Using Census to Estimate Mortality Differentials: an Application to Rural-urban Adult Mortality in Colombia in 2018

José H C Monteiro da Silva*

Population Studies Center, University of Pennsylvania*

November 1, 2023

(Do not cite without author's permission)

Abstract

Urban-rural mortality differentials in Low and Middle Income countries have been broadly investigated for younger ages, but not as much for adults. Most studies conducted so far on urban-rural adult mortality, however, are comparative analyses that do not investigate socioeconomic and regional disparities within-countries. Lack of reliable good quality data is one of the main reasons for the few studies in Latin America, in particular. The present study aims to fill this gap for Colombia, the second largest country in population size of South America, and a country severally impacted by a long-term internal armed conflict. By using data from the 2018 national censuses, I could estimate death rates by sex, region, area of residence, and socioeconomic deprivation status of the household (using the unmet basic needs index). I show that, in general, there is a rural mortality penalty in Colombia, which is more pronounced in regions with higher levels of socioeconomic deprivation (*Pacífico* and *Amazonía*). Nevertheless, socioeconomic disparities account for about 20-25% of the differences observed between male and female adult mortality rates in urban and rural areas.

1 Introduction

There has been a growing effort over the last years to document urban-rural adult mortality differentials in Low and Middle Income countries (LMCIs) (Menashe-Oren and Stecklov, 2018; Menashe-Oren and Masquelier, 2022). The majority of the earlier work published is dedicated to the assessment of mortality in the initial stages of life (infant and under-five mortality) (Fink et al., 2016; Garcia, 2020; Kumar et al., 2022; Van De Poel et al., 2009; Bocquier et al., 2011; Moultrie et al., 2013). Overall, these studies report an urban mortality advantage for younger ages (Garcia, 2020; Van De Poel et al., 2009). However, a similar advantage is not always present among adults (Menashe-Oren and Masquelier, 2022; Menashe-Oren and Stecklov, 2018). For example, in Latin America, studies have reported an urban adult mortality advantage in Bolivia and Peru (Menashe-Oren and Masquelier, 2022), but a rural advantage in Brazil (Albuquerque, 2019).

Despite these recent contributions, urban-rural adult mortality differentials have not been studied in depth in Latin America. These studies mostly report the adult morality differentials at the national level and do not discuss the role of within-country socioeconomic and regional disparities. Taking into account the heterogeneity of mortality and epidemiological transition profiles that characterizes Latin America (Frenk et al., 1991; Palloni and Pinto-Aguirre, 2011), it is likely that the mechanisms behind the urban-rural adult mortality differentials differ across and within countries.

For instance, Colombia is a case for which not much is known regarding urban-rural mortality differentials. Recent findings have documented an infant mortality advantage in urban areas (Garcia, 2020), but a not significant urban advantage for adult mortality Menashe-Oren and Masquelier (2022). However, these two studies focus on cross-country analyses and do not go in depth into assessing within country disparities. There are several reasons for studying the country in respect to urban-rural adult mortality differentials. First, Colombia is home of approximately 51 million inhabitants, according to the recent estimates from the World Population Prospects (United Nations, 2022), which places it as the second largest country (in population size) of South America, just behind Brazil. Second, the country has been exposed to an internal armed conflict involving political parties, drug traffickers, *guerrillas*, the military, and paramilitary groups for more than 70 years (LeGrand, 2003). Despite harming both urban and rural residents of the country, the conflict hot spots are mostly concentrated in rural municipalities in the countryside of Colombia. As a consequence of this long-term conflict, these areas experience population displacements, lower access to health services, worsened health outcomes, and higher homicide rates (especially among the male population) (Franco et al., 2006; Vital Strategies, 2020; Urdinola, 2004; Urdinola-Contreras, 2018; Urdinola et al., 2016; Vallejo et al., 2018; Hinestroza et al., 2021). In this sense, documenting and understanding the mechanisms behind these urban-rural disparities within the country gain relevance.

The lack of good quality data can be mentioned as one of the main reasons for having few studies of urban-rural adult mortality differentials in Colombia and in Latin America. Civil registration and vital statistics (CRVS) systems are the gold standard data collection for assessing vital events, but only few countries (Argentina, Chile, Costa Rica, Cuba, and Uruguay) in the region have reached high standards in terms of completeness and quality of mortality data (Fabiana Del Popolo and Guiomar Bay, 2021; Palloni and Pinto-Aguirre, 2011). Even though there have been significant improvements over the last decades, the quality and completeness of death registration in Colombia are heterogeneous across regions (Vital Strategies, 2020; Urdinola et al., 2016). Therefore, much of the adult mortality knowledge relies on estimates from household surveys or on adjustments using demographic methods (Palloni and Pinto-Aguirre, 2011). In this regard, retrospective questions on household deaths in national censuses offer an opportunity for studying sub-national level adult mortality, especially in contexts where CRVS systems are affected by coverage or completeness issues (Queiroz and Sacco, 2018; Queiroz and Sawyer, 2012; Borges, 2018; United Nations, 2017). In addition, these inquiries have the advantage of collecting socioeconomic status (SES), person-years of exposure and death counts information from the same data source, reducing biases that emerge when estimating mortality rates with denominators and numerators from different sources (Queiroz and Sacco, 2018).

This paper has the intention to address the existing gaps in the literature regarding urban-rural mortality differentials in Colombia using household deaths from the 2018 census. I show that the census mortality data is likely to be more reliable in rural areas than the CRVS data. I estimate age-standardized deaths rates (ASDR) for adults (ages 15-59) by region and area of residence, using both census and CRVS (adjusted and unadjusted) deaths to support this argument. Further, I use the index of Unsatisfied Basic Needs (UBN) at the household level (Mancero and Feres, 2001) to estimate SES through a degree of deprivation of rural and urban areas and, then I decompose rural-urban differences in ASDR in terms of deprivation and rate effects (Kitagawa, 1955). Finally, I estimate Poisson log-linear models for checking the robustness of the decomposition estimates.

The results show that rural adult mortality is consistently higher than urban adult mortality in Colombia across all regions. The highest differences were observed in the *Pacifico* and in the *Amazonia* regions. Further, I estimate that only about 20-25% of the differences between rural and urban ASDR (for both men and women) can be explained by differences in the composition of the populations relative to the UBN index across these two areas. Hence, most of the differences in adult mortality between rural and urban areas of the country are a result of factors not captured by differences between regions and unmet basic needs.

2 Background

2.1 Urban-rural mortality differentials: what do we know?

During the advent of mortality transition, High Income Countries (HICs) used to experience higher mortality rates in urban areas than in rural areas (Woods, 2003). Since public health infrastructure (such as access to water, sanitation and sewage systems) were not broadly available in cities and towns, the higher population density brought by the rapid urbanization process facilitated the spread of infectious diseases among urban areas (Horiuchi, 1999; Reher, 2001). Nevertheless, urban areas were the first ones to benefit from mortality decrease due to later advances in public health (access to clean water and sanitation systems, and spread of personal hygiene practices) and in medical technologies (vaccines and antibiotics) (Vallin and Meslé, 2004; Cutler and Miller, 2005; Easterlin, 2004; Davis, 1965).

Recent findings from HICs have shown that differentials in urban-rural mortality are heterogeneous, and the advantage varies according to local contexts, sex, socioeconomic status, and age-groups. For instance, in the United States, urban and rural areas used to experience similar mortality levels around 1970s and 1980s, however, by the end of 1980s and beginning of 1990s, a rural mortality penalty emerged and the gap between urban and rural mortality has been continuously increasing since then (Cosby et al., 2019; Singh and Siahpush, 2014a). Socioeconomic characteristics, such as racial and educational disparities, and access to health care in rural areas are among the factors that seem to mediate the urban mortality advantage in the US (Singh and Siahpush, 2014b; Spencer et al., 2018). In England and Wales, on the other hand, higher urbanization level is associated with higher mortality, a relationship that is partially, but not fully, explained by socioeconomic status (Allan et al., 2019). For older age groups in Germany, a rural mortality advantage is observed for the younger elderly (60-79), but not for those above 80, and a similar pattern is observed in England and Wales (Ebeling et al., 2022).

In LMICs, studies regarding rural-urban mortality differentials have mostly addressed under-five and infant mortality. An urban advantage was documented in Latin America (Garcia, 2020; Behm, 1980), in Sub-Saharan Africa (Yaya et al., 2019; Van De Poel et al., 2009), and in India (Kumar et al., 2022; Saikia et al., 2013). Access to economic resources, education of mother, access to health care, and community infrastructure are commonly cited among the mechanisms of this gap in infant and under-five mortality (Yaya et al., 2019; Van De Poel et al., 2009; Bocquier et al., 2011).

Findings from research on adult mortality differentials in LMICs, on the other hand, are far from uniform. The heterogeneity across countries in terms of the size and direction of the urban-rural adult mortality differentials are particularly pronounced in LMICs and vary according to different mortality and epidemiological transition stages. Menashe-Oren and Masquelier (2022) documented the urban-rural mortality differentials by age group for multiple countries and reported a pattern of urban mortality advantage in countries with lower mortality (such as Peru and Bolivia) and a rural mortality advantage in countries with higher mortality (such as Mozambique, Mali and Malawi). Additionally, the authors verified that, in those countries that observe an urban penalty in adult mortality, the shift from an urban advantage to a disadvantage occurs between ages 15-29, and is particularly pronounced for ages 30-44.

