Empirical Economics
https://doi.org/10.1007/s00181-021-02115-2

Does economic inequality breed murder? An empirical investigation of the relationship between economic inequality and homicide rates in Canadian provinces and CMAs

Livio Di Matteo[1] · Robert Petrunia[1]

Received: 25 March 2020 / Accepted: 9 August 2021

Abstract
National and international research documents a relationship between greater economic inequality and higher homicide rates. However, much of this work uses simple cross sections at high levels of aggregation rather than longer time series of cities or districts and lacks controls for a more substantial range of confounding factors. Using longitudinal Canadian provincial-level data over the period 1982 to 2017, we occasionally find a positive correlation between inequality and homicides rates. However, the relationship between income inequality and homicide rates in Canada reverses to become negative when looking at Canadian census metropolitan areas (CMAs). Moreover, the province-level result between greater inequality and homicide rates also appears to break down once accounting for regional effects. We conclude that much of the literature that finds a relationship between greater economic inequality and homicide rates needs to be re-examined within a longer time and more disaggregated framework.

Keywords Economic inequality · Homicide rates · Regional analysis

✉ Robert Petrunia
rpetruni@lakeheadu.ca

Livio Di Matteo
ldimatte@lakeheadu.ca

[1] Lakehead University, Thunder Bay, Canada

1 Introduction

Economic inequality is now a leading topic of scrutiny in social science research as well as media attention. Widespread media coverage accompanies the release of international asset distribution data,[1] with an incessant preoccupation with the wealth and income shares of the top one percent.[2] A concern of economists is what connection exists between economic inequality and social outcomes such as crime rates, education and health status. Rowlingson (2011), Subramanian and Kawachi (2004), Lynch et al., (2001), Lobmayer and Wilkinson (2000) and Wolfson et al., (1999) tie income inequality to higher mortality rates as well as health and social problems. Wilkinson and Pickett (2009) link reduced life expectancy, lower educational attainments, more crime, higher teenage pregnancy rates and higher incidences of mental health problems to income inequality within countries. More recently, Daly (2016) argues that economic inequality elicits greater violence and crime particularly among males. This paper focuses on this latter relationship and specifically examines the empirical homicide rate-inequality link.

Previous national- and international-level work in this area documents a relationship between greater economic inequality and higher homicide rates, but this work generally uses simple cross sections rather than longer time series and often does not control for a substantial range of confounding factors. Our analysis indicates a positive relationship between economic inequality and homicide rates when looking at Canadian provinces over the period 1987 to 2017. However, this finding is suspect for a couple of reasons. First, using panel data regression techniques on data from Canadian census metropolitan areas (CMAs) over the period 2000 to 2017 that also control for confounding factors and endogeneity issues, we do not find a positive relationship between income inequality and homicide rates in Canada. In fact, the results show a statistically significant negative relationship at the CMA level.

Second, we find that the province-level relationship between economic inequality and homicides breaks down once one accounts for regional effects. We conclude that much of the literature that finds a relationship between greater economic inequality and homicide rates needs to be re-examined within a longer time framework and with tighter specifications. Regression results using only a single cross section or a short time series, and aggregated data are likely to be spurious. Thus, data having a localized area for the unit of observation and a fairly long time series, along with accounting for various confounding factors, matter when determining the relationship between economic inequality and homicide rates. A more localized analysis provides a better match for assessing the crime–economic inequality relationship. Institutional, economic and political differences are more specific to sub-national levels such as CMAs or provinces. Finally, crime tends to be local in nature.

Indeed, we believe our results are an empirical manifestation of the Yule–Simpson paradox or what is also termed aggregation bias. Yule (1903) and Simpson (1951)

[1] See, for example, the *Global Wealth Report 2020,* done by Credit Suisse, analyzed the wealth held by 5.2 billion adults around the world. See: https://www.credit-suisse.com/about-us/en/reports-research/global-wealth-report.html.

