Millennium Development Goals (MDGs): Measuring Within-Country Inequalities for Selected Indicators for South America using IPUMS-International Data (1990-2010)

Ana Cuesta[1] **and Rodrigo Lovatón**[2]

Draft as of March 24nd, 2014

The objective of this paper is to determine if achieving the Millennium Development Goals (MDGs) for a country has an impact on the geographical disparities, for some selected indicators, within the country. IPUMS-I data provide variables comparable between countries (while the sources traditionally used by the United Nations are not) and allow examining differences for lower levels of geography and between urban and rural settings. We measure indicators related to education, gender equality, and maternal health. For all of them, we will be primarily interested in disparities by gender across geographical units. Furthermore, we will track changes for these indicators for all censuses available since the 1990's round. Finally, we explore demographic factors related to higher disparities for these indicators.

1. Introduction

The Millennium Development Goals (MDGs) are a set of objectives that expresses the commitment to improve well-being for all persons around the world, especially of those residing in developing countries. This initiative started with the United Nations Millennium Declaration agreed by leaders of 189 nations in 2000. The MDGs include eight goals with over 40 targets that are expected to be met by 2015. A set of indicators were developed to assess the progress for each of these goals between 1990 and 2015.

[1] Minnesota Population Center, University of Minnesota, 225 19th Ave S # 50 Minneapolis, MN 55455, USA.
[2] Minnesota Population Center, University of Minnesota, 225 19th Ave S # 50 Minneapolis, MN 55455, USA.
Corresponding Author:
Rodrigo Lovatón., Minnesota Population Center, University of Minnesota, 225 19th Ave S # 50 Minneapolis, MN 55455, 612.624.4389, Email: lovat003@umn.edu

Phenomena such as poverty and hunger, education, health, and environmental sustainability are monitored through the evolution of these targets. This paper will focus on those indicators related to gender equality and development. Goal 3 aims to "*Promote gender equality and empower women*" and Goal 5 aims to "*Improve maternal health*". It is well known that women are participating more in the labor force and that girls are attending school more than decades ago, but the geography dimension has not been often taken into account and this progress might be unequal within countries.

These indicators have been traditionally measured at a country level. The recent global policy agenda acknowledges the importance of including different levels of geography in order to understand the degree of development of regions. The 2009 World Development Report recognizes the importance of including different dimensions of geography in the analysis of economic phenomena: density, distance, and division. Each of these dimensions contributes to narrow or broaden the development gaps not just between countries but within them.

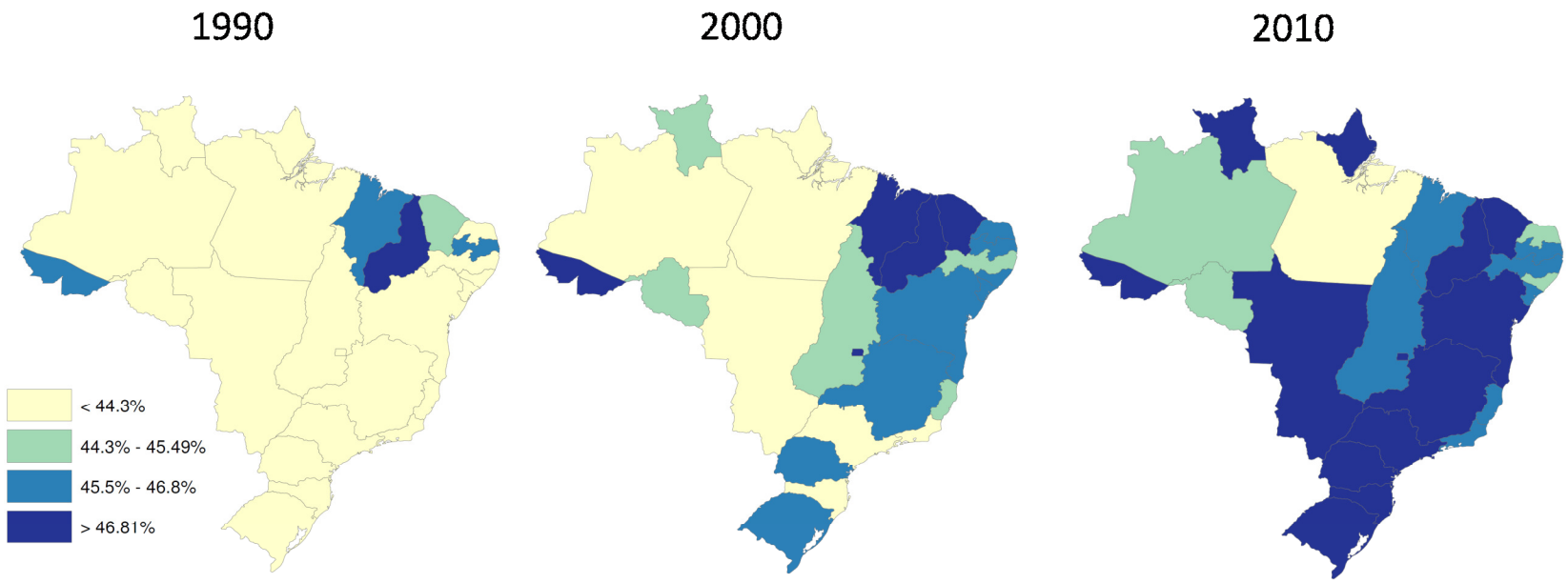
In order to illustrate the previous point, Figure 1 presents the evolution of the share of women in wage employment in the non-agricultural sector in Brazil between 1990 and 2010, which is one of the indicators within Goal 3. In the 1990's, in most of the country, women that were economically active had a participation rate of less than 45% in wage employment for the non-agricultural sector. It is clear, from Figure 1 that women's participation in the wage employment has increased dramatically in the past 20 years. However, even though the evolution of the indicator is different for each state, we observe dropping geography inequalities overall, both calculated through the range and the coefficient of variation measures. This finding requires further investigation by analyzing even smaller geographical units. But the bottom line is that incorporating the geography dimension into the analysis of the Millennium Development Goals (MDG) unveils potential within country disparities.

In this paper, we identify a set of MDG indicators feasible to be implemented with IPUMS-I census data. Our analytical approach will be to estimate some basic measures of inequality between geographical units using these indicators and explore potential demographic factors correlated to them. The primary interest relies on gender disparities by geography.

