Canadian Studies in Population

Official Journal of the
Canadian Population Society

Editor
Frank Trovato, University of Alberta

Assistant to the Editor
Ksenia Maryniak, University of Alberta

Vol. 43, No. 3–4 (Fall/Winter) 2016

Population Research Laboratory, University of Alberta

© 2016 by the Canadian Population Society and the Population Research Laboratory, University of Alberta, Edmonton
All rights reserved. Published 2016

ISSN 0380-1489
Frank Trovato, Editor
Ksenia Maryniak, Assistant Editor

Canadian Studies in Population (*CSP*) is indexed in the *Web of Science*, the largest and most comprehensive citation index available. All articles published in *CSP* since 2009 are indexed and abstracted in the Social Sciences Citation Index, Journal Citation Reports/Social Sciences Edition, and in Current Contents/Social and Behavioral Sciences.

The *Canadian Studies of Population* journal are available online at http://web.uvic.ca/~canpop/journal/.

Canadian Studies in Population was founded by Dr. P. Krishnan, Professor Emeritus at the University of Alberta, and is published with the financial support of the Society of Edmonton Demographers, Population Research Laboratory, Canadian Population Society, and the Department of Sociology, University of Alberta.

Funding for *Canadian Studies in Population* is provided by the Social Sciences and Humanities Research Council of Canada (SSHRC).

Contents

Canadian Studies in Population 43, no. 3–4 (2016): 175–202.

Regional variations of 1932–34 famine losses in Ukraine

Oleh Wolowyna[1]
Serhii Plokhy
Nataliia Levchuk
Omelian Rudnytskyi
Alla Kovbasiuk
Pavlo Shevchuk

Abstract

Yearly estimates of urban and rural direct losses (excess deaths) from the 1932–34 famine are presented for the oblasts of Soviet Ukraine. Contrary to expectations, the highest losses are not found in the grain-producing southern oblasts, but in the north-central Kyiv and Kharkiv oblasts. Several hypotheses are proposed and tested to explain this finding. No single hypothesis provides a comprehensive explanation. Losses in some oblasts are due to specific factors, while losses in other oblasts seem to be explained by a combination of economic and political factors. Quantitative analyses are presented of resistance and Soviet repressions in 1932, and effects of the food assistance program and historical-political factors on direct losses in 1933 are analyzed.

Keywords: 1932–33 famine losses by oblast; Holodomor; regional Holodomor losses; Ukrainian famine; urban and rural Holodomor losses.

Résumé

Des estimations annuelles de pertes (décès excédentaires) directement attribuables à la famine de 1932–34 sont présentées pour les zones urbaines et rurales d'Ukraine soviétique. Contrairement aux attentes, les pertes les plus importantes n'étaient pas dans la région méridionale productrice de grain, mais plutôt dans la région du nord-centre, soit Kiev et Kharkiv. Plusieurs hypothèses sont proposées et mises à l'épreuve pour vérifier cette conclusion. Cependant, aucune hypothèse, à elle seule, ne fournit une explication complète. Dans certaines régions, les pertes sont causées par des facteurs précis, alors que dans d'autres, les pertes sont expliquées par une combinaison de facteurs économiques et politiques. Des analyses quantitatives sont présentées sur la résistance et les répressions soviétiques en 1932. L'effet du programme d'assistance alimentaire et les facteurs politico-historiques attribuables directement aux pertes en 1933 est également analysé.

Mots-clés : pertes de la famine de 1932–33 par région; holodomor; pertes régionales de l'holodomor; famine en Ukraine; pertes urbaines et rurales de l'holodomor.

Introduction

The 1932–34 famine in Ukraine, also known as the *Holodomor* (death by hunger), is an extreme example of a man-made famine that resulted in millions of losses.[2] As a result of our research, Holodomor losses have been estimated at 4.5 million, with 3.9 million excess deaths and 0.6 million lost births (Rudnytskyi et al. 2015). *Direct losses* or *excess deaths* (these terms will be used interchangeably) are additional deaths caused by the famine; *indirect*

1. Oleh Wolowyna, Center for Slavic, Eurasian and East European Studies, University of North Carolina at Chapel Hill, 935 White Cross Rd., Chapel Hill, NC 27516 USA, e-mail: olehw@aol.com; Serhii Plokhy, Director, Harvard Ukrainian Research Institute; and Nataliia Levchuk, Omelian Rudnytskyi, Senior Researchers, and Alla Kovbasiuk and Pavlo Shevchuk, Researchers, Ptoukha Institute of Demography and Social Studies, at the National Academy of Sciences of Ukraine (NASU), Kyiv.

2. The widely accepted period for the Holodomor is 1932–33, but our research shows that there were also famine-related losses in 1934.

losses or *lost births* are births that did not occur due to the famine, i.e., they would have occurred had there been no famine. In this article, we present estimates of yearly direct Holodomor losses by oblast for urban and rural areas, and propose explanations for the differences found.

While numerous studies have attempted to estimate Holodomor losses for Ukraine, estimates at the regional level are scarce. S. Kulchytskyi (2003) and S. Maksudov (2012) analyzed mortality differentials at the oblast level, and Wheatcroft and Garnaut (2013) did the same at the raion level. However, these studies were based on registered deaths and did not attempt to estimate direct or indirect losses. Estimation of regional Holodomor losses is important for several reasons. First, it shows that the average national and urban-rural estimates hide significant regional differences. Second, it quantifies the losses in each region. Third, these data provide the demographic underpinnings necessary for historical analyses of the Holodomor and its consequences at the subnational level. Fourth, it helps us to better understand the dynamics of the Holodomor and its consequences.

The analysis presented here is based on our previous work on yearly estimates of direct Holodomor losses in Ukraine, by urban and rural areas and by age and sex (Rudnytskyi et al. 2015), and on a discussion of regional differences in direct Holodomor losses that is based on maps posted as part of 'The Great Famine' component of the *Mapa: Digital Atlas of Ukraine* program developed by the Harvard Ukrainian Research Institute (Plokhy 2016).

Oblast losses are presented without age and sex detail, mainly because estimates by age at the oblast level are based on relatively small numbers of registered deaths, which affects reliability. We provide a brief discussion of urban losses, but the emphasis is on rural losses. The dynamics of urban excess deaths are quite different from rural dynamics, and require a separate analysis. Estimates of oblast urban and rural losses are adjusted to the national urban and rural estimates presented in our previous work.

The loss estimates cover the administrative structure of Soviet Ukraine at the time of the famine, i.e., seven oblasts (Vinnytsia, Kyiv, Chernihiv, Kharkiv, Donetsk, Dnipropetrovsk, and Odesa), and also the Moldavian Autonomous Soviet Socialist Republic (ASSR). Our analysis shows distinct regional patterns in the spatial distribution of the direct losses.

NOTE: The Moldavian ASSR is included in our analysis because it was part of the Ukrainian SSR during the famine period. It was separated from Soviet Ukraine in 1940 and was not a part of it thereafter. To simplify the presentation, in some cases we will refer in the text to 'seven oblasts' instead of 'eight regions.'

Changes in administrative-territorial structure

Our estimation of losses is based on reconstruction of yearly populations for the eight regions of the Ukrainian SSR during the 1926–39 intercensal period. Several changes in the administrative structure during this interval had to be taken into account. The country was divided into 40 districts called *okrugs*[3] during the 1926–30 period; then the province-type *oblasts* were created to replace them, increasing progressively in number from 7 in 1932 to 15 in 1939. Also, an additional structure of six economic-geographical areas was in place during 1924–31 (Polissia, Right Bank, Left Bank, Dnipropetrovsk, Mountain Region, and Steppe). Furthermore, the country was divided into constantly fluctuating county-type *raions* during the whole 1926–39 period.

As all demographic data were recorded according to the administrative structure in place at the time of their collection, it was necessary to recalculate the data from different years to the seven-oblast-plus-Moldavia structure that was current during the 1932–34 famine period.

Recalculation of data into the seven-oblast structure

The seven oblasts of the Ukrainian SSR were in place between October 1932 and January 1937, and the territory of the Moldavian ASSR did not change during the 1926–39 period; thus, it was necessary for us to estimate transition coefficients from the other administrative structures to the seven-oblast structure for the other years in the research period. These coefficients were applied in order to recalculate population by age and sex,

3. The Russian term *okrug* (pl. *okruga*) has entered English usage and is therefore used here in roman type and pluralized accordingly. The equivalent Ukrainian term is *okruha* (pl. *okruhy*). In 1930–2 the okrugs were abolished and, after a further consolidation, replaced by oblasts.

births, deaths, and migration data. In the end, specific transition coefficients were estimated for the following periods: 1926–28, 1929–31, 1937, 1938, and 1939.

1926–28: From 40 okrugs to seven oblasts

Our transition coefficients from 40 okrugs to seven oblasts are based on detailed maps for the above-mentioned six economic areas that were published as part of the 1926 census. These maps show raion and okrug borders and allowed us to construct the oblasts based on these smaller units.

1929–31: From six economic-geographic zones to seven oblasts

Besides the normal tabulation by okrugs, the Central Statistical Administration of Soviet Ukraine (CSA UkrSSR) tabulated vital statistics by six economic-geographic areas during 1924–31. As vital statistics were not available by okrugs for the 1929–31 period—only for the six areas—our recalculation of vital statistics for this period was done in two steps: (1) estimation of transition coefficients from the six economic-geographical areas to the 40 okrugs; and (2) use of the transition coefficients estimated for the previous period, from the 40 okrugs to the seven oblasts.

The years 1937, 1938, and 1939

Two sets of transition coefficients were estimated for 1939: (1) for total populations—from 15 oblasts, as published for the 1939 census, to seven oblasts (Poliakov 1992; Korchak-Chepurkivskyi 1962); and (2) for populations by age and sex—from the 17-oblast structure in place in 1969 to seven oblasts (CSA USSR 1969). The first set of coefficients was used to estimate total populations for each oblast, yearly births, deaths by age and sex, and net migration. The second set of coefficients was used to estimate population by age and sex in 1939. Transition coefficients from 15 to seven oblasts were based on populations by raion that were published in the 1939 census. A similar methodology was used to estimate transition coefficients for 1937, from 11 to seven oblasts (five oblasts were subsequently added to the seven oblasts, on 22 September 1937), and for 1938 from 12 to seven oblasts. These coefficients were used to estimate population by age and sex for 1937, as well as births, deaths by age and sex, and net migration for 1937 and 1938.

Data and methods

Our reconstruction of the yearly demographic dynamics of the eight regions in the Ukrainian SSR for the 1926–39 intercensal period was based on the following data: 1926, 1937, and 1939 population censuses, 1931 urban count, rural-urban reclassification of population settlements, yearly numbers of births and deaths, and migration statistics.

Vital statistics are from the Russian State Archive of the Economy (RSAE) and the personal archives of the Ukrainian demographer Yuri Korchak-Chepurkivskyi. They include: (a) yearly births by sex, deaths by age and sex, and infant deaths by month of death—for urban and rural areas in 1927–29 and 1933–38; (b) only total number of births and deaths by sex for 1932; and (c) total number of births and deaths only, with no details by sex or urban/rural subpopulations, for 1930–31 (CSA UkrSSR 1927–32; ANER 1933a; RSAE 1562/20/41, 43, 46, 49, 59, 61, 62, 80, 86, 88, 121, 125, 153, 155; RSAE 1562/329/18, 20, 22, 33, 54, 56, 57, 114, 254, 261, 263, 264). That is, we have full yearly series of total number of births and deaths for all UkrSSR oblasts by urban and rural areas; what is missing are selected detailed data by age and sex (as well as rural-urban breakdown) for 1930, 1931, and 1932. It is important to note that complete birth and death data are available for the critical 1933 year.

Migration was registered only in urban areas during the 1920s and 1930s. Thus, the following information was used in our analysis: (a) yearly net migration for 1927–38; (b) yearly number of net migrants by sex and age and rural-urban migration streams for 1932–38 (RSAE 1562/20/22, 27, 29, 30, 38, 73, 75, 76, 118, 145). Our estimation of rural migration was based on a detailed compilation of various forced and voluntary migration streams from and to rural areas.

Estimation of overall Holodomor losses was based on a detailed reconstruction of the yearly populations in the eight regions during the 1926–39 period. The actual population dynamics were calculated by making relevant adjustments to census data, vital statistics, estimations of migration, and urban-rural reclassifications; the yearly populations were then calculated based on these components.

Adjustments of the three Soviet censuses are described below. We also adjusted the official 1931 urban count for the Ukrainian SSR to compensate for the undercount of children aged 0–4 years (ANER 1933b). The undercount for urban areas was distributed proportionately to the respective populations in the eight regions. After all the adjustments were made, we shifted populations from the dates of the three censuses and the urban count to the closest January 1 date.

Adjustment of 1926, 1937, and 1939 censuses

The 1926 and 1937 censuses are considered to be of good quality; the 1939 census, on the other hand, was deliberately falsified to cover up the huge population losses due to the Holodomor and other repressive measures revealed by the 1937 census (Andreev et al. 1990; Tolts 1995; Zhiromskaia 1990). Before using their data, we needed to make minor adjustments to the 1926 and 1937 censuses and major adjustments to the 1939 census. These corrections were applied to the official urban and rural population figures of the eight regions by sex and age, as published by the CSA USSR. The general methodology we used to make these adjustments is the same as the one we used in our previous work on Ukraine (Rudnytskyi et al. 2015); here we describe only the additional steps needed for adjustments at the regional level.

1926 census

We made two adjustments to the official 1926 census figures (CSA USSR 1929): redistribution of armed forces and adjustment of under-reporting for children aged 0–4 years. The census counted military personnel at the garrisons where they were stationed—mostly located in urban areas—thus introducing a significant distortion in the age structure of the urban population. We estimated the total number of armed forces stationed in the Ukrainian SSR at 121,200, by applying the proportion of the civilian population in Ukraine to the total USSR civilian population, 19 per cent, to the total armed forces in the Soviet Union. This estimate was distributed in the eight regions, proportionately to their urban and rural populations.

Our adjustment of the undercount of children aged 0–4 years was done using a methodology developed by Korchak-Chepurkivskyi (1928) for Ukraine. Adjustment coefficients were estimated for each region and were applied to the urban and rural areas of the region. The overall average adjustment for Soviet Ukraine was 0.8 per cent, with the following breakdown for the eight regions: 1.3 per cent in Dnipropetrovsk and Odesa oblasts, 1.2 per cent in Moldavia, 1.0 per cent in Donetsk oblast, 0.8 per cent in Kyiv oblast, 0.6 per cent in Kharkiv and Vinnytsia oblasts, and 0.5 per cent in Chernihiv oblast.

1937 census

As stated in our paper on Ukraine (Rudnytskyi et al. 2015: 57):

> The 1937 census was the first census conducted after the Great Famine, and it documented large population losses in Ukraine. It showed the total civilian population of Ukraine to be significantly lower than projected by central planners (the Central Economic Survey Administration of the USSR) and lower than in 1926. Given these unexpected results, the government declared the census 'defective' and its organizers were executed or exiled (Tsaplin 1989; Volkov 1990). Some of the 1937 census documents were destroyed, and the remaining results discredited because of supposedly flawed methods and organizational failures. Only in the late 1980s did the data from the 1937 census become available (Poliakov 1992), and it was shown that the census was executed correctly (Tolts 1989; Volkov 1990; Livshits 1990).

Population data at the oblast level are available for the urban and rural areas by sex, but oblast data by age were never tabulated (Poliakov 2007). As with the 1926 Soviet census, we made two adjustments: redistribution of armed forces and compensation of census undercount. Estimation of the armed forces in 1937 and their

distribution among the eight regions was done using the same methods as for the 1926 census. Thus, the total number of armed forces in Ukraine in 1937 was estimated at 346,800, based on the proportion of the civilian population in Ukraine relative to the USSR total. The undercount of the 1937 census was estimated by Andreev et al. (1990) at 0.43 per cent for the whole Soviet Union. As we did not have elements for estimating undercount for Ukraine and its oblasts, the same percent was used for each oblast.

1939 census

It was discovered in 1990 that the 1939 census, considered for many years a model for Soviet censuses, was seriously flawed. A sophisticated falsification plan had been implemented to hide large population losses that were already documented in the 1937 census (Zhiromskaia 1990). Our adjustments to the census data at the regional level were made using the same methodology as the overall adjustments for Ukraine (Rudnytskyi et al. 2015). They included the elimination of two types of falsification: (1) inflated undercount and inflated adjustment factors for control forms; and (2) reassignment—to place of residence at time of census—of the census forms of persons in forced labour camps, "special groups," and military personnel, which had been arbitrarily assigned to different parts of Ukraine.

Table 1. Adjustment steps for 1939 census populations of Ukrainian SSR, by region (in 1,000s)

Region	Reported civilian population	Special subpopulations			Corrections			Adjusted census population = (2) + (3) + (4) + (5) + (6) + (7)	Official census figures	% Adjustment (8) / (9)
		Army	Civilian population related to NKVD	Groups[1] A, B, C	Correction for undercount	Correction for control forms	Correction for 'unknown difference'			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ukraine total	29,269.2	380.7	8.0	194.3	82.4	113.4	94.6	30,142.6	30,946.2	−2.6
Vinnytsia	3,967.1	51.6	0.5	12.9	4.6	15.3	12.8	4,064.8	4,193.0	−3.1
Kyiv	5,084.9	66.1	1.6	39.8	16.2	19.7	16.5	5,244.9	5,394.0	−2.8
Chernihiv	2,564.4	33.4	0.4	10.2	4.9	9.9	8.3	2,631.4	2,721.3	−3.3
Kharkiv	5,352.7	69.6	1.7	41.5	12.3	20.8	17.3	5,516.0	5,654.7	−2.5
Donetsk	4,704.4	61.2	1.1	25.9	27.4	18.2	15.2	4,853.4	4,941.4	−1.8
Dnipropetrovsk	3,662.9	47.6	0.9	21.4	8.7	14.2	11.8	3,767.5	3,871.4	−2.7
Odesa	3,358.5	43.7	1.6	40.0	2.2	13.1	10.9	3,470.0	3,571.3	−2.8
Moldavian ASSR	574.2	7.5	0.1	2.6	6.1	2.2	1.9	594.6	599.2	−0.8

[1] A = NKVD; B = prisoners; C = forced resettlements

Sources: Poliakov (1991, 1992), Simchenko (1990), Kokurin and Petrov (2000), and authors' calculations.

We redistributed the armed forces, estimated at 380,700 for the Ukrainian SSR, among the oblasts by rural and urban areas using the same methodology as in the 1926 and 1937 censuses. Next, data on "special groups"—NKVD personnel, prisoners, and forced settlers—were available only for the whole country (Poliakov 1992). However, the distribution of these special contingents by oblast and rural-urban areas was published in the 1937 census (Poliakov 1991), and we used these 1937 proportions to redistribute the total numbers of special contingents in Ukraine by oblast in 1939. Data on the civilian NKVD staff contingent, also available only for Soviet Ukraine overall, was distributed by oblast and rural-urban areas proportionately to the oblast distributions of the special contingents. Comparing our resulting adjusted figures with the official census figures, we arrived at an overall inflation factor for Ukraine of 2.6 per cent; at the oblast level, the inflation factors vary between 0.8 per cent for Moldavia to 3.3 per cent for Chernihiv oblast (see Rudnytskyi et al. 2015 for more details).

Our adjustment of the 1939 populations by age and sex is based on official data published by the CSA USSR in 1969 and our transition coefficients from 17 to seven oblasts (see discussion in the previous section). These official data include 383,600 individual census records of prisoners living outside the Ukrainian SSR that had been arbitrarily redistributed in the rural areas of five oblasts in Ukraine—Vinnytsia, Kyiv, Odesa, Chernihiv, and Kharkiv—undoubtedly to artificially boost populations in rural areas that were decimated by the Holodo-

mor (Simchenko 1990). As the official civilian population by age and sex contained these extra 383,600 prisoners residing outside Ukraine, we subtracted them from the rural populations of the five oblasts, using the age-sex structure of the labour camp populations (Kokurin and Petrov 2000).

Adjustment of vital statistics for under-registration

Registered numbers of births and deaths were distorted by different degrees of under-registration during the intercensal period; during the famine years, levels of under-registration reached extremely high proportions. The general adjustment approach used for urban and rural areas of Soviet Ukraine is described in Rudnytskyi et al. (2015), and the same approach was used for each of the eight regions. Only a brief conceptual description of the adjustment methodology is presented here.

We made adjustments along three dimensions: (1) crisis (1932–34) and non-crisis (1927–31 and 1935–39) periods; (2) urban, rural and total; and (3) three vital events: births, infant deaths, and deaths after one year of age. The adjustment methods for the three vital events differed, depending on the related dimension. Namely, the same adjustment methods were applied to the three events in urban areas during crisis and non-crisis periods, while different methods were applied (to the three vital events) during crisis years for rural and total populations. In almost all cases, adjustments for rural areas were calculated as the difference between total and urban values. Ukrainian demographers did extensive research on this topic in the 1930s, and we took full advantage of their work in our adjustment methodology.

Estimation of net migration by oblast, 1927–38

Migration is difficult to estimate, as migration statistics are incomplete and fragmentary. Estimation of migration for urban areas is less problematic than for rural areas, as there was a migration registration system in place in cities during this period, while no such system existed in rural areas. In 1932 the urban registration system was improved by the introduction of registration cards for all arrivals and departures in most cities of Soviet Ukraine (Popov 1995). However, urban migration statistics are problematic, requiring systematic evaluation and adjustments for under-registration. Rural migration estimates had to be pieced together using different statistical sources and archival documents. Once the yearly rural and urban migration was estimated for each region, the respective totals were adjusted to the yearly net migration data for urban and rural areas of the Ukrainian SSR (Rudnytskyi et al. 2015).

Urban migration

The following data sources were used for estimating urban migration by oblast: (1) numbers of net migrants for Ukraine, 12 separate oblasts and the Moldavian ASSR for 1927–38, compiled by ANER without sex, age, and flow details (RSAE 1562/20/73); and (2) number of migrants by arrivals and departures, as well as net migrants, during 1933–38, by sex and age for all urban centers in each oblast (RSAE 1562/20/30, 38, 75, 75, 18, 145). We also had migration data for 1932 by arrivals and departures by sex, age and migration flows (RSAE 1562/20/27). The yearly numbers of net migrants, calculated by ANER for the 1927–38 period and for 12 oblasts, were recalculated by us for the seven oblasts using the appropriate transition coefficients. We then used these estimates of net migrants as the basis for estimating yearly numbers of urban migrants.

Our estimation of numbers of net urban migrants in each oblast was done for three separate periods, 1927–30, 1931–36, 1937–38, using the same method as for urban Ukraine (Rudnytskyi et al. 2015). Yearly disaggregation of the net migrants was done proportionately to the yearly number of registered net migrants. Thus, the total number of net urban migrants for the 1927–38 period is 3,792,200 (see Table 2).

Vinnytsia was the only oblast with negative net urban migration for the 1927–38 period. Migration was a significant factor in the urban growth of Donetsk and Dnipropetrovsk oblasts, with total net migration equal to 159 per cent and 142 per cent of their respective 1927 urban populations. For the other oblasts, this figure varied between 36 and 63 per cent; the contribution of net migration to Moldavia's urban growth was negligible.

Rural migration

We classified rural migration into two types of streams: internal and external. The main internal stream is *rural-to-urban migration for the 1927–38 period*. According to the urban registration system, of the total net urban migration about 81 per cent were rural-to-urban migrants, i.e., 3,085,800. We did a yearly distribution of this total according to the yearly distribution of rural-to-urban migrants in the urban registry system. Then, within each year we distributed the number of migrants among the oblasts proportionately to their rural population size.

The second internal stream is *organized inter-oblast migration from rural to rural areas during 1934–35*. Data on these migration streams can be found in Iefimenko (2013) and Iukhnovskyi et al. (2008). In 1934, 16,200 families were resettled from Vinnytsia, Kyiv, and Chernihiv oblasts to Odesa, Kharkiv, Donetsk, and Dnipropetrovsk oblasts; in 1935, 9,800 families were resettled from Vinnytsia and Kyiv oblasts to Donetsk, Dnipropetrovsk, and Kharkiv oblasts.

External rural migration is composed of nine streams, seven out-migration and two in-migration streams:

1. *Persons sent to labour camps (gulags) and working colonies, 1929–38*. Sources for these data are: Nikolskyi 2001; Mozokhin nd; Zemskov 2005; and Tronko et al. 1994–2011; the total number of prisoners is 284,600. Detailed information on this type of emigration by oblast is available for 1937 and 1938; about 90 per cent of these migrants were males (Golotik and Minaev 2004). The 64,300 sent to penal camps in 1937–38 were distributed by oblast as follows: 24,200 in Vinnytsia, 10,700 in Kyiv, 8,200 in Donetsk, 6,300 in Kharkiv, 5,700 in Odesa, 5,200 in Dnipropetrovsk, 2,600 in Chernihiv, and 1,400 in Moldavia.

   For the other years, 1929–36, reliable statistics are available only for Ukraine, and yearly estimates for Ukraine were calculated in Rudnytskyi et al. (2015). These yearly numbers of migrants were distributed by oblast using the proportions available for 1937–38.

2. *Eviction of kulaks,*[4] *1930–33*. Data on this migration stream can be found in: SARF 9414/1/1943, 1944; SARF 9479/1/2; Yakovlev et al. 2005; and Bugai 2013; the total number of evicted kulaks is 364,500. Detailed information on this migration is available for 1930 by 40 okrugs, and we recalculated the data for the seven oblasts. The 111,400 kulaks evicted in 1930 are distributed by oblast as follows: 24,400 in Odesa, 19,600 in Kharkiv, 18,100 each in Vinnytsia and Kyiv oblasts, 17,500 in Dnipropetrovsk, 7,200 in Donetsk, 3,600 in Chernihiv, and 3,100 in Moldavia. Total numbers for 1931, 1932, and 1933 were only available for Ukraine, and we redistributed them by oblast using the proportions from 1930. The yearly totals of evicted kulaks were 194,100 in 1931, 15,000 in 1932, and 44,000 in 1933.

3. *Forced emigration of peasants, 1929–33*. Statistics on this migration category are fragmentary and unreliable, as most of it took place during the Holodomor (RSAE 1562/20/22, 29, 30, 73; Vynnychenko 1994). The estimate is 532,200; yearly estimates were taken from Rudnytskyi et al. (2015), and numbers by oblast were distributed proportionately to the respective rural populations.

4. *Organized mass resettlements of peasants, 1927–30*. These resettlements were a continuation of previous campaigns to resettle peasants from Soviet Ukraine to Siberia and the Far East. A total of 120,000 were resettled in 1927–30 (Hirshfeld 1930; Platunov 1976; Rybakovskii 1990). The yearly overall numbers of migrants for Ukraine were distributed by oblast proportionately to their rural population.

5. *Deportation of Poles and Germans to Kazakhstan in 1936*. From areas in Vinnytsia and Kyiv oblasts bordering Polish-occupied Galicia and Volhynia, 14,900 peasant families (or 60,000 persons) were deported in 1936 to Kazakhstan (Yakovlev et al. 2005; Bugai 2013; Iefimenko 2013). We distributed this number by oblast proportionately to their rural populations.

6. *Emigration of Jews, 1929–38*. An ethno-demographic balance methodology was used by Rudnytskyi et al. (2015) to estimate the number of Jews who emigrated from Ukraine during this period. The number of

4. The Russian term *kulak* has entered English usage and is therefore used here in roman type and pluralized accordingly. The equivalent Ukrainian term is *kurkul* (pl. *kurkuli*).

Jewish emigrants thus estimated for the intercensal period was distributed yearly, based on information from the following sources: Hirshfeld 1930; Weitsblit 1930; Vynnychenko 1994; Leskova 2005; Rudnik 2006. Approximately 57,000 Jews emigrated from Ukraine during this period, and this number was distributed among the oblasts, proportionately to the number of Jews in each oblast.

7. *Peasants hired to work on construction projects in other parts of the Soviet Union, 1935–38***.** Information about this migration stream is fragmentary (Kozin 1936; Vynnychenko 1994; RSAE 1562/20/73, 75, 76, 118, 143, 145). According to our calculations, about 170,500 peasants from Vinnytsia, Kyiv, Chernihiv, and Odesa oblasts, as well as Moldavia, were involved in this state-run initiative. The yearly estimates calculated in Rudnytskyi et al. (2015) were distributed proportionately to the rural population of the oblasts and Moldavian ASSR.

8. *Resettlement of peasants from Belarus and Russia to Ukraine during 1933–34***.** The 1932–34 famine left many villages in Ukraine practically empty, and the Soviet government decided to settle these villages with peasants from Belarus and Russia. Sources on these resettlements are: Iefimenko (2013) and CSANO (1/2/6583–85, 6392). During the second half of 1933 a total of 27,100 families (137,800 persons) were resettled from Belarus and from the following four regions of the RSFSR: Gorky (Nizhnii Novgorod) Krai and Yaroslavl, Western, and Central Black Earth oblasts. They were resettled in the following oblasts of the UkrSSR: 44,300 in Kharkiv, 39,600 in Dnipropetrovsk, 34,600 in Odesa, and 19,300 in Donetsk. However, a portion of these settlements turned out to be temporary; Iefimenko presents data that by March 1935, at least half of the settlement populations had left (2013: 143–48).

9. *Resettlement of kulaks from Central Asia to Ukraine in 1931*. The policy of destroying kulaks as a class was not limited to the European regions of the Soviet Union; it also affected wealthy farmers in Central Asia. More than 40 villages in Odesa oblast were recipients of peasants from Uzbekistan branded as kulaks (Vynnychenko 1994; Smolii et al. 2003; Zemskov 2005). According to our calculations, this contingent had about three thousand families (16,000 persons).

Our team systematized, evaluated, analyzed, and organized all these data into yearly numbers of net migrants by rural area in each oblast, resulting in −4,400,300 total net rural migrants. Net rural migration for the 1927–38 period was negative for all oblasts. Dnipropetrovsk oblast had the largest net migration, with −1,017,300. Adding net urban and rural migration, we obtained −608,100 net migrants for Ukraine (Table 2).

More detailed information discovered about urban-rural reclassification during our oblast estimates resulted in some changes in our total urban and rural net migration numbers compared to previous estimates for Ukraine (Levchuk et al. 2015; Rudnytskyi et al. 2015). The previous number of 4,108,000 net urban migrants changed to 3,792,200, and the previous number of −4,826,400 net rural migrants changed to −4,400,300. There were also minor adjustments to total number of deaths, from 8,519,600 to 8,640,100 in rural areas, and from 1,650,000 to 1,639,400 in urban areas, with a total of 10,279,500 deaths in Soviet Ukraine. These adjustments resulted in minor changes to the yearly balances of urban and rural areas by oblast. As a result of these changes, the number of direct losses for urban areas decreased by 2 per cent, and the number of direct losses for rural areas increased by 0.2 per cent; the total number of direct losses remained the same.

Population reconstruction by oblast, 1927–39

Our reconstruction of yearly populations for urban and rural areas of the eight regions was done in two steps: first for total population and then by age and sex. Having adjusted the census populations as well as numbers of births, deaths, and net migration, we needed one more element to reconstruct the urban and rural populations for the seven oblasts and Moldavia—namely, rural-urban reclassifications. They were implemented in 1930, 1936, and 1938, were fairly extensive, and had significant implications for the sizes of the urban and rural areas in many oblasts.

Table 2. Total population balance for Ukrainian SSR and its region, by urban-rural areas, 1927–38 (in 1,000s)

A – Total

Region	Population on 1 Jan. 1927	Births	Deaths	Net migration	Population on 1 Jan. 1939	% annual natural rate	% annual total rate
Ukraine total	29,316.3	11,685.0	10,279.5	−608.1	30,113.8	0.42	0.22
Vinnytsia	4,405.1	1,678.2	1,542.7	−486.3	4,054.3	0.29	−0.69
Kyiv	5,877.6	2,182.0	2,328.8	−495.4	5,235.4	−0.17	−0.96
Chernihiv	2,812.6	1,010.8	850.3	−339.6	2,633.5	0.49	−0.55
Kharkiv	5,784.4	2,074.7	2,195.0	−159.2	5,504.9	−0.14	−0.41
Donetsk	3,007.5	1,746.8	1,128.6	1,221.1	4,846.9	1.57	3.98
Dnipropetrovsk	3,548.9	1,509.6	1,077.8	−203.2	3,777.6	0.98	0.52
Odesa	3,302.7	1,220.4	950.9	−103.1	3,469.1	0.68	0.41
Moldavian ASSR	577.5	262.5	205.4	−42.4	592.2	0.82	0.21

B – Urban

Region	Population on 1 Jan. 1927	Births	Deaths	Net migration	Urban–rural reclassification	Population on 1 Jan. 1939	% annual natural rate
Ukraine total	5,322.4	2,463.2	1,639.4	3,792.2	1,103.6	11,041.8	1.20
Vinnytsia	537.2	154.7	104.4	−28.9	−18.8	539.2	0.75
Kyiv	1,065.5	321.2	250.1	376.8	2.2	1,515.1	0.54
Chernihiv	344.5	101.0	75.7	215.9	−138.9	447.0	0.59
Kharkiv	981.5	357.2	261.8	585.6	238.4	1,900.8	0.77
Donetsk	942.8	889.3	521.1	1,498.2	760.6	3,570.7	2.75
Dnipropetrovsk	573.4	357.8	210.5	814.1	204.8	1,740.5	1.91
Odesa	797.8	251.8	198.5	328.3	27.3	1,206.3	0.54
Moldavian ASSR	79.6	30.1	17.3	2.1	27.9	122.3	1.24

C – Rural

Region	Population on 1 Jan. 1927	Births	Deaths	Net migration	Urban-rural reclassification	Population on 1 Jan. 1939	% annual natural rate
Ukraine total	23,994.0	9,221.8	8,640.1	−4,400.3	−1,103.6	19,071.8	0.20
Vinnytsia	3,867.9	1,523.5	1,438.2	−457.5	18.8	3,514.4	0.18
Kyiv	4,812.1	1,860.8	2,078.7	−872.2	−2.2	3,719.7	−0.39
Chernihiv	2,468.1	909.8	774.6	−555.5	138.9	2,186.6	0.44
Kharkiv	4,802.9	1,717.5	1,933.2	−744.8	−238.4	3,604.1	−0.38
Donetsk	2,064.8	857.5	607.6	−277.1	−760.6	1,277.0	0.95
Dnipropetrovsk	2,975.6	1,151.8	867.3	−1,017.3	−204.8	2,037.9	0.76
Odesa	2,504.8	968.6	752.4	−431.4	−27.3	2,262.3	0.69
Moldavian ASSR	497.9	232.4	188.1	−44.5	−27.9	469.8	0.71

In 1930 the CSA UkrSSR approved a new classification of urban settlements, with significant differences compared to the classification used for the 1926 census. In most oblasts, many urban settlements were reclassified as rural settlements. Criteria used in this reclassification were: small population size and high percentage of the population in agriculturally related occupations. Donetsk and Dnipropetrovsk oblasts were the exceptions to this pattern, where many rural settlements were reclassified as urban settlements. The net result of rural-urban reclassifications in 1930 was an increase in the rural population by 194,100.

Subsequently, in preparation for the 1937 and 1939 censuses, new rounds of rural-urban reclassification were implemented in 1936 and 1938. In both cases, the reclassification was in one direction for all oblasts, from rural to urban, and it was quite extensive. In 1936 the reclassification process increased the urban population of Ukraine by 1,077,700, and in 1938 by 219,900. Two-thirds of these increases occurred in Donetsk and Dnipropetrovsk oblasts (ANER 1933b; UCEC 1933, 1936; CANER 1936; SS USSR 1938, 1939).

Final reconstructed total populations of the eight regions for 1927 and 1939 are presented in Table 2. During the intercensal period, the Ukrainian SSR had an overall annual average natural exponential growth rate (births minus deaths) of 0.4 per cent. Two oblasts, Kyiv and Kharkiv, lost population, while the other oblasts had an annual average natural growth rate of less than one per cent, except Donetsk, with a yearly rate of 1.6 per cent.

The urban population of Ukraine grew at an annual average natural growth rate of 1.2 per cent, and urban areas in all oblasts had positive growth. Donetsk had the highest annual average natural growth rate, with 2.7 per cent, followed by Dnipropetrovsk oblast with 1.9 per cent and Moldavia with 1.2. For the other oblasts, rates varied between 0.5 and 0.8 per cent. The average annual natural rate of growth for all rural areas was only 0.2 per cent, and Kyiv and Kharkiv oblasts had negative yearly natural growth of −0.4 per cent each. For all the other oblasts, the yearly natural rate of growth was positive, albeit quite small.

Direct Holodomor losses

Direct losses are estimated as the difference between the number of deaths occurring during the famine years, and the hypothetical number of deaths had there been no famine during the same period. We estimated the number of hypothetical deaths had there been no famine using linearly extrapolated age-specific deaths rates between 1931 and 1935, i.e., years before and after the famine, at a time when mortality was considered 'normal.'

Our results for direct losses by oblast are presented in Table 3 in three panels: total, urban, and rural. The total number of direct losses for Ukraine is estimated at 3.9 million, with 250,000 in 1932, 3,529,000 in 1933, and 163,000 in 1934 (panel A). Most of the losses occurred in 1933 in all oblasts. In five oblasts, 90 per cent or more of the losses occurred in 1933, and the percentages for Donetsk and Chernihiv oblasts and Moldavia are 76, 81, and 85, respectively.

Kyiv oblast had the highest number of losses with 1,111,000, followed by Kharkiv with 1,038,000 and Vinnytsia with 546,000. Direct losses for Chernihiv, Donetsk, Dnipropetrovsk, and Odesa oblasts vary between 254,000 and 368,000, and Moldavia has the smallest number of direct losses, 68,000.

As the number of losses is directly related to population size, to make valid comparisons among oblasts it is necessary to control for population size. Focusing first on 1933, the total number of excess deaths for Soviet Ukraine in 1933 is 119 deaths per 1,000 population. Kyiv and Kharkiv oblasts have the highest losses, with 179 excess deaths per 1,000 population each, followed by Vinnytsia oblast with 115 and Moldavia with 102; Donetsk oblast has the lowest value, 41 direct losses per 1,000 population. The relative number of losses for Ukraine was lower in 1934 than in 1932, and this is the case for all except Chernihiv oblast, where relative losses were higher in 1934 than in 1932.

The ratios in the last column provide a summary of relative losses for the three famine years, calculated as the total number of losses for the three years, divided by the 1933 mid-year population and multiplied by 1,000. If we divide this indicator by 10, it can be used as an approximation of the per cent of the 1933 population that died due to famine. Thus, we can say that the total number of excess deaths constituted about 13 per cent of the 1933 UkrSSR population; Kyiv oblast had the highest value with 20 per cent, and Donetsk oblast had the lowest with 5 per cent.

The total urban excess deaths in Soviet Ukraine were 293,000, with 49,000 in 1932, 194,000 in 1933 and 51,000 in 1934 (Table 3, panel B); 66 per cent of all urban excess deaths occurred in 1933. Regarding excess deaths per 1,000 population, the yearly ratios are 7.0, 26.8, and 6.9 respectively, and almost 40 for the 1932–34 period. The concentration of urban direct losses in 1933 varies from around 80 per cent in Kharkiv and Odesa oblasts to 41 per cent in Donetsk oblast.

Of the total 3.9 million excess deaths, 3.6 million occurred in rural areas, and 91 per cent of them occurred in 1933. This high concentration of rural excess deaths in 1933 is found in all oblasts, from 82 per cent in Chernihiv oblast to 94 per cent in Kharkiv oblast. Yearly direct losses per 1,000 population are 8.3, 149.4, and 5.5, totalling almost 164 for the 1932–34 period.

Table 3. Direct losses (excess deaths) from the Holodomor in Ukrainian SSR, in numbers and by 1,000 population, by region and rural-urban areas

A – Total								
Region	Thousands				Per 1,000 population			
	1932	1933	1934	1932–34	1932	1933	1934	1932–34*
Ukraine total	250.0	3529.2	163.3	3942.5	8.0	119.3	5.9	133.3
Vinnytsia	27.0	497.3	21.1	545.5	5.9	114.6	5.2	125.6
Kyiv	83.9	991.5	35.4	1110.8	13.7	178.7	7.0	200.3
Chernihiv	17.6	205.4	31.2	254.2	6.0	73.7	11.9	91.3
Kharkiv	46.9	969.9	20.8	1037.6	7.8	178.9	4.2	191.4
Donetsk	28.7	175.2	26.9	230.8	7.0	41.1	6.4	54.2
Dnipropetrovsk	20.6	331.3	16.5	368.4	5.4	91.6	4.7	101.9
Odesa	19.6	300.3	7.0	326.9	6.1	98.8	2.4	107.6
Moldavian ASSR	5.7	58.2	4.4	68.3	9.6	102.4	8.1	120.2
B – Urban								
Region	Thousands				Per 1,000 population			
	1932	1933	1934	1932–34	1932	1933	1934	1932–34*
Ukraine total	48.8	193.9	50.6	293.4	7.0	26.8	6.9	39.7
Vinnytsia	3.1	14.6	2.0	19.7	7.6	35.8	5.0	48.4
Kyiv	11.2	44.4	1	65.8	11.5	46.7	10.5	69.2
Chernihiv	2.6	11.1	3.8	17.4	9.0	39.6	13.7	62.4
Kharkiv	6.2	45.7	5.0	56.9	5.0	36.3	4.0	45.3
Donetsk	15.2	24.1	19.6	58.9	7.3	10.4	8.4	25.6
Dnipropetrovsk	5.0	23.3	7.1	35.4	5.0	22.7	6.7	34.5
Odesa	4.7	29.0	2.8	36.6	5.0	30.7	3.0	38.7
Moldavian ASSR	0.7	1.9	0.0	2.7	9.3	24.5	0.4	34.4
C – Rural								
Region	Thousands				Per 1,000 population			
	1932	1933	1934	1932–34	1932	1933	1934	1932–34*
Ukraine total	201.2	3335.3	112.7	3649.1	8.3	149.4	5.5	163.7
Vinnytsia	23.9	482.8	19.1	525.7	5.7	122.7	5.2	133.6
Kyiv	72.7	947.1	25.2	1045.1	14.2	206.0	6.2	227.3
Chernihiv	15.0	194.4	27.4	236.8	5.7	77.6	11.7	94.5
Kharkiv	40.8	924.2	15.7	980.7	8.6	222.0	4.3	235.6
Donetsk	13.4	151.2	7.3	171.9	6.6	77.3	3.8	87.9
Dnipropetrovsk	15.6	308.0	9.3	333.0	5.5	118.9	3.8	128.5
Odesa	14.9	271.2	4.2	290.3	6.5	129.6	2.1	138.7
Moldavian ASSR	5.0	56.3	4.4	65.7	9.6	114.7	9.4	133.7

Note: Summary indicator * represents total number of direct losses in 1932–34/1933 mid-year population.

Discussion

This section is based on the discussion of regional differences of Holodomor direct losses posted on the website *Mapa: Digital Atlas of Ukraine* (Plokhy 2016). We systematized and quantified arguments presented in that paper, and elaborated the discussion with new elements.

As indicated above, the Holodomor dynamic in urban areas is very different from the one in rural areas and requires a separate, more detailed analysis. We present here a brief discussion of the urban losses, and then proceed to analyze in detail the spatial distribution of the rural direct losses.

Spatial distribution of excess deaths in urban areas

Research on the Holodomor has focused mainly on rural areas; our research shows that urban areas were also significantly affected by this famine. Relative 1932–34 direct losses represent 6.9 per cent and 6.2 per cent

of the urban populations in Kyiv and Chernihiv oblasts, respectively, while in the other oblasts they vary between 2.6 per cent in Donetsk and 4.5 per cent in Vinnytsia oblasts (Table 3).

Rural losses are expected to be always higher than urban losses. However, we see that urban losses are higher than rural losses in 1934, namely, 6.9 and 5.5 per 1,000 persons, respectively. This surprising result can be explained by considering the systemic relationship between urban and rural areas during the Holodomor.

As a result of the Soviet government's policy to control agricultural production through the collectivization of farms, the state assumed direct responsibility for providing food to the urban population. Due to increasing shortages of food in cities, in 1931 the Politburo approved the resolution 'On the introduction of a single system of supply for the working population by ration books.' Key elements this resolution and its consequences are described as follows (emphasis added):

> Only those who worked in the state sector of the economy (industrial factories, state and military organizations and departments, and state farms) and their families received ration cards. Peasants and the politically disenfranchised were left out of the state food supply system. These people made up more than 80 per cent of the total population. Even ration sizes depended on how important people were to the industrialization process… From the beginning of 1931 there were four types of rations throughout the country: *special, first, second,* and *third*. They were called *city lists*, but in reality they were groupings of enterprises and organizations, because factories in the same city could be on different supply lists. The special and first lists had priority and included key industries in Moscow, Leningrad, Baku, the Donbas, Karaganda, Eastern Siberia, the Far East, and the Urals. Constituting only 40 per cent of the total number of people on rations, they received nearly 80 per cent of all the food supplies. The second and third lists included smaller and non-industrial cities' (Osokina 2001: 61–62).

The ration system was affected by two opposite processes during the famine years: a rapid increase of the urban population triggered by Stalin's industrialization policy, and diminishing food production due to peasant opposition to collectivization and increasing mismanagement of the agricultural sector (Levchuk et al. 2015). The result was a gradual diminishing of the official food ration amounts, especially on the lower ration lists; thus, an increasing proportion of the urban population ended up without any food assistance. By 1934, starvation reached critical levels in many cities, resulting in higher relative excess deaths in urban than in rural areas.

Spatial distribution of excess deaths in rural areas

In this section we discuss different factors that may explain the variable and unexpected regional distribution of direct losses caused by the Holodomor, as shown in Map 1.

We start by presenting four hypotheses that have been suggested for the expected distribution of losses. Next, as the number of excess deaths experienced a drastic change between 1932 and 1933, we first examine factors related to the onset of the famine and regional distribution of direct losses in 1932. At the end of 1932, the north-central oblasts of the Ukrainian SSR had lower levels of collectivization and higher levels of grain quota fulfillment than the southern oblasts. This apparent contradiction leads us to examine, in the third section, the resistance to collectivization and grain procurement, and the repression of this resistance by the Soviet government. The explosion of excess deaths during the first half of 1933, and its relationship with the food 'assistance' program, are examined in the fourth section.

Four hypotheses

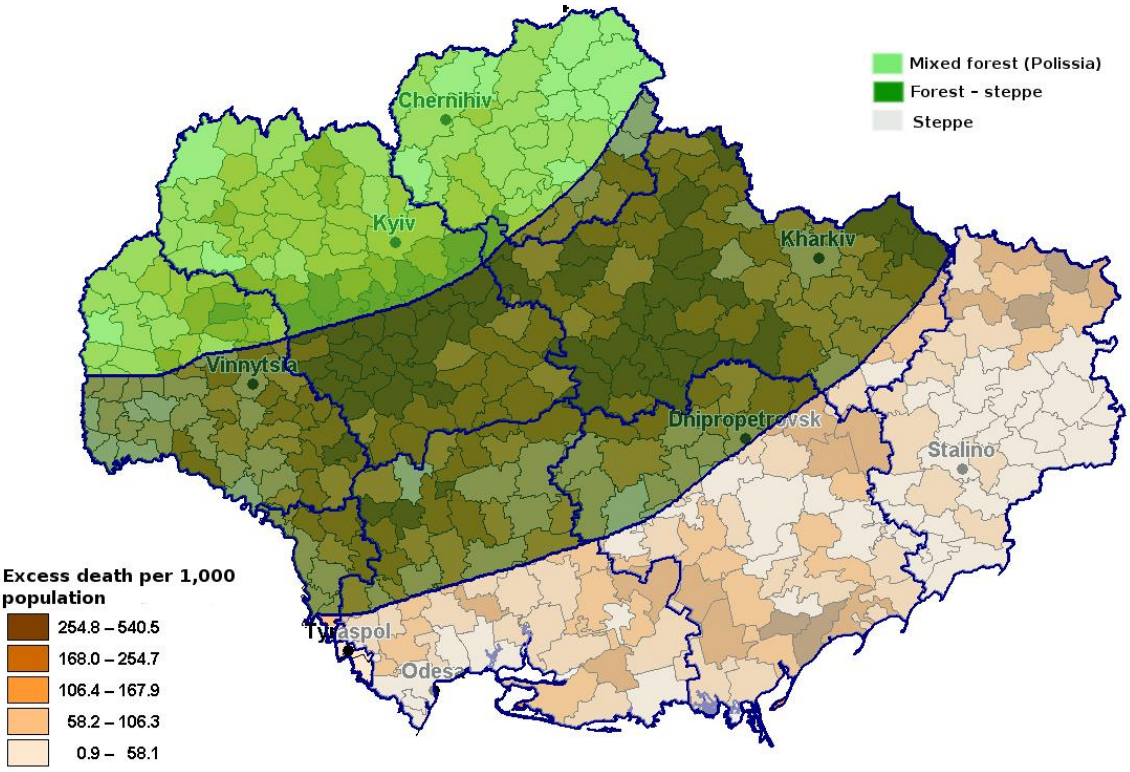
Several hypotheses have been suggested to explain the spatial variation in levels of excess deaths in rural areas of Ukraine: historical, ecological, border, and economic.

Historical hypothesis. The 1921–23 famine affected the southern grain-growing regions of Ukraine, and it was natural to assume that the 1932–34 famine would also have a more pronounced effect on these regions (Plokhy 2016: 378). As can be seen on Map 1, this was not the case for the 1932–34 famine; the highest losses are found in the north-central oblasts of Kyiv and Kharkiv.

Ecological hypothesis. Ukraine can be divided into three natural zones: mixed forest or *Polissia*, forest-steppe, and steppe. Polissia is covered with forests and wetlands and has rich natural vegetation, the steppe region is a vast plain covered with grass and little or no trees, and the forest-steppe zone is a transition zone, with forests

Map 1. Number of rural excess deaths per 100 population by oblast, Ukrainian SSR, 1932–34.

in the north and steppes in the south. According to the ecological hypothesis, the expectation is that the relative number of excess deaths should be lowest in the Polissia zone and highest in the steppe zone. The rationale is that once most of the grain, and in many cases all food, was confiscated by the Soviet government, people in Polissia could find some food in the forests and swamps, while no alternative food was to be found in the steppe zone. Losses in the forest-steppe zone are expected to be in the middle range.



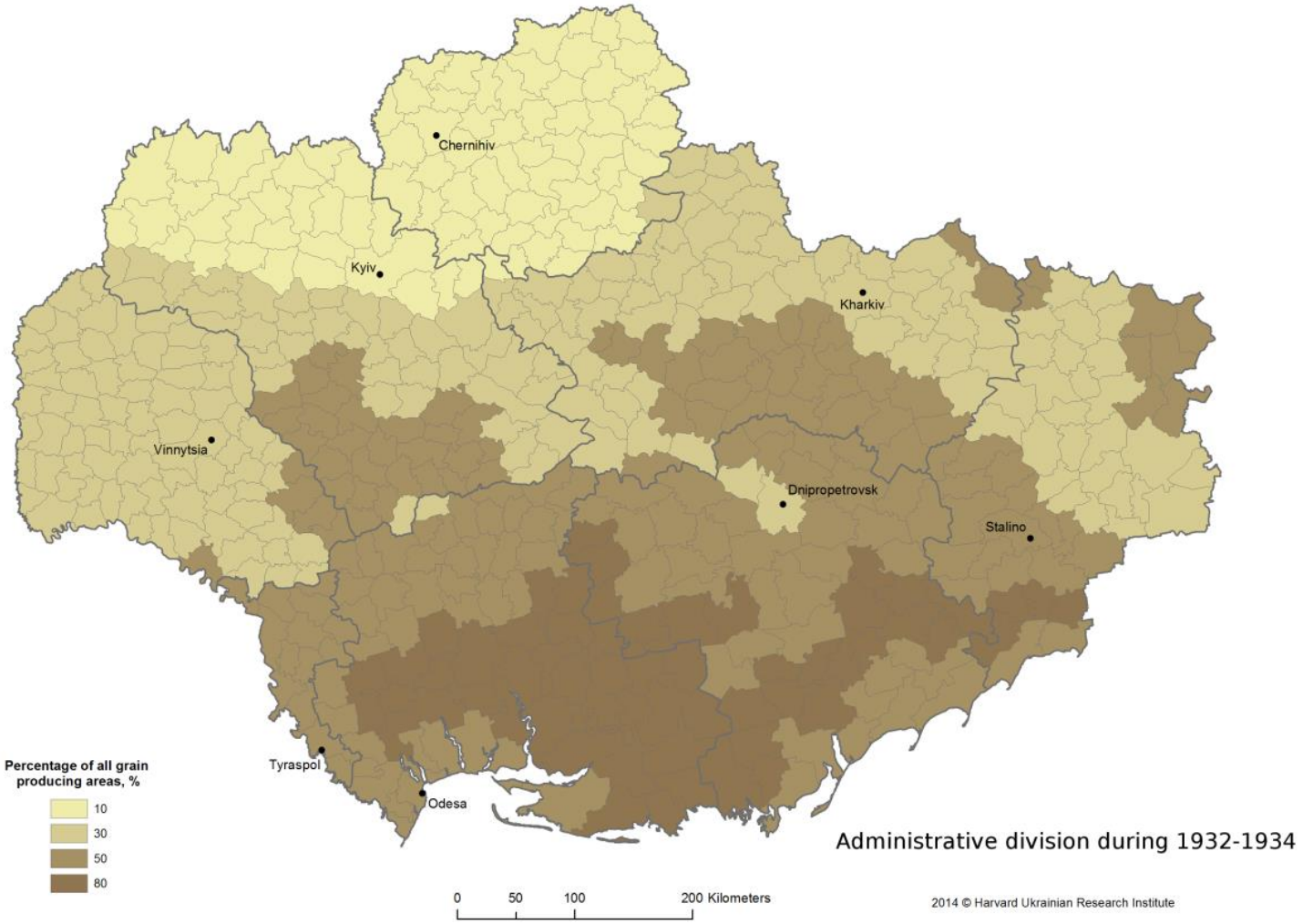
Map 2. Number of rural excess deaths per 1,000 population by raion, Ukrainian SSR, 1932–34.

Unfortunately, the seven-oblast administrative structure of the Ukrainian SSR at that time does not provide a clear picture of the situation in the ecological zones, as they cut across oblast borders. However, detailed vital statistics available for 1933 allow us to estimate excess deaths at the raion level for that year, and Map 2 provides the resulting details. It is clear from this map that the ecological hypothesis does not explain the regional variations in levels of excess deaths, as the highest relative direct losses are mainly in the forest-steppe zone, not in the steppe zone.

Border hypothesis. Map 2 also shows that the relative number of direct losses becomes lower the closer one gets to the international borders of Kyiv, Vinnytsia, and Moldavia with Poland and Romania. This pattern is consistent with the border hypothesis formulated by Shlyakhter (ch. 10, p. 1–2):

> Exploring some of the striking regional variations in the famine's severity, this paper argues that these differences resulted from a combination of official policies and the survival strategies of border strip inhabitants… In addition to offering an explanation for the lower mortality in Ukraine's border districts during the Holodomor, this analysis also views the famine as a window onto Soviet security and revealshowcasing policies in the border strip, peasant survival strategies, and the interplay between the two.

Economic hypothesis. Given the failure of the ecological hypothesis to explain the spatial variations in excess deaths, Plokhy (2016: 379) suggests that 'on the eve and in the course of the Great Ukrainian Famine, environmental factors influenced human actions, particularly government policies that eventually contributed to the death toll.' Specifically, Moscow focused its attention on the grain-producing areas of southern Ukraine, as they had the optimal capacity to produce the grain needed for the implementation of Stalin's policies, while other regions were left basically to their own devices.



Map 3. Wheat growing areas of the Ukrainian SSR, 1937.

Map 3 shows the distribution of wheat growing areas in 1937, and it can be used as an approximation of grain growing areas in 1932–34. Major grain growing areas are located in Odesa, Dnipropetrovsk, and Donetsk oblasts and Moldavia, while only 10 per cent of Chernihiv oblast is dedicated to grain crops. The agriculture of Kyiv, Kharkiv, and Vinnytsia oblasts is more diversified, with sugar beet, potatoes, and legumes besides grains. Thus, the policy of favouring the southern oblasts makes economic sense.

In a situation of generalized agricultural crisis like the one in 1932 (see discussion below), decisions had to be made about the priorities of resources, and Moscow's more favourable treatment of the grain-producing oblasts was expected to result in lower relative direct losses in these oblasts than in the rest of the Ukrainian SSR. Comparing Maps 1 and 3, we see that this is only partially true. Although in general the southern oblasts have lower relative losses than the northern oblasts, there are exceptions. In the steppe zone, Donetsk oblast has significantly lower relative direct losses than Dnipropetrovsk, Odesa, and Moldavia. In the forest-steppe zone, Vinnytsia oblast has much lower losses than Kyiv and Kharkiv oblasts, and the level of its losses is similar to that of most oblasts in the steppe zone. Chernihiv oblast also does not conform to the economic hypothesis, as its relative direct loss is as low as in Donetsk oblast.

1932: Early manifestations of the Famine

To better understand the reasons for the differences in relative direct losses among the different oblasts, it is necessary to examine separately what happened in 1932 and 1933, as the dynamics of the Holodomor changed drastically between 1932 and 1933.

Regional differences in rural direct losses were already present in 1932. Kyiv oblast had the highest number of excess deaths per 1,000 population, with 14.2, followed by Moldavia with 9.6 and Kharkiv with 8.6; losses in the other oblasts vary between 5.5 and 6.6 excess deaths per 1,000 population (Table 3). Some of the reasons for this situation are described in detail in a letter to Stalin from the head of the Council of People's Commissars of the Ukrainian SSR, Vlas Chubar, in June 1932, which is quoted by Plokhy (2016: 382):

> The failure of legume and spring crops in those raions, above all, was not taken into account, and the insufficiency of those crops was made up with foodstuffs to fulfill the grain requisition plans. Given the overall impossibility of fulfilling the grain requisition plan, the basic reason for which was the lesser harvest in Ukraine as a whole and the colossal losses incurred during the harvest (a result of weak economic organization of the collective farms and their utterly inadequate management from the raions and from the center), a system was put in place of confiscating all grain produced by individual farmers, including seed stocks, and almost complete confiscation of all produce from the collective farms… In addition to grain procurements, the same methods were applied to potato and, especially, meat procurements.

The situation in Kharkiv oblast was no better. After his tour of Kharkiv oblast, Hryhorii Petrovsky, head of the Communist Party's Central Executive Committee for the UkrSSR, wrote to Stalin in June 1932 that 'famine has engulfed a good part of the countryside… It will take a month or a month and a half for new grain to appear… This means that famine will intensify' (Plokhy 2016: 383). In a list of raions most affected by the famine, compiled by Party officials in Kharkiv in June 1932, Kyiv and Vinnytsia oblasts had 10 and 11 raions, respectively, while the number of affected raions in the southern oblasts was much smaller. The critical situation in Kyiv and Kharkiv in 1932 is confirmed by the high relative direct losses in these two oblasts;[5] the lower level in Vinnytsia oblast (an international border oblast) was likely due to the lower mortality in the border areas (border oblast).

