organizations for the most part do not break down the labor force by age, sex and educational attainment at the same time. If microdata are accessible, the necessary tabulations can be done on one's own. To a certain degree, the data situation does dictate the choice of methods that can be applied to perform education-specific labor force projections. The more elaborate the method, the more extensive the data requirements in terms of break-down and time-horizon that is covered. A side-issue that is addressed is the problem with open-ended highest age-groups when comparing labor force data, particularly in aging societies.

[1] Researcher at Institute for Human Capital and Development, Vienna, Austria. Academic Researcher at College of Population Studies, Chulalongkorn University, Bangkok, Thailand.

I. Introduction

1. In order to be able to estimate future labor supply, labor force projections by age and sex are regularly performed by national and as well as international bodies. Some have the purpose to profect as accurately as possible the short- to medium-term development of age- and sex-specific labor supply, whereas others are rather aiming at estimating the outcome range for a longer time-horizon. The reason to disaggregate participation by age and sex is the fact that participation does vary significantly over the life course and between men and women.

2. Another characteristic that has been associated with different levels of participation is education: persons with differing highest levels of educational attainment show different labor market attachment. These differentials are the result of a complex combination of factors that affect the supply and demand of labor and are usually larger for women than for men. Also, the direction of education differentials can change in the course of countries' economic development and are not uniform in size across age. Hence, incorporating information about educational attainment can be worthwhile endeavor in order to capture another dimension of heterogeneity in participation besides the one captured by age and sex.

3. At the last Joint Eurostat/UNECE Work Session on Demographic Projections in Rome in 2013, I presented results for education-specific labor force projections for the European Union (Loichinger 2015). Since then, I have been exploring the idea of extending the regional scope to countries outside of Europe. In this overview paper, I discuss the possibilities and constraints of doing so. The biggest constraint is data availability; readily available data from national statistical offices and international organizations do not consistently break down the labor force by age, sex and educational attainment at the same time. Specific data requests are not always successful. If microdata about economic activity are accessible, the tabulations can be done on one's own. To a certain degree, the data situation does dictate the choice of methods that can be applied to perform education-specific labor force projections. The more elaborate the method, the more extensive the data requirements in terms of break-down and time-horizon that is covered. A side-issue that I address is the problem with open-ended age-groups when compiling and comparing labor force participation rates across space and time, particularly in aging societies.

II. Labor force projections and educational attainment

4. The consistent addition of educational attainment to labor force projections has both, a methodological and a qualitative justification. Given that labor force participation (LFP) rates differ across education levels and that populations' educational compositions are changing, adding education to the projections picks up on an important variation in addition to the differences already captured by age and sex. Second, it allows to estimate the education composition of the future workforce, and with education being a proxy for individual productivity, this can be used to gauge future aggregate labor productivity. In doing labor force projections by educational attainment it is possible to explicitly show changes in the composition of labor supply – overall and by age-groups. This is useful because

it illustrates the shift in educational attainment along cohort lines and what that means for the composition of the adult population. This clearly shows that in the great majority of countries, educational expansion has led to a situation where younger adults have obtained higher levels of education that those born before them. With them replacing these older cohorts, the composition of the working-age population is seeing an "upgrade" in its human capital, which has crucial implications for labor supply and demand. Particularly in countries with rapidly ageing populations, having information about not only the prospects of labor supply in terms of absolute numbers of persons but additionally also in terms of their education level is beneficial; smaller labor forces might just be as productive as larger ones if labor productivity increases due to increased levels of human capital. Also, information about educational attainment is also of interest when the role of female labor supply is being discussed. Education is strongly associated with female economic activity (Besamusca et al. 2015; Pena-Boquete 2016, Verick 2014), and breaking down female labor supply by educational attainment can capture this better than just looking at data by age.

5. There seems to be a more or less universal pattern of development, where the education gradient in LFP is negative at first – when the agricultural sector still plays a large role in the overall economy – but then turns positive over time, as labor demand shifts away from agriculture and returns to education increase (Yubiku and Schlabach 2009). However, when and under what conditions this happens depends on the specific country context. There is no automatism that translates higher levels of educational attainment in higher levels of economic activity, as theory might predict. If there is a (skill) mismatch between the supply of those with higher education and labor demand, then higher levels of education will not translate into higher economic activity. When this happens, the structure of the economy and consequently employment has not (yet) caught up with educational expansion. This can be seen in countries with large cohorts of young, educated adults that are trying to enter the labor market but are not able to find adequate jobs.