The document is organized as follows. Section 2 presents a review of the existent literature on inequality, emphasizing the studies using techniques which allow the analysis of inequality for lower levels of geography. Section 3 provides definitions for each of the selected indicators based on United Nations documentation and will describe how these could be estimated. In most cases, the indicator using IPUMS-I data is an approximation of the one proposed by United Nations, based on data availability. Additionally, it describes the methods used in order to compute the inequality measures. Section 4 describes the data that will be used to estimate the selected indicators. Finally, section 5 presents some preliminary results for the different inequality measures for each indicator and country.

Figure 1

Goal 3:Brazil, Share of women in wage employment in the non-agricultural sector (1990-2010)



Source: Author's calculations using Integrated Public Use Microdata Series (IPUMS) International.

2. Literature Review

There is an extended literature on the analysis and measures of inequality. Inequality has been traditionally examined using measures such as the Gini-coefficient of inequality, Theil index, and others, depending on the variable of interest. For example, the World Bank's Poverty Analysis and Equity group uses the decile dispersion ratio and the share of income/consumption of the poorest x% of the population as measures to analyze poverty.[3] In the context of the debate on inequality in education, Jacob and Holsinger (2009) report measures of an Education Gini Coefficient. In addition, Ibourk and Amaghouss (2012) use a Gini Index of education and standard deviation of schooling to conclude that in 2010 the Middle East and North Africa countries "the education distribution was more unequal in the middle-income countries than in the higher-income countries".

The analysis of inequalities has started to incorporate the geography dimension by analyzing geographic units at different levels. Barro and Sala-i-Martin (1991) reintroduced the concept of a region and convergence in the macroeconomic debate and, since then, research taking into account the geographical dimension has increased in the economics literature. The inequality literature has incorporated geography and has moved forward on the analysis by decomposing the Gini Coefficient and the Theil Index. Bellu and Liberati (2006) provide an accessible and step by step description of the decomposition of both the Gini Coefficient and the Theil Index.

Several authors have performed empirical illustrations of decompositions of inequality measures based on geographic variables. In research regarding income inequalities in the US, Rey (2006) uses a decomposition of the Theil index to analyze the importance of spatial dependence and scale when understanding the income inequality in the US from 1929 to 2000. The author decomposes the Theil index into the between and within-groups which correspond to Regions and States. Silva and Leichencko (2004) studied the impact of trade on income inequality across and within States in the US; in their study, they use a decomposed Theil index to estimate income inequality. Both papers provide a two geographic level analysis. Akita (2003) analyses the regional income inequality in China and Indonesia including an additional geography level (region, province, and district) using a two-stage nested Theil decomposition method. He found that the within-province component explained most of the regional inequality in China, but was not as determining in Indonesia.

The analysis of inequality by geography also includes research on its determinants. Peters (2012) identifies socioeconomic factors explaining income inequality in the U.S. His findings suggest that higher and growing inequality is related to both low-skill and high-skill services jobs as well as employment in the agricultural and industrial sectors. The author also concludes that the inequality outcomes could differ when using different geographic aggregations.

[3] For further reference: http://www.worldbank.org/en/topic/poverty

Most of the inequality research is focused on income inequalities and poverty. There is a branch of the literature which has addressed the inequalities in education research. The spatial dimension has been incorporated as well as the decomposition of the inequality measures to better understand what is happening between and within regions. There is a lack of studies which analyze inequalities in access to education and labor force participation from a gender perspective, especially at lower geographical levels. These are the main contributions of this study.

3. Methodology

In this section we will provide an overview of the development goals that will be analyzed and the different methods that will be used in measuring inequality. We selected the goals related to gender equality that are feasible to compute using IPUMS-I data. The selected inequality measures include some that will be used to describe the extent of inequality at a specific geography level (standardized range and coefficient of variation) and others that are suitable to be decomposed into between and within components and that are transversal to more than one geography level.

MDGs Indicators

As mentioned before, Goals 3 and 5 are the ones related to gender equality and maternal health. The indicators contained in these goals will let us understand the evolution of gender disparities by geography in the South American countries, covering areas such as education, work, and maternal health. Table 1 below presents the summary of the indicators which will be estimated. Appendix I describes how the indicators are computed using IPUMS-I data.

Table 1: Selected MDGs Indicators to measure gender equality using IPUMS-I

Goal	Indicators
3. Promote gender equality and empower women	3.1A. Ratio of girls to boys in primary, secondary, and tertiary education 3.1B. Ratio of literate women to men, 15-24 years old 3.2. Share of women in wage employment in the non-agricultural sector
5. Improve maternal health	5.4. Adolescent birth rate

Measuring Inequality

Our inequality estimations will use some traditional measures to examine differences across geographic units. A first set of measures will provide information of inequalities at each of the

different levels of geography; while a second group will allow for the decomposition of inequality by geography and to analyze differences in inequality between and within levels of geography. We take advantage of the availability of different geography levels in IPUMS-I data to examine the extent of inequalities across them.

In the first group of measures we include the standardized range and the coefficient of variation. The range is defined as the difference between the highest and lowest values for the estimated indicators. Our estimations will use the standardized range, which is the range divided by the mean. That is:

$$R = (\frac{1}{\mu})(X_{MAX} - X_{MIN}) \qquad (1)$$

where μ is the mean, X_{MAX} is the maximum, and X_{MIN} is the minimum value. This provides an initial approximation to the dispersion of the estimated gender indicators across geographical units.

The coefficient of variation is defined as the standard deviation of the indicator divided by the mean. That is:

$$CV = (\sigma/\mu) \qquad (2)$$

where μ is the mean and σ is the standard deviation. Even though this is a simple measure of inequality, it has desirable properties, including the Pigou-Dalton or transfer principle.[4]

The second group of inequality measures corresponds to those that make use of different levels of geography. In this family we find the various generalized entropy indexes, from which we will focus our analysis on the Theil index. Following Akita (2003), the Theil index could be decomposed into different components to include information on different geographic levels.[5] His model considers a three-level hierarchical structure for a country, where inequality is measured by a Theil index based on the lowest level of geography means. For instance, in the case of Brazil the index could be decomposed using the State – Mesoregion – Municipality structure and it would provide the following inequality components: between-state, between-mesoregion, and within-municipality. This same framework can be adapted to countries with only two levels of geography, such as the case of Colombia, whose structure comprises departments and municipalities.