A key factor at the beginning of the famine was the grain procurement plan for 1932 (Table 4). It documents the expectations of the Soviet government regarding Ukraine's contribution to Stalin's overall procurement plan, and provides a fairly good understanding of the conditions in the different oblasts. The total 1932 quota for the Ukrainian SSR was 5,831,000 tons of grain. This target seems reasonable, as it constituted 90 per cent of the amount collected from the 1931 crop. The relative allocation of this quota among the different

5. Although we were not able to find official documents about the situation in Moldavia, the high losses estimated are consistent with the fact that repeated allotments of food were provided for this autonomous republic starting as early as March 1932 (RSASH 17/167/35, List 4: #44, #72).

Table 4. Grain procurement quotas for Ukrainian SSR in 1932, by region

Region	1932 grain procurement quotas, % of 1931 quota	1932 grain procurement quotas		
		tons	% of other crops (non-grain and forage)	% of quota for independent farmers
Ukraine total	90.0	5,831,000	9.7	17.1
Vinnytsia	88.0	639,000	22.4	40.2
Kyiv	65.5	511,000	26.0	41.1
Kharkiv	74.5	1,212,000	11.9	23.8
Dnipropetrovsk	90.0	1,441,000	5.3	6.8
Odesa	140.0	1,376,000	1.7	6.7
Donetsk	95.0	583,000	7.4	5.1
Moldavian ASSR	46.0	69,000	2.9	30.4

Note: Chernihiv oblast was created later in 1932.
Source: Pyrih 2007: 242.

oblasts favoured the forest-steppe oblasts of Kyiv, Kharkiv, and Vinnytsia, at the expense of the steppe oblasts. Compared to what was collected in 1931, the amounts allocated to the steppe oblasts are higher than to the forest-steppe oblasts. The plan also takes into account the mixed-crop composition of the forest-steppe zone, with much higher allocations to these crops for the oblasts in this zone than for the oblasts in the steppe zone. It also acknowledges the fact that the proportion of independent farmers was much higher in the forest-steppe than in the steppe zone, and their grain quotas are much higher in the former than the latter.

The official procurement plan corroborates, at a more general level, Chubar's impressions about the situation in Kyiv and Vinnytsia oblasts. It provides credence to Chubar's statement that the unexpected failure of the non-grain crops and the heavy reliance of the official grain procurement plan on these crops had dire consequences. The crop failure led to widespread famine in the forest-steppe zone, forcing the government to confiscate most of the grain at kolkhozes and impose even harsher confiscation measures on individual farmers.

The extreme famine conditions in many areas of the forest-steppe zone, and to a lesser degree in the steppe zone, forced the Ukrainian SSR government in Kharkiv[6] to petition repeatedly for some relief from the grain procurement quotas. After strong resistance, Stalin had to accept reality, and grain procurement quotas were reduced three times during 1932: two significant reductions in August and October, and a more modest reduction at the end of the year.

Table 5. Successive reductions of 1932 grain quotas for Ukrainian SSR, by region

Region	Original quota		% reduction			January 1933 quota		
	million poods	% distr.	August 1932	October 1932	January 1933	million poods	% overall reduction	% distr.
Ukraine total	356	100	11	25	29	210	41	100
Vinnytsia	39	11	23	12	0	26.5	32	13
Kyiv	31	9	35	30	0	14	54	7
Kharkiv	74	21	11	41	3.4	35.5	52	17
Dnipropetrovsk	88	25	4.5	20	12	55.5	37	26
Odesa	84	24	2.3	17	12	56	33	27
Donetsk	36	10	14	33	2	19	47	9
Moldavian ASSR	4	1	12	22	0	3	29	1

Note: Chernihiv oblast was created later in 1932.
Source: Pyrih 2007: 242, 298, 303–04, 355–56, 601–02.

The first round of reductions favoured heavily the forest-steppe zone at the expense of the steppe zone. Kyiv oblast received the largest reduction, with 35 per cent, followed by Vinnytsia oblast with 23 per cent and

6. Kharkiv was the capital of the Ukrainian SSR until 1934.

Kharkiv with 11 per cent, while reductions for Odesa and Dnipropetrovsk oblasts were in the 2.3–4.5 per cent range. Donetsk oblast received a reduction of 14 per cent, significantly higher compared to the other two steppe oblasts; this was repeated also during the next round of reductions. (The special status of Donetsk oblast will be further discussed below.) During the second round of reductions, Kharkiv and Kyiv again received large reductions, which prompted the steppe oblasts to demand significant reductions as well.

Overall, the grain procurement quota for the Ukrainian SSR was reduced by 41 per cent. Kyiv and Kharkiv oblasts had their original quotas reduced by more than half, and Vinnytsia oblast by one-third. The reduction for Odesa and Dnipropetrovsk oblasts was about one-third, and for Donetsk oblast it was close to half.

Table 6. Percent fullfillment of grain quotas by region in Ukrainian SSR, as of 1 Jan. 1933

Region	Kolkhozes	Sovkhozes	Independent farmers	Total	% collectivized as of 1 Oct. 1932
Ukraine total	78	86	72	77	69
Chernihiv	92	96	68	78	47
Vinnytsia	100	95	100	100	59
Kyiv	100	101	90	100	67
Kharkiv	85.5	92	44	77	72
Dnipropetrovsk	70	82	54	69.5	85
Odesa	73	70	57	72	84
Donetsk	76	77	85	76	84
Moldavian ASSR	89	40.5	108	93	68

Sources: Pyrih 2007: 571–72; ANER 1935: 205.

The grain quota fulfillment results and collectivization levels shown in Table 6 are surprising, if not puzzling. By October 1932 the steppe oblasts had reached very high levels of collectivization, while levels of collectivization in the forest-steppe and Chernihiv oblasts were significantly lower. In contrast, by the end of 1932 Kyiv and Vinnytsia had fulfilled 100 per cent of the grain procurement quotas, and Kharkiv close to 80 per cent, while the average for the forest oblasts was around 75 per cent. The collectivization levels are consistent with the official objective of faster collectivization of the grain-producing steppe region. The grain quota fulfillment data merit a more detailed analysis.

Fulfillment data is available for three groups: *kolkhoz, sovkhoz,*[7] and *independent farmers*. For the kolhozes and sovkhozes, per cent fulfillment is similar for all oblasts within each zone; per cent fulfillment is higher among the forest-steppe zone oblasts than among the steppe zone oblasts. The differences between the forest-steppe and steppe oblasts are mainly due to the performance of the independent farmers. Although independent farmers fulfilled over half of their quotas in Dnipropetrovsk and Odesa, and 85 per cent in Donetsk, this had little impact on the overall quota, due to the small proportion of independent farmers in these oblasts. The low performance of Kharkiv oblast, on the other hand, is due exclusively to the very low output fulfillment per cent from the independent farmers.

Resistance and repressions in 1932

Why is it that in spite of their relatively lower level of collectivization, the forest-steppe oblasts of Soviet Ukraine, except the independent farmers in Kharkiv oblast, show such extraordinary levels of compliance with the grain requisition plan? One possible answer is that these oblasts had been granted substantial reductions in their grain quotas (Table 5). Another possibility is the 'ruthless efficiency of the local Party machine in requisitioning grain from the peasantry' in Kyiv and Kharkiv oblasts, as a reaction to active and passive resistance (Plokhy 2016: 389).

7. The Russian terms *kolkhoz* (collective farm) and *sovkhoz* (state farm or plantation) have entered English usage and are therefore used here in roman type and pluralized accordingly. The equivalent Ukrainian terms are *kolhosp* and *radhosp*.

Table 7. Selected indicators of resistance and repression in Soviet Ukraine during the Holodomor, by region

Region	number of petitions to leave kolkhozes			registered 'terror' acts	fines in kind		# brigades requisitioning grain from indep. farmers
	individuals	farms	raions		number	% fines of indep. farmers	
	(1)	(2)	(3)	(4)*	(5)*	(6)	(7)
Ukraine total	14,095	475	111	73	1,791	n/a	n/a
Vinnytsia	5,800	219	42	75	150	90	51
Kyiv	3,320	75	21	79	70	99	65
Kharkiv	3,892	137	36	81	658	123**	84
Dnipropetrovsk	269	17	5	49	263	90	19
Odesa	191	7	4	94	344	97	24
Donetsk	–	n.d.	–	38	14	59	26
Moldavian ASSR	623	20	3	126	291	7	0

Notes: Chernihiv oblast is not listed as it was created in 1932 and some indicators are missing;
* indicators standardized by size of oblast's rural population; (1)…(3) June 1932;
(4) 1 Jan. 1932–31 Jan. 1933; (5)–(6) 5 Dec. 1932;
(7) 5 Dec. 1932; ** error in original data
Source: Pyrih 2007: 250, 445, 456, 631.

The following factors of resistance and repression are quantified in Table 7: exodus from the kolkhozes, acts of 'terror,' total fines, including in kind and percentage of independent farmers fined, and number of Communist Party grain-search 'brigades.' The flight from kolkhozes was quite extensive in the forest-steppe oblasts, but negligible in the steppe oblasts. While the relative number (standardized by the rural population of each oblast) of registered acts of 'terror' was very high in Odesa oblast, on average this indicator was higher in the forest-steppe than in the steppe oblasts.

The picture regarding number of fines in kind, also standardized by the rural population in each oblast, is less clear-cut. This indicator was extremely high in Kharkiv oblast, quite low in Vinnytsia and Kyiv, and very low in Donetsk oblast. In all oblasts except Donetsk, the great majority of fines in kind were applied to independent farmers.

On 11 November 1932, the Central Committee of the Communist Party of the Ukrainian SSR ordered the creation by December 1 of at least 1,000 brigades to search for hidden grain among the independent farmers. The proposed number of brigades was much higher for the forest-steppe oblasts than for the steppe oblasts: 200, 300, and 350 for Vinnytsia, Kharkiv, and Kyiv oblasts, respectively, and 50 each for the three steppe oblasts; these proportions are maintained when the numbers are standardized by the rural population of each oblast. The higher number of brigades for the forest-steppe oblasts was due, in part, to the fact that these oblasts had more independent farmers. The very high percentage of grain procurement quotas for independent farmers in Vinnytsia and Kyiv oblasts (Table 4), and the fact that independent farmers in these oblasts had the highest percent fulfillment of these quotas (Table 6), tend to support the 'ruthless efficiency' argument.

Further evidence about the more aggressive grain requisition practices in Kyiv and Kharkiv oblasts during 1932 is provided in a report on the fulfillment of seed grain quotas for the 1933 harvest. As of 10 December 1932, only 20.5 per cent and 16.5 per cent of the quotas were filled in Kyiv and Kharkiv oblasts, respectively, while 40 per cent of the quota was filled in Dnipropetrovsk, 28 per cent in Donetsk, and 22 per cent in Odesa oblasts. These numbers support the hypothesis that most of the grain was already taken away in Kyiv and Kharkiv oblasts due to more aggressive requisition, while there was still a fair amount of grain left in the steppe oblasts. More updated data for Kharkiv oblast tends to confirm this hypothesis. Namely, it was reported that by 15 February 1933, only 35.6 per cent of the seed grain quota was fulfilled, and that the campaign was facing strong resistance (Pyrih 2007: 697).

The data tend to support the hypothesis that there was higher resistance to collectivization and grain procurements in the forest-steppe oblasts, especially in Kyiv and Kharkiv, than in the steppe oblasts, and that these

oblasts were consequently subject to harsher repressions. The evidence may not be conclusive, as there is no certainty that the documents found so far are representative of the total picture in each oblast. Nevertheless, they show a correlation that is quite suggestive.

1933: Famine as terror

The number of relative rural losses presented in Map 1 is for the whole 1932–34 period. As 90 per cent of all losses occurred in 1933, the level of these losses is determined to a great extent by what happened in that year. In rural areas, two processes were happening in 1933: (1) extraordinary increase in monthly registered deaths during the first 6–7 months (Wolowyna 2013); and (2) implementation of a food aid program by Moscow as a reaction to this critical situation.

Between January and June 1933, the number of *registered rural deaths* increased by 11 times in Kyiv and Kharkiv oblasts, and eightfold in Vinnytsia oblast; in Odesa, Dnipropetrovsk, and Donetsk oblasts the increases ranged from fourfold to sevenfold, and in Moldavia rural registered deaths increased by half. These extraordinary increases were the result of several measures implemented by the Soviet government in late 1932 and early 1933.

First, two of these measures prevented peasants from travelling in search of food: (1) the introduction in December 1932 of domestic identity documents ("passports") only for city residents, limiting the peasants' ability to travel to cities in search of food; and (2) the closing of borders between Ukraine (as well as the Northern Caucasus) and Russia in January 1933, stopping the flow of Ukrainian peasants to Russia in search of food. Thousands of Ukrainian peasants were arrested in Russia and returned to their villages (CC ACP 2001).

Second, Stalin's directive dated 1 January 1933 reiterated the penalties outlined in the decree dated 7 August 1932, for 'stealing' stalks from the fields or hiding grain from the State, and harsh penalties in kind (meat and potatoes) introduced on 18 and 20 November 1932 for independent farmers and kolkhozes that did not fulfill their grain quotas.

Third, numerous brigades of Communist Party activists descended towards the end of 1932 and beginning of 1933 on villages to confiscate hidden grain, although most of it had been already seized, especially in Kyiv and Kharkiv oblasts. According to thousands of testimonies, even if no grain was found, in many instances every last scrap of food was confiscated (see also Chubar's letter to Stalin above).

Fourth, a system of blacklists was instituted in November 1932 against kolkhozes, entire villages, and in some cases raions that failed to fulfill their grain quotas, and was gradually expanded to the whole country. 'For a village to be blacklisted meant that: (1) all stores would be closed and supplies removed from the village; (2) all trade was prohibited, including trade in food or grain; (3) all loans and advances were called in, including grain advances; (4) the local Party and collective farm organizations were purged, and usually subject to arrest; (5) food and livestock would be confiscated as a 'penalty'; and (6) the territory would be sealed off by OGPU (secret police) detachments' (Andriewsky 2015). In other words, a death sentence was imposed on the population of the given kolkhoz, village, or raion.

Once Moscow realized the catastrophic nature of the famine, a program of food aid was implemented during the first half of 1933. The program entailed loans that the oblasts were required to pay back from the next harvest with 10 per cent interest, and had other strong restrictions. Boriak (2012) documents in detail the characteristics of this program: (1) the food was to be given mainly to members of kolkhozes who were willing and able to work, and to independent farmers willing to join the kolkhozes and work; (2) instructions for the administration of the program show clearly that its main objective was not to prevent starvation but to provide badly needed aid in order to save the next sowing season; (3) a good part of the food provided came from internal reserves (in Ukraine), that had been requisitioned from Ukrainian farmers in 1932 and were now being given back to them as 'assistance', with selective distribution.

A total of 176,000 tons of food, mainly grain, was distributed to the eight regions of Ukraine between February and July 1933 (169,800 tons allocated to specific regions, plus 6,200 unallocated tons for selective distribution):

	Dnipropetrovsk	Odesa	Kharkiv	Kyiv	Vinnytsia	Donetsk	Chernihiv	Moldavia
tons food aid	56,200	49,400	29,900	19,900	9,600	3,300	1,200	300
kg per person	20.5	22.3	6.4	3.9	2.3	1.6	0.5	0.6

The data illustrate the importance of using relative indicators when making comparisons. In absolute numbers, the bulk of the food aid went to Dnipropetrovsk and Odesa oblasts, with sizeable contributions also to Kharkiv and Kyiv oblasts. However, standardizing by the size of the respective rural populations introduces significant changes in the distribution. For example, the ranking between Dnipropetrovsk and Odesa oblasts is reversed, and more importantly, the difference in food aid amounts between Dnipropetrovsk and Odesa and the forest-steppe oblasts becomes much more pronounced. Thus, the actual amount to Dnipropetrovsk oblast is three times that given to Kharkiv oblast, instead of just under double as per the unadjusted figures.

To illustrate the devastating effect of Stalin's measures in late 1932 and early 1933 on the level and distribution of monthly losses in these oblasts in 1933, we selected two oblasts from the forest-steppe region, Kyiv and Kharkiv, and two from the steppe region, Odesa and Dnipropetrovsk. We show the relationship between the volume and timing of this food aid, and the number and monthly pattern of excess deaths in each of these oblasts.

The oblasts in the forest-steppe and those in the steppe region have very different patterns of monthly excess deaths in 1933 (Figure 1). Kyiv and Kharkiv experienced a sharp increase in monthly excess deaths between January and June, and then a sharp decrease. The rate of increase for Odesa and Dnipropetrovsk was somewhat smaller than for Kyiv and Kharkiv oblasts, with the peak in June being much lower and the decrease during the second half of 1933 being much less pronounced. The ratio of direct losses between the peak month of June and January of 1933 is even higher than the ratio of registered deaths. During the first half of 1933, the number of excess deaths increased by 14–15 times in Kharkiv and Kyiv, and by 7–8 times in Dnipropetrovsk and Odesa oblasts.

Figure 2 shows the timing and volume of food distributed to the different oblasts, in tons per 1,000 rural population. The graph shows very clearly that Odesa and Dnipropetrovsk oblasts received much more food aid than Kyiv and Kharkiv oblasts, and that this assistance started to arrive much earlier.

Comparing the two figures, we see a strong relationship between the food aid dynamics and the patterns of monthly excess deaths. The volume and timing of food distributed are clearly reflected in the two distinct patterns of monthly direct losses. The large amounts of food sent to Dnipropetrovsk and Odesa oblasts in February and March had two effects: it slowed down the monthly increase of direct losses and resulted in much lower peaks in June. The absence of practically any food aid to Kyiv oblast before March, or to Kharkiv oblast before April, resulted in faster rates of increase and much higher peaks in direct losses for these two oblasts.

One can also detect specific effects of the food assistance on the distribution of excess deaths in certain oblasts. For example, the rate of increase in monthly excess deaths slowed down between March and April in Kyiv oblast compared to Kharkiv oblast, and in Dnipropetrovsk oblast compared to Odesa oblast. This is likely related to the large amount of food aid sent to Kyiv oblast in mid-March, and larger amounts of food aid provided to Dnipropetrovsk oblast than to Odesa oblast in February and March.

It is clear that the food aid program saved many lives in Odesa and Dnipropetrovsk oblasts. However, the main goal of the program was to save the 1933 harvest, and thus the assistance was targeted at specific oblasts and groups. As a result, many more peasants were condemned to death by starvation in Kyiv and Kharkiv oblasts than in the strategically more important oblasts of Odesa and Dnipropetrovsk. Although the number of excess deaths was significantly lower in the steppe than in the forest-steppe oblasts, the rate of monthly increase and maximum levels of death in Odesa and Dnipropetrovsk were still extremely high.

Historical legacy of peasant uprisings

For the sake of completeness, we shall also examine a hypothesis that suggests links between high regional direct losses and past events in those regions. It posits that the degree of resistance and resulting persecutions in certain regions, described in the section 'Resistance and repressions in 1932,' is related to different types of peasant revolts having occurred in those places during the preceding period (1918–31). Thus, regions with strong resistance to collectivization and grain procurement in 1932 had a history of rebellions in the past, of which the Soviet regime was keenly aware—especially in relation to the great social and national uprising of the spring

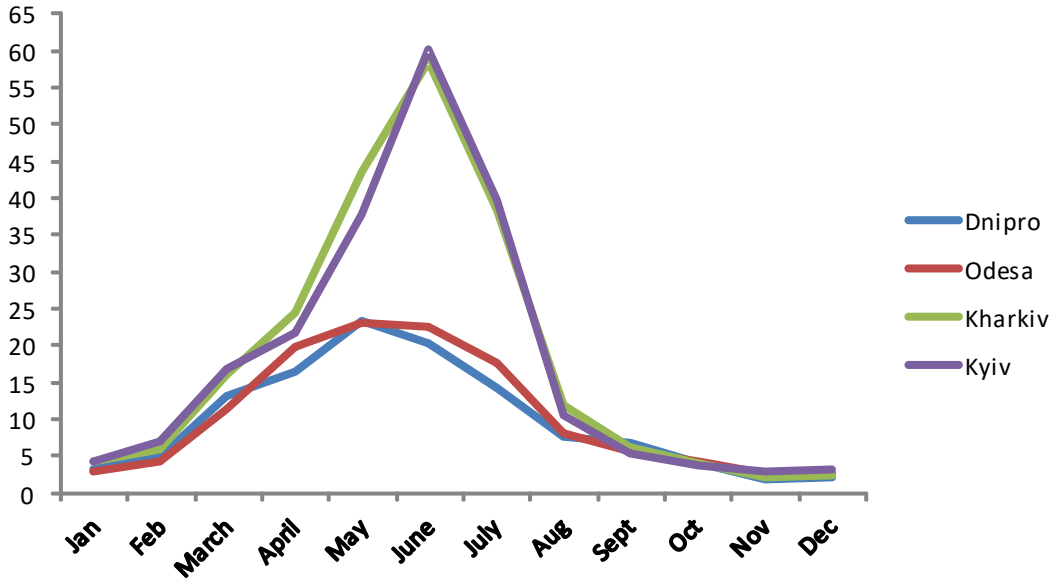


Figure 1. Monthly direct losses (per 1,000 rural population) for four oblasts of the Ukrainian SSR, 1933.
Source: Authors' calculations.

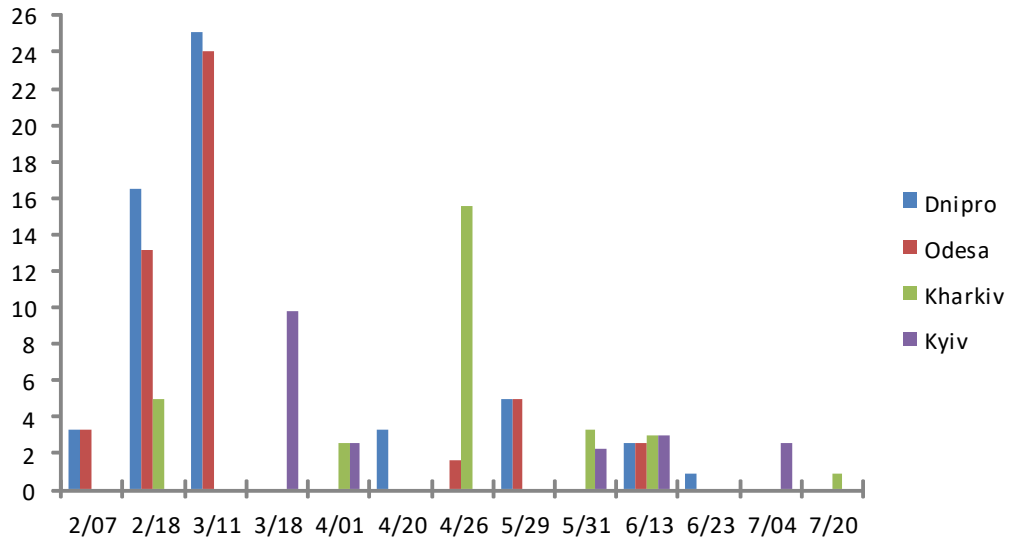


Figure 2. Food aid to four oblasts of the Ukrainian SSR (tons/1,000 rural population), 1933.
Source: Davies and Wheatcroft 2009, pp. 481–4.

and summer of 1919, which had forced them out of Ukraine, and in particular out of its two capitals (Kyiv and Kharkiv).[8] This historical memory resulted, first, in stronger repressions and thus higher excess deaths in 1932, and then in a decision, taken in late 1932 and applied during the following months, to use hunger as a tool to eradicate the possibility of a new general uprising, and to deprive the Ukrainian national movement of its social base, which Stalin had identified as being the villages (Graziosi 2015).

If this hypothesis is correct, the effects of the food aid program on 1933 direct losses, as described in the section '1933: Famine as terror,' need to be compared to the effects of the punitive policy in the different regions. Testing this hypothesis requires two elements: a map depicting the historical revolts at the raion level, and

8. This may have been a factor in the decision, taken in 1929, not to discontinue the extant state indigenization program (*korenizatsiia*, or, in the case of the Ukrainian SSR, *ukraïnizatsiia*) during collectivization—precisely because of the awareness of the need to prevent a repetition of social and national elements combining to engender peasant revolts, as had occurred in Ukraine in 1919.

1933 estimates of rural direct losses at the raion level. We have calculated the estimates and hope that a map will be found to allow us to test this hypothesis.

We do have, however, some elements that permit testing the first part of this hypothesis, i.e., that uprisings during the 1918–31 period are linked to areas where stronger repressions were applied in 1932. Viola (1996) and Graziosi (1996) documented widespread peasant rebellions in different Soviet republics, starting in 1918, but their data is at the republic and large-region levels. On the other hand, recently discovered documents in Ukraine's archives provide more information about these movements in specific regions of Ukraine (Krutsyk 2011).

First, however, we have to deal with a technical problem. The data on the historical peasant rebellions are for nine gubernias, while our estimates of direct losses are for seven oblasts. Due to problems with vital statistics for this period, it is impossible to make estimates of excess deaths for the nine gubernias. However, we can approximate the nine gubernias with the 17-oblast structure in 1939 (15 oblasts plus Cherkasy and Kherson, which were created in 1944 and 1954, respectively). We estimated direct losses for these 17 oblasts, and then calculated direct losses for the nine gubernias based on the losses for the 17 oblasts. Table 8 shows the equivalence between the nine gubernias and 17 oblasts, and the 1933 rural relative direct losses for the nine gubernias, as well as the 1933 losses.

The following indicators are presented in absolute and relative numbers (per one million rural population): number of peasant uprisings, number of clandestine organizations, and number of rebel groups; all indicators are for the period 1918–32, and the rural population is given as of 1 January 1927. We see that neither absolute numbers nor indicators standardized by the rural populations of respective gubernias show a relationship between the intensity of rebellion indicators and relative numbers of direct losses in 1932. The highest 1932 rural relative losses are in Kyiv gubernia, while the highest absolute and relative values for the three indicators are found mostly in other gubernias. The same applies to 1933 losses, with Poltava gubernia having the highest losses.

There are several problems with this test: (1) the data are for different periods, and it is difficult to establish common standards with the gubernias and their raion structures, as the administrative structures changed repeatedly during this period; (2) as witnessed by differing numbers on the different types of peasant resistance presented by Viola (1996) and Graziosi (1996) under different labels, there seems to be a lack of established definitions for concepts describing these events; and (3) the reliability of official statistics has not been evaluated. A key problem is that only macro-level data (for gubernias) is available, while a more valid test would require data at the raion level, especially of the uprisings in 1919. The fact that currently available data does not support this hypothesis does not mean that the hypothesis is incorrect; further research is needed before a more definite judgment can be made.

Table 8. Indicators of peasant resistance movements in Soviet Ukraine, by gubernia, 1917–32

9 gubernias	17 oblasts (equivalent)	Absolute numbers			Per one million peasants*			% of 1932 rural direct losses	% of 1933 rural direct losses
		Number of peasant uprisings, 1918–32	Number of clandestine organizations	Number of rebel groups	Number of peasant uprisings, 1918–32	Number of clandestine organizations	Number of rebel groups		
Volyn + Podillia	Zhytomyr + Vinnytsia + Khmelnytskyi	32	120	300	8	27	68	9	12
Kyiv	Kyiv + Cherkasy	40	107	296	9	25	69	11.5	22
Poltava	Poltava	29	103	165	9	32.5	52	6	24
Katerynoslav	Dnipropetrovsk + Zaporizhia + Kherson	21	57	104	8	21.5	39	5	10
Odesa	Odesa + Mykolaiv + Kirovohrad	57	72	188	22	28	73	8	11
Chernihiv	Chernihiv	32	64	137	16	33	70	6	8
Kharkiv	Kharkiv + Sumy	34	65	133	14	26	54	6.5	12
Donetsk	Donetsk + Luhansk	23	54	112	11.5	27	56	10	9

* As of 1 January 1927.

Sources: Krutsyk 2011 and authors' calculations.

Summary and conclusions

Our analysis has shown significant variation in Holodomor-caused direct losses at the oblast level in Soviet Ukraine. Several hypotheses about these differences have been evaluated, but no single hypothesis provides a comprehensive explanation. As pointed out by Plokhy (2016), the solution probably lies in a composite of several hypotheses. The direct loss levels in three oblasts—Chernihiv, Vinnytsia, and Donetsk—can be explained as specific cases.

Firstly, Chernihiv oblast is part of the Polissia region, and the only oblast that satisfies the ecological hypothesis. Besides the ecological advantage of having food available in the forests and wetlands, Chernihiv did not fall under the close scrutiny of the Soviet government, as it had the smallest land area dedicated to grain production. Thus, Chernihiv oblast was probably least affected by the searches for hidden grain in late 1932 and early 1933, which likely explains its low level of rural direct losses in 1933, and thus for the whole 1932–34 period (Table 3).

Next, the lower level of direct losses in Vinnytsia oblast, compared to Kyiv and Kharkiv oblasts, can be explained to some degree by the border hypothesis, as the lower levels of direct losses in border raions bring down the oblast average. Thirdly, the low level of direct losses in Donetsk oblast is due to several unique characteristics. This oblast had the lowest percentage rural population, and moreover it received special assistance from Moscow due to the strategic importance of its industrial infrastructure; workers in these enterprises belonged to a privileged group that received adequate food rations, and this probably allowed them to help their families in the countryside. If we exclude Moldavia, three indicators support Donetsk oblast's privileged position: (1) the overall reduction of its grain quota was the largest among the steppe oblasts (Table 5); (2) it had the lowest number of registered 'terror' acts; and (3) it had the lowest number of in-kind fines, including among independent farmers (Table 7).[9]

Table 9. Comparison of oblasts with high (Kyiv and Kharkiv) and low (Odesa and Dnipropetrovsk) rural relative excess deaths during Holodomor

#	Indicator	High losses	Low losses
	1932–34 rural excess deaths/100 population	23	13
	A – Background indicators		
1	1932 grain quotas: % other crops	16%	4%
2	1932 grain quotas: % independent farmers	29%	7%
3	% overall reduction of 1932 grain quotas	53%	35%
4	% fullfilment 1932 grain quotas, 1/1933	81%	70%
	B – Resistance and repressions indicators		
5	% grains collected of 1933 sawing quota	18%	32.5%
6	# of petitions to leave kolkhozes, 1932:		
	- individuals	7,212	460
	- farms	212	24
	- raions	57	9
7	# of fines in kind, 1932*	350	299
8	% independent farmers among all fined, 1932	99**	93
9	% of registered 'terror' acts, 1932*	80	69
	C – Situation in 1933		
10	1933 excess deaths: June/January	14	8
11	1933 food assistance (kg per rural inhabitant)	5.1	21.3

* per 1,000,000 rural population

** for Kyiv only; there is an error in the original data for Kharkiv

Source: Authors' calculations.

9. Moldavia does not seem to fit a pattern and we excluded it from our analysis. Given its small size and the fact that it was part of Ukraine only during a limited period (1924 to 1940), this exclusion has little effect on understanding the regional dynamics of Holodomor losses.

We are left with having to explain the levels of excess deaths differences between Kyiv and Kharkiv oblasts and Odesa and Dnipropetrovsk oblasts.[10] The evidence summarized in Table 9 shows that the much higher levels of direct losses in Kyiv and Kharkiv oblasts than in Odesa and Dnipropetrovsk oblasts can be explained by a combination of the economic hypothesis, significantly higher levels of resistance and repressions in the first two oblasts and selective implementation of the food assistance program in 1933.

A second result is the elaboration and quantification of the already known fact that the dynamic of the Holodomor was very different in 1932 than in 1933. The onset of the famine is characterized by regional differences in collectivization, grain quota fulfillment, opposition to collectivization and grain procurement, and levels of repressions against this opposition. The sudden explosion of deaths, and thus direct losses, during the first half of 1933 can only be explained as the result of the actions implemented towards the end of 1932 and beginning of 1933, as listed above in the section '1933: Famine as Terror.' Although no document has been found with a general directive to confiscate not only all grain but also other foodstuffs during the searches for 'hidden' or 'stolen' grain, the demographic evidence does not leave room for any other explanation. The regional differences in direct losses found in 1933 are to a great extent a function of selective implementation of the food aid program, and it remains to be seen whether a set of politically-motivated actions rooted in the 1919 Soviet experience in Ukraine, and the active resistance in the 1920s and early 1930s, constituted additional factors.

Our analysis documents the complex dynamics of the Holodomor and shows that there are still quite a few unanswered questions. Examples of areas that require systematic research are: (1) the possible link between peasant uprisings during the 1918–31 period and the level of excess losses in 1932, and the more specific link between the 1919 uprisings and direct losses in 1933; (2) more systematic research on the searches carried out for hidden food in late 1932 and early 1933; and (3) the role of the nationality factor in chances of survival.

Finally, we observe that research on the 1932–34 famine in Soviet Ukraine has been pursued independently along two disciplines: demography and history. Our approach of addressing both demographic analysis and historical research illustrates the importance of such a combined strategy. Often historical evidence is needed to explain demographic results, and demographic techniques can be used to test hypotheses suggested by historical analysis, or hypotheses derived from historical analysis can suggest specific demographic analyses. The complementarity of the two disciplines provides a more fruitful strategy for researching the Holodomor.

Acknowledgements

This study was made possible by a Fulbright grant to Wolowyna and funding from the Harvard University Research Institute and the Ukrainian Studies Fund. Map 3 is used with permission of the Harvard Ukrainian Research Institute. We would like to express our appreciation to Prof. Andrea Graziosi for his extremely helpful comments, and to an anonymous reviewer for his comments and suggestions.

References

Andreev, E.M., L.E. Darskii, and T.L. Kharkova. 1990. *Istoriia naseleniia SSSR, 1920–1959 gg.* [History of the Population of the USSR, 1920–59]. Moscow: Goskomstat SSSR. Special monograph issue of *Istoriia statistiki* [History of Statistics] 3/5(1). (in Russian)

10. Wheatcroft suggests that the high levels of mortality in Kyiv oblast are due to the fact that, unlike the industrial centers in the east and south, the cities of Kyiv oblast received little or no food from central depositories, and that Kyiv oblast authorities were forced to take away food from the oblast's rural areas in order to feed the cities (Wheatcroft and Garnaut 2013). There are several problems with Wheatcroft's hypothesis. First, as shown in Map 2, 1933 relative rural excess deaths at the raion level are not uniformly distributed in Kyiv oblast; they vary from the highest to the lowest levels for the whole country. Thus, if Wheatcroft's hypothesis is correct, the requisition of food in rural areas was very selective and his hypothesis would require further elaboration. Second, it would make sense to collect the extra grain from areas with the highest potential grain production. Comparing Map 2 with Map 3, areas with the highest levels of excess deaths are quite different from areas with the highest percent of potential grain production. Third, losses in Kharkiv oblast were as high as in Kyiv oblast, with the distribution of losses in rural areas by raion being similar to that in Kyiv oblast; Wheatcroft does not comment on this.

Andriewsky, O. 2015. Towards a decentred history: The study of the Holodomor and Ukrainian historiography, in *Contextualizing the Holodomor: The Impact of Thirty Years of Ukrainian Famine Studies*, edited by A. Makuch and F.E. Sysyn. Toronto: University of Toronto Press, p. 4–48.

ANER. 1933a. (Administration of National Economic Records of the Ukrainian SSR [Upravlinnia narodnohospodarskokho obliku URSR]). *Naslidky reproduktsii liudnosti USRR za roky 1926–1931* [Results of Human Reproduction in UkrSSR in the Years 1926–31]. Kharkiv: Narodne Hospodarstvo ta Oblik. (in Ukrainian)

———. 1933b. *Pidsumky obliku mis'koï liudnosti URSR 1931 roku* [Summary of the Urban Population Count of the UkrSSR in 1931]. Kharkiv: Narodne Hospodarstvo ta Oblik. (in Ukrainian)

———. 1935. *Narodne hospodarstvo URSR: Statystychnyi dovidnyk* [National Economy of the UkrSSR: A Statistical Guide]. Kyiv: Narodne Hospodarstvo ta Oblik. (in Ukrainian)

Boriak, T. 2012. Prodovol'cha dopomoha Kremlia iak instrument Holodomoru v Ukraïni [Kremlin food aid as an instrument of the Holodomor in Ukraine], in *Zlochyny totalitarnykh rezhymiv v Ukraïni: Naukovyi ta osvitnii pohliad* [Crimes of Totalitarian Regimes in Ukraine: A Scholarly and Educational Perspective], edited by K.C. Berkhoff, A. Podol's'kyi, and M. Tiahlyi. Kyiv: Centre for Holocaust and Genocide Studies of the Royal Netherlands Academy of Arts and Sciences, p. 10–33. (in Ukrainian)

Bugai, N.F. 2013. *Poliaki Rossii: Poisk istiny (prinuditel'noe pereselenie, vozvrashchenie, sud'by)* [Poles in Russia: The Search for Truth (Forced Resesettlement, Return, Destinies)]. Moscow: RAS Institute of Russian History. (in Russian)

CC ACP (Central Committee of All-Union Communist Party (Bolshevik) [TsK VKP(b)]). 2001. Direktiva №65/III ot 22 ianvaria 1933 goda 'V sviazi s massovym vyezdom krest'ian za predely Ukrainy' [Directive No. 65/III dated 22 January 1933 'In conjunction with the mass exodus of peasants out of Ukraine'], in *Tragediia sovetskoi derevni: Kollektivizatsiia i raskulachivanie, dokumenty i materialy* [Tragedy of the Soviet Countryside: Collectivization and Dekulakization, Documents and Materials], vol. 3, edited by V. Danilov, R. Manning, and L. Viola. Moscow: ROSSPEN, p. 634. (in Russian)

CANER (Central Administration of National Economic Records of the USSR [Tsentralnoie upravleniie narodno-khoziaistvennogo ucheta SSSR]). 1936. *Raionnye tsentry i administrativno-territorial'noe delenie SSSR na 1 oktiabria 1936* [Raion Centres and Territorial-Administrative Structure of the USSR as of 1 October 1936]. Moscow: TsUNKhU. (in Russian)

CSA UkrSSR (Central Statistical Administration of the Ukrainian SSR [Tsentralne statystychne upravlinnia URSR]). 1927–32. *Pryrodnyi rukh naselennia Ukraïny: 1924–1929* [Vital Statistics of Ukraine's Population: 1924–29]. Kharkiv: TsSU URSR and Hospodarstvo Ukrainy. (in Ukrainian) (also publ. as *Statystyka Ukraïny* 106, 117, 154, 169, 193, and 213)

CSA USSR (Central Statistical Administration of the USSR [Tsentralne statystychne upravlinnia SRSR]). 1929. *Vsesoiuznyi perepys liudnosti 1926 roku: Ukraïns'ka RSR.* [1926 All-Union Population Census: Ukrainian SSR]. Vol. 11, 12, 13. Moscow: TsSU SRSR. (in Ukrainian)

———. 1970. *Raspredelenie naseleniia USRR po vozrastu i polu za perepysiu 1939 goda v administrativno-territorial'nom delenii 1970 goda* [Population Distribution of the UkrSSR by Sex and Age at the 1939 Census According to the 1970 Administrative-Territorial Structure]. Moscow: TsSU SSSR. (in Russian)

CSANO (Central State Archive of Non-Government Organizations in Ukraine [Tsentralnyi derzhavnyi arkhiv hromadskykh obiednan Ukrainy]). 1/2/6583–85, 6392.[11]

Davies, R.W., and S. Wheatcroft. 2009. *The Years of Hunger: Soviet Agriculture, 1931–1933*. London: Palgrave Macmillan.

Graziosi, A. 1996. *The Great Soviet Peasant War: Bolsheviks and Peasants, 1917–33*. Harvard Papers in Ukrainian Studies. Cambridge, MA: Harvard University Press.

11. Archive numbering indicates *fond* 'collection' / *opys* 'series' / *sprava* 'file' (Ukrainian).

———. 2015. The impact of Holodomor studies on the understanding of the USSR, in *Contextualizing the Holodomor: The Impact of Thirty Years of Ukrainian Famine Studies*, edited by A. Makukh and F.E. Sysyn. Toronto: University of Toronto Press, p. 49–75.

Golotik, S., and V. Minaev. 2004. *Naselenie i vlast': Ocherki demograficheskoi istorii SSSR 1930 godov* [Population and Power: Outline of the Demographic History of the USSR in the 1930s]. Moscow: Izd-vo Ippolitova. (in Russian)

Hirshfeld, A. 1930. *Migratsiini protsesy na Ukraïni (v svitli perepysu 1926 roku)* [Migration Processes in Ukraine: In Light of the 1926 Census]. Kharkiv: Hospodarsto Ukrainy. (in Ukrainian)

Iefimenko, H. 2013. Pereselennia ta deportatsii v postholodomorni roky (1933–1936): Porayonnyi zriz [Resettlements and Deportations the Post-Holodomor Years (1933–36): Cross-Section by Raion]. http://gis.huri.harvard.edu/images/pdf/Relocation-1933-1936.pdf. (in Ukrainian)

Iukhnovs'kyi, I., V. Verstiuk, V. Borysenko, H. Boriak, L. Bur'ian, O. Veselova, L. Hrynevych, H. Iefimenko, S. Kul'chyts'kyi, N. Lisunova, V. Marochko, O. Movchan, H. Papakin, I. Rykun, V. Telishchak, and N. Topishko. 2008. *Natsional'na knyha pam'iati zhertv Holodomoru 1932–1933 rokiv v Ukraïni* [National Book of Memory of the 1932–33 Holodomor Victims in Ukraine]. 18 vols. Kyiv: Olena Teliha Publishers. (in Ukrainian)

Kokurin, A., and N. Petrov. 2000. GULAG: Struktura i kadry [The GULAG: Structure and cadres]. *Svobodnaia mysl'* [Independent Thought] 3:105–23. (in Russian)

Korchak-Chepurkivs'kyi, Iu. nd. Personal papers. Kyiv: NASU Institute of Demography and Social Sciences.

———. 1928. Sproba analizu perepysnykh chysel malykh ditei na Ukraïni [Attempt at an analysis of the census numbers of small children in Ukraine]. *Visnyk statystyky Ukraïny* [Bulletin of Statistics of Ukraine] 2:153–58. (in Ukrainian)

———. 1962. Estimation of the 1926 Census population by oblasts and raions in Soviet Ukraine according to the administrative structure of the 1939 Census, in Personal papers of Iu. Korchak-Chepurkivs'kyi. (in Ukrainian)

Kozin, I.S. 1936. *Dovidnyk po orhanizovanomu naboru robochoï syly* [Guide to Organized Workforce Recruiting]. Kyiv: Radianske Budivnytstvo i Pravo. (in Ukrainian)

Krutsyk, R. 2011. *Narodna viina, 1917–1932: Putivnyk do ekspozytsii* [The People's War, 1917–32: Guidebook to the Exhibition]. Kyiv: Memorial Society. (in Ukrainian)

Kul'chyts'kyi, S.V. 2003. Heohrafiia Holodu [Geography of the Famine], in *Holod 1932–1933 rokiv v Ukraini: Prychyny ta naslidky* [Famine in 1932–33 in Ukraine: Causes and Consequences], edited by V.A. Smolii, V.M. Lytvyn et al. Kyiv: Naukova Dumka, p. 509–13. (in Ukrainian)

Leskova, T.A. 2005. Gosudarstvennoe regulirovanie migratsionnykh protsessov na iuge Dal'nego Vostoka Rossii v 20–30-e gg. XX v. [State Regulation of Migration Processes in the South of Russia's Far East in the 1920s–30s]. Candidate's diss., Blagoveshchensk. (in Russian)

Levchuk, N., T. Boriak, O. Volovyna [Wolowyna], O. Rudnyts'kyi, and A. Kovbasiuk. 2015. Vtraty mis'koho ta sil's'koho naselennia Ukraïny vnaslidok Holodomoru v 1932–1934 rr.: novi otsinky [Urban and rural population losses in Ukraine as a result of the Holodomor in 1932–34: New estimates] *Ukraïns'kyi istorychnyi zhurnal* [Ukrainian Historical Journal] 4:84–112. (in Ukrainian; English translation, with title as given above, to be published in a forthcoming book by CIUS Press)

Livshits F.D. 1990. Perepis' naseleniia 1937 goda [Population Census of 1937], in *Demograficheskie protsessy v SSSR* [Demographic Processes in the USSR], edited by A.G. Volkov. Moscow: Nauka, p. 174–207. (in Russian)

Maksudov, S. 2012. Geografiia poter' ot goloda 1932–1934 godov na Ukraine [Geography of losses from the 1932–34 famine in Ukraine], in *Demograficheskie aspekty sotsial'no-ekonomicheskogo razvitiia* [Demographic Aspects of Socioeconomic Development], edited by M.B. Denisenko. Moscow: MAKS-Press, p. 268–78. (in Russian)

Mozokhin, O.B. nd. Statistika repressivnoi deiatel'nosti organov bezopasnosti SSSR (1921–1953) [Statistics on Repressive Activity of Security Organs of the USSR (1921–53)]. Publ. online at http://istmat.info/node/255. (in Russian)

Nikol's'kyi, V.M. 2001. Natsional'ni aspekty politychnykh represii 1937 roku v Ukraïni [National aspects of political repressions in Ukraine in 1937]. *Ukraïns'kyi istorychnyi zhurnal* 2:74–89. (in Ukrainian)

Osokina, E. 2001. *Our Daily Bread: Socialist Distribution and the Art of Survival in Stalin's Russia, 1927–1941.* Armonk, NY, and London (UK): M.E. Sharpe.

Platunov, N.I. 1976. *Pereselencheskaia politika Sovetskogo gosudartsva i ee osushchestvlenie v SSSR (1917 – iiun' 1941)* [Resettlement Policy of the Soviet State and Its Implementation in the USSR: 1917 to June 1941]. Tomsk: Tomsk State University. (in Russian)

Plokhy [Plokhii], S. 2016. Mapping the Great Famine, in *The Future of the Past: New Perspectives on Ukrainian History*, edited by S. Plokhy. Cambridge, MA: Harvard Ukrainian Research Institute.

Poliakov, Iu.A. (ed.). 1991. *Vsesoiuznaia perepis' naseleniia 1937 goda: Kratkie itogi* [1937 All-Union Population Census: Brief Results]. Moscow: AS USSR Institute of USSR History. (in Russian)

———. 1992. *Vsesoiuznaia perepis' naseleniia 1939 goda: Osnovnye itogi* [1939 All-Union Population Census: Main Results]. Moscow: Nauka. (in Russian)

———. 2007. *Vsesoiuznaia perepis' naseleniia 1937 goda* [1937 All-Union Population Census]. Moscow: ROSSPEN. (in Russian)

Popov, V.P. 1995. Pasportnaia sistema v SSSR (1932–1976 gg.) [Passport system in the USSR (1932–76)]. *Sotsiologicheskie issledovaniia* [Sociological Research] 8:3–14 and 9:3–13. (in Russian)

Pyrih, R. (ed.). 2007. *Holodomor 1932–1933 rokiv v Ukraïni: Dokumenty i materialy* [Holodomor 1932–33 in Ukraine: Documents and Materials]. Kyiv: Kyiv-Mohyla Academy Publishing House. (in Ukrainian)

RSAE (Russian State Archive of the Economy [Russkii gosudarstvennyi arkhiv ekonomiki]). 1562/20/22, 27, 29, 30, 38, 41, 43, 46, 49, 59, 61, 62, 73, 75, 76, 80, 86, 88, 118, 121, 125, 143, 145, 153, 155.[12]

———. 1562/329/18, 20, 22, 33, 54, 56, 57, 114, 254, 261, 263, 264.[12]

RSASH (Russian State Archive of Sociopolitical History [Russkii gosudarstvennyi arkhiv sotsial'nopoliticheskoi istorii]). 17/167/35, List 4.[12]

Rudnik, V.V. 2006. Migratsionnye protsessy i ikh rol' v formirovanii naseleniia Sibiri [Migration Processes and Their Role in the Balance of Siberia's Population]. Candidate's diss., Omsk. (in Russian)

Rudnytskyi, O., N. Levchuk, O. Wolowyna, P. Shevchuk, and A. Kovbasiuk. 2015. Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932–1933. *Canadian Studies in Population* 42(1–2):53–80.

Rybakovskii, L.L. 1990. *Naseleniie Dal'nego Vostoka za 150 let* [Population of the Far East for 150 Years]. Moscow: Nauka. (in Russian)

SARF (State Archive of the Russian Federation [Gosudarstvennyi arkhiv Rossiiskoi Federatsii]). 9414/1/1943, 1944.[12]

———. 9479/1/2.[12]

Simchenko, Iu. B. 1990. *Perepis' 1939 goda: Dokumental'nye istochniki* [1939 Census: Documentary Sources]. 15 vols. Moscow: AS USSR Institute of Ethnology and Ethnic Anthropology. (in Russian)

Shlyakhter, A. nd. Smugglers and Commissars: Contraband Trades, Soviet Solutions, and the Making of the Soviet Border Strip, 1917–1939. PhD diss., University of Chicago. Forthcoming publ., incl. ch. 10, Borderness and famine: Why did fewer people starve to death in Soviet Ukraine's western border districts during the Holodomor, 1932–34?

Smolii, V.A., V.M. Lytvyn et al. (eds.). 2003. *Holod 1932–1933 rokiv v Ukraini: Prychyny ta naslidky* [Famine in 1932–33 in Ukraine: Causes and Consequences]. Kyiv: Naukova Dumka (in Ukrainian).

12. Archive numbering indicates *fond* 'collection' / *opis* 'series' / *delo* 'file' (Russian).

SS USSR (Supreme Soviet of the USSR [Verkhovnyi sovet SSSR]). 1938. *Administrativno-territorialnoe deleniie soiuznykh respublik na 1 oktiabria 1938* [Administrative-Territorial Structure of the Union Republics as of 1 October 1938]. Moscow: Vedomosti Verkhovnogo Soveta RSFSR. (in Russian)

———.1939. *Administrativno-territorialnoe deleniie soiuznykh respublik na 1 marta 1939* [Administrative-Territorial Structure of the Union Republics as of 1 March 1939]. Moscow: Vedomosti Verkhovnogo Soveta RSFSR. (in Russian)

Tol'ts, M. 1989. Repressirovannia perepis' [Repressed census]. *Rodina* [Homeland] 11:56–61. (in Russian)

———. 1995. The Soviet Censuses of 1937 and 1939: Some Problems of Data Evaluation. Paper presented at the International Conference on the Soviet Population in the 1920s and 1930s. Toronto: Univ. of Toronto, 27–29 January 1995. (in Russian; revised as of 30 May 2009)

Tron'ko, P.T., O.P. Reient, S.A. Kokin et al. (eds.). 1994–2001. *Reabilitovani istoriieiu* [Rehabilitated by History]. 27 vols. Kyiv: Instytut Istorii Ukrainy NANU. (in Ukrainian) (Available online at www.reabit.org.ua)

Tsaplin, V. 1989. Statistika zhertv stalinizma v 30-e gody [Statistics on the Victims of Stalinism in the 1930s]. *Voprosy istorii* 4:175–81. (in Russian)

UCEC (Ukrainian Central Executive Committee [Vseukrainskyi tsentralnyi vykonavchyi komitet, or VUTsVK]). 1933. *Administratyvno-terytorial'nyi podil USRR za stanom na 1 hrudnia 1933 roku* [Administrative-Territorial Structure of the UkrSSR as of 1 December 1933]. Kharkiv: Radianske Budivnytstvo i Pravo. (in Ukrainian)

———. 1936. *Dovidnyk adminterpodilu USRR za stanom na 15 hrudnia 1935 roku* [Administrative-Territorial Structure of the UkrSSR as of 15 December 1935]. Kharkiv: Radianske Budivnytstvo i Pravo. (in Ukrainian)

Viola, L. 1996. *Peasant Rebels Under Stalin: Collectivization and the Culture of Peasant Resistance.* New York: Oxford University Press.

Volkov, A.G. 1990. *Perepis' naseleniia SSSR 1937 goda* [1937 USSR Population Census]. Vol. 3–5, part II. Moscow: Informtsentr Goskomstata SSSR. (in Russian)

Vynnychenko, I. 1994. *Ukraïna 1920–1980 rokiv: Deportatsii, zaslannia, vysylannia* [Ukraine in the 1920s–80s: Deportations, Exile, Resettlement]. Kyiv: Rada. (in Ukrainian)

Weitsblit [Veitsblit], I. 1930. *Rukh ievreis'koï liudnosti na Ukraïni* [Vital Statistics of Jews in Ukraine]. Kyiv: Proletar. (in Ukrainian)

Wheatcroft [Uitkroft], S.G., and A. Garnaut. 2013. Poteri naseleniia v otdel'nykh raionakh SSSR (1929–1934): Statistika, karty i sravnitel'nyi analiz (osoboie polozhenie Ukrainy) [Population losses in selected districts of the USSR (1929–34): Statistics, maps and comparative analysis (specific situation in Ukraine)], in *Holod 1993: Ukraïnitsi* [1933 Famine: The Ukrainians]. Kyiv: Ukrainian Institute of National Memory, p. 376–91 (in Russian).

Wolowyna [Volovyna], O. 2013. Pomisiachnyi rozpodil demohrafichnykh vtrat vnaslidok holodu 1933 roku v Ukraïni [Monthly distribution of demographic losses resulting from the 1933 Famine in Ukraine], in *Holod v Ukraïni u pershii polovyni XX stolittia: Prychyny ta naslidky (1921–1923, 1932–1933, 1946–1947)*, proceedings of an International Scholarly Conference in Kyiv, 20–21 November 2013. Kyiv: NASU M.V. Ptoukha Institute of Demography and Social Studies, NASU Institute of History of Ukraine, Kyiv Shevchenko National University, National University of 'Kyiv-Mohyla Academy,' p. 233–42.

Yakovlev [Iakovlev], A.N., N.L. Pobol', and P.M. Polian (eds.). 2005. *Stalinskie deportatsii: 1928–1953* [Stalin's Deportations: 1928–53]. Moscow: Mezhdunarodnyi Fond 'Demokratiia' (Aleksandr N. Yakovlev Foundation) and Izd-vo Materik. (in Russian)

Zemskov, V.N. 2005. *Spetsposelentsy v SSSR, 1930–1960* [Forced Resettlements in the USSR, 1930–60]. Moscow: Nauka. (in Russian)

Zhiromskaia, V. 1990. Vsesoiuznye perepisi naseleniia 1926, 1937, 1939 godov: Istoriia podgotovki i provedeniia [All-Union Censuses in 1926, 1937, 1939: History of preparations and implementation]. *Istoriia SSSR* 3:84–103. (in Russian)

Canadian Studies in Population 43, no. 3–4 (2016): 203–14.

Using the probabilistic fertility table to test the statistical significance of fertility trends

Nan Li[1]

Abstract

At below replacement level, fertility changes are subtle and complex; and distinguishing statistically significant trends from random shifts is becoming a relevant issue. The probabilistic fertility table describes the uncertainty of the childbearing process, and provides a significance test for the annual changes of various fertility measures, which is essential for distinguishing between a statistically significant change from a random fluctuation. This paper provides an analytical model for the total fertility of the probabilistic fertility table, and extends the significance test to period trends that include multiple annual changes. The extended significance test indicates that complex annual changes could accumulate to become a significant trend. Applying the analytical model and extended test to the total fertility of Canada, it indicates that the 2000–11 upward trend is statistically significant and, therefore, supports recently projected future increases of total fertility.

Keywords: childbearing uncertainty, probabilistic fertility table, fertility trend, significance test, Canada.

Résumé

En-dessous du seuil de remplacement des générations, les changements à la fertilité sont subtils et complexes. Aussi, il est devenu pertinent de pouvoir distinguer les tendances significatives au plan statistique des écarts aléatoires. Le tableau probabiliste de fertilité décrit l'incertitude liée au processus de reproduction et fournit un critère de signification des changements annuels dans les diverses mesures de fertilité, élément essentiel pour distinguer un changement important au plan statistique des fluctuations aléatoires. Cet article fournit un modèle analytique pour l'ensemble du tableau probabiliste de fertilité et élargit la portée de cette mesure aux tendances dans le temps incluant les multiples changements annuels. Ce critère élargi indique que les changements complexes annuels peuvent représenter une tendance significative. En appliquant le modèle analytique et le critère au tableau de fertilité du Canada, on constate que la tendance à la hausse de 2000–11 est importante au plan statistique et, par conséquent, augure des hausses futures dans la fertilité totale.

Mots-clés : incertitude relative à la procréation, tableau probabiliste de fertilité, tendance en fertilité, critère de signification.

Introduction

Total fertility in more developed regions declined across replacement level in the middle of the 1970s, and has stayed below that level since then (United Nations 2015). This unprecedented phenomenon caused the 'Low Fertility Trap' hypothesis (Lutz, Skirbekk, and Testa 2006). Starting from year 2000, many developed countries experienced slight increases in their fertility rates (Goldstein, Sobotka, and Jasilioniene 2009; Bongaarts and Sobotka 2012). After 2008, however, some developed countries have seen minor declines in their fertility levels (Goldstein et al. 2013). At above replacement level, remarkable annual declines in total fertility often demon-strate obvious downward trends. At levels below replacement, however, fertility changes are much more subtle and complex. As a result, whether the successive annual changes in a certain period compose a genuine trend or a random shift is becoming an important question.

1. Nan Li, DC2-1938, United Nations Population Division, New York, NY 10017. E-mail: li32@un.org.

To provide a statistical answer to this question, analysis is needed of the uncertainty of the childbearing process. The conventional total fertility (TF) is defined as the sum of age-specific fertility rates that do not distinguish the order of births and the parity of women, and therefore do not offer a basis to carry out probabilistic analysis. Standard errors of TF are estimated, for example, in the Demographic and Health Surveys (e.g., Statistics Indonesia et al. 2013) to measure sampling errors. These standard errors, however, do not indicate the uncertainty of the childbearing process for which the data collected from a whole country are available.

In order to investigate the uncertainty of the childbearing process, an analytic model for the total fertility of the probabilistic fertility table[2] is proposed in this paper. Based on this analytic model, the statistical significance test is extended from an individual annual change to a period trend that includes multiple annual changes. Furthermore, the period trend significance test indicates that multiple insignificant and complex annual changes could gradually accumulate to a significant trend, thus providing important insight for the analysis of fertility change and for fertility projections.

An analytic model of total fertility

The probabilistic fertility table describes the uncertain childbearing process of a hypothetical cohort of women who are subject to neither to mortality nor migration. Following the rationale of developing probabilistic life tables (Li 2015a), the number of women in the hypothetical cohort is specified to minimize the differences between the hypothetical cohort and the observed population. Let the minimal and maximal reproductive age be $a_{\min}$ and $a_{\max}$,[3] respectively; and let the number of women at age $a_{\min}$ be $l_0(a_{\min})$, where subscript 0 refers to having zero children. Then, the number of women is $l_0(a_{\min})$ at all reproductive ages, because of the absence of mortality and migration. Denote the number of observed female population at age a by $f_p(a)$. Then, minimizing $\sum_{a=a_{\min}}^{a_{\max}}[l_0(a_{\min}) - f_p(a)]^2$ leads to $l_0(a_{\min})$ being the average of the observed female population over reproductive ages:

$$l_0(a_{\min}) = \frac{\sum_{a=a_{\min}}^{a_{\max}} f_p(a)}{a_{\max} - a_{\min} + 1}. \tag{1}$$

The uncertain childbearing process can be simulated by assuming that a woman delivers children independently from the others according to the Bernoulli distribution and the probability of delivering children by age and parity[4] in a certain year, starting from the minimal reproductive age $a_{\min}$. With a value of $l_0(a_{\min})$, this simulation can be repeated for each of the $l_0(a_{\min})$ women; and a sample of the childbearing process of the hypothetical cohort is obtained, which provides a sample fertility table. A large number of sample fertility tables then comprises a probabilistic fertility table (Li 2015b).