III. Adding education as an additional dimension to labor force projections

6. In order to project a country's future labor force, two things are required: information on economic activity, along the relevant dimensions (age, sex, and whichever other dimension might be included), as well as population projections, along these same dimensions. Both, participation rates and population numbers, are then multiplied to obtain the future labor force. Projecting the labor force by age and sex alone is not constrained through the availability of the necessary population projections, because even if there are no country-specific projections performed by any national entity there are always the UN data to resort to (UN 2015). When adding education, constraints appear with both components, i.e. education-specific participation rates as well as education-specific population projections.

7. The efforts by the ILO to estimate past and present labor force participation rates by age and sex give a good indication on global data availability concerning economic activity when education is not considered. According to the ILO, data availability of past and present (1980 to 2012) labor force participation rates varies greatly by country groupings (ILO 2013). For this whole period, only 31% of all potential data points were available, and the missing data points had to be modeled (ILO 2013: 11). It can be expected that the result would be significantly lower if the education dimension was to be included for the same number of countries and time period.

a) ***Data sources***

8. The best starting point for already aggregated data related to economic activity is the ILO. The ILO's central statistical database *ilostat* contains detailed aggregate data for selected yearly indicators, among them also absolute numbers of both, the population and the economically active population, by age, sex, and educational attainment. In case the information for the economically active population and the population in itself stems from the same base data (i.e. is based on the same survey, same census etc.) it is possible to calculate education-specific labor force participation using the absolute numbers provided by *ilostat*. This still re quires a lot of data cleaning, since breaks in series and changes in definitions have to be considered in the construction of time-series data for any country. Also, the breakdown by age-groups varies across time and space and data for 10-year age-bands has at times to be converted into 5-year age-groups.

9. If the data are not readily available in aggregate format in any database – be it as readily available rates or absolute numbers that can be used to calculate the rates – calculations can be done using microdata. The data sources are basically the same as for labor force projections by age and sex, predominantly labor force surveys, census data, household surveys, or administrative data. Breaking down the population by the additional education variable means however that sample sizes that were large enough to calculate participation rates by age and sex might be getting too small to obtain reliable estimates. If the data source are labor force surveys, this is usually no problem, but if other household surveys or subsamples of larger surveys (e.g. available census samples on IPUMS) are used it might become an issue. A good starting point for access to labor force surveys is the list of surveys provided by the ILO (http://www.ilo.org/dyn/lfsurvey/lfsurvey.home).

10. Various issues emerge that are not restricted to education-specific analyses but that apply already when rates are calculated by age and sex. Concerning the coverage of the survey, some labor force surveys are constrained and only include the population in urban areas or the population in the larger metropolitan areas. In general, for any country, data on economic activity based on census data does usually not match the results

obtained from labor force surveys, and labor force data should be the preferred choice if available. In the example of Australia, some of the reasons for this discrepancy are differences in sampling methodology (self-enumeration vs. face-to-face interview), the number of questions used to determine labor force status, the in- or exclusion of members of the armed forces, and the treatment of non-response, to name just a few (Australian Bureau of Statistics, 2012). Besides labor force surveys and census data, other surveys that collect data on economic activity and individual characteristics can be used as well, for example national socio-economic household surveys or internationally planned comparable data collections (e.g. EU SILC, currently covering EU countries plus Iceland, Turkey, Norway and Switzerland). An increasing number of aging surveys are available (e.g. SHARE, ELSA, SAGE, KLoSA, CHARLS) and these data can complement the data on economic activity of persons aged 50 or 60+, in case this information is incomplete or unreliable elsewhere. The suitability of each data source has to be tested (coverage, sampling frame, definitions, treatment of non-response etc.). In countries with a comprehensive registration system that also collects data on economic activity and other individual characteristics administrative records can also be used, but this possibility applies to only a small number of countries.