However, the mechanisms that explain this disparities in LMICs do not seem to resemble those observed in HICs. I highlight two particular features that are distinguished between these two groups that might play a role in determining the urban-rural mortality differentials: the epidemiological and mortality transitions and the relationship between socioeconomic status and mortality.

Considering the theory of epidemiological transition, LMICs went through heterogeneous transition paths and they did not necessarily follow the initially proposed linear framework model based on the trends observed for Western European countries (Vallin and Meslé, 2004; Sudharsanan et al., 2022; Bygbjerg, 2012). Latin America, in particular, has a distinguished epidemiological transition profile described as a *prolonged polarized model*, characterized by a long term overlap of profiles from different stages of the transition (high incidence of both communicable and non-communicable diseases), by reverse transitions (re-emergence of infectious diseases such as *dengue* and malaria), and by an epidemiological polarization (where different social groups are in different stages of the transition) (Frenk et al., 1991; Palloni and Pinto-Aguirre, 2011). Additionally, the region - in particular Brazil, Mexico, and Colombia - is exposed to the high burden of external causes of death that disproportionately affect the male population (Canudas-Romo and Aburto, 2019; Calazans and Queiroz, 2020; Hinestroza et al., 2021; Acosta and Prieto, 2014). Hence, each of these characteristics affect differently urban and rural areas, since they differ in terms of environment conditions, socioeconomic status and age profiles.

Finally, the documented gradient observed between SES and mortality in HICs (where higher SES is associated with lower mortality) does not necessarily hold in LMICs (Sudharsanan et al., 2020). Differences in health behaviors between higher SES groups from LMICs and HICs might be hindering the potential advantages of the first group in some cases (Sudharsanan, 2020; Vallin, 1980). In addition, socioeconomic deprivation might represent different mortality outcomes in urban and rural areas (Carvalho and Wood, 1978).

2.2 Sources of data for urban-rural adult mortality estimation

The estimation of adult mortality in LMICs is particularly challenging for two main reasons. First, several of these countries do not have functioning CRVS systems to collect the data, or, when they do, they are inaccurate or incomplete (Mikkelsen et al., 2015; Castanheira and Monteiro da Silva, 2022; Peralta et al., 20219). Second, deaths of adults, when compared to infant or older age deaths, are rare events. Therefore, large samples are required for them to be captured by household surveys (Timæus, 1991). Hence, most studies on adult mortality differentials in LMICs have relied on data from Demographic and Health Surveys (DHS) and on indirect methods, such as the sibling survival method or model life tables (Menashe-Oren and Masquelier, 2022; Carvalho and Wood, 1978; Moultrie et al., 2013).

Over the last three decades, the data quality of CRVS systems have been improving in Latin America both in terms of completeness, and in terms of quality of its records (Fabiana Del Popolo and Guiomar Bay, 2021; Mikkelsen et al., 2015). However, these data constraints the type of research questions we can answer for several reasons. First, SES information, such as race and educational attainment is not always reliable, which might incur in mismatches between numerators and denominators when computing mortality rates (Queiroz and Sacco, 2018). Second, Latin American countries show high within-country heterogeneity in terms of data quality and completeness, which impose some limitations to the study of mortality disparities (Castanheira and Monteiro da Silva, 2022; Peralta et al., 20219; Queiroz et al., 2020; Urdinola et al., 2016; Diógenes et al., 2022). Third, even countries such as Brazil and Colombia, with well functioning systems, still face problems with ill-definition of causes of death (Mikkelsen et al., 2015). Finally, the use of CRVS data for investigating sub-national level mortality disparities without accounting for differential coverage and completeness across country regions may result in biased estimates and unrealistic conclusions (Arriaga, 1967).

In this regard, the availability of questions on household deaths in national censuses represents an opportunity for capturing adult mortality in countries where CRVS systems are absent or not reliable (United Nations, 2017) or in areas where the registration of deaths is still deficient (Diógenes et al., 2022). The sample size and the fact of having both numerator and denominator from the same source is an advantage of census for mortality estimation (Queiroz and Sacco, 2018). The inclusion of this inquiry in Brazil, for example, brought up the opportunity for conducting several studies relevant for public health policy regarding: mortality of indigenous and non-indigenous children and adolescents by area of residence (Campos et al., 2017; Santos et al., 2020), educational gradients in mortality (Silva et al., 2016), and mortality differentials across urban and rural areas (Albuquerque, 2019). Hence, the example of the 2010 census mortality questionnaire in Brazil has proven to be a successful case in which censuses can be used to provide reliable mortality information (Borges, 2018; Queiroz and Sawyer, 2012), especially in the more isolated areas of the country (Diógenes et al., 2022). Questions on household deaths over a specific reference period are also included in household surveys and censuses of other LMICs and have been proven useful for documenting health and mortality disparities between urban and rural areas and SES groups in countries such as India, and Indonesia (Sudharsanan, 2017; Vyas et al., 2022; Sudharsanan, 2020).

3 Data and Methods

3.1 Data

I use the publicly available microdata from the 2018 national census of Colombia for computing both numerators (deaths) and denominators (exposure) for mortality rates. Data for the UBN index, is also retrieved at the dwelling level using the census.

The question about household deaths was first included in the 2005 national census. In 2018, the census asked if any person had died in the dwelling throughout 2017 (1/1/2017 to 12/31/2017). If it was the case, interviewers also asked the sex, age at death, and if a death certificate was issued for the deceased (DANE, 2018).

Additionally, I also use the microdata from the CRVS system compiled by the national statistics office of Colombia, DANE (2023), for calculating numerators of mortality rates by region, sex, and area of residence. Recent assessments of data quality show that the data is of fair quality, but the completeness and coverage vary across regions (Urdinola et al., 2016; Cendales and Pardo, 2018; Adair and Lopez, 2018).

3.2 Rural-Urban Classification

I classify urban and rural areas according to the criteria of degrees of urbanization recommended by the United Nations Statistical Commission (Dijkstra et al., 2021). I adopt this strategy in order to make cross-country comparisons of the reported results more straightforward. This strategy also minimizes mismatches from calculating rates from two different sources for numerators (CRVS) and denominators (census).

The following categories were constructed based on population size and density, using the census 2018 geographies:

1. City (urban): municipalities that have a population of at least 50,000 inhabitants and a population density of more than 1,500 inhabitants per km^2;
2. Towns (urban): municipalities with at least 5,000 inhabitants and with a density of at least 300 inhabitants per km^2;
3. Rural: low density municipalities (less than 300 inhabitants per km^2).

Given the similarity of data on deaths between cities and towns and in order to allow a smoother dis-aggregation of data between geographic regions and SES, I grouped these two into a single urban category.

3.3 Geographic regions

I used the following geographic division for Colombia, based on the aggregation of the departments of the country (in accordance with Acosta and Prieto (2014)):

1. **Caribe and San Andrés**: San Andrés, La Guajira, Cesar, Magdalena, Atlántico, Bolívar, Sucre, and Córdoba;

2. **Pacífico**: Chocó, Cauca, Nariño, and Buenaventura (municipality);

3. **Andes Orientales**: Norte de Santander, Santander, Boyacá, Cundinamarca, Bogotá, Tolima, and Huila;

4. **Andes Occidentales**: Antioquia, Caldas, Risaralda, Quindío, and Valle del Cauca (without the municipality of Buenaventura);

5. **Orinoquía**: Arauca, Casanare, Meta, and Vichada;

6. **Amazonía**: Guaviare, Guainía, Caquetá, Vaupés, Putumayo, and Amazonas.

3.4 Socioeconomic Status: Unmet Basic Needs Index (UBN)

The UBN index is widely used by the Economic Commission for Latin America and the Caribbean (ECLAC) for measuring poverty and vulnerability using census and survey data (Mancero and Feres, 2001). Among the advantages of this index, the most evident is its flexibility, which allows for the users to define the dimensions and adjust the criteria according to the country's characteristics.

I choose 7 dimensions related to 5 basic needs at the household level:

1. Housing quality: (1) household is considered in deprivation if the materials of walls and floor are inadequate;

2. Housing overcrowding: (2) household is considered in deprivation if there are more than 2.5 people per dormitory in the household;

3. Access to services: household is considered in deprivation if it is not provided with the following services: (3) access to water, (4) sewage system, and (5) electricity;

4. School attendance of children: (6) if the household has children 5-14, and at least one do not attend school, the household is considered in deprivation;

5. Educational attainment of household head: (7) if the household head has an educational attainment level lower to the equivalent of at least 8 years of education, the household is considered in deprivation in that dimension.

I assign equal weights to each of these 7 dimensions and count the number of unmet basic needs (minimum 0 and maximum 7 - or 6 if there is no children 5-14 in the household). Then, I create three categories for each household: 0, only 1, and 2 or more UBNs.

3.5 Methodological procedure

I divided this analysis into three stages. In a first step, I calculate the age-specific mortality rates in five-year age groups (${}_5m_x$) for adult ages from 15 to 59 for each region, data source, and area (rural-urban) by sex using the formula:

$$ {}_5m_x^{R,j,A} = \frac{{}_5D_x^{R,j,A}}{{}_5N_x^{R,A} + 0.5 \times_5 D_x^{R,j,A}}. \tag{1} $$

Where ${}_5D_x^{R,j,A}$ are the number of deaths in region R, source j (CRVS or census), and residence area A (urban or rural) between ages x to $x+5$ in 2017, and ${}_5N_x^{R,A}$ is the population counts in region R and area A between ages x to $x+5$ from the 2018 census. The census was conducted between January and October of 2018, and there is no exact reference date for it. Therefore, I assume the population as the population of January 2018 and sum it to half of 2017 deaths in the denominator in order to get an estimate of the population at the middle of 2017[1].