Focusing on total fertility, an analytical model can be derived as below. Let the number of children of the *jth* woman at age $a_{\max}$ be a random variable, X_j. Then, the total fertility of the probabilistic fertility table (TF_f), which is defined as the average number of children per woman at age $a_{\max}$, is

$$TF_f = \frac{\sum_{j=1}^{l_0(a_{\min})} X_j}{l_0(a_{\min})}. \tag{2}$$

Further, let the probability for a woman to have *i* children at age $a_{\max}$ be p_i. Then, the mean and variance of X_j are $\sum_i i \cdot p_i$ and $\sum_i (i - \sum_k k \cdot p_k)^2 \cdot p_i$. According to (2), the mean and variance of TF_f are therefore

2. The quantitative values of conventional total fertility are often close to that of the probabilistic fertility table total fertility.
3. In this paper, a_{min} and a_{max} are taken as the commonly employed 15 and 50 years, respectively.
4. The birth parity of a woman refers to the number of children she has delivered.

$$\mu = Mean(TF_f) = \sum_i i \cdot p_i \tag{3}$$

and

$$\sigma^2 = Var(TF_f) = \frac{\sum_i (i - \sum_k k \cdot p_k)^2 \cdot p_i}{l_0(a_{\min})} \tag{4}$$

Formulas (3) and (4) indicate that the mean of TF_f is independent from $l_0(a_{\min})$, but the variance of TF_f is inversely proportional to $l_0(a_{\min})$. In other words, the uncertainty of TF_f is smaller when the population size is larger, and vice versa.

Finally, when μ is not close to zero and $l_0(a_{\min})$ is larger than 30 (see Agresti and Finlay 1997: 104), the law of large numbers provides an analytical model for TF_f as

$$TF_f \sim N(\mu, \sigma^2)\,. \tag{5}$$

Using observed data, the mean and variance of TF_f are estimated according to the formulas in the appendix, as $\hat{\mu}$ and $\hat{\sigma}^2$.

Significance tests of the changes and trends of TF_f

The statistical significance of an annual change of TF_f

When total fertilities are forecasted by time-series models, they are correlated over time because of containing the same modelling errors of previous years. It is worth noting that the uncertainty is introduced from the errors of modelling the over-time changes in total fertility. Moreover, in a time-series model, uncertainty cannot be assigned to total fertility in the initial years, because there are no modelling errors.

On the other hand, in the probabilistic fertility table, TF_f is uncertain in any year. The uncertainty does not come from modelling errors but from the uncertain childbearing process of the hypothetical cohort, in which a woman's childbearing behaviour is assumed to be independent from the others. Subsequently, the uncertainties in $TF_f(\mathrm{t})$ and $TF_f(\mathrm{t}+1)$ are caused by the uncertain childbearing processes of two hypothetical cohorts. Since each woman's childbearing behaviour is assumed to be independent from the others, the childbearing processes of the two hypothetical cohorts are consequently independent.

Let the mean and variance of $TF_f(t)$ and $TF_f(t+1)$ be $\mu(t)$, $\mu(t+1)$, $\sigma^2(t)$, and $\sigma^2(t+1)$, respectively. Then, setting the null hypothesis as

$$H_0: \quad \mu(t) = \mu(t+1)\,, \tag{6}$$

and noting that $TF_f(t)$ and $TF_f(t+1)$ are independent, and that $l_0(a_{\min})$ is large (so $\hat{\sigma}^2$ is close to σ^2), we obtain

$$Z(t) = -\frac{TF_f(t) - TF_f(t+1)}{\sqrt{\hat{\sigma}^2(t) + \hat{\sigma}^2(t+1)}} \sim N(0,1)\,, \tag{7}$$

where a negative sign is used to make a positive Z, representing an increase in total fertility.

If the estimated value of $Z(t)$, namely,

$$\hat{z}(t) = -\frac{\hat{\mu}(t) - \hat{\mu}(t+1)}{\sqrt{\hat{\sigma}^2(t) + \hat{\sigma}^2(t+1)}}\,, \tag{8}$$

is found to be outside (−1.96, 1.96), which occurs with a probability smaller than 0.05 according to the null hypothesis, then the null hypothesis is rejected, implying that the change in TF_f is statistically significant. Other-

wise, the change in TF_f cannot be concluded to be statistically significant. The above procedure can also be used to test the significance of the difference between the total fertility in two separate years or of two populations, and in general could be called a two-point significance test.

The statistical significance of a period trend of TF_f

For a period that includes multiple years, namely from year 1 through year *t*, the significance of the difference in TF_f between year 1 and year *t* can be tested using the above two-point significance test. When the difference between year 1 and year *t* is insignificant, there is no significant trend in period $[1, t]$, because a significant trend should not lead to an insignificant difference. When the difference between year 1 and year *t* is significant, however, the trend in period $[1, t]$ may not necessarily be significant. A simple example is that the annual change between year 1 and year 2 is just significant, and there is no change later. In this example, then, the difference between year 1 and year *t* is significant, because the distance between the two points does not matter in a two-point significance test. But intuitively there is no significant trend when *t* is large, because among the multiple annual changes only the first one is just significant and all others are zero. In real situations, there may be significant and insignificant annual changes over a certain time interval. These changes may not be exactly zero and they may cancel each other. In these situations, whether there is a significant trend depends on the details of the annual changes.

To test the significance of a period trend that includes multiple annual changes, the difficulty is that in calculating the average change over a period, the middle values of TF_f will cancel each other and only the first and last values matter. A solution to overcome this difficulty is to construct the average of odd and even ranked changes. Here, *odd rank* signifies that the earlier year of each annual change being an odd number; and *even rank* is defined analogously. Subsequently, the odd and even ranked average changes, namely Y_1 and Y_2, are constructed as:

$$\begin{aligned} Y_1 &= -\frac{[TF_f(1)-TF_f(2)]+[TF_f(3)-TF_f(4)]+\ldots+[TF_f(t_1-1)-TF_f(t_1)]}{t_1/2}, \\ Y_2 &= -\frac{[TF_f(2)-TF_f(3)]+[TF_f(4)-TF_f(5)]+\ldots+[TF_f(t_2-1)-TF_f(t_2)]}{(t_2-1)/2}. \end{aligned} \tag{9}$$

Noting that $TF_f(i)$ and $TF_f(j)$ are independent and that $l_0(a_{\min})$ is large, we obtain the following relations for the variances:

$$Var(Y_1) = \frac{\sum_{i=1}^{t_1}\sigma^2(t)}{t_1^2/4} \approx \frac{\sum_{i=1}^{t_1}\hat{\sigma}^2(t)}{t_1^2/4}, \quad Var(Y_2) = \frac{\sum_{i=2}^{t_2}\sigma^2(t)}{(t_2-1)^2/4} \approx \frac{\sum_{i=2}^{t_2}\hat{\sigma}^2(t)}{(t_2-1)^2/4}. \tag{10}$$

To test the statistical significance of period trends, the null hypothesis can be set as *no trend*,

$$H_0: \quad Mean(Y_1) = Mean(Y_2) = 0, \tag{11}$$

and the alternative hypothesis can be set, for even and odd ranked trends both exist and do not cancel each other, as

$$H_a: \quad Mean(Y_1) > 0, \quad Mean(Y_2) > 0, \quad or \quad Mean(Y_1) < 0, \quad Mean(Y_2) < 0. \tag{12}$$

According to the null hypothesis, there are

$$Z_1 = \frac{Y_1}{\sqrt{Var(Y_1)}} = -\frac{[TF_f(1) - TF_f(2)] + ... + [TF_f(t_1 - 1) - TF_f(t_1)]}{\sqrt{\sum_{t=1}^{t_1} \hat{\sigma}^2(t)}} \sim N(0,1),$$
$$Z_2 = \frac{Y_2}{\sqrt{Var(Y_2)}} = -\frac{[TF_f(2) - TF_f(3)] + ... + [TF_f(t_2 - 1) - TF_f(t_2)]}{\sqrt{\sum_{t=2}^{t_2} \sigma^2(t)}} \sim N(0,1). \quad (13)$$

Subsequently, the corresponding sample values of Z_1 and Z_2 are

$$\hat{z}_1 = -\frac{[\hat{\mu}(1) - \hat{\mu}(2)] + ... + [\hat{\mu}(t_1 - 1) - \hat{\mu}(t_1)]}{\sqrt{\sum_{t=1}^{t_1} \hat{\sigma}^2(t)}},$$
$$\hat{z}_2 = -\frac{[\hat{\mu}(2) - \hat{\mu}(3)] + ... + [\hat{\mu}(t_2 - 1) - \hat{\mu}(t_2)]}{\sqrt{\sum_{t=2}^{t_2} \hat{\sigma}^2(t)}}. \quad (14)$$

Then, if

$$\hat{z}_1 \geq 1.96,\ \hat{z}_2 \geq 1.96, \text{ or } \hat{z}_1 \leq -1.96,\ \hat{z}_2 \leq -1.96, \quad (15)$$

the null hypothesis is rejected and the alternative hypothesis is in favour, which indicates that both the even and odd ranked trends are statistically significant and they do not cancel each other. In other words, the whole trend is statistically significant. On the other hand, if (15) does not stand, then the whole trend cannot be concluded to be statistically significant, although the null hypothesis could still be rejected.[5]

Multiple insignificant and complex changes could accumulate to a significant trend

Consider now the formulas in (14). Compared to the differentials that enlarge the effect of random fluctuation in the numerators, the over-time changes in $\hat{\sigma}(t)$ in the denominators are negligible in common situations (e.g., Figure 3). Thus, $\hat{\sigma}(t)$ can be approximately replaced by their average, $\overline{\sigma}$. Subsequently, denoting the over-time average of odd ranked $\hat{z}(t)$ by $\overline{z}_1$, the following relations hold:

$$\hat{z}_1 = -\frac{[\hat{\mu}(1) - \hat{\mu}(2)] + ... + [\hat{\mu}(t_1 - 1) - \hat{\mu}(t_1)]}{\sqrt{\sum_{t=1}^{t_1} \hat{\sigma}^2(t)}} \approx -\frac{[\hat{\mu}(1) - \hat{\mu}(2)] + ... + [\hat{\mu}(t_1 - 1) - \hat{\mu}(t_1)]}{\overline{\sigma}\sqrt{t_1}}$$
$$= \frac{\sqrt{2}}{\sqrt{t_1}} \cdot \{-\frac{[\hat{\mu}(1) - \hat{\mu}(2)]}{\overline{\sigma}\sqrt{2}} - ... - \frac{[\hat{\mu}(t_1 - 1) - \hat{\mu}(t_1)]}{\overline{\sigma}\sqrt{2}}\} \quad (16)$$
$$\approx \frac{\sqrt{t_1}}{\sqrt{2}} \cdot \{\frac{\hat{z}(1) + ... + \hat{z}(t_1 - 1)}{t_1/2}\} = \sqrt{\frac{t_1}{2}} \cdot \overline{z}_1, \quad t_1 \geq 4.$$

For the same reason,

$$\hat{z}_2 \approx \sqrt{\frac{t_2 - 1}{2}} \cdot \overline{z}_2, \quad t_2 \geq 5, \quad (17)$$

where $\overline{z}_2$ is the over-time average of the even-ranked $\hat{z}(t)$.

5. Similar to a one-sided significance test ($H_0: \mu = 0$, $H_a: \mu > 0$), here the null and alternative hypotheses are not complementary; rejecting the null hypothesis does not lead to accepting the alternative hypothesis.

Now, consider a case in which the annual change of $TF_f(t)$ is linear and insignificant ($-1.96 < \hat{z}(t) = \overline{z}_1 = \overline{z}_2 < 1.96$). In this case, all individual changes are insignificant, but the period trend is significant when the number of years is large, because (16) and (17) indicate that large values of t_1 and t_2 will make $\hat{z}_1, \hat{z}_2 \geq 1.96$ or $\hat{z}_1, \hat{z}_2 \leq -1.96$. In real situations, the changes in $TF_f(t)$ are not linear; $\hat{z}(t)$ are not constant and could be positive in one year but negative in another. Nonetheless, (16) and (17) still indicate that even if individual annual changes are insignificant and contain cancellations over time, $\overline{z}_1$ and $\overline{z}_2$ could still have the same sign, and therefore the period trend could still be significant when the number of years is large.

Although insignificant and complex changes could accumulate to a significant trend, it is not guaranteed. Using (16) and (17), a condition for a trend to be significant is obtained as

$$\overline{z}_1 \geq \frac{1.96}{\sqrt{t_1/2}}, \quad \overline{z}_2 \geq \frac{1.96}{\sqrt{(t_2-1)/2}}, \quad or \quad \overline{z}_1 \leq \frac{-1.96}{\sqrt{t_1/2}}, \quad \overline{z}_2 \leq \frac{-1.96}{\sqrt{(t_2-1)/2}}. \tag{18}$$

Different from the two-point significance test, in which the distance between the two points does not matter, in (18) the length of the period (t_1 or t_2) matters: the trend is more likely significant when the length of the period is longer, given the values of $\overline{z}_1$ and $\overline{z}_2$.

Applications

An analytical model

Based on the law of large numbers, (5) indicates that the probability distribution of TF_f is approximately normal, of which the mean and variance are estimated according to the formulas in the appendix. Using the latest (year 2011) data on age-parity-specific fertility rates of Canada in the Human Fertility Database (MPIDR and VID 2013), the normal distribution of TF_f is computed and shown by the solid curve in Figure 1. Compared to the numerical distribution that is computed through simulation using 1,000 sample fertility tables (Li 2015b) and described by the squares in Figure 1, we see that the analytical model works well.

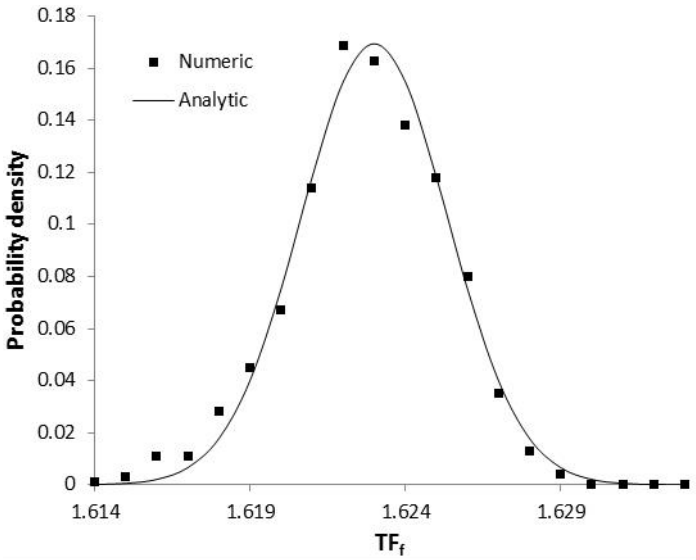


Figure 1. Probability distribution of total fertility, Canada 2011.

Changes and trends in total fertility

Since the year 2000, total fertility has started to increase in many low-fertility countries (Goldstein, Sobotka, and Jasilioniene 2009). Bongaarts and Sobotka (2012) explained the main reasons as the consequence of pro-family policies and a diminishing pace of the postponement of childbearing. After 2008, some developed

countries have seen minor declines in their fertility levels, which may be caused by the global financial crisis (Goldstein et al. 2013). These changes appeared also in Canada, as can be seen in Figure 2, in which the values of fertility table total fertility TF_f and conventional total fertility TF are compiled from the Human Fertility Database up to 2011.

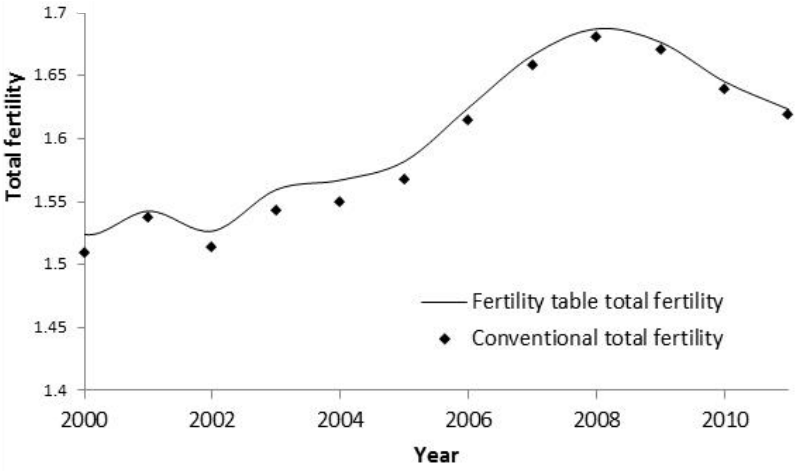


Figure 2. Total fertility of Canada, 2000–11.

The differences between the values of TF_f and TF are negligible. The reason could be that the age distributions of woman by birth parity approached that of the hypothetical cohort. Regardless of using TF_f or TF, the changes in Canadian total fertility are typical among developed countries: increased slightly from 2000 or 2002; and started to decline after 2008. Starting from 2000 or some later years, the overall trend is an increasing one. The overall increase trend has been used as the basis of fertility and population projections by many low-fertility countries, including Canada (Bohnert et al. 2015). Because these trends are subtle and include offsets, whether they are statistically significant becomes an important question; and an answer for Canada is provided below.

Results of significance test

Applying (14) to the data for Canada in 2000–11, the results are $\hat{z}_1 = 9.4$ and $\hat{z}_2 = 3.3$, both greater than 1.96. Thus, the overall upward trend in 2000–11 is statistically significant. Because projections are believed to be better when based on a longer period, it is not practically useful to test the trends in shorter periods starting later than 2000. Given the annual declines in total fertility, especially 2008–11, how can the overall upward trend in 2000–11 be statistically significant? It can be explained using Figure 3 and the condition in (18).

In general, complex annual changes that include offsets could accumulate to a statistically significant trend under certain conditions. In the common situation that the standard deviation of TF_f is approximately constant compared to the annual changes in TF_f, a simple condition for complex annual changes to compose a statistically significant trend is found as (18).

Although the changes in total fertility are declines in 2001–02 and 2008–11, they are increases in 7 other years. Figure 3 indicates that Canada is in the common situation so that condition (18) applies. Note also that with

$$\bar{z}_1 = 3.7 > \frac{1.96}{\sqrt{t_1/2}} = \frac{1.96}{\sqrt{9/2}} \text{ and } \bar{z}_2 = 2 > \frac{1.96}{\sqrt{(t_2-1)/2}} = \frac{1.96}{\sqrt{9/2}},$$ (18) is satisfied and indicates that after off-

setting with the declines, the increases still accumulate to a statistically significant trend. Nonetheless, if annual declines after 2011 occurred dominantly, $\bar{z}_1$ and $\bar{z}_2$ would be reduced to break (18), and the annual changes starting from 2000 would not accumulate to a statistically significant trend.

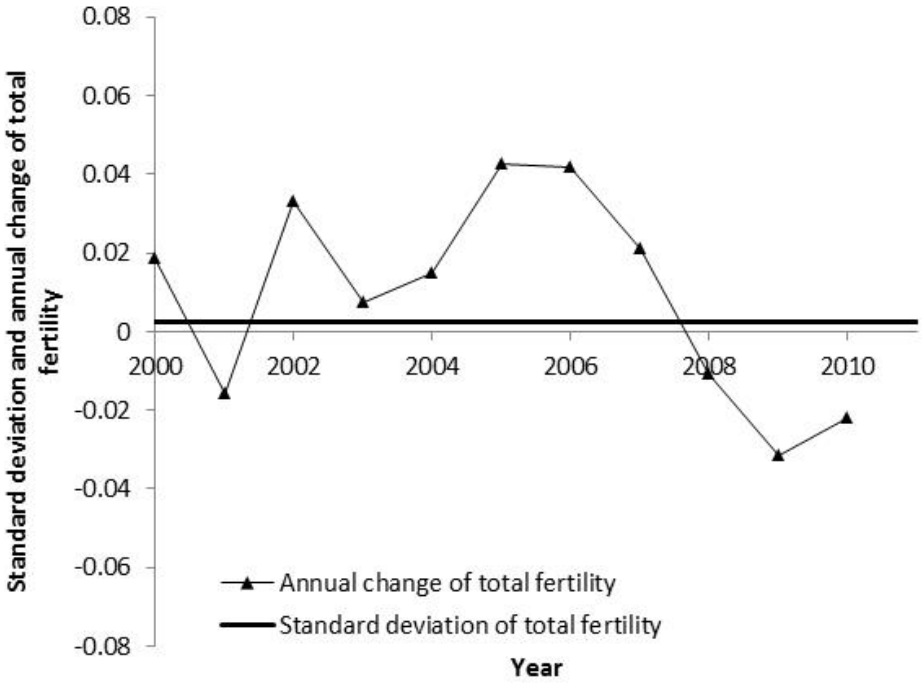


Figure 3. Standard deviation and annual change of total fertility, Canada 2000–11.

At below replacement level, the annual changes in total fertility are subtle and complex; and whether they compose a genuine trend or a random shift is an important question. This question is common among low-fertility countries, among which Canada is not an exception: of the 11 recent annual changes, 7 are increases and 4 are declines. Differing from the other countries for which this question remains open, in Canada we see that the 11 recent annual changes have accumulated to a genuine trend.

Basing fertility projections on the trend of a recent period is a common practice. This basis is obviously sounder when the period is longer, and should be more reasonable when the trend is statistically significant. An 11-year increase of trend in total fertility is proper to empirically support the 10-year increase of total fertility in the medium projections of Statistics Canada (Bohnert et al. 2015). Confirming the statistical significance suggests that this 11-year upward trend for Canada is genuine, and provides statistical support to these medium projections.

Summary

The probabilistic fertility table describes the uncertainty of the childbearing process, and hence provides significance tests for an annual change in various fertility variables. On the other hand, the probabilistic fertility table requires immense calculation. Moreover, how to test the statistical significance of a fertility trend that includes multiple annual changes is still an open question. The purposes of this paper have been to simplify the application of the probabilistic fertility table, and to extend the test of significance from an annual change to a period trend that includes multiple annual changes.

Using the law of large numbers, the total fertility of the probabilistic fertility table is found to obey the normal distribution approximately, whose mean and variance can be estimated using analytical formulas. This analytical model substantially simplifies significance tests, in both description and calculation. It should be mentioned that in the probabilistic fertility table, variables other than total fertility may not obey normal distribution, or may not be described by analytical distributions.

Constructing the even and odd ranked average changes in total fertility over a period, the obstacle of the middle values of TF_f offsetting each other is avoided. Using the normal distribution of total fertility, the even

and odd ranked average changes are found to also obey normal distributions. Thus, the null hypothesis that the mean values of the average changes are zero, or there is no trend over the period, can be tested. Furthermore, the test procedure indicates that multiple insignificant and complex annual changes may accumulate to a significant period trend.

Finally, applying the analytical model and extended test to the fertility data of Canada, the results indicate that the 11 recent annual changes in total fertility, of which 7 are increases and 4 are declines, accumulated to a statistically significant trend. This 11-year statistically significant trend supports Statistics Canada's recently projected future increases of total fertility.

Acknowledgments

The views expressed in this paper are those of the author and do not necessarily reflect those of the United Nations. The author thanks two reviewers, as well as the editor of the journal, for their useful comments and suggestions.

References

Agresti, A., and B. Finlay. 1997. *Statistical Methods for the Social Sciences.* Upper Saddle River, NJ: Prentice Hall.

Bohnert, N., J. Chagnon, S. Coulombe, P. Dion, and L. Martel. 2015. Population Projections for Canada (2013 to 2063), Provinces and Territories (2013 to 2038): Technical Report on Methodology and Assumptions. Statistics Canada. Available at http://www.statcan.gc.ca/pub/91-620-x/91-620-x2014001-eng.htm

Bongaarts, J., and T. Sobotka. 2012. A demographic explanation for the recent rise in European fertility. *Population and Development Review* 38(1):38–120.

Goldstein, J.R., M. Kreyenfeld, A. Jasilioniene, and D.K. Örsal. 2013. Fertility reactions to the "Great Recession" in Europe: Recent evidence from order-specific data. *Demographic Research* 29:85–104.

Goldstein, J.R., T. Sobotka, and A. Jasilioniene. 2009. The end of lowest-low fertility? *Population and Development Review* 35(4):663–99.

Li, N. 2015a. The probabilistic life table and its applications to Canada. *Canadian Studies in Population* 42(1–2):117–29.

———. 2015b. The probabilistic Fertility Table and Its Applications. Paper presented at the 2015 Annual Meeting of the Population Association of America, San Diego.

Lutz, W., V. Skirbekk, and M.R. Testa. 2006. The low fertility trap hypothesis: Forces that may lead to further postponement and fewer births in Europe. *Vienna Yearbook of Population Research* 4:167–92. Vienna: Austrian Academy of Sciences Press.

MPIDR and VID (Max Planck Institute for Demographic Research [Germany] and Vienna Institute of Demography [Austria]). 2013. Human Fertility Database. Available at www.humanfertility.org (downloaded in June 2016)

Preston, S.H., P. Heuveline, and M. Guillot. 2001. *Demography, Measuring and Modeling Population Processes.* Oxford (UK): Blackwell Publishers Ltd.

Statistics Indonesia (Badan Pusat Statistik, or BPS), National Population and Family Planning Board (BKKBN), and Kementerian Kesehatan (Kemenkes MOH), and ICF International. 2013. *Indonesia Demographic and Health Survey 2012*. Jakarta: BPS, BKKBN, Kemenkes, and ICF International. Available at http://dhsprogram.com/publications/publication-fr275-dhs-final-reports.cfm

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

Appendix

To simplify the equations in this appendix, all the variables are used to represent the corresponding estimated values. For the hypothetical cohort, let the number of women having $(i-1)$ children at age a be $l_{i-1}(a)$, and the number of children delivered by these women at ages $[a,a+1)$ be $b_i(a)$. Then, the probability of delivering the *ith* child in age interval $[a,a+1)$, namely, $q_i(a)$, can be defined as

$$q_i(a)=\frac{b_i(a)}{l_{i-1}(a)},\quad q_{m+}(a)=\frac{b_{m+}(a)}{l_{(m-1)+}(a)}, \tag{A.1}$$

where $(m-1)+$ indicates the open parity of having the $(m-1)th$ and higher-order children. Using definition (A.1), the childbearing process is written as

$$\begin{aligned} l_{i-1}(a) &= \begin{cases} l_{i-1}(a-1)[1-q_i(a-1)], & i=1, \\ l_{i-1}(a-1)[1-q_i(a-1)]+l_{i-2}(a-1)q_{i-1}(a-1), & 1<i<m, \end{cases} \\ l_{(m-1)+}(a) &= l_{(m-1)+}(a-1)+b_{(m-1)}(a-1). \end{aligned} \tag{A.2}$$

Using population data of census and estimates, and data on births of vital registrations, the values of the age-parity-specific fertility rate for a certain time interval can be computed as:

$$M_i(a)=\frac{\textit{Number of the ith births delivered by below women}}{\textit{Person-years of women having } (i-1) \textit{ children at ages } [a,a+1)}. \tag{A.3}$$

In (A.3), both the numerator and the denominator refer to a certain time interval, which may or may not be a calendar year. It should be mentioned that, although the age interval must be one year for fertility table, the time interval to which a fertility table refers can be flexible such as 5 years. This is important for small populations, of which a longer time interval should contain more births and hence make the age-parity-specific fertility rates more robust.

Because fertility may change only slightly in one year interval of age and a moderate time interval, there is approximately (see Preston, Heuveline, and Guillot, 2001)

$$M_i(a)=m_i(a), \tag{A.4}$$

where $m_i(a)$ represents the age-parity-specific fertility rate of the hypothetic cohort, and is defined as

$$m_i(a)=\frac{b_i(a)}{L_{i-1}(a)}, \tag{A.5}$$

where $L_{i-1}(a)$ represents the person-years of the $(i-1)th$ parity in $[a,a+1)$:

$$L_{i-1}(a)=\int_{y=a}^{a+1} l_{i-1}(y)dy, \tag{A.6}$$

and $l_{i-1}(y)$ represents the number of women of parity $(i-1)$ at age y.

Using $m_i(a)$, $q_i(a)$, and $l_{(i-1)}(a)$ can be computed. For $i=1$, $L_0(a)$ is the population exposed to the chance of having the first child at ages $[a,a+1)$. For $i>1$, however, $L_{i-1}(a)$ is not the population exposed to the chance of having the *ith* child at ages $[a,a+1)$, because some women entered parity $(i-1)$ by bearing the $(i-1)th$ child at ages $[a,a+1)$ and thence are not exposed to the chance of having the *ith* child within the rest of the calendar year, according to the assumption that a woman can bear at most one birth in one year.

Under the assumption that the births occur evenly in each age interval, both the decline (due to delivering the *ith* child) and the increase (due to delivering the $(i-1)th$ child) of $l_{i-1}(a)$ are linear functions of age. Thus, $l_{i-1}(a)$ changes with a linearly, and therefore

$$L_{i-1}(a) = 0.5 \cdot [l_{i-1}(a) + l_{i-1}(a+1)]. \tag{A.7}$$

For the case of $i=1$, (A.7) leads to

$$m_1(a) = \frac{b_1(a)}{L_0(a)} = \frac{q_1(a) l_0(a)}{0.5 \cdot [l_0(a) + l_0(a+1)]} \ . \tag{A.8}$$

Using the first line of (A.2), (A.7) is rewritten as

$$m_1(a) = \frac{q_1(a)}{0.5 \cdot [1 + l_0(a+1)/l_0(a)]} = \frac{q_1(a)}{0.5 \cdot [1 + (1 - q_1(a))]}, \tag{A.9}$$

which yields

$$q_1(a) = \frac{m_1(a)}{1 + 0.5 \cdot m_1(a)}. \tag{A.10}$$

Equation (A.10) is identical to the corresponding formula in life tables, because $L_0(a)$ is the population exposed to the chance of having the first child at ages $[a,a+1)$.

For the cases of $m>i>1$, (A.7) still yields

$$m_i(a) = \frac{b_i(a)}{L_{i-1}(a)} = \frac{q_i(a) l_{i-1}(a)}{0.5 \cdot [l_{i-1}(a) + l_{i-1}(a+1)]}, \tag{A.11}$$

but now the second line of (A.2) applies, and leads to

$$m_i(a) = \frac{b_i(a)}{L_{i-1}(a)} = \frac{q_i(a) l_{i-1}(a)}{0.5 \cdot [l_{i-1}(a) + l_{i-2}(a) q_{i-1}(a) + l_{i-1}(a)(1 - q_i(a))]}. \tag{A.12}$$

Rewriting (A.12), we obtain

$$q_i(a) = \frac{m_i(a)}{1 + 0.5 \cdot m_i(a)} [1 + \frac{0.5 \cdot l_{i-2}(a) q_{i-1}(a)}{l_{i-1}(a)}], \quad i > 1. \tag{A.13}$$

The difference between (A.10) and (A.13) is caused by that, although $L_{i-1}(a)$ is still the person years, it is no longer the population exposed to the chance of having the *ith* child at ages $[a,a+1)$ for $i>1$. This can be explained as below. For $i>1$, the $(i-1)th$ births make $L_{i-1}(a)$ larger than the population exposed to the chance of bearing the *ith* birth, and the $m_i(a)$ smaller, comparing to that of $i=1$. Thus, as a compensation, (A.13) includes an additional term, compared to (A.10). This additional term makes the calculation slightly complicated.

In (A.13), $q_i(a)$ and $l_{i-1}(a)$ are unknown, and can be solved iteratively together with the second line of (A.2):

$$\begin{aligned} l_{i-1}(a) &= l_{i-2}(a-1) q_{i-1}(a-1) + l_{i-1}(a-1)[1 - q_i(a-1)], \\ q_i(a) &= \frac{m_i(a)}{1 + 0.5 \cdot m_i(a)} [1 + \frac{0.5 \cdot l_{i-2}(a) q_{i-1}(a)}{l_{i-1}(a)}]. \end{aligned} \tag{A.14}$$

The iteration starts from $i=2$, of which $q_1(a)$ and $l_0(a)$ for all a are already computed as the result of $i=1$. Using the first line of (A.14), $l_1(a_{min}+1)$ is obtained, because it is known that $q_2(a_{min})=0$, according to the

assumption that a woman could deliver only one child in a year. Subsequently, $q_2(a_{min}+1)$ is obtained from the second line of (A.14). When $q_2(a_{min}+1)$ is known, $l_1(a_{min}+2)$ is obtained from the first line of (A.14), and so is $q_2(a_{min}+2)$ from the second line of (A.14). Repeating this process, $q_2(a)$ and $l_1(a)$ for all a are obtained. Now the iteration reaches $i=3$, of which $q_2(a)$, $q_1(a)$, $l_1(a)$, and $l_0(a)$ are already computed. Here, $q_3(a)$ and $l_2(a)$ for all a can be computed in the way similar to that of $i=2$, starting from $q_3(a)=0$ for $a \leq (a_{min}+1)$ according to the assumption that a woman could deliver only one child in a year. Repeating the process, $q_1(a)$ and $l_{i-1}(a)$ for all $i \leq (m-1)$ are obtained.

For the open parity $q_{m+}(a)$, the assumption that the births occur evenly in each age interval still leads to

$$m_{m+}(a) = \frac{b_{m+}(a)}{L_{(m-1)+}(a)} = \frac{q_{m+}(a) l_{(m-1)+}(a)}{0.5 \cdot [l_{(m-1)+}(a) + l_{(m-1)+}(a+1)]} . \tag{A.15}$$

Using the third line of (A.2), we obtain

$$m_{m+}(a) = \frac{b_{m+}(a)}{L_{(m-1)+}(a)} = \frac{q_{m+}(a) l_{(m-1)+}(a)}{0.5 \cdot [l_{(m-1)+}(a) + l_{(m-1)+}(a) + l_{(m-2)}(a) q_{(m-1)}(a)]} , \tag{A.16}$$

which leads to

$$q_{m+}(a) = m_{m+}(a)[1 + \frac{0.5 \cdot l_{m-2}(a) q_{m-1}(a)}{l_{(m-1)+}(a)}] . \tag{A.17}$$

Different from the case of $i<m$, in which $q_i(a)$ are computed iteratively, all $q_{m+}(a)$ can be computed by (A.17); this is because $l_{(m-1)+}(a)$ can be calculated given that the hypothetical cohort subjects neither mortality nor migration:

$$l_{(m-1)+}(a) = l_0(a_{\min}) - \sum_{i=0}^{m-2} l_i(a) . \tag{A.18}$$

After obtaining $q_1(a)$ and $l_{(i-1)}(a)$, the probabilities of having i children at the maximal reproductive age are obtained as

$$p_i = \frac{l_i(a_{\max})}{l_0(a_{\min})}, \quad p_{(m-1)+} = \frac{l_{(m-1)+}(a_{\max})}{l_0(a_{\min})} , \tag{A.19}$$

Finally, the mean (μ) and variance (σ^2) of total fertility are estimated as

$$\hat{\mu} = \sum_{i=1}^{(m-2)} i \cdot p_i + p_{(m-1)+} \cdot [(m-1) + \frac{\sum_{a=a_{\min}}^{a_{\max}} q_{m+}(a) \cdot l_{(m-1)+}(a)}{l_{(m-1)+}(a_{\max})}] \tag{A.20}$$

and

$$\hat{\sigma}^2 = \sum_{i=1}^{(m-2)} (i-\mu)^2 \cdot p_i + p_{(m-1)+} \cdot \{[(m-1) + \frac{\sum_{a=a_{\min}}^{a_{\max}} q_{m+}(a) \cdot l_{(m-1)+}(a)}{l_{(m-1)+}(a_{\max})}] - \mu\}^2 , \tag{A.21}$$

respectively.

Canadian Studies in Population 43(3–4): 215–33.

Changes in cause-specific mortality among the elderly in Canada, 1979–2011

Marie-Pier Bergeron-Boucher[1]
Robert Bourbeau
Jacques Légaré

Abstract

The structure of causes of death in Canada has been changing since the onset of the "cardiovascular revolution." While mortality due to cardiovascular diseases has been declining, mortality due to other causes of death, such as cancers and Alzheimer's disease has been increasing. Our research investigates how these changes have re-modeled life expectancy at age 65 and age 85, and what specific causes of death are involved. We distinguish between premature and senescent deaths in Canada, using a cause-specific age structure. Our results suggest that although a decline in premature deaths has contributed to increasing life expectancy in recent years, most of the gains in life expectancy at age 65 and 85 have resulted from a decline in senescent deaths. We also find a decline in mortality due to the main causes of death, leading to a greater diversification of causes.

Keywords: Causes of death, life expectancy, age patterns of mortality, Canada, decomposition.

Résumé

Depuis le début de la révolution cardiovasculaire, le Canada a connu d'importants changements dans la distribution des décès selon la cause. La mortalité par maladies cardiovasculaires a connu une importante diminution alors que les taux de mortalité pour les cancers et pour la maladie d'Alzheimer ont augmenté. Cet article examine comment ces changements ont influencé les tendances de l'espérance de vie à 65 et à 85 ans et quelles causes de décès spécifiques furent impliquées. Une distinction entre les décès prématurés et les décès liés à un processus de sénescence est réalisée, se basant sur deux indicateurs de variations par âge des causes de décès. Nos résultats suggèrent que la majorité des gains en espérance de vie à 65 et 85 ans proviennent d'une plus faible mortalité par cause de décès sénescente. De plus, une diminution des principales causes de décès chez les personnes âgées de 65 ans et plus laisse place à une plus grande diversification de causes aux grands âges.

Mots-clés : Causes de décès, espérance de vie, variations par âge, Canada, décomposition.

Background

Life expectancy has undergone an important increase over the last 170 years in many industrialized countries (Oeppen and Vaupel 2002). Changes in age and cause of death structures of mortality account for this development. The epidemiological transition, as formulated by Omran (1971), summarizes these changes. The epidemiological transition involves a shift from a high to a low stable level of mortality, in which degenerative and man-made diseases replace infectious diseases as the main causes of death (Omran 1971). It also involves an important transformation of the age structure of mortality, i.e., the average age at death shifts from young ages toward older ages. The epidemiological transition was observed in many industrialized countries until the 1970s

1. Marie-Pier Bergeron-Boucher, Max-Planck Odense Center on the Biodemography of Aging, University of Southern Denmark, Department of Epidemiology, Biostatistics and Biodemography, J.B. Winsløws Vej 9, 5000 Odense C, Denmark, email: mpbergeron@health.sdu.dk; co-authors Robert Bourbeau and Jacques Légaré, Department of Demography, Université de Montréal.

(Meslé and Vallin 2000). However, in most of these countries, mortality levels from degenerative and man-made diseases did not stabilize, as suggested by Omran (1971), but a further decline in mortality due to cardiovascular diseases was observed (Meslé and Vallin 2000). These changes in the structure of causes of death have resulted in a further increase in life expectancy for many industrialized countries.

Between 1921 and 2011, life expectancy at birth in Canada increased from 56.0 to 79.5 years for males and from 58.2 to 83.7 for females (CHMD 2015). Progress in life expectancy at birth was initially driven by a reduction in infant mortality. However, since the middle of the twentieth century, most of this increase has resulted from mortality reduction at higher ages (Decady and Greenberg 2014). These changes have enabled life expectancy at birth to keep rising, but have also caused life expectancy at higher ages to increase over time (Vallin and Meslé 2010). For example, between 1950 and 2011, life expectancy at age 65 in Canada increased from 13.3 to 18.9 years for males and from 15.0 to 21.8 for females (CHMD 2015).

Reductions in cause-specific mortality occurring at higher ages, as in the case of cardiovascular diseases, have been the main drivers of life expectancy increase in Canada since the 1950s (Decady and Greenberg 2014). Death rates from heart and cerebrovascular diseases decreased by 37.4 and 39.2 per cent, respectively, between 1981 and 2007 (Milan 2011). However, mortality due to other causes of death rose over the same period. Mortality due to cancers increased until the middle of the 1990s but has since begun to decrease. An increase in mortality due to diabetes and Alzheimer's disease has also been observed (Milan 2011). However, less is known about how these changes have influenced Canadian life expectancy trends, and how the cause of death structure has changed since the cardiovascular revolution.

Decady and Greenberg (2014) state that most of the progress in life expectancy in Canada in the past 90 years has come from a decline in premature deaths. *Premature deaths* are defined by the authors as deaths occurring among individuals aged less than 75 years. Other indicators of premature deaths are also often used, such as potential years of life lost (PYLL), which also uses a cutoff age (e.g., 70 or 75) under which a person's death will be considered premature (OECD 2016). The concept of premature deaths refers to deaths occurring at younger ages, which are before their time and could, a priori, have been preventable. Premature deaths are often contrasted with *senescent deaths*, which tend to result from a more natural aging process and to occur at older ages (Brody and Schneider 1986; Horiuchi 2007). Senescent mortality here refers to the increase over age in the risk of dying, resulting from gradual physical deterioration with age (Bongaarts 2005; Vaupel 2010).

Instead of defining premature deaths as deaths occurring before a fixed age, some authors have preferred to look at the age structure of cause-specific mortality (Brody and Schneider 1986; Horiuchi 2007; Horiuchi and Wilmoth 1997). Causes of death have their own specific age structure: some tend to occur during childhood, others around middle age, and some at older ages. The age structure of causes of death reveals information about physiological and biological changes with age, but also about individual differences in longevity (Horiuchi 2007). Age-variation analyses of causes of death have also been useful in understanding how diseases develop: some diseases have been shown to develop prematurely and selectively among certain individuals, whereas others appear at old or very old ages and may reflect a senescence process (Brody and Schneider 1986; Horiuchi 2007).

Our research aims to analyze changes in causes of death at age 65 and older in Canada since 1979 via three objectives. We first look at changes in life expectancy trends at ages 65 and 85 years. Given these changes in life expectancy trends, we determine different historical time periods and calculate, for each of these periods, the age- and cause-specific contribution to life expectancy increase, in order to evaluate emergent and declining causes of death and their impact on general mortality. Finally, we establish age variation patterns of causes of death, to distinguish between premature and senescent causes of death in Canada.

Data

Data on causes of death are taken from the Canadian Vital Statistics from 1979 to 1999, provided by Statistics Canada, which publishes causes of death data after collating and grouping information from the provinces and territories where data are collected. Causes of death data from 2000 to 2011 were extracted from CANSIM tables published by Statistics Canada (Statistics Canada 2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g). The

data are available for 5-year age groups, with the last available age-group being 90 years and older (90+). The data on population and exposure to risk were extracted from the Canadian Human Mortality Database (CHMD), which provides historical data based on population surveys and censuses in Canada (CHMD 2015).

Deaths are classified according to the International Classification of Diseases (ICD), which has changed over the years. In the period from 1979 to 2011, the ICD was changed only once in Canada, in 2000, from ICD-9 (WHO 1977) to ICD-10 (WHO 2010). However, this change was one of the most major that the ICD has undergone, leading to important discontinuities in some causes of death time-series, e.g., those for pneumonia and Alzheimer's disease (Geran et al. 2005; Meslé and Vallin 2008). To avoid such discontinuities, data on causes of death before 1999 and after 2000 are analyzed as two different series.

The ten main causes of death at age 65 and older in Canada are selected (Table 1). In 2011, these causes represented 75.7 per cent of all deaths at age 65 and older. Neoplasms are divided into five subgroups of causes: cancers of the lung, bronchitis and trachea (LBT), colon and rectum (CR), breast, prostate, and other cancers.

Table 1. ICD-9 and ICD-10 codes for the selected causes of death and their proportions (%) at age 65 years and older in 1979, 1995, and 2011, Canada

Causes of death	ICD-9	ICD-10	% 1979	% 1995	% 2011
1 Malignant neoplasms	140–208	C00–C97	21.4	25.4	27.8
Lung, bronchitis, and trachea cancer	162	C33–C34	4.4	6.5	7.4
Colon and rectum cancer	153–154	C18–C21	3.2	2.9	3.3
Breast cancer	174–175	C50	1.4	1.8	1.6
Prostate cancer	185	C61	1.7	2.2	1.8
Other cancers	140–152; 155–161; 163–173; 176–184; 186–208	C00–C17; C22– C32; C35–C49; C51–C60; C62–C97	10.7	12.0	13.7
2 Heart diseases	390–398; 402; 404; 410–429	I00–I09; I11; I13; I20–I51	40.2	30.6	21.3
3 Cerebrovascular diseases	430–438	I60–I69	11.8	8.8	6.3
4 Chronic lower respiratory diseases	490–94; 496	J40–J47	3.4	5.2	5.4
5 Diabetes mellitus	250	E10–E14	2.0	2.8	3.1
6 Alzheimer's disease	331.0	G30	0.0	1.7	3.3
7 Influenza and pneumonia	480–487	J10–J18	3.3	4.3	2.7
8 Accidents	E800–E869; E880–E929	V01–X59; Y85–Y86	2.4	2.1	3.2
9 Nephritis, nephrotic syndrome and nephropathy	580–589	N00–N07; N17–N19; N25–27	1.0	1.4	1.6
10 Parkinson's diseases	332	G20–G21	0.3	0.7	1.0
11 Other diseases	–	–	14.2	17.0	24.3

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), and Geran et al. (2005).

Methods

Changes in life expectancy at ages 65 and 85 years: Segmented regression

To evaluate changes in life expectancy trends, a segmented regression methodology is used. The segmented regression method was introduced by Muggeo (2003) and applied in the life expectancy context by Camarda et al. (2012). This method aims to find the slopes and breakpoints of a piecewise linear regression, formed of two or more linear segments connected at unknown values (Camarda et al. 2012). Ouellette et al. (2014) also used this method to find discontinuities in general, and in age-and cause-specific mortality trends. When there is only one breakpoint, the method corresponds to:

$$e_x(t) = \alpha_x + \beta^1_x\, t + \beta^2_x\, (t - \psi_x)^+,$$

where $e_x(t)$ is the life expectancy at age x and year t, β^1_x is the slope of the first segment, β^2_x is the difference between the slope of the first and the second segment, and ψ_x is the breakpoint. Then, $(t-\psi_x)^+ = (t-\psi_x)^* I(t>\psi_x)$, where I is an indicator function equal to 1 when the year t is higher than the breakpoint year $(t>\psi_x)$ (Camarda et al. 2012).

Contribution to changes in life expectancies

Evaluation of age and cause of death contributions to changes in life expectancy is done using standard decomposition methods (Arriaga 1984; Preston et al. 2001). The contributions are calculated for different periods, based on the breakpoints resulting from the segmented regression.

Age variation

The age profile of each cause of death is based on two indicators. The first is the proportion between age x and $x+n$ of cause i, ${}_np^i_x(t)$, calculated as

$$ {}_np^i_x(t) = {}_nD^i_x(t) / {}_nD_x(t), $$

where ${}_nD_x(t)$ and ${}_nD^i_x(t)$ are the total and cause i number of observed deaths between ages x and $x+n$ for year t, respectively. This indicator shows the relative importance of causes of death at specific ages (Horiuchi 2007).

The second indicator is the life table aging rate (LAR). The LAR is interpreted as the slope of an exponential curve. Constant LARs through ages thus represent an exponential increase with age of the death rates (Horiuchi and Wilmoth 1997). This indicator allows us to see whether and how the increase of the death rates for a specific cause is accelerating or decelerating with age. The LAR for a cause of death i at age x is estimated as

$$ LAR^i_x(t) = [ln({}_nM^i_x(t)) - ln({}_nM^i_{x-n}(t))] / n, $$

where ${}_nM^i_x(t)$ is the death rate of cause i between ages x and $x+n$ for year t.

These indicators thus represent two different aspects of causes of death age variation: the proportions show the relative importance of causes of death at each age and the LARs show if the risk of dying from a specific cause of death accelerates or decelerates with age, and whether it increases (positive LARs) or decreases (negative LARs). Both indicators are compared to validate each other's results.

Results

Breakpoints in life expectancy at ages 65 and 85 years

Life expectancy at age 65 and 85 years (e_{65} and e_{85}) has undergone a major increase in Canada over the twentieth century. However, Figure 1 shows that this progress has not been linear and a number of different breaks in the trends can be observed. Three breakpoints in life expectancy at age 65 years for females were detected since 1921 (1943, 1987, and 1998), but only two for males (1972 and 1998). For trends in life expectancy at age 85, there were three breakpoints for each sex: 1949, 1989, and 1998 for females and 1947, 1984, and 1999 for males.

Before the first breakpoint, in the mid-1940s, progress in life expectancy at age 65 and 85 years was modest. Based on the maximum life expectancy observed each year, Vallin and Meslé (2010) determined that the increase in life expectancy after the first breakpoint was the result of the cardiovascular revolution. Reductions in cardiovascular mortality enabled life expectancy at higher ages to increase. In Canada, these progresses only became evident on e_{65} after 1972 for males.

In the middle of the 1980s, the rate of increase in e_{65} for females and e_{85} for both sexes decelerated. For the first time, life expectancy for males began to increase faster than for females. Since 1998–99, life expectancies for both sexes and at both ages have increased at a faster pace than in any other observed period. In the next section, we present the age and cause contributions to these changes in life expectancy.

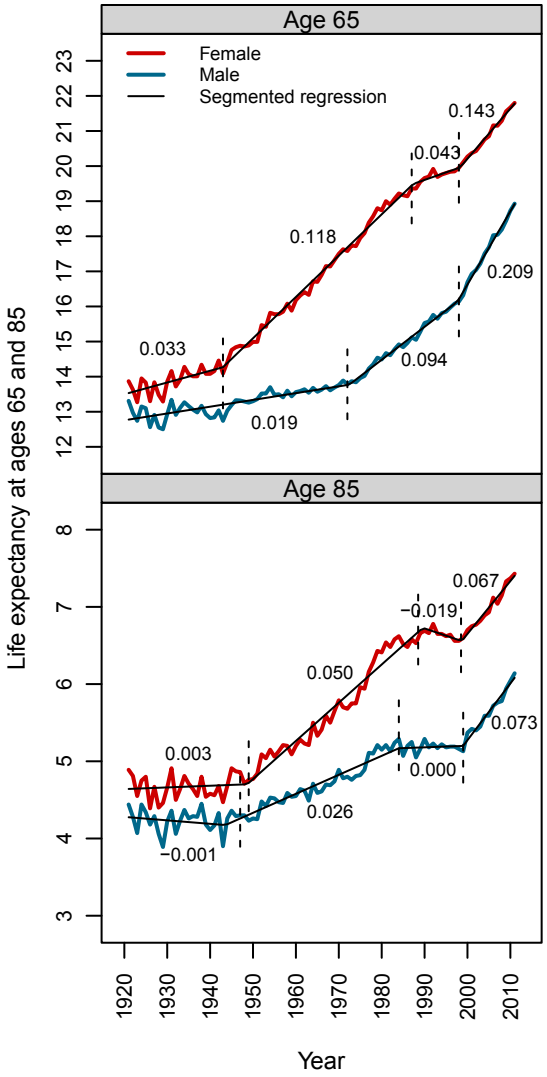


Figure 1. Canadian life expectancy at ages 65 and 85 years by sex, associated segmented regressions, and respective slopes, 1921–2011.

Sources: CHMD (2015) and authors' own calculations.

Age and cause of death contributions to changes in life expectancy

Based on the breakpoints identified in the previous section and the causes of death data constraints, we determined three historical periods of change: 1979–85, 1985–99, and 2000–11. In these three periods, life expectancy at age 65 increased, respectively, by 0.32, 0.88, and 1.59 years for females and by 0.29, 1.46, and 2.31 years for males.

Age contributions

Figure 2 shows the annualized age- and cause-specific contribution to changes in e_{65}. For the period 1979–85, 88 per cent (for males) and 51 per cent (for females) of the changes in life expectancy at age 65 were the result of lower mortality between ages 65 and 74 years. A large part of the increase in life expectancy therefore came about because of a decline in premature mortality, as defined by Decady and Greenberg (2014). However, these proportions decreased to 44 and 31 per cent, respectively, in 2000–11. Changes taking place at older ages are thus playing a greater role in the increase in life expectancy in recent years.

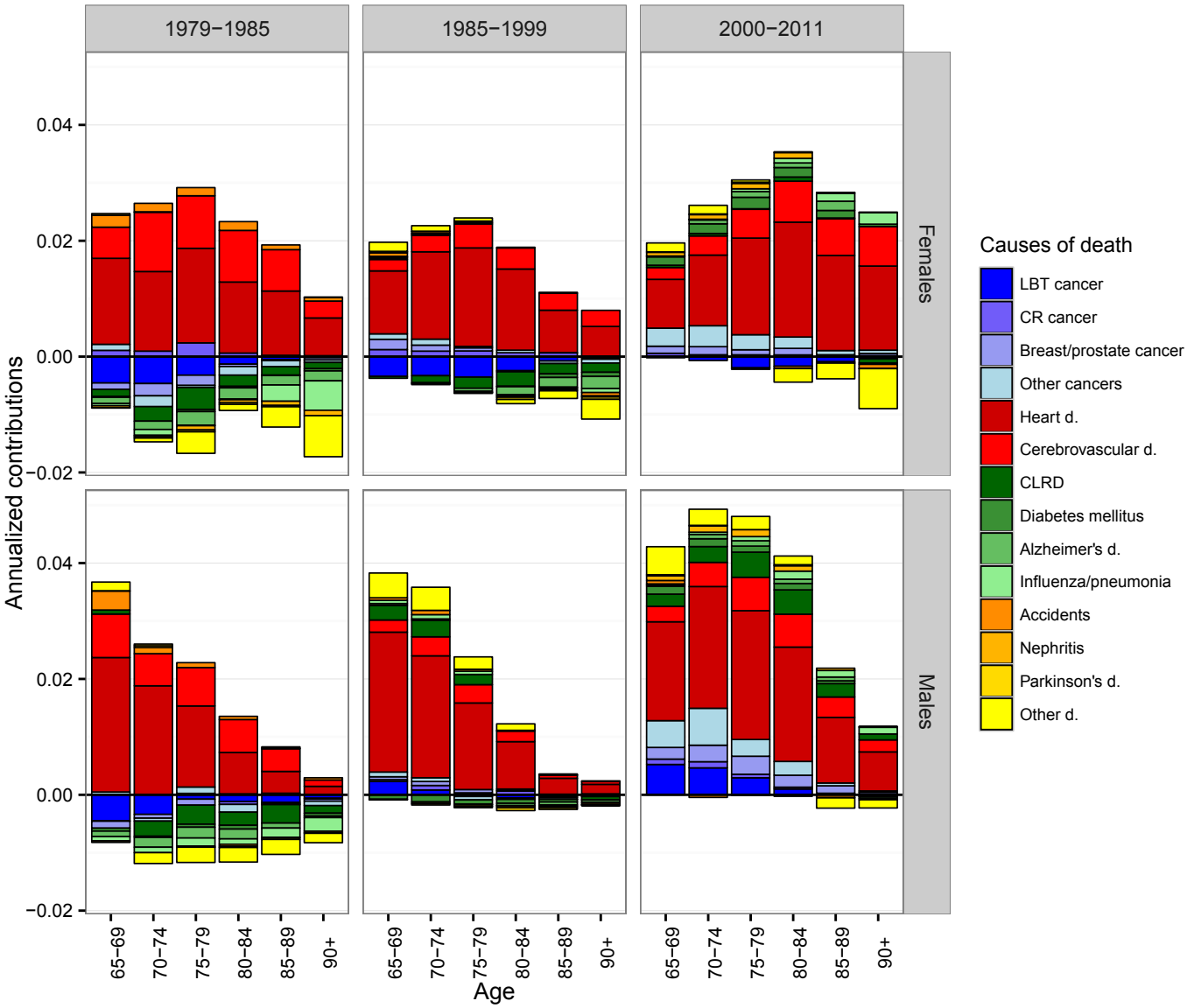


Figure 2. Annualized age and causes of death contribution to the increase in Canadian life expectancy at age 65 by sex, period 1979–85, 1985–99, and 2000–11.

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

Cause contributions

Regardless of the period, the increase in life expectancy at age 65 is mainly the result of a decrease in mortality due to cardiovascular diseases (CVD), which include heart and cerebrovascular diseases (Figure 2). For the period 1979–85, most of the other selected causes contributed to reduce e_{65}, especially cancers, chronic lower respiratory diseases, Alzheimer's disease, and influenza and pneumonia.

Over the period 1985–99, the main causes of death were now contributing positively, or made only very small negative contributions, to the e_{65} increase for males. However, the increase in life expectancy for females continued to be slowed down by lung, bronchitis, and trachea (LBT) cancer, chronic lower respiratory diseases (CLRD), Alzheimer's disease, and the residual group of causes. Females experienced a loss of 0.43 years in life expectancy due to these last four causes, while males gained 0.27 years as a result of changes in those same causes (see Appendix A). These differences explain why female e_{65} increased less than that of males during this period, together with a greater contribution from heart diseases for males. Smaller gains in cerebrovascular diseases were also observed for both sexes over this same period.

In the period 2000–11, e_{65} for both sexes increased at a faster pace than in the previous periods. Most causes of death were now showing positive or null gains in terms of life expectancy, with the exception of LBT cancer for females, and accidents and "other causes" at older ages for both sexes. Appendix B shows that the death rates at 65 years and older by cause of death, with exception of the three previously mentioned causes, have been decreasing or stayed approximately constant since 2000.

Given the increasing importance of older age-groups in life expectancy changes, we also looked at the cause-specific contributions to changes in e_{85}. Appendix A shows similar results for the decomposition of life expectancy at age 85 and age 65 years. The deceleration in the e_{85} trend observed between 1985 and 1999 (Figure 1) is partially explained by a lower annualized gain in terms of cerebrovascular diseases compared to the previous period, and by a negative contribution for most of the other causes.

Age variation profile of causes of death

As suggested by Horiuchi (2007), we use the age of 85 years to separate "younger-old age" from "oldest-old age" mortality, when looking at the proportions by age. We distinguish between causes of death that are more prevalent before (downward trends) or after (upward trends) age 85 (see Figures C1 and C2 in Appendix C).

The age variation of causes of death using LARs shows that the increase in some cause-specific death rates decelerates (decreasing LARs) with age, while it accelerates (increasing LARs) for other causes (see Figures C3 and C4 in Appendix C). The LARs even become negative after a certain age for some causes, meaning that the death rates are decreasing with age. Based on this indicator, we separate cause of death LARs that decrease with age from those which stay constant or increase (see Appendix D for more details on the classification criteria).