11. The definition of educational attainment is another aspect to consider. Besides the national classifications that exists and that capture past developments and present situation of countries' education systems, data are often provided also according to the ISCED 97 classification. This classification has recently been revised (ISCED 11). In many instances, data are still available according to ISCED 97, sometimes complemented by the newer definition. In those cases where data are only available in the national classification and where the conversion to international definitions are done on one's own, it is important to consider possible past changes in countries' education systems in order to label the attainment of all age-groups accurately.

12. As stated earlier, education-specific labor force projections are only possible if population projections are available along the same dimension and with the same education categories. Performing population projections by age, sex and education simultaneously is a very time and data intensive task; not only does it require baseline data that is broken down by educational attainment, in addition to age and sex, but it also requires future assumptions about all three demographic parameters (fertility, mortality and migration) to be education-specific. With the launch of the updated dataset on education-specific population projections for 195 countries (Lutz, Butz and K.C. 2014), the performance of global education-specific labor force projections has become possible, given that the respective data on labor force participation can be collected. The education categories that the WIC education projections are broken down by are no education, incomplete primary education, primary education (ISCED 1), lower secondary education (ISCED 2), upper secondary education (ISCED 3) and post-secondary education (ISCED 4,5,6).[2] This selection of categories allows a detailed distinction at the lower end of

[2] The complete dataset, based on ISCED 97, can be found at http://www.oeaw.ac.at/vid/dataexplorer/

educational attainment; however, in the most advanced economies, it does unfortunately now allow to distinguish between persons that have completed tertiary education (ISCED 5 and 6) and those who have post-secondary non-tertiary (ISCED 4) education as their highest level of educational attainment.

b) *Methods*

13. There are four principal approaches to project labor force participation by age and sex: time-series extrapolation, regressions, a qualitative/target approach, and cohort analysis (Houriet-Segard and Pasteels 2011; ILO 2013). There is no principal reason why these methods should not be used with a further break-down by education. Besides methodological considerations, the choice of one or the other method is largely data-driven. More elaborate methods require more detailed and comprehensive data and also for a longer time horizon. In general, it is not necessary to have longitudinal data for any method but repeated cross-sectional observations are sufficient. As it is the case with the projections of the economic active population by the ILO, it is possible to combine approaches (ILO 2013).

14. A crucial consideration is the selection of education categories to break the data down into. This decision depends on mainly two factors (besides the consideration of the categories used in the education-specific population projections): the distribution of the population by educational attainment and the observed differentials in labor force participation. A complicating factor is the fact that the education composition changes across age-groups and with small population numbers in specific age-sex-education combinations, it becomes difficult to estimate reliable participation rates. Figure 1 exemplifies this point: in Spain as well as in Thailand, the share of those with post-secondary education above age 50 is relatively small, making it difficult to estimate education-specific participation rates for men and women at these ages and this education level. Also in Spain, the share of the population that possesses at most primary education is much lower for young than older age-groups, again posing a problem with the estimation of the participation rate for these ages. A possible solution is to collapse education categories and to combine this group with the next lower education level. In general, deciding on the number of education categories is a trade-off between capturing the education dimension as detailed as possible while at the same time keeping the categories manageable.