[2] See, for example: Yalnizyan (2010), Wolff (2016, 2010), Oxfam (2015), Macdonald (2014), Jackson (2015) and Freund and Oliver (2016).

described a statistical anomaly whereby aggregated data reveal a trend or result that reverses or disappears when examining data from sub-groups. This 'reversal effect' is generally seen as a paradox with examples found in areas such population or health sciences and public safety.[3] The presence of different results when two heterogeneous populations are aggregated into one can be viewed as a form of spurious correlation.[4] Resolution of the paradox[5] is beyond the scope of this paper, but its presence is an important one and should be acknowledged when examining casual relationships in data.[6]

Furthermore, given that Canadian crime trends broadly mirror those of the USA, Canadian empirical work on both crime trends and the relationship between crime and inequality is of interest to policy makers in both countries. As Curry et al. (2016) note, Canadian crime trends have moved in the same direction as the USA and exhibit the same decline in the 1990s. At the same time, while income inequality has grown in Canada, it has been more muted than the USA.[7] Canada currently has less income inequality than the USA, but Heisz (2016: 87) shows that the top 1% share of market income for Canada closely followed that of the USA from the 1930s to the early 1980s, dropping from about 18 to about 8 percent. However, Heisz also shows that by 2010 the comparison had altered radically, with the USA top 1% income share again close to 20 percent, but Canada's just over 10 percent.

Saez and Veall (2005) show a similar pattern of top income share performance, with increasing income inequality in both countries after 1975 but with the USA again showing greater income inequality growth than Canada. Further, while wealth inequality is substantially less in Canada compared to the USA, Canadian wealth inequality actually rose faster than US wealth inequality from 1984 to 2012, but since 2016 has risen faster in the USA (Davies and Di Matteo, 2021).[8] Thus, Canada offers an interesting parallel to the USA given similar crime trends but somewhat dissimilar trends in economic inequality.

The rest of the paper is organized as follows. Section 2 discusses previous literature and provides a further discussion about aggregation. Section 3 presents the data. Section 4 provides the econometric specification and results. In Sect. 5, we discuss econometric issues and provide robustness checks. Section 6 concludes this study.

[3] See Cohen (1986) and Davis (2004).

[4] Pearl (2018: 70–71).

[5] The reversal effect in Simpson's paradox can be viewed as a numerical phenomenon arising from a reversal in the relative frequency of a particular event across several different samples after being merged. One solution is to better understand the process of causation in a particular relationship so as to decide on the appropriate level of aggregation and avoid the paradox. Pearl (2018: 200–2017).

[6] For a statistical discussion of the paradox, see Pearl (2011).

[7] For a recent detailed discussion of Canadian economic inequality, see Osberg (2018).

[8] Wealth shares of the top 1 percent of Canadian families rose from 22.9 percent in 1970 to 28.7 percent in 2012 and then remained at 28.7 percent in 2016. Meanwhile, the top 1 percent of US households accounted for 31.1 percent of wealth in 1970, 35.2 percent in 2012 and 38.6 percent by 2016 Davies and Di Matteo (2021).

2 Literature and context

The substantial literature on the general determinants of violent crime and its links with economic inequality focuses on both the proximity of low-income individuals to higher-income ones and the expected gains from criminal activity, as well as the social frustration associated with relative economic failure.[9] Becker (1968) provides a seminal neoclassical theoretical framework based on economic costs and benefits to criminal activity and suggests that crime increases with inequality. The mechanism driving this result is from the income gap between the poor (potential criminals without resources) and the rich (with resources attractive to potential criminals). A greater income gap causes the net benefits of crime to increase. Thus, higher inequality raises the incentive for criminal activity, which creates the inequality and crime rate link.

The general links between economic inequality and violence have also been examined in a wide range of socioeconomic literature that places emphasis on what is perceived as the damaging role of economic competition. While economists generally view the free market and laissez-faire competitive economic structures as producing socially optimal results, Frank (2012) argues that Adam Smith's invisible hand, in which self-interest promotes the common good and positive social outcomes, is really a special case and that excessive economic competition generally wastes resources, harms group outcomes, and that any gains are relative and often only short term.