Two different denominators were used for the CRVS data: 1) the unadjusted death counts (${}_5D_x^{R,CRVS,A}$), as reported by DANE (2023), and 2) the adjusted count ${}_5D_x^{R,CRVSadj,A}$, created using the census information about death certification. This second count of deaths is estimated by dividing the reported deaths of the CRVS system by the proportion of deaths in the census for which a death certificate was issued (${}_5PDC_x^{R,A}$) (by age, sex, region, and area of residence):

[1]The 2018 national census microdata does not have single age information publicly available. Therefore, I could not simply draw a lexis diagram and get the population from 2018 to middle 2017. As an alternative, I applied demographic age graduation methods (Sprague) to retrieve the population in 2018 in single age groups, then I shifted the ages of individuals to the middle of 2017, and finally I regrouped then into five year age groups - at the national level. This strategy did not change the main findings.

$$ {}_5D_x^{R,CRVSadj,A} = \frac{{}_5D_x^{R,CRVS,A}}{{}_5PDC_x^{R,A}}. \tag{2} $$

Afterwards, in the second step, I calculate the age-standardized death rates between ages 15 and 59 for each region R, data source j, and area of residence A, by sex:

$$ {}_{45}ASDR_{15}^{R,j,A} = \sum_{x=15,20,...}^{55} {}_5c_x^S \times_5 m_x^{R,j,A}, \tag{3} $$

where ${}_5c_x^S$ is the proportional age-distribution for the standard population, which I chose as the both-sex population distribution of the country as a whole. Then, I calculate the ratio between ${}_{45}ASDR_{15}^{R,j,A}$ and a national reference value by sex, the ${}_{45}ASDR_{15}^{National}$ (both rural and urban areas):

$$ ASDRRatio^{R,j,A} = \frac{{}_{45}ASDR_{15}^{R,j,A}}{{}_{45}ASDR_{15}^{National}}. \tag{4} $$

By assessing this ratio we can distinguish, from a common standard, which areas are in advantaged or in disadvantaged position in relation to the national reference level.

In the third step, I decompose the differences in ASDR by region between rural and urban areas in terms of differences in deprivation level composition (ΔUBN), and of differences in rate (ΔRate). Following the formulation from Kitagawa (1955) decomposition,

$$ \begin{aligned} \Delta ASDR = & ASDR^{Rur} - ASDR^{Urb} = \\ & \sum_{UBN} W_{UBN}^{Rur} \times ASDR_{UBN}^{Rur} - \sum_{UBN} W_{UBN}^{Urb} \times ASDR_{UBN}^{Urb} = \\ & \sum_{UBN} \left(W_{UBN}^{Rur} - W_{UBN}^{Urb}\right) \times \left(\frac{ASDR_{UBN}^{Rur} + ASDR_{UBN}^{Urb}}{2}\right) + \\ & \sum_{UBN} \left(ASDR_{UBN}^{Rur} - ASDR_{UBN}^{Urb}\right) \times \left(\frac{W_{UBN}^{Rur} + W_{UBN}^{Urb}}{2}\right), \end{aligned} \tag{5} $$

where, W_{UBN}^{A} is the proportion of population of area A at the respective UBN deprivation category. Therefore, the deprivation level component of the difference is defined as:

$$ \Delta \text{UBN} = \sum_{UBN} \left(W_{UBN}^{Rur} - W_{UBN}^{Urb}\right) \times \left(\frac{ASDR_{UBN}^{Rur} + ASDR_{UBN}^{Urb}}{2}\right), \tag{6} $$

and the rate component as:

$$\Delta\text{Rate} = \sum_{UBN} \Big(ASDR^{Rur}_{UBN} - ASDR^{Urb}_{UBN}\Big) \times \Big(\frac{W^{Rur}_{UBN} + W^{Urb}_{UBN}}{2}\Big). \tag{7}$$

Finally, as a second part of the third step analysis, I estimate log-linear Poisson regression models using the multi-way contingency tables with the information of death counts and exposure by region (R), deprivation level (U), area (A), sex (S), and age (X) (Rodríguez, 2007). I consider several models interacting all five covariates to find a model that fits the data as well as the saturated model while being more parsimonious. Such a model allows us to more concisely describe the patterns observed in the data and to assess which difference by region, deprivation level, area, sex, and age are statistically significant. (Powers and Xie, 2008). I use both Bayesian Information Criterion (BIC) and Deviance measures to assess the goodness of fit.

4 Results

4.1 Number of adult deaths in urban and rural areas by data source

I start the analysis by looking at the number of deaths reported by degree of urbanization across different data sources for the adult population.[2] Table 1 shows the total number of deaths reported by the census and by the CRVS system in urban and rural areas by sex and region of Colombia. We notice a consistent gradient between cities, towns and rural areas: CRVS deaths are higher or similar to deaths reported by the census in towns and cities, but they are much lower in rural areas. This difference in reporting is noteworthy in lower density regions, such as *Pacífico*, *Orinoquía*, and the *Amazonía*. The census reported more deaths for the adult age population than the CRVS system in every region (except for males in *San Andrés*). At the national level, the census reported 25% more deaths than the CRVS system and 60% more deaths in rural areas (both sexes combined).

[2] In Appendix 6 I include additional analyses of the total number of deaths by source and the differences in reporting for all five-year age groups from 0 to 100+.

Table 1: Total adult age (15-59) population (2018) and death counts (2017) by region, area, source (census vs CRVS), and sex. Source: DANE (2018, 2023).

		Males				Females			
Region	Area	Population	Census	CRVS	Ratio	Population	Census	CRVS	Ratio
San Andrés									
	Town	14038	35	67	0.52	15077	17	16	1.06
	Rural	1410	10	7	1.43	1529	1	0	-
	Total	15448	45	74	0.61	16606	18	16	1.12
Caribe									
	City	515456	1044	1579	0.66	552406	579	886	0.65
	Town	543158	1280	1543	0.83	575793	677	835	0.81
	Rural	1919086	7309	4207	1.74	1948016	4237	2433	1.74
	Total	2977700	9633	7329	1.31	3076215	5493	4154	1.32
Pacífico									
	Town	212869	911	612	1.49	233348	419	360	1.16
	Rural	793318	4950	2323	2.13	817060	2332	1094	2.13
	Total	1006187	5861	2935	2.00	1050408	2751	1454	1.89
Andes Orientales									
	City	2886135	5679	5883	0.96	3120257	3109	3402	0.91
	Town	1004949	2631	2387	1.10	1076674	1445	1421	1.02
	Rural	1442991	5188	3865	1.34	1404878	2670	1953	1.37
	Total	5334075	13498	12135	1.11	5601809	7224	6776	1.07
Andes Occidentales									
	City	1835036	6063	6021	1.01	2029571	2720	2972	0.92
	Town	633727	2447	2098	1.17	686439	1076	965	1.11
	Rural	1229857	6325	4540	1.39	1264498	2787	1900	1.47
	Total	3698620	14835	12659	1.17	3980508	6583	5837	1.13
Orinoquía									
	Town	146737	430	434	0.99	150664	205	224	0.92
	Rural	374503	1536	1031	1.49	359298	763	472	1.62
	Total	521240	1966	1465	1.34	509962	968	696	1.39
Amazonía									
	Rural	271478	1442	927	1.56	259576	797	475	1.68
	Total	271478	1442	927	1.56	259576	797	475	1.68
National									
	City	5236627	12786	13483	0.95	5702234	6408	7260	0.88
	Town	2555478	7734	7141	1.08	2737995	3839	3821	1.00
	Rural	6032643	26760	16900	1.58	6054855	13587	8327	1.63
	Total	13824748	47280	37524	1.26	14495084	23834	19408	1.23

Table 2: Census reported adult age (15-59) deaths with a death certificate in 2017 by region and area of residence. Source: DANE (2018).