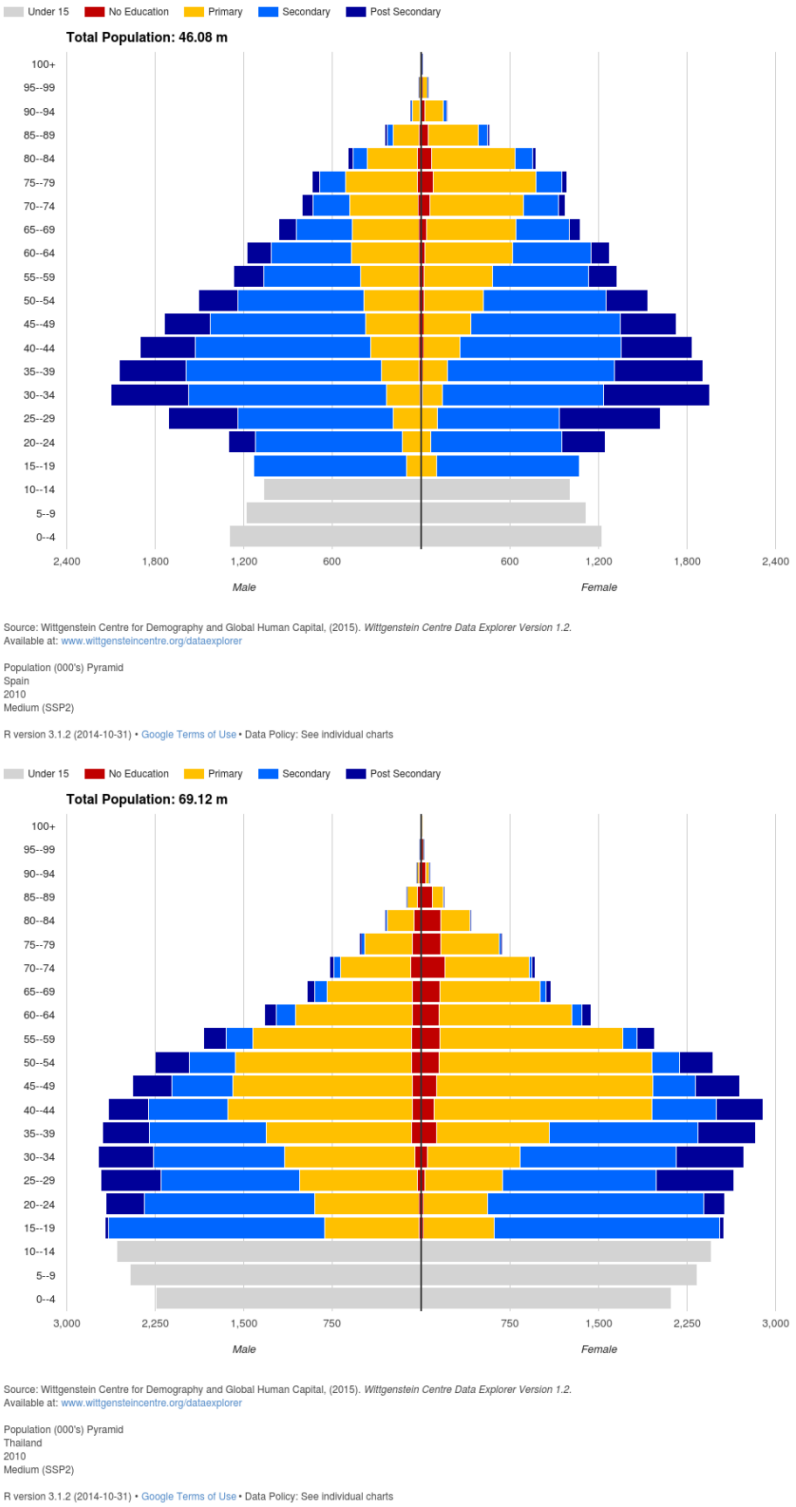

Figure 1: Population pyramids by education, 2010, Spain (top) and Thailand (bottom) (source: Wittgenstein Centre for Demography and Global Human Capital, (2015). *Wittgenstein Centre Data Explorer Version 1.2.* Available at: http://www.wittgensteincentre.org/dataexplorer)

IV. Issues with open-ended age-groups

15. The term "(total) labor force participation rate" is often used without explicit definition about which ages the given values refer to, be it in academic writing or in the media. Only too often developments over time or comparisons between countries are presented, for both sexes or men and women separately, without stating what ages the given numbers specifically refer to. Using the case of Japan as an example, I will exemplify that it does actually matter significantly whether figures are presented for all ages, e.g. 15+, or for a broader age interval, e.g. 25 to 64, and that the interpretation of changes in participation across countries and over time is directly connected to this.

16. Using data for ages 15+ or 15-64 does not matter that that much in countries where there are only small proportions of the population above age 65. However, this is not the case anymore in a lot of countries, and will be even less so in the future. When showing participation rates for open-ended or large age-groups, developments over time or comparisons across countries can be flawed because changes in age-specific participation patterns are confounded with changes in the age-composition of the adult population. Whether this is a problem depends on the interpretation of observed changes in aggregate participation. Comparing LFP 15+ over time and interpreting a decline as a decline in LFP is potentially wrong – age-specific participation rates could have not changed at all but the population structure has shifted towards age-groups with lower levels of participation. Even constant age-specific participation rates will inevitably lead to a decline in LFP 15+ and LFP 25+ in an aging population. If the goal is to show that the share of the adult population that is economically active is going down, irrespective whether this is due to changes in behavior or changes in the population structure, than LFP 15+ is still the right indicator.