[4] The four properties that any measure of inequality should satisfy are: anonymity, scale independence, population independence and transfer principle (also known as the Pigou-Dalton principle).

[5] On the data section we will list the different levels of geography of the South American samples. Brazil and Chile are the samples with the largest number of levels of geography (three levels). For the Brazilian case, those levels correspond to State-Mesoregion-Municipality and for the Chilean case they correspond to Region-Province-Municipality.

If we consider only two levels of geography, following Akita (2003), the Theil index could be decomposed as follows. For descriptive purposes, we will identify the first or higher level of geography as "department" and the second or lower level of geography as "municipality."

$$T_m = \sum_i \sum_j \left(\frac{Y_{ij}}{Y}\right) log\left(\frac{Y_{ij}/Y}{N_{ij}/N}\right) \qquad (3)$$

where Y_{ij} is the outcome of interest of municipality *j* in department *i*, Y is the total outcome of interest for all municipalities, N_{ij} is the population of municipality *j* in department *i,* and N is the total population for all municipalities. Equation (3) could be decomposed into:

$$T_m = T_{WD} + T_{BD} \qquad (4)$$

where the within-department component (T_{WD}) becomes:

$$T_{WD} = \sum_i \left(\frac{Y_i}{Y}\right) \sum_j \left(\frac{Y_{ij}}{Y}\right) log\left(\frac{Y_{ij}/Y}{N_{ij}/N}\right) \qquad (5)$$

and the between-department component (T_{BD}) is represented by:

$$T_{BD} = \sum_i \left(\frac{Y_i}{Y}\right) log\left(\frac{Y_i/Y}{N_i/N}\right) \qquad (6)$$

Additionally, in order to explore factors correlated to higher disparities, we will estimate models for each of the selected indicators and use as controls a vector of demographic factors averaged for each geographic unit, such as population density (per square kilometer), number of own family members in the household, proportion of female-headed households, and educational attainment for the household head. Furthermore, in order to have a proxy for socioeconomic status, we will calculate an asset-based wealth index for those samples with information available on assets, utilities, and housing characteristics. We will also explore the inclusion of fixed effects for higher-level geographic units, to control for other unobserved characteristics, and also the possibility of effects between neighboring geographic units.

4. Data

The IPUMS-International project is the largest database of census microdata from around the world, which currently includes 238 census samples (74 countries) from 1960 to present. The IPUMS-I project contains microdata that can be used to measure progress for some MDG indicators. An important advantage of census data over other sources is that progress can be measured not only for the country but also for smaller geographical units (usually up to two levels of geography for each country), thus providing richer information for analytical purposes.

Moreover, variables available through the IPUMS-I project are comparable across countries, unlike some data sources traditionally used to calculate the MDG indicators.

This paper focuses on the South America region, given that IPUMS-I has data for nine countries and for most of the latest census rounds. The MDGs measure the progress since 1990 and will be measured until 2015. Hence, in this paper, we classify census samples into three rounds: 1990s, 2000s, and 2010s. Table 2 presents the data available for the South American countries.

Table 2: Census data years for South American countries in IPUMS-I

Census round	1990-1999	2000-2009	2010
Argentina	1991	2001	2010
Bolivia	1992	2001	N.A
Brazil	1991	2000	2010
Chile	1992	2002	N.A
Colombia	1993	2005	N.A
Ecuador	1990	2001	2010
Peru	1993	2007	N.A
Uruguay	1996	2006 1/	N.A
Venezuela	1990	2001	N.A

Source: Integrated Public Use Microdata Series (IPUMS) International.

1/ Household survey. N.A=Not available.

Given the goals that will be considered in this paper, the analysis will use some demographic and education harmonized IPUMS-I variables widely available, such as: age and sex of the respondent (demographic), and school attendance, literacy and educational attainment (education). Additionally, we will include work variables for the economically active population, such as: class of worker and industry in which the respondent works. Finally, we considered a fertility variable reporting the total number of children ever born to a woman. The full description of the harmonized variables is available in Appendix II.

A key part of the analysis will be based on the different levels of geography available in the microdata. The IPUMS-I dataset includes a harmonized variable (Geolev1) which usually corresponds to the first subnational geographic level or major administrative unit in which the household was enumerated. Other lower geographic administrative divisions are also available through IPUMS-I and will be incorporated to the estimations. For example, Brazil includes three geographic levels: State (which is harmonized in GeoLevel1), Mesoregion, and Municipality. Appendix III describes the different geographic levels for the South American countries included in the analysis.

5. Preliminary results[6]

The analysis will be presented by goal as follows. The results will be analyzed using one or more geographic levels. First, the range and coefficient of variation will be reported and analyzed for the different geographic levels for each country. The tables with detailed results are included in Appendix IV. Second, the inequality decomposition by geography will be performed for each country. The evolution of the indicators will be analyzed in more detail for Argentina, Brazil, and Ecuador, since these countries have data available for all the Census rounds.

5.1. Ratio of girls to boys in primary, secondary, and tertiary education

Gender equality is almost achieved for most countries for the primary and secondary levels (ratios are very close to 1), which is associated with universal primary access and other advances in compulsory basic schooling in the region. Deviations from equality are larger for tertiary education. For example, the ratio of girls to boys is 0.955 in primary, 1.058 in secondary, and 1.323 in tertiary for the 2010 round for Argentina. This indicator shows that there are more boys than girls in primary while the opposite happens in secondary and tertiary, with the exception of Bolivia and Peru where there are more boys attending school at all three levels. Overall, the largest differences in the ratio of girls to boys are observed for tertiary education for Uruguay and Venezuela, followed by Argentina and Brazil.

The standardized range and coefficient of variation show that inequalities by geography are more often larger for tertiary compared to secondary and for secondary compared to primary (see tables in Appendix IV). For instance, the standardized range in the 2010 round for the higher geography level is 0.35 for tertiary, 0.18 for secondary, and 0.06 for primary for Argentina. The largest inequalities for this indicator are observed for Brazil and Colombia. As it is expected, the inequalities measured both by the standardized range and the coefficient of variation are larger for the lower geography level for all countries, about 3 to 10 or more times larger with respect to the higher geography units. All results are consistent across time, even though measured inequalities have been declining for more recent census rounds.