		Is there a death certificate for the deceased? Total (%)					
			Males			**Females**	
Region	Area	Yes	No	Not informed	Yes	No	Not informed
San Andrés							
	Town	26 (74.3)	6 (17.1)	3 (8.6)	14 (82.4)	1 (5.9)	2 (11.8)
	Rural	9 (90.0)	0 (0.0)	1 (10.0)	0 (0.0)	0 (0.0)	1 (100.0)
	Total	35 (77.8)	6 (13.3)	4 (8.9)	14 (77.8)	1 (5.6)	3 (16.7)
Caribe							
	City	896 (85.8)	65 (6.2)	83 (8.0)	500 (86.4)	26 (4.5)	53 (9.2)
	Town	1037 (81.0)	118 (9.2)	125 (9.8)	543 (80.2)	48 (7.1)	86 (12.7)
	Rural	5062 (69.3)	1403 (19.2)	844 (11.5)	2806 (66.2)	910 (21.5)	521 (12.3)
	Total	6995 (72.6)	1586 (16.5)	1052 (10.9)	3849 (70.1)	984 (17.9)	660 (12.0)
Pacífico							
	Town	759 (83.3)	70 (7.7)	82 (9.0)	361 (86.2)	23 (5.5)	35 (8.4)
	Rural	3394 (68.6)	795 (16.1)	761 (15.4)	1591 (68.2)	360 (15.4)	381 (16.3)
	Total	4153 (70.9)	865 (14.8)	843 (14.4)	1952 (71.0)	383 (13.9)	416 (15.1)
Andes Orientales							
	City	5008 (88.2)	324 (5.7)	347 (6.1)	2770 (89.1)	158 (5.1)	181 (5.8)
	Town	2228 (84.7)	187 (7.1)	216 (8.2)	1255 (86.9)	84 (5.8)	106 (7.3)
	Rural	4092 (78.9)	506 (9.8)	590 (11.4)	2116 (79.3)	226 (8.5)	328 (12.3)
	Total	11328 (83.9)	1017 (7.5)	1153 (8.5)	6141 (85.0)	468 (6.5)	615 (8.5)
Andes Occidentales							
	City	5233 (86.3)	404 (6.7)	426 (7.0)	2353 (86.5)	173 (6.4)	194 (7.1)
	Town	2103 (85.9)	171 (7.0)	173 (7.1)	951 (88.4)	56 (5.2)	69 (6.4)
	Rural	5133 (81.2)	576 (9.1)	616 (9.7)	2220 (79.7)	273 (9.8)	294 (10.5)
	Total	12469 (84.1)	1151 (7.8)	1215 (8.2)	5524 (83.9)	502 (7.6)	557 (8.5)
Orinoquía							
	Town	348 (80.9)	36 (8.4)	46 (10.7)	174 (84.9)	13 (6.3)	18 (8.8)
	Rural	1155 (75.2)	167 (10.9)	214 (13.9)	569 (74.6)	80 (10.5)	114 (14.9)
	Total	1503 (76.4)	203 (10.3)	260 (13.2)	743 (76.8)	93 (9.6)	132 (13.6)
Amazonía							
	Rural	967 (67.1)	235 (16.3)	240 (16.6)	564 (70.8)	110 (13.8)	123 (15.4)
	Total	967 (67.1)	235 (16.3)	240 (16.6)	564 (70.8)	110 (13.8)	123 (15.4)
National							
	City	11137 (87.1)	793 (6.2)	856 (6.7)	5623 (87.7)	357 (5.6)	428 (6.7)
	Town	6501 (84.1)	588 (7.6)	645 (8.3)	3298 (85.9)	225 (5.9)	316 (8.2)
	Rural	19812 (74.0)	3682 (13.8)	3266 (12.2)	9866 (72.6)	1959 (14.4)	1762 (13.0)
	Total	37450 (79.2)	5063 (10.7)	4767 (10.1)	18787 (78.8)	2541 (10.7)	2506 (10.5)

One can get an estimate of completeness of the CRVS system through looking at the proportion of deaths in the census that were issued a death certificate. Table 2 reports the number of deaths in the census for which there was issued a death certificate. Rural areas have lowest proportion of deaths with an issued death certificate: 74% for males and 72.6% for females against around 85-87% for urban areas. I then use this information by age, sex, and region to provide an adjusted count for the CRVS system reported deaths.

For the following sections, given its low number of deaths, I grouped *San Andrés* with the *Caribe* region. Also, I grouped cities and towns as an unique urban area group, since they show similar adult mortality profiles.

4.2 Adult mortality by area of residence, sex, and region: differences by data source

In Figures 1 and 2, I show the age-specific mortality rates for urban (cities and towns combined) and rural areas by data source and the ratios between rates from these two areas. From the census reports, we see that rural areas have higher death rates for every age group in all regions and at the national level. On the other had, the CRVS data suggests smaller differences between areas, or even an urban advantage for *Amazonía*, *Orinoquía*, *Caribe* and *Pacífico* regions combined in some age groups.

Figure 1: Age-specific mortality rates by region and area, Colombia 2018 census. Source: DANE (2018, 2023)

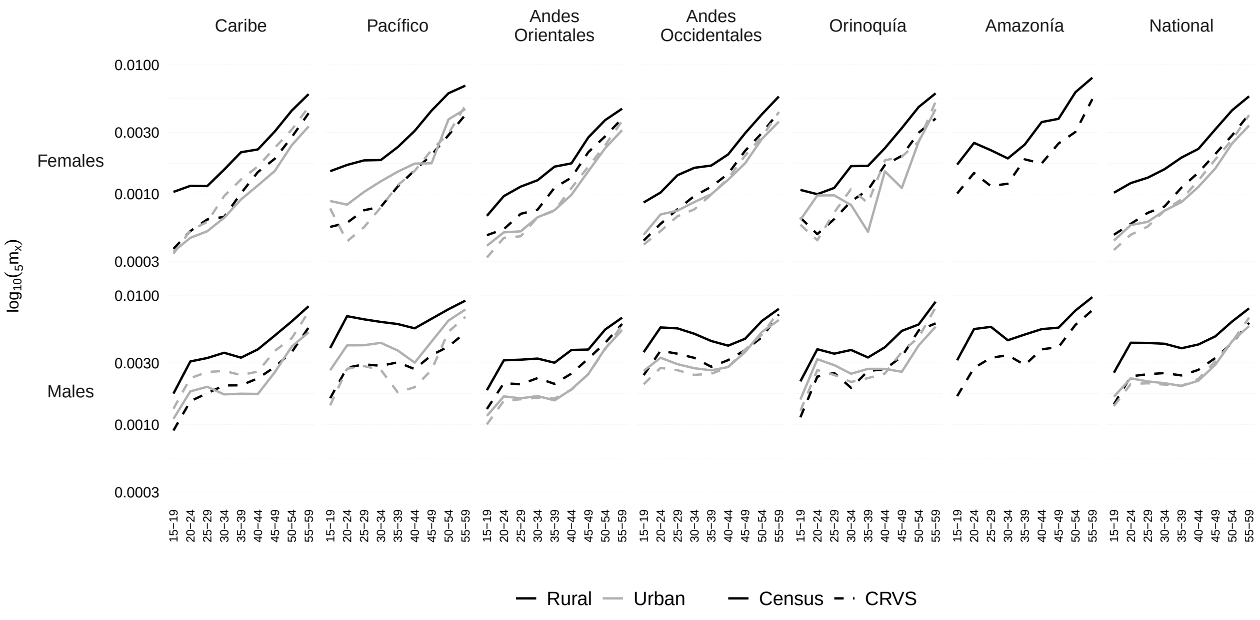


Figure 2: Rural to urban ratio of age-specific mortality rates by region and area, Colombia 2018 census. Source: DANE (2018, 2023)

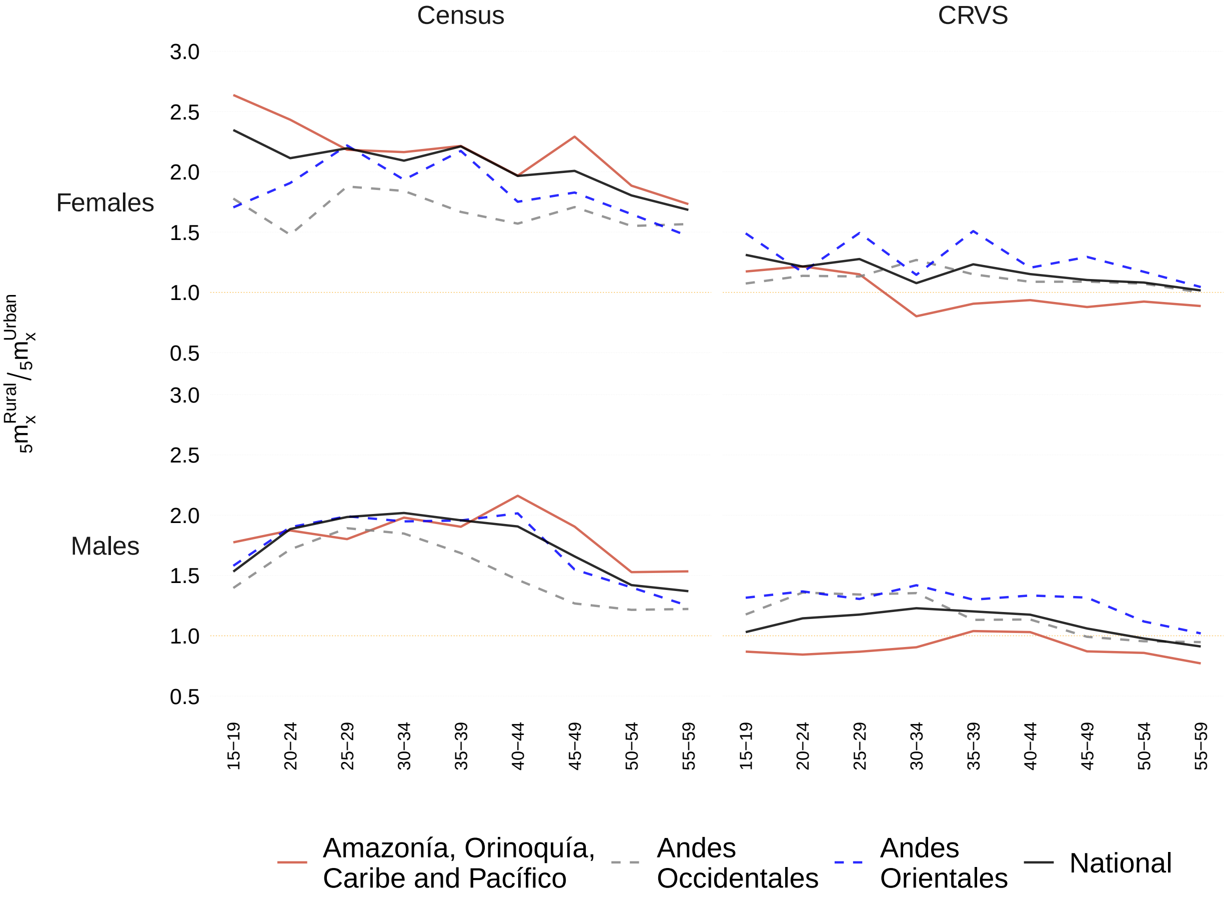
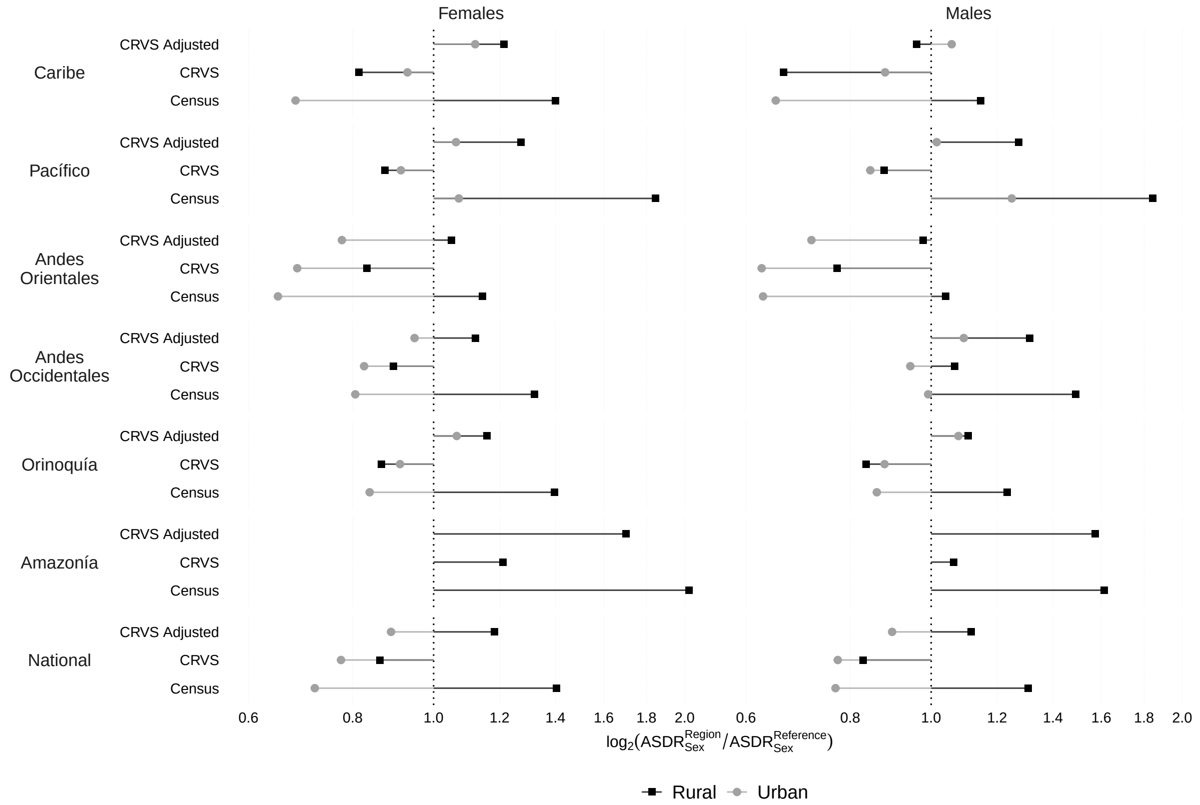