17. Using Japan – a prime example of an aging country – I illustrate my point. Figure 2 shows age-specific labor force participation patterns for Japanese men and women for the years 1984 and 2014. These 2 points in time that are 30 years apart show only minimal differences in participation rates of men. For women, on the other hand, participation rates increased for every age-group with the exception of the youngest and oldest women.

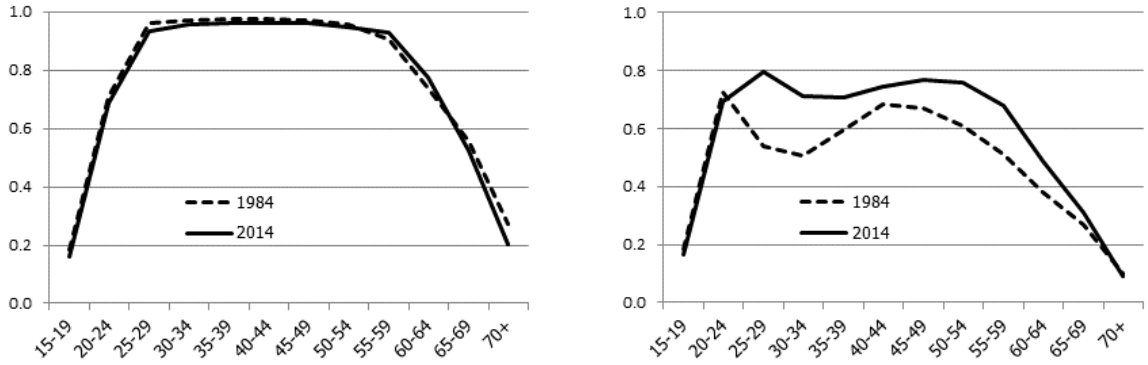

Figure 2: Labor force participation rates, Japan, 1984 and 2014, men (left) and women (right). (source: Statistics Japan, historical labor force Survey data by 5-year age-groups and sex, http://www.stat.go.jp/english/data/roudou/lngindex.htm)

18. How do the patterns in Figure 2 square with the results in Table 1 that participation rates for ages 15+ and 25+ did decrease for Japanese men and stay constant for Japanese women? The answer is that the age-composition of the population changed significantly during this period, showing strong increases in the older population, in relative as well as absolute terms. This can be verified in two simple ways (and more elaborate decomposition techniques would lead to the same conclusion). First, instead of using open-ended age-groups, participation rates are calculated for 2 broad closed age-groups, 20-64 and 15-69. This reveals that participation on the aggregate level did actually stay constant for men between 1984 and 2014, and increased significantly for women (from 0.59 to 0.70 for ages 20-64) (see Table 1). Second, it is possible to control for the effect of changes in the age-composition of the population by applying age-specific participation rates for 2014 to the population composition in 1984. In doing so, any difference of these results with the actually observed aggregate participation rate for ages 15+ in 2014 must be due to changes in the age composition. For men, the thus calculated LFP 15+ in 2014 is only a little lower than in 1984 (0.77 compared to 0.79), whereas for women, it is significantly higher (0.57 compared to 0.49).

age-groups	**men**		**women**	
	1984	**2014**	**1984**	**2014**
15+	0.79	0.70	0.49	0.49
25+	0.87	0.74	0.50	0.50
20-64	0.92	0.91	0.59	0.70
15-69	0.83	0.82	0.53	0.62
	LFP rate in 2014 when LFP rates of 2014 are applied to population of 1984			
15+		0.77		0.57

Table 1: Labor force participation rates, Japan, 1984 and 2014, by sex, for selected age-groups (source: Statistics Japan, historical labor force Survey data by 5-year age-groups and sex, own calculations. http://www.stat.go.jp/english/data/roudou/lngindex.htm)

19. The same argument means that changes in the participation rate of the aggregate age-group 65+ have to be interpreted carefully. In addition, a related but slightly different reasoning applies to the open ended highest age-group, and it is here even more important in the context of labor force projections: with increasing shares of persons falling into the highest age-group – be it 65+, 70+ or 75+ – participation rates for this age-group are applied to a very different age-composition of the elderly in the future compared to today. Over time, the age composition is shifting towards higher ages within this highest age-group, meaning that even if participation rates for this last age-group are for example kept constant in the projection, this directly implies that people of higher and higher ages are considered to be economically active. This might be a valid assumption, it has just to be spelled out clearly and interpreted correctly.