5.2. Ratio of literate women to men, 15-24 years old

The ratio of literate women to men 15 to 24 years old indicates a high level of equality by gender for the countries under analysis. The largest deviations from equality correspond to Bolivia, which had a ratio of 0.95 in the 1990 round and 0.97 in the 2000 round. Furthermore, literacy is generally higher among women, except for Bolivia, Peru, and the 1990 round for

[6] This section currently presents only preliminary results, mostly focused on the state and evolution of gender indicators and inequalities based on the standardized range and coefficient of variation. The objective is to extend the current analysis and include decompositions by geography and also demographic factors related to inequality, as stated in the methodology section. Updated results along with an updated version of the paper will be sent to the session discussant.

Ecuador. These results are translated into relatively smaller inequalities measured at the higher geography level, particularly for the most recent census rounds. In fact, even the largest value for the standardized range and coefficient of variation at the higher geography level, observed for the Brazil 1990 census round, are relatively small when compared to measured inequalities for other indicators. However, these inequalities are larger if we analyze the lower geography level and achieve a moderately high value for Bolivia, Brazil, and Colombia. In addition, we observe that in most cases the size of inequalities is smaller for the more recent census rounds.

5.3. Share of women in wage employment in the non-agricultural sector

The share of women in wage employment in the non-agricultural sector suggests varying degree of progress in the region, with values ranging mostly between 35% and 45% and with a clear increasing trend for the more recent census rounds. The lowest shares are observed for Bolivia, Chile, Ecuador, and Peru, although there is no data for the most recent census round for the first two countries. The largest inequalities for this indicator are observed for Bolivia, Ecuador, and Peru for the higher geography level, both considering the standardized range and coefficient of variation. Measured inequalities are larger at the lower geography level, as expected. Bolivia, Ecuador, and Peru remain with relatively high inequalities at this lower level, in addition to Chile and Colombia. Inequalities are consistently decreasing over time for all countries at all geography levels, with the exception of Uruguay which has larger inequalities in the 2010 with respect to the 2000 round.

5.4. Adolescent birth rate

The proportion of adolescent women that already had a child varies for the countries under analysis, ranging from about 10% to 20%. As it was previously mentioned, this indicator was approximated using information on children ever born, given that the question on births during the last year was not available for any of the selected countries. This may explain the relatively high rates observed. Furthermore, there is no clear pattern in the evolution of the adolescent birth rate over time: even though some countries have a decreasing trend (Colombia, Peru, and Venezuela), others seem to be increasing (Argentina, Bolivia, Chile, and Ecuador) or do not have a clear pattern. Inequalities for this indicator are relatively large for most countries (the largest among all indicators analyzed), are generally higher for the lower geography levels, and are surprisingly increasing over time for Argentina and Colombia while these are not decreasing significantly for other countries.

References

Akita, T.(2003). Decomposing Regional Income Inequality in China and Indonesia using Two-Stage Nested Theil Decomposition Method. *Annals of Regional Science*, Vol.37, No.1, pp.55-77.

Barro, Robert J. and Sala-i-Martin, X. (1991). Convergence across states and regions. *Brookings papers on Economic Activity*, 1:107-182.

Bellu, G. and Liberati, P. (2006). Decomposition of Income Inequality by Subgroups. *Policy Impacts on Inequality.* Model 052.

Cowell, Frank (1998). "Measurement of Inequality". London: London School of Economics and Political Science, Discussion paper No. DARP/36.

Haughton, Jonathan and Shahidur Khandker (2009). "Handbook on Poverty + Inequality." Washington D.C.: World Bank.

Holsinger, Donald and James Jacob, editors (2008). "Inequality in Education: Comparative and International Perspectives." Hong Kong: Comparative Education Research Centre-University of Hong Kong, CERC Studies in Comparative Education No. 24.

Ibourk, A. and Amaghouss, J. (2012). Measuring Education Inequalities: Concentration and Dispersion-Based Approach. *World Journal of Education.*Vol. 2, No. 6, pp 51-65.

Jacob, W. James, & Donald B. Holsinger. (2009). Inequality in Education: A Critical Analysis. In Donald B. Holsinger and W. James Jacob (Ed.), *Inequality in Education: Comparative and International Perspectives* (pp. 1–33). Dordrecht, The Netherlands: Springer/CERC.

Nolan, Brian; Salverda, Wiemer; and Timothy Smeeding (2011). "The Oxford Handbook of Economic Inequality". New York: Oxford University Press.

Peters, David. (2012). Income Inequality across Micro and Meso Geographic States in the Midwestern United States, 1979-2009. *Rural Sociology.* 77 (2): 171-202.

Silva, J. and R. Leichenko. (2004). Regional Income Inequality and International Trade. Economic Geography 80: 261-286.

Rey, S. J. (2006). Spatial Analysis of Regional Income Inequality. In *Spatially Integrated Social Science*, 280–99, edited by M. F. Goodchild and D. Janelle. Oxford, UK: Oxford University Press.

United Nations (2003). "Indicators for Monitoring the Millennium Development Goals: Definitions, Rationale, Concepts, and Sources." New York: United Nations.

World Bank (2009). "World Development Report: Reshaping Economic Geography". Washington D.C.: World Bank.

Appendix I: Selected MDGs Indicators to measure gender equality using IPUMS-I

This section describes how the indicators are computed using IPUMS-I data. The variable names in capital letters correspond to the integrated variables from IPUMS-I that would be necessary for the estimation of the selected indicators. The description includes treatment of special values (unknown and not in universe) and specific formulas which define the indicators.

Goal 3: Promote gender equality and empower women
Indicator: Ratio of girls to boys in primary, secondary, and tertiary education

> *U.N. Definition*: "Ratio of girls to boys in primary, secondary and tertiary education is the ratio of the number of female students enrolled at primary, secondary and tertiary levels in public and private schools to the number of male students." (United Nations, 2003)
>
> *IPUMS-I Operationalization*: Ratio of girls to boys (SEX=2/SEX=1) who are currently attending school (SCHOOL=1) and that have not completed primary (less than primary completed or EDATTAN=1), secondary (primary complete or EDATTAN=2), or tertiary (secondary complete or EDATTAN=3). Persons with unknown school attendance (SCHOOL=9) or educational attainment (EDATTAN=9) or outside the universe for the questions of school attendance (SCHOOL=0) or educational attainment (EDATTAN=0) are not considered in the calculation. The proportion of unknown cases for these integrated variables is small and the education census questions typically include all persons in school age for primary, secondary, or tertiary, so these should not affect the results.