Table 2. Causes of death classified according to their age variation using proportions and life table aging rates (LARs), Canada, 1979–2011

	Lower proportions before age 85 (Upward trends)	Higher proportions before age 85 (Downward trends)
Decreasing LAR	Cerebrovascular diseases (Male) Alzheimer's disease	Lung, bronchitis and trachea cancer Colon and rectum cancer (Male) Prostate cancer Other cancers Chronic lower respiratory diseases Diabetes mellitus Parkinson's disease
Constant or increasing LAR	Heart diseases Cerebrovascular diseases (Female) Influenza and pneumonia Accidents Nephritis, nephrotic syndrome and nephropathy Other diseases	Colon and rectum cancer (Female) Breast cancer

According to these definitions, two main patterns of causes can be distinguished, as shown in Table 2. The first pattern shows higher proportions before age 85 years and decreasing LARs over age. These causes have a higher relative importance at youngest-old ages than at oldest-old ages, and the risk of dying from these causes decelerates or decreases with age. The decrease in the proportions and LARs with age could come from (1) a tendency of these causes to develop early in life; and/or (2) other causes of death tending to take over mortality at higher ages. This pattern is consistent with Horiuchi's (2007) findings for the US White population, where this category represents causes of death that tend to develop prematurely and selectively among certain individuals. They are also often associated with certain risk factors, such as smoking, diet, heredity, etc.

The second pattern shows higher proportions after age 85 years, and increasing or constant LARs over age. These causes have a higher relative importance at the oldest-old ages, and the risk of dying from them increases and often accelerates with age. Causes included in this group thus occur at older ages, and tend to be more associated with a process of senescence. As Horiuchi and Wilmoth (1997) also pointed out, competing risk among

causes of death is more important after a certain age, so that as a result of senescence, people become more susceptible to death from causes that develop more rapidly than from other causes which take longer to result in death. An example for males is given in Figure 3 for LBT cancer and for accidents, representing the two age patterns, respectively (see Appendix C for the other causes).

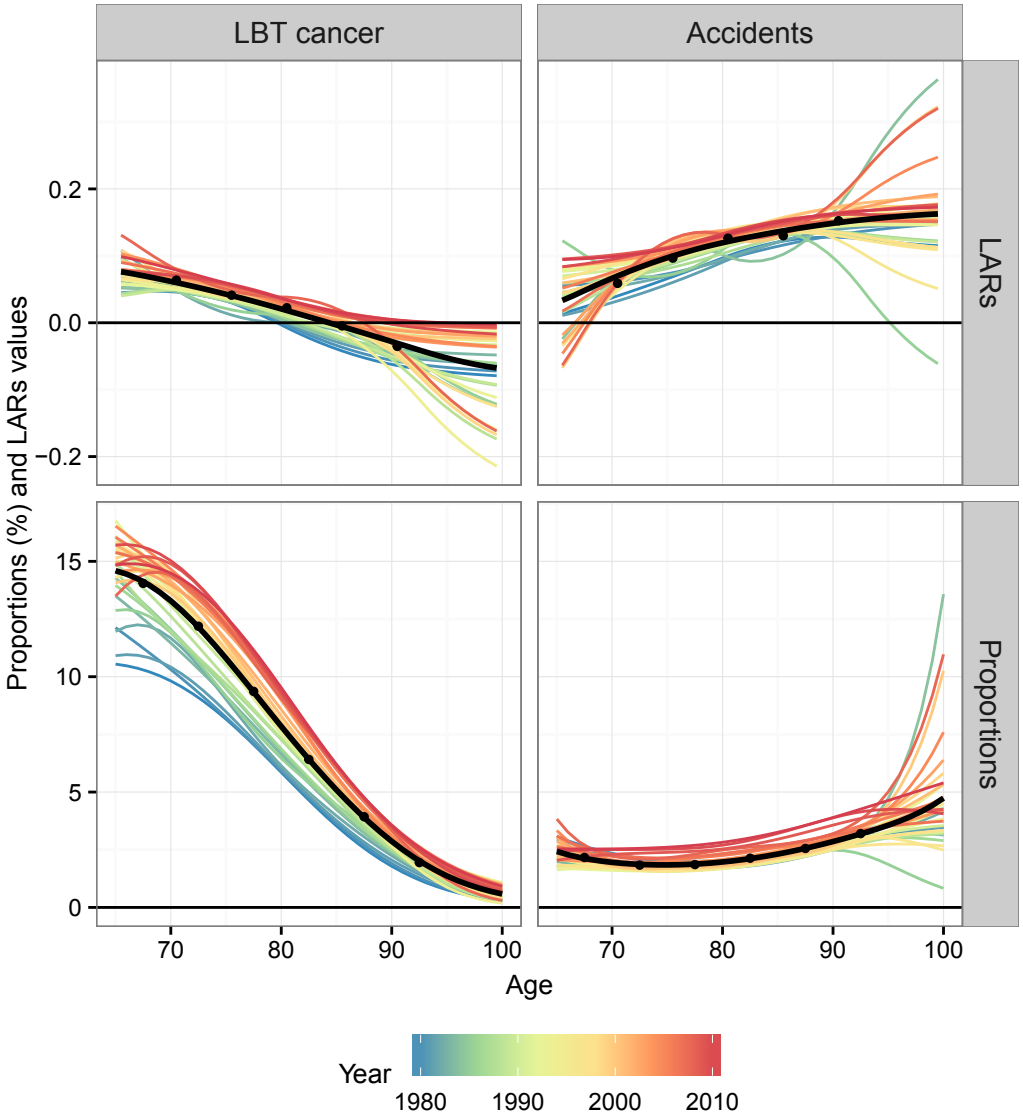


Figure 3. Age variation for Lung, bronchial, and trachea cancer and Accidents using proportions and life table aging rates (LARs), observed (dot) and smoothed (line), Males, Canada, 1979–2011 and age-specific average over time (in black).

Note: The trends are smoothed using a P-spline smoothing procedure for Poisson death counts (Camarda 2012).

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

There are, however, some exceptions to these two patterns. Alzheimer's and cerebrovascular diseases (male) have more importance at higher ages, although their LARs decrease with age. Breast cancer and CR cancer (female) have more importance at younger ages; however, their LARs do not decrease with age. These results may be due to inadequate classification criteria or specification of the cause of death, but more research would be needed to explain them fully.

Using the classification of deaths discussed earlier, in 1979 premature deaths represented about 21 and 30 per cent of total deaths at age 65 years and older for females and males, respectively, compared with 31 and

41 per cent in 2011, including CR cancer deaths as premature deaths for females. This change in proportions may be a result of the decline in deaths due to cardiovascular diseases, classified as senescent, and the increase in cancer deaths observed up to the beginning of the 21st century. Between 1979 and 1985, premature deaths contributed to a loss of 0.17 years of life expectancy at age 65 for females and 0.18 years for males. However, between 2000 and 2011 they contributed to increasing e_{65} by 0.20 and 0.75 years for females and males, respectively, representing 13 and 32 per cent of the total increase in that period.

These results point to some particular characteristics of the Canadian population and of the health conditions in Canada. The main drivers of the loss of premature years in life expectancy at age 65 years in the period 1979–85 were increases in mortality from lung, bronchitis, and trachea cancers and from chronic lower respiratory diseases (Appendix A). These two last causes of death are highly influenced by tobacco consumption (CDC 2016). A decline in smoking among the Canadian population could have enabled more of these deaths to be prevented in more recent years, especially for males (CCSSC 2010).

However, changes in premature deaths have not been the main drivers of mortality changes since 1979. Gains in life expectancy at age 65 years have mainly resulted from a decrease in senescent deaths, including heart diseases. Vaupel (2010) has argued that although senescence is inevitable for human populations, it can nonetheless be postponed, and this has been made possible by ongoing improvements in health care. It has led to declines in senescence-associated causes of death, with the result that mortality decrease is observed at higher and higher ages.

Discussion

In Canada, deaths classified as premature (first age pattern) represented about 26 per cent of all deaths at age 65 and older in 1979 and about 36 per cent in 2011, for both sexes combined. Most of these deaths are considered to be less senescence-related, and so to be untimely and potentially preventable. An increase in premature deaths might reflect a greater burden of certain risk factors, e.g., smoking (lung, bronchitis and trachea cancer, and chronic lower respiratory diseases) and diet (diabetes). This information is potentially useful for prevention and treatment programs. And in fact, a reduction in tobacco consumption in Canada has resulted in decreased mortality due to cancer of the lung, trachea, and bronchitis and to chronic lower respiratory diseases among males. Female mortality due to these causes has, however, increased since 1979, because of a time lag in female patterns of tobacco consumption relative to males (CCSSC 2010).

But the main influence on changes in life expectancies at age 65 and 85 years since 1979 has been a decrease in senescent deaths. As one important explanatory variable for these causes is age, the increasing number of people reaching higher ages should have enhanced mortality pressure due to causes that occur at very advanced ages. However, it has been possible to successfully reduce death rates for these diseases, as shown by the major decline in mortality from cardiovascular diseases. Still, even if the senescence process can be postponed (Vaupel 2010), it cannot be avoided, and all individuals eventually become vulnerable to multiple diseases (Horiuchi 2007). The decrease in the main causes of death and the increase in age at death might lead toward a greater diversification of causes of death at old ages.

The decrease in mortality from the main causes of death, with no apparent mortality increase in other specific-causes since 2000 (Appendix B), except for LBT cancer for females and accidents for both sexes, is likely to lead to a greater diversification of causes of death. Mortality due to predominant causes is decreasing, whereas the other causes of death are tending to increase in relative importance—without, however, becoming predominant causes (Table 1). This diversification is also reflected in an increase in mortality and proportions for the residual group of causes. Causes included in this group are very diverse, and none of them accounted for more than 2 per cent of the deaths in 2011. Increases in life expectancy and advancing age come with an increasing vulnerability to multiple diseases, reflecting a "long-term accumulation of unrepaired damages, leading to simultaneous deterioration of many physiological functions" (Horiuchi 2007: 231). The decrease in mortality due to the main causes of death, together with the increase in life expectancy, means that a diversification of causes of death at old ages is likely to take place.

The use of age 85 as a threshold between youngest-old and oldest-old ages may be questionable. For example, Kannisto (1994) defined the oldest-old as people aged 80 years and higher. Using this threshold instead would not have changed much our results, with the exception of Parkinson's disease. The 80- and 85-year thresholds are often used, but do not reflect the dynamic process of aging itself (Tomassini 2005). Other, more dynamic indicators, such as modal age at death, might be more appropriate.

Another potential limitation to this study is that the selected period of 1979-85 may not be the most relevant in terms of comparing a previous period with the deceleration in e_{65} and e_{85} of 1985–99, because the period is short and may already embody some features of the cause of death changes of the subsequent period.

It is also important to note that there are underlying problems with cause of death data, which may bias the results ((Désesquelles and Meslé 2004; Geran et al. 2005; Meslé and Vallin 2008). After a certain age, competing risks between different causes of death grow, as the chances of dying from various diseases increase. This may lead to misreporting on death certificates. Some causes of death seem to be more often declared as the primary cause of death, while other causes will be considered as secondary causes (Désesquelles and Meslé 2004). This is the case with diabetes, for example. Diabetes seems to contribute frequently to the morbid process among the elderly, but is rarely stated as the primary cause of death (Désesquelles and Meslé 2004).

The impact of the ICD changes is also not well understood, and may have led to greater discontinuities than expected (Geran et al. 2005). As shown in Appendix B, there are important discontinuities in the time trends of influenza and pneumonia. The drastic decrease in the death rates of influenza and pneumonia is related to an important change in the rule to classify pneumonia: in ICD-10, pneumonia is now accepted as a complication of any disease (WHO 2010). Pneumonia is thus considered more often as a secondary cause of death than a primary one with the ICD-10 than with the ICD-9. This has led to an important decrease in the number of deaths due to pneumonia and to a corresponding increase in the related primary conditions (Geran et al. 2005; Meslé and Vallin 2008).

Another underlying problem with the causes of death series is changes in the perception and knowledge of a disease. As an example, Alzheimer's disease was considered as a rare disease until the end of the 1970s, but a better understanding of the disease and its delineation from dementia increased the number of diagnoses and deaths classified as Alzheimer's since (Boller and Forbes 1998). Progress in perception and knowledge of this disease is still ongoing. An increase in the death rates of Alzheimer's disease from ICD-9 to ICD-10 is observed as dementia in and due to Alzheimer, classify as Senile and pre-senile organic psychotic conditions in ICD-9, are classify to Alzheimer's disease in ICD-10 (Geran et al. 2005; Meslé and Vallin 2008).

Although we study causes of death trends classified with ICD-9 and ICD-10 as two different series in this article, these discontinuities in the trends and changes in disease knowledge could affect the results. However, despite these limitations, cause of death data still inform us regarding changes in cause of death structure in Canada.

Conclusion

Changes in the cause of death structure are ongoing in Canada, and the main causes of death among people aged 65 years and older have been decreasing in importance since 2000. A decrease in premature deaths is observed, but the main gains in life expectancy at age 65 and 85 years come from a decline in senescence-related causes of death. These changes have led to a major increase in life expectancy at age 65 and 85 years. Furthermore, as a result of the decline in the main causes of death and the corresponding increase in life expectancy, a greater diversification of causes of death may take place at older ages.

Acknowledgments

This paper is based on results from Marie-Pier Bergeron-Boucher's Master's thesis research, carried out under the supervision of Robert Bourbeau and Jacques Légaré, of the Department of Demography at the Université de Montréal. The authors are grateful to Vladimir Canudas-Romo, of the Max-Planck Odense Center on the Biodemography of Aging, for his useful comments.

References

Arriaga, E.E. 1984. Measuring and explaining the change in life expectancies. *Demography* 21(1):83–96.

Boller, F., and M.M. Forbes. 1998. History of dementia and dementia in history: An overview. *Journal of Neurological Sciences* 158:125–33.

Bongaarts, J. 2005. Long-range trends in adult mortality: Models and projection methods. *Demography* 42(1):23–49.

Brody, J.A., and E.L. Schneider. 1986. Diseases and disorders of aging: An hypothesis. *Journal of Chronic Diseases* 39(11):871–76.

Camarda, G., J. Vallin, and F. Meslé. 2012. Identifying the Ruptures Shaping the Segmented Line of Secular Trends in Maximum Life Expectancies. Paper presented at the 2012 European Population Conference: Gender, Policies, and Population, Stockholm, Sweden (June 13–16).

Camarda, G. 2012. MortalitySmooth: An R package for smoothing Poisson counts with P-splines. *Journal of Statistical Software* 50(1):1–24.

CCSSC (Canadian Cancer Society's Steering Committee). 2010. *Canadian Cancer Statistics 2010*. Toronto: Canadian Cancer Society.

CDC (Centers for Disease Control and Prevention). 2016. Health Effects of Cigarette Smoking [Fact sheet]. Retrieved from https://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/effects_cig_smoking/

CHMD. 2015. Canadian Human Mortality Database. Department of Demography, University of Montreal (Canada). Available at www.demo.umontreal.ca/chmd/(data downloaded on March 2015).

Decady, Y., and L. Greenberg. 2014. Ninety years of change in life expectancy. Health at a Glance Report (Cat. no. 82-624-X). Ottawa: Statistics Canada.

Désesquelles, A., and F. Meslé. 2004. Intérêt de l'analyse des causes multiples dans l'étude de la mortalité aux grands âges : l'exemple francais. *Cahier québécois de démographie* 33(1):83–116.

Geran, L., P. Tully, P. Wood, and B. Thomas. 2005. Comparability of ICD-10 and ICD-9 for Mortality Statistics in Canada. Health Statistics Division Report (Cat. No. 84-548-XIE). Ottawa: Statistics Canada.

Horiuchi, S. 2007. Causes of death among the oldest-old: Age-related changes in the cause-of-death distribution, in J.-M. Robine, E. Crimmins, S. Horiuchi, and Z. Yi (eds.), *Human Longevity, Individual Life Duration, and the Growth of the Oldest-Old Population*. International Studies in Population 4. Netherlands: Springer, p. 215–35.

Horiuchi, S., and J.R. Wilmoth .1997. Age patterns of the life table aging rate for major causes of death in Japan, 1951–1990. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 52A*(1): B67–B77.

Kannisto, V. 1994. *Development of Oldest-Old Mortality, 1950–1990: Evidence from 28 Developed Countries*. Monographs on Population Aging 1. Odense (Denmark): Odense University Press.

Meslé, F., and J. Vallin. 2000. Transition sanitaire : tendances et perspectives. *Médecine/Science* 16(11):1161–71.

Meslé, F., and J. Vallin. 2008. The effect of ICD-10 on continuity in cause-of-death statistics: The example of France. *Population* (English edn) 63(2):347–59.

Milan, A. 2011. Mortality: Causes of Death, 2007. Report on the Demographic Situation in Canada (Component of Cat. No. 91-209-X). Ottawa: Statistics Canada.

Muggeo, V.M.R. 2003. Estimating regression models with unknown break-points. *Statistics in Medicine* 22(19):3055–71.

OECD. 2016. Potential Years of Life Lost (Indicator). doi: 10.1787/193a2829-en (Accessed on 2 December 2016)

Oeppen, J., and J.W. Vaupel. 2002. Broken limits to life expectancy. *Science* 296(5570):1029–31.

Omran, A.R. 1971. The epidemiologic transition: A theory of the epidemiology of population change. *The Milbank Memorial Fund Quarterly* 49(4):509–38.

Ouellette, N., M. Barbieri, and J.R. Wilmoth. 2014. Period-based mortality change: Turning points in trends since 1950. *Population and Development Review* 40(1):77–106.

Preston, S., P. Heuveline, and M. Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishing.

Statistics Canada. 2015a. Table 102-0522 – Deaths, by cause, Chapter II: Neoplasms (C00 to D48), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015b. Table 102-0524 – Deaths, by cause, Chapter IV: Endocrine, nutritional and metabolic diseases (E00-E90), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015c. Table 102-0526 – Deaths, by cause, Chapter VI: Diseases of the nervous system (G00-G99), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015d. Table 102-0529 – Deaths, by cause, Chapter IX: Diseases of the circulatory system (I00 to I99), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015e. Table 102-0530 – Deaths, by cause, Chapter X: Diseases of the respiratory system (J00 to J99), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015f. Table 102-0534 – Deaths, by cause, Chapter XIV: Diseases of the genitourinary system (N00-N99), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Statistics Canada. 2015g. Table 102-0540 – Deaths, by cause, Chapter XX: External causes of morbidity and mortality (V01-Y89), age group and sex, Canada, annual (number), 2000 to 2011, CANSIM (database).

Tomassini, C. 2005. The demographic characteristics of the oldest old in the United Kingdom. *Population Trends* 120:15–22.

Vallin, J., and F. Meslé. 2010. Will life expectancy increase indefinitely by three months every year? *Population and Societies* 473:1–4.

Vaupel, J.W. 2010. Biodemography of human ageing. *Nature* 464:536–42.

WHO (World Health Organization). 1977. *Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death*. Ninth rev edn, vol. 1. Geneva: World Health Organization.

———. 2010. *ICD-10. International Statistical Classification of Diseases and Related Health Problems*. Tenth rev edn, vol. 2. Geneva: World Health Organization.

Appendix A. Cause of death contributions to change in life expectancy at age 65 and 85 years

Table A1. Cause of death contributions to the increase in Canadian life expectancy at ages 65 and 85 years, by sex, for the periods 1979–85, 1985–99, and 2000–11

Causes of death	e_{65}						e_{85}					
	Males			Females			Males			Females		
	1979–1985	1985–1999	2000–2011	1979–1985	1985–1999	2000–2011	1979–1985	1985–1999	2000–2011	1979–1985	1985–1999	2000–2011
Total	0.29	1.46	2.31	0.32	0.88	1.59	−0.16	0.08	0.87	0.00	0.03	0.81
Lung, bronchi, and trachea cancer	−0.07	0.03	0.15	−0.09	−0.19	−0.06	−0.04	−0.01	0.00	−0.01	−0.03	−0.03
Colon and rectum cancer	0.00	0.03	0.03	0.03	0.06	0.02	−0.01	0.01	0.00	0.00	0.02	0.00
Breast cancer	–	–	–	−0.04	0.04	0.06	–	–	–	−0.01	0.00	0.01
Prostate cancer	−0.02	0.02	0.13	–	–	–	−0.01	−0.02	0.05	–	–	–
Other cancers	0.00	0.01	0.19	−0.02	0.02	0.14	−0.01	−0.02	0.02	−0.02	−0.03	0.03
Heart diseases	0.41	1.02	1.08	0.45	0.97	0.97	0.12	0.22	0.54	0.22	0.35	0.62
Cerebrovascular diseases	0.18	0.16	0.26	0.26	0.26	0.34	0.11	0.05	0.17	0.13	0.16	0.26
Chronic lower respiratory diseases	−0.07	0.09	0.18	−0.07	−0.13	0.01	−0.10	−0.03	0.10	−0.03	−0.10	−0.01
Diabetes mellitus	−0.01	−0.06	0.06	−0.01	−0.01	0.09	−0.01	−0.03	0.01	−0.01	−0.04	0.03
Alzheimer's disease	−0.04	−0.01	0.03	−0.06	−0.08	0.04	−0.03	−0.04	0.01	−0.04	−0.11	0.03
Influenza and pneumonia	−0.05	0.02	0.05	−0.06	−0.01	0.05	−0.09	0.00	0.07	−0.10	−0.02	0.07
Accidents	0.04	0.02	−0.01	0.05	0.01	−0.02	0.01	−0.02	−0.02	0.02	−0.02	−0.02
Nephritis, nephrotic syndrome, and nephrosis	0.00	−0.02	0.05	−0.02	−0.01	0.04	−0.01	−0.01	0.01	−0.02	−0.01	0.00
Parkinson's disease	0.00	−0.02	0.00	−0.01	−0.02	0.01	−0.01	−0.02	0.00	0.00	−0.02	0.00
Other diseases	−0.06	0.16	0.09	−0.09	−0.03	−0.09	−0.09	0.01	−0.10	−0.13	−0.13	−0.19

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

Appendix B. Changes over time of death rates by cause of death

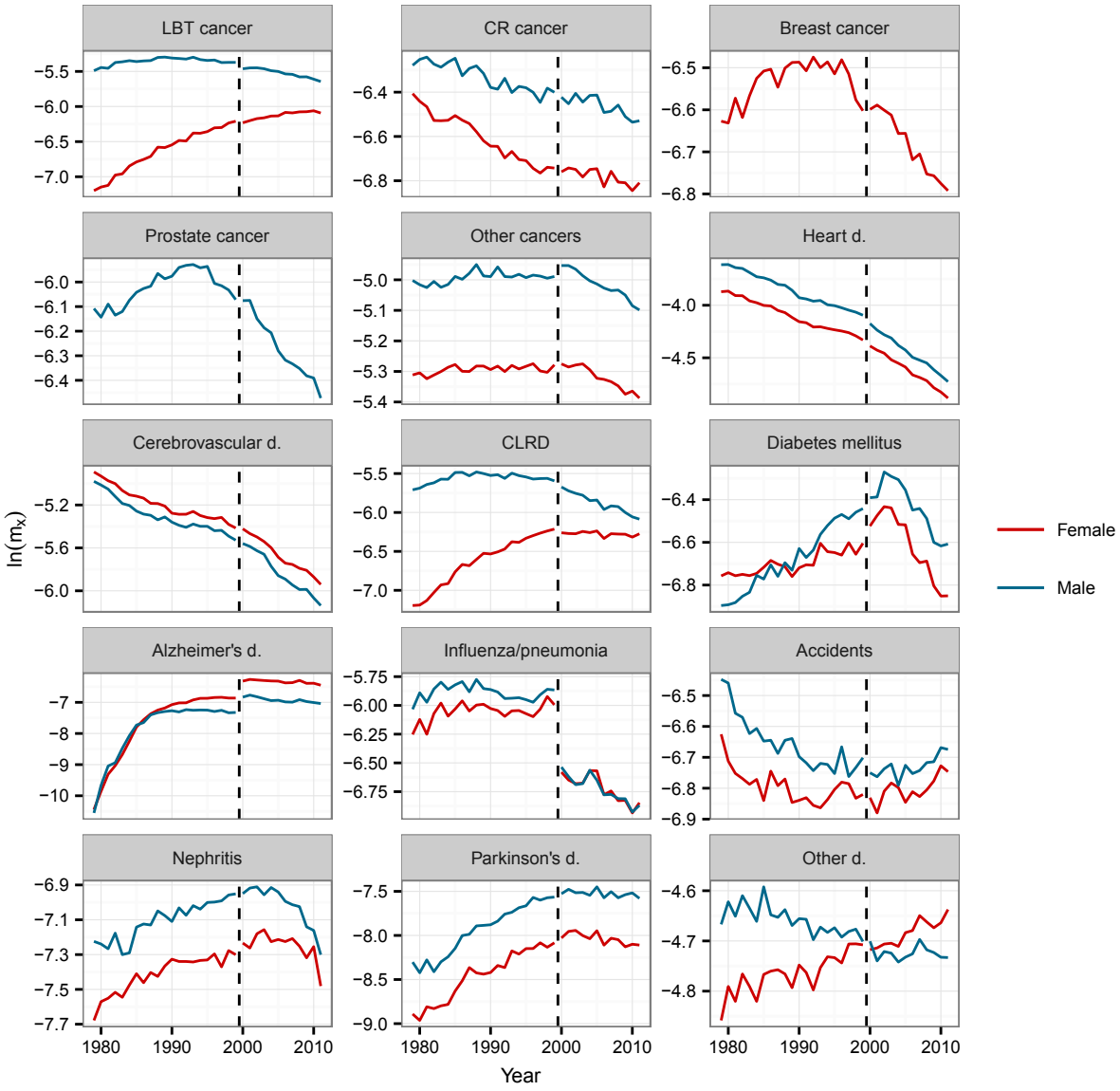


Figure B1. Age-standardized death rates for ages 65 years and older by cause and sex, 1979–2011, Canada.

Note: y-scales might differ from one graph to another to better visualize the time-trends; the standard population is the year 2000.

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

Appendix C. Cause-specific age variation: proportions and life table aging rates (LARs)

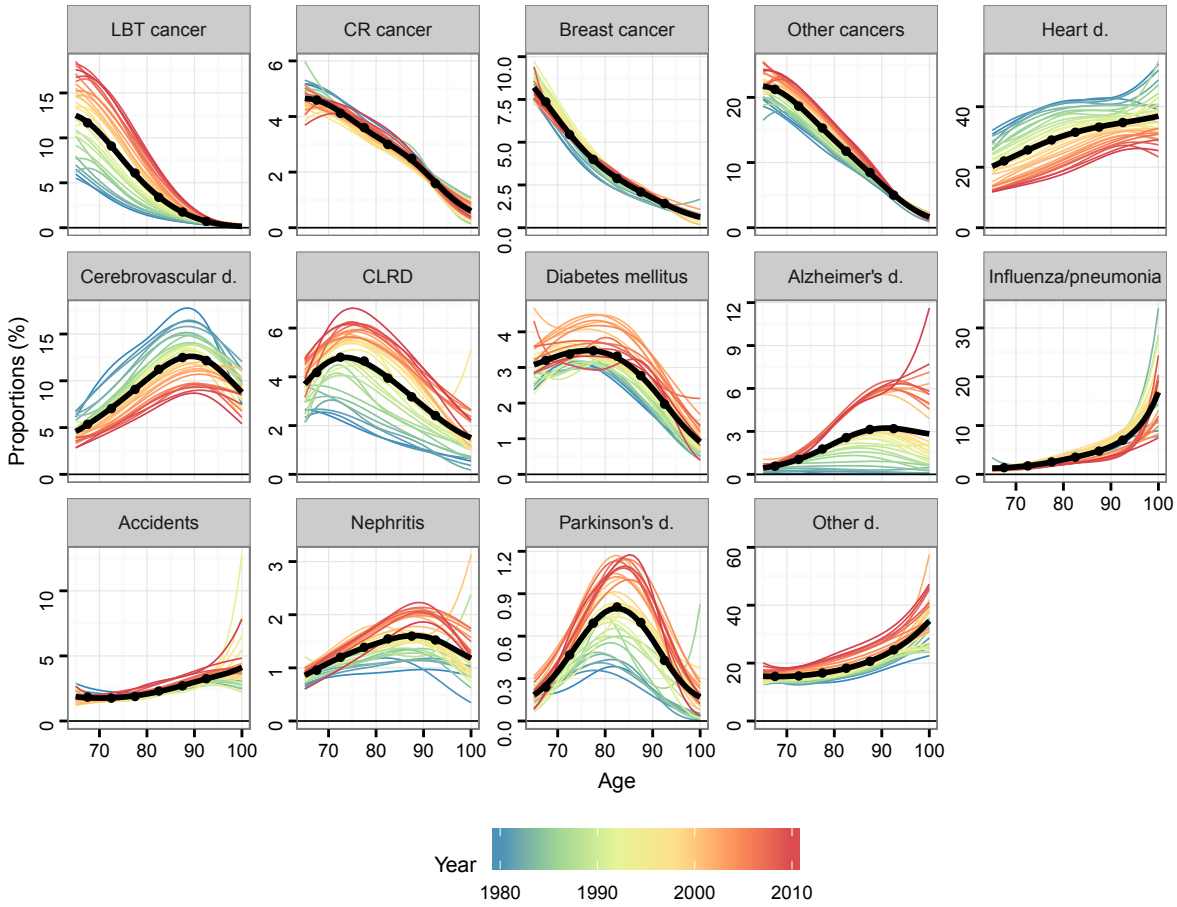


Figure C1. Age variation by cause of death using proportions (expressed in %) over the period 1979–2011 and age-specific average over time (in black), observed (dot) and smoothed (line), Females, Canada.

Note: y-scale might differ from one graph to another to better visualize the age pattern; the trends are smoothed using a P-spline smoothing procedure for Poisson death counts (Camarda 2012).

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

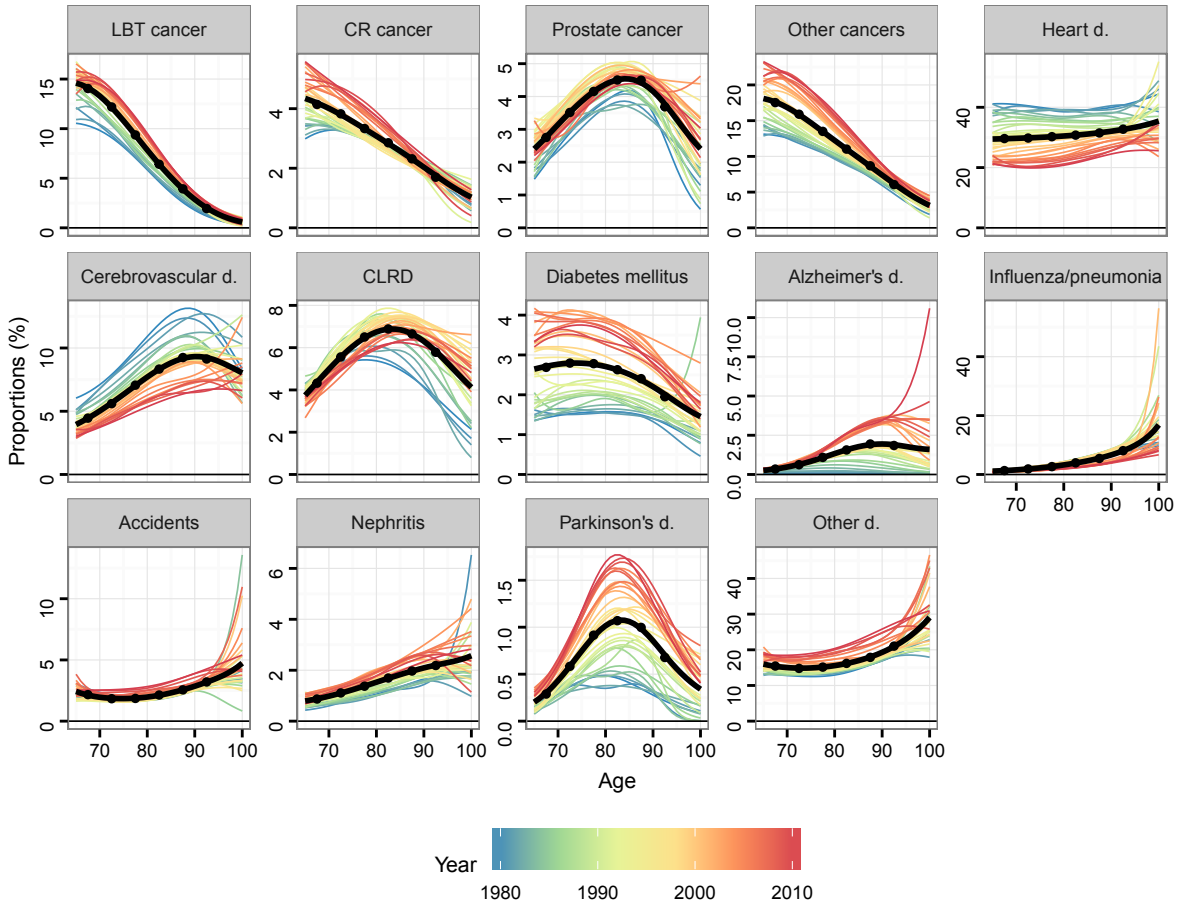


Figure C2. Age variation by cause of death using proportions (expressed in %) over the period 1979–2011 and age-specific average over time (in black), observed (dot) and smoothed (line), Males, Canada.

Note: y-scale might differ from one graph to another to better visualize the age pattern; the trends are smoothed using a P-spline smoothing procedure for Poisson death counts (Camarda 2012).

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015) and authors' own calculations.

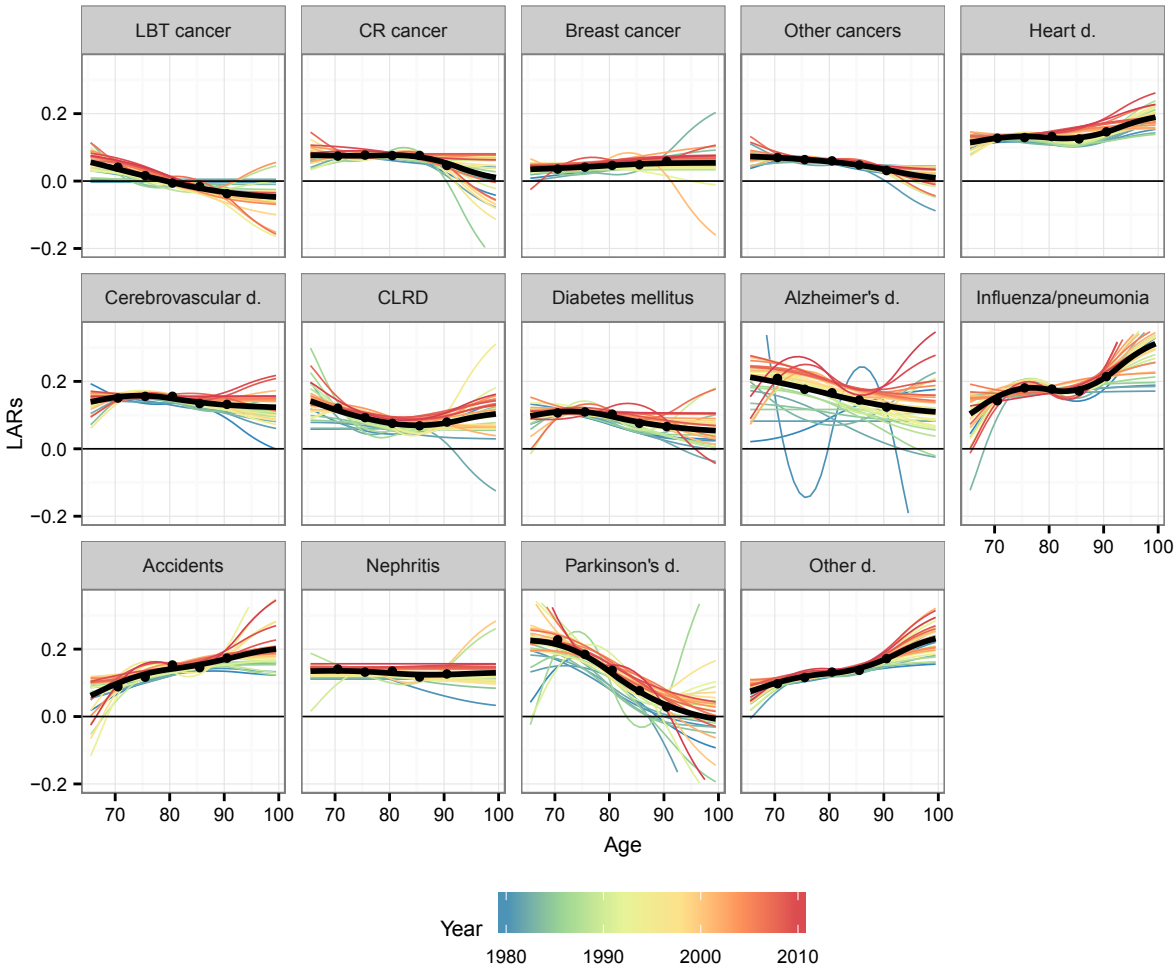


Figure C3. Life table aging rates (LARs) by cause of death over the period 1979–2011 and age-specific average over time (in black), observed (dot) and smoothed (line), Females, Canada.

Note: The trends are smoothed using a P-spline smoothing procedure for Poisson death counts (Camarda 2012).

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015) and authors' own calculations.

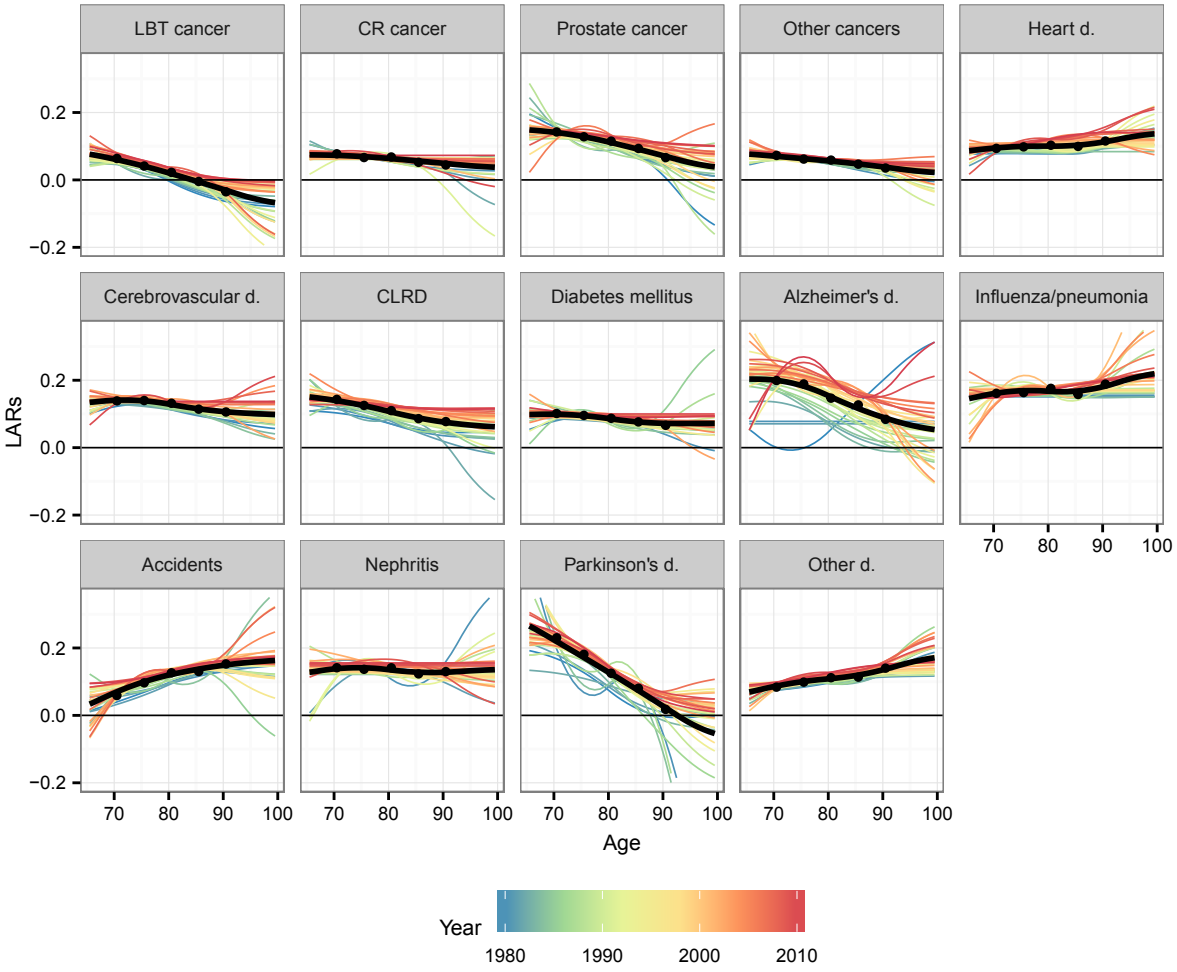


Figure C4. Life table aging rates (LARs) by cause of death over the period 1979–2011 and age-specific average over time (in black), observed (dot) and smoothed (line), Males, Canada.

Note: The trends are smoothed using a P-spline smoothing procedure for Poisson death counts (Camarda 2012).

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

Appendix D. Causes of death classification criteria

In the section "Age variation profile of causes of death," we classify the different causes of death based on two age profile indicators: proportions and LARs. For data availability reasons, we only have 6 age groups for our analysis. To help visualize the age pattern, a P-spline smoothing procedure for Poisson death counts is applied to our data, using the Camarda (2012) MortalitySmooth R Package. The data are interpolated and extrapolated for 1-year age groups for ages from 65 to 100 years.

When looking at the proportions by age we used age 85 to separate *younger-old age* from *oldest-old age* mortality. We aim to see if causes of death are more prevalent before or after age 85. For some causes, as LBT cancer, the visualization of the age pattern could be enough to classify the cause, i.e., if the trends are downward, the proportions are higher at younger-old age, and if the trends are upward, the proportions are higher at oldest-old age. However, the age pattern is sometimes bell-shaped, as in the case of Parkinson's disease. In order to classify the causes between more or less prevalent before age 85 years, we use a ratio of the proportion of death for the 80–84-year age group to the proportion of death for the 85–89-year age group. If the ratio is higher than 1, the peak of the bell-shaped distribution will occur before age 85.

To determine if the LARs are decreasing, increasing, or staying constant over age ranges, we applied a linear regression to the LARs, for the non-smoothed trends. A negative or positive slope parameter indicates if the trend is decreasing or increasing. The significance level of the slope parameter (p-value) indicates if the decrease or increase is significant. If the p-value is higher than 10 per cent, the LARs are considered to be constant over age.

Table D1. Classification criteria for causes of death according to their age variation using proportions and life table aging rates, Canada, average for the period 1979–2011

	Proportions		LARs	
	Proportion of age group 80–84 / Proportion of age group 85–89		Direction and p-value of the slope parameter of a linear regression	
Cause of death	Female	Male	Female	Male
Lung, bronchitis and trachea cancer	1.94	1.63	N***	N***
Colon and rectum cancer	1.20	1.23	N	N**
Breast cancer	1.38	–	P**	–
Prostate cancer	–	1.01	–	N**
Other cancers	1.38	1.27	N**	N**
Heart diseases	0.95	0.98	P	P
Cerebrovascular diseases	0.90	0.90	N	N*
Chronic lower respiratory diseases	1.24	1.04	N	N***
Diabetes mellitus	1.20	1.09	N*	N***
Alzheimer's disease	0.82	0.80	N**	N**
Influenza and pneumonia	0.74	0.73	P	P
Accidents	0.85	0.83	P*	P**
Nephritis, nephrotic syndrome and nephropathy	0.97	0.86	N	N
Parkinson's disease	1.16	1.07	N***	N***
Other diseases	0.88	0.91	P**	P**

Note: Notation for slope parameter direction: N: negative, P: positive. Notation for slope parameter significance level: *** $p<0.001$; ** $p<0.01$; * $p<0.05$; $p<0.1$

Sources: Canadian Vital Statistics 1979–99, Statistics Canada (2015a, 2015b, 2015c, 2015d, 2015e, 2015f, 2015g), CHMD (2015), and authors' own calculations.

Canadian Studies in Population 43(3–4): 234–48.

Type and timing of first union formation in Québec and the rest of Canada: Continuity and change across the 1930–79 birth cohorts

Laura Wright[1]

Abstract

Trends in age at marriage have been well documented, but less is known about age at first union among recent cohorts of Canadians. Using the 2011 GSS, I document changes in the type and timing of first union formation among Canadians born over five decades, and examine how regional differences in partnering behaviours have changed over time. The trend away from entering marriage directly has continued among Canadians born in the 1970s, but Québec-Canada differences have narrowed. The trend towards later marriage has continued, but age at first union has not changed across the five cohorts under study.

Keywords: first union, marriage, cohabitation, Québec, Canada.

Résumé

Les tendances en ce qui a trait à l'âge au mariage ont été bien documentées, mais moins est connu quant à l'âge à la première union chez les dernières cohortes de Canadiens. À partir de données provenant de l'ESG de 2011, je documente les changements au niveau du type et du moment de la formation de premières unions auprès de Canadiens nés au cours de cinq décennies et j'examine en quelle mesure les différences régionales quant aux comportements liés au partenariat ont changé au fil du temps. La tendance à délaisser l'entrée directe au mariage a continué chez les Canadiens nés dans les années 1970, mais l'écart Québec-Canada a diminué. La tendance vers le mariage plus tardif a continué, mais l'âge à la première union n'a pas changé à travers les cinq cohortes à l'étude.

Mots-clés : Première union, mariage, union libre, Québec, Canada.

Introduction

Patterns of union formation have been changing in Canada and other Western nations over the last six decades. Past cohorts of Canadians have been delaying or forgoing marriage, and have increasingly formed nonmarital cohabitations, either as a step in the marriage process or as a union separate from marriage (e.g., Kerr et al. 2006; Le Bourdais and Lapierre-Adamczyk 2004). Differences between Québec and Canada in the type of first union, either marriage or cohabitation, widened between 1960 and 2000 as people in Québec became increasingly less likely to directly marry than other Canadians (Le Bourdais and Lapierre-Adamczyk 2004). Have these diverging trends continued, or has the gap narrowed with the rise of cohabitation in the rest of Canada among the newest birth cohorts? It is also less clear whether recent cohorts of Canadians are delaying all types of unions or whether the rise of cohabitation has offset delays of marriage. Studies of older Canadian cohorts suggest that median age at first partnership has not increased to the same extent as median age at first marriage (Ravanera et al. 2002). In other words, has age at first union increased along with age at first marriage in Canada, or do Canadians born in the 1970s continue to form their first romantic co-residential unions at the same age as their parents' and grandparents' generations?

1. Assistant Professor, Department of Sociology, University of Saskatchewan College of Arts & Science, Room 1019 Arts, 9 Campus Drive, Saskatoon SK S7N 5A5; email: ldw960@mail.usask.ca.

Understanding the changing partnership behaviours of recent cohorts of young Canadians is important because these changes are part of a much larger and significant transformation of family behaviours (Lesthaeghe 1995). There are also widespread institutional and individual implications for changes in partnership behaviours, including delayed and lower fertility, changes in union stability and the family contexts in which children are reared, the length of time spent as a dependent in the parental home, and intergenerational resource transfers (e.g., Bumpass et al. 1991; Kerr et al. 2006; Wu and Balakrishnan 1995). It is important to examine partnering behaviour in the culturally distinct region of Québec separately from the rest of Canada, because changes in partnering patterns have taken a different trajectory in the two regions (Hamplová et al. 2014; Laplante 2014). I seek to add to our broad understanding of family transformation and to provide stimulus for future research on the implications of these recent family changes.

Drawing on the 2011 General Social Survey, I update and extend past research on the changing patterns of union formation in Canada by examining how the type and timing of first union formation have changed across cohorts of Canadians born between 1930 and 1979. I examine changes across cohorts in the proportion of men and women choosing cohabitation rather than marriage as their first union type, changes in the proportion who are ever-partnered by age 35, and changes in median age at first marriage and at first partnership. I examine these aspects of union formation by region to investigate whether differences between those in Québec and in the rest of Canada are continuing to grow or if the differences are narrowing as the rest of Canada continues on the trend toward increased cohabitation and declines in marriage.

Background

Changes in union type

There has been a significant transformation in demographic behaviour in Canada and other Western countries since about the 1960s, characterized in part by increased flexibility in union formation (Kerr et al. 2006; Lesthaeghe 1995). These changes are due to a combination of ideational shifts, including secularization and increased individualism, which created new norms regulating union formation, and structural changes, including the increased labour force participation of women (Lapierre-Adamczyk and Charvet 2000; Lesthaeghe 1995; Oppenheimer 1997).

Recent cohorts of Canadians have been delaying marriage compared to cohorts who came of age in the decades following WWII (e.g., Kerr et al. 2006), and an increasing proportion are forming non-marital cohabitations (Le Bourdais and Lapierre-Adamczyk 2004). In Canada, the median age at first marriage among women reached the lowest point in the 20th century in the 1960s, at around 21 years. Since then, the median age at first marriage has been increasing dramatically; in 2008, the average first-time Canadian bride was 29 years old (ESDC 2016). At the same time, the marriage rate in Canada has been decreasing, reaching only 4.4 marriages per 1,000 people in 2008 (Statistics Canada 2012).

The trend towards delayed or forgone marriage may be offset by increases in non-marital cohabitation, which has largely become an accepted and normalized part of the transition to partnership (Guzzo 2014; Settersten and Ray 2010). Cohabiting couples accounted for 6.3 per cent of co-residential Canadian couples in 1985, 10 per cent of couples in 1995 (Wu and Balakrishnan 1995), and nearly 17 per cent of Canadian couples in 2011 (Statistics Canada 2012). The percentage of Canadians who have ever cohabited has increased over time, as has the proportion of first unions that are non-marital cohabiting relationships. Using the 1984 Canadian Fertility Study, Rao (1990) found that 20.6 per cent of Canadian women cohabited outside of marriage with their first partner. Dumas and Belanger (1997) updated this research using the 1995 General Social Survey, and found that of the Canadians who entered a first union between 1990 and 1994, 57 per cent formed a cohabiting union. The most recent information to date on the proportion of Canadians starting conjugal life through cohabitation is derived from life table estimates using the 2001 Census, which finds that 53 per cent of Canadian women born in the 1970s can expect to cohabit as a first union (Le Bourdais and Lapierre-Adamczyk 2004).

This past work has shown that the prevalence of cohabitation is increasing in Canada, but because each study uses different samples, measures, and methodologies, it is difficult to explicitly examine changes over time.

For instance, some studies have examined cohabiting unions formed in a given year (e.g., Dumas and Belanger 1997; Manning et al. 2014), some use cross-sectional data to determine how many Canadians are currently cohabiting (e.g., Wu and Balakrishnan 1995), and some estimate the proportion of people who have ever-cohabited, regardless of the order of the union (e.g., Brown et al. 2012; Bumpass and Lu 2000; Bumpass et al. 1991). In this paper, I use retrospective data to examine the differences in the type of first union across five birth cohorts of Canadians, and whether increases in cohabitation have offset declines in direct marriage.

First union timing

Median age at first marriage has been increasing in Canada, and the prevalence of cohabitation generally, and as a first union, has also increased. Yet, less is known about median age at first union when considering both marriage and cohabitation as possible first union types, especially in Canada. Manning et al. (2014) show that in the United States, median age at first union has not increased; Americans were partnering at roughly the same age between 1988 and 2010. They also show that the proportion of people who have ever partnered has also stayed relatively stable during this period.

Few studies have compared increases in the median age at marriage and median age at first union in Canada. Past studies of women born in the 1960s and earlier have found no significant changes in the median age at first partnership over time (Ravanera et al. 1998). However, since these studies, Canadians born in the 1970s have entered early adulthood and little is known about their partnering behaviour, despite many claims in the popular media that the new generations of Canadians are delaying or forgoing monogamous relationships. In this paper, I draw from the most recent available data to examine whether these trends have continued among the most recent Canadian cohort to enter into early adulthood.

Regional differences

The meaning and prevalence of cohabitation differ greatly between Québec and the rest of Canada (Hamplová et al. 2014). Quebecers tend to have more liberal perspectives on family issues than other Canadians (Wu 2000). Cohabitation has become a socially acceptable alternative to marriage in Québec, but is more likely to be a childless prelude to marriage in the rest of Canada (Hamplová et al. 2014; Kerr et al. 2006; Le Bourdais and Lapierre Adamcyk 2004). In 1981, only 7 per cent of couples in Québec were cohabiting, compared to 29.8 per cent in 2001 (Kerr et al. 2006), and 38 per cent in 2011 (Hamplová et al. 2014). There were also increases in the proportion of couples that were cohabiting in the rest of Canada during this period, but these increases were not as rapid or dramatic as those seen in Québec. In the rest of Canada, the prevalence of cohabitation increased from 5 per cent of couples in 1981 to 12 per cent in 2001 and only 14 per cent in 2011 (Hamplová et al. 2014).

The differences in union formation behaviour between people in Québec and the rest of Canada are far greater than the differences between the other Canadian provinces (Pollard and Wu 1998). In fact, the marriage rates of all of the Canadian provinces, excluding Québec, became more similar over the course of the 20th century (Wu and Balakrishnan 1995), reaching 608 per 1,000 women outside of Québec and only 373 per 1,000 women in Québec in 1994 (Pollard and Wu 1998). Moreover, the gap between the proportion of women in Québec and the rest of Canada expected to ever-marry has widened from the 1960s to the 2000s, with 40 per cent of Québec women expected to marry compared to 60 per cent of other Canadian women (Le Bourdais and Lapierre-Adamcyk 2004). In a recent study of how the competing risks of forming a first union through marriage or cohabitation have changed across cohorts of Canadians born between 1911 and 1971, Laplante (2014) found that the difference in the risks of cohabitation between French-speaking Québec Catholics and other groups of Canadians became dramatic starting in the 1951–60 birth cohort. He also finds that the risks of entering marriage have declined for all Canadians across cohorts, but that the category "French-speaking Québec Catholics and Atheists" has shown the greatest decline.

Differences in union formation behaviour between Québec and the rest of Canada have been best explained by referencing differences in normative cultures that are bounded by language, region, and religion (Laplante 2006, 2014). The demographic behaviour of French-speaking Catholic Québecois in particular diverged from other groups as the norms governing these behaviours diverged (Laplante 2006). Canadian researchers have argued that Québec experienced a "quiet revolution" in the 1960s whereby ideologies, values, and norms

changed rapidly towards individualism, secularism, and gender equality, which led to the creation of a unique regime of union formation (Laplante 2014; Pollard and Wu 1998; Wu and Baer 1996).

It is less clear whether the differences in union formation patterns between Québec and the rest of Canada have continued to increase for those born in the 1970s, or whether there has been some convergence over time as cohabitation has become increasingly popular in non-Québec Canada. Laplante's (2014) study is the most recent to examine some of these questions; however, he does not explicitly examine changes over time in the proportion of ever-partnered, or patterns in median ages at first marriage and first union over time.

Present study

Past research provides insight into the union formation behaviours of Canadians, but relies on data from 2006 or earlier. In this paper, I add to our understanding of ongoing changes in union formation by using the most recent Canadian data available on cohabitation and marriage formation. By using retrospective data on union histories reported in the 2011 GSS, I am able to build on the approaches and findings of past research by analyzing the union formation patterns of birth cohorts rather than period changes in union formation. I am also able to analyze trends over a very wide range of birth cohorts born between 1930 and 1979, which will provide a better understanding of long-term trends in marriage and cohabitation.

Research questions

In this paper I address three research questions. First, how has the type of first union that Canadians form changed over time, and have the differences in the type of first union formed between those in Québec and the rest of Canada widened for those born in the 1970s? Second, is the decline in marriage over time being offset by increases in cohabitation for Canadians in Québec and other parts of Canada? Finally, has cohabitation been delayed to the same extent as marriage across cohorts of men and women in Québec and other parts of Canada, or has earlier cohabitation offset delays in marriage?

Data

I use the 2011 General Social Survey (GSS) to examine changes in union formation across five birth cohorts in Canada. The Canadian GSS is a cross-sectional survey conducted by Statistics Canada every year since 1985, with a specific thematic focus each year. The data for this study come from Cycle 25, the fifth and most recent GSS to focus on families. The GSS uses a stratified clustered sample and is representative of the non-institutionalized population aged 15 years or older living in the 10 Canadian provinces. It was conducted by computer-assisted telephone interviews between February and November 2011, and had a response rate of 65.8 per cent. The 2011 GSS is ideal for this study because it includes detailed retrospective information on both marriage and cohabitation histories for respondents born between 1911 and 1996, which allows for an examination of long-term trends in the changes of timing and type of union formation over many birth cohorts in Canada. These data are also the most recent available on Canadian families, and cover the most recent Canadian cohorts that have reached adulthood, born in the 1970s.

Sample

I restrict my analysis to respondents with complete information who were born after 1929 and before 1980. I exclude immigrants who migrated to Canada after age 15 (n = 2,859), in order to maintain the comparability of the Québec population and the rest of Canada population, since patterns of immigration have changed dramatically in the last 100 years (Boyd and Vickers 2000), and to ensure that all of the partnering events that were included in the analyses occurred in Canada. These restrictions resulted in a sample of 15,941 respondents, reduced from the original sample of 22,435.

Measures

The GSS uses an inclusive measure of cohabitation and allows respondents to self-classify their unions as cohabitation regardless of the length of co-residence. The English version of the GSS asks respondents if they are, or have been, in a "common-law relationship, even if for less than one year." The French version asks the same questions but uses the term *union libre*. Québec follows the civil law tradition, whereas the rest of Canada is based on the common law tradition, which has resulted in different legal definitions of *unions de libres* in Québec and *common law unions* in the rest of the country (Beaujot et al. 2013). This measure of cohabitation is therefore inclusive of both definitions, used by both Anglophone and Francophone Canadians. I use the term *cohabitation* to encompass both common law unions formed outside of Québec and *unions de libres* in Québec.

I construct two separate dependent variables that reflect time to first union and time to first marriage. I use the age the respondent reported beginning their first union, either marriage or cohabitation, to construct the *age at first partnership* measure. I use respondents' reports of their age at their first marriage, regardless of any previous non-marital unions, to construct the *age at first marriage* measure. To construct my key independent variables, I group respondents by decade of birth, encompassing respondents born in the 1930s through the 1970s. I group these respondents into two categories based on their province of residence: those living in Québec and those living in the rest of Canada (ROC).

Finally, I control for several sociodemographic factors in the multivariate analyses. *Language* and *religion* have been found to predict union formation behaviours (e.g., Eggebeen and Dew 2009; Rao 1990), largely as markers of separate cultures, with different normative systems (Lapierre-Adamczyk and Charvet 2000; Laplante 2104). I classify the respondents into English-speakers, French-speakers, and others, based on the language they speak most frequently at home. I group respondents by religious affiliation, which includes categories for no religion, Catholic, Protestant, and other. Past research has also found that the more highly educated are more likely to marry (e.g., Sassler and Goldscheider 2004; Wu and Pollard 2000), and the less educated are more likely to cohabit, but that this association is weaker in Québec (Kerr et al. 2006). *Highest educational attainment* is coded as less than a high school diploma, a high school diploma, some postsecondary education including college, trades certificate or incomplete bachelor's degree, and bachelor's degree or higher. Lastly, individuals with divorced parents have been shown to be less likely to marry (Lapierre-Adamczyk and Charvet 2000), at least partly due to different attitudes towards marriage (Axinn and Thornton 1996). To control for these differences, I include an indicator for whether the respondent *lived with two parents* until at least age 15 years.