Figure 3 shows the results for the ratios between age-standardized death rates (by source, region, and area) and the national reference age-standardized death rate (from the census data) by sex in the log_2 scale. Values below one represent a lower mortality, and values above one represent a mortality disadvantage in relation to the national level for each sex. For census mortality data, all regions show a clear rural mortality penalty. At the national level, women have 40% higher ASDR in rural areas compared to the ASDR observed for women of the country as a whole. Men show a lower disadvantage, around 30% higher than the male national average from the census, but this average is also higher than the one observed for females.

Figure 3: Ratio between region age-standardized death rates and Reference standard age-standardized death rate by region, sex and area, Colombia 2018 census. Source: DANE (2018, 2023)



This graph also shows that the choice of the source matters for the direction of the gap. Using the unadjusted CRVS data can lead to concluding that there is an urban penalty in *Orinoquía* and in the *Caribe* for males and females, and in the *Pacífico* for females. The *Caribe* was the only case in which both adjusted and unadjusted CRVS data showed an urban penalty. I highlight the observed results of *Amazonía* and *Pacífico*, for which the female standardized rates were twice as large as the national reference, and the males were 1.6 and 1.8 times larger.

4.3 Adult mortality differentials by area of residence and household deprivation status

Table 3 reports the population composition by deprivation status by region and area. We verify that urban residents are in general less deprived than rural residents. While around half of residents of urban areas do not have any unmet basic need, around half of rural residents have at least 2 unmet basic needs. Given the size of the samples, for the decomposition exercise I grouped *Caribe*, *Pacífico*, *Orinoquía*, and the *Amazonía* regions.

Table 3: Population by level of deprivation (number of UBNs), by region and area of residence. Source: DANE (2018).

Region	Area	Population	UBN (%) 0	1	2+
Caribe					
	Urban	3430024	40.2	36.5	23.2
	Rural	6422526	18.7	27.8	53.5
	Total	9852550	26.2	30.9	43.0
Pacífico					
	Urban	662262	42.8	38.6	18.6
	Rural	2611128	12.3	23.9	63.8
	Total	3273390	18.5	26.9	54.6
Andes Orientales					
	Urban	11967058	53.0	36.8	10.1
	Rural	4624895	21.7	36.8	41.5
	Total	16591953	44.3	36.8	18.9
Andes Occidentaless					
	Urban	7741824	49.2	40.4	10.4
	Rural	3950131	24.5	38.9	36.5
	Total	11691955	40.9	39.9	19.2
Andes Orinoquía					
	Urban	435286	45.0	39.9	15.1
	Rural	1137462	22.7	32.6	44.7
	Total	1572748	28.9	34.6	36.5
Amazonía					
	Rural	821589	16.7	24.8	58.5
	Total	821589	16.7	24.8	58.5
National					
	Urban	24236454	49.6	38.0	12.4
	Rural	19567731	19.9	31.8	48.3
	Total	43804185	36.3	35.3	28.4

Figure 4 shows the mortality rates by area of residence and deprivation status at the national level by sex. We see that the rural mortality penalty holds within each of the deprivation statuses. Mortality rates for groups with at least one unmet basic need are the highest and for urban residents with no deprivation are the lowest, for both sexes. For ages 15-44 rural mortality rates for the least deprived group is lower than for urban groups with at least one unmet basic need. Nevertheless, an rural disadvantage emerges from ages 45-59 even for the least deprived rural population.

Figure 4: Age-specific mortality rates by sex, area of residence, and deprivation status. Source: DANE (2018).

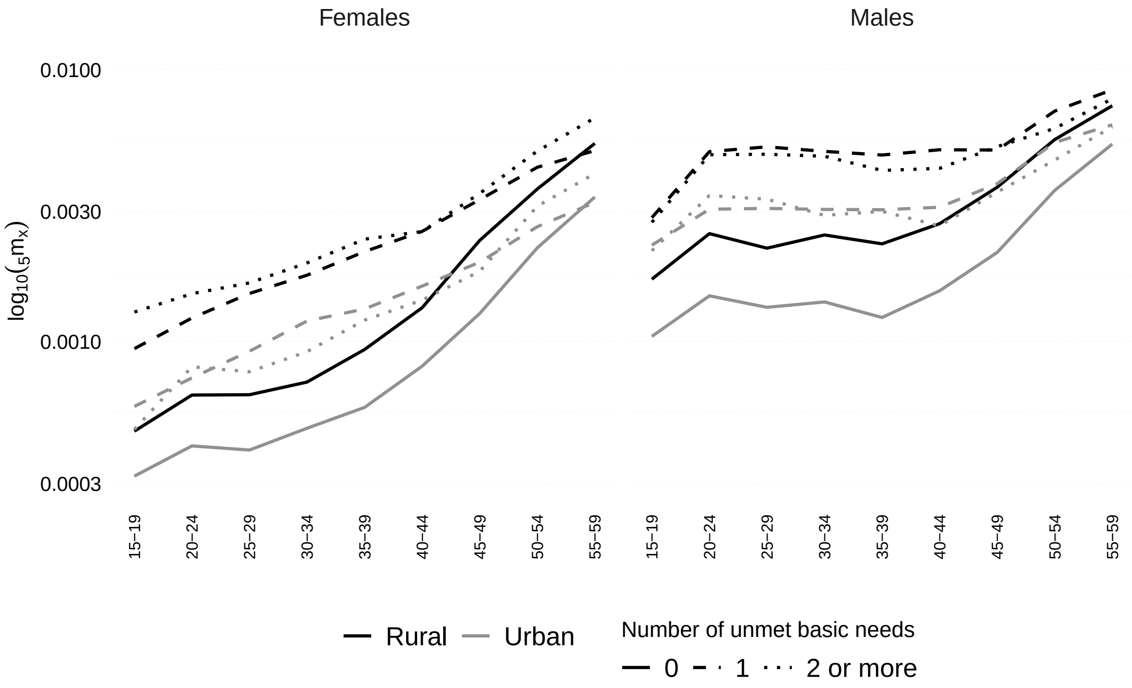


Table 4: Kitagawa (1955) decomposition of differences in age-standardized death rates between rural and urban areas.

Region	Sex	$\Delta ASDR$	ΔRate (%)	ΔUBN (%)
Caribe, Pacífico, Orinoquía, Amazonia				
	Females	1.29	83.11	16.89
	Males	2.09	84.86	15.14
Andes Orientales				
	Females	0.79	75.80	24.20
	Males	1.48	69.99	30.01
Andes Occidentales				
	Females	0.85	77.79	22.21
	Males	1.81	72.30	27.70
National				
	Females	1.11	77.60	22.40
	Males	1.91	74.99	25.01

From the Kitagawa (1955) decomposition exercise reported in Table 4, we see that the differences in deprivation status composition of urban and rural areas explain around 25% of differences in rates for men and 22% for women at the national level. In the most vulnerable areas (*Caribe*, *Pacífico*, *Orinoquía*, *Amazonía*), the share explained by differences in rates is higher: 83% for women, and around 85% for men. In the Andean regions (*Andes Orintales* and *Andes Occidentales*), on the other hand, deprivation composition explains a higher share of the differences in rates for both men (27-30%) and women (22-24%).

Finally, we run several Poisson Log-linear Models interacting the five available covariates (region (R), area of residence (A), sex (S), household deprivation status (U), and age (X) in order to find a more parsimonious model that still fits the data. Table 5 reports the results of goodness-of-fit statistics. We can see that no single model was able to fit the data. Taking into account the simplest model (A, only area of residence) and the best fitting one ($A + RSXU$, area of residence plus the interaction between the other covariates), we verify a reduction in the coefficient for urban-rural area from 0.57 to 0.38 (reduction of 33%, similar to the one observed between the simplest and the additive model $R + A + U + S + X$). Therefore, these results suggest that we cannot explain the differences observed in our data without accounting for the area of residence.

Table 5: Deviances and BIC measures for Poisson Log-linear Models fitted to mortality rates by region (R), area of residence (A), sex (S), household deprivation status (U), and age (X). Source: DANE (2018).

Model	d.f	Deviance	BIC	p-value
Null	647	38248.00	41819.00	0.000
A	646	32533.00	36110.00	0.000
$R + U + S + X$	631	6583.00	10257.00	0.000
$R + A + U + S + X$	630	4722.00	8403.00	0.000
$A + RU + S + X$	620	4641.00	8386.00	0.000
$A + R + US + X$	628	4639.00	8333.00	0.000
$A + U + RS + X$	625	4485.00	8198.00	0.000
$A + RUS + X$	603	4303.00	8158.00	0.000
$A + U + S + RX$	590	4151.00	8090.00	0.000
$A + R + S + UX$	614	3538.00	7322.00	0.000
$A + R + U + SX$	622	3084.00	6817.00	0.000
$A + S + RUX$	484	2800.00	7426.00	0.000
$RUSX$	324	2471.00	8133.00	0.000
$A + RSX + U$	537	2135.00	6417.00	0.000
$A + RSXU$	323	561.00	6229.00	0.000
Saturated Model				
$RAUSX$	0	0.00	7759.00	1.000

5 Discussion

This work reported a consistent rural penalty for all regions of Colombia. Additionally, about 20-25% of this penalty seems to be explained by differences in the household deprivation level compositions across urban and rural areas. Mortality rates for rural areas are higher than in rural areas at each level of household socioeconomic deprivation.

Hence, these results place Colombia within the *low mortality* group, according to the theoretical framework suggested by Menashe-Oren and Masquelier (2022) regarding urban-rural mortality differentials. For this group of countries, which are in a more advanced stage of the epidemiological transition, the urban advantage persists from younger to adult age groups. However, the case of Colombia suggests that the armed conflict might be playing an important role for defining such large differentials.

I highlight the differentials observed among the most vulnerable departments of the country - *Pacífico* and *Amazonía*. These two areas are highly affected by the armed conflict, and they are home of vulnerable social groups and ethnic minorities that are undeserved by the public services of the country (Álvaro Turriago-Hoyos et al., 2020; Vallejo et al., 2018; Acosta and Prieto, 2014; Hinestroza et al., 2021).

The armed conflict in rural areas impose direct and indirect negative effects on the population mortality (Franco et al., 2006). First, it directly influences by the disproportional numbers of homicides within the affected areas (Hinestroza et al., 2021; Acosta and Prieto, 2014; Vallejo et al., 2018). Second, it disrupts the health services, imposing barriers for the population that requires health care (Franco et al., 2006). Finally, I also mention a possible long-term effect of the conflict over the inhabitants of the area. Taking into account that mortality is a product of life course experiences (Elo and Preston, 1992; Elo, 2009), the worsened birth outcomes of those who are born in areas of conflict (Urdinola-Contreras, 2018; Urdinola, 2004) might suggest that individuals born in areas of conflict are more likely to experience worsened health outcomes in adulthood.

The fact that most of differences in rates could not be explained by differences in socioeconomic deprivation indicates that factors not included in this analysis is playing a role in determining urban-rural mortality differentials in Colombia. For instance, differences in smoking (Macías et al., 2013) and drinking behavior might be playing a role. I can also mention the lack of public health infrastructure among rural areas, which is suggested by the high under-reporting of vital events in these areas (Vital Strategies, 2020). The lack of good quality civil registration data limits my analysis, since an assessment of cause of death profiles by area could shed some light on the observed differentials.

Beyond the main purpose of documenting the urban-rural differentials in adult mortality and showing how socioeconomic deprivation is related to the observed gap, this study also advocates for the use of household deaths from 2018 national census to study adult mortality in Colombia. Up to now, no studies in Colombia have used this information for investigating mortality disparities, even though the data appearing to be more reliable in more remote areas of the country as suggested by this analysis. We notice a consistent gradient in death reporting across different areas by data source. CRVS death counts are higher or similar to those reported from censuses in urban areas, but lower in rural areas. Similar results have been observed in Brazil, where more isolated areas had higher deaths reported in the 2010 national census (Diógenes et al., 2022). The higher reporting of CRVS for cities in some regions might indicate a limitation of the census inquiry in detecting deaths when it results in household dissolution - deaths that occur in one person households (Borges, 2018). The main conclusion from the comparison of these two data sources is that one cannot use the CRVS data without accounting for differential reporting across areas when doing sub-national level analysis (Arriaga, 1967). The suggested approach of correcting the CRVS data using the census question about the emission of death certificate might not be sufficient for properly correcting the death counts, but they might give a direction towards the true values.

5.1 Limitations and Next Steps

This work does not come without limitations, which I will try to address accordingly.

First of all, in this work I am using unadjusted death counts from the 2018 Colombian National census. Even though I have shown that the census death counts are likely to have better coverage and completeness levels than the CRVS deaths, these numbers are still likely to be under-estimated given the issues related to this type of data collection (Timæus, 1991). Maternal mortality assessments using similar information from Asian and African countries have shown high levels of under-reporting of census household deaths using death distribution methods (Hill et al., 2018). However, in Brazil, where the few studies that assessed the under-reporting of census deaths from this type of inquiry in Latin America were conducted, it was found a completeness level of approximately 80% of completeness for the 2010 census household deaths (Queiroz and Sawyer, 2012). Borges (2018) suggested an adjustment of the data by removing from the denominator of mortality rates people that lived in one person households for which no deaths have been reported and found values similar to the ones reported by the CRVS system of the country for adult ages. Therefore, the reliability of this kind of information might not be comparable across countries from different regions. In preliminary checks with the last estimates from the World Population Prospects for Colombia, I found mortality levels similar between those reported by the **?** and those estimated using the 2018 census. As a next step regarding this limitation, I will implement two different approaches for adjusting household deaths: 1) I will implement the one person household adjustment used by Borges (2018) in the Brazilian census, and 2) I will use the death distribution methods for assessing the completeness by *departamento* (Hill et al., 2009).

Second, the lower number of deaths in urban areas for the census might suggest an over-estimation of the rural penalty. I believe household dissolution might be playing a role for these observed differences (Timæus, 1991). I will address this by using Borges (2018) approach described above.

Third, migration is likely to be taking place, especially from rural areas affected by the armed conflict (Franco et al., 2006). For assessing the influence of migration, I will remove households that have any recent migration history from the analysis to check if the results hold.

Fourth, the mechanisms regarding SES and mortality are more complex and are not the same for different characteristics such as employment, education and household conditions (Elo, 2009). In this work, I am estimating deprivation through seven dimensions of unmet basic needs at the household level. Further, the index might be masking the different mechanisms from each of its dimensions, and, also, they do not necessarily reflect the social capital of the individual itself. However, given the high colinearity of each of these dimensions, and the fact that we do not have SES information for the deceased, the UBN index seems as a good alternative for assessing deprivation in this context (Mancero and Feres, 2001). As a next step, I will replicate the analysis using single dimensions and using an index built from principal component analysis to check if the conclusions do not change.

At last, there is likely to be some bias in the age reporting or in the reference period of household deaths. Using five-year age groups might reduce this bias, and also, restricting the analysis for adult mortality limits the potential errors of age exaggeration that occur in advanced age groups (Palloni and Pinto-Aguirre, 2011).

6 Conclusion

To my knowledge, this is the first work to document urban-rural mortality differentials in Colombia. Much of the work done before in the country focus on urban areas or specific causes of deaths and use CRVS data, but few do account for geographic differences in completeness (Urdinola et al., 2016). I show that the 2018 census death counts can be more reliable in rural areas and I explore its use to estimate urban-rural mortality differentials.

As expected from previous studies (Urdinola et al., 2016; Hinestroza et al., 2021; Acosta and Prieto, 2014; Vallejo et al., 2018), I find a rural mortality penalty in Colombia. The highest penalties were found for the most vulnerable regions that are affected by the armed conflict (*Pacífico* and *Amazonía*).

By decomposing the differences in age-specific death rates by UBN index composition, and rate effects, I found that, at the national level, about 20-25% of the differences between urban and rural can be explained by differences in socioeconomic deprivation composition. Therefore, most of the urban-rural gap is driven by the rate-effect or by characteristics/factors not included in my analysis, such as exposure to the conflict zones, behaviors (smoking and drinking), and lack of public health infrastructure in more isolated rural areas.

References

Acosta, K. and Prieto, J. R. (2014) Cambios recientes en las principales causas de mortalidad en colombia. *Documentos de trabajo sobre Economía Regional 209*, Centro de Estudios Económicos Regionales (CEER), Banco de La República, Cartagena.

Adair, T. and Lopez, A. D. (2018) Estimating the completeness of death registration: An empirical method. *PLOS ONE*, **13**, 1–19.

Albuquerque, F. R. P. C. (2019) Tábuas de Mortalidade segundo a situação do domicílio utilizando-se o Censo Demográfico de 2010. *Informe de Previdência Social 4*, Ministério da Economia.

Allan, R., Williamson, P. and Kulu, H. (2019) Gendered mortality differentials over the rural-urban continuum: The analysis of census linked longitudinal data from england and wales. *Social Science Medicine*, **221**, 68–78.

Arriaga, E. E. (1967) Rural-urban mortality in developing countries: An index for detecting rural underregistration. *Demography*, **4**, 98–107.

Behm, H. (1980) Socio-economic determinants of mortality in latin america. *Population Bulletin*, 1–15.

Bocquier, P., Madise, N. J. and Zulu, E. M. (2011) Is There an Urban Advantage in Child Survival in Sub-Saharan Africa? Evidence From 18 Countries in the 1990s. *Demography*, **48**, 531–558.

Borges, G. M. (2018) *Consistent Population Estimates: an Application to Brazil*. Ph.D. thesis, University of California, Berkeley.

Bygbjerg, I. C. (2012) Double burden of noncommunicable and infectious diseases in developing countries. *Science*, **337**, 1499–1501.

Calazans, J. A. and Queiroz, B. L. (2020) The adult mortality profile by cause of death in 10 latin american countries (2000–2016). *Rev Panam Salud Publica*, **44**.

Campos, M. B., Borges, G. M., Queiroz, B. L. and Santos, R. V. (2017) Diferenciais de mortalidade entre indígenas e não indígenas no brasil com base no censo demográfico de 2010. *Cadernos de Saúde Pública*, **33**.

Canudas-Romo, V. and Aburto, J. M. (2019) Youth lost to homicides: disparities in survival in latin america and the caribbean. *BMJ Global Health*, **4**.

Carvalho, J. A. M. and Wood, C. H. (1978) Mortality, income distribution, and rural-urban residence in brazil. *Population and Development Review*, **4**, 405–420.

Castanheira, H. C. and Monteiro da Silva, J. H. C. (2022) Examining sex differences in the completeness of peruvian crvs data and adult mortality estimates. *Genus*, **78**.

Cendales, R. and Pardo, C. (2018) Colombian death certificate data quality. *Colombia Médica*, **49**, 121–127.

Cosby, A. G., McDoom-Echebiri, M. M., James, W., Khandekar, H., Brown, W. and Hanna, H. L. (2019) Growth and persistence of place-based mortality in the united states: The rural mortality penalty. *American Journal of Public Health*, **109**, 155–162.

Cutler, D. and Miller, G. (2005) The role of public health improvements in health advances: The twentieth-century united states. *Demography*, **42**, 1–22.

DANE (2018) Censo Nacional de Población y Vivienda 2018.

— (2023) Demografía y población, Nacimientos y Defunciones. URL: `https:`

//www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/nacimientos-y-defunciones.

Davis, K. (1965) The urbanization of the human population. *Scientific American*, **213**, 40–53.

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M. and Schiavina, M. (2021) Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*, **125**.

Diógenes, V. H. D., Pinto Júnior, E. P., Gonzaga, M. R., Queiroz, B. L., Lima, E. E. C., Costa, L. C. C. d., Rocha, A. S., Ferreira, A. J. F., Teixeira, C. S. S., Alves, F. J. O., Rameh, L., Flores-Ortiz, R., Leyland, A., Dundas, R., Barreto, M. L. and Ichihara, M. Y. T. (2022) Differentials in death count records by databases in brazil in 2010. *Revista de Saúde Pública*, **56**, 92.

Easterlin, R. A. (2004) How beneficent is the market? a look at the modern history of mortality. In *The Reluctant Economist: Perspectives on Economics, Economic History, and Demography*, 101–138. Cambridge: Cambridge University Press.

Ebeling, M., Rau, R., Sander, N., Kibele, E. and Klüsener, S. (2022) Urban–rural disparities in old-age mortality vary systematically with age: evidence from germany and england wales. *Public Health*, **205**, 102–109.

Elo, I. T. (2009) Social class differentials in health and mortality: Patterns and explanations in comparative perspective. *Annual Review of Sociology*, **35**, 553–572.

Elo, I. T. and Preston, S. H. (1992) Effects of early-life conditions on adult mortality: A review. *Population Index*, **58**, 186–212.

Fabiana Del Popolo and Guiomar Bay (2021) Las estadísticas de nacimientos y defunciones en américa latina con miras al seguimiento de la agenda 2030 para el desarrollo sostenible y del consenso de montevideo sobre población y desarrollo. *Serie Población y Desarrollo 134*, Comisión Económica para América Latina y el Caribe (CEPAL), Santiago.

Fink, G., Günther, I. and Hill, K. (2016) Urban mortality transitions: The role of slums. In *New Approaches to Death in Cities during the Health Transition* (eds. D. Ramiro Fariñas and M. Oris), 197–216. Cham: Springer International Publishing.

Franco, S., Suarez, C. M., Naranjo, C. B., Báez, L. C. and Rozo, P. (2006) The effects of the armed conflict on the life and health in colombia. *Ciência e Saúde Coletiva*, **11**.

Frenk, J., Frejka, T., Bobadilla, J., Stern, C., Lozano, R., Sepúlveda, J. and José, M. (1991) [the epidemiologic transition in latin america]. *Boletin de la Oficina Sanitaria Panamericana. Pan American Sanitary Bureau*, **111**, 485—496. URL: http://europepmc.org/abstract/MED/1838685.

Garcia, J. (2020) Urban–rural differentials in Latin American infant mortality. *Demographic Research*, **42**, 203–244.

Hill, K., Johnson, P., Singh, K., Amuzu-Pharin, A. and Kharki, Y. (2018) Using census data to measure maternal mortality: A review of recent experience. *Demographic Research*, **39**, 337–364.

Hill, K., You, D. and Choi, Y. (2009) Death distribution methods for estimating adult mortality: Sensitivity analysis with simulated data errors. *Demographic Research*, **21**, 235–254.

Hinestroza, B. C., Sánchez, H. B., Aidar, T. and Palloni, A. (2021) Mortality rates estimates in the context of high homicide levels and poor data quality on deaths: The case of the Pacific region in Colombia between 1993 and 2013. *Revista Latinoamericana de Población*, **15**, 140–165.

Horiuchi, S. (1999) Epidemiological transitions in human history. In *Health and Mortality Issues of Global Concern*. New York: Department of Economic and Social Affairs, Population Division, United Nations.

Kitagawa, E. M. (1955) Components of a difference between two rates. *Journal of the American Statistical Association*, **50**, 1168–1194.

Kumar, C., Piyasa and Saikia, N. (2022) An update on explaining the rural-urban gap in under-five mortality in india. *BMC Public Health*, **22**, 2093.

LeGrand, C. C. (2003) The Colombian crisis in historical perspective. *Canadian Journal of Latin American and Caribbean Studies / Revue canadienne des études latino-américaines et caraïbes*, **28**, 165–209.

Macías, F., Malmusi, D. and Borrell, C. (2013) Different patterns by age-group and gender of socioeconomic inequalities in smoking in colombia. *Nicotine Tobacco Research*, **15**, 1745–1755.

Mancero, X. and Feres, J. C. (2001) El método de las necesidades básicas insatisfechas (NBI) y sus aplicaciones en América Latina. *Serie estúdios estadísticos y prospectivos 7*, Comisión Económica para América Latina y El Caribe - División de Estadística y Proyecciones Económicas, Santiago, Chile.

Menashe-Oren, A. and Masquelier, B. (2022) The shifting rural–urban gap in mortality over the life course in low- and middle-income countries. *Population Studies*, **76**, 37–61.

Menashe-Oren, A. and Stecklov, G. (2018) Urban-rural disparities in adult mortality in sub-saharan africa. *Demographic Research*, **39**.

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., de Savigny, D., Lozano, R. and Lopez, A. D. (2015) A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *The Lancet*, **386**, 1395–1406.

Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I. and Zaba, B. (2013) *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population (IUSSP).

Palloni, A. and Pinto-Aguirre, G. (2011) Adult mortality in latin america and the caribbean. In *International Handbook of Adult Mortality* (eds. R. G. Rogers and E. M. Crimmins), 101–132. Dordrecht: Springer Netherlands.

Peralta, A., Benach, J., Borrell, C., Espinel-Flores, V., Cash-Gibson, L., Queiroz, B. L. and Marí-Dell'Olmo, M. (20219) Evaluation of the mortality registry in Ecuador (2001–2013) – social and geographical inequalities in completeness and quality. *Population Health Metrics*, **17**.

Powers, D. A. and Xie, Y. (2008) *Statistical Methods for Categorical Data Analysis*. Emerald Group Publishing.

Queiroz, B. L., Gonzaga, M. R., Vasconcelos, A. M. N., Lopes, B. T. and Abreu, D. M. X. (2020) Comparative analysis of completeness of death registration, adult mortality and life expectancy at birth in brazil at the subnational level. *Population Health Metrics*, **18**, 11.

Queiroz, B. L. and Sacco, N. (2018) Es relevante incorporar la medición de la mortalidad en los censos de américa latina y el caribe? *Revista Brasileira de Estudos de População*, **35**.

Queiroz, B. L. and Sawyer, D. O. T. (2012) O que os dados de mortalidade do Censo de 2010 podem nos dizer? *Revista Brasileira de Estudos de População*, **29**, 225–238.

Reher, D. S. (2001) In search of the 'urban penalty': exploring urban and rural mortality patterns in spain during the demographic transition. *International Journal of Population Geography*, **7**, 105–127.

Rodríguez, G. (2007) *Lecture Notes on Generalized Linear Models*. URL: `https://grodri.github.io/glms/notes/`.

Saikia, N., Singh, A., Jasilionis, D. and Ram, F. (2013) Explaining the rural-urban gap in infant mortality in India. *Demographic Research*, **29**, 473–506.

Santos, R. V., Borges, G. M., de Campos, M. B., Queiroz, B. L., Coimbra, C. E. and Welch, J. R.

(2020) Indigenous children and adolescent mortality inequity in brazil: What can we learn from the 2010 national demographic census? *SSM - Population Health*, **10**, 100537.

Silva, L. E. d., Freire, F. H. M. d. A. and Pereira, R. H. M. (2016) Diferenciais de mortalidade por escolaridade da população adulta brasileira, em 2010. *Cadernos de Saúde Pública*, **32**.

Singh, G. K. and Siahpush, M. (2014a) Widening rural–urban disparities in all-cause mortality and mortality from major causes of death in the usa, 1969–2009. *Journal of Urban Health*, **91**, 272–292.

— (2014b) Widening rural–urban disparities in life expectancy, u.s., 1969–2009. *American Journal of Preventive Medicine*, **46**, e19–e29.

Spencer, J. C., Wheeler, S. B., Rotter, J. S. and Holmes, G. M. (2018) Decomposing mortality disparities in urban and rural u.s. counties. *Health Services Research*, **53**, 4310–4331.

Sudharsanan, Nikkiland Ho, J. Y. (2020) Rural–urban differences in adult life expectancy in indonesia: A parametric g-formula–based decomposition approach. *Epidemiology*, **31**.

Sudharsanan, N. (2017) The Association Between Socioeconomic Status and Adult Mortality in a Developing Country: Evidence From a Nationally Representative Longitudinal Survey of Indonesian Adults. *The Journals of Gerontology: Series B*, **74**, 484–495.

Sudharsanan, N., Aburto, J. M., Riffe, T. and van Raalte, A. (2022) Large variation in the epidemiological transition across countries: is it still valuable as a mortality theory? *International Journal of Epidemiology*, **51**, 1057–1061.

Sudharsanan, N., Zhang, Y., Payne, C. F., Dow, W. and Crimmins, E. (2020) Education and adult mortality in middle-income countries: Surprising gradients in six nationally-representative longitudinal surveys. *SSM - Population Health*, **12**.

Timæus, I. M. (1991) Measurement of adult mortality in less developed countries: A comparative review. *Population Index*, **57**, 552–568.

Álvaro Turriago-Hoyos, Mateus, W. A. M. and Thoene, U. (2020) Spatial analysis of multidimensional poverty in colombia: Applications of the unsatisfied basic needs (ubn) index. *Cogent Economics & Finance*, **8**.

United Nations (2017) Principles and recommendations for population and housing censuses. *Tech. Rep. 3*, United Nations, New York.

— (2022) World Population Prospects.

Urdinola, B. P. (2004) *Could political violence affect infant mortality? The Colombian case*. Ph.D. thesis, University of California, Berkeley.

Urdinola, B. P., Avilés, F. T. and Velasco, J. A. (2016) The homicide atlas in colombia: Contagion and under–registration for small areas. *Cuadernos de Geografía: Revista Colombiana de Geografía*, **26**, 101–118.

Urdinola-Contreras, B. P. (2018) War is not healthy. political violence and infant health outcomes in colombia. *Revista de Salud Pública*, **20**.

Vallejo, K., Tapias, J. and Arroyave, I. (2018) Trends of rural/urban homicide in colombia, 1992-2015: Internal armed conflict and hints for postconflict. *BioMed Research International*, **2018**.

Vallin, J. (1980) Socio-economic determinants of mortality in industrialized countries. *Population Bulletin of the United Nations*, **13**, 26–41.

Vallin, J. and Meslé, F. (2004) Convergences and divergences in mortality: A new approach of health transition. *Demographic Research*, **S2**, 11–44.

Van De Poel, E., O'donnell, O. and Van Doorslaer, E. (2009) What explains the rural-urban gap in infant mortality: Household or community characteristics? *Demography*, **46**, 827–850.

Vital Strategies (2020) Colombia rural vital: A strategy to strengthening civil registration and vital statistics in rural municipalities and in ethnic communities. *Case study*, Vital Strategies. Civil Registration Vital Statistics Program.

Vyas, S., Hathi, P. and Gupta, A. (2022) Social disadvantage, economic inequality, and life expectancy in nine indian states. *Proceedings of the National Academy of Sciences*, **119**, e2109226119.

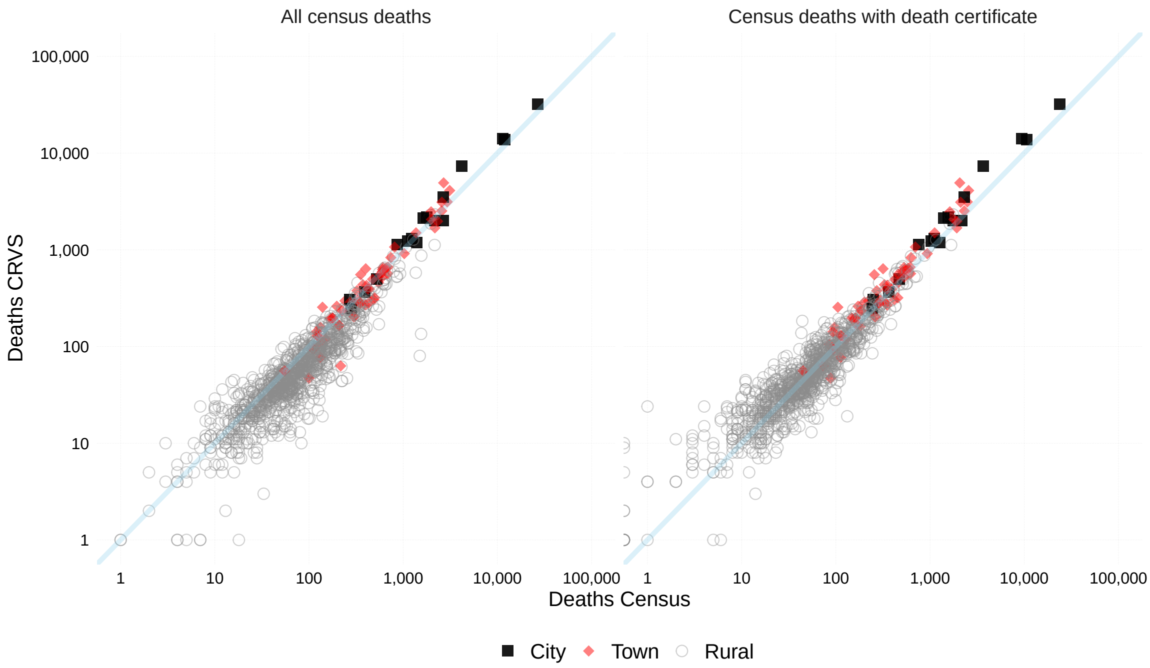
Woods, R. (2003) Urban-rural mortality differentials: An unresolved debate. *Population and Development Review*, **29**, 29–46.

Yaya, S., Uthman, O. A., Okonofua, F. and Bishwajit, G. (2019) Decomposing the rural-urban gap in the factors of under-five mortality in sub-saharan africa? evidence from 35 countries. *BMC Public Health*, **19**, 616.

Appendix A

In Figure 5, we show the total death counts in the CRVS system in the y-axis and the number of deaths reported by the census in 2017 in the x-axis. We see that both urban areas (cities and towns) fall close to the diagonal line, showing that they report similar numbers in both data sources. For rural areas, on the other hand, we see a high heterogeneity, with most areas falling below the line, indicating higher death counts in the Census than in the CRVS for rural areas.

Figure 5: Census reported deaths (all census deaths and only census deaths with death certificate) vs CRVS registered deaths in 2017 by municipality size.



When we look at age groups at the national level, we again see a heterogeneous profile regarding the census vs the CRVS reporting. In Figure 6, we see that the census outnumbers the civil registration by a factor of 3 for younger age groups (¡ 10) for both men and women in both urban and rural areas. For adult ages, on the other hand, the numbers are more similar in urban areas, and they get to be a bit higher for more advanced age groups (60-89) in the CRVS. For rural areas, more deaths are reported by the census in every age group, but the ratio diminishes with age.

Figure 6: Ratio between census reported deaths (all census deaths and only census deaths with death certificate) and CRVS registered deaths in 2017 by degree of urbanization and age group. Sources: DANE (2018, 2023).