Frank and Cook (1995) maintain that "winner-take-all" processes in which first-place finishers get all of the prize, and the remainder get nothing or very little, misallocate economic resources and cause serious income inequality. Frank (2013) follows up on the effects of economic inequality by arguing that rising income and wealth inequality—especially concentrations at the very top of the economic pyramid—have set off "expenditure cascades" that raise the cost of achieving middle class goals. The middle class is losing the race for economic status that comes from the consumption of what Frank terms "positional goods" imposing important psychological costs as well as economic costs as resources are diverted into a race for status marked by what essentially amounts to conspicuous consumption.[10]

2.1 Canadian evidence

Daly (2016) builds on some of these concepts as he builds an important link between economic competition and homicides. Daly maintains that most homicides result between young men from competitive conflicts driven by competition over status and position, rivalry over women, or issues of respect and deference. Indeed, the great majority of killers and their victims are men. Rising economic inequality creates more competition for material goods, and as a result, Daly posits a link between economic inequality and homicide rates and musters evidence from the literature as well as empirical evidence to support his claim. Moreover, Daly emphasizes that it is inequality rather than poverty as captured by absolute income levels that matters. Much of

[9] For a discussion, see Coccia (2017).

[10] Conspicuous consumption refers to the public display of economic prowess via spending on luxury items and was coined by Thorstein Veblen (1899).

this evidence, however, consists of cross-sectional plots and correlations of homicide rates in a given year against some measure of economic inequality.

Daly, Wilson and Vasdev (2001) plot homicide rates in 50 US states and 10 Canadian provinces circa 1990 against Gini coefficients for total household income and find a positive relationship. They also provide some regression results beyond a simple cross section regressing Canadian provincial homicide rates for the period 1981 to 1996 on the Gini coefficients of after-tax household income inequality, median after-tax household income, province-specific dummies and annual time dummies. They find the Gini coefficient to be a statistically significant predictor of homicides estimated as a panel and the estimated coefficients of median income to be positive and not significant, with the regression demonstrating an r-squared of 83 percent. They conclude on the basis of these results that "the degree to which resources are unequally distributed is a stronger determinant of levels of lethal violence in modern nation states than is the average level of material welfare."[11] While this result does control for provinces and time and is a more flexible specification, the use of annual time dummies rather than a time trend and province-specific rather than regional dummies (e.g., West, Atlantic, etc.) reduces the degrees of freedom for the regression.

Kennedy et al., (1991) use census information on 24 Canadian CMAs for 1971, 1976 and 1981 to test the relationship between family income inequality measured using Gini coefficients,[12] homicide, social disorganization[13] and regional location. The results find that homicide in Canada exhibits a regional pattern, rising from east to west. There is a reduction in regional effects via a convergence in homicide rates between eastern, central and western Canada in census metropolitan areas combined with higher levels of inequality and social disorganization. The effect of inequality on the total homicide rate is evident for 1972–1976, but not for 1977–1981, with family income inequality in 1972–1976 positively associated with homicide. As well, the proportion of young males within a community also is found to be associated with higher homicide rates, while the unemployment rate is inversely related to homicide rates.

2.2 International evidence

There is also a substantial amount of international research on the topic. Fajnzylber, Lederman and Loayza (2002) study the correlation between the Gini index and homicide and robbery rates within and between countries. Using non-overlapping panel data consisting of 5-year averages for 39 countries during 1965–1995 for homicides and for 37 countries during 1970–1994 for robberies, they find crime rates and inequality positively correlated within countries and, particularly, between countries. Moreover, this correlation reflects causation from inequality to crime rates, even after controlling

[11] Daly, Wilson, Vadeva (2001: 231).

[12] Kennedy et al., (1991) construct their inequality measures using the Canadian census (Statistics Canada, 1973; 1982) with supplemental information gathered from the Labour Force Survey (Statistics Canada, 1977). They also use an income dispersion measure whereby the distribution of family incomes within a CMA is compared to the distribution of family income for all of Canada.

[13] Social disorganization theory argues that poorly integrated communities with breakdowns in community control have higher rates of crime. See Kennedy et al. (1991) for an overview. See also Boustan et al. (2013)

for other aggregate crime determinants. At the same time, they also find that violent crime rates decrease when economic growth improves and that faster poverty reduction leads to a decline in crime rates. However, the level of income is not a significant predictor—it is the change in income that matters.

Harris and Vermaak (2015) examine sub-Saharan African countries and South Africa's 52 districts for an association between income inequality and violence. While there is no evidence of such a relationship in sub-Saharan Africa, there was evidence of a significant positive relationship between homicide rates and inequality in South Africa. A one percent increase in inequality was found to be associated with an increase in the homicide rate of 2.3 to 2.5 percent. Moreover, this relationship remained significant even after controlling for other district characteristics. Examining data for Latin America and Caribbean countries for the period 1950 to 2010, Baier (2014) finds no link between crime rates and income inequality once accounting for possible spurious correlation due to common trends in the time series of the two variables. Enamorado et al. (2016) obtain a substantially positive relationship between drug-related homicides and income inequality measures in Mexico during the 1990–2010 period.