20. The arguments presented here do apply just the same when data are additionally broken down by education. Maybe even more so, since participation rates by education particularly for e.g. ages 60+ will be influenced by the education composition within this highest age-group. This has to be kept in mind, even more so when changes in participation over time are interpreted and assumptions about future participation rates are made.

V. Conclusion

21. This is mostly an overview paper. I presented the benefits, possibilities and challenges of performing education-specific labor force projections. I also discuss what data sources there are besides labor force surveys and talked about the choices of methods. Naturally, adding a further characteristic – even if it captures some heterogeneity that otherwise goes unnoticed – also introduces more uncertainty. The extent of this potential additional error depends on the quality of the data that participation information by age, sex and education are based on.

22. Before any education-specific projections can be done on a global scale, a lot of time and effort will need to go into compiling data for the baseline: age-, sex- and education-specific participation rates for as many countries as possible and for as long a time-series as possible. Once this data is compiled, it could be used for the analysis of other aspects related to labor force participation as well. For example, analyses like the one performed by Gaddis and Klasen (2014) who investigat in depth the empirical evidence in favor of or against the often claimed U-shaped relationship between GDP per capita and female labor force participation could be done again, now disaggregating female labor force participation by education.

23. As I pointed out, there is nothing in principle wrong with calculating participation rates for open-ended age-groups, as for ages 15+, as long as the notion is to show what share of the population, men or women, above that age is economically active. However, as it happens in a lot of instances, the (aggregate) LFP rate of men and women ages 15+ is used as an indicator of the development of (age-specific) participation rates , and then it is not representing developments in an adequate way since shifts in the population structure towards older age-groups are distorting the picture. Consequently, more attention should be paid when comparing labor force participation over time or between countries, since the age-groups that are looked at can determine the outcome of the analysis. More attention should be given to underlying age-specific developments of participation. Population ageing will aggravate this issue, since more persons will occupy higher age-groups.

References

Australian Bureau of Statistics (2012). Differences between the LFS and the Census. Accessed March 20, 2016 http://www.abs.gov.au/websitedbs/censushome.nsf/home/factsheetslfsc?opendocument&navpos=450

Besamusca, J., Tijdens, K., Keune, M., and Steinmetz, S. 2015. Working Women Worldwide. Age Effects in Female Labor Force Participation in 117 Countries. *World Development*, 74: 123-141.

Gaddis, I. and Klasen, S. 2014 Economic development, structural change, and women's labor force participation: A reexamination of the feminization U hypothesis. *Journal of Population Economics*, 27(3): 639-681.

ILO (2013). *ILO Estimates and Projections of the Economically Active Population: 1990-2030 (2013 Edition). Methodological Description*. Geneva: ILO, Department of Statistics.

Houriet-Segard, G. and Pasteels, J.-M. 2012. *Projections of Economically Active Population. A Review of National and International Methodologies*. Geneva: ILO, Department of Statistics Working Paper Nr. 4.

Loichinger, E. 2015. Labor Force Projections up to 2053 for 26 EU countries by Age, Sex, and Highest Level of Educational Attainment. *Demographic Research*, 32(15): 443-486.

Lutz, W., Butz, W.P., and K.C., S., (eds.). 2014. *World Population and Human Capital in the 21st Century*. Oxford University Press.

Pena-Boquete, Y. 2016. Further developments in the dynamics of female labour force participation. *Empirical Economics*, 50: 463-501.

Verick, S. 2014. *Female labor force participation in developing countries*. IZA World of Labor 2014: 87.

UN. 2015. *World Population Prospects: The 2015 Revision*. United Nations, Department of Economic and Social Affairs, Population Division.

Yabiku, S. T. and Schlabach, S. 2009. Social Change and the Relationship Between Education and Employment. *Population Research and Policy Review* 28: 533-549.