Analytic strategy

I begin by charting changes in the proportion of Canadian women and men, in Québec and in other parts of Canada, who enter their first union through marriage, through cohabitation, or who remain unpartnered by age 35. For these analyses I exclude respondents born in 1977 or later, because they had not yet reached age 35 at the time of the survey.

Next, I examine how changes across cohorts in the age at first union compare to changes in the age at first marriage, while controlling for other factors that influence the timing of union formation. To do this, I use discrete time logistic regression models to examine the risks of: 1) forming a first union and 2) entering legal marriage—separately for Québec and the ROC, and by sex, and include controls for language, religion, education, and childhood family structure. Respondents enter the risk set at age 15 and exit at age of first marriage/first union or the date of the survey, whichever occurs first. I created a person-period data file consisting of 245,941 person-years for the time to first marriage analyses, and 200,019 person-years for the time to first partnership analyses.

Unlike Cox-proportional hazards models, discrete time models require that the shape of the baseline hazard (the duration dependence) be specified (Box-Steffensmeier and Jones 2004; Jenkins 2005). Rather than assume a theoretical shape of the hazard function, I use a piecewise linear spline, with knots at each quintile of survival time, to model the duration dependence.[2] Within each linear segment, the hazard rate is assumed to be constant

2. I specified the baseline hazard in several different ways, and a 5-point piecewise spline specification was the best characterization of the baseline hazard according to a variety of fit statistics, including the BIC and AIC (Singer and Willett 2003).

but is allowed to vary across segments. This approach has the advantage of allowing the shape of the hazard function to be determined empirically without burdening the model with dummy variables for every unit of time (Singer and Willett 2003).

I then use these discrete time logistic regression models to estimate men's and women's median survival times to two events: (1) first marriage and (2) first partnership, by birth cohort and by region, while controlling for relevant sociodemographic variables. All estimates are derived using sample weights, to ensure they are representative of the population.

Results

Description of the sample

Table 1 presents characteristics of the full analytic sample. The left panel of the table provides the number of women in each birth cohort and the percentage of each cohort living in Québec and in other parts of Canada; the right panel displays the same information for men. Roughly 30 per cent of women and men born between 1930 and 1959 were living in Québec, and this proportion decreased slightly for those born between 1960 and 1979.

Table 1. Characteristics of the Sample 2011 General Social Survey Cycle 25 (Family), n=15,941

Women				Men			
Birth cohort	n	Region	%	Birth cohort	n	Region	%
1930–39	1,673	ROC	71.7	1930–39	962	ROC	72.3
		Québec	28.3			Québec	27.7
1940–49	1,717	ROC	71.7	1940–49	1,359	ROC	73.3
		Québec	28.3			Québec	26.7
1950–59	2,256	ROC	73.0	1950–59	1,766	ROC	71.4
		Québec	27.0			Québec	28.6
1960–69	1,858	ROC	75.5	1960–69	1,504	ROC	77.0
		Québec	24.5			Québec	23.0
1970–79	1,560	ROC	76.5	1970–79	1,286	ROC	77.0
		Québec	23.5			Québec	23.0

Notes: Proportions are weighted to be representative of the Canadian population. "ROC" stands for "rest of Canada," and signifies respondents living in provinces outside of Québec.

Proportion marrying, cohabiting, and never-partnered by age 35

To examine changes over time in the type of first union that Canadians form, I plot the proportion of each cohort that entered into marriage directly, the proportion whose first union was a non-marital co-residential partnership, and the proportion unpartnered by age 35. Figures 1 and 2 display these proportions for women and men, respectively, by region. The solid lines represent the proportion whose first union was legal marriage, the dashed lines represent the proportion who cohabited with their first partner, and the dotted lines represent the proportion who had never partnered by age 35. The grey lines show the trends for those living in Québec, and the black lines show the trends for those in the rest of Canada (ROC). In light grey, I provide the 95% confidence intervals derived from two-sample tests of equality of proportions for each point estimate.

Figure 1 shows that women in Québec are less likely to enter into direct marriage than women in the ROC across all birth cohorts, and the decline in the proportion entering marriage directly across birth cohorts is more dramatic in Québec than for women in the ROC. Approximately 87 per cent of women living in Québec born in the 1930s entered marriage directly, compared to approximately 93 per cent of their counterparts living in the ROC. The regional difference in the propensity to enter marriage directly is similar among women born in the 1940s. However, there is no significant regional difference in the proportion of women who cohabit with their

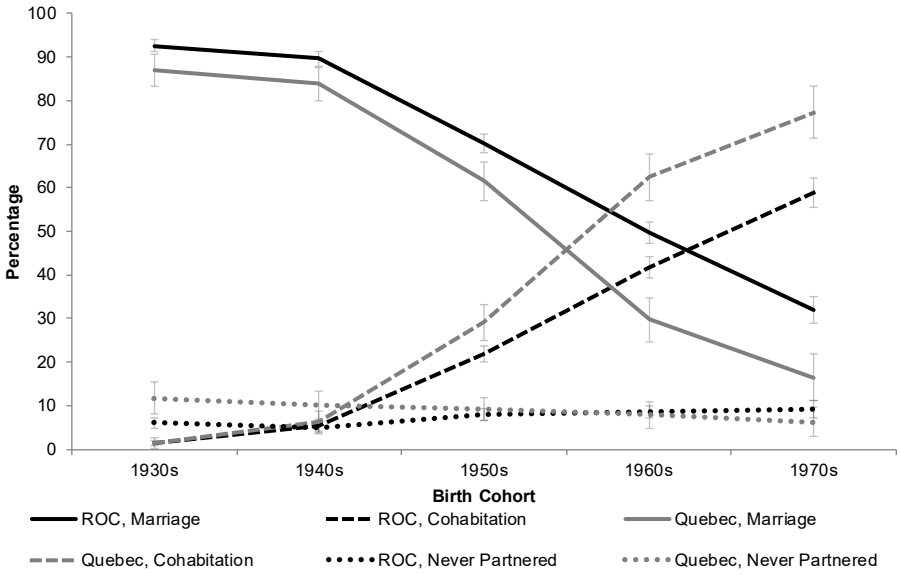


Figure 1. Percentage of women whose first union was marriage, cohabitation, and who never-partnered by age 35, by region, 1930s–70s birth cohorts, with 95% confidence intervals

first partner in these early birth cohorts; in both regions, fewer than 2 per cent of women born in the 1930s and 1940s entered their first unions through cohabitation. The difference in these cohorts is that women in Québec were less likely to be in any type of partnership than those in the ROC (10–12 per cent of those in Québec compared to 5–6 per cent of those in the ROC).

Regional differences in the proportion of women entering conjugal life through cohabitation become significant in the 1950s birth cohort, and increase among those born in the 1960s (Figure 1). Approximately 22 per cent of women in the ROC born in the 1950s cohabited with their first partner, compared to 29 per cent of women in Québec. By the 1960s birth cohort, the majority of women in Québec cohabited as their first union (62 per cent), as did a large minority of women in the ROC (42 per cent). The trend toward cohabitation and away from direct marriage continues for women born in the 1970s; however, the regional divergence in first union type that had been increasing since the 1940s birth cohort appears to have stalled among this birth cohort. Women in Québec

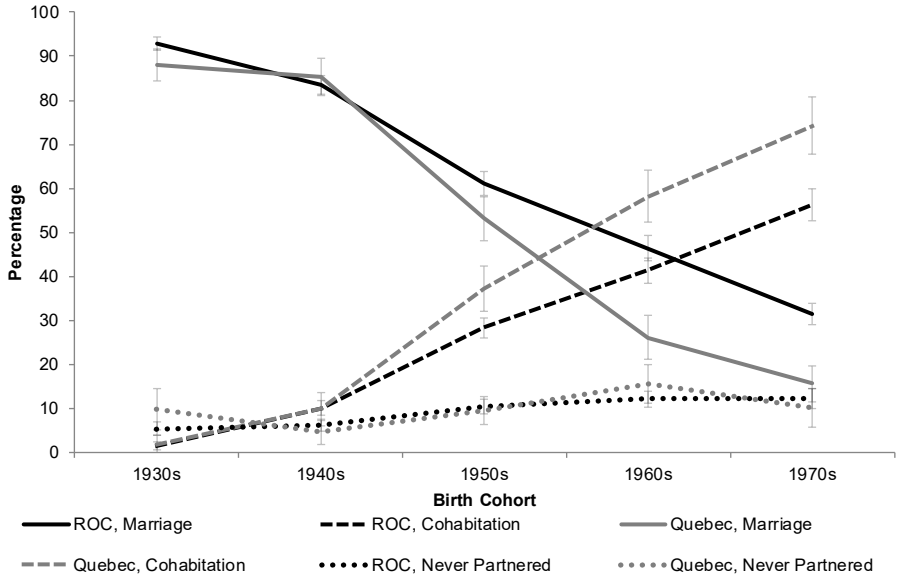


Figure 2. Percentage of men whose first union was marriage, cohabitation and who never-partnered by age 35, by region, 1930s–70s birth cohorts, with 95% confidence intervals

born in the 1970s continue to be more likely to cohabit with their first partner than their counterparts born in the ROC, but the difference is smaller than in the previous birth cohort. Approximately 16 per cent of women in Québec born in the 1970s entered directly into marriage, compared to roughly 32 per cent of women in the ROC, which indicates that large regional differences remain despite the stalled divergence over time.

The proportion of women never-partnered by age 35 has stayed relatively stable across the birth cohorts of women born in the ROC, increasing slightly, albeit statistically insignificantly, from 5–6 per cent of those born in the 1930s and 40s to 8–9 per cent of those born in the 1950s, 60s, and 70s. There are also no statistically significant changes in the proportion never-partnered by age 35 across birth cohorts among women in Québec, which decreased slightly, from 10–11 per cent of those born in the 1930s and 40s to 6–7 per cent of those born in the 1960s and 70s.

The proportions of men who enter directly into marriage, who cohabit with their first partner, and who remained unpartnered by age 35 are shown in Figure 2. The overall and region-specific patterns in men's first union types are similar to those found for women; however, those born in the 1940s and 50s were slightly less likely to enter directly into marriage than their female counterparts. Approximately 59, 44, and 28 per cent of men living in the ROC born in the 1950s, 60s, and 70s, respectively, entered directly into marriage with their first partner, compared to 53, 26, and 16 per cent of their counterparts in Québec. Regional differences in the proportion of men entering conjugal life through cohabitation first appear in the 1950s birth cohort, which is the same pattern as seen for women. As was the case for women, the difference between men born in Québec and the ROC in the propensity to marry directly also increases between the 1940s, 50s, and 60s birth cohorts, but the regional divergence slows among the most recent birth cohort, who were born in the 1970s and came of age in the 1990s, albeit to a lesser degree than is the case for women. The proportion of men born in the ROC who never partnered by age 35 has increased across birth cohorts, ranging from approximately 5–6 per cent in the early cohorts to around 12 per cent in the most recent birth cohort of men born in the 1970s, and the difference between the those born in the 1930s and 40s and those born after is statistically significant. Among Québec-born men however, there are no statistically significant changes in the proportion never-partnered by age 35 across the birth cohorts (Figure 2).

Age at first marriage vs. age at first union across cohorts

Next, I examine cohort differences in the risk of entering first marriage, and the risk of entering a first union of any kind, in order to predict the timing of union formation. Tables 2 and 3 present odds ratios from discrete-time logistic regression models for women and men, respectively. The left panel of each table displays models predicting first marriage separately by region, and the right panel displays the same for first union.

The parameter estimates describing the duration dependence shown in the bottom of Tables 2 and 3 represent the fitted baseline hazard function (transformed into odds ratios) (Jenkins 2005; Singer and Willett 2003). The baseline represents English-speaking respondents born in the 1950s, who have no religious affiliation, graduated high school, and who lived with both parents until at least age 15. Across all models, the odds of experiencing a partnering event peak in the second quintile of survival time (ages 18–21 years for the marriage models and 18–19 years for the first-union models). After this point, the baseline hazard decreases over survival time.

The marriage models for women in the left panel of Table 2 show that there are significant differences in the odds of marriage across birth cohorts for both women in Québec and women in the ROC. Women in Québec born in the 1970s have 3.2 times lower odds of marriage compared to those born in the 1950s (0.31 odds ratio). For women in the ROC, the difference is smaller; those born in the 1970s have 1.89 times lower odds of marriage compared to the 1950s birth cohort. Francophone women in Québec have lower odds of marriage than Anglophones in Québec, but there are no differences by language in the ROC. In the ROC, any religious affiliation is associated with higher odds of marriage, but Protestants have the largest difference, with odds of marriage 1.24 times that of women with no religious affiliation. In Québec, Protestant and other religious groups are more likely to marry, but Catholics are not statistically different in their odds of marriage compared to the unaffiliated. In both regions of Canada, women with higher levels of education have lower odds of marriage. Women with a bachelor's degree or higher have lower odds of marriage than the

high school educated (1.56 times lower for women in Québec, and 1.85 times lower for women in the ROC). Finally, women who grew up without two parents in the home had 1.18 times lower odds of marriage than respondents who lived with two parents, regardless of region.

Table 2. Odds ratios from discrete-time logistic regression models predicting entry into first marriage and first partnership, women, n=9,064

		First marriage		First union	
Model		Que	ROC	Que	ROC
Birth cohort (1950s)					
1930s		1.10	1.12+	0.69***	0.84**
1940s		1.29**	1.39***	0.88	1.13*
1960s		0.64***	0.67***	1.04	0.79***
1970s		0.31***	0.53***	1.08	0.73***
Language (English)					
French		0.84+	1.03	1.30*	1.14
Other		0.96+	0.98	0.94	0.99
Religion (None)					
Catholic		1.14	1.13*	0.97	0.87**
Protestant		1.47+	1.24***	1.12	0.95
Other		1.83+	1.22*	1.14	0.88
Education (High School)					
Less than H.S.		1.10	1.01	1.04	1.17*
PSE less than BA		0.88	0.76***	0.84	0.81***
BA or higher		0.64***	0.54***	0.62***	0.50***
Family structure (2 parents)					
Other family type		0.85+	0.84***	1.02	1.11*
Duration dependency					
Marriage	Union				
15–17	15–17	0.76***	0.80***	0.76***	0.80***
18–21	18–19	2.12***	1.80***	3.64***	2.84***
22–25	20–23	0.98	1.02	1.27***	1.18***
26–34	24–29	0.85***	0.92***	0.88***	0.95***
35+	30+	0.88***	0.86***	0.88***	0.87***

Note: *** p<0.001; ** p<0.01; * p<0.05; +p<0.10

Quintiles of survival time to first marriage and first partnership are used to model the duration dependency; these coefficients form the baseline hazard. Reference categories are in parentheses.

Source: 2011 General Social Survey (Cycle 25).

The right panel of Table 2 shows the estimates from the first-union models, which treat either marriage or cohabitation, whichever occurs earlier, as a first partnering event. Cohort differences in the odds of partnering are far smaller than differences in the odds of marriage, and are statistically insignificant for women in Québec (except women in Québec born in the 1930s, who are 1.45 times less likely to form any sort of partnership compared to those born in the 1950s). Cohort differences in the ROC are statistically significant, but women born in the 1970s have only 1.37 times lower odds of partnering than the 1950s birth cohort (compared to 1.89 times lower in the case of marriage). Francophone women in Québec have odds of partnering that are 1.3 times higher than Anglophone women in Québec; and, similar to the models of marriage, these are the only significant language differences. The only statistically significant difference by religious affiliation is for Catholic women in the ROC, who have 1.15 times lower odds of partnering than the unaffiliated. Educational patterns in the risks of partnership are very similar to those for risks of marriage: the more highly educated have lower odds of partnering. Finally, living in a home without two parents during childhood is associated with 1.11 times higher odds of partnering for women in the ROC, but there are no differences by childhood family structure for women in Québec.

Table 3. Odds ratios from discrete-time logistic regression models predicting entry into first marriage and first partnership, men, n=6,877

		First marriage		First union	
Model		Que	ROC	Que	ROC
Birth cohort (1950s)					
1930s		1.54***	1.57***	0.80*	1.15*
1940s		1.73***	1.41***	1.08	1.19***
1960s		0.50***	0.71***	0.79*	0.88*
1970s		0.29***	0.64***	0.90	0.84**
Language (English)					
French		0.71*	1.09	1.05	1.08
Other		1.04	0.95	1.05	0.90
Religion (None)					
Catholic		1.51**	1.12*	1.31**	0.92+
Protestant		1.38	1.36***	1.30	1.12*
Other		3.07**	1.06	1.52	0.89
Education (High School)					
Less than H.S.		0.92	0.93	1.17	1.06
PSE less than BA		0.93	1.05	1.16	1.06
BA or higher		1.00	0.96	1.03	0.86*
Family structure (2 parents)					
Other family type		0.81+	0.88*	1.08	1.07
Duration dependency					
Marriage	Union				
15–17	15–17	0.68***	0.69***	0.71*	0.71***
18–21	18–19	2.65***	2.48***	3.12***	3.97***
22–25	20–23	1.27***	1.16***	1.53***	1.41***
26–34	24–29	0.86***	0.97***	0.98	1.03**
35+	30+	0.90***	0.86***	0.88***	0.89***

Note: *** p<0.001; ** p<0.01; * p<0.05; +p<0.10
Quintiles of survival time to first marriage and first partnership are used to model the duration dependency; these coefficients form the baseline hazard. Reference categories are in parentheses.
Source: 2011 General Social Survey (Cycle 25).

Table 3 presents the same discrete-time logistic regression models predicting marriage and partnership, by region, for men. The patterns for men's odds of marriage are largely the same as for women, but there are a few exceptions (left panel of Table 3). Catholic men in Québec have 1.51 times higher odds of marrying compared to those with no religious affiliation, and there are no significant educational differences in the odds of marriage for men, in either Québec or the ROC. The right panel of Table 3 presents the models for partnering, either marriage or cohabitation. Unlike for women in Québec, there is no statistically significant difference between the odds of partnering for Francophone men compared to Anglophone men in Québec. There are also differences by religious affiliation for men: Catholic men in Québec have 1.31 times higher odds of partnering than the unaffiliated. Catholic men in the ROC however, have 1.09 times lower odds of partnering, and Protestant men have 1.12 higher odds than those with no religious affiliation. There are also few educational differences in the odds of first union for men. The only statistically significant difference is for men in the ROC who have a bachelor's degree or higher, who have odds of partnering 1.16 lower than those with a high school diploma. Family structure in childhood does not appear to be associated with the risks of partnership for men.

Median survival time to first marriage and to first union for each cohort, by region, regardless of union type and controlling for other sociodemographic factors, are presented in the top panel (for women) and bottom panel (for men) of Table 4. The estimates in these tables are derived from the discrete-time logistic regression models in Tables 2 and 3. The age at which half of the women in the ROC are estimated to form a first marriage increased, from a low of 21.2 years for women born in the 1930s to a high of 26.1 years for women born in the 1970s (top panel of Table 4). Over this time, women in the ROC have delayed their first marriage by five years,

Table 4. Fitted median times to first marriage and first union, across birth cohorts, by place of residence.

	Birth Cohort				
	1930–39	1940–49	1950–59	1960–69	1970–79
Women (n=9,064)					
Province other than Québec					
Age at first marriage	21.2	21.6	22.2	24.2	26.1
Age at first union	21.4	20.8	21.3	22.1	22.5
Québec					
Age at first marriage	21.8	21.6	22.4	24.5	–
Age at first union	23.3	22.6	22.3	23.3	22.3
Men (n=6,877)					
Province other than Québec					
Age at first marriage	23.8	24.3	25.5	27.7	28.7
Age at first union	23.6	23.6	24.2	24.8	25.1
Québec					
Age at first marriage	23.9	23.5	25.1	31.5	–
Age at first union	24.4	23.4	23.7	24.6	24.2

Note: Estimates derived from discrete-time logistic regression models predicting entry into first marriage and first partnership, displayed in Tables 2 and 3.
Dash – indicates that less than half of the subgroups are predicted to experience the partnering event by the last time in the life table, so an estimate is not available.
Source: 2011 General Social Survey (Cycle 25).

even after controlling for differences in education, religion, childhood family structure, and the other controls in Table 2. However, Canadian women are not delaying partnering to nearly the same degree as marriage. Half of all women in the ROC born in the 1930s are estimated to form their first union by age 21.4, and among those born in the 1970s, half are estimated to form their first union by age 22.5—a difference of only 1.1 years. Typical ages at first marriage and first union corresponded quite closely in the earlier cohorts, in which marriage was by far the most likely way to form a first partnership. These ages began to diverge across the cohorts, especially for the 1960s birth cohort, which entered adulthood in the 1980s, as cohabitation became an increasingly common way to form a first union.

These trends are even more pronounced among women in Québec. Half of the women in Québec born in the 1930s, 40s, and 50s are estimated to have married between age 21.6 and 22.4 years. By the 1960s birth cohort, the estimated median survival time to first marriage increased to 24.5 years. The median survival time to first marriage could not be estimated for Québec women born in 1970 or, later because less than 50 per cent of this birth cohort is predicted to be married by age 40 years, indicating a continued delaying or forgoing of marriage among this cohort. Median estimated survival time to first partnership, however, has stayed remarkably stable across the cohorts of women in Québec, even with controls. The estimated median age at first union varies by only 1 year (ranging from age 22.3 years for those born in the 1950s and 70s to 23.3 years for those born in the 1930s and 60s). The same patterns in fitted median survival times to first marriage and first partnership can be seen for men in the bottom panel of Table 2, but men tend to enter marriage and form their first partnerships 2–3 years later than their female counterparts.

Discussion and conclusion

The rise in cohabitation and delaying of marriage are two of the most important changes in union formation patterns that have occurred in Canada over the last 50 years. In this paper, I have documented these well-known known trends in older Canadian birth cohorts, and have updated previous analyses by using the most recent Canadian data available to examine the most recent cohorts of Canadians. I have also documented long-term trends in median age at first union across birth cohorts, which has been far less studied than median age at

marriage. The results contribute to our understanding of the way in which increases in cohabitation have offset the decline and delay of marriage as a first partnership for the newest cohort of Canadians entering adulthood.

Consistent with past research (e.g., Laplante 2014; Le Bourdais and Lapierre-Adamczyk 2004), I found that across birth cohorts, an increasing number of Canadian men and women are choosing to form non-marital cohabiting unions rather than marriages as they enter conjugal life. Marriage as a first partnership type has continued to decline among the most recent birth cohorts of Canadians. The decline in the proportion of Canadians whose first union was marriage, however, has been largely offset by an increase in the formation of cohabiting relationships. The proportion of women and men in Québec and women in the ROC forming any type of union by age 35 years has remained stable over the birth cohorts. Among the men in other parts of Canada, however, increases in the proportion forming cohabiting unions have not kept pace with decreases in marriage formation for the most recent birth cohorts, leading to a slight increase in the proportion never-partnered by age 35 across the birth cohorts.

I find that the trend towards delayed marriage in Canada, which began in earnest among those born in the 1960s, who came of age in the 1980s, has continued for both men and women born in the 1970s. The typical age at first partnership, when both marriage and cohabitation are considered, however, has not changed much over the course of the 50 years under study. This is further evidence that the rise in cohabiting unions has indeed offset the delays in marriage. Canadians born in the 1970s continue to form their first unions at approximately the same age as their parents' and grandparents' generations, but the type of first union they form is different.

I also examined differences in first union formation behaviours between Canadians in Québec and the ROC to determine if the disparity in the preferred type of first union that has been growing since the 1940s birth cohort has continued among the most recent birth cohorts, who came of age in the 1990s. Consistent with past research (e.g., Laplante 2014; Le Bourdais and Lapierre-Adamczyk 2004), I found that the pattern of increased preference for cohabitation and decreased preference for marriage as a first union type is more dramatic among men and women in Québec and less dramatic for those in other parts of Canada. Across all cohorts, men and women in Québec are less likely to marry their first partner. However, the this trend toward an ever-decreasing proportion of marriages as first union has slowed for the most recent cohort in Québec, while it continued for the most recent cohort in the rest of Canada, especially for men. This means that the difference in choice of first union type between the Quebecers and other Canadians, which has been growing since at least the 1940s birth cohort, has stabilized among the youngest Canadians included in this study. This provides some evidence that the meaning and place of cohabitation in the union formation process in the rest of Canada may be becoming more like that found in Québec. Further examination of the characteristics and outcomes of cohabiting unions of recent birth cohorts in the two regions is needed to fully address this question.

Québec also displays a more dramatic pattern of change in age at first marriage and first partnership over time than the ROC. Age at first marriage has increased to a greater extent in Québec, but age at first partnership continues to be younger in Québec than in the rest of Canada. Quebecers are increasingly moving away from marriage, but not only are they still partnering, in more recent birth cohorts they are doing so earlier than other Canadians.

The picture that these results reveal about how the role of cohabitation differs in Québec and the rest of Canada is clear. On the one hand, Canadians outside of Québec seem to be catching up to those in Québec in terms of their propensity to start their conjugal lives through cohabitation. This indicates that cohabitation as a first union type is perhaps on its way to near universality among non-immigrant Canadians. However, this says little about whether these first cohabiting unions, or cohabiting unions in general, have replaced marriage or whether they are better conceived as a stage in the marriage process. The large differences in age at first marriage do, however, provide some evidence that marriage is still much more common among men and women in other parts of Canada than it is among those in Québec who are delaying and increasingly forgoing marriage altogether.

This study has many advantages, including the use of the most recent available Canadian data on union formation, and the inclusion of a wide range of birth cohorts of Canadians born between 1930 and 1979. However, it is not without its limitations. One limitation is that it excludes other determinants of type of first union choice and timing of first union that have been shown to be important in past research, including the conception, birth, and presence of children, income, and work status (e.g., Eggebeen and Dew 2009; Kerr et al.

2006; Rao 1990). The 2011 GSS includes retrospective information about fertility and work histories, so future work could include these measures to further the results of this study. Unfortunately, the data do not include time-varying measures of income, so a different data source is necessary to directly examine hypotheses about association between delayed or forgone marriage and income.

A second limitation is the reliance on retrospective data regarding union histories. As with all retrospective data, these data are subject to recall and mortality biases (Hassan 2005). Recall bias is likely less of a problem when studying significant life course events, such as marriage and cohabitation, that this paper addresses, than it may be for more mundane or more frequently occurring events (Freedman et al. 1988). Past research has also shown that many couples "slide" into cohabitation (Manning and Smock 2005; Stanley et al. 2006), so assessing the exact timing of cohabitation union start may be difficult. This is not a major limitation in this study, because I use age at union start, which is easier to determine than the specific date that the union began. The mortality bias introduced by the data is likely more serious for the earlier birth cohorts under examination. Respondents born in the 1930s and 40s were between 62 and 81 years old at the time of the survey, and only individuals who survived to this age could be sampled. The median survival time to first marriage for these birth cohorts found in this study corresponds closely with past studies of these cohorts (e.g., Pollard and Wu 1998; Rao 1990; Ravanera et al. 2002), so it appears that the mortality bias is not a large concern.

Despite its limitations, this paper contributes to our understanding of the first partnering behaviours of recent cohorts of Canadians. The widespread changes in union formation that have occurred in Canada over the last 50 years are continuing among the newest generation of Canadians to come of age. Cohabitation is increasingly becoming the most common way to form a first union, and marriage is being delayed even longer and is increasingly forgone, especially in Québec. Yet the more things change, the more they seem to stay the same. Namely, the proportion of Canadians that have formed any type of union by age 35 has not declined along with the decline in marriage for those in Québec, or women in the ROC, and has declined only marginally for men in the ROC. The typical age at first union has stayed remarkably stable across the cohorts of Canadians born between 1930 and 1979. However, the changes in the types of unions that Canadians are forming may have further implications. For instance, if cohabiting relationships continue to be less stable than marriages (Bumpass and Lu 2000), and if unions formed at younger ages are more likely to dissolve, we can expect that more recent cohorts of Canadians will experience more turbulent partnership trajectories than past generations. This paper serves as the foundation for future studies on the explanations and consequences of the partnership behaviours of Canadians born after the 1970s.

Acknowledgements

I am grateful to Rachel Margolis, Lisa Strohschein, and the anonymous reviewers for their valuable comments on this manuscript, and to Jean-François Nault for translating the abstract into French. This research was supported by the Social Sciences and Humanities Research Council of Canada.

References

Axinn, W.G., and A. Thornton. 1996. The influence of parents' marital dissolutions on children's attitudes toward family formation. *Demography* 33(1):66–81.

Beaujot, R., C.J. Du, and Z. Ravanera. 2013. Family policies in Québec and the rest of Canada: Implications for fertility, child-care, women's paid work, and child development indicators. *Canadian Public Policy* 39(2):221–39.

Box-Steffensmeier, J.M., and B.S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.

Boyd, M., and M. Vickers. 2000. 100 years of immigration in Canada. *Canadian Social Trends*. Catalogue no. 11-008.

Bumpass, L.L., and H. Lu. 2000. Trends in cohabitation and implications for children's family contexts in the United States. *Population Studies* 54(1):29–41.

Bumpass, L.L., J.A. Sweet, and A. Cherlin. 1991. The role of cohabitation in declining rates of marriage. *Journal of Marriage and Family* 53(4):913–27.

Brown, S.L., J. Roebuck, and G.R. Lee. 2012. Transitions into and out of cohabitation in later life. *Journal of Marriage and Family* 74(4):774–93.

Dumas, J., and A. Belanger. 1997. *Report on the Demographic Situation in Canada 1996*. Cat. 91-209E. Ottawa: Statistics Canada.

Eggebeen, D., and J. Dew. 2009. The role of religion in adolescence for family formation in young adulthood. *Journal of Marriage and Family* 71:108–21.

ESDC (Employment and Social Development Canada). 2016. Family life – marriage. *Indicators of Well-being in Canada*. Accessed Feb 2016 from http://well-being.esdc.gc.ca/misme- iowb/.3ndic.1t.4r@-eng.jsp?iid=78.

Freedman, D., A. Thornton, D. Camburn, D. Alwin, and L. Young-DeMarco. 1988. The life history calendar: A technique for collecting retrospective data. *Sociological Methodology* 18:37–68.

Guzzo, K.B. 2014. Trends in cohabitation outcomes: Compositional changes and engagement among never-married young adults. *Journal of Marriage and Family* 76(4):826–42.

Hamplová, D., C. Le Bourdais, and É. Lapierre-Adamczyk. 2014. Is the cohabitation-marriage gap in money pooling universal? *Journal of Marriage and Family* 76:983–97.

Hassan, E. 2005. Recall bias can be a threat to retrospective and prospective research designs. *The Internet Journal of Epidemiology* 3(2).

Jenkins, S.P. 2005. Survival analysis. Institute for Social and Economic Research, University of Essex. Available online at http://www.iser.essex.ac.uk/survival-analysis.

Kerr, D., M. Moyser, and R. Beaujot. 2006. Marriage and cohabitation: Demographic and socioeconomic differences in Québec and Canada. *Canadian Studies in Population* 33(1):83–117.

Lapierre-Adamcyk, É., and C. Charvet. 2000. Cohabitation and marriage: An assesment of research in demography. *Canadian Studies in Population* 27(1):239–54.

Laplante, B. 2006. The rise of cohabitation in Quebec: Power of religion and power over religion. *Canadian Journal of Sociology* 31:2–24.

Laplante, B. 2014. Normative groups: The rise of the formation of the first union through cohabitation in Québec, A comparative approach. *Population Research and Policy Review* 33:257–85.

Le Bourdais, C., and É. Lapierre-Adamcyk. 2004. Changes in conjugal life in Canada: Is cohabitation progressively replacing marriage? *Journal of Marriage and Family* 66:929–42.

Lesthaeghe, R. 1995. The Second Demographic Transition in Western countries: An interpretation, in *Gender and Family Change in Industrialized Countries*, edited by K. Oppenheim Mason and A-M. Jensen. Oxford: Clarendon Press, p. 17–62.

Manning, W.D., and P.D. Smock. 2005. Measuring and modeling cohabitation: New perspectives from qualitative data. *Journal of Marriage and Family* 67:989–1002.

Manning, W.D., S.L. Brown, and K.K. Payne. 2014. Two decades of stability and change in age at first union formation. *Journal of Marriage and Family* 76:247–60.

Oppenheimer, V.K. 1997. Women's employment and the gain to marriage: The specialization and trading model. *Annual Review of Sociology* 23:431–53.

Pollard, M.S., and Z. Wu. 1998. Divergence of marriage patterns in Québec and elsewhere in Canada. *Population and Development Review* 24(2):329–56.

Rao, K.V. 1990. Marriage risks, cohabitation and premarital births in Canada. *European Journal of Population* 6:27–49.

Ravanera, Z., F. Rajulton, and T.K. Burch. 1998. Early life transitions of Canadian women: A cohort analysis of timing, sequences, and variations. *European Journal of Population* 14: 179–204.

Ravanera, Z., F. Rajulton, T.K. Burch, and C. Le Bourdais. 2002. The early life courses of Canadian men: Analysis of timing and sequences of events. *Canadian Studies in Population* 29:293–312.

Sassler, S., and F. Goldscheider. 2004. Revisiting Jane Austen's theory of marriage timing: Changes in union formation among American men in the late 20th century. *Journal of Family Issues* 25(2):139–66.

Settersten, R.A., and B. Ray. 2010. What's going on with young people today? The long and twisting path to adulthood. *The Future of Children* 20(1):19–41.

Singer, J.D., and J.B. Willett. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press.

Stanley, S.M., G.K. Rhoades, and H.J. Markman. 2006. Sliding versus deciding: Inertia and the premarital cohabitation effect. *Family Relations* 55:499–509.

Statistics Canada. 2012. *Portrait of Families and Living Arrangements in Canada: Families, Households and Marital Status, 2011 Census of Population*. Catalogue no. 98-312-X2011001. Ottawa: Statistics Canada.

Wu, Z. 2000. *Cohabitation: An Alternative Form of Family Living*. Don Mills, ON: Oxford University Press.

Wu, Z., and D.E. Baer. 1996. Attitudes toward family and gender roles: A comparison of English and French Canadian women. *Journal of Comparative Family Studies* 27:437–52.

Wu, Z., and T.R. Balakrishnan. 1995. Dissolution of premarital cohabitation in Canada. *Demography* 32(4):521–32.

Wu, Z., and M.S. Pollard. 2000. Economic circumstances and the stability of nonmarital cohabitation. *Journal of Family Issues* 21(3):303–28.

Canadian Studies in Population 43(3–4): 249–63.

The persistent caste divide in India's infant mortality: A study of Dalits (ex-untouchables), Adivasis (indigenous peoples), Other Backward Classes, and forward castes

Bali Ram[1]
Abhishek Singh
Awdhesh Yadav

Abstract

Using data from two national surveys, this paper examines caste differences in infant mortality in India. We find that children from the three lower caste groups—Dalits (ex-untouchables), Adivasis (indigenous peoples) and Other Backward Classes—are significantly more likely than forward-caste children to die young. While this observation largely mirrors caste differences in socioeconomic conditions, low socioeconomic status is found to be only a partial explanation for higher infant mortality among lower castes. Higher mortality risks among backward-class children are almost entirely attributable to background characteristics. However, Dalit children are most vulnerable in the neonatal period even when all background characteristics are taken into account, whereas Adivasi children remain highly vulnerable in the post-neonatal period.

Keywords: infant mortality, neonatal mortality, post-neonatal mortality, caste, India's ex-untouchables.

Resume

Au moyen des données provenant des deux enquêtes nationales, cet article examine les différences dans la mortalité infantile par caste en Inde. Nous constatons que, par rapport aux enfants des castes élevées, ceux des trois castes inférieures, notamment les dalits (les ex-intouchables), les adivasis (peuples indigènes) et autres classes défavorisées (plusieurs castes désignées comme appartenant à un groupe défavorisé) courent un risque beaucoup plus grand de mourir jeunes. Bien que cette observation reflète largement les différences entre les castes sur le plan socioéconomique, le faible niveau socioéconomique n'explique qu'en partie le taux de mortalité plus élevé chez les castes inférieures. Les risques de mortalité des enfants des castes inférieures étaient presque entièrement attribuables aux caractéristiques des antécédents de la mère. Cependant, les enfants dalits demeurent les plus vulnérables pendant la période néonatale, bien que le risque de mortalité demeure le même que celui des enfants des castes supérieures pour la période post-néonatale. L'inverse est vrai pour les enfants adivasis : les caractéristiques des antécédents expliquent leur plus grande vulnérabilité pendant la période néonatale, mais pas pendant la période post-néonatale.

Mots-clés : mortalité infantile, mortalité néonatale, mortalité post-néonatale, caste, ex-intouchables en Inde.

Introduction

Caste has been a major foundation of the Indian social structure and stratification system since ancient times. In India there are thousands of castes, which have been classified into broad social groups and ranked according to social status, power, and prestige, emanating mainly from their own and their ancestors' occupations. That people from many "lower castes" have been oppressed, disadvantaged, and discriminated in social

1. Corresponding author: Bali Ram, Department of Sociology and Anthropology, B750 Loeb Building, Carleton University, 1125 Colonel By Drive, Ottawa, ON, Canada, K1S 5B6; email: bram42@yahoo.com. Co-authors Abhishek Singh, International Institute for Population Sciences, Mumbai, India, and Awdhesh Yadav, Public Health Foundation of India, Gurgaon, Haryana, India.

and economic spheres of life, and even treated as "untouchables" for centuries is well documented in popular and scholarly literature (Desai and Dubey 2012; Desai and Kulkarni 2008; Dommaraju et al. 2008). Since the introduction of its own constitution in 1950, India has prohibited discrimination based on caste and legally abolished the notion of untouchability, and special initiatives have been developed whereby people belonging to certain lower castes and tribes are provided special status and privileges in certain sectors such as education and employment. However, there is ample evidence that most people belonging to lower castes—ex-untouchables (Dalits hereafter), indigenous peoples (Adivasis hereafter), and other backward classes (OBCs hereafter)—still lag behind those from "forward castes" (sometimes referred to as "upper castes") on almost all social indicators (Desai and Dubey 2012; Desai and Kulkarni 2008; Kumar et al. 2008; Majid 2012). This paper examines caste differences in infant mortality, a highly sensitive indicator of population health and wellbeing in less industrialized countries. We do so by investigating the relationship between young children's risks of dying and their caste background, while controlling for pertinent socioeconomic characteristics. Specifically, there are three main objectives of this study: (1) determine the extent to which caste differences in infant mortality can be explained by socioeconomic background; (2) find out if caste differences in mortality risks vary by the stage of infancy; and (2) examine to what extent caste differences in infant mortality have converged over time.

Background and conceptual framework

It is well known that higher mortality rates in minority populations in most parts of the world are at least partly associated with their lower socioeconomic levels. A number of studies, originating particularly from the United States, show that racial and ethnic differences in mortality can be largely explained by socioeconomic characteristics (Crimmins et al. 2004; Hummer 1996; Williams and Collins, 1995). However, there is no dearth of counter-evidence, according to which substantial racial and ethnic differences in mortality persist even when socioeconomic status, living conditions, life style, and neighbourhood are controlled (Brown et al. 2012; Hayward et al. 2000; Williams and Sternthal 2010). At every level of socioeconomic status, blacks, for example, have poorer health outcomes and higher mortality than whites, suggesting that socioeconomic status does not fully explain racial differences in mortality. Some researchers find racism, discrimination, and stressful life events to be largely responsible for this phenomenon (Bratter and Gorman 2011; Williams 1999). According to another exception called the "epidemiological paradox" or "Hispanic paradox" (Ruiz et al. 2013), people of Hispanic descent in the United States exhibit higher life expectancy and lower infant mortality than their non-Hispanic white counterparts, despite disadvantaged socioeconomic conditions and poor access to health services. While there is no consensus on the explanations of the paradox, some researchers point to strong social ties, certain cultural practices, selective immigration, and a healthy lifestyle as protective factors (Fenelon 2013). Although it is highly presumptuous to extrapolate American experiences to Indian society, it is reasonable to hypothesize that socioeconomic background does not provide a complete explanation for understanding higher mortality among the lower castes in India.

There are hundreds of small-scale sociological and anthropological studies investigating aspects of socioeconomic life among people from certain castes and tribes in India; however, there is a relative dearth of research that focuses on caste differences in health, illness, and mortality, despite their important humanitarian and policy relevance. This is largely due to the lack of reliable data by caste at the state, regional, and national levels. However, in recent years, some large-scale national sample surveys have allowed researchers to conduct descriptive and comparative analyses of the four broad caste groups: "scheduled castes" or Dalits, "scheduled tribes" or Adivasis, "backward class" (OBCs), and upper or forward castes (Baru et al. 2010; IIPS 2000, 2007; Nayar 2007; Pandey et al. 1998). These studies show that while there has been an overall decline in infant mortality over the last fifty years, caste disparities in infant mortality persist.

These surveys have also allowed researchers to carry out micro-level multivariate analyses, with the aim of disentangling the effects of caste and socioeconomic background on mortality (Das et al. 2010; Dommaraju et al. 2008; June et al. 2011; Mohindra et al. 2006; Nguyen et al. 2013; Singh-Manoux et al. 2008; Subramanian et al. 2006a, 2006b). These studies typically use logit models, with infant and child mortality as the outcome variable,

caste as the primary independent variable, and various parental characteristics as control variables. In general, these studies find that caste differences in infant and child mortality are substantially reduced after parental socioeconomic characteristics are held constant. Subramanian and colleagues (2006a, 2006b; June et al. 2011; Singh-Manoux et al. 2008) analyze mortality data from the NFHS-2 and Human Development Survey, 2004–05, and attribute caste differences in infant mortality to primarily caste differences in socioeconomic wellbeing. They also find that socioeconomic variables are largely responsible for mortality differences between Adivasi and non-Adivasi populations, although Adivasi children are still at a significantly greater risk of dying during the early childhood period (i.e., between the ages of 2 and 5 years). A major limitation of these studies is that they do not make the distinction between children who die within the first few weeks of life and those who die in later infancy. Using data from the NFHS-2, Dommaraju et al. (2008) examine the effects of caste on child mortality and find that caste differences in mortality cannot be attributed to socioeconomic factors alone. They find that socioeconomic factors play a more important role in explaining the differences between Dalits/Adivasis and OBCs than between forward and lower castes. This study groups Dalits and Adivasis in one category, in spite of considerable differences in their social, cultural, and geographic backgrounds. By analyzing rural mortality, Das et al. (2010) reinforce the findings of earlier studies and show that Adivasi children are less likely than Dalit children to die during infancy, but more likely to die during early childhood, particularly by the time they are five years old. Nguyen et al. (2013) focus on just two states, Odisha and Madhya Pradesh. By using a number of data sets, they find that in Madhya Pradesh the under–age 5 mortality among Dalits and Adivasis has fallen at a faster pace compared with that among the other caste groups, whereas in Odisha the converse is true. The study also finds that in Odisha, for both groups the neonatal mortality rate has declined at a steady pace, while in Madhya Pradesh it has stagnated. This study does not differentiate between backward and upper castes.

In the present study, we contribute to the understanding of caste differences in mortality in three ways. First, unlike some previous studies, we focus on the four caste groups—Dalits, Adivasis, OBCs, and forward castes—separately, for which pertinent data are available. Merging two caste groups into one category is of limited significance as far as policy implications are concerned.

Second, considering the fact that the measure of overall infant mortality analyzed in previous studies masks much of the fine differences in health, illness, and mortality among various population groups at different stages of life, we distinguish between mortality that occurs in the first month of life (neonatal mortality) and mortality that occurs in the following eleven months of life (post-neonatal mortality). Since causes of death in these two periods are quite different, it is useful to examine these components of infant mortality separately (Bicego and Boerma 1993; Lawn et al. 2005; Shryock et al. 1973: 405–06). Neonatal deaths are largely attributable to *endogenous* factors—perinatal and biologic-genetic causes—such as pre-term birth complications, low birth weight, asphyxia, congenital anomalies, diarrhea, tetanus, and severe infections, whereas post-neonatal deaths result from *exogenous* factors, such as poor hygiene, communicable diseases, malnutrition, and unintentional injuries, which are generally caused by socio-environmental conditions that arise after delivery. Usually, changes in socioeconomic and environmental conditions, including improvements in sanitation and public hygiene, improved nutrition, and increased availability of vaccines and antibiotics, contribute to a reduction in mortality among older children compared to younger ones. These factors are influenced by the family, community, or public policy measures as they affect conditions that arise after childbirth, when both mothers and children have survived the physiologically most vulnerable stage of life. In contrast, reducing mortality among very young infants is a more arduous task, which can only be achieved by improvements in prenatal care, health care facilities, and mothers' nutritional status, as well as reductions in infectious diseases. Thus, it is reasonable to hypothesize that net of socioeconomic factors, caste differences in mortality would be smaller during the post-neonatal period than during the neonatal period.

Third, we chart out temporal patterns in caste differences in the above measures of mortality. We hypothesize that with an increased emphasis on maternal and child health care services and special privileges for lower castes, the mortality gap between upper and lower castes would be reduced with the passage of time. Previous studies have analyzed data for just one period, and have made speculations about the convergence between various caste groups; however, with the availability of comparable data sets for various points in time, this study examines caste differences in mortality, controlling for relevant variables over time.

Data and method

This study analyses micro-data files obtained from the last two waves of the National Family Health Survey (NFHS-2, 1998–99, and NFHS-3, 2005–06). These surveys were designed to provide estimates on various aspects of demographic behaviour, including mortality and health. They were conducted by the International Institute for Population Sciences, Mumbai, under the stewardship of the Ministry of Health and Family Welfare, Government of India, and with technical assistance from ORC Macro (now known as ICF International) in Calverton, Maryland, USA. We base our analysis on information from weighted samples of births which occurred during the five years preceding the surveys.

In both surveys, response rates among women interviewed were quite high: 95.5 per cent in NFHS-2 and 94.5 per cent in NFHS-3. These surveys adopted a two-stage sampling design in rural areas and a three-stage design in urban areas. In rural areas, villages were selected in the first stage using a probability proportional to size (PPS) sampling scheme. Households were selected in the second stage using a systematic sampling scheme. In urban areas, wards were selected in the first stage using a PPS sampling scheme. Census enumeration blocks (having approximately 150–200 households) were selected in the second stage using PPS. Households were selected in the third stage using systematic sampling (IIPS 2000, 2007).

The risk of children dying before reaching their first birthday (infant mortality) is the major dependent variable in this study. Data for estimating the risks of dying were based on the number of children who were born during the five years preceding the survey (56,259 in NFHS-2; 51,172 in NFHS-3). Risk is categorized in two ways: the risk of dying in the first month of life (neonatal mortality) and the risk of dying after the first month of life but before the first birthday (post-neonatal mortality). It may be useful to note that age at death was recorded in days for children who died in the neonatal period and in months for children who died in the post-neonatal period (IIPS 2000, 2007).

The primary independent variable is the caste group of the child's mother. As mentioned before, there are four caste groups: Dalits, Adivasis, OBCs, and forward castes. Dalits include castes which were formerly labelled *untouchables* and are now classified as "Scheduled Castes" (SC) by the Government of India. Adivasis include indigenous or aboriginal peoples, who are labelled "Scheduled Tribes" (ST). OBCs—"Other Backward Classes" (OBCs) in Government of India documents—is a somewhat poorly defined category, which includes a number of educationally and socially disadvantaged castes. "Forward caste" is a remainder category, which usually consists of Brahmins, Kshatriyas, and some Vaishya castes.

There are two sets of control variables. The first set includes four demographic variables: mother's age, child's sex, region of residence (South, North, Central, East, Northeast and West), and place of residence (urban and rural). The second set includes two measures of socioeconomic status: mother's education and the standard of living index (SLI). Mother's age and child's sex are the two most important covariates of mortality among children. Studies show that children born to younger and older mothers are more likely to die than those born to middle-aged mothers (Mathews and MacDorman 2013). Generally, younger women have little knowledge, experience, or resources for parenting and are less likely to use either antenatal care or delivery care, or to have their infants immunized, whereas older women are not only at an increased risk of having adverse medical conditions such as hypertension and diabetes, but also tend to lack the required time to care for their later-born children (Sharma et al. 2008). While it is a well documented fact that female infants have a biological survival advantage over males, this is not necessarily true in the Indian sociocultural context. Usually, mortality is significantly lower among girls during the first month of life—the neonatal period, which is indicative of the biological superiority of baby girls (Ulizzi and Zonta 2002); however, the picture is reversed during the post-neonatal period, when mortality becomes susceptible to "societal manipulation" (Das Gupta 1987; Lahiri et al. 2011). Region of residence and place of residence are important from the viewpoint of geographic distribution by caste; Dalits and Adivasis are heavily concentrated in rural areas. There are important regional differences in infant mortality in India, with southern states showing lower mortality levels than northern states (Pandey et al. 1998; Ram et al. 2013). The regional differences may reflect different effects of state government policies. Mortality levels are also higher in rural than urban areas (Pandey et al. 1998; Singh et al. 2013). Basically, these two variables act as proxies for the availability and accessibility to health care facilities.

Maternal education is by far one of the most important predictors of mortality in less industrialized countries (Basu and Stephenson 2005; Caldwell 1979) and also a variable that explains much of the ethnic and cultural differences in mortality (Antai 2011). It is known to be a valid proxy for life style which in turn influences various risk factors, such as smoking, alcohol use, limited or no breastfeeding, and obesity, which are associated with health and infant mortality. Furthermore, education enables mothers to process information regarding healthy behaviours and to better utilize existing medical facilities (Vikram et al. 2012). Educational attainment, as measured by years of schooling, is known to be superior to other dimensions of socioeconomic status in an agricultural economy such as India because it can be better ascertained, with reliable accuracy, from self-reports. Also, unlike other measures such as occupation and income, it is "cumulative and irreversible" and is an important determinant of those measures. We also include a standard of living index (SLI) as a control variable, which is a summary measure of household quality of life and economic wellbeing. This index was calculated by adding scores for the following eleven variables: dwelling type, toilet facility, source of lighting, main fuel of cooking, source of drinking water, a separate kitchen, ownership of a house, ownership of agricultural land, ownership of irrigated land, ownership of livestock, and ownership of durable goods. Index scores range from 0–14 for a low SLI to 15–24 for a medium SLI and 25–67 for a high SLI (IIPS 2000, 2007).

We use the Cox proportional hazards model to estimate the net effects of caste and its covariates on the three measures of mortality described above. This model, unlike the logit model used in previous studies (cf. Dommaraju et al. 2008), allows for the inclusion of censored data on children who could not complete the exposure period at the time of interview. We present three models. The first model controls for two fundamental demographic covariates: mother's age and child's sex. The second model adds region of residence, place of residence (rural-urban), and maternal education. The third model includes SLI as an additional control. The upper/forward caste—the lowest-mortality group—is the reference category. *Infant mortality* refers to the probability of newborn children dying before reaching their first birthday. *Neonatal* and *post-neonatal mortality* refer to the probability of dying in the first and the next eleven months of life, respectively. Death is a dichotomous variable, where '0' indicates that the child survived the period under study and '1' indicates otherwise (i.e., the child died before reaching their first birthday in the case of infant mortality). We focus primarily on hazards ratios and their 95% confidence intervals (CI). If a hazard ratio (HR) is greater than 1, the relationship is positive, and if it is less than 1, the relationship is negative.

It may be useful to state at the outset that this study is cross-sectional, and therefore it would be somewhat presumptuous to infer causal relationships between various variables. Considering that the dependent variables are derived from information on births that occurred during the five years preceding the survey, while the independent variables refer to the survey date, there is a possibility of causality running in a reverse direction. However, retrospective information is likely to circumvent this problem to a large extent.

Characteristics of the study sample

Table 1 presents the distribution of the cases in the two samples, by independent variables, included in the study. As expected, in both surveys, the samples include a slightly larger proportion of children who are male (52 per cent). In NFHS-3, about 7 per cent of the mothers belong to the 15–19 age group, slightly lower than that in NFHS-2 (9 per cent); the proportion belonging to the age group 30 and over in NFHS-3 is similar to that in NFHS-2 (25 versus 24 per cent). Once again, this is to be expected, considering that the average maternal age has increased during the inter-survey period. The distribution of samples by rural-urban residence is also consistent with the expectation. The overwhelming majority of the sample cases live in rural areas, although their proportion is slightly lower in NFHS-3 than that in NFHS-2 (75 versus 78 per cent). In NFHS-2, the majority (57 per cent) of mothers are illiterate; in NFHS-3, this figure is significantly lower (50 per cent). Conversely, over the survey period there is a substantial increase in the proportion of mothers who have attained 9 years or more schooling (from 17 to 22 per cent). Consistent with this trend, we also find a sharp rise in the proportion of mothers with high a standard of living (16 per cent in NFHS-2 compared with 32 per cent in NFHS-3). The regional distributions of the sample cases in the two surveys are generally comparable. The distributions of sample cases by caste groups are not highly comparable between the two surveys, which may have happened due to the deliberate misreporting of

caste and the reallocation of certain castes from one group to another during the inter-survey period. Thus, about 37 per cent of the cases in NFHS-2, but only 26 per cent in NFHS-3, belong to forward castes, while about 32 per cent in NFHS-2, and 40 per cent in NFHS-3, belong to OBCs. Proportions of cases for Dalits (20 per cent) and Adivasis (10 per cent) groups are highly comparable between the surveys. In a small proportion of the cases, caste is not reported in both surveys (1 per cent in NFHS-2 and 3 per cent in NFHS-3).

Table 1. Characteristics of the samples, NFHS-2 and NFHS-3

	NFHS-2		NFHS-3	
	Percentage*	Number**	Percentage*	Number**
Caste				
Dalits	19.8	10,353	20.7	15,074
Adivasis	9.8	8,478	9.6	9,167
OBCs	32.0	15,956	40.3	8,386
Forward castes	37.1	21,424	26.4	16,746
Don't Know	1.3	523	3.0	2,182
Mother's age				
< 20 years	9.0	4,142	6.6	2,677
20–29 years	67.6	37,969	68.7	34,495
30+ years	23.5	14,623	24.7	14,383
Child's sex				
Male	51.7	29,478	52.1	26,799
Female	48.3	27,256	47.9	24,756
Region				
South	18.9	7,587	15.8	7,232
North	12.8	13,321	13.0	9,286
Central	29.5	12,526	29.7	11,659
East	22.0	9,817	25.2	8,126
Northeast	3.7	7,872	3.8	9,655
West	13.2	5,611	12.5	5,597
Place of residence				
Rural	78.1	42,210	74.7	32,072
Urban	21.9	14,524	25.3	19,483
Mother's education				
No education	57.1	30,298	50.2	21,125
1–8 years	26.2	15,642	27.8	15,337
9+ years	16.7	10,767	22.0	15,092
SLI				
Low	37.8	18,741	33.5	12,224
Medium	46.7	27,169	35.0	16,326
High	15.5	10,098	31.5	18,350

Note: * weighted percentage; ** unweighted number

Multivariate analysis

Higher infant mortality rates among lower-castes are often associated with their disadvantaged socioeconomic background. In order to address this hypothesis, we apply the Cox Proportional Hazards Model for estimating the effects of caste on infant mortality by controlling for potential confounding variables. Results presented in the basic model 1A of Table 2, which includes two fundamental demographic variables—mother's age at the time of the survey and sex of the child—as controls, show that in NFHS-2 caste differences in the likelihood of children dying during infancy are large and highly significant. Compared with the mortality risk for forward-caste children (reference category), the risks are 23 per cent greater for OBC children (HR = 1.23; CI = 1.13, 1.34), 32 per cent greater for Dalit children (HR = 1.32; CI = 1.20, 1.45), and 27 per cent greater for Adivasi children (HR = 1.27; CI = 1.15, 1.41). Surprisingly, the picture does not change much over the next seven years. Thus, in NFHS-3 (Model 1B) the relative risks for OBC (HR = 1.25; CI = 1.13, 1.39) and Adivasi children (HR = 1.28; CI = 1.13, 1.46) remain virtually unchanged, whereas the risk for Dalit children worsens somewhat (HR = 1.42; CI = 1.26, 1.59).

Table 2. Partial results of Cox Proportional Hazards Model of infant mortality, NFHS-2 and NFHS-3

	Model 1A NFHS-2		Model 1B NFHS-3		Model 2A NFHS-2		Model 2B NFHS-3		Model 3A NFHS-2		Model 3B NFHS-3	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Caste												
Forward(R)												
Dalits	1.32***	(1.20, 1.45)	1.42***	(1.26, 1.59)	1.18***	(1.07, 1.30)	1.19***	(1.06, 1.34)	1.12**	(1.01, 1.23)	1.17**	(1.03, 1.32)
Adivasis	1.27***	(1.15, 1.41)	1.28***	(1.13, 1.46)	1.18***	(1.06, 1.32)	1.10	(0.96, 1.26)	1.13**	(1.01, 1.26)	1.08	(0.94, 1.24)
OBCs	1.23***	(1.13, 1.34)	1.25***	(1.13, 1.39)	1.15***	(1.06, 1.26)	1.09	(0.97, 1.21)	1.13***	(1.03, 1.23)	1.09	(0.97, 1.22)
Don't know	1.63***	(1.22, 2.16)	1.09	(0.87, 1.36)	1.25	(0.93, 1.67)	1.06	(0.84, 1.32)	1.15	(0.86, 1.55)	1.09	(0.87, 1.36)
Child's sex												
Male(R)												
Female	0.96	(0.90, 1.03)	0.95	(0.88, 1.03)	0.95	(0.89, 1.02)	0.95	(0.87, 1.02)	0.95	(0.89, 1.02)	0.95	(0.88, 1.03)
Mother's age												
< 20 years												
20–29 years	0.30***	(0.27, 0.34)	0.27***	(0.23, 0.31)	0.33***	(0.29, 0.36)	0.29***	(0.25, 0.33)	0.33***	(0.30, 0.37)	0.29***	(0.25, 0.34)
35+ years	0.26***	(0.22, 0.28)	0.19***	(0.16, 0.23)	0.26***	(0.23, 0.29)	0.20***	(0.17, 0.24)	0.26***	(0.23, 0.29)	0.21***	(0.18, 0.25)
Region of residence												
South(R)												
North					1.39***	(1.22, 1.58)	1.36***	(1.15, 1.60)	1.49***	(1.31, 1.70)	1.41***	(1.19, 1.66)
Central					1.55***	(1.36, 1.75)	1.84***	(1.59, 2.13)	1.61***	(1.41, 1.82)	1.84***	(1.59, 2.13)
East					1.21***	(1.06, 1.38)	1.45***	(1.23, 1.70)	1.19***	(1.05, 1.36)	1.40***	(1.19, 1.64)
Northeast					1.18**	(1.02, 1.38)	1.36***	(1.15, 1.62)	1.17**	(1.01, 1.37)	1.385***	(1.17, 1.64)
West					1.08	(0.92, 1.28)	1.16	(0.96, 1.40)	1.08	(0.91, 1.27)	1.18*	(0.97, 1.43)
Place of residence												
Rural(R)												
Urban					0.86***	(0.78, 0.94)	0.90**	(0.82, 0.98)	0.89**	(0.81, 0.98)	0.94	(0.86, 1.03)
Mother's education												
No education												
1–8 years					0.81***	(0.74, 0.88)	0.85***	(0.77, 0.93)	0.87***	(0.80, 0.95)	0.89**	(0.81, 0.98)
9+ years					0.56***	(0.49, 0.63)	0.58***	(0.51, 0.66)	0.68***	(0.59, 0.78)	0.65***	(0.57, 0.74)
SLI												
Low(R)												
Medium									0.88***	(0.81, 0.95)	0.91*	(0.82, 1.00)
High									0.61***	(0.53, 0.70)	0.77***	(0.67, 0.87)

Note: * significant at 0.05 level, ** significant at 0.01 level, *** significant at 0.001 level, R: reference category.