Ratio of girls to boys in primary:

$$Formula = \frac{\text{Girls (SEX = 2) currently attending school (SCHOOL = 1) that have not completed primary (EDATTAN = 1)}}{\text{Boys (SEX = 1) currently attending school (SCHOOL = 1) that have not completed primary (EDATTAN = 1)}}$$

Ratio of girls to boys in secondary:

$$Formula = \frac{\text{Girls (SEX = 2) currently attending school (SCHOOL = 1) that have not completed secondary (EDATTAN = 2)}}{\text{Boys (SEX = 1) currently attending school (SCHOOL = 1) that have not completed secondary (EDATTAN = 2)}}$$

Ratio of girls to boys in tertiary:

$$Formula = \frac{\text{Girls (SEX = 2) currently attending school (SCHOOL = 1) that have not completed teriary (EDATTAN = 3)}}{\text{Boys (SEX = 1) currently attending school (SCHOOL = 1) that have not completed tertiary (EDATTAN = 3)}}$$

> *IPUMS-I Integrated variables*: SEX, SCHOOL, and EDATTAN.

Goal 3: Promote gender equality and empower women
Indicator: Ratio of literate women to men, 15-24 years old

U.N. Definition: "The ratio of literate women to men, 15–24 years old (literacy gender parity index) is the ratio of the female literacy rate to the male literacy rate for the age group 15–24." (United Nations, 2003)

IPUMS-I Operationalization: The literacy rates are defined as in the previous literacy indicator, but in this case it is necessary to define a ratio based on the person's gender. This indicator is calculated as the ratio of women's to men's (SEX=2/SEX=1) literacy rate (LIT=2) for ages 15-24 (AGE≥15 and AGE≤24). Similarly, persons with unknown literacy (LIT=9), sex (SEX=9), or age (AGE=999) or outside the universe for the literacy question (LIT=0) are not considered in the calculation. The proportion of unknown cases for these integrated variables is small and the question for literacy always includes persons in the relevant age range (15 to 24 years old), so these should not affect the results.

$$Formula = \frac{\dfrac{\text{Women (SEX = 2) that are literate (LIT = 2) and ages 15 - 24 (AGE} \geq 15 \text{ \& AGE} \leq 24)}{\text{Women (SEX = 2) ages 15 - 24 (AGE} \geq 15 \text{ \& AGE} \leq 24)}}{\dfrac{\text{Men (SEX = 1) that are literate (LIT = 2) and ages 15 - 24 (AGE} \geq 15 \text{ \& AGE} \leq 24)}{\text{Men (SEX = 1) ages 15 - 24 (AGE} \geq 15 \text{ \& AGE} \leq 24)}}$$

IPUMS-I Integrated variables: AGE, SEX, and LIT.

Goal 3: Promote gender equality and empower women
Indicator: Share of women in wage employment in the non-agricultural sector

U.N. Definition: "The share of women in wage employment in the non-agricultural sector is the share of female workers in the non-agricultural sector expressed as a percentage of total employment in the sector." (United Nations, 2003)

IPUMS-I Operationalization: Proportion of female workers (SEX=2) in the non-agricultural sector (INDGEN≥20 and INDGEN≤130) that are in wage employment (CLASSWK=2). The IPUMS-I industry general recode (INDGEN) includes agriculture, fishing, and forestry in the same category, so this is an approximate figure. That is, fishing and forestry are also excluded from the "non-agricultural" sector.

$$Formula = \frac{\text{Female workers (SEX = 2) in the non - agricultural sector (INDGEN} \geq 20 \text{ \& INDGEN} \leq 130) \text{ and in wage employment (}CLASSWK = 2)}{\text{Persons in the non - agricultural sector (INDGEN} \geq 20 \text{ \& INDGEN} \leq 130) \text{ and in wage employment (}CLASSWK = 2)}$$

IPUMS-I Integrated variables: SEX, CLASSWK, and INDGEN.

Goal 5: Improve maternal health
Indicator: Adolescent birth rate

U.N. Definition: "The adolescent birth rate measures the annual number of births to women 15 to 19 years of age per 1,000 women in that age group. It is also referred to as the age-specific fertility rate for women aged 15-19." (United Nations, metadata)

IPUMS-I Operationalization: Proportion of women (SEX = 2) ages 15 to 19 (AGE≥15 and AGE≤19) who have ever had a live birth (CHBORN ≥1).[7] Women with unknown number of live births (CHBORN=98) or outside the universe for the fertility question (CHBORN=99) are not considered in the calculation. The proportion of unknown cases for these integrated variables is small so these should not affect the results.

$$Formula = \frac{\text{Women who had live births last year (CHBORN} \geq 1\text{) and ages 15 - 19 (AGE} \geq 15 \text{ \& AGE} \leq 19\text{)}}{\text{Women ages 15 - 19 (AGE} \geq 15 \text{ \& AGE} \leq 19\text{)}}$$

IPUMS-I Integrated variables: CHBORN, AGE and SEX.

[7] The UN definition uses the number of births last year, unfortunately this variable is not available in most of the South American IPUMS-I samples.

Appendix II: IPUMS-I variables used in the analysis

The name of the variables correspond to the harmonized IPUMS-I variables available in the website.[8] These set of variables are person level ones.

Demographic variables

- SEX reports the sex (gender) of the respondent.
- AGE gives age in years as of the person's last birthday prior to or on the day of enumeration.

Fertility variables

- CHBORN indicates the number of children ever born to a woman. Only live births are counted

Education variables

- SCHOOL indicates whether or not the person attended school at the time of the census or within some specified period of time prior to the census.
- LIT indicates whether or not the respondent could read and write in any language. A person is typically considered literate if he or she can both read and write. All other persons are illiterate; including those who can either read or write but cannot do both.
- EDATTAN records the person's educational attainment in terms of the level of schooling completed (degree or other milestone). The emphasis on level completed is critical: a person attending the final year of secondary education receives the code for having completed lower secondary only -- and in some samples only primary.