Coccia (2017) examines the effects of both climate and inequality as explanations of violent crime using a sample of 191 countries and finds that after controlling for thermal climate (heat hypothesis)[14] and other factors, socioeconomic inequality is positively associated with violent crime. In particular, Coccia argues the results support the hypothesis that differences between countries in intentional homicides (per 100,000 people) can be explained by the level of income inequality alone, and not thermal climate as a second predictor or the interaction of income inequality and thermal climate.

Ouimet (2012), using international data for 165 countries in 2010, demonstrates that the stage of economic development, income inequality as measured by the Gini coefficient, and poverty as measured by excess infant mortality are all significant predictors of the homicide rate for all countries. Moreover, it is income inequality and not economic development or poverty that predicts homicide for countries with a medium level of human development.

Szwarcwald et al. (1999) look at the effect of income inequality on homicide rates in the state of Rio de Janeiro, Brazil, using 1991 census data for both municipalities and administrative regions, and multiple regression analysis with controls for other socioeconomic indicators. For the municipalities of the state of Rio de Janeiro, the regression results found no association between homicide rates and income inequality, but there was a correlation at the more aggregate administrative region level highlighting the difference in results based on aggregation levels. This difference, however, was attributed to degrees of urbanization as more urbanized areas had higher homicide and inequality rates.

In another study, Roberts and Willits (2015) note that choices of inequality measures and homicide types may account for what are often mixed findings on the income inequality–homicide link. Using various inequality measures for 208 US cities, they

[14] The "heat" hypothesis argues that hot temperatures increase aggression and violent behaviour, while colder temperatures and more seasonal weather variation result in societies with more emphasis on self-control. See Coccia (2017) for an overview.

nevertheless find that all measures of income inequality had significant and positive associations with both overall and specific homicide rates.

Bailey (1984), in a critique of previous studies, uses US city-level data rather than more aggregated SMSA data, and several cross sections of data (1950, 1960, 1970), rather than one year of data, and finds a positive relationship between poverty levels and homicide rates, and only a slight and insignificant relationship between relative economic deprivation (income inequality) and homicides. Bailey notes that there is only a weak theoretical linkage between homicide and relative economic deprivation and is not surprised by the lack of a significant empirical relationship between inequality and homicide rates. Examining US county-level FBI crime data in 1991, Kelly (2000) finds a positive link between violent crime and income inequality, but no link between property crime and inequality.

Hicks and Hicks (2014) examine the impact of inequality in consumption expenditures on criminal activity using data on US states for the 1986–2001 period. The paper looks at various measures of violent and property crimes and distinguishes between income inequality, consumption inequality and visible consumption inequality.[15] The authors demonstrate that a positive relationship exists between visible consumption and measures of violent crime, but it fails to hold for measures of property crimes. Further, income inequality and total consumption inequality do little to explain the various measures of violent and property crimes.

Summarizing these studies suggests that much of the literature finds that especially at the national or regional level, inequality affects homicide rates, whereas a society's average income level does not. In other words, it is relative rather than absolute economic deprivation that is the key determinant of homicide rates. However, this relationship is not always apparent when more disaggregated data at the city or municipal level are used. Indeed, the choice of data in terms of its level of aggregation, regional groupings, time span and geographic coverage does seem to be factors in the results.

2.3 Aggregation

The level of aggregation best suited to studying the relationship between crime and inequality and empirical relationships in general is an important question and one raised in other literatures. In health economics, for example, estimates of the income elasticity of health costs and spending can vary with data at the national or regional level reporting high-income elasticities and individual-level data reporting more inelastic estimates (Di Matteo, 2003). Getzen (2000, 2006) examines the extent to which variation in health costs differs by the level at which observations are made. He finds that individual expenditure variation within a particular healthcare system is largely due to differences in health status rather than income, but across systems, morbidity has almost no effect on costs while income is more important. For nations, differences in per capita income explain over 90 percent of the variation in both time series and

[15] The authors use household characteristics and annual consumption expenditure data from the NBER Consumer Expenditure Survey (CEX) family level extracted for the period 1986 to 2001 to construct Gini coefficients by state year for income and expenditures.

cross section. He ultimately concludes that units of observation used for analysis of healthcare spending must be matched to the units at which decision making occurs.