In these models, mother's age is a highly significant covariate of infant mortality, showing that children born to younger women (15–19-year age group) are more than three times as likely to die before reaching their first birthday as those born to women in the older age groups (20–29 years and 30+). Surprisingly, child's sex fails to emerge as a significant covariate. However, as shown later, an overall measure of infant mortality is deceptive at times for studying the relationship between child's sex and mortality. Gender differences in mortality are more significant in the neonatal period, suggesting that boys are significantly more likely than girls to die in the first month of life, while the converse is usually true in the post-neonatal period.

When we extend the analysis by adding three variables—region of residence, place of residence, and maternal education—the results change substantially (Models 2A and 2B). Region of residence and place of residence carry highly significant coefficients, showing that children from rural areas and from all non-Southern regions, except for the West, are at a greater risk of dying early; however, they do not influence the relationship between caste and infant mortality in a significant way. Consistent with previous research (Basu and Stephenson 2005; Caldwell 1979; Cleland and Van Ginneken 1988; Singh-Manoux et al. 2008), maternal education emerges as a very important predictor. In both NFHS-2 and NFHS-3, children born to women with 9 years or more of schooling are slightly less than half as likely to die in the first year of life as those born to women with no education (NFHS-2: HR = 0.56; CI = 0.49, 0.63 and NFHS-3: HR = 0.58; CI = 0.51, 0.66). Maternal education exerts an important influence on the relationship between caste and infant mortality, while also explaining some of the effects of other control variables. In NFHS-2, the differences in mortality risk between the three lower caste groups and the forward castes are minimized considerably. In NFHS-3, mortality differences between forward-caste, OBC, and Adivasi children are fully accounted for by maternal education, although Dalit children still remain at higher risk (HR = 1.19; CI = 1.06, 1.34). With the addition of SLI, which is also highly correlated with maternal education, the relative risk for Dalit children is further reduced, but still statistically significant (NFHS-2: HR = 1.12; CI = 1.01, 1.23; NFHS-3: HR = 1.17; CI = 1.03, 1.32). This implies that the two socioeconomic variables—maternal education and SLI—are strong predictors of infant mortality, but do not fully explain the excess mortality among Dalits.

Given the limitations of the measure of infant mortality as discussed earlier, we carry out separate analyses for mortality during the neonatal and post-neonatal periods (Table 3). Consistent with previous research (Choe et al. 1995; Modin 2002), we find that girls are significantly less likely than boys to die in the neonatal period, which is indicative of their innate biological survival advantage. However, this is not necessarily true in the post-neonatal period, when the effects of sociocultural milieu overshadow the effects of biology, and girls become increasingly exposed to various societal factors affecting health, illness, and death. This observation holds particularly true in the Indian context (Das Gupta 1987; Subramanian et al. 2006a). We also find that children from rural areas are at greater risk of dying young than those from urban areas; however, they remain more vulnerable in the neonatal period even when other background characteristics are controlled. Somewhat similar findings emerge in the case of region of residence. Children from non-southern regions are significantly more likely than those from southern regions to die young. In contrast to the pattern of relationship between pertinent background characteristics and infant mortality described above, we find that maternal education is more strongly related to post-neonatal than neonatal mortality. In NFHS-3, for example, children born to women with 9 years or more of schooling are at a 29 per cent lower risk of dying in the neonatal period, compared with those born to illiterate women (HR = 0.71; CI= 0.61, 0.82). The gap between the two groups of children is much larger (64 per cent) in the post-neonatal period (HR = 0.36; CI = 0.28, 0.45). SLI is also a more powerful predictor of post-neonatal than of neonatal mortality.

Results presented in Table 3 show that among the lower caste groups, OBCs have made the most impressive gains in reducing infant mortality. In both study waves, the baseline model (Models 1A and 2A) shows that OBC children are slightly more than 20 per cent as likely as forward-caste children to die in the neonatal period (NFHS-2: HR = 1.23; CI = 1.11, 1.36; NFHS-3: HR = 1.24; CI = 1.09, 1.40). When background characteristics (except for SLI) are held constant, this gap is reduced to 16 per cent (HR = 1.16; CI = 1.04, 1.29) in NFHS-2 (Model 3A). However, the gap disappears altogether in the NFHS-3 (Model 3B), suggesting that the excess neonatal mortality among OBC children is entirely attributable to background characteristics—maternal education in particular. The picture is even more dramatic in the post-neonatal period. After adjusting for the above-mentioned variables, there is a trivial mortality gap between OBCs and forward-caste children in both NFHS-2 and NFHS-3.

Table 3. Partial results of Cox Proportional Hazards Model of neonatal and post-neonatal mortality, NFHS-2 and NFHS-3

	NFHS-2				NFHS-3			
	Model 1A		Model 1B		Model 2A		Model 2B	
	Neonatal mortality		Post-neonatal		Neonatal mortality		Post-neonatal	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Caste								
Forward(R)								
Dalits	1.30***	(1.16, 1.46)	1.35***	(1.15, 1.59)	1.39***	(1.21, 1.60)	1.49***	(1.20, 1.85)
Adivasis	1.14**	(1.00, 1.30)	1.53***	(1.30, 1.80)	1.10	(0.94, 1.29)	1.73***	(1.39, 2.14)
OBCs	1.23***	(1.11, 1.36)	1.23***	(1.06, 1.43)	1.24***	(1.09, 1.40)	1.28**	(1.05, 1.55)
Don't know	1.26	(0.85, 1.86)	2.33***	(1.54, 3.52)	1.06	(0.82, 1.39)	1.15	(0.77, 1.71)
Child's sex								
Male(R)								
Female	0.86***	(0.79, 0.94)	1.18***	(1.06, 1.32)	0.87***	(0.79, 0.95)	1.16**	(1.01, 1.33)
Mother's age								
< 20 years								
20–29 years	0.29***	(0.26, 0.34)	0.31***	(0.26, 0.38)	0.26***	(0.22, 0.31)	0.30***	(0.23, 0.40)
35+ years	0.24***	(0.21, 0.28)	0.26***	(0.21, 0.32)	0.17***	(0.14, 0.21)	0.25***	(0.19, 0.34)
Region of residence								
South(R)								
North								
Central								
East								
Northeast								
West								
Place of residence								
Rural(R)								
Urban								
Mother's education								
No education								
1–8 years								
9+ years								
SLI								
Low(R)								
Medium								
High								

Note: * significant at 0.05 level, ** significant at 0.01 level, *** significant at 0.001 level, R: reference category.

Mortality experiences of Dalit children are worse. Not only are they more likely than forward-caste children to die at both stages of infancy, but also appear to have become increasingly vulnerable over time. In the baseline model in NFHS-2, they are at a 30 per cent greater risk in the neonatal period and at a 35 per cent greater risk in the post-neonatal period, compared with forward-caste children. In the NFHS-3, these risks are higher, at 39 per cent and 49 per cent, respectively. The socioeconomic disadvantage of Dalits mirrors these patterns. Thus, when background characteristics are held constant, the gap in post-neonatal mortality risks between Dalit and forward-caste children is substantially reduced in both NFHS-2 and NFHS-3. However, this does not quite happen in the case of neonatal mortality. As shown in Models 3A and 4A, Dalit children remain vulnerable in both the NFHS-2 (HR = 1.18; CI = 1.05, 1.33) and the NFHS-3 (HR = 1.21; CI = 1.05, 1.40). The inclusion of SLI in the model lessens the Dalit effect slightly (Model 5A: HR = 1.12 CI = 1.00, 1.27 and Model 6A: HR = 1.19; CI = 1.03, 1.38).

Surprisingly, Adivasi children are in much better condition than Dalit children in the neonatal period. They are only slightly more vulnerable than forward-caste children in terms of mortality risk in this period, even when no background characteristics are taken into account. However, they are highly vulnerable in the post-neonatal period, when all background characteristics are accounted for. In both surveys, their excess mortality is quite

Table 3. (continued)

	NFHS-2				NFHS-3			
	Model 3A		Model 3B		Model 4A		Model 4B	
	Neonatal mortality		Post-neonatal		Neonatal mortality		Post-neonatal	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Caste								
Forward(R)								
Dalits	1.18***	(1.05, 1.33)	1.17*	(0.99, 1.38)	1.21***	(1.05, 1.40)	1.15	(0.92, 1.43)
Adivasis	1.09	(0.95, 1.26)	1.35***	(1.13, 1.61)	1.01	(0.85, 1.19)	1.29**	(1.03, 1.63)
OBCs	1.16***	(1.04, 1.29)	1.15*	(0.99, 1.34)	1.11	(0.97, 1.26)	1.04	(0.85. 1.27)
Don't know	0.99	(0.66, 1.49)	1.70**	(1.12, 2.59)	1.06	(0.81, 1.39)	1.05	(0.71, 1.57)
Child's sex								
Male(R)								
Female	0.85***	(0.79, 0.93)	1.17***	(1.05, 1.31)	0.86***	(0.79, 0.95)	1.15*	(1.00, 1.32)
Mother's age								
< 20 years								
20–29 years	0.32***	(0.28, 0.36)	0.34***	(0.28, 0.42)	0.27***	(0.23, 0.32)	0.33***	(0.25, 0.44)
35+ years	0.25***	(0.22, 0.29)	0.26***	(0.21, 0.33)	0.18***	(0.15, 0.22)	0.26***	(0.19, 0.35)
Region of residence								
South(R)								
North	1.22**	(1.05, 1.43)	1.83***	(1.44, 2.32)	1.38***	(1.13, 1.67)	1.31*	(0.97, 1.77)
Central	1.40***	(1.21, 1.63)	1.92***	(1.53, 2.42)	1.87***	(1.57, 2.23)	1.78***	(1.36, 2.33)
East	1.11	(0.94, 1.30)	1.48***	(1.16, 1.89)	1.52***	(1.26, 1.84)	1.29*	(0.96, 1.73)
Northeast	1.03	(0.85, 1.24)	1.59***	(1.21, 2.07)	1.26**	(1.02, 1.55)	1.58***	(1.17, 2.13)
West	1.07	(0.88, 1.30)	1.12	(0.82, 1.53)	1.25*	(1.00, 1.56)	0.96	(0.67, 1.39)
Place of residence								
Rural(R)								
Urban	0.85***	(0.76, 0.95)	0.87*	(0.75, 1.02)	0.86***	(0.77, 0.96)	0.99	(0.85, 1.16)
Mother's education								
No education								
1–8 years	0.82***	(0.74, 0.91)	0.79***	(0.69, 0.91)	0.92	(0.82, 1.03)	0.71***	(0.61, 0.84)
9+ years	0.68***	(0.58, 0.78)	0.35***	(0.27, 0.45)	0.71***	(0.61, 0.82)	0.36***	(0.28, 0.45)
SLI								
Low(R)								
Medium								
High								

apparent. After controlling for all background variables, including maternal education and SLI, in both NFHS-2 and NFHS-3 Adivasi children are slightly more than 25 per cent as likely as forward-caste children to die in the post-neonatal period (Model 5B: HR = 1.28; CI = 1.07, 1.53; Model 6B: HR = 1.27; CI = 1.00, 1.60).

Discussion and conclusions

Using data from the latest two waves of India's National Family Health Survey (NFHS-2: 1998–99 and NFHS-3: 2005–06), this study examines the relationship between young children's risk of dying and their caste background. Consistent with previous research (Dommaraju et al. 2008; June et al. 2011; Mohindra et al. 2006; Nguyen et al. 2013; Singh et al. 2013; Singh-Manoux et al. 2008; Subramanian et al. 2006a, 2006b), we find that despite large improvements in health conditions and reductions in mortality in India in recent years, children from lower castes continue to experience higher mortality than those from forward castes. Estimates obtained from the Cox Proportional Hazards Model show that in NFHS-2, with mother's age and child's sex controlled, OBC, Dalit, and Adivasi children are 23, 32, and 27 per cent more likely, respectively, than forward-caste children to die in the first year of life. In NFHS-3, the relative risks remain virtually unchanged for Adivasi and OBC chil-

Table 3. (continued)

	NFHS-2				NFHS-3			
	Model 5A		Model 5B		Model 6A		Model 6B	
	Neonatal mortality		Post-neonatal		Neonatal mortality		Post-neonatal	
	HR	95% CI	HR	95% CI	HR	95% CI	HR	95% CI
Caste								
Forward(R)								
Dalits	1.12*	(1.00, 1.27)	1.10	(0.94, 1.30)	1.19**	(1.03, 1.38)	1.11	(0.89, 1.39)
Adivasis	1.05	(0.92, 1.21)	1.28***	(1.07, 1.53)	0.99	(0.84, 1.18)	1.27**	(1.00, 1.60)
OBCs	1.13**	(1.01, 1.25)	1.12	(0.97, 1.31)	1.11	(0.97, 1.27)	1.04	(0.84, 1.27)
Don't know	0.9	(0.60, 1.36)	1.59**	(1.04, 2.43)	1.11	(0.85, 1.46)	1.03	(0.68, 1.57)
Child's sex								
Male(R)								
Female	0.86***	(0.79, 0.93)	1.16**	(1.04, 1.30)	0.87***	(0.79, 0.96)	1.16**	(1.00, 1.33)
Mother's age								
< 20 years								
20–29 years	0.32***	(0.28, 0.36)	0.36***	(0.29, 0.43)	0.280***	(0.23, 0.33)	0.34***	(0.25, 0.45)
35+ years	0.25***	(0.22, 0.29)	0.27***	(0.22, 0.34)	0.19***	(0.15, 0.23)	0.28***	(0.20, 0.38)
Region of residence								
South(R)								
North	1.29***	(1.10, 1.51)	2.05***	(1.61, 2.60)	1.44***	(1.19, 1.76)	1.33*	(0.98, 1.80)
Central	1.45***	(1.24, 1.68)	2.02***	(1.60, 2.54)	1.88***	(1.57, 2.24)	1.76***	(1.35, 2.31)
East	1.10	(0.93, 1.29)	1.46***	(1.14, 1.86)	1.49***	(1.23, 1.80)	1.22	(0.91, 1.64)
Northeast	1.03	(0.85, 1.24)	1.55***	(1.19, 2.03)	1.30**	(1.06, 1.60)	1.55***	(1.15, 2.10)
West	1.07	(0.88, 1.30)	1.1	(0.81, 1.51)	1.27**	(1.02, 1.59)	0.97	(0.67, 1.40)
Place of residence								
Rural(R)								
Urban	0.87**	(0.78, 0.98)	0.94	(0.80, 1.10)	0.89**	(0.80, 1.00)	1.06	(0.90, 1.26)
Mother's education								
No education								
1–8 years	0.86***	(0.78, 0.96)	0.88*	(0.76, 1.02)	0.96	(0.85, 1.07)	0.77***	(0.65, 0.92)
9+ years	0.78***	(0.67, 0.92)	0.47***	(0.36, 0.62)	0.77***	(0.66, 0.91)	0.41***	(0.32, 0.54)
SLI								
Low(R)								
Medium	0.91**	(0.82, 1.00)	0.83***	(0.73, 0.94)	0.94	(0.83, 1.06)	0.85*	(0.72 ,1.02)
High	0.69***	(0.58, 0.81)	0.48***	(0.37, 0.62)	0.80***	(0.68, 0.93)	0.71***	(0.56, 0.90)

dren, whereas Dalit children are at an enhanced risk (42 per cent). Ironically, these regressive incidents happened at a time when many governments were formed on the basis of caste, politics and various programs and policies were introduced for the uplift of the lower castes. It appears that maternal and child health care programmes such as the Child Survival and Safe Motherhood Programme, and National the Maternity Benefit Scheme, did not fully reach the most vulnerable sections of society.

The research presented here indicates that background characteristics such as maternal age, child's sex, region of residence, rural-urban residence, maternal education, and SLI account for much of the caste differences in infant mortality, although they do not adequately address certain anomalies. When these characteristics—particularly maternal education and SLI—are controlled, differences between OBCs and forward castes in infant mortality in both neonatal and post-neonatal periods are diminished in NFHS-2 and eliminated in NFHS-3. Socioeconomic background characteristics also account for much of the excess mortality among Dalit children, although they still have a 19 per cent higher risk of death (in NFHS-3) in the neonatal period, compared with forward-caste children. It is possible that in addition to individual-level factors, macro-level characteristics such as physical environment, residential segregation, social inequalities and discrimination, which "get under the skin," result in higher neonatal mortality among Dalits, especially those at the bottom of the social and economic hierarchy. Presumably, these factors discourage many Dalits to avail and/or utilize antenatal and prenatal ser-

vices and healthcare facilities for child delivery (Chalasani 2012; McKinnon et al. 2014; Paudel et al. 2013), and also result in stress-related birth outcomes, such as preterm deliveries (Pike 2005) and low birth weight (Ellen 2000). Certain traditional and cultural practices and beliefs associated with pregnancy and childbirth may also be responsible for excess neonatal mortality among Dalits (Ghosh 2012; Kesterton and Cleland 2009).

In stark contrast, Adivasi children are at most risk in the post-neonatal period, even when all background characteristics are taken into account, although they are at a much lower risk in the neonatal period. The reasons for this epidemiological paradox are not clear. It is possible that certain cultural practices among the Adivasis protect newborn children from infections and other factors responsible for deaths in the neonatal period (Das et al. 2010), whereas harsh environmental conditions, malnutrition, lack of medical facilities in remote and rural areas, discriminatory medical practices, and persistent socioeconomic deprivation result in elevated health and mortality risks in later childhood. Underreporting of births and neonatal deaths may also be partially responsible for lower neonatal mortality among the Adivasis compared with other caste groups. These are mere speculations which need to be examined in future studies. In sum, the Dalit experience suggests that programmes and policies that focus on increased access to prenatal and antenatal care and healthcare facilities for child delivery need to be intensified in less advantaged sections of society, whereas the Adivasi experience suggests, however, that these efforts would be incomplete if the post-neonatal stage of life is neglected.

Our results also suggest that a global measure of infant mortality would be misleading to fully understand mortality dynamics in a less industrialized country such as India. Wherever possible, analyses need to differentiate between neonatal and post-neonatal mortality. Since neonatal and post-neonatal deaths are affected by a different set of factors, policymakers need to adopt different strategies to deal with mortality in these two stages of life. We find that in both surveys, Dalit children are significantly more likely than forward-caste children to be at risk of dying in the first year of life. This finding is true in the case of neonatal mortality but not post-neonatal mortality. In fact, mortality risks for Dalit children are not significantly different from forward-caste children's in the post-neonatal period. We also find that the overall infant mortality risks for Adivasi children are no different from forward-caste children. However, the differences between the two groups are significant in the post-neonatal period, but not in the neonatal period.

The evidence presented in this study suggests that including Dalits and Adivasis in a single category hides the fine differences that exist between the two groups in terms of socioeconomic status, health, and mortality. Although historically both groups have been at the bottom of the social hierarchy in India, and have faced continuous discrimination in their social and economic lives, they are highly different from each other due to their dissimilar residential locations and cultural practices. They are also different in terms of the pace of socioeconomic progress over time. Thus, as discussed earlier, while Dalit children are at a greater risk of dying during early infancy, Adivasi children are at a greater risk of dying during later childhood. Further research is needed to identify how sociocultural forces shape individual beliefs and health-related behaviours among the Dalits and Adivasis.

A few comments on the limitations of this study are in order. First, the data used for the analysis are dated. Infant mortality statistics in NFHS-2 refer to approximately the 1994–98 period, while statistics in NFHS-3 refer to the 2002–06 period. They may not capture much of the developments that have taken place in recent years, in terms of reductions in mortality and policy initiatives such as the National Rural Health Mission (NRHM) and Janani Suraksha Yojana (Mother Security Scheme)—a conditional cash transfer programme—which were launched in 2005. Second, the control variables included in this study do not fully capture the socioeconomic status that may be responsible for caste differences in mortality. If we were to include some macro-level variables (such as neighbourhood, residential segregation), various cultural practices, and programmes and policies, caste differences in mortality may have been more adequately addressed. Third, data on the background variables included in this study were measured on the survey date, while the dependent variable—infant mortality—was derived from the number of children who were born during the five years preceding the survey. However, we do not expect this to be an important bias, because none of the characteristics (except perhaps for SLI) is likely to be influenced by infant mortality. Fourth, data on infant mortality in a less industrialized country such as India is suspect to recall lapse, failure to report a sad event, and, consequently, underreporting of deceased children. This may have happened more among the Adivasis, who report significantly lower neonatal mortality

than Dalits, much closer to that among OBCs. Finally, the four caste groups included in this study present just a global picture. In each group, there are hundreds of castes, some of which are better-off and less discriminated than others. Among the Dalits, for example, some lower castes have been recently identified as "Maha Dalits" (or extremely oppressed) or "extremely backward", while some others are in much better conditions in certain geographic regions. Hopefully, the 2011 Census of India, which collected data for all castes, will enable researchers and policymakers to go beyond simplistic generalizations based on broad categories of caste.

Acknowledgment

Acknowledgment is due Shefali S. Ram, A.K. Sharma, R.D. Sharma, Satish Sharma, and two anonymous referees for their comments on earlier versions of the paper.

References

Antai, D. 2011. Inequalities in under-5 mortality in Nigeria: Do ethnicity and socioeconomic position matter? *Journal of Epidemiology* 21:13–20.

Baru, A.B., S. Acharya, A.K. Shiva Kumar, and K. Nagraj. 2010. Inequities in access to health services in India: Caste, class, and religion. *Economic and Political Weekly* 45:49–38.

Basu, A.M., and R. Stephenson. 2005. Low level of maternal education and the proximate determinants of childhood mortality: A little learning is not a dangerous thing. *Social Science and Medicine* 60:2011–23.

Bicego, G.T., and J.T. Boerma. 1993. Maternal education and child survival: A comparative study of survey data from 17 countries. *Social Science and Medicine* 36:1207–27.

Bratter, J., and B. Gorman 2011. Is discrimination an equal opportunity Risk? Racial experiences, socioeconomic status, and health status among black and white adults. *Journal of Health and Social Behavior* 52:365–82.

Brown, T.H., A.M. O'Rand, and D.E. Adkins. 2012. Race-ethnicity and health trajectories: Tests of three hypotheses across multiple groups and health outcomes. *Journal of Health and Social Behavior* 53:359–77.

Caldwell, J.C. 1979. Education as a factor in mortality decline: An examination of Nigerian data. *Population Studies* 33:395–413.

Chalasani, S. 2012. Understanding wealth-based inequalities in child health in India: A decomposition approach. *Social Science and Medicine* 75:2160–69.

Choe, M.K., H. Ho, and F. Wang 1995. Effects of gender, birth order, and other correlates on childhood mortality in China. *Social Biology* 42:50–64.

Cleland, J.G., and J.K. Van Ginneken. 1988. Maternal education and child survival in developing countries: The search for pathways of influence. *Social Science and Medicine* 27:1357–68.

Crimmins, E.M., M.D. Hayward, and T.E. Seeman. 2004. Race/ethnicity, socioeconomic status, and health, in *National Research Council (US) Panel on Race, Ethnicity, and Health in Later Life*, edited by N.B. Anderson, R.A. Bulatao, and B. Cohen. Washington, DC: National Academic Press.

Das, M.B., S. Kapoor, and D. Nikitin. 2010. A Closer Look at Child Mortality among Adivasis in India. Policy Research Working Paper 523. The World Bank, South Asia Region, Human Development Department.

Das Gupta, M. 1987. Selective discrimination against female children in rural Punjab, India. *Population and Development Review* 13:77–100.

Desai, S., and A. Dubey. 2012. Caste in 21st-century India: Competing narratives. *Economic and Political Weekly* 46:40–49.

Desai, S., and V. Kulkarni 2008. Changing educational inequalities in India in the context of affirmative action. *Demography* 45:245–50.

Dommaraju, P., V. Agadjanian, and S. Yabiku. 2008. The pervasive and persistent influence of caste on child mortality in India. *Population Research and Policy Review* 27:477–95.

Ellen, I.G. 2000. Is segregation bad for your health? The case of low birth weight, in *Brookings-Wharton Papers on Urban Affairs*, edited by W.G. Gale and P. Rothenberg, Jr. Washington, DC: Brookings Inst. Press, p. 203–29.

Fenelon, A. 2013. Revisiting the Hispanic paradox in the United States: The role of smoking. *Social Science and Medicine* 82:1–9.

Ghosh, R. 2012. Child mortality in India: A complex situation. *World Journal of Pediatrics* 8:11–18.

Hayward, M.D, E.M. Crimmins, T.P. Miles, and Y. Yu. 2000. The significance of socioeconomic status in explaining the racial gap in chronic health conditions. *American Sociological Review* 65:910–30.

Hummer, R.A. 1996. Black-white differences in health and mortality: A review and conceptual model. *Sociological Quarterly* 37:105–25.

IIPS (IIPS and Macro International). 2000. *National Family Health Survey (NFHS-2), 1998–99: India.* Vol. 1. Mumbai: International Institute for Population Sciences.

———. 2007. *National Family Health Survey (NFHS-3), 2005–06: India.* Vol. 1. Mumbai: International Institute for Population Sciences.

June, Y., T. Po, and S.V. Subramanian. 2011. Mortality burden and socioeconomic status in India. *PLoS*. DOI: 10.1371.

Kesterton, A.J., and J. Cleland. 2009. Neonatal care in rural Karnataka: healthy and harmful practices, the potential for change. *BMC Pregnancy and Childbirth* 9:20 (Published online 20 May 2009, doi: 10.1186/1471-2393-9-20).

Kumar, S., M.S. Alam, and D. Joshi. 2008. Caste dynamics and political process in Bihar. *Journal of Indian School of Political Economy* 20:1–32.

Lahiri, S., A. Hazra, and A. Singh 2011. Sex differentials in childhood mortality in Punjab and Haryana: Are they reality? *Population Studies* 43:71–98.

Lawn, J.E., S. Cousens, and The Lancet Neonatal Survival Steering Team. 2005. 4 Million Neonatal Deaths: When? Where? Why? www.thelancet.com (March).

Majid, A. 2012. Future of untouchables in India: A case study of Dalits. *South Asian Studies* 27:263–85.

Mathews, T.J., and M.F. MacDorman. 2013. Infant mortality statistics from the 2010 period linked birth/infant death data set. *National Vital Statistics Reports.* 62(8):NVSS.

McKinnon, B., S. Harper, J.S. Kaufman, and Y. Bergevin. 2014. Socioeconomic inequality in neonatal mortality in countries of low and middle income: A multicountry analysis. Lancet, 2, e165-e173 (DOI: http://dx.doi.org/10.1016/S2214-109X(14)70008-7).

Modin, B. 2002. Birth order and mortality: A life-long follow-up of 14,200 boys and girls born in the early 20th century Sweden. *Social Science and Medicine* 54:1051–64.

Mohindra, K.S., S. Haddad, and D. Narayana. 2006. Women's health in a rural community in Kerala, India: Do caste and socioeconomic position matter? *Journal of Epidemiology and Community Health* 60:1020–26.

Nayar, K.R. 2007. Social exclusion, caste, and health: A review based on social determinants framework. *Indian Journal of Medical Research* 126:355–63.

Nguyen, K.H., E. Jimenez-Soto, P. Dayal, and A. Hodge. 2013. Disparities in child mortality trends: What is the evidence from disadvantaged states in India? The Case of Orissa and Madhya Pradesh. *International Journal for Equity in Health* 12:45–59.

Pandey, A., M.K. Choe, N.Y. Luther, D. Sahu, and L. Chand. 1998. *Infant and Child Mortality in India.* National Family Health Survey Subject Reports 11. Mumbai: IIPS and Honolulu:East-West Center.

Paudel, D., I.B. Shrestha, M. Siebeck, and E. Rehfuess. 2013. Neonatal health in Nepal: Analysis of absolute and relative inequalities, and impact of current efforts to reduce neonatal mortality. *BMC Public Health* 3:1239.

Pike I.L. 2005. Maternal stress and fetal responses: Evolutionary perspectives on preterm delivery. *American Journal of Human Biology* 17:55–65.

Ram, B., A. Singh, and A. Yadav. 2013. Regional disparity in infant mortality in India. *Demography India* 42:35–54.

Ruiz, J.M., P. Steffen, and T.B. Smith. 2013. Hispanic mortality paradox: A systematic review and meta-analysis of the longitudinal literature. *American Journal of Public Health* 103: e1–9.

Sharma, V., J. Katz, L.C. Mullany, S.K. Khatry, S.C. LeClerg, S.R. Shrestha, G. Darmstadt, and J.M. Tielsch. 2008. Young aternal age and the risk of neonatal mortality in rural Nepal. *Archives of Pediatric and Adolescent Medicine* 162:828–35.

Shryock, H.S., and J.S. Siegel and Associates. 1973. *The Methods and Materials of Demography.* Vol. 2. Washington, DC: U.S. Census Bureau.

Singh, A., A. Kumar, and A. Kumar. 2013. Determinants of neonatal mortality in rural India, 2007–2008. *Peer J.* 1:e75.

Singh-Manoux, A., A. Dugravot, G.D. Smith, M. Subramanyam, and S.V. Subramaian. 2008. Adult education and child mortality in India: The influence of caste, household wealth, and urbanization. *Epidemiology* 19:294–301.

Subramanian, S.V., S. Nandy, M. Irving, D. Gordon, H. Lambert, and G.D. Smith. 2006a. The mortality divide in India: The differential contributions of gender, caste, and standard of living across the life course. *American Journal of Public Health* 96:818–25.

Subramanian, S.V., G.D. Smith, and M. Subramanyam. 2006b. Indigenous health and socioeconomic status in India. *PLoS Med* (10): e421. doi: 10.1371/journal.pmed.0030421.

Ulizzi, I., and I.A. Zonta. 2002. Sex differential patterns in perinatal deaths in Italy. *Human Biology* 74:879–88.

Vikram, K., R. Venneman, and S. Desai. 2012. Linkages between maternal education and childhood immunization in India. *Social Science and Medicine* 75:331–39.

Williams, D.R. 1999. Race, socioeconomic status, and health: The added effects of racism and discrimination. *Annals of New York Academy of Sciences* 998:173–88.

Williams, D.R., and C. Collins. 1995. US socioeconomic and racial differences in health: Patterns and explanations. *Annual Review of Sociology* 21:349–86.

Williams D.R., and M. Sternthal. 2010. Understanding racial-ethnic disparities in health: Sociological contributions. *Journal of Health and Social Behavior* 51 (Suppl):S15–27.

Canadian Studies in Population 43(3–4): 264–71.

Data and discrimination: A research note on sexual orientation in the Canadian labour market

Nicole Denier[1]
Sean Waite

Abstract

Growing interest in the labour market outcomes of sexual minorities presents novel methodological and theoretical challenges. In this note, we outline important challenges in the study of wage inequality between sexual minorities and heterosexuals in Canada. We discuss the current state of available data on sexual orientation and economic outcomes in Canada, and further evaluate how estimates of sexual orientation wage gaps differ across earnings definition and sample composition. Our analysis of the 2006 Census shows considerable heterogeneity in point estimates of wage disadvantage across definitions of earnings and sample selections; however, all estimates show that gay men suffer labour market penalties and lesbians experience wage premiums.

Keywords: sexual orientation; earnings; gay; lesbian; Canada.

Résumé

L'intérêt grandissant pour la situation des minorités sexuelles sur marché du travail soulève de nouveaux enjeux méthodologiques et théoriques. Dans ce commentaire, nous soulignons les enjeux importants que présente l'étude des inégalités salariales entre minorités sexuelles et hétérosexuels au Canada. Nous discutons de la disponibilité actuelle de données sur l'orientation sexuelle et le revenu au Canada et évaluons la manière selon laquelle les écarts salariaux varient en fonction de la définition de revenu et la composition de l'échantillon. Notre analyse du recensement de 2006 indique une hétérogénéité considérable des estimations ponctuelles de l'écart salarial à travers différentes définitions de revenu et différentes sélections d'échantillon. Cependant, toutes les estimations indiquent que les hommes gays sont désavantagés sur le marché du travail et que les lesbiennes obtiennent des salaires supérieurs.

Mots-clés : orientation sexuelle; revenu; gays; lesbiennes; Canada.

Interest in the labour market outcomes of gay men and lesbian women has grown considerably over the past decade. While previous research was limited to small convenience samples, new population data at last includes information on sexual minorities in Canada. With this data, researchers have begun enumerating previously undocumented aspects of labour market stratification by sexual orientation, including the presence of wage disparities between gay men, lesbians, and heterosexual men and women. With few exceptions, the growing international literature has generally found that gay men earn less than heterosexual men and lesbians earn more than heterosexual women, but still less than all men (see Klawitter 2015 for a review and meta-analysis of this research).[2] This expanding field presents novel methodological and theoretical challenges that pertain to studying the populations at hand.

In this research note, we outline important challenges in the study of wage inequality between sexual minorities and heterosexuals in Canada. We first outline the current state of available data on sexual orientation and economic outcomes in Canada. We then turn to evaluating how estimates of sexual orientation wage gaps differ across definitions of earnings that are consistent with differences in earnings variables provided in available data sources. We discuss what divergent results across definitions of earnings and sample selection criteria indicate for the performance of sexual minorities in the Canadian labour market and how they relate to the conclusions reached in recent research. Our analysis shows considerable heterogeneity in point estimates of wage disadvantage across definitions of earnings and sample selections; however, all estimates are consistent with the growing international literature finding that gay men suffer labour market penalties and lesbians experience wage premiums. We argue that Census data provides a robust estimate of wage inequality, as it offers a sufficiently large sample of sexual minorities. Unfortunately, the Census remains limited, due to its failure to identify unmarried LGBTQ persons.

Data deficiencies

Identifying sexual minorities

Three sources of population data have been used to study earnings differences between sexual minorities and heterosexual Canadians: the General Social Survey (GSS), Canadian Community Health Survey (CCHS), and the Census. While each provides

1. Corresponding author: Nicole Denier, Colby College, 4000 Mayflower Hill Dr., Waterville, ME, USA 04901; email: nicole.denier@colby.edu. Sean Waite, Memorial University, St. John's, NL.
2. Notable exceptions include Carpenter (2005) using the California Health Interview Survey and Mueller (2014) drawing from the Canadian General Social Survey.

information on the sexual orientation of respondents, measures of sexual orientation vary across the surveys, and in ways that affect which sexual minorities are identified. The GSS, CCHS, and the Census all allow researchers to identify sexual orientation through partnership status. Namely, gay men and lesbian women are identified by their common-law or marital partnership with a person of the same self-identified sex. This method excludes unmarried people and ignores bisexuality. The CCHS and GSS further ask a direct question on sexual identity, asking the respondent, "Do you consider yourself to be… (1) heterosexual (sexual relations with people of the opposite sex), (2) homosexual, that is, lesbian or gay (sexual relations with people of your own sex); [or] (3) bisexual (sexual relations with people of both sexes)." While the GSS and CCHS characterize the question as one of identity and not behaviour, the wording of the clarifications defines identity through sexual practice/behaviour (Carpenter 2008). Previous research shows that individuals are more likely to report same-sex sexual behaviour rather than a same-sex identity (Badgett 2009).[3] Thus, the question, in addition to including single gay men and women, may also include individuals who may not necessarily identify as gay but do so because they have engaged or continue to engage in occasional same-sex sexual relations. Conversely, it may exclude people who have only ever engaged in same-sex sexual relations but nevertheless identify as heterosexual.

Why do these definitional differences matter? In the case of the Census, clearly a major issue is that information on unmarried sexual minorities is lost—and inferences about the sexual minority population as a whole are then trickier, requiring some information about how differences in selectivity into partnership vary by sexual orientation and how this selectivity may relate to earnings (Carpenter 2008). On the other hand, using partnership status may offer some advantages (Klawitter 2015). Those who are in long-term same-sex relationships may be less willing and/or less able to conceal their sexual orientation. Some individuals in same-sex partnerships may also have incentives to disclose their sexual identity, in order to receive employment-provided fringe benefits, like dental insurance, for their partner. If discrimination is a key mechanism producing earnings disadvantage, individuals must somehow reveal their sexual orientation to bosses and co-workers. Single people have less of an incentive to disclose their identity in the workplace, and individuals that engage occasionally in same-sex sexual relations may not convey a non-heterosexual identity at all. Thus, there is reason to believe that across surveys, the population identified varies, and in ways that shape the mechanisms which impact earnings disadvantage.[4]

Earnings three ways

Both the couples approach and the GSS/CCHS question identify important aspects of sexual orientation that may be relevant in influencing labour market outcomes. But what both the CCHS and GSS definitely do not measure is earnings, instead providing (at times crude) indicators of income. The GSS contains a categorical variable on total individual income from all sources, ranging in values of 1–12 and representing incomes of "no income or loss" up to "$100,000 or more." The top-coding of income will curtail wage disadvantage if high-earning heterosexuals earn more than high-earning sexual minorities; this pattern is documented in Waite and Denier (2015). The CCHS, on the other hand, provides a continuous income variable for some respondents; those who do not provide an exact value are then probed with a series of categories that their income may fall into. The CCHS income variable is tangential to the main aims of the health survey, and survey documentation warns that it should be used with caution and as a control variable. Thus, studies using the CCHS income variable are prone to measurement error in the dependent variable. It is not clear if these errors vary across sexual orientation, but if this is the case, estimates of wage gaps may be biased. The Census, however, offers fairly high-quality earnings and income data. Starting in 2006, Canadians had the option of linking their Census responses to their tax records, with over 80 per cent of respondents allowing the linkage (Statistics Canada 2008). For those who did not give permission, the questionnaire asked about detailed income components, divided into various sources to facilitate accurate recall (Statistics Canada 2008). Thus, the Census data offers superior data on both earnings and income.

Using individual income to study earnings also poses an issue, as it often includes non-wage income sources; paramount among them are government transfers, which ultimately depend on family relationships. There is wide variability in the receipt of government transfers across the earnings distribution (Heisz 2007), which may systematically impact the income of gay and heterosexual individuals. Given that lower-wage workers tend to have lower incomes at similar hours worked, a larger portion of low-wage workers will receive income that is not wage and salary income, inflating the "earnings" of low earners relative to high earners in the sample. Depending on the program, government transfers may be means-tested to total household income. Thus, the lower-earning partner in a high-earning household may have lower individual income than the same low earner in a low-earning household. For example, lesbians may be eligible for more transfers—given they are partnered with another woman, who in general is paid less in the Canadian labour market—than a woman would be who is partnered with a high-earning man (married heterosexual women earn less, but their husbands earn more). Lesbian incomes may be inflated precisely because their position in a lower-earning all-female household allows them to qualify for more non-wage income. At the same time, heterosexual women are more likely to receive transfers that are targeted at families with children, since coupled heterosexual women

3. In 2009 the Sexual Minority Assessment Research Team (SMART) at the Williams Institute UCLA collaborated with over twenty-five experts to prepare *Best Practices for Asking Questions about Sexual Orientation on Surveys* (see Badgett 2009).
4 Klawitter's (2015) meta-analysis of studies mainly from the U.S. indicates that wage penalties for gay men identified by behavior are actually slightly larger than those estimated from a sample of couples.

are more likely to have children than coupled lesbians (Waite and Denier 2015).[5] Thus, reliance on an income variable may bias estimates of earnings differences by either understating or overstating actual labour market earnings.

Current estimates

Using these data sources, five studies have provided evidence of an earnings gap for sexual minorities in Canada.[6] Table 1 presents findings from the most fully specified model in each of the three studies, and reports estimates for coupled sexual minorities relative to coupled heterosexuals.[7] At first glance, the estimates vary widely across studies (part of the impetus for formulating this note). Two papers draw samples of couples to estimate sexual minority earnings disadvantage. Mueller (2014) examines the 2006–10 GSS, limiting his sample to those likely earning most of their total income, and finds no wage disadvantage for gay men but a large wage advantage of about 16 per cent for lesbian women. Waite and Denier (2015), on the other hand, find using Census data that gay men earn 5.1 per cent less than heterosexual men and lesbians earn 8 per cent more than heterosexual women. The other three studies, using the CCHS, include both singles and individuals in couples. Carpenter (2008) acknowledges that he is examining income and includes a broad sample, taking care to avoid relating income differences directly to labour market processes. He finds that on average, gay men have total incomes that are about 12 per cent lower than heterosexual men, while lesbians have incomes about 15 per cent higher. Carpenter (2008) further shows that these wage differences are larger when restricting the sample to those in couples, with the income penalty for gay men around 20 per cent for partnered gay men relative to partnered heterosexual men, and the income advantage at about 43 per cent for lesbians relative to coupled heterosexual women. LaFrance, Warman, and Wooley (2009) are primarily interested in how wage differentials vary across partnership status. Limiting their CCHS sample to individuals who work 30+ hours a week, they find that gay men in a married/common-law relationship make about 20 per cent less than married (not cohabiting) straight men, while lesbian women in marital/common-law relationships make about 10 per cent more than married straight women. Single gay men make about 24 per cent less than married heterosexual men, while single heterosexual men make about 14 per cent less. Single lesbians and single heterosexual women both have incomes that are about 10 per cent higher than married heterosexual women. These differences hold even when including only people whose main source of income is wages and salaries or who only receive income from wages and salaries. Cerf (2016) uses the 2000–09 CCHS and finds that partnered gay men have incomes about 13 per cent less than heterosexual men, while partnered lesbian women have incomes about 8 per cent higher than partnered straight women. Notably, in contrast to LaFrance et al. (2009) and consistent with Carpenter (2008), he finds no wage difference for single gay men and women. The variability in estimates for gay men suggests substantially different conclusions, ranging from no disadvantage to considerable earnings gaps, even after accounting for work effort and occupation and industry choice. For lesbians, the magnitude of the wage advantage over heterosexual women also remains unclear. In the following section, we attempt to uncover some of the sources of these differences.

Reconciling results

The available data present challenges to identifying sexual minority earnings gaps, as evidenced by the breadth of previous findings. Perhaps the single greatest challenge is that most surveys measuring sexual orientation do not measure earnings (or, conversely, most high-quality labour market studies do not measure sexual orientation). A second practical challenge is that researchers often specify different analytic samples, making it difficult to pinpoint whether it's the data or the sample that is driving the result. In order to better understand how sample composition and variable definitions impact estimates of sexual minority wage gaps, we use couples data from the 2006 Census to replicate the sample selection criteria and earnings/income variables used in some previous research.

We are interested in two types of comparisons: across variable definition and across sample specification. For variable definitions, we are primarily focused on how different income and earnings concepts change our understanding of pay (dis)advantage. We generate annual and hourly earnings variables, which directly reflect labour market processes. We further examine hourly and annual income measures, like those that would be found in the CCHS.[8] Finally, we generate a series of "discrete" income and earnings measures that reflect the type of imputation strategy required in surveys that have categorical measures of income, like the GSS. We follow Mueller (2014) and take the midpoint of the 12 income categories available in the GSS (from $0 to $100,000+) and assign those values to the corresponding continuous income/earnings levels of our respondents in the Census. We calculate average earnings/income differences between coupled gay men and coupled heterosexual men, and between coupled lesbian women and coupled heterosexual women, using OLS regressions with robust standard errors. We control for age, education, potential work experience, common-law status, presence of children in the household, rural residence, and province of residence. For annual earnings and income models, we further control for weeks worked and part-time status. We also present models controlling

5. Practically, the transfers are commonly assigned to the adult female in the household.
6. While Carpenter (2008) sheds light on the economic situation of gay and lesbian Canadians, he is focused on identifying income differences, and thus does not comment on labour market dynamics, particularly discrimination.
7. Waite (2015) explored whether sexual minority wage gaps attenuated between 2001 and 2011 using Canadian census and survey data. We do not include this study in our table since the sample, methodologies and point estimates are comparable to Waite and Denier (2015).
8. This is not to say that the CCHS variable will be as high-quality as that in the Census, which is drawn largely from tax data.

Table 1. Sexual orientation wage gaps in Canada

Author	Year	Data	Dependent variable	Same-sex definition	Sample specification	Coupled gay pay gap	Coupled lesbian pay gap
Carpenter	2008	CCHS (2003–05)	Annual income	Identifies as gay/lesbian	Aged 18–55	−0.115 (single & coupled) −0.210 (coupled)	0.154 (single & coupled) 0.359 (coupled)
LaFrance, Warman, and Wooley	2009	CCHS (2003–07)	Annual income	Identifies as gay/lesbian	Aged 25–59, works 30+ hrs/ week	−0.217	0.100
Cerf	2016	CCHS (2003–09)	Hourly income: annual income divided by 50 × hrs worked/wk	Identifies as gay/lesbian and member of a same-sex household	Aged 18–65, Canadian-born, non-Aboriginal, not self-employed full-time, and not bisexual	−0.130	0.079
Mueller	2014	GSS (2006–10)	Hourly income: mid-point values of each category for personal income divided by annual hrs worked	Same-sex partnership	Aged 20–60, not attending school full-time, earned between $5 and $500/hr and claimed employment or self-employment income as main source of income	−0.060 (n.s.)	0.163
Waite and Denier	2015	Census (2006)	Annual earnings	Same-sex partnership	Aged 25–64, Canadian-born, non-Aboriginal, non-visible minority, annual earnings $1000+, working for wages and salaries	−0.051	0.079

occupation and industry of employment in Appendix A. Occupation is coded using the National Occupational Classification for Statistics major groups, and industry is coded using the North American Industry Classification System at the sector level. Appendix B further presents annual earnings and income differences unadjusted for labour supply.

We compare these earnings/income differences across two samples. With our data, we are only able to address differences across a coupled sample, and therefore focus on the samples of two studies that use data on couples. The first replicates Waite & Denier (2015), examining a sample of Canadian-born, non-visible minority, non-aboriginal employees between the ages of 25 and 64, with at least $1,000 in annual earnings. The second approximates Mueller (2014) by focusing on a sample aged 20–60, not in school, with discrete hourly incomes between $5 and $500. The major differences across the samples will reflect the combined effect of changes in age and the inclusion of immigrant and aboriginal populations. We are interested in this combined impact, particularly as Mueller (2014) reports no significant wage disadvantage for gay men using a coupled sample, a finding that contradicts previous research.

Taken together, our comparisons reconcile divergent findings in two previous studies that identify sexual orientation through partnership with a member of the same sex (Mueller 2014; Waite and Denier 2015). They further illustrate how studies using an income variable, like those that draw on the CCHS (LaFrance, Warman, and Wooley 2009; Cerf 2016), relate to estimates using an earnings variable. We cannot directly reconcile the results of all the previous studies reported in Table 1, as we lack a data source that includes a measure of sexual orientation (and thus identifies both singles and couples), as well as both earnings and income data. Making a direct comparison between estimates from our coupled sample and estimates derived from both singles and couples would be imprudent; the populations potentially differ in ways we are unable to quantify. This would not only directly impact the average wage differences between sexual minorities and heterosexuals, but could also potentially indirectly impact estimates by modifying the relationship between important control variables and sexual orientation wage gaps.[9] Instead, we focus on how the use of an income variable may generally affect the conclusions drawn in those studies (LaFrance, Warman, and Wooley 2009; Cerf 2016).

Tables 2 and 3 present estimates by variable definition and sample specification for gay men compared to heterosexual men, and lesbian women compared to heterosexual women. Our results suggest that the definition of earnings used introduces nuanced differences in the estimates. Comparing first earnings and income concepts in a single sample across the row (using the Waite and Denier 2015 sample), for gay men both the annual and the hourly income disadvantage is larger than the annual and

9. This could be important if, for instance, the impact of variables like age or education on the wages of sexual minorities and heterosexuals varies based on their relationship status. For example, it may be that older heterosexual men who remain single possess characteristics that make them both less attractive partners and less attractive workers, weakening the positive relationship between age/potential experience and earnings. Older gay men who remain unmarried may have done so as a result of discriminatory barriers uncorrelated with their productive capabilities. Such compositional changes to the sample would yield a lower pay gap for gay men. Comparing estimates drawn from a sample of singles and couples to one drawn only from couples would not be able to identify these types of differences—specifically, whether it is due to the changing composition of heterosexuals or sexual minorities present in the sample.

Table 2. Estimates of wage gaps for coupled gay men relative to coupled heterosexual men

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income	Continuous hourly income	Discrete hourly income	Continuous hourly earnings	Discrete hourly earnings
a. Waite & Denier (2015)								
	−0.104***	−0.087***	−0.126***	−0.110***	−0.097***	−0.080***	−0.076***	−0.059***
	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)
b. Mueller (2014)								
1. No additional controls	−0.114***	−0.089***	−0.124***	−0.101***	−0.093***	−0.066***	−0.093***	−0.067***
	(0.012)	(0.013)	(0.010)	(0.008)	(0.010)	(0.008)	(0.012)	(0.013)
2. + Additional controls	−0.097***	−0.072***	−0.104***	−0.084***	−0.075***	−0.051***	−0.076***	−0.051***
	(0.012)	(0.013)	(0.010)	(0.008)	(0.010)	(0.008)	(0.012)	(0.013)

Notes: *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

Table 3. Estimates of wage gaps for coupled lesbian women relative to coupled heterosexual women

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income	Continuous hourly income	Discrete hourly income	Continuous hourly earnings	Discrete hourly earnings
a. Waite & Denier (2015)								
	0.093***	0.095***	0.083***	0.080***	0.062***	0.058***	0.075***	0.077***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.010)	(0.010)	(0.010)	(0.010)
b. Mueller (2014)								
1. No additional controls	0.066***	0.058***	0.071***	0.069***	0.042***	0.037***	0.043***	0.035**
	(0.011)	(0.012)	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)	(0.012)
2. + Additional controls	0.070***	0.063***	0.076***	0.074***	0.048***	0.041***	0.040***	0.040**
	(0.011)	(0.012)	(0.011)	(0.008)	(0.011)	(0.008)	(0.010)	(0.012)

Notes: ** P ≤ .01; *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

hourly earnings disadvantage. This means that using an income variable, like that available in the CCHS, would likely overstate gay men's earnings disadvantage. The second important definitional distinction is that between continuous measures and discrete measures based on an imputation of categorical income measures, like those available in the GSS. In every instance, the discrete earnings/income measure understates the wage disadvantage of gay men. This is likely a result of top coding in the dependent variable, and suggests that a meaningful portion of the gay wage penalty and lesbian wage advantage emerges at the top of the earnings distribution. This is consistent with the larger wage disadvantage that Waite and Denier (2015) observed for gay men in the tenth percentile of the wage distribution. For lesbian women, on the other hand, differences in the discrete and continuous earnings/income measures are not large. Nevertheless, like for gay men, there are differences in the magnitude of advantage across earnings and income measures; the use of income rather than earnings actually understates the lesbian wage premium observed in the labour market.

We then turn to differences across sample specification; here we compare estimates across the samples for similar earnings concepts (i.e., a comparison down the column). We focus on the two dependent variables used by Waite and Denier (2015) and Mueller (2014): continuous annual earnings and discrete hourly income, respectively. For gay men, what is striking is the sensitivity of the results to the inclusion/exclusion of the aboriginal and immigrant populations, depending on the measure of earnings. The Mueller (2014) sample produces slightly lower estimates of continuous annual earnings, but a full 3 percentage point difference in the pay gap based on the discrete hourly income measure. Considering a broader range of measures, the lower wage disadvantage reported using the Mueller (2014) seems to obtain in particular with measures of income and with measures based on discrete transformations of the variable. Notably, the models with additional controls account for both aboriginal group membership and immigration status, suggesting that these groups may be underrepresented among coupled gay men, or that they may modify the impact of other control variables in mediating the relationship between sexual orientation and earnings. These changes in estimates across the samples, conditional on the dependent variable, may help explain the null finding reported in Mueller (2014). The results for lesbians, in Table 3, similarly vary by sample specification. However, for all measures, the lesbian wage advantage is lower using the Mueller (2014) sample than the Waite and Denier (2015) sample—opposite the pattern

in the published studies, in which Waite and Denier (2015) reported a lower lesbian wage advantage. This suggests that perhaps the couples in the GSS and Census are qualitatively different.

Appendix A presents the results controlling for occupation and industry of employment, two important mechanisms accounting for observed (dis)advantage for gay men and lesbians, as documented in previous studies. The general shape of changes across definition and sample remains true when controlling for occupation and industry. For gay men, income measures tend to generate larger earnings disadvantage than earnings measures. And again, when using the Mueller (2014) type sample and an income measure, the wage disadvantage of gay men is lower—by more than half when considering discrete hourly income, and reduced to non-significance and approaching zero when examining discrete hourly earnings. For lesbian women, with controlling for occupation and industry, differences between discrete and continuous measures are not large. However, the controls do seem to reduce differences across the samples, particularly for the annual measures of income and earnings, suggesting that some unique aspects of the two samples may be accounted for with observable differences.

Conclusion

Recent research has generated interesting and important questions about the role of sexual orientation in labour market outcomes. This research has also generated a wide range of estimates of the wage penalties for gay men relative to heterosexual men and the wage premiums for lesbian women relative to heterosexual women in Canada. In this note, we provided evidence on the likely sources of some of these disparities, and our findings point to a few key reasons.

First, the use of total income rather than earnings can distort pictures of earnings inequality. For unadjusted estimates, a continuous income variable (like that in the CCHS) produces larger, although consistent, estimates of earnings disparities. However, the relationship between key explanatory variables, particularly occupation and industry, varies considerably across the income and earnings specification for gay men. Total income that is top coded (like that in the GSS) introduces larger distortions, particularly for gay men, because an important part of the gay pay penalty emerges at the top of the earnings distribution.

Second, using a younger sample that also includes the aboriginal and immigrant populations is associated with a lower estimate of the gay pay gap. This may be for a number of reasons. Immigrants, in particular, are less likely in the Census sample to be in same-sex couples than native-born Canadians (this may be for a variety of reasons, including previous immigration rules that may not have allowed same-sex couples to migrate together or cultural/religious intolerance towards homosexuality within certain immigrant communities). Immigrants are also more likely than the native-born to earn less at similar levels of education and potential work experience; thus including the immigrant population may lower the average wages of the heterosexual population relatively more than the gay population. Similarly, a younger sample may lead to lower estimates of disadvantage for gay men, particularly if older gay men gained much of their labour market experience during a time when there was less social acceptance of the LGBTQ community. Waite (2015), however, documents a larger earnings penalty for *younger* gay men. He offers that older gay men may have been more likely to conceal their sexual orientations in the past, perhaps subjecting them to less overt discrimination.

Taken together, this analysis helps to reconcile some differences observed in previous estimates. Specifically, Mueller's (2014) finding that there is no wage gap for gay men is likely influenced by low sample sizes in the GSS and a combination of sample and variable definition—difficulties unique to studying the population at hand. For example, estimates for gay men are more sensitive to sample specification than are those for lesbian women. Consideration of such issues will benefit future research.

Moreover, we outlined important deficiencies in identifying sexual minorities in population-based data in Canada. The CCHS and GSS do well in providing a question on sexual orientation. Yet, the question conflates sexual identity and sexual behaviour (Carpenter 2008). This distinction is important in identifying the mechanisms that may lead to disadvantage, particularly in teasing out whether it is choice or constraint that is responsible for leading to observed differences. Additionally, for many substantive outcomes of interest, these surveys do not provide a large enough sample of sexual minorities. The Census, which does provide large sample sizes of sexual minorities, does not ask a question about sexual identity. This omission continues to limit research on the economic lives of gay and lesbian Canadians.

References

Badgett, M.V.L. 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys*. Los Angeles: The Williams Institute, University of California Los Angeles. (http://escholarship.org/uc/item/706057d5)

Carpenter, C.S. 2005. Self-reported sexual orientation and earnings: evidence from California. *Industrial and Labor Relations Review* 58(2):258–73.

———. 2008. Sexual orientation, work, and income in Canada. *Canadian Journal of Economics* 41(4):1239–61.

Cerf, B. 2016. Sexual orientation, income, and stress at work. *Industrial Relations: A Journal of Economy and Society* 55(4):546–75.

Heisz, A. 2007. *Income Inequality and Redistribution in Canada: 1976 to 2004*. Analytic Studies Branch Research Paper Series 11F0019 No. 298. Ottawa: Statistics Canada.

Klawitter, M. 2015. Meta-analysis of the effects of sexual orientation on earnings. *Industrial Relations: A Journal of Economy and Society* 54(1):4–32.

LaFrance, A., C. Warman, and F. Woolley. 2009. Sexual Identity and the Marriage Premium. Working Paper 1219. Kingston, ON: Department of Economics, Queen's University.

Mueller, R.E. 2014. Wage differentials of males and females in same-sex and different-sex couples in Canada, 2006–2010. *Canadian Studies in Population* 41(3–4):105–16.

Statistics Canada. 2008. *Income and Earnings Reference Guide, 2006 Census*. Document No. 97-563-G. Ottawa: Statistics Canada.

Waite, S. 2015. Does it get better? A quasi-cohort analysis of sexual minority wage gaps in Canada. *Social Science Research* 54:113–30.

Waite, S., and N. Denier. 2015. Gay pay for straight work: Mechanisms generating disadvantage. *Gender & Society* 29(4):561–88.

Appendix A

Table A1. Estimates of wage gaps for coupled gay men relative to coupled heterosexual men, controlling for occupation and industry of employment

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income	Continuous hourly income	Discrete hourly income	Continuous hourly earnings	Discrete hourly earnings
a. Waite & Denier (2015)								
	−0.052***	−0.036***	−0.076***	−0.061***	−0.060***	−0.045***	−0.036***	−0.021*
	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
b. Mueller (2014)								
1. No additional controls	−0.086***	−0.061***	−0.095***	−0.071***	−0.058***	−0.034***	−0.051***	−0.024
	(0.013)	(0.013)	(0.011)	(0.009)	(0.010)	(0.008)	(0.012)	(0.012)
2. + Additional controls	−0.068***	−0.041**	−0.074***	−0.053***	−0.042***	−0.021**	−0.037**	−0.009
	(0.013)	(0.013)	(0.011)	(0.008)	(0.009)	(0.008)	(0.012)	(0.012)

Notes: * P ≤ .05; ** P ≤ .01; *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

Table A2. Estimates of wage gaps for coupled lesbian women relative to coupled heterosexual women, controlling for occupation and industry of employment

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income	Continuous hourly income	Discrete hourly income	Continuous hourly earnings	Discrete hourly earnings
a. Waite & Denier (2015)								
	0.099***	0.101***	0.086***	0.082***	0.049***	0.046***	0.062***	0.065***
	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
b. Mueller (2014)								
1. No additional controls	0.096***	0.091***	0.091***	0.086***	0.033**	0.027***	0.037***	0.033**
	(0.012)	(0.013)	(0.012)	(0.009)	(0.011)	(0.008)	(0.011)	(0.012)
2. + Additional controls	0.099***	0.094***	0.095***	0.088***	0.036***	0.030***	0.040***	0.036**
	(0.012)	(0.013)	(0.012)	(0.008)	(0.011)	(0.008)	(0.010)	(0.012)

Notes: * P ≤ .05; ** P ≤ .01; *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

Appendix B

Table B1. Estimates of wage gaps for coupled gay men relative to coupled heterosexual men, unadjusted for labour supply

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income
a. Waite & Denier (2015)				
	−0.111***	−0.094***	−0.131***	−0.114***
	(0.012)	(0.011)	(0.011)	(0.010)
b. Mueller (2014)				
1. No additional controls	−0.145***	−0.124***	−0.150***	−0.123***
	(0.014)	(0.015)	(0.012)	(0.010)
2. + Additional controls	−0.122***	−0.101***	−0.125***	−0.102***
	(0.014)	(0.015)	(0.012)	(0.009)

Notes: *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

Table B2. Estimates of wage gaps for coupled lesbian women relative to coupled heterosexual women, unadjusted for labour supply

Sample	Continuous annual earnings	Discrete annual earnings	Continuous annual income	Discrete annual income
a. Waite & Denier (2015)				
	0.119***	0.121***	0.106***	0.102***
	(0.012)	(0.011)	(0.010)	(0.010)
b. Mueller (2014)				
1. No additional controls	0.108***	0.099***	0.107***	0.101***
	(0.013)	(0.014)	(0.013)	(0.009)
2. + Additional controls	0.114***	0.105***	0.112***	0.106***
	(0.013)	(0.014)	(0.013)	(0.009)

Notes: *** P ≤ .001. Standard errors given in parentheses. Model 1 controls age, education, work experience, common-law status, presence of children in the household, rural residence, and province of residence. Models for annual earnings and income further control weeks worked and part-time status. Models for Mueller (2014) with additional controls also include controls for aboriginal status and immigration status.