Work variables

- INDGEN recodes the industrial classifications of the various samples into twelve groups that can be fairly consistently identified across all available samples. The groupings roughly conform to the International Standard Industrial Classification (ISIC). The third digit of INDGEN retains important detail among the service industries that could not be consistently distinguished in all samples.
- CLASSWK refers to the status of an economically active person with respect to his or her employment -- that is, the type of explicit or implicit contract of employment with other persons or organizations that the person has in his/her job. In general, the variable indicates whether a person was self-employed, or worked for someone else, either for pay or as an unpaid family worker.

[8] Source: https://international.ipums.org/international/

Appendix III: Geography levels for the countries analyzed

	Level 1[a]		Level 2		Level 3	
Argentina	Province	24	Department	309	-	
Bolivia	Department	9	Province	84	-	
Brazil	State	25	Mesoregion	159	Municipality	1,524
Chile	Region	9	Province	44	Municipality	178
Colombia	Department	25	Municipality	532	-	
Ecuador	Province	20	Canton	141	-	
Peru	Region	25	Province	176	-	
Uruguay	Department	19	-	-	-	
Venezuela	State	23	Municipality	243	-	

a: The number of provinces, departments, regions or states may differ from the offical major administrative areas given that some of the units were combined because of confidentiality.

Appendix IV: Results for the Range and Coefficient of Variation

Argentina

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	0.9556	0.9172	0.9767	0.0627	0.0003	0.6963	1.1921	0.5200	0.0046
Round 2000	0.9700	0.9104	1.0142	0.1074	0.0006	0.7500	1.2174	0.4828	0.0049
Round 1990	0.9685	0.9222	1.0057	0.0865	0.0006	0.7669	1.2308	0.4783	0.0049
Ratio of girls to boys in secondary education									
Round 2010	1.0586	0.9702	1.1647	0.1826	0.0018	0.8285	1.5043	0.6213	0.0113
Round 2000	1.0374	0.9754	1.1507	0.1674	0.0021	0.7931	1.8421	0.9909	0.0143
Round 1990	1.0601	0.9601	1.1730	0.1994	0.0019	0.7452	1.6122	0.7855	0.0153
Ratio of girls to boys in tertiary education									
Round 2010	1.3230	1.1825	1.6750	0.3492	0.0101	0.8364	2.8571	1.4007	0.0449
Round 2000	1.3582	1.1143	1.7892	0.4666	0.0186	0.5000	3.4000	1.8225	0.1198
Round 1990	1.2447	1.0789	1.9041	0.5947	0.0267	0.0000	14.0000	7.4894	0.7695
Ratio of literate women to men, 15-24 years old									
Round 2010	1.0038	1.0003	1.0111	0.0108	0.0000	0.9820	1.0492	0.0669	0.0001
Round 2000	1.0040	0.9975	1.0157	0.0181	0.0000	0.9232	1.0502	0.1264	0.0001
Round 1990	1.0050	0.9961	1.0238	0.0276	0.0000	0.9398	1.0737	0.1329	0.0002
Share of women in wage employment in the non-agricultural sector									
Round 2010	-								
Round 2000	0.4403	0.3920	0.5041	0.2552	0.0018	0.3097	0.5539	0.5590	0.0049
Round 1990	0.4192	0.3649	0.4769	0.2656	0.0021	0.1846	0.6456	1.1001	0.0094
Adolescent birth rate									
Round 2010	-								
Round 2000	0.1250	0.0500	0.2043	1.1026	0.0075	0.0179	0.3357	2.1829	0.0174
Round 1990	0.1194	0.0401	0.1877	1.0356	0.0083	0.0338	0.3140	1.8812	0.0167

Bolivia

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	-								
Round 2000	0.9743	0.9345	1.0014	0.0694	0.0007	0.8131	1.1618	0.3625	0.0051
Round 1990	0.9471	0.8838	1.0144	0.1381	0.0015	0.6281	1.4220	0.8566	0.0161
Ratio of girls to boys in secondary education									
Round 2010	-								
Round 2000	0.8951	0.8276	0.9891	0.1805	0.0033	0.2895	1.2308	1.1905	0.0369
Round 1990	0.8951	0.8017	1.0703	0.2927	0.0095	0.1316	1.1846	1.3902	0.0666
Ratio of girls to boys in tertiary education									
Round 2010	-								
Round 2000	-								
Round 1990	0.8010	0.6500	0.9935	0.4148	0.0136	0.0000	2.6667	3.5149	0.4157
Ratio of literate women to men, 15-24 years old									
Round 2010	-								
Round 2000	0.9772	0.9205	1.0080	0.0897	0.0007	0.6653	1.0156	0.3668	0.0033
Round 1990	0.9521	0.8743	0.9789	0.1110	0.0015	0.6005	1.0109	0.4541	0.0106
Share of women in wage employment in the non-agricultural sector									
Round 2010	-								
Round 2000	0.3859	0.2935	0.4457	0.4035	0.0048	0.1361	0.4965	1.0788	0.0192
Round 1990	0.3320	0.2524	0.4078	0.4670	0.0064	0.1250	0.5700	1.5276	0.0363
Adolescent birth rate									
Round 2010	-								
Round 2000	0.1891	0.1180	0.4384	1.3793	0.0582	0.0947	0.5510	1.6681	0.0481
Round 1990	0.1708	0.1088	0.4833	1.7621	0.0672	0.0476	0.5818	2.1033	0.0537