Wilkinson and Pickett (2006) examine the relationship between population health and income inequality, and a large majority of studies (70 percent) suggest that health is poorer in societies with large income differences. However, there were substantial differences in results based on whether the inequality was measured in a large or small area with the positive inequality–population health link stronger in larger areas of aggregation. Their reasons for the unsupportive findings include small areas may be too small to reflect the scale of social class differences in society or the controlled factors may not be true confounders but merely other reflections of the scale of social stratification.

The level of aggregation also appears as a concern in the natural resource curse literature. James and Aadland (2011) point out that there are fewer "institutional or political differences" at more local levels such as city or state/provincial level. Further, they also identify that some economic policies are set by city or province planners. These factors are also relevant when considering the crime–inequality relationship. Incentives to commit criminal activity because of inequality reasons are more relevant at a local level. Criminal activities like homicide do mostly occur at local level. Inequality in one city or province is unlikely to drive criminal activity at another city or province in a country.

As a result, it is important in studies of the relationship to include varying levels of aggregation with consistent results across these levels as the strongest support for the resulting relationships. Ultimately, the choice of the level of aggregation cannot always provide a clear and definitive answer with interpretation of the results ultimately requiring judgment based on knowledge of the source of the behavioral decisions being analyzed as well as the availability of the data. Based on the principle of subsidiarity, one should drill down to the lowest level at which a decision is made and for which data are available.

Economic inequality and crime statistics in Canada can be found at a CMA, provincial and national level and in the case of our study would at minimum require the study of the relationship between homicides and inequality be at the CMA level. Canadian CMAs are defined as having a population of at least 100,000 which we believe provides a large enough population to provide income and social diversity. Therefore, we perform our analysis both at the CMA level and at the provincial level in an effort to gauge if consistent results in the relationship between inequality and homicides can be found. We do not use national-level data as the sample size of the annual time-series data available for Canada would be exceedingly small.

3 Data and trends

3.1 Data

The homicide and economic data for this study are for 26 Canadian census metropolitan areas (CMAs) and ten provinces. We limit the sample to 26 CMAs since these are the ones with the most complete data over the time span. The data are panel data and include several measures of crime as other recent economic studies dealing with crime in Canada have employed.[16] All data come from Statistics Canada's Web site and include:

(1) Provincial unemployment rate, employment levels, employment rate—Table 14–10-0018–01 (1981–2017)
(2) CMA unemployment rate, employment levels, employment rate—Table 14–10-0144–01 (1987–1995), Table 14–10-0092–01 (1996–2000), Table 14-10-0096-01 (2001–2017).
(3) Provincial Gini coefficients (based on total income)—Table 11-10-0134-01 (1981–2017).
(4) Median income level excluding zeros (2015 constant dollars)—Table 11-10-0191-01 (1981–2017).
(5) Median income level by top 1 percent, bottom 50 percent, top 5 percent—Table 11-10-0056-01 (1981–2017).
(6) Homicide rate per 100,000 people—Table 35-10-0071-01 (1981–2017).
(7) Total, violent and property crime rate per 100,000 people—Table 35-10-0177-01 (1998–2017).
(8) Police per 100,000 people—Table 35-10-0077-01—police personnel and selected crime statistics, municipal police services, annually (2000–2017) and Table 35-10-0076-01—police personnel and selected crime statistics, Canada, provinces and territories, annually (1986–2017).
(9) Provincial percentage of persons in low income after tax[17]—Table 11-10-0136-01 (1981–2017).
(10) CMA percentage of persons in low income after tax—after-tax low-income status of tax filers and dependents based on census family low-income measure (CFLIM-AT), by family type and family-type composition, annually, provincial and by CMA—Table 11-10-0018-01 (2000 to 2017).
(11) Provincial total and male population aged 15–24 years—Table 17-10-0005-01 (1981–2017).
(12) Provincial total immigrants by year—Table 17-10-0014-01 (1981–2017).
(13) Minimum wage—provincial minimum wages. Historical minimum wage rates in Canada, Employment and Social Development Canada, Open Government Licence—Canada, https://open.canada.ca/data/en/dataset/390ee890-59bb-4f34-a37c-9732781ef8a0.