Canadian Studies in Population 43(3–4): 272–86.

The provision of unpaid care across cohorts and genders: A Research Note

Christine Proulx[1]

Abstract

In this study, the caregiving history collected in the 2007 General Social Survey (GSS) is used to document the provision of care since the age of 15 years, the number of people helped, and the relationship with the care recipients. Using life tables, we confirm an upward trend in caregiving across birth cohorts. Unexpectedly, the findings also show that providing care starts at earlier ages in more recent cohorts—a result that appears partly linked to the emergence of new care relationships—and that the gender gap in the provision of care has widened over time.

Keywords: Canada; caregiving; gender; cohorts; life tables.

Résumé

Dans le cadre de cette étude, nous examinons les soins fournis depuis l'âge de 15 ans, le nombre de personnes aidées et la nature des liens avec ces personnes en nous basant sur l'historique des soins prodigués recueilli dans l'Enquête sociale générale (ESG) de 2007. À l'aide de tables d'extinction simple, nous confirmons l'augmentation de la prestation de soins au fil des cohortes de naissance. Par ailleurs, les soins commencent de plus en plus tôt, notamment grâce à l'apparition de nouveaux types de relations d'aide, et l'écart dans la prestation de soins se creuse entre les hommes et les femmes.

Mots-clés : Canada; prestation de soins; genre; cohortes; tables d'extinction simple.

Introduction

In 2012, 8.1 million Canadians representing 28 per cent of the population aged 15 years and over and living in the community reported providing some type of care to relatives or friends (Sinha 2013). Age-related needs of the care recipient were by far the reason most frequently cited for the assistance provided by individuals. With the aging of the population, the need for care is expected to increase in the future. In 2036, one in four Canadians will be aged 65 years or older, a marked increase from their 16 per cent share of the population in 2014 (Statistics Canada 2010, 2014a). The proportion of the population aged 75 years and older will increase from 8 per cent in 2014 to 13 per cent in 2036. A large majority of care recipients receive help from family members, relatives, or friends who are not paid for the care they provide (Sinha and Bleakney 2014). Considering the greater care needs at older ages and government policies prioritizing aging in the community and care in the home (Keefe et al. 2005; Lavoie 2012), the reliance on unpaid caregivers is unlikely to fade in years to come.

Our understanding of how common the provision of unpaid care is in Canada as well as how it has evolved over time is limited due to the absence of data collected prior to 1996. The number of caregivers appears to have increased over the past two decades. Is the increase linked to a similar increase across cohorts of the proportion of those who provide care or is it also related to a change in *tempo*, that is, the age at which individuals start providing care? The analysis of individuals' care histories, which can be approached with retrospective survey data, can shed light on the life stage at which care responsibilities arise which is informative of the responsibilities generally associated with that life stage and of the potential consequences that providing care might have regarding other domains of their life, such as the caregiver's health, employment and family life. The

1. Research carried out while completing doctoral studies at McGill University. Email: cproulx05@gmail.com.

repercussions of caregiving in the lives of individuals may vary depending upon the timing over the life course (Elder 1998). This no doubt constitutes valuable information for policymakers who may develop services to the population according to caregivers' needs at different stages of their lives. In addition, gaining a better understanding of past trends in caregiving by observing the experiences of successive birth cohorts could shed light on whether these trends will be maintained or not.

The 2007 General Social Survey (GSS) constitutes the very first attempt in Canada at gathering the complete caregiving history of individuals. The survey provided information collected about a total of five care episodes experienced by respondents since the age of 15, elaborating as well on whether or not they had provided care in the twelve months prior. The reconstruction of the respondents' care history allows us to estimate the percentage of individuals who have provided care at least once in their life, the age at which they started doing so and the extent to which these have changed across cohorts. In addition, this paper examines differences in the cumulated probabilities of men and women to provide care and their variations across birth cohorts. A convergence across cohorts in the proportions of men and women who provide care at least once over the course of their lives would support expectations of an increasing similarity in gender roles over time. However, a divergence would challenge our ideas about the evolution of gender roles and motivate a search for alternative explanations.

Literature review

The increasing proportion of caregivers in the Canadian population

Statistics Canada collected information on caregiving for the first time as part of the 1996 GSS (Statistics Canada 2013). Cycles on social support were repeated in 2002, 2007 and 2012, but changes in the population surveyed and definitions of care make it difficult to readily compare published estimates. With these limitations in mind, a close look at the proportions of caregivers in each of these years suggests that they have been increasing over the past 20 years. Table 1 presents a summary of these estimates.

In 1996, 12,756 Canadians aged 15 years and over and living in private households in one of the ten provinces were asked whether they had provided care to persons with a long-term health condition or physical limitation in the year preceding the survey. The population of caregivers was estimated at 2.85 million, of which nearly half were less than 45 years old (Fast et al. 2011; Fast and Keating 2001). According to this survey, 10 and 14 per cent of Canadian men and women aged 15 and over, respectively, had been providing care to at least one person during the year prior to the survey (Cranswick 1997). A large share of caregivers provided care to a senior. In all, 2.1 million Canadians, representing 11 per cent of the population aged 15 years and over, helped at least one person over the age of 65 (Fast and Keating 2001).

In 2002, the target population was restricted to individuals aged 45 years and over and caregiving questions concerned only assistance provided to seniors. According to this survey, 2 million Canadians aged 45 years and over—19 per cent of men and 18 per cent of women—had provided eldercare for a long-term health problem within the year prior to the survey, an increase from the approximate number of 1 million people of that age group estimated in the 1996 GSS (Cranswick and Dosman 2008).

The range of information collected on caregiving was far more extensive in the 2007 GSS. A wide array of questions were asked of respondents aged 45 and over about the care for a long-term health condition or physical limitation provided in the twelve months preceding the survey. Further questions were also asked about the episodes of care that respondents provided to up to five persons since the age of 15. Studies using this dataset suggested that 3.8 million—or 29 per cent of the Canadian population aged 45 and over—had provided care in the year prior to the 2007 GSS, among whom 2.7 million, 19 per cent of men and 22 per cent of women, provided care to a senior (Cranswick and Dosman 2008; Fast et al. 2011). These studies thus suggest that the number of people 45 years and over who provide care to seniors increased by approximately 700,000 between 2002 and 2007 and that the increase was due primarily to the addition of women, the proportion of female caregivers rising from 18 to 22 per cent between the two surveys. The focus on the assistance provided in a given year does not give a good indication of how common the provision of care can be over one's life course. Indeed,

Table 1. Estimates of the proportion of caregivers in the population published using data from the 1996, 2002, 2007, and 2012 General Social Surveys (GSS)

	1996	2002	2007	2012
% provided care in past 12 months	10% of men 14% of women (Cranswick 1997)	–	29% (Fast et al. 2011)	28% (Sinha 2013)
% provided care since age 15	–	–	40% of men 52% of women (Fast et al. 2013)	46% (Sinha 2013)
% provided eldercare in past 12 months	11% (Fast and Keating 2001)	19% of men 18% of women (Cranswick and Dosman 2008)	19% of men 22% of women (Cranswick and Dosman 2008)	–
Age restriction of sample	15+	45+	45+	15+
Definitions of care	Assisted someone with child care, meal preparation and clean-up, house cleaning, laundry, sewing, house maintenance, outdoor work, shopping for groceries and other necessities, transportation or personal care because of a long-term health condition or physical limitation in the past 12 months	Assisted someone over the age of 65 with meal preparation and clean-up, house cleaning, laundry, sewing, house maintenance, outdoor work, shopping for groceries and other necessities, transportation or personal care because of a long-term health condition or physical limitation in the past 12 months (child care assistance only asked to respondents 65 years and older)	Assisted someone with meal preparation and clean-up, house cleaning, laundry, sewing, house maintenance, outdoor work, shopping for groceries and other necessities, transportation, personal care, coordination of caregiving tasks, medical treatments or managing finances because of a long-term health condition or physical limitation in the past 12 months AND Assisted someone since age 15 (up to 5 care recipients listed)	Assisted someone with meal preparation and clean-up, house cleaning, laundry, sewing, house maintenance, outdoor work, shopping for groceries and other necessities, transportation, personal care, coordination of caregiving tasks, medical treatments or managing finances because of a long-term health condition, physical limitation or aging needs in the past 12 months AND Assisted someone since age 15 (up to 5 care recipients listed)

more than half of women (52 per cent) and 40 per cent of men 45 years and over provided care at least once at some point in their lives since the age of 15, a much higher proportion than 29 per cent characterizing those who assisted someone in the previous year (Fast et al. 2013).

In 2012, the most recent cycle of the GSS to focus on caregiving and care-receiving, 28 per cent of individuals aged 15 years and over provided care, a sharp increase from the percentage of 10–14% derived from the 1996 GSS (Cranswick 1997; Sinha 2013). Expanding the definition of care to include aging needs in addition to long-term health conditions or disabilities probably accounts for part of the increase observed but surely cannot explain its entirety. Again, a number of individuals (46 per cent) had provided care at some point during their lives (Sinha 2013). Caregiving appears to be on the rise and the increase described above took place over a relatively short period of time.

Whether this trend was already ongoing prior to the 1990s remains unclear since, to our knowledge, no Canadian study used the retrospective information on caregiving available in the last two cycles of the GSS on social support to examine the prevalence of care across cohorts. Using this information would increase the period over which caregiving could be studied to the latter half of the 20th century and beginning of the 21st century. A closer look at past trends in caregiving may help us understand where trends may head in the future. The next few paragraphs review other studies' findings regarding the evolution of caregiving across cohorts and provide reasons to expect similar developments in Canada.

Changes in proportions of caregivers across birth cohorts

One U.S. study of married women living in a mid-sized community in upstate New York used such retrospective information about the care provided over the life course (Robison et al. 1995). Four birth cohorts were compared and an increase in the proportions of women who provided care to someone of an older generation, a person of the same generation or a person of a younger generation was witnessed across cohorts. For example, 17 per cent of women born between 1905 and 1917 had provided care to an older person who was sick or infirm at some point in their lives compared to 33 per cent of women born between 1927 and 1934.

Another study, using a quasi-cohort approach with four cycles (1985, 1990, 1995 and 2000) of the British General Household Survey (GHS), also found that mid-life men and women increasingly took on caregiving responsibilities for a sick, disabled or elderly person across birth cohorts (Evandrou and Glaser 2002). Between the ages of 55 to 59 years old, 19 per cent of women and 12 per cent of men born between 1926 and 1930 reported providing some type of care compared to 28 and 20 per cent, respectively, in the 1941–45 cohort. These substantially lower proportions than in the U.S. study described above can be attributed to the cross-sectional nature of the GHS.

A number of reasons lead us to believe that an increase in caregiving over the life course could also be witnessed across cohorts in Canada. Firstly, the de-institutionalization of the elderly and of other persons with disabilities that took place in the 1980s shifted more of the responsibility for care to family members (Lavoie 2012). Home care services were implemented during the 1980s to help seniors remain in their home but the resources for those services remained limited compared to the services provided in institutions in the 1960s and 1970s (Lavoie 2012; Protecteur 2012) These institutional changes likely increased the proportion of individuals who had to provide care without pay.

Secondly, given the steady increase in life expectancy over the past 90 years (Decady and Greenberg 2014), it has become increasingly common for multiple generations of a family to be able to maintain relationships over several years. In a seminal lecture, Bengtson (2001: 12) used the expression "longer years of shared lives" to describe this phenomenon, which he asserts is likely to yield new forms of interactions. We would argue that the increasing involvement of younger generations in relations of care may be such a form of interaction. Providing care to a grandparent might have been a rare event for the parents of the baby boomers, who were less likely than their children and grandchildren to have known their grandparents. These relationships, although not as frequent as other care relationships such as the aging parent-adult child dyad, are becoming more common. For example, in 2000 it was estimated that grandchildren accounted for 8 per cent of caregivers in the United States (Foundation 2003, as cited in Fruhauf et al. 2006). Young adults and children sometimes act as an additional source of care to that provided by their parents, and they may find in this activity a way to return to their grandparents the support they received from them earlier in their lives (Fruhauf et al. 2006; Szinovacz 2008). It is unknown whether the provision of care by grandchildren displaces the assistance provided by other members of the family, or rather adds to the latter, thus increasing the size of the care network of elderly persons. The latter is plausible, since many grandchildren caregivers appear to act as secondary sources of support (Fruhauf et al. 2006).

Third, new forms of care relationships may have emerged as a result of the increasing conjugal instability observed since the late 1960s (Ménard and Le Bourdais 2012). While it was relatively common in the past for individuals to rely on a spouse or partner when in need of assistance, turning to friends for support may be increasingly common, especially following separation or divorce. The data published until now does not support an increase in friend caregiving. Sixteen percent of unpaid caregivers provided care to a friend, colleague or neighbour in 2012 compared to 24 and 13 per cent for instrumental and personal care, respectively, in 1996 (Cranswick 1997; Sinha 2013). In 2012, only 8 per cent helped a partner (Sinha 2013). However, it is two different things to compare proportions of caregivers who helped a friend and to scrutinize proportions of the population who were friend caregivers. It is possible that the relative share of non-kin caregivers decreased because the number of parent caregivers, for example, substantially increased, but that their share in the general population (including individuals not providing care) still increased. Friend caregivers report spending on

average fewer hours per week providing care than spouses and partners or other close relatives (Himes and Reidy 2000; Sinha 2013). However, research has shown that the number of hours of assistance received is not affected by the proportion of kin versus non-kin in the family/friend care network (Fast et al. 2004). Perhaps, non-kin networks are larger in size than kin-only networks, which would explain why care recipients do not differ in the amount of care received depending upon the relationship composition of their network and why a decrease in partner caregiving could be accompanied by an increasing involvement of others than partners, including friends.

Lastly, the proportion of caregivers reported in surveys might have increased in part due to a change in perceptions about the caregiver role. Guberman et al. (2011) suggest that baby boomers consider their caring responsibilities as a role in itself, a shift from the perspective of their parent generation who perceived these as an aspect of their natural roles of mother, daughter, or wife. Furthermore, baby boomer caregiving women voice the importance to maintain their other activities, paid work and social activities, and not having to sacrifice those to take on caregiving, although, in practice, many do make a considerable amount of changes in their professional and personal lives as a consequence of caregiving. They tend to find the use of formal services as acceptable, services they believed were not available to their mothers who had no choice but to provide care to other family members. Because of their perception of caregiving as a role in itself, it is thus reasonable to expect that baby boomers will be more likely to report periods of caregiving than previous generations.

Gender differences in caregiving

Women have traditionally engaged in caring work and housework and they continue to form the majority of caregivers (Cranswick and Dosman 2008; Fast et al. 2011; Sinha 2013). Women also spend on average more hours per week in that role and are more involved in the provision of personal care and housework than are men, tasks that need to be done on a regular schedule. Only in one domain, that of household maintenance and outdoor work do men invest more time than women. Nonetheless, in 2012, men accounted for 46 per cent of all caregivers aged 15 and over (Sinha 2013). This proportion is substantial and probably higher than it used to be 50 or 60 years ago, when gender roles were more segregated. Yet, caregiving still characterizes the lives of greater proportions of women than of men. Indeed, 52 per cent of women and 40 per cent of men 45 years and over in 2007 had provided care at least once since age 15 (Fast et al. 2013).

A study on the provision of care over the life course conducted with British data suggested that women start assuming caregiving responsibilities for a first time at a faster rate than men almost throughout their lives (Henz, 2004). The hazard rate of first caring increases steadily from the ages of 15 to 40, but men's rate is consistently half that of women. Then, the rate of women peaks at age 50 when 23 women per 1000 who had not yet provided care enter that role. After that age, women's rate falls, even briefly below men's, which remains steady at 10 per 1000 between the ages of 54 and 65. However, hazard rates present another side of the story than cumulative probabilities over time. Although women's hazard rates are below men's during two years at ages 63 and 64, the proportion of women who provided care at least once is surely greater than men's since they have been entering first care at a higher rate throughout their life before that age.

Apart from gender role explanations, there does not seem to be many theories on why women would be more involved in caregiving than men. A qualitative study of siblings whose parents were at least 75 years old suggests that brothers and sisters have different approaches in how they interact with their parents, differences which may become more pronounced as parents' needs increase (Matthews 2002). On the one hand, brothers tend to wait for their parents' requests for help as they want to encourage them to be self-reliant. On the other hand, sisters are generally proactive and offer assistance regularly without waiting for their parents requests for help. In families where there are brothers and sisters, brothers generally provide more assistance than in only-sons or brothers-only families, possibly because their sisters who hold a familial view of responsibility pressure them into doing so. Despite their implication, sisters tend to perceive their brother's contribution as not sufficient.

Converging gender roles and norms

Although men and women may differ in their approaches to care, with women more often involved in tasks that need to be done on a regular basis, the gender differences in the proportion of caregivers in the population could be less pronounced today than in the past. With the profound transformations in gender roles and norms observed over the past few decades, proportions of men and women who engage in family or friend caregiving could be converging. Although estimates are difficult to compare, there is some evidence of a convergence between 2007 and 2012. In 2007, 43.5 per cent of caregivers aged 45 and over were men compared to 46 per cent of those 15 years and over in 2012 (Fast et al. 2011; Sinha 2013). This convergence is likely to originate from a relatively greater increase in the number of men involved in caregiving than in that of women.

One of the major components of the observed social transformation in gender roles is the greater participation of women in the sphere of paid employment, which has increased dramatically, from below 40 per cent in the 1960s to 62 per cent in 2013 (Engelhardt and Prskawetz 2004; Statistics Canada 2014b). It is now common for both members of a couple to engage in paid employment (Marshall 2006). However, the narrowing of the gender gap in unpaid labour has occurred much less rapidly than in paid work. Hochschild and Machung (1997) described how women engaged in paid work dealt with a *second shift* of housework and childcare once they returned home. Whereas a gap of one hour per day still exists between men and women in the amount of time spent doing unpaid work in the home, time-use surveys show that,between 1986 and 2005, men aged 25 to 54 increased their number of hours, while their female counterparts saw theirs diminished (Marshall 2006). Arrangements in the division of paid and unpaid work among couples have become more diversified, with the traditional model losing ground and couples sharing roles becoming more common (Ravanera et al. 2009). The gender differences noted in unpaid work, and the observed evolution in gender roles and norms, motivate us to conduct a separate analysis for men and women, in order to assess the extent of changes across cohorts in the provision of caregiving over a life course.

Research questions

The objective of this study is to examine the population of men and women who have experienced caregiving over the course of their lives. More specifically, we aim to estimate (1) the proportion of individuals who have provided care at least once since age 15; (2) the number of people to whom they have provided care; and (3) the nature of the relationship with the persons whom they have helped. We expect the proportion of individuals who have been caregivers at least once in their lives to be higher than the 29 per cent observed by Fast et al. (2011) for the year 2007, and closer to 40 and 52 per cent, the respective proportions of men and women aged 45 and over who had provided care to a family member or a friend at some point in their lives (Fast el al. 2013).

Using the life table method to estimate the cumulated probabilities of men and women to experience a first care episode across four birth cohorts, we examine (4) whether or not providing care has become more common among more recent cohorts; and (5) whether or not the timing of the first care episode has changed across cohorts. Our separate analysis by gender seeks to estimate (6) whether or not the gap in the experience of caregiving separating men and women has narrowed over time. Finally, we aim to (7) observe any possible emergence among later cohorts of such new care relationships as care to grandparents and to friends.

Data and method

Data

The General Social Survey (GSS) is an annual survey conducted by Statistics Canada and is representative of the Canadian population living in private households in the ten provinces. We use cycle 21 collected in 2007 because it was the first GSS to collect retrospective information about the experience of caregiving. It constitutes

the best opportunity to go further back in time to study changing trends in caregiving. The 2007 GSS collected information on family, social support and retirement from a sample of 23,404 respondents aged 45 years and over who were interviewed by phone. Two modules assessed the provision of unpaid care to family members, relatives and friends.

Participants were first asked if they had provided assistance to a person because of a long-term health condition or physical limitation in the twelve months prior to the survey. If they reported giving help for at least one care task—transportation, shopping, banking and bill paying, meal preparation and housework, house maintenance and outdoor work, personal care, medical care, coordination of caregiving tasks and management of finances—they were then asked to report the number of persons to whom they provided assistance. They also had to respond to a series of questions regarding the main care recipient. These included the latter's age and gender, the relationship respondents had with this person, the date at which they started providing care, the frequency with which assistance was provided for each task, the residential proximity and the type of dwelling of the care recipient.

In a further section of the survey, respondents were asked to indicate the number of people they had helped for a period of six months or longer, excluding those helped in the past twelve months, since the age of 15. The start and end dates of up to five care episodes were collected, along with the relationship to each care recipient. Besides the date at onset of the episode of care, only relationship to each care recipient was collected in both modules. In order to determine which care episode came first in the lives of respondents to examine the timing of the first caregiving episode, we had to merge the two modules: the episode with the primary care recipient that occurred in the previous twelve months and all the other episodes that occurred since the age of 15. In order to exclude the episodes recorded in the previous year of the survey that lasted less than six months, we retained only those starting in 2006 or earlier. Respondents' complete caregiving histories could thus contain up to a maximum of six caregiving episodes that lasted at least six months.

Sample

The sample of the present study includes individuals who never provided care and those who did at least once over the course of their lives. In order to establish the timing of the first caregiving episode among caregivers, we needed their complete history of caregiving. Missing dates for even one caregiving episode prevented us from ordering all episodes and consequently from studying the timing of first care. Of the original sample of 23,404 respondents, we excluded 1,011 individuals who were not sure whether or not they had helped anyone. The caregiving history was de facto incomplete for another 2,299 participants who had provided care to more than one person in the twelve months preceding the survey but for whom only the information on the care provided to the main care recipient was collected by Statistics Canada. Finally, 1,063 respondents were excluded because of missing information on start/end dates or relationships for some care episodes or because they had provided care to more than five persons since the age of 15. The start date of the sixth and subsequent episodes were thus missing and we could not ascertain that these episodes followed chronologically the first five episodes. The analytic sample thus contains 19,031 respondents, some of whom provided care over the course of their lives and others who did not.

Because the large majority of excluded cases reported having provided care at least once in their life but had missing information in their caregiving history, the proportions of lifetime caregivers that we estimated are conservative. A detailed analysis showed that women are slightly overrepresented among the excluded cases, which is not surprising given their greater propensity to provide care relative to men. The way the survey was designed made it impossible for the large number of respondents who helped more than one person in the year preceding the survey to report the start date and their relationship with all those they helped. In fact, they were only asked the date at which they started providing care to their main care recipient, but not for the others they helped during that year. The exclusion of these cases is thus not due to a recall bias. If it had been, we should have found a greater proportion of older than younger individuals excluded from the sample, which is clearly not the case here.

Method

We use life tables to estimate the cumulated probability, by age of respondent, of experiencing the onset of a first caregiving episode. The youngest age at which individuals could experience this transition was fixed at 15 years by Statistics Canada.[2] In the life table method, subjects are considered at risk of experiencing a transition, that is, of becoming a caregiver, until they experience that transition or until they are no longer under observation, in which case they become censored (i.e., retrieved from the risk group). In other words, respondents who never had provided unpaid care contributed to the calculation of the cumulated probabilities of providing care from age 15 until the time of the survey in 2007. Hence, all respondents are observed until at least age 45.

The cumulated probabilities of providing care for the first time were estimated separately for men and women among four birth cohorts. All but the earliest birth cohort span ten years. The earliest cohort, born prior to 1933, includes respondents who were aged 75 years and over at the time of survey in 2007; the other three cohorts were born between 1933 and 1942 (aged 65–74 years), 1943 and 1952 (aged 55–64 years), and 1953 and 1962 (aged 45–54 years).

Frequency tables are used to provide, separately for men and women, the distribution of respondents' birth cohort, number of care recipients since age 15, and relationship with care recipients. Another table presents, separately by cohorts, the respondents' relationship with their first care recipient for care episodes starting before age 45. Population weights are applied to percentages and cumulated probabilities.

Results

Table 2 presents some characteristics of the 8,250 men and 10,781 women aged 45 years and older in 2007 whose data was retained for the analysis. Given their greater life expectancy, it is not surprising to find that a larger percentage of women than men were born prior to 1933. In contrast, a larger share of men than women belongs to the Baby Boom cohorts of 1943–52 and 1953–62.

Caregiving is a common experience. Table 2 shows that 38.7 per cent of them provided care at least once for a period of six months or longer since the age of 15. Women were more likely to engage in the provision of care compared to their male counterparts, and to provide care to a larger number of recipients over the course of their lives. Twenty-eight percent of women provided care to one person, 11 per cent helped two persons, and 6 per cent assisted at least three persons. Although lower, the men's contribution should not be understated, with 22 per cent who helped one person, 8 per cent who assisted two persons, and 3 per cent who provided care to three or more persons since they were 15 years old.

Consistent with the literature, helping a mother was the most frequent care relationship observed, with 13 per cent of men and nearly 24 per cent of women ever providing care to their mother. More than one in ten women provided assistance to another relative or to their father, compared to 8 per cent of men. Providing care to a spouse or partner appears to be less common. Nearly 5 per cent of men and 7 per cent of their female vis-à-vis helped their ill or disabled partner at some point in their lives, percentages that are similar to those observed for the provision of care to a non-relative.

The experience of caregiving is common, and it has become even more common over time, as can be seen in Figure 1, which depicts the cumulated probabilities of men and women who provided care for the first time by birth cohort. Cohorts differ significantly in their experience of caregiving. The earliest cohort (represented by a dash-dot line) exhibits the lowest probability of providing care at any age throughout the life course, and the progression of experiencing caregiving is constant across cohorts. Hence, the cohort representing those born between 1953 and 1962 (solid line) presents the highest cumulated probability of having ever provided care at any age up until age 54 years (the maximum age of this cohort at survey). For instance, at age 40 years,

2. Even if the question asked respondents to report only those episodes that occurred since age 15, a very small number of respondents reported care episodes starting earlier. These episodes were not considered in the analysis, but the respondents remained in the analysis.

6 per cent of men born prior to 1933 had provided care for the first time compared to 17 per cent of those born between 1953 and 1962. A greater proportion of women had already been caregivers at that age, with one woman in ten from the earliest cohort and 23 per cent of those from the latest cohort having done so. In other words, the percentage of respondents who had ever provided care by age 40 years more than doubled, from the earliest to the latest cohorts of women, and nearly tripled among the cohorts of men.

Table 2. Characteristics of Canadian men and women aged 45 years and over in 2007

	Men	Women	Total
Birth cohort***			
Born prior to 1933	13.1	16.8	15.0
1933–42	17.3	18.5	17.9
1943–52	30.2	28.4	29.3
1953–62	39.4	36.4	37.8
Number of care recipients since age 15***			
0	67.8	55.2	61.3
1	21.6	28.0	24.9
2	7.6	10.8	9.3
3 or more	3.0	6.0	4.5
Ever provided care to…			
A spouse/partner***			
No	95.3	92.9	94.1
Yes	4.7	7.1	5.9
A mother***			
No	87.4	76.5	81.8
Yes	12.6	23.5	18.2
A father***			
No	91.9	89.2	90.5
Yes	8.1	10.8	9.5
A parent-in-law			
No	95.4	95.4	95.4
Yes	4.6	4.6	4.6
Another relative***			
No	92.3	88.1	90.1
Yes	7.7	11.9	9.9
A non-relative***			
No	94.4	92.9	93.6
Yes	5.6	7.1	6.4
N (unweighted)	8,250	10,781	19,031

Note: Chi-square tests significance levels: † $p<.10$; *$p<.05$; ** $p<.01$; *** $p<.001$. Some percentages do not add up to 100 due to rounding. The percentages are based on population weight data.

Source: General Social Survey 21.

The proportion of respondents who will have ever provided care by the end of their life is increasing across cohorts. Whereas the highest proportion (nearly 35 per cent) of men from the earliest cohort who reported having ever provided care was reached around age 85 years, a higher fraction of men had already done so at younger ages in the more recent cohorts. For example, 43 per cent of the men born in 1943–52 had experienced caregiving by age 64 years (the maximum age in this cohort), and this percentage is likely to keep increasing as they get older. A similar pattern can be observed among women. Slightly over 40 per cent of the earliest cohort of women reported ever providing care to someone; nearly 60 per cent of them had done so by age 64 years in the 1943–52 cohort.

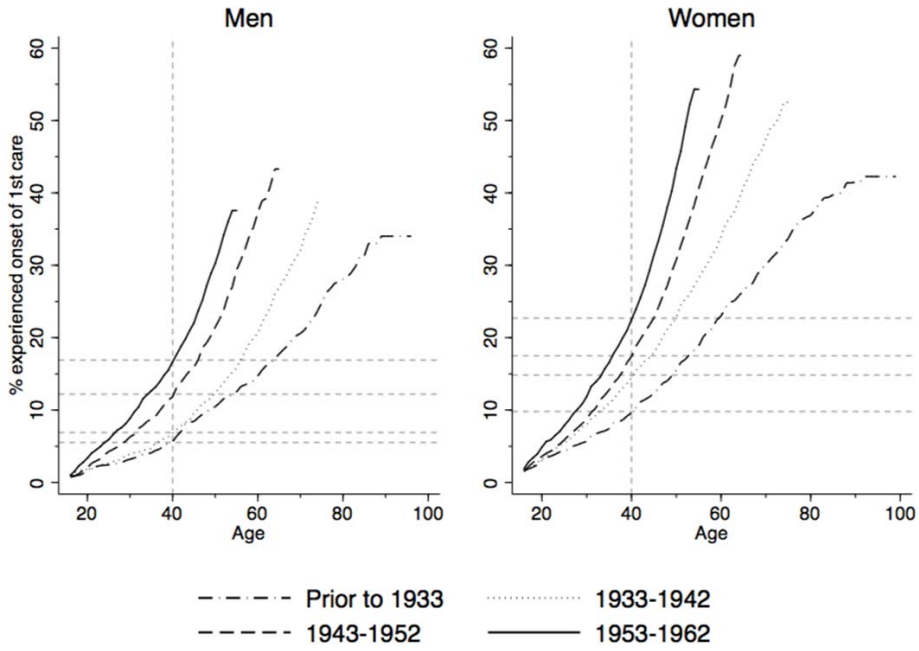


Figure 1. Cumulated probabilities of starting to provide care for the first time according to sex and birth cohort.

Source: General Social Survey 21. The cumulated probabilities are based on population weight data.

The differences observed between men and women in the proportions who ever provided care remain significant across all cohorts. Women are more likely than men to provide care over their life course. This pattern becomes obvious when comparing the horizontal lines denoting the proportions of individuals who had provided care by age 40. No signs of a narrowing gap between genders are found across cohorts. The gap between men's and women's proportions of first caregiving by age 40 passes from 4 percentage points in the earliest cohort to 6 percentage points in the latest cohort, a 50 per cent increase. In fact, the gap at the time of survey between men and women appears to be increasing even more, from 8 percentage points among the cohort born prior to 1933 to 14 percentage points in the 1933–42 cohort, 16 percentage points in the 1943–52 cohort and finally 17 percentage points in the latest (1953–62) cohort. This corresponds to a 112.5 per cent increase from the earliest to the most recent cohort. In other words, the gender gap that had already existed among the earliest cohort widened due to the larger increase over time of the proportion of women than men who ever provided care. Furthermore, a major part of the widening of the gap occurred past age 40.

The shift of the curves towards the left indicates that the onset of a first care episode occurs at a younger age in later than in earlier cohorts. This finding is surprising, considering the increase in life expectancy over the past few decades. Indeed, one might have expected that care needs would come later in the lives of individuals and that their caregivers would be correspondingly older. Therefore, the years gained in life expectancy may be disabled years. As well, this could suggest that caregiving is not confined within conjugal and parent-adult child relationships, but may span more than two generations, leading young adults to be providing care to elderly persons.

Table 3 examines this issue by comparing the proportions of respondents who provided care for the first time before age 45 years across cohorts, depending on their relationship with the care recipient. The relative prevalence of the episodes of care to spouses or partners, to parents or parents-in-law, and to other relatives has all declined across cohorts. Part of the decline in the provision of care to a partner may be linked to the better health status of individuals under the age of 45 years. In addition, the greater level of conjugal instability experienced by the two most recent cohorts compared to their predecessors may have limited the possibility to rely on a partner when in need of assistance. Even though the relative prevalence of providing care to parents or parents-in-law has declined over time, it still constitutes the majority of first care episodes occurring early in life.

Table 3. Respondents aged 45 years and over in 2007 who provided care to their first care recipient before the age of 45 (%), according to birth cohort and relationship with the care recipient

	Born prior to 1933	1933–42	1943–52	1953–62	Total
Spouse/partner	9.8	6.5	5.8	4.8	5.7
Parent/parent-in-law	67.1	66.8	64.9	58.5	62.0
Grandparent	1.2	4.4	5.6	10.1	7.4
Other relative	16.4	15.7	15.0	14.4	14.9
Non-relative	5.4	6.6	8.7	12.3	10.0
N (unweighted)	350	552	1,178	1,775	3,855

Note: Significant at $p < 0.001$. Some percentages do not add up to 100 due to rounding. The percentages are based on population weight data.

Source: General Social Survey 21.

In contrast, the proportions of respondents who have helped a grandparent or a non-relative have risen significantly over time. The former, which constituted only 1 per cent of all first care episodes starting before age 45 years for the earliest cohort, increased to 10 per cent among the most recent cohort. Similarly, the prevalence of help to non-relatives more than doubled, going from 5 per cent to 12 per cent of all first care episodes reported before age 45 years. The finding regarding the care to grandparents suggests that as life expectancy increases, relationships, including care relationships, can be maintained over a longer period of time across multiple generations of families—a social trend already noted by Bengtson (2001). Friends form the majority of non-relative caregivers. This category has probably increased in relative importance, because friends might in part compensate for the absence of support from partners or other relatives that is associated with the profound changes that families have experienced over the last 40 years.

Discussion

As already noted above, the provision of unpaid care is a very common experience. Using the 2007 GSS, Fast et al. (2013) estimated the proportion of individuals who had ever provided care at 40 and 52 per cent, respectively, for men and women aged 45 years and older. In this paper, we showed that 39 per cent of Canadians aged 45 years and over have provided care at some point in their lives for a period of at least six months since the age of 15 years. Among women, the proportion is higher with 45 per cent who assisted at least one person since age 15, although not as high as the proportion reported by Fast et al. (2013). As mentioned in the Data and Methods section, our calculations underestimate the proportion of Canadians who ever provided care since we had to exclude all those who had provided care, but who could not provide the start dates of all their care episodes. A large number of cases were excluded because they had provided care to more than one person in the past 12 months, but only the start date of the care episode for the main care recipient was collected by Statistics Canada. We had to exclude those cases because we needed start dates from all care episodes to order them correctly and be able to capture the timing of the first care episode. In addition, we only considered as caregivers individuals who reported providing care at least once for at least six months. A number of respondents had just started providing care in the year of the survey and as such, remained in our sample, but were not considered to have been caregivers at least once for six months or longer. Despite our conservative estimates, we do not expect them to affect the cohort differences observed in Figure 1, since we did not find any significant differences between cohorts in the percentages of cases with missing information on caregiving variables.

The use of life tables clearly showed that the provision of care has increased over time. The proportion of women in the two most recent cohorts who had provided care at least once in their lives by the time they participated in the survey is already higher than the percentages discussed above. These individuals were all under 64 years of age at the time of survey. Those who had not yet provided care over the course of their lives could eventually take on caregiving responsibilities as they continue to age. The deinstitutionalization of the elderly in the public health system may have sustained the trend in first care witnessed across cohorts. Although, the

possibility that caregiving is perceived differently—with later cohorts more likely to report episodes of care than earlier cohorts—clearly cannot be minimized in accounting for the trend observed.

The life table analysis also showed that not only are respondents from more recent birth cohorts more likely to have ever been a caregiver, they also started doing so for the first time at a younger age. This last finding was counter-intuitive given the marked increase in life expectancy over the past century. If care needs occur later in life, we could have expected caregivers, who oftentimes are the adult children of the care recipient, to be older as well. However, living a longer life does not necessarily mean that it will be exempt of illnesses or disabilities. The increase in life expectancy witnessed over the past several decades in the U.S. and in Canada is linked to increased survival with disease (Crimmins 2004; Mandich and Margolis 2014). The health of seniors generally improved in the 1980s and 1990s although they live longer with diseases. In the 1970s, there was an increase in mild disability among the elderly followed by a decrease in the 1980s and 1990s (Crimmins 2004). However, rates of severe disability increased among the U.S. working-age population over these last two decades (Bhattacharya et al. 2005). The latter trend has been associated to the greater prevalence of a number of chronic diseases and of rising disability among chronically ill individuals (Bhattacharya et al. 2005).

Severe disability is still more characteristic of the oldest old than the working-age population. Research using the 2009–10 Canadian Community Health Survey showed that, while severe disability occurs on average around the age of 77 years, moderate disability—the inability to perform some activities because of a limitation in vision, hearing, speech, mobility, dexterity, feelings, cognition or pain—usually appears around age 40 (Decady and Greenberg 2014). Hence, almost a third of all individuals aged 15 years and over who received care in Canada in 2012 were aged 45–64 years old (Sinha and Bleakney 2014). The fact that a greater share of people survive into older ages but are not necessarily healthy during these extra years may have contributed to the increase in the proportions of Canadians who provided care. That a good share of those who need assistance are middle-aged adults, a proportion that might have risen given the increased prevalence of chronic diseases and related disabilities within this population perhaps helps explain the fact that over the last few decades, Canadians have started providing care at ever-younger ages.

Another part of the explanation for the observed trend in the timing of the first care episode lies in the diversification of the relationships of care. While providing assistance to a parent still remains the most common care relationship, even among first episodes starting before age 45, baby boomers were found to be involved to a greater extent in the care of their grandparents than their parent generation was. Undoubtedly, this finding is associated with the increase in life expectancy that makes possible interactions and exchanges across more than two generations within families. Helping an ill or disabled grandparent was almost unheard of in the earlier cohorts, but is becoming an increasingly common phenomenon. The likelihood of starting to provide care for the first time before the age of 45 to a non-relative has also increased across cohorts, while the likelihood of helping a spouse or partner has decreased. This suggests that social support may be sought in friendships when other sources of support, such as the care from a partner, may be lacking. Further research in this area is needed in order to assess the role that non-relatives are likely to play in the future.

Given the trend towards more egalitarian gender norms and the estimates published in studies using GSS data, we were surprised to find that the proportion of women who engaged in caregiving increased more rapidly across cohorts than it did for their male counterparts. Whether this result reflects a stable trend or a difference in how men and women perceive, recall and report episodes of care is unclear, but it challenges our expectations and motivate us to find an explanation. The gap between the proportions of women and men who engage in caregiving is larger when we consider the care provided at any point over the life course since age 15 rather than the care provided in the past 12 months (see Table 1). Our results suggest that the small differences found at every age cumulate to yield larger gaps after several years of observation. The gender gap in the probability of providing care already widens across cohorts by age 40, but it increases even more after that age.

Guberman and her colleagues (2011) suggested that the identity of baby boomers regarding caregiving responsibilities might have been different than their parents' identity. Women constituted the overwhelming majority of their sample and family responsibilities have traditionally been considered part of women's role. Could it be possible that men and women of different birth cohorts see things differently? At the time when gender roles were much more segregated, men who engaged in caregiving might have been very much aware

that these responsibilities were unusual for men of their time whose lives were primarily defined by paid labour. This may have led them to recall those moments of their lives vividly and to report them once questioned on them, even several years later. On the other hand, women of the earliest cohort could have minimized the care they provided in their lives, as this was just a natural part of their roles of mothers, daughters and sisters. As gender roles became more similar, women may have recognized the care they provided as such. Later cohorts of women and men may thus have reported their episodes of care in a way that is more similar. In a sense, the increasing gender gap in first caregiving revealed by our results does not necessarily mean that genders have become more unequal in the domain of unpaid caregiving. Additional research would be needed to examine more closely how gender differences evolve across the life course and over historical time in terms of the provision and meaning of unpaid care. Interviews with men and women from the generations preceding the Baby Boom could shed light on this issue.

All things considered, our analysis suggests that the provision of care increased over time and that social support will continue to be an important part of the lives of Canadians in decades to come. It should thus remain an important area of study. The decreasing age at the onset of care raises a number of questions regarding the consequences that these responsibilities might have on the life course of individuals. A large proportion of the later cohorts had already provided care to at least one person before reaching the age of 45 years. Compared to earlier cohorts, these individuals have also delayed childbearing. This suggests that unpaid caregivers are increasingly likely to juggle multiple responsibilities in the future, raising young children while being employed. More refined studies of the repercussions that the provision of care exerts on individuals' employment, conjugal and family relations and health are needed in order to develop policies that meet the needs of these individuals. In addition to providing care at earlier ages, many will be caregivers more than once over the course of their lives and may need different kinds of support at different life stages.

The 2007 GSS is a useful data set to study caregiving over the life course, but it is not without limitations. Various pieces of information about the care provided by respondents were not collected retrospectively for the episodes of care that occurred since the age of 15. For instance, the survey did not contain information about the tasks accomplished, their frequency, and the number of hours per week taken up by caregiving responsibilities. Furthermore, a large number of cases had to be excluded, because information was not collected about the provision of care in the twelve months prior to the survey to individuals other than the main care recipient. Without the start dates of all episodes, we could not assume that the care provided to the main care recipient was the first episode to occur in the life of these respondents; thus, they could not be included in studying the timing of the first care episode. Collecting more information on all the episodes that occurred in the twelve months preceding the survey could be an interesting supplement to add to future collections of the GSS on social support and should not be too difficult to do as it used to be done in 1996 and 2002. At the least, Statistics Canada could collect information on the relationship with the care recipient as well as the start and end dates of each episode, as it does in the care history section. In addition, gathering information on whether other people provide assistance to the care recipient, or support the caregiver by taking on some of his or her tasks, could help us to understand whether the involvement of grandchildren or of friends compensates for or complements the care that would otherwise be provided by other family members.

Acknowledgments

Support for this research was provided by the Fonds de recherche du Québec – Société et culture (FRQSC), the Social Sciences and Humanities Research Council (SSHRC) and the McGill Canada Research Chair on Social Statistics and Family Change. The analysis presented in this paper was conducted at the Quebec Interuniversity Centre for Social Statistics (QICSS), which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the QICSS are made possible by financial or in-kind support from SSHRC, the Canadian Institutes of Health Research (CIHR), the Canada Foundation for Innovation (CFI), Statistics Canada, FRQSC and the Québec universities. The views expressed in this paper are those of the author, and not necessarily those of the CRDCN or its partners. I would like to thank Céline Le Bourdais for her continued support and Roderic Beaujot for his helpful comments.

References

Bengtson, V.L. 2001. Beyond the nuclear family: The increasing importance of multigenerational bonds. *Journal of Marriage and Family* 63(1):1–16.

Bhattacharya, J., K. Choudhry, and D. Lakdawalla. 2005. Chronic disease and trends in severe disability in working age populations. *Trends Working Paper Series* (05-4):1–30.

Cranswick, K. 1997. Canada's caregivers. *Canadian Social Trends* 47:2–6.

Cranswick, K., and D. Dosman. 2008. Eldercare: What we know today. *Canadian Social Trends* 86:48–56.

Crimmins, E.M. 2004. Trends in the health of the elderly. *Annual Review of Public Health* 25:79–98.

Decady, Y., and L. Greenberg. 2014. Ninety Years of Change in Life Expectancy. Health at a Glance (Cat. No. 82-624-X). Ottawa: Statistics Canada.

Elder, G.H. 1998. The life course as developmental theory. *Child Development* 69(1):1–12.

Engelhardt, H., and A. Prskawetz. 2004. On the changing correlation between fertility and female employment over space and time. *European Journal of Population* 20(1):35–62.

Evandrou, M., and K. Glaser. 2002. Changing economic and social roles: The experience of four cohorts of mid-life individuals in Britain, 1985–2000. *Population Trends* 110:19–30.

Fast, J., and N. Keating. 2001. *Informal Caregivers in Canada: A Snapshot.* Research on Aging, Policies and Practice Series (RAPP). Edmonton, AB: University of Alberta.

Fast, J., K. Duncan, C. Dunlop, J. Eales, N. Keating, D. Lero, and S. Yoshino. 2011. Gender Differences in Family/Friend Caregiving in Canada.http://sociology.uwo.ca/cluster/en/-PolicyBriefs.html (retrieved 6 September 2013).

Fast, J., N. Keating, L. Derksen, and P. Otfinowski. 2004. Characteristics of family/friend care networks of frail seniors. *Canadian Journal on Aging* 23(1):5–19.

Fast, J., N. Keating, D. Lero, J. Eales, and K. Duncan. 2013. *The Economic Costs of Care to Family/Friend Caregivers: A Synthesis of Findings*. Research on Aging, Policies and Practice Series (RAPP). Edmonton, AB: University of Alberta.

Foundation (Foundation for Accountability and The Robert Wood Johnson Foundation). 2003. *A Portrait of Informal Caregivers in America, 2001*. Portland: Foundation for Accountability.

Fruhauf, C.A., S.E. Jarrott, and K.R. Allen. 2006. Grandchildren's perceptions of caring for grandparents. *Journal of Family Issues* 27(7):887–911.

Guberman, N., J-P. Lavoie, and I. Olazabal. 2011. Baby-boomers and the 'denaturalisation' of care-giving in Quebec. *Ageing and Society* 31(7):1141.

Henz, U. 2004. The effects of informal care on paid-work participation in Great Britain: A lifecourse perspective. *Ageing and Society* 24(6):851–80.

Himes, C.L., and E.B. Reidy. 2000. The role of friends in caregiving. *Research on Aging* 22(4):315–36.

Hochschild, A., and A. Machung. 1997. *The Second Shift*. New York: Avon Books.

Keefe, J., P. Fancey, and S. White. 2005. *Consultation on Financial Compensation Initiatives for Family Caregivers of Dependent Adults: Final Report.* Halifax: Maritime Data Centre for Aging Research and Policy Analysis.

Lavoie, J-P. 2012. Entre État, famille, tiers secteur et marché: Les politiques québécoises à l'égard des personnes âgées ayant des incapacités, in *Le vieillissement de la population et les politiques publiques: enjeux d'ici et d'ailleurs*, edited by P. Marier. Québec: Les Presses de l'Université Laval, p. 129–51.

Mandich, S., and R. Margolis. 2014. Changes in disability-free life expectancy in Canada between 1994 and 2007. *Canadian Studies in Population* 41(1–2):192–208.

Marshall, K. 2006. Converging gender roles. *Perpectives on Labour and Income* 7(7):5–17.

Matthews, S.H. 2002. Brothers and parent care: An explanation for sons' underrepresentation, in *Men as Caregivers: Theory, Research, and Service Implications*, edited by B.J. Kramer and E.H. Thompson. New York: Springer, p. 234–49.

Ménard, F-P., and C. Le Bourdais. 2012. Diversification des trajectoires familiales des Canadiens âgés de demain et conséquences prévisibles sur le réseau de soutien. *Cahiers québécois de démographie* 41(1):131–61.

Protecteur (Le Protecteur du citoyen). 2012. *Chez soi: Toujours le premier choix? L'accessibilité aux services de soutien à domicile pour les personnes présentant une incapacité significative et persistante.* Québec: Le Protecteur du citoyen.

Ravanera, Z., R. Beaujot, and J. Liu. 2009. Models of earning and caring: Determinants of the division of work. *Canadian Review of Sociology* 46(4):319–37.

Robison, J., P. Moen, and D. Dempster-McClain. 1995. Women's caregiving: Changing profiles and pathways. The Journals of Gerontology Series B. *Psychological Sciences and Social Sciences* 50(6):S362–73.

Szinovacz, M.E. 2008. Children in caregiving families, in *Caregiving Contexts: Cultural, Familial, and Societal Implications*, edited by M.E. Szinovacz and A. Davey. New York: Springer, p. 161–91.

Sinha, M. 2013. Portrait of Caregivers, 2012. Spotlight on Canadians: Results from the General Social Survey (Cat. No. 89-652-X-001). Ottawa: Statistics Canada.

Sinha, M., and Bleakney, A. 2014. Receiving Care at Home. Spotlight on Canadians: Results from the General Social Survey (Cat. No. 89-652-X-002). Ottawa: Statistics Canada.

Statistics Canada. 2010. Population projections for Canada, provinces, and territories 2009 to 2036 (Cat. No. 91-520-X). Ottawa: Statistics Canada.

———. 2013. General Social Survey: An Overview, 2013 (Cat. No. 89F0115X). Ottawa: Statistics Canada.

———. 2014a. Table 051-0001: Population by sex and age group, CANSIM (database). http://-www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo10a-eng.htm (retrieved 13 October 2014).

———. 2014b. Table 282-0002: Labour force survey estimates (LFS), by sex and detailed age group, annual (persons unless otherwise noted), CANSIM (database). http://www5.statcan.-gc.ca/cansim/ (retrieved 22 September 2014).

Population Change in Canada

by Don Kerr and Roderic Beaujot
3rd edn, Don Mills, ON: Oxford University Press 2016
ISBN 978-0-19-900262-7
Softcover $64.95, 448 pp.

Reviewed by Anthony C. Masi
Desautels Faculty of Management, McGill University

The authors designed this monograph primarily as a supplementary text for courses dealing with population studies or as complementary reading in social sciences or public policy offerings that would benefit from content on Canada's demography. It is well-suited for both purposes. Indeed, many individual chapters would make excellent required reading in specialist courses on a variety of topics.

Following a now-characteristic formula of OUP textbooks, the Kerr and Beaujot volume has a decidedly pedagogical style: each chapter starts with an "at a glance" box that in addition to providing highlights of what is to be presented also serves implicitly to outline specific learning objectives. Then the authors provide carefully researched and clearly written arguments for the topic at hand, and conclude by posing questions for critical thought, providing annotated recommended readings and pointing to related websites. Throughout the book, words and phrases placed in italics are (usually) defined in the glossary. The bibliography included at the end of the book is a masterful compilation of materials that have either been cited directly or have influenced the development of the substantive topics covered.

After two brief introductory chapters, one on the study of population change and the other on Canada's population prior to the 20th century, the volume is divided into three parts: population processes (chapters on mortality, fertility, immigration); growth, distribution, and aging (chapters on population distribution, internal migration and the regions, the implications of changing growth patterns, and population aging); and the consequences of population change (chapters on families and households, sociocultural and socioeconomic composition, the demography of Canada's Aboriginal peoples, and population and the environment). Following a summary concluding chapter, there are four appendices (the life table; rates and standardization; the 2014 world population data sheet; and population estimates and demographic accounts), the aforementioned glossary, references, and a comprehensive subject index.

This volume is more than a simple update of the second edition (Beaujot and Kerr 2004), which appeared over a decade ago. It contains new materials on Canada's pre-Confederation demography, the demographic history of First Nations peoples, and a detailed consideration of the interplay between population growth and the environment that places Canada's story in global context. Another feature is that topics are consistently framed in relation to population change: processes undergo variations, leading to structures and compositions that deviate from previous states, leading to processes with new values and changed context.

Faithful to the intent of the first edition of this book (Beaujot 1991), which carried the subtitle "challenges of policy adaptation," this third edition also explores appropriate policy responses on the part of governments when forced to "consider the impact of demographic change on fundamental aspects of Canadian society" (p. xi). Indeed, sensitivity to socioeconomic and political matters is what renders the Kerr-Beaujot volume a potentially interesting choice for courses across the spectrum of the social sciences. Any courses dealing with First

Nations issues would be well advised to assign Chapter 11, "The Demography of the Aboriginal Population of Canada," and pages 16 to 23 of Chapter 2 (pre-contact native population issues) as required readings.

Focusing attention on processes, structures, and consequences enables Kerr and Beaujot to treat each demographic topic in a discursive, non-technical manner, while being able to refer interested readers to the appendices, supplementary materials, or related websites for required details. Further, within the well-designed and clearly written chapters, the authors employ another standard OUP textbook device to address interesting, difficult, or controversial subjects: the sidebar box. Here are some examples: avoidable mortality (Box 3.2, p. 48–52), attitudes toward having children (Box 4.2, p. 84–86), the economic adjustment of immigrants (Box 5.6, p. 129–130), immigration and labour force growth (Box 7.1, p. 180–81), university credential, over-qualification, and the Canadian labour force (Box 10.2 p. 268–69), and the number of status Indians (Box 11.1, p. 282–83). In addition, they also use this tool to touch upon vital theoretical and/or methodological issues without losing the reader in mechanical and technical details, as in: the cohort as a concept in studying social change (Box 8.1, p. 195–97), immigration as a solution to population aging (Box 8.2, p. 198–200), and defining "family" (Box 9.1, p. 227).

Chapter 7, "Changing Growth Patterns and Their Implications," although not citing it directly, and Chapter 2, "The Population of Canada before the Twentieth Century," which does cite it, recall some of the themes presented in Beaujot and McQuillan's 1982 book, *Growth and Dualism: The Demographic Development of Canadian Society*.

Chapter 12, "Population and the Environment," is another example of a section of this book that can be employed standalone as required reading in courses dealing with (the human impact on) the health of the planet Earth. Here again, Kerr and Beaujot use the sidebar to get at topics not usually covered in introductory courses in population studies, e.g., "Box 12.1: Taking a Closer Look at Canada's Record on GHG Emissions" (p. 307–10) and "Box 12.2: How Large is Our Ecological Footprint?" (pp. 312–15).

The treatment of some of topics will not be to the liking of everyone, but each topic is covered in a way that allows for discussion, debate, and dissent. In addition, the volume could have been shortened by eliminating the appendices and simply referring readers to other print or digital sources. Greater attention might have been given to the design and presentation of some tables, with an eye to eliminating clutter and redundancy. These, however, are small quibbles that do not detract from the overall quality of the effort and output.

This volume also marks an impressive milestone for Professor Roderic Beaujot—an important book dealing with Canadian demography in every one of the last four decades.

References

Beaujot, R. 1991. *Population Change in Canada: The Challenges of Policy Adaptation*. Toronto: Oxford University Press.

Beaujot, R., and D. Kerr. 2004. *Population Change in Canada*. 2nd edn. Toronto: Oxford University Press.

Beaujot, R., and K. McQuillan. 1982. *Growth and Dualism: The Demographic Development of Canadian Society*. Toronto: Gage.

Applied Multiregional Demography: Migration and Population Redistribution

by Andrei Rogers
Dordrecht, Heidelberg, London, New York: Springer 2016
ISBN 978-3-319-22317-9
Softcover $54.99, 114 pp.

Reviewed by David A. Swanson
Department of Sociology, University of California Riverside

As I read through *Applied Multiregional Demography: Migration and Population Redistribution*, I repeatedly stopped to compare it to a book I read by Frans Willekens: *Multistate Analysis of Life Histories with R* (for a review, see Swanson 2015). I find Willekens' book much more innovative and useful than this book by Rogers.

Willekens' book is much more innovative because he discusses the data problems faced when considering multi-state demography, and he offers ways to overcome them (e.g., with micro-level data). Rogers however, focuses on the use of aggregated data and gives the problems associated with aggregated data very short shrift. Instead, he spends too much of the book complaining about the shortcomings of other approaches. Rogers also spends too much of the book on another complaint—the use of "net migration" as a concept in demographic analysis. I find his complaint ironic given that in this book, there are 26 instances of the use of "natural increase," which is itself a net measure, conceptually the same as net migration. Unlike "Requiem for the net migrant" (Rogers 1990), I have never run across a work by Rogers titled "Requiem for the natural increment."

I find the discussion of another construct also to be ironic. Toward the end of the book, Rogers notes that "double-entry bookkeeping" is conceptually similar to "multi-regional migration." Of course it is! But what Rogers fails to bring into this discussion is the fact that for all businesses, including those who use double-entry bookkeeping and those that do not, the bottom line is "profit," a measure that is conceptually the same as net migration. As is the case with the "natural increment," I have never read something by Rogers entitled "Requiem for profit."

The actual examples of applications provided by Rogers in terms of multi-regional demography all involve large pieces of geography. There is nothing on small pieces of geography, such as counties, much less census tracts. Is this because of the difficulty encountered in getting the data needed to implement the multi-regional approach for small areas such as counties, census tracts, and the like? Even for the examples he provides, there is absolutely no mention of the issues affecting the validity and reliability of the data required. This is especially important in the US, where the problematic "long form" data needed to use a multi-regional perspective in the past have been replaced by the far more problematic "American Community Survey" data. As such, I find the examples to be not very useful.

For what it does cover—large geographic areas, aggregate data, and a list of the problematic issues associated with net migration—the book is fine. However, if you are not interested in polemics on net migration, but are interested in actually implementing the multi-regional approach, especially for small geographic levels using micro-level data, I recommend instead the book by Willekens, who provides a much more innovative perspective, along with useful examples. In a similar vein, another useful example is found in the work of Wilson (forthcoming), who uses the "multi-regional" perspective to effect dynamic changes in ethnic identity in a micro-simulation projection.

References

Rogers, A. 1990. Requiem for the net migrant. *Geographical Analysis* 22:283–300.

Swanson, D.A. 2015. Review of *Multistate Analysis of Life Histories with R*, by Frans Willekens. *Canadian Studies in Population* 42(3–4):80–81.

Wilson, T. nd. The future of Australia's indigenous population. *Population Studies*. Forthcoming.

Reproducing Citizens: Family, State, and Civil Society

edited by Sasha Roseneil, Isabel Crowhurst,
Ana Cristina Santos, and Mariya Stoilova
London and New York: Routledge (Taylor and Francis Group) 2016
ISBN 978-1-138-94229-5
Hardcover $160, 106 pp.

Reviewed by Anthony C. Masi
Desautels Faculty of Management, McGill University

This short volume was originally published as a series of articles in *Citizenship Studies*, volume 17, issue 8 (December 2013), and the reader is informed that when this material is cited, the original page numbering should be used. In this review, however, for the sake of simplicity, I use the page numbers in the book when referring to the specific passages cited.

The twelve contributors to the seven chapters (including the editorial introduction) of this book come from the following fields of study: sociology, criminology, education, communications studies, psychology, history and contemporary studies, law, and social work. The chapters present case studies of various issues, dealing with the following countries: Mexico, Sweden, Italy (2), Australia, and the United Kingdom.

From this array of disciplines and geographic locales, the reader is invited to consider reproduction in relation to citizenship as these variables intersect with diverse dichotomies such as public/private, political/personal, rational/emotional, mind/body and with the "biological, sexual, and technological realities of natality" (p. 1). Thus, this is a book about citizenship, not about demography per se, but it does play nicely on the ambiguity of the term "reproducing": (i) creating new life, (ii) perpetuating social inequalities across groups and generations, and (iii) socializing individuals to expectations regarding rights and obligations, of which the ability to have children should be seen as fundamental to citizenship. In this regard, the editorial introduction plays the crucial role of providing the reader with sufficient background to appreciate the other contributions in the volume.

For demographers, two novel aspects of these chapters/articles will be (a) the categorization of different dimensions of citizenship (political, social, economic, multicultural, bodily, and intimate); and (b) framing the relationship between the citizenship and the state's regulation of reproduction through the promulgation of procreative standards that are not always available to citizens who do not fit the norm. While dealing with widely varying settings and alternative theoretical perspectives, the principal (and principle) concern of this volume is the extent to which reproductive self-determination has been, can be, or should be secured. In other words, the chapters in this book provide commentary (sometimes explicitly, sometimes implicitly) on population policies, laws, and other regulations, and on the impact that these norms have on citizens' participation, or lack thereof, in procreation.

In reading these chapters, the quantitatively-minded will be challenged to think about demographic variables without numbers and asked to appreciate the more qualitative aspects of the issues: abortion questions in Mexico (p. 12–27), motherhood for the transgendered in Sweden (p. 28–41), lack of access to assisted reproduction in Italy (p. 42–55), media representations and "reproductive vulnerabilities" of those from Australia who seek off-shore surrogacy (p. 56–69), Kurdish migrant women's attitudes and practices regarding motherhood in London (UK) (p. 70–84), and intergenerational relations within the biological family for non-straight individuals in (again) Italy (p. 85–99).

Those who already have a strong qualitative orientation will read the chapters and immediately understand the saliency of the categorical questions raised by the authors; therefore, they would benefit by making the effort to assess the numerical import of the arguments, such as their actual or potential impact on population birth rates or on individual fertility. In both instances, approaching the studies from the perspective of the other should make the volume more enjoyable than just reading the words as written.

The individual chapters are comprehensible, but often the theoretical discussions are verbose, convoluted, and not always suitably linked to the empirical evidence presented for the case studies that form the backbone of each chapter. These articles/chapters do form a coherent narrative on the theme of citizenship and reproduction (in the three senses of the word noted above), thanks to the introductory essay by the editors and also to the fact that the authors of the chapters all cite directly, or draw on indirectly, the essay on citizenship, reproduction and the state by B.S. Turner (2008) or his earlier piece (2001) on the erosion of citizenship.

In summary, this edited volume (originally a special issue of *Citizenship Studies*) provides some thought-provoking ideas about natality and citizenship from theoretical and policy perspectives, but it is lacking the kind of demographic data that could actually inform or sustain the arguments that are made. These are analyses of numerators and the authors appear not to be interested in the associated denominators.

References

Turner, B.S. 2008. Citizenship, reproduction, and the state: international marriage and human rights. *Citizenship Studies* 12(1):45–54.

———. 2001. The erosion of citizenship. *The British Journal of Sociology* 52(2):189–209.

States and Markets: Sociology of Public Policy in Canada

by Susan A. McDaniel and Seong-gee Um
Don Mills, ON: Oxford University Press 2015
ISBN 978-0-19-544280-9
Softcover $58.95, 256 pp.

Reviewed by Anthony C. Masi
Desautels Faculty of Management, McGill University

This monograph, in the OUP series "Themes in Canadian Sociology," is well-suited as a companion text in introductory sociology courses or in upper-division offerings in any discipline dealing with public and social policies. While prevalently Canadian material is presented, there is sufficient information on other countries to place Canada in comparative perspective, which the authors often do.

McDaniel and Um have designed the volume with a decidedly pedagogical approach. Each chapter starts with specific learning objectives, briefly introduces its topic, delves into the arguments, sums up, provides questions for critical thought, and concludes with suggested readings and websites. Throughout the text, words and phrases in bold are defined fully in the glossary. In addition, there is a set of references that have been directly cited in the book or influenced its development of arguments.

The eight chapters follow a logical path: (1) sociological views of public policy, (2) theories of public policy, (3) present challenges framed by the past and an anticipated future, (4) the meaning of "security" for Canadians, (5) states and markets as forces for change, (6) equality and inclusion under policy retrenchment, (7) creating a new policy agenda, and (8) summary of themes and predicting challenges.

Readers are reminded, early and often, that Canada's "public policy architecture" is framed by a macro-level context: globalization, trends in industrial and occupational structures, changes in the economy and labour market, demographic shifts, and fluctuating discourses and paradigms regarding approaches to public and social policies. The authors are equally clear in pointing out that their book is not about the "policy process," does not go into depth about the political economy, political sociology, economics, social work, or political science of public and social policies. Neither does it contain comprehensive micro-level analyses of specific policies. Rather, the text contains illustrative examples of particular policies that highlight the sociological points under discussion.

McDaniel and Um (p. 2) consider that public policy "encompasses all actions taken by states (governments writ large) presumably in the public interest. It can include everything from economic or monetary policy to the regulation of pets." They stick to the middle ground. Social policies are treated as a subset of more general public policies. The authors define "the state" as the only legitimate institution that can tax the citizenry and use force to ensure order, both to the benefit of society. States accumulate resources and then reallocate them for the public good. The "welfare state" is taken to mean the sum of economic policies and social programs. It is variable across time and national boundaries, and as in the case in Canada, across provincial jurisdictions. The other institution that accumulates and allocates resources is, of course, "the market". This juxtaposition allows the authors to contrast social versus market citizenship and to outline the conflicts that emerge in formulating public policies.

For the authors, theories are simply explanations of why things happen, so they quickly move on to concrete expressions of what public policies actually are: reflections of a society's goals, what it values from its citizens,

and its orientation to how to achieve both. Policies are the collective products of belief systems, negotiations, bargaining, trading interests, compromises, and political processes among a society's actors and institutions (p. 19).

Complex sociological approaches are condensed into comprehensible paragraphs, leaving ample room for criticism that they have given the short-shrift to this or that perspective. In fact, by focusing on some of the grander schemes, the authors have missed opportunities to tie public policy to theories of the middle-range.

The book contains a considerable dissection of the extensive, yet still very inconclusive, debate about "varieties of capitalism." In addition to states and markets, other social actors play a role in determining the existence and shape of public policies: families, voluntary organizations, and (some aspects of) enterprises. A further useful pedagogical tool that McDaniel and Um employ throughout the volume is "the box" or sidebar. There is a worthwhile discussion, in the box and in the text (p. 31), that conceptualizes service provision as a "care diamond": purchased (markets), reciprocal (families/kin), collective (communities/voluntary associations), and solidarity (state). That diamond is extendable to a "welfare pentagon" when enterprises provide supplementary services not elsewhere available.

McDaniel and Um develop the recurrent problem of tension between the egalitarian ideal of citizenship in democracies and the reality of (growing) economic inequalities under free-market capitalism. In the face of uncertainty, what do citizens expect from their politicians? For them, globalization has driven a wedge between a once dense network of social welfare policies and the requirements of a more flexible workforce. Canada's experience, while less dramatic than elsewhere in the Anglophone world, has been what the authors characterize as a neoliberal reduction in the apparatus of the state. Enter the counter-ideology.

This retreat from Keynesianism, that had emphasised risk protection and redistribution in favour of relying on competitive market forces for the efficient allocation of resources, forms the core of the substantive chapters of this volume. McDaniel and Um develop three lines of argument. First, the "caring and sharing" approach, based on reciprocity, inclusion, and trust, which the governments took over from churches and other voluntary associations to form "welfare state programs," has been under attack since the 1970s. Neoliberals consider it to be the cause of, not the solution to, inequalities. Second, the authors insist that the "lack of policy," itself a public policy stance (p. 41), is being increasingly employed by (conservative) governments. Third, and following Stiglitz (2012), over the last four decades or so, too little attention has been given to the following facts: (a) markets have not been stable, efficient, or self-regulating, (b) states have been unable to provide appropriate correctives to market failures, and (c) citizens in the advanced democracies, including Canadians, have become increasingly skeptical regarding the fairness of their political and economic systems and the people who lead them.

The discussion of "embedded taxation" (p. 61–74) tries to link public tax and transfer policies to differences in well-being among Canadians in different provincial jurisdictions, as well as between Canada and other OECD countries. These are complex matters. How are tax burdens, ratios of taxes to gross domestic product (GDP), ratios of debt to GDP, and differences in powers and responsibilities between federal and provincial governments related to a variety of outcome measures? Can the latter be mediated by a variety of policies and/or local expression of intentions? The potential for confusion is illustrated by the data on international tax burden (Table 4.1, p. 63). These data deal only with "central governments," making Canada appear rather moderate in generating revenues from its citizens and residents. However, given the constitutional division of powers, when this information is supplemented with provincial income taxes (p. 66), the burden on Canadians increases. Further, no mention is made of the fact that in some central, rather than federal systems, property taxes are included in those tax rates. In Canada, of course, housing taxes are levied at an even lower level, being generally a municipal matter.

There is a discussion in Chapter 4 of the differences between "Thatcherism" (lower social expenditures, lower taxes, privatization, smaller government) and "Reaganomics" (reduced growth in government expenditures, reduced marginal tax rates, reduced regulation, and reduced inflation). McDaniel and Um link these approaches to the Mulroney, Chrétien, and Harper modifications of several Canadian federal programs, including employment insurance, as well as to some provincial attempts at workfare. In their exegesis of "devolution" (p. 96), the authors missed the opportunity to contrast it to the principle of subsidiarity, or to more fully engage with the notion of trade-offs when resources are truly scarce.

The authors return to their theoretical approaches (Marx, Weber, Durkheim, Marshall) in discussing poverty, income inequality, and socio-economic mobility. They also include a set empirical indicators (low-income cut-offs, low income measures, market basket measures) and provide illustrations of various policy instruments (Old Age Security, Canada/Quebec Pension Plan, Guaranteed Income Supplement, social security programs). Barriers are outlined that seem to prevent the policies from more fully adjusting the outcomes: technological advances, differences in jurisdictional minimum wage rates, performance bonuses, declining unionization, and various socio-demographic changes like family structure and assortative mating patterns.

McDaniel and Um (p. 170–71, 182) use the 2012 Quebec student protest against proposed increases in university tuition as an illustration of the "vox populi" ensuring continued accessibility to higher education in that Province. As noted above, the authors did not intend to do microanalysis of any specific policy, but the facts of this case clearly deserved more attention than contained in their illustration. Tuition rates did not go up, but neither did government support, forcing universities to cut budgets and/or to seek alternative sources of revenue. Access was not enhanced. Fact: Quebec remains with very low tuition rates, but still has very low university participation rates!

There are several demography-related entries in the index: population ageing, generation boomerang, assortative matting, First Nations life expectancy, Census of Canada (2011 long-form controversy). In the book, demographers will also find several policy examples of direct relevance: ageing population structures and pension policies, international and inter-provincial migration, childcare and other family measures, life course transitions. Each would have benefited from more data, better illustrations, and fuller discussion.

Notwithstanding some shortcomings, this is book nicely fills a gap by providing an introduction to the sociology of public policy, illustrated with Canadian examples, some important ones that are discussed in detail.

References

Stiglitz, J.E. 2012. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York: W.W. Norton & Company.

Mortality in an International Perspective

edited by Jon Anson and Marc Luy
New York: Springer 2014
ISBN 978-3-319-03028-9 (print), 978-3-319-03029-6 (online)
Hardcover $179, e-book $139, 359 pp.

Reviewed by Sulaiman Bah
Department of Health Information Management and Technology,
Imam AbdulRahman Alfaisal University,[1] Dammam, Saudi Arabia

The subject of mortality is a meeting ground for diverse disciplines, and hence it is not surprising to see it being approached from remarkably different perspectives. The book has 13 chapters, covering a wide range of topics, using data from different geographic areas. The chapters can be grouped into three broad themes: mortality estimation and projections (chapters 2, 3, and 5), explanation of trends in mortality and causes of death (chapters 4, 8, 9, 11, and 13), and measurement of impact of determinants (chapters 6, 7, 10, and 12).

Chapter 1, by Jon Anson and Marc Luy, adequately summarizes all the chapters in the book. It is clear that the aim of the book is to put together 'state of the art' methods used in mortality and morbidity. However, it is in the chapters dealing with mortality estimation and projections that "cutting edge" methods are used. The remaining chapters, interesting as they are, used methods that would fall under "normal science" rather than cutting-edge methods.

Chapter 2 is authored by Peter Congdon, a pioneer in the analysis of small area mortality and the author of books on Bayesian statistical modelling. In this chapter, he exploited correlations between adjacent ages and areas with Bayesian modelling and applied it to data of over 3,000 US counties. He found that "whereas there is little gain in life expectancy in the lowest income counties, high income counties showed expectancy improvements exceeding the US average." This new approach is an improvement on standard conventional life table methods used in small area mortality that overlook spatial or age correlations.

Chapter 3, by Joroen Spijker, is clearly the most ambitious chapter in the book. He uses data from 21 countries over the period from 1980 to 2000 to model death rates for 11 causes of death. The model used allows for the simulataneous analysis of inter-country and inter-temporal variations in mortality. As a departure from other models based on extrapolation, this model included data on some known socioeconomic determinants of mortality. The model was validated and then used to produce short-term projections of rates due to causes of death. This is a significant contribution in an area that is still in its youthful stage of development.

Chapter 4, by Katalin Kovács, thoroughly reviews the different variants of Epidemiological Transition Theory and the Nutritional Transition Theory. Using causes of death data from Hungary, Kovács tries to group the different causes of death in such a way as to allow her to see the role of the different theories in explaining inequalities in mortality between the less educated and the more educated. Her conclusion was that "nutrition transition theory provides a very plausible explanatory framework for the growth of mortality inequalities."

Chapter 5, by Sarinapha Vasunilashorrn and others, attempts to predict mortality from profiles of biological risk and performance measures of functioning. They were able to get a rich set of data by linking a national US survey data with the causes of death data contained in the National Death Index. According to the authors,

1. Formerly University of Dammam.

this was the first study "that methodologically examines the relationship of both biomarkers and indicators of functioning using the latent class approach."

Chapter 6, by V. Semoyonova and others, attempts to estimate alcohol-related losses in the Russian population. It is interesting to note that there is marked similarity between the findings in the chapter with those found in some countries affected by high HIV/AIDS-related deaths. In both cases, the underestimation comes from multiple stakeholders, including the state (to avoid negative publicity), next-of-kin (to "save face"), and certifiers and coders (innocently or in collusion with the two aforementioned stakeholders).

Chapter 7, by Maria Pizarro and others, analysed infant mortality in the context of the International Convention of the Rights of the Child (CRC). They show that the main causes of infant deaths could be prevented by simple interventions, and hence, if those causes are addressed, they would help to reduce regional inequalities and hence help Argentina to meet the CRC and related goals. The authors are skeptical about the possibility of a country presenting a "less than truthful" report to the CRC and getting away with it. Indeed, it is not easy for a country to do so, as the CRC allows for NGOs to independently present their own report (alternative report) to the CRC at the same forum that the national committee is presenting its report.

Chapter 8, by İlknur Yüksel-Kaptanoğlu and others, looks at maternal mortality in Turkey. They lament the quality of vital registration data in the country, as well as the conflicting maternal mortality estimates that are published. They try to address these shortcomings by using data obtained from a survey designed after RAMOS (Reproductive Age Mortality Survey). Their research allowed them to obtain estimates of maternal mortality, as well as identify "which avoidable factors are most important for elimination of the maternal deaths in Turkey." The study found that "household and community factors" were the most important of the avoidable factors. This is a surprising finding, given that, as the authors observe "there are no legal barriers to women's participation in education, labour force and political life…"

Chapter 9, by Rosa Gómez Redondo and others, studies Spanish old age mortality over the period 1975 to 2006, using overall and cause- and sex-specific standardized death rates. Based on the reduction in death rates for specific chronic diseases among males, they predict convergence in male and female mortality rates in the future. Such a convergence could result in revision of some of the theories on male-female mortality differentials.

Chapter 10, by Anne Herm and others, asked a very important question: whether or not institutionalization increases or decreases the risk of mortality. Using Hungarian census data, they report, "our results indicate very clearly that, except at very old ages, it is preferable for people to remain at home…" This has major implications for health care provision for the elderly. One policy implication of this finding, at least for Hungary, is the shift in emphasis from nursing homes to home health care. Replication of such a study in other countries will contribute to evidence-based decision making for health care among the elderly.

Chapter 11, by Madelin Gómez León and Esther Maria León Diaz, is on decomposition of the sex differential in life expectancy in Cuba. Close to half of the results section is used to show that "Cuba is in an advanced stage of the demographic transition." But that could have been proven in far fewer pages. Nevertheless, the authors use methods of decomposition of life expectancy by causes of death, and arrive at the conclusion that the four main causes that contribute most to the high sex differential in mortality are heart diseases, external causes of death, cancer, and cerebrovascular disease. If these diseases are addressed, life expectancy could still increase in Cuba.

Chapter 12, by T. Sabgayda and others, addresses the issue of avoidable mortality in Russia. By analyzing data in over 80 Russian regions, they found similar patterns in the spatial variation of avoidable and unavoidable mortality. After doing several comparisons, they reach the conclusion that the "larger increase in unavoidable, compared to avoidable, mortality during the post-Soviet period suggests that the social and political upheavals had a greater influence of the mortality of the Russian population than did the activities of health institutions." This innovative approach at explaining determinants of mortality change is applicable to other areas that have undergone social and political turmoil.

The last chapter, by Zhongwei Zhao and others, looks at long-term mortality changes in East Asia. They look at countries that underwent "rapid mortality transition" (Omran's accelerated model), including, Japan, Hong Kong, and Taiwan. The aim of the chapter is to draw lessons from the experience of those countries.

According to the authors, the early detection and effective prevention of specific diseases (e.g., neoplasms and CVD) helped them to achieve rapid improvements in mortality over a relatively short period of time.

In all, three chapters dealt with the concept of "preventable/avoidable" causes of death (chapters 7, 8, and 12). All three chapters defined the concept differently. This means that if all three concepts are applied to the same data, different results will be obtained. There is clearly a need for standardization of definitions in this area. In the chapters dealing with causes of death, there is a glaring omission of qualification of the cause of death. Is it a principal cause, underlying cause, or multiple cause? The reader is left to guess. Further, there is no chapter dealing with multiple causes of death or emerging infectious diseases.

While there are no obvious grammatical or typological flaws, there are a few slips. For example, the text makes reference to a blue line on a black and white graph (Fig 11.5) and to green and yellow bars in black and white graphs (Fig 11.12). As usual with similar books published by Springer, there is no list of tables or list of figures.

The book makes a useful contribution to new methods in the field of mortality and morbidity in general, and to the elucidation of trends and determinants. It is recommended for libraries and graduate programs covering the topic of mortality and morbidity. As the scope of chapters is diverse, a researcher interested in only a few of the topics has the option of purchasing only those specific chapters individually from Springer.

Population Ageing in India

edited by G. Giridhar, K.M. Sathyanarayana,
Sanjay Kumar, K.S. James, and Moneer Alam
Delhi: Cambridge University Press 2014
ISBN 978-1-107-07332-6
Hardcover $151.73, 226 pp.

Reviewed by Bali Ram
Department of Sociology and Anthropology, Carleton University, Ottawa
and Department of Sociology, Western University, London ON

Until recently, population aging—usually defined as the increasing share of older persons age 60 or older (and sometimes age 65 or older) in the population—was known as a phenomenon of the industrialized world. Now it is recognized as a global phenomenon. Almost all countries in the world, with the exception of the African continent, are experiencing growth in the number and proportions of older people in their populations. Population in the industrialized countries has been aging for quite some time. Today, more than one in four people in a number of countries are over the age of 60 years, with Japan (33 per cent), Germany (28 per cent), Italy (28 per cent), and Finland (27 per cent) being in the lead. Almost one-third of the population in the industrialized world is expected to be in the older age groups by 2050. India is still far behind in this race. Today, just about 10 per cent of India's population is 60 years or older, and this figure is not expected to exceed 20 per cent by 2050. However, the sheer number of older people—about 315 million—will have a profound impact on nearly all facets of society, particularly health and long-term care and old-age income security systems, which are still relatively underdeveloped in India. Given the relative dearth of Indian studies that would help develop evidence-based policies to meet the challenges of population aging, this book is a welcome addition to the growing body of literature on the subject.

This volume begins with a general overview chapter by Giridhar, Sathyanarayana, and James, which largely sets the tone of the chapters that follow. According to these authors, the book is intended to bring together the available evidence from existing secondary data for developing a good knowledge base on the lives and living conditions of the elderly and to provide useful policy and program insights. Also, it is supposed to focus on the vulnerabilities of older people, especially women. To some extent, the book endeavours to address these issues, though not necessarily in a comprehensive manner. In their lead article on demographics of population aging, Subaiya and Bansod (Chapter 1), present detailed Census data and population projections that offer a numerical picture of changing patterns of aging at the national and state levels. This 8-page article, which contains 4 tables, 9 figures, and 12 Appendix tables is highly simplistic in describing and explaining major trends of population aging. One key piece of demographic information that is not adequately addressed concerns the roles of declines in fertility and increases in life expectancy—the two major determinants of aging process. The fact that fertility decline is the primary factor behind population aging in India, while improved longevity plays a minor role is discussed only in passing.

Selected economic characteristics of the older population are examined in Chapter 2. Using data from various rounds of the Employment and Unemployment Surveys (National Sample Survey), Selvaraj, Karan, and Madheswaran highlight the economic vulnerability of older people. For example, they show that well over half of men and one-fifth of women aged 60 years and over are in the labour force—a phenomenon that has

changed little over the past three decades, particularly in rural areas where the rates are even higher (64 per cent for males and 25 per cent for females in 2004–05). The vulnerability of older people is also apparent in wage statistics. While the daily real wage rate (in constant 1993–94 prices) for older men increased from Rs. 59 in 1987–88 to Rs. 90 in 2004–05, it remained virtually unchanged for women at around Rs. 35 during the same period. This chapter does not provide any information at the state level.

Sathyanarayana, Kumar, and James (Chapter 3) present a highly illuminating picture of changes in the living arrangement patterns of the elderly. We learn that population aging is particularly pronounced among women, resulting in the feminization of aging and preponderance of older women—often widows—living alone. We also learn about the declining prevalence of older couples living with their children and grandchildren, and the increasing prevalence of older people living with a spouse only. The authors use only first and third waves of the National Family Health Survey and ignore the second wave. The data for all three waves would have been more useful in charting out living arrangement trends. As the authors state, they do not use Census data—an ideal source for studying trends and patterns of living arrangement— because unit-level Census data are not available. However, they could have used Census data on sex ratio among older people for examining feminization of aging, and on marital status for studying the rise in widowhood in older ages. Census data could have also allowed the analysis of widowhood and feminization of aging at the state level.

In Chapter 4, Alam and Karan make diligent use of data from the National Sample Surveys in describing and explaining various facets of health status of older people. Surprisingly, they do not provide any data at the state level, despite the fact there is enormous regional disparity in health conditions in India. Also, they do not discuss many important statistics by gender, such as death rates, life expectancy, incidence of diseases, self-perceived health status, or health care services utilization.

Rajan and Mishra (Chapter 5) present a critique of the National Policy for Older Persons (NPOP) formulated in 1999. The NPOP has "the primary objectives of encouraging individuals to make provision for their own as well as their spouse's old age; encouraging families to take care of their older family members; and creating in the elderly persons an awareness of the need to develop themselves into independent citizens" (p. 136). As the authors rightly point out, the NPOP has a "broad sweep" and is "difficult to implement with the limited organizational, financial and management resources" (p. 151). Thus, the inclusion of an article by Shankardass (Chapter 6) with a review of policy initiatives in a number of Asian countries such as Japan, Singapore, China, Malaysia, and Thailand is highly appropriate for appraising India's programs and policies on aging in a relative context. This chapter would certainly be beneficial to scholars and policy analysts, but it would have been much more useful had the author devoted some space to the exposition of demography and living conditions of older people in those countries. In the final Chapter 8, Siva Raju presents a comprehensive review of the literature on aging in India in order to show the gaps in this area of research. For example, he states that a majority of the studies conducted so far are "exploratory and descriptive," and "localized and based on sample surveys on specific segments of elderly." This chapter should have appeared immediately after the Introduction, or could have been part of the introduction.

Overall, this volume has four major shortcomings. First, most of the articles are highly descriptive, lacking in-depth analysis. The 226-page book includes 65 tables and 34 figures and is much like a compendium of statistics on aging in India. How aging and socioeconomic changes affect one another is discussed only superficially. Second, there is a serious lack of analyses based on individual-level data, although public-use microdata files from various surveys (e.g., National Family Health Survey) are readily available. In the absence of multivariate analyses based on individual-level data, it becomes difficult to disentangle underlying mechanisms through which the aging process and public policy affect demographic, economic, and health behaviours of the elderly. Third, there is a neglect of the use of available data in describing and explaining the conditions of the elderly. Except in Chapter 1, Census data have not been utilized in an appropriate manner. Data from the 2011 Census may not have been easily available to the authors by the time of the publication of this volume, but I wonder why data from the previous censuses were completely ignored, even when describing historical trends. Recently, the Government of India's Central Statistical Organization (2011) produced a very detailed study of the demographic, social, economic, and health conditions of the elderly, along with a description of national policies and programs for welfare of the elderly. Although the book under review was published three years later, it does not

make concerted effort to make use of the rich data source. Finally, except for the chapter on policy initiatives in selected Asian countries, no other chapter includes any discussion on international comparison of population aging. So much international data on older people are now available from organizations such as United Nations, World Bank, and World Health Organization. Despite flaws and weaknesses, however, this book should serve well as a sourcebook for students, policy analysts, and policymakers. Hopefully, it will encourage researchers to undertake comprehensive studies on various facets of population aging.

Reference

Central Statistical Organization, Government of India. 2011. *Situation Analysis of the Elderly in India*. New Delhi: Central Statistical Office, Government of India.

Analysing China's Population: Social Change in a New Demographic Era

edited by Isabelle Attané and Baochang Gu
Dordrecht: Springer Science & Business Media 2014
ISBN 978-94-017-8986-8
Hardcover $129, 270 pp.

Reviewed by Yan Wei
University of Finance and Economics, School of Public Administration (PRC)

Along with its rapid socioeconomic development over the past few decades, China experienced profound demographic transitions, and has entered into a new stage of demographic development. In the past half-century, China has completed the demographic transition and has become a country with low population growth. China's 2010 census confirms the new demographic era, characterized by prolonged low fertility, persistently elevated sex ratios, rapid aging, massive urbanization, and widespread geographic redistribution (Cai 2013). The results of 2010 census confirm a series of demographic changes that had been largely foreseen by demographers and reveal few unexpected trends. Isabelle Attané and Baochang Gu's collection of essays titled *Analysis China's Population: Social Change in New Demographic Era* aims to address various defining patterns of China's demographic landscape in the early twenty-first century, some of which pose severe challenges to China's government.

The 297-page collection includes 13 papers in three parts. In the first chapter, titled "China's demographic in a changing society: Old problem and new challenge," Isabelle Attané and Baochang Gu give a comprehensive review of demographic vicissitudes and social change in last few decades, and also a brief introduction of the 12 papers collected in the book. The six papers of the first part, entitled "China's low fertility: Facts and correlates," focus on concerns relating to recent fertility trends. In chapter 2, titled "China's low fertility: Evidence from the 2010 census", Zhigang Guo and Baochang Gu argue that the 2010 census reflects the true level of fertility which is far from adhering to the official TFR of 1.8. In Chapter 3, titled "Changing pattern of marriage and divorce in today's China," Jiehua Lu and Xiaofei Wang state that first marriage is increasingly delayed for both men and women, the age-specific proportions of unmarried people are growing, and divorce is now better accepted socially. In Chapter 4, titled "Education in China: Uneven progress," Qiang Ren and Ping Zhu indicate that China has achieved significant improvement in education, while progress is uneven and gaps remain, in particular between gender, ethnic groups, provinces, and place of residence. In Chapter 5, titled "The male surplus in China's marriage market: Reviews and prospects," Shuzhuo Li, Quanbao Jiang, and Marcus W. Feldman estimate the male surplus in the population and investigate the possible social and individual consequences of the male-biased sex structure. In Chapter 6 titled "Being a woman in China today: A demography of gender," Isabelle Attané draws up a socio-demographic inventory of the situation of Chinese women in demographic and socioeconomic transition.

The three papers of the second part, entitled "Modernization, social change, and social segregation," focuses on various dimensions of social inequality that have emerged or grow more acute with the transformation of economic system. In chapter 7, titled "Are China's minority nationalities still on margin?", Dudley Poston and Qian Xiong conclude that Chinese minorities are socially different from the Han majority due to centuries of spatial segregation. In chapter 8, titled "Demographic and social impact of internal migration in China," Delia Davin focuses on rural-to-urban migration flows and their impact on age and sex structure, people left behind

in village, gender relationship. In chapter 9, titled "China's urbanization: A new 'leap forward,'" Guixin Wang argues that sustainable urbanization calls for more balanced development between urban and rural economics, and the necessary integration of rural migration into urban life.

The three papers of the third part, entitled "Changing age structure, labour force and the older population," address what will probably be the greatest challenge in the coming decades: the shrinkage of the working-age population that will accompany population aging. In chapter 10, titled "Urban-rural housing inequality in transitional China," Yanjie Bian and Chuntian Lu review the recent reforms in housing allocation system and argue that housing is an important constituent of inequality in both rural and urban China. In chapter 11, titled "Mortality in China: Data source trends and patterns," Zhongwei Zhao, Wei Chen, Jiaying Zhao, and Xianling Zhang address the important issue of mortality in the context of epidemiological transition. In chapter 12, titled "Dwindling labour supply in China: Scenarios for 2010–2060," Michele Bruni projects future trends between 2015 and 2030, and finds that China will be affect by a much sharper decline in its working-age population. In chapter 13, titled "The economic support system and changing age structure in China," Sanh-Hyop Lee and Qiulin Chen provide insight into some important features of recent changes in intergenerational resource allocation in China.

This collection of papers features comprehensiveness and authoritativeness. The editors clearly attempt to include the research on China's population in all aspects and the effect of population on social economy in a new situation. The essays collected are all recent research by scholars at home and broad who have made great achievement in the research of China's issues. The data utilized, primarily based on China's sixth population census data in 2010, is supplemented by all previous census data to analyze all aspects in detail, containing fertility, marriage patterns, internal migration, mortality patterns, urbanization, gender structure, and age structure. Meanwhile, minority and female populations are also analyzed, and some issues such as social inequality are discussed. The collection remains comprehensive and close to practice in analyzing and judging problems of population society, and it is informative and reliable in the data materials, thus fulfilling the intentions of the editors. The collection of papers serves as a high reference value regarding overall trends and characteristics of China's population as well as consequence of social economy, and also as an important book with regard to the study of China's population development and social transformation.

However, as a whole, this collection contains various papers but with numerous instances of jumbled and unclear categorization. Though editors try to connect socio-economic development with all chapters in the opening chapter, the 12 collected papers divided into 3 parts mostly deal with formal demography, also researches on sub-population, social inequality, and economics. There seems to be no definite standard for selecting papers and no careful classification of these papers.

In addition, several population issues touching on the Chinese circumstances are not involved in this book. From the point of view of demography, there is almost no deep research on two aspects. The first one is Chinese aging population, which is an existing problem but also a new one. Aging is slightly mentioned in many chapters of the book, but there is no special section for the Chinese aging problem, which includes, for instance, the recent situation and prediction of aging population, the features and distribution of the elderly, or the social, economic, and health conditions of the elderly. The second one is changes in Chinese family composition. The results of China's 2010 census show that the average size of family households was 3.10 persons, compared with 3.44 and 3.96 persons in the 2000 and 1990 population censuses, respectively. Family is one of the significant factors of population processes, whose size, structure, and type determine the dynamic trends and characteristics of population. In addition, other demographic issues with Chinese characteristics, such as the one-child family, parents who lost their only child, and transformation of fertility attitudes are worth further study.

After three decades of strict birth control, the rapid growth of China's population has been curbed. On 29 October 2015, the Chinese government declared full implementation of a universal two-child policy, which has drawn a heated discussion among media and individuals (Jiang and Cullinane 2015; Xinhua 2015). It is the second adjustment of fertility policy, after a selective two-child policy announced in November 2013, which means the end of 35 years of strict family planning policies. The universal two-child policy can mitigate the problems mentioned in this collection of essays, such as sex ratio, aging of population, and labour shortage.

However, it may take some time before the policies gradually take effect, due to the longevity of the population process. The two-child policy may help China to solve some population-related challenges, but cannot reverse the general demographic trends. More integrated social policies should be designed and implemented at the same time, as a more effective means for China to tackle future development challenges, including population problems (Peng 2013). The challenges and opportunities for China's population with the new fertility policy certainly warrant a look.

References

Cai, Y. 2013. China's new demographic reality: Learning from the 2010 census. *Population and Development Review* 39(3):371–96.

Peng, X. 2013. China's demographic challenge requires an integrated coping strategy. *Journal of Policy Analysis and Management* 32(2):399–406.

Jiang, S., and S. Cullinane. 2015. China's one-child policy to end. CNN, 29 October 2015. http://edition.cnn.com/2015/10/29/asia/china-one-child-policy/

Xinhua. 2015. China to allow two children for all couples. *Xinhua*, 29 October 2015. http://www.chinadaily.com.cn/china/2015-10/29/content_22312495.htm

Les théories de la fécondité

edited by Henri Leridon
Paris: INED 2014
ISBN 978-2-7332-2024-5
Softcover $43.95, 512 pp.

Reviewed by Roderic Beaujot
Department of Sociology, Western University, London (Canada)

This volume on fertility theories is an excellent addition to INED's *Série des Textes fondamentaux*, which aims to highlight fundamental contributions to theoretical and conceptual development in the various areas of demography. As in the other volumes in this collection, there is a substantial integrative introduction, followed by 23 chapters taken from previously published journal articles or book chapters. The individual chapters are divided into three sections: (1) precursors (Malthus, Dumont, Landry, and Ryder), (2) structured analytic frameworks (Davis-Blake framework, Henry on natural fertility, Coale on the conditions for fertility decline, and Bongaarts on intermediate factors), and (3) thematic theoretical developments (with sub-sections entitled: historical, anthropological, institutional/political, micro-economic, sociocultural and values, birth control and family planning, gender, and other). Five of the chapters were previously published in French, and the other 18 are translated from English. Two of the 23 chapters were originally published in 1817 and 1890, another in 1909, then nine are from 1956–80, and eleven were published in the period 1981–96. While it must have been difficult to make a selection of 23 texts, this reviewer finds that excellent choices were made. However, if the collection had been put together in 2015, I would have added Goldscheider, Bernhardt, and Lappegard's "The gender revolution: Understanding changing family and demographic behavior" (2015). Besides covering the field and having survived the test of time, the chapters themselves are mostly succinct (an average of 20 pages, including references).

As editor, Henri Leridon is particularly conscientious to select theoretical statements that are useful in the context of empirical analyses. For instance, he makes only passing reference to what the French have called "Maltusianisme de pauvreté" (in a context of acceptance of birth control, reductions in living standards can bring fertility decline), because this has only been used in Latin America. He also gives little credit to aggregate-level studies that may explain a large proportion of the variance across areas, but with widely different factors entered into regression equations from study to study. Leridon (p. 13) asks how it can be that two studies each explain 90 per cent of the variation, but each with different factors. That is, while the focus is on concepts and theory, the editor pays considerable attention to developments in data collection, measurement and methods: from period to cohort and parity, from vital statistics to surveys on Knowledge Aptitude and Practice (KAP), later the World Fertility Surveys, and the Demographic and Health Surveys. However, the Gender and Generation Surveys are not mentioned.

Leridon's own introduction is also succinct (30 pages), plus 150 references, of which 20 were published since the latest article (1996). He also demonstrates a healthy acceptance of theoretical diversity in the discipline, and for lack of theoretical closure. For instance, he ends with two puzzles that he feels have not been answered: the baby boom and variations in low fertility. Elsewhere, this reviewer has suggested theoretical explanations of these two puzzles. With the benefit of hindsight, might the baby boom not be understood as a period between two transitions, where the opportunity structures of young adults were favourable to young ages at marriage,

and gender structures involved the breadwinner model, in the context of limited access to sexual intimacy outside of marriage and inefficient contraception (Kerr and Beaujot 2016: 229). In understanding low fertility "lite" in Canada, I have proposed that we should consider the "U.S. model" in Alberta and the "Nordic model" in Quebec, with variations of these two models across the country (Beaujot and Wang 2010; Beaujot et al. 2013). That is, in some contexts, the opportunity structures of young adults permit labour force withdrawals for childbearing, in spite of limited parental leave benefits. In Quebec, and to a lesser extent elsewhere, the Nordic model may be operating (Roy and Bernier 2006), with family support in terms of parental leave, child care, and direct benefits to parents (Child Tax Benefits, Universal Child Care Benefit). More generally, with the help of studies based on the Gender and Generation Surveys, might we not say that the variation in low fertility is a function of: gender structures, opportunity structures of young adults, and extent of state support for family benefits that encourage a two-income model.

While I see Leridon's lack of theoretical closure as a healthy attitude, my own orientation is to view the demographic transition as a viable unifying framework, especially if it integrates demographic, structural, and cultural factors (Kerr and Beaujot 2016: 76–79). The demographics of the demographic transition include improved child survival and contraceptive efficiency. The structural elements include the extent to which families are units for economic production and units for the security of dependents, in contrast to other economic entities and the state (e.g., family policy). These structural questions include family nucleation à la Caldwell, opportunity structures of young adults, along with gender structures associated with women's agency, and the gender division of paid and unpaid work. MacDonald (2000) develops this concept, arguing that fertility will be particularly low if women have equity in access to the public sphere (especially for education and employment), but family structures force women to be responsible for unpaid work and the care of dependents. The cultural elements in the demographic transition include "moral restraint" à la Malthus, the acceptance of individual agency in the control of childbearing, individualism à la Lesthaeghe, and values on the side of diversity in accepting alternate family forms (including childlessness). These ideas are also rendered by Coale's necessary conditions for fertility decline, that Lesthaeghe and Vanderhoeft (2001) have rendered as "ready" given structural conditions, "willing" given the cultural context, and "able" given contraceptive efficiency and the associated steep learning curve.

This text should be required reading for students of demography. Henri Leridon's introduction is food for thought to all persons interested in understanding fertility. Since most of the articles are already available in English, let us hope that the introductory chapter will soon be translated.

References

Beaujot, R., C. Du, and Z. Ravanera. 2013. Family policies in Quebec and the rest of Canada: Implications for fertility, child care, women's paid work and child development indicators. *Canadian Public Policy* 39(2):221–39.

Beaujot, R., and J. Wang. 2010. Low fertility in Canada: The Nordic Model in Quebec and the U.S. Model in Alberta. *Canadian Studies in Population* 37(3–4):411–43.

Goldscheider, F., E. Bernhardt, and T. Lappegard. 2015. The gender revolution: Understanding changing family and demographic behavior. *Population and Development Review* 41(2):207–39.

Kerr, D., and R. Beaujot. 2016. *Population Change in Canada*. 3rd edn. Don Mills, ON: Oxford University Press.

Lesthaeghe, R., and C. Vanderhoeft. 2001. Ready, willing and able: A conceptualization of transitions to new behavioral forms, in *Diffusion Processes and Fertility Decline: Selected Perspectives*, edited by John B. Casterline. Washington: The National Academies Press, p. 240–64.

McDonald, P. 2000. Gender equity in theories of fertility. *Population and Development Review* 26:427–39.

Roy, L., and J. Bernier. 2006. *La politique familiale, les tendances sociales et la fécondité au Québec : Une expérimentation du modèle nordique?* Québec : Ministère de la Famille, des Aînés et de la Condition féminine.

The International Handbook on Gender, Migration and Transnationalism: Global and Development Perspectives

edited by Laura Oso and Natalia Ribas-Mateos
Cheltenham (UK) and Northampton, MA: Edward Elgar 2013
ISBN 978-1-78195-146-0
Softcover $70, 512 pp.

Reviewed by Meng Yu
Department of Sociology, Memorial University of Newfoundland

As the editors remark, the scope of this book is ambitious. The book attempts to provide a comprehensive discussion of gender, migration, development, and transnationalism by bridging studies of the topic. In addition to a thorough introduction, the publication consists of six parts with twenty articles, the majority of which are reprints of the authors' previously published papers. To link globalization, gender, and migration together, the book adopts Saskia Sassen's conceptual framework. According to Sassen, starting as early as the 1980s, economic globalization processes (structural adjustment programs, opening up to foreign capital, and removal of state subsidies in the Global South) led to a sharp rise in female migration. At the end of the twentieth century, transnationalism reached a particularly high degree of intensity on a global scale, as a result of globalization and technological changes.

Prior to the 1990s, gender was relatively overlooked in studies of the links between migration and development. Although the compilation's first three parts all have the word "development" in them, the concept of development is not clearly defined. As pointed out in the book, the European conceptions of the link between migration and development were based on a modernizing vision of this relationship, which defines development exclusively within economic paradigms. Although both the collection's editors and some authors lament the lack of adoption of the concept of "human development" in relevant research, "human development" is not clearly defined in the book, either. Chapter four may shed some light on this concept, when it attempts to explain what "smart economics" is. "Smart economics" rationalizes "investing" in women and girls for more effective development outcomes, rather than "promoting women's rights for their own sake" (Chant 2013: 97).

The first chapter in Part One, by Beneria and associates, elaborates on the "feminization" of international migration framework. Around the year 2000, the average emigration rate of tertiary-education women from non-OECD countries was 17.6 per cent, exceeding that of men. The authors point out that although much of the early literature on international migration was focused on the "brain drain," little attention was given to the gender dimensions of the phenomenon. However, readers who expect to read about the gender aspect of the "brain drain" would be disappointed. This collection includes only one case study on skilled immigrant women, those who migrate from Poland to the UK.

Although Part Two of the book is devoted to "new theoretical and methodological issues in the study of female migration and development," the two articles in it do not seem to have covered methodological issues. Nonetheless, a strength of this book is the diverse methodologies used in its case studies: quantitative, qualitative, and mixed methods. Qualitative methods used in the book include in-depth interviews and ethnography. Among the quantitative methods, some articles use rather simple descriptive statistical analyses, and some adopt such complex models as Bayesian networks. However, a common problem with these quantitative articles is

their lack of explanation as to why the methods are used. Another noteworthy point is that the majority of the case studies are from either Spain or Latin America, probably due to the fact both editors are based in Spain.

Part Three is composed of three case studies which examine immigrants and their children's transnational engagement. The authors point out that maintaining a transnational household and engaging in pendulum migration behaviour are strategies immigrants use to cope with structured discrimination and barriers to integration in the host society. An intriguing case study is on West African immigrants in the US, who send their children back to Africa to be raised in order to solve problems of discipline and avoid what they see as harmful interference by legal authorities. Still, none of the three articles seems to have a clear gender focus. One possible reason could be that the articles were either originally published somewhere else or written for a different project, and when they were revised, a critical gender perspective was not really adopted.

The authors in Part Four have different opinions on what transnationalism means to immigrants and their children. Catarino and Morokvasic believe that transnational spaces can be both obstacles and sources of opportunity. However, according to Hondagneu-Sotelo, transnationalism or post-nationalism does not provide a viable framework for immigrant rights. With two theoretical pieces and three case studies, using both quantitative and qualitative methods, Part Four is devoted to transnationalism. This section raises and tests some important theories on the gendered dynamics of integration and transnational engagement especially among second-generation immigrants. According to the article by Bachmeier et al., when compared with visiting parents' country of origin and remitting, the intention to return to live in parents' origin country constitutes the greatest attachment to the origin country. The results of their research show that second-generation women in Europe are less likely to entertain the possibility of return migration and second-generation men who are highly integrated into the host-country economy are the most likely to entertain the possibility of return.

Part Five of the book is titled "Global Production." However, its three articles are each on female immigration to Spain, remittances sent to Ecuador, and remittances from Spain to Ecuador. None of the articles seems to have a strong theoretical base on the relationship between gender and migration.

The book concludes with a section on the "global care chains." According to Setien and Acosta, "the concept of global care chains has come to problematize the issue of replacement as one of the strategies used by women to solve the care crisis," and "it has also been identified as a structural cause of gender inequality." Their article examines caregivers' access to social rights and social citizenship, an important issue in the feminization of international migration. However, another article on the temporary migrations of Riffian women between Morocco and Europe, based on ethnography, does not seem to give a very clear account of the subject matter.

Two themes that could have been developed more in-depth in this book are (1) human development and how that is related to gender justice in migration, and (2) whether transnationalism represents opportunities or obstacles for migrants of different genders. A lack of discussion on methodologies used in migration studies, and how they can apply to the study of feminization of international migration, is a weakness in this book. Last but not least, this international handbook on gender, migration, and transnationalism could become a stronger tool for students of migration if it were to include case studies from more diverse geographic locations.

Publications received

Adida, C.L., D.D. Laitin, and M-A. Valfort. 2016. *Why Muslim Integration Fails in Christian-Heritage Societies*. Harvard University Press. ISBN 978-0-674-50492-9.

Alexius, S., and K.T. Hallström (eds.). 2014. *Configuring Value Conflicts in Markets*. Edward Elgar Publishing. ISBN 978-1-782-54 446-3.

Anson, J., and M. Luy (eds.). 2014. *Mortality in an International Perspective*. European Studies of Population 18. Springer. ISBN 978-3-319-03028-9.

Artal-Tur, A., G. Peri, and F. Requena-Silvente (eds.). 2014. *The Socio-Economic Impact of Migration Flows: Effects on Trade, Remittances, Output, and Labour Markets*. Springer. ISBN 978-3-319-04077-6.

Attané, I., and B. Gu (eds.). 2014. *Analyzing China's Population: Social Change in a New Demographic Era.* Population Studies 3. Springer. ISBN 978-9-4017-8986-8.

Bacas, J.L., and W. Kavanagh (eds.). 2016. *Border Encounters: Asymmetry and Proximity at Europe's Frontiers*. Berghahn. ISBN 978-1-78533-219-7.

Beckford, J.A. (ed.). 2016. *Migration and Religion.* Vol I & II. International Library of Studies on Migration 18. Edward Elgar. ISBN 978-1-78347-257-4.

Bronstein, J.M. 2016. *Preterm Birth in the United States: A Sociocultural Approach*. Nature Series. Springer. ISBN 978-3-319-32713-6.

Carmichael, G.A. 2016. *Fundamentals of Demographic Analysis: Concepts, Measures and Methods*. Demographic Methods and Population Analysis 38. Springer. ISBN 978-3-319-26254-6.

Connell, J., and R.P.C. Brown (eds.). 2015. *Migration and Remittances*. International Library of Studies on Migration 17. Edward Elgar. ISBN 978-1-78347-946-7.

Darroch, G. (ed.). 2014. *The Dawn of Canada's Century: Hidden Histories*. McGill-Queen's University Press. ISBN 978-0-7735-4252-5.

Dicken, P. 2015. *Global Shift: Mapping the Changing Contours of the World Economy*. 7th edn. Guilford Press. ISBN 978-1-4625-1955-2.

Dufty-Jones, R., and J. Connell (eds.). 2014. *Rural Change in Australia: Population, Economy, Environment*. Surrey (UK): Ashgate. ISBN 978-1-4094-5204-1.

Eisenberg, C.; translated from the German by D. Cohen. 2016. *The Rise of Market Society in England, 1066–1800*. Berghahn. ISBN 978-1-78553-217-3.

Figart, D.M., and T.L. Warnecke. 2015. *Handbook of Research on Gender and Economic Life*. Edward Elgar. ISBN 978-1-78254-749-9.

Freeman, G.P., and N. Mirilovic (eds). 2016. *Handbook on Migration and Social Policy.* Edward Elgar. ISBN 978-1-78347-628-2.

Frey, W. 2015. *Diversity Explosion: How New Racial Demographics Are Remaking America.* Brookings Institution Press. ISBN 978-0-81-572398-1

González de Molina, M., and V.M. Toledo. 2014. *The Social Metabolism: A Socio-Ecological Theory of Historical Change.* Environmental History 3. Springer. ISBN: 978-3-319-06357-7.

Greely, H.T. 2016. *The End of Sex and the Future of Human Reproduction*. Harvard University Press. ISBN 978-0-674-72896-7.

Greenwood, M., S. de Leeuw, N.M. Lindsay, and C. Reading (eds.). 2015. *Determinants of Indigenous Peoples' Health in Canada and Beyond the Social.* Canadian Scholars Press. ISBN 978-1-55130-732-9.

Grossbard, S. 2015. *The Marriage Motive: A Price Theory of Marriage, How Marriage Markets Affect Employment, Consumption, and Savings*. Springer. ISBN 978-1-4614-1622-7.

Howell, F.M., J.R. Porter, and S.A. Matthews (eds.). 2016. *Recapturing Space: New Middle-Range Theory in Spatial Demography*. Springer. ISBN 978-3-319-22809-9.

Ike, S. 2016. *Fertility Decline and Background Independence: Applying a Reaction-Diffusion System as a Stochastic Process.* Springer. ISBN 978-4-431-55108-8.

Kaiser, W., S. Krankenhagen, and K. Poehls (transl. from German). 2016. *Exhibiting Europe in Museums: Transnational Networks, Collections, Narratives, and Representations*. Berghahn. ISBN 978-1-78533-260-9.

Kettunen, P., S. Michel, and K. Peterson (eds.). 2015. *Race, Ethnicity and Welfare States an American Dilemma?* Globalization and Welfare Series. Edward Elgar Publishing. ISBN 978-1-784-71536-6.

King, S., and A. Winter. 2016. *Migration, Settlement and Belonging in Europe, 1500–1930s*. Berghahn. ISBN 978-1-78533-218-0.

Kline, R.B. 2016. *Principles and Practice of Structural Equation Modeling*. 4th edn. Guilford Press. ISBN 978-1-4625-2334-4.

Knoepfler, P. 2016. *The Life-Changing Science of Designer Babies*. World Scientific. ISBN 978-9-814-67853-7.

Lamond, I.R., and L. Platt. 2016. *Critical Event Studies Approaches to Research*. Palgrave Macmillan. ISBN 978-1-137-52384-6.

Lebow, R.N., P. Schouten, and H. Suganami (eds.). 2016. *The Return of the Theorists: Dialogues with Great Thinkers in International Relations*. Palgrave Macmillan. ISBN 978-1-349-57788-0.

Leridon, H. 2014. *Les théories de la fécondité*. Manuels. INED. ISBN 978-2-7332-2024-5.

Levy, G., and B. Levin. 2014. *The Biostatistics of Aging: From Gompertzian Mortality to an Index of Aging-Relatedness*. Wiley. ISBN 978-1-118-64585-7.

Lundh, C., and S. Kurosu (eds.). 2014. *Similarity in Difference: Marriage in Europe and Asia, 1700–19*. MIT Press. ISBN 978-0-262-02794-6.

Manso, B.P. 2016. *Homicide in São Paulo: An Examination of Trends from 1960–2010*. Springer. ISBN 978-3-319-13164-1.

Masood, E. 2016. *The Great Invention: The Story of GDP and the Making and Unmaking of the Modern World*. Pegasus Books. ISBN 978-1-68177-137-3.

Morgan, S.L., and C. Winship. 2015. *Counterfactual and Causal Inference: Methods and Principles for Social Science Research*. 2nd edn. Cambridge University Press. ISBN 978-1-107-69416-3.

Morland, P. 2014. *Demographic Engineering: Population Strategies in Ethnic Conflict*. International Population Studies. Routledge. ISBN 978-1-4724-4164-5.

Neusser, K. 2016. *Econometrics*. Time Series. Springer. ISBN 978-3-319-32861-4.

Novak, M., L. Campbell, H.C. Northcott. 2014. *Aging and Society: Canadian Perspectives*. 7th edn. Nelson Education. ISBN 978-0-17-656226-7.

Oso, L., and N. Ribas-Mateos. 2015. *The International Handbook on Gender Migration and Transnationalism: Global and Development Perspectives*. Edward Elgar. ISBN 978-1-78195-146-0.

Peet, R., and E. Hartwick. 2015. *Theories of Development: Contentions, Arguments, Alternatives*. 3rd edn. Guilford Press. ISBN 978-1-4625-1957-6.

Penman, J. 2015. *Biohistory*. Cambridge Scholars. ISBN 978-1-4438-7165-5.

Pérouse de Montclos, M-A., E. Minor, and S. Samrat (eds.). 2016. *Violence, Statistics, and the Politics of Accounting for the Dead*. Springer. ISBN 978-3-319-12035-5.

Philipov, D., A.C. Liefbroer, and E.J. Klobas (eds.). 2015. *Reproductive Decision-Making in a Macro-Micro Perspective*. Springer. ISBN 978-94-017-9400-8.

Phillimore, J. (ed). 2015. *Migration and Social Policy*. International Library of Studies on Migration 19. Edward Elgar. ISBN 978-1-78347-104-1.

Raymo, J.M., and M. Iwasawa. 2017. *Diverging Destinies: The Japanese Case*. Nature Series. Springer. ISBN: 978-9-81100-184-0.

Rieder, T.N. 2016. *Toward a Small Family Ethic: How Overpopulation and Climate Change Are Affecting the Morality of Procreation*. Springer. ISBN 978-3-319-33869-9.

Rindfuss, R.R., and M.K. Choe (eds.). 2015. *Low and Lower Fertility Variations across Developed Countries*. Springer. ISBN 978-3-319-21482-5.

Rogers, A. 2016. *Applied Multiregional Demography: Migration and Population Redistribution*. Springer. ISBN 978-3-319-22317-9.

Rohrbasser, J-M., and M. Rousso-Rossmann. 2015. *1939–1945: Une démographie dans la tourmente*. INED. ISBN 9-782-733-21060-4.

Rothstein, F. (2016). *Mexicans on the Move: Migration and Return in Rural Mexico*. Palgrave Macmillan. ISBN 978-1-137-55993-7.

Seguy, I., and L. Buchet. 2013. *Handbook of Palaedemography*. Springer. ISBN 978-3-319-01552-1.

Shanahan, M.J., J.T. Mortimer, and M. Kirkpatrick Johnson (eds.). 2016. *Handbook of the Life Course*. Vol. 2. Springer. ISBN 978-3-319-20879-4.

Simon, P., V. Piché, A.A. Gagnon (eds.). 2015. *Social Statistics and Ethnic Diversity: Cross-National Perspectives in Classifications and Identity Politics*. IMISCOE Research Series. Springer. ISBN 978-3-319-20094-1.

Swanson, D.A. 2016. *The Washington State Census Board and Its Demographic Legacy*. Springer. ISBN 978-3-319-25947-5.

Tepperman, L., and S. Chiang. 2015. *FLOWS: A Network Approach to Social Inequality*. Rock's Mills Press. ISBN 978-1-77244-003-4.

Ther, P. 2016. *The Dark Side of Nation-States: Ethnic Cleansing in Modern Europe*. Berghahn. ISBN 978-1-78533-195-4.

Thornhill, R., and C.L. Fincher. 2014. *The Parasite-Stress Theory of Values and Sociality, Infectious Disease, History and Human Values Worldwide*. Science & Business Media Series. Springer. ISBN 978-3-319-08039-0.

Tian, X. 2014. *The Hope of the Country with a Large Population: Theories and Practices of China's Population Transformation*. Social Sciences Academic Press and Springer-Verlag. ISBN 978-3-642-40831-1.

Volchenkov, D. 2016. *Survival under Uncertainty: An Introduction to Probability Models of Social Structure and Evolution*. Springer. ISBN 978-3-319-39419-0.

Waters, M.C., and M. Gerstein Pineau (eds.). 2015. *The Integration of Immigrants into American Society*. National Academies of Sciences, Engineering, and Medicine. ISBN 978-0-309-37398-2.

Weisz, G. 2014. *Chronic Disease in the Twentieth Century: A History*. John Hopkins University Press. ISBN 978-1-42141-303-7

Wilson, T., E. Charles-Edwards, and M. Bell (eds.). 2016. *Demography for Planning and Policy: Australian Case Studies*. Applied Demography Series 7. Springer. ISBN 978-3-319-22134-2.

Woodwell, G.M. 2016. *A World to Live in: An Ecologist's Vision for a Plundered Planet*. MIT Press. ISBN 978-0-262-03407-4.

Zhang, L., R. LeGates, and M. Zhao (eds.). 2016. *Understanding China's Urbanization: The Great Demographic, Spatial, Economic, and Social Transformation*. Edward Elgar. ISBN 978-1-78347-473-8.

Reviewers for Volume 43

Sulaiman Bah, Imam AbdulRahman Alfaisal University (formerly University of Dammam), Dammam (Saudi Arabia)

Alaka Basu, Cornell University, Ithaca NY (USA)

Roderic Beaujot, Western University (formerly University of Western Ontario), London (Canada)

Martin Daly, McMaster University, Hamilton ON (Canada)

Barry Edmonston, University of Victoria (Canada)

Amir Erfani, Nipissing University, North Bay ON (Canada)

Dean Eurich, University of Alberta, Edmonton (Canada)

Janet Fast, University of Alberta, Edmonton (Canada)

Nora Galbraith, Statistics Canada, Ottawa

Benjamin C. Harris, U.S. Census Bureau, Washington DC (USA)

Feng Hou, Statistics Canada, Ottawa

Mazaruhl Islam, Sultan Qaboos University, Muscat (Oman)

Aiva Jasilioniene, Max Planck Institute for Demographic Research, Rostock (Germany)

Nora Keating, University of Alberta, Edmonton (Canada)

Don Kerr, Western University (formerly University of Western Ontario), London (Canada)

Hiroaki Kuromiya, Indiana University, Bloomington (USA)

Kevin McQuillan, University of Calgary (Canada)

Richard Mueller, University of Lethbridge (Canada)

Richard A. Nisbett, Vanderbilt Institute for Global Health, Nashville TN (USA)

Herb Northcott, University of Alberta, Edmonton (Canada)

Robert M. O'Brien, University of Oregon, Eugene (USA)

Walter Omariba, Statistics Canada, Ottawa

Eric M. Ossiander, Washington State Department of Health, Spokane (USA)

Roland Rau, Universitat Rostock (Germany)

Anatole Romaniuc, University of Alberta, Edmonton (Canada)

Christoph Schimmele, University of Victoria (Canada)

Alan B. Simmons, York University, Toronto (Canada)

Martin Spielauer, Statistics Canada, Ottawa

Zongli Tang, Auburn University at Montgomery (USA)

Lise Thibodeau, University of Montreal (Canada)

Frank Trovato, University of Alberta, Edmonton (Canada)

Ravi Verma, Ottawa

Russell Wilkins, Statistics Canada, Ottawa

Zheng Wu, University of Victoria (Canada)

Suryakant Yadav International Institute for Population Sciences, Mumbai (India)