Brazil

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	0.9410	0.9013	0.9631	0.0661	0.0002	0.6457	1.2907	0.6888	0.0080
Round 2000	0.9382	0.9086	0.9742	0.0700	0.0004	0.6685	1.2905	0.6627	0.0079
Round 1990	0.9957	0.8756	1.1454	0.2666	0.0054	0.6019	2.1620	1.5364	0.0219
Ratio of girls to boys in secondary education									
Round 2010	1.1230	1.0267	1.3173	0.2526	0.0059	0.6808	2.6420	1.6650	0.0461
Round 2000	1.1437	1.0171	1.4378	0.3432	0.0127	0.6268	3.2268	2.0693	0.0810
Round 1990	1.2669	1.0691	1.6315	0.4206	0.0172	0.3575	11.1897	7.0329	0.4317
Ratio of girls to boys in tertiary education									
Round 2010	1.3129	1.1241	1.7390	0.4305	0.0177	0.0000	21.2345	11.0770	1.2358
Round 2000	1.3205	1.1538	1.7645	0.4219	0.0201	0.0000	24.8501	12.8423	1.5255
Round 1990	1.1794	0.8003	1.7117	0.7408	0.0417	0.0000	19.0612	11.2386	1.8235
Ratio of literate women to men, 15-24 years old									
Round 2010	1.0154	0.9982	1.0491	0.0500	0.0003	0.9079	1.2140	0.2992	0.0011
Round 2000	1.0303	1.0037	1.1221	0.1138	0.0016	0.8205	1.8204	0.9538	0.0049
Round 1990	1.0565	0.9997	1.2723	0.2512	0.0067	0.8642	2.4115	1.3925	0.0247
Share of women in wage employment in the non-agricultural sector									
Round 2010	0.4703	0.4433	0.4882	0.0963	0.0003	0.2698	0.6728	0.8576	0.0044
Round 2000	0.4491	0.4217	0.4784	0.1249	0.0005	0.1465	0.7071	1.2295	0.0076
Round 1990	0.4035	0.3783	0.4776	0.2349	0.0018	0.1751	0.8153	1.5047	0.0182
Adolescent birth rate									
Round 2010	0.1181	0.0838	0.2060	0.9087	0.0095	0.0159	0.4815	3.5840	0.0248
Round 2000	0.1480	0.1182	0.2709	0.9091	0.0093	0.0268	0.4195	2.4205	0.0202
Round 1990	0.1255	0.0965	0.3175	1.4560	0.0181	0.0000	0.4443	3.1107	0.0281

Chile

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of literate women to men, 15-24 years old									
Round 2010	-								
Round 2000	1.0044	0.9982	1.0101	0.0119	0.0000	0.9751	1.0328	0.0575	0.0001
Round 1990	1.0059	0.9978	1.0162	0.0183	0.0000	0.9728	1.0649	0.0914	0.0002
Share of women in wage employment in the non-agricultural sector									
Round 2010	-								
Round 2000	0.3925	0.3230	0.4056	0.2200	0.0023	0.1574	0.6285	1.2630	0.0093
Round 1990	0.3587	0.2914	0.3734	0.2432	0.0036	0.1423	0.6697	1.5612	0.0151
Adolescent birth rate									
Round 2010	-								
Round 2000	0.1544	0.1379	0.1881	0.3108	0.0023	0.0212	0.3537	1.9999	0.0154
Round 1990	0.1477	0.1051	0.1768	0.4706	0.0036	0.0294	0.3867	2.1434	0.0175

Colombia

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	0.9387	0.8870	1.0326	0.1549	0.0009	0.7230	1.2705	0.5911	0.0053
Round 2000	-								
Round 1990	0.9688	0.9439	1.1894	0.2494	0.0031	0.5000	1.6140	1.1537	0.0133
Ratio of girls to boys in secondary education									
Round 2010	1.0479	0.8849	1.1076	0.2138	0.0029	0.7078	1.5809	0.8162	0.0145
Round 2000	-								
Round 1990	1.1227	1.0000	1.2592	0.2289	0.0039	0.5489	2.0000	1.2744	0.0371
Ratio of girls to boys in tertiary education									
Round 2010	1.2292	0.7265	1.6941	0.7737	0.0220	0.2778	6.0000	3.9780	0.2083
Round 2000	-								
Round 1990	1.2121	0.7778	1.5250	0.6093	0.0183	0.0000	13.0000	8.5725	0.7153
Ratio of literate women to men, 15-24 years old									
Round 2010	1.0141	0.9816	1.0386	0.0563	0.0001	0.8397	1.1480	0.3023	0.0008
Round 2000	-								
Round 1990	1.0168	0.9811	1.0539	0.0720	0.0002	0.9366	1.1833	0.2414	0.0011
Share of women in wage employment in the non-agricultural sector									
Round 2010	0.4415	0.3756	0.5064	0.3045	0.0019	0.0690	1.0000	2.2588	0.0291
Round 2000	-								
Round 1990	0.4106	0.3435	0.4766	0.3302	0.0027	0.1525	0.7391	1.4838	0.0181
Adolescent birth rate									
Round 2010	0.1513	0.1111	0.2816	0.9618	0.0125	0.0389	0.4769	2.4047	0.0238
Round 2000	-								
Round 1990	0.2041	0.1615	0.3566	0.8360	0.0137	0.0000	0.5273	2.3260	0.0327

Ecuador

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	0.9821	0.8999	1.1085	0.2124	0.0018	0.7847	1.2326	0.4589	0.0061
Round 2000	0.9762	0.9193	1.0093	0.0924	0.0004	0.8176	1.2367	0.4297	0.0052
Round 1990	0.9719	0.8732	1.0294	0.1619	0.0016	0.6667	1.1686	0.5220	0.0074
Ratio of girls to boys in secondary education									
Round 2010	0.9962	0.8592	1.1278	0.2711	0.0047	0.7075	1.3768	0.6657	0.0136
Round 2000	1.0112	0.8259	1.1537	0.3325	0.0055	0.5088	1.5204	1.0027	0.0259
Round 1990	1.0601	0.8532	1.2003	0.3363	0.0078	0.6909	1.4571	0.7366	0.0227
Ratio of girls to boys in tertiary education									
Round 2010	1.2223	1.0187	2.1429	0.8340	0.0548	0.4615	3.1000	1.8540	0.1276
Round 2000	1.1455	0.8254	1.4858	0.5841	0.0305	0.2143	4.0000	3.1661	0.1903
Round 1990	1.0734	0.6029	1.5938	0.8809	0.0590	0.0000	3.0000	2.5420	0.2318
Ratio of literate women to men, 15-24 years old									
Round 2010	1.0028	0.9854	1.0220	0.0365	0.0001	0.9572	1.0425	0.0850	0.0002
Round 2000	1.0012	0.9726	1.0211	0.0486	0.0001	0.8632	1.0762	0.2124	0.0007
Round 1990	0.9906	0.9337	1.0070	0.0748	0.0005	0.8304	1.0694	0.2431	0.0014
Share of women in wage employment in the non-agricultural sector									
Round 2010	0.3936	0.3130	0.4297	0.3026	0.0026	0.1614	0.5615	1.0590	0.0095
Round 2000	0.3804	0.2488	0.4274	0.4820	0.0050	0.1500	0.5660	1.1411	0.0125
Round 1990	0.3431	0.1900	0.4189	0.7153	0.0097	0.1273	0.5800	1.3696	0.0186
Adolescent birth rate									
Round 2010	0.1761	0.1261	0.3132	0.9872	0.0145	0.0799	0.3636	1.3848	0.0194
Round 2000	0.1677	0.0993	0.2744	0.9940	0.0143	0.0709	0.3688	1.5440	0.0214
Round 1990	0.1490	0.1049	0.2644	0.9781	0.0148	0.0667	0.4659	2.1972	0.0267