Data are available from 1982 to 2017 for provinces and 2000 to 2017 for CMAs with the following exceptions.[18] First, the police per 100,000 people variable is available from 1986 to 2017 at the provincial level. Second, total, violent and property crime rates are available from 1998 for provinces and CMAs. Also, we note that information for Gini coefficients, total and male 15–24 population, and immigrants is available at the provincial level only. Finally, the province sample is balanced and the CMA

[16] See, for example, Curry et al., (2016).

[17] Low-income measure after tax (low-income measures (LIMs) are relative measures of low income, set at 50% of adjusted median household income).

[18] We perform analysis over the 2000 to 2017 period for the provincial data for comparison purposes and to make sure any results are not an artifact of the sample period.

sample is unbalanced. Table 7 in appendix contains a list of the CMAs along with their years of coverage and number of observations. This table also contains the average number of observations across the CMAs.

In terms of empirical methodology, the use of CMA rather than province-level data is superior as it adds substantially more observations and degrees of freedom to the analysis, as well as allowing for greater diversity than more aggregated provincial- or national-level data. As well, limiting the study to the Canadian federation controls for institutional and data collection differences that might affect results in the cases where international data are used.

The income data used are median total income by tax filer for all tax filers, the top 1 percent of tax filers, the top 5 percent, the top 10 percent and the bottom 50 percent. As Gini coefficients at the CMA level for tax filer income are not available, we construct three alternate measures of income inequality by taking the ratio of median total income for the top 1, 5 and 10 percent against the median total income of the bottom 50 percent. Gini coefficients are available over the 1982–2017 period at the provincial level of aggregation.[19] Although we do not have CMA-level Gini coefficients, we also consider the effect of provincial Gini coefficients on homicide rates within CMAs by assigning CMAs their respective provincial Gini coefficients.

As a further point with respect to empirical methodology, we note that the relationship between income levels, poverty and inequality, trends in income and income inequality are broadly consistent with trends in poverty with reductions in inequality usually accompanied by reductions in poverty rates. However, given that inequality is linked to other social characteristics it may be difficult to isolate its effect on outcomes such as health or crime (Lynch et al. 2004). Deaton (2003) notes that there is a concave relationship between health and income illustrating that health is negatively associated with income inequality in a nonlinear fashion. By extension, this concavity issue can affect other relationships such as that between crime and inequality.

This suggests the need to separate out the influences that may affect both crime and inequality either separately or jointly—in other words, controlling as best as possible for confounding third factors. In essence, there are direct and indirect effects of inequality on crime, and they may manifest themselves via the effects of absolute income levels as well as relative income and poverty rates. Therefore, in assessing the relationship between crime and inequality we should attempt an effort to concern ourselves with the effects not only of income, but also of poverty.

To control for poverty, we make use of the provincial percentage of persons in low income after tax as one of our variables. In Connor et al. (2019), poverty rates are also a useful proxy to control for confounding income effects and the potential for a concave, nonlinear relationship between crime and income. In the Canadian setting, another poverty variable to consider is the proportion of families below the low-income cut-offs as defined by Statistics Canada. The low-income cut-off is defined as the income threshold below which a family will likely devote a larger share of its

[19] A ratio approach to inequality measures has been used in the literature. For example, Harris and Vermaak (2015) use the ratio of the average income of the richest 10 percent of the population to the average income of the poorest 10 percent. Lobmayer and Wilkinson (2000) use the 50:10 centile ratio and disposable household income. For example, the income at the 50th centile (the median income) was twice as high as at the 10th centile (the top of the bottom decile); the 50:10 centile ratio would equal 2.0.

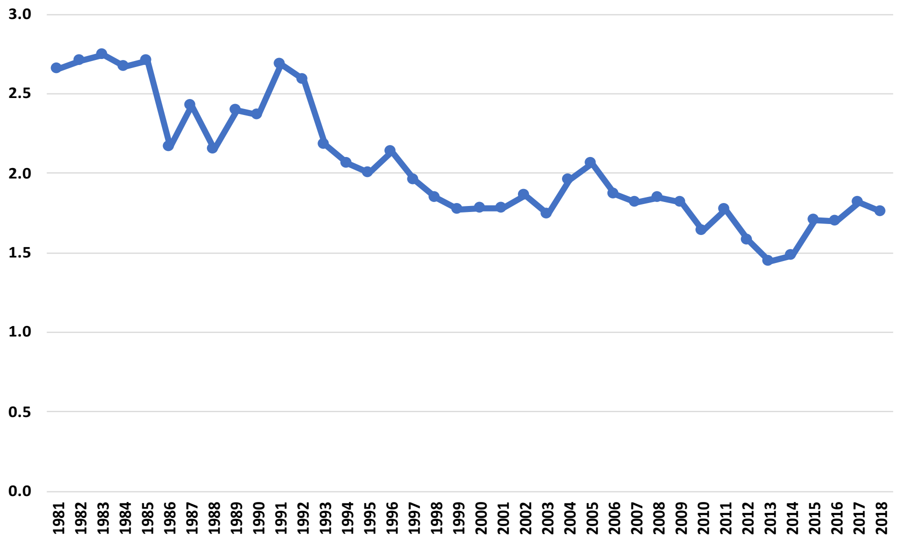

Fig. 1 Homicide rate per 100,000 population, Canada 1981 to 2018 *Source*: Statistics Canada

income on the necessities of food, shelter and clothing than the average family with this threshold estimated at 20 percentage points more than the average family. The percent of persons in each province and CMA who are below this family low-income cut-off is also used as a poverty variable.

3.2 Trends

Figures 1, 2, 3, 4 and 5 provide an overview of the data and some of the relationships we seek to explore. Figure 1 plots Canada's homicide rate per 100,000 population from 1981 to 2018, and it reveals a long-term decline. From a homicide rate of 2.7 per 100,000 in 1981, it reached a rate of 1.8 by 2018 for an overall decline in the rate of 34 percent.[20] This decline in homicide rates parallels the overall decline in crime rates in Canada, which saw the peak reached in 1991, with decline afterward. Between 1991 and 2012, the crime rate in Canada as measured by major police reported criminal code incidents fell from 10,341.1 incidents per 100,000 to 5,588.0 per 100,000—a decline of 46 percent.[21]

The factors behind this decrease in crime have been the subject of some debate, and explanations including rising prison populations, the legalization of abortion eliminating generations of future criminals, the waning crack epidemic, more effective policing, an aging population and the performance of the economy.[22] Indeed, with

[20] At the same time, the period since 2013 has actually seen a small increase in Canada's homicide rate from 1.5 to 1.8.

[21] See Di Matteo (2014: Fig. 1a).

[22] See Di Matteo (2014: 6–7), Levitt (1999a, b), Levitt (2004) and Scheider, Spence and Mansourian (2012).

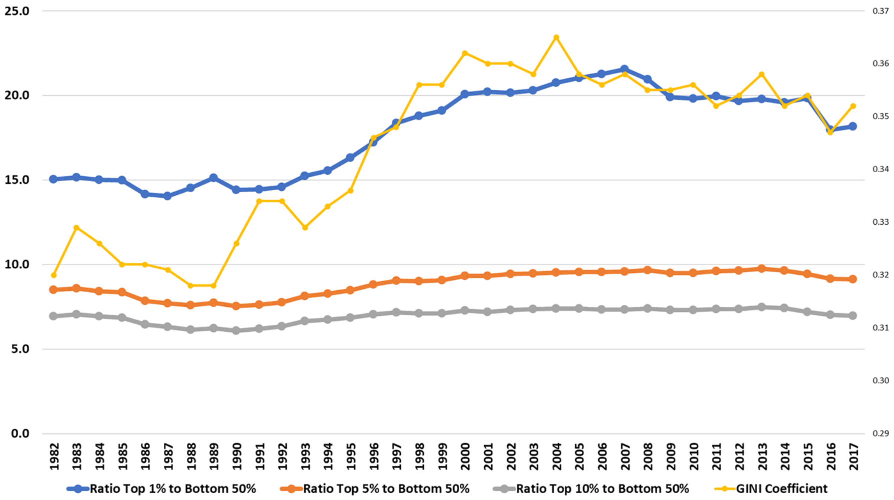

Fig. 2 Income inequality ratios, Canada, 1982 to 2017: ratios of median total income by taxfiler income group and Gini coefficient *Source* statistics Canada *Note*: For the vertical axis, left axis captures range of ratios of median total income measures, while right axis captures range of Gini coefficient values over time