Peru

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	0.9658	0.9294	0.9856	0.0584	0.0003	0.8182	1.1958	0.3916	0.0039
Round 2000	-								
Round 1990	0.9608	0.9124	1.0061	0.0980	0.0009	0.7488	1.1784	0.4566	0.0051
Ratio of girls to boys in secondary education									
Round 2010	0.9445	0.8226	1.0866	0.2825	0.0035	0.6233	1.1475	0.5831	0.0124
Round 2000	-								
Round 1990	0.9292	0.6744	1.0359	0.4067	0.0111	0.3594	1.1575	0.9959	0.0372
Ratio of girls to boys in tertiary education									
Round 2010	0.9798	0.7611	1.1350	0.3940	0.0101	0.2093	1.4898	1.4998	0.0550
Round 2000	-								
Round 1990	0.9945	0.6403	1.1880	0.5656	0.0162	0.2308	2.6000	2.5385	0.1214
Ratio of literate women to men, 15-24 years old									
Round 2010	0.9908	0.9627	1.0008	0.0386	0.0001	0.8624	1.0192	0.1602	0.0008
Round 2000	-								
Round 1990	0.9672	0.8535	0.9953	0.1486	0.0020	0.6590	1.0052	0.3733	0.0061
Share of women in wage employment in the non-agricultural sector									
Round 2010	0.4006	0.2841	0.4245	0.3740	0.0033	0.1016	0.5051	1.1604	0.0181
Round 2000	-								
Round 1990	0.3430	0.2357	0.3965	0.4941	0.0044	0.0939	0.6230	1.6182	0.0220
Adolescent birth rate									
Round 2010	0.1168	0.0690	0.2394	1.2465	0.0185	0.0495	0.4259	2.4417	0.0318
Round 2000	-								
Round 1990	0.1209	0.0734	0.3119	1.5203	0.0333	0.0329	0.5037	2.5825	0.0416

Uruguay

	National	Geo Level 1			
Round	Average	Min.	Max.	Range	CV
Ratio of girls to boys in primary education					
Round 2010	0.9479	0.8518	1.0846	0.2486	0.0029
Round 2000	0.9669	0.8646	1.1051	0.2466	0.0048
Round 1990	0.9419	0.8142	1.0055	0.2040	0.0025
Ratio of girls to boys in secondary education					
Round 2010	1.0995	0.9611	1.3170	0.3080	0.0074
Round 2000	1.2119	1.1047	1.4533	0.2771	0.0089
Round 1990	1.3307	1.0500	1.7767	0.5288	0.0274
Ratio of girls to boys in tertiary education					
Round 2010	1.5936	1.3984	2.8261	0.7382	0.0823
Round 2000	1.5862	0.5714	4.1111	1.5764	0.2695
Round 1990	0.8599	0.5944	2.0000	1.5474	0.1126
Ratio of literate women to men, 15-24 years old					
Round 2010	1.0082	0.9962	1.0170	0.0207	0.0000
Round 2000	1.0086	0.9995	1.0286	0.0287	0.0001
Round 1990	1.0078	0.9968	1.0333	0.0361	0.0001
hare of women in wage employment in the non-agricultural secto					
Round 2010	0.4843	0.4257	0.5294	0.2193	0.0011
Round 2000	0.4314	0.3759	0.4568	0.1975	0.0008
Round 1990					
Adolescent birth rate					
Round 2010	0.0791	0.0560	0.1595	1.1085	0.0067
Round 2000	0.1388	0.1051	0.2342	0.7856	0.0052
Round 1990	0.0872	0.0380	0.1673	1.2620	0.0077

Venezuela

	National	Geo Level 1				Geo Level 2			
Round	Average	Min.	Max.	Range	CV	Min.	Max	Range	CV
Ratio of girls to boys in primary education									
Round 2010	-								
Round 2000	0.9490	0.9260	0.9751	0.0516	0.0002	0.7438	1.1339	0.4144	0.0050
Round 1990	0.9576	0.8753	1.0254	0.1572	0.0011	0.7555	1.2860	0.5580	0.0093
Ratio of girls to boys in secondary education									
Round 2010	-								
Round 2000	1.0974	1.0227	1.2024	0.1611	0.0028	0.6100	1.5795	0.8558	0.0203
Round 1990	1.1719	1.0796	1.3412	0.2170	0.0040	0.6533	2.1712	1.2377	0.0383
Ratio of girls to boys in tertiary education									
Round 2010	-								
Round 2000	1.5294	1.3803	1.9804	0.3662	0.0187	0.4571	4.1667	2.0146	0.1318
Round 1990	1.3623	1.2243	1.8533	0.4306	0.0259	0.2500	8.5000	4.6054	0.5408
Ratio of literate women to men, 15-24 years old									
Round 2010	-								
Round 2000	1.0179	1.0048	1.0469	0.0412	0.0001	0.9324	1.1349	0.1977	0.0006
Round 1990	1.0176	0.9399	1.0490	0.1071	0.0004	0.9399	1.1480	0.2019	0.0011
Share of women in wage employment in the non-agricultural sector									
Round 2010	-								
Round 2000	0.4174	0.3731	0.4907	0.2773	0.0020	0.2712	0.5541	0.6970	0.0077
Round 1990	0.3799	0.3295	0.4402	0.2912	0.0024	0.1676	0.6000	1.1961	0.0124
Adolescent birth rate									
Round 2010	-								
Round 2000	0.1575	0.1236	0.2718	0.8759	0.0092	0.0365	0.4227	2.1634	0.0225
Round 1990	0.1740	0.1386	0.3096	0.8702	0.0130	0.0128	0.4404	2.0566	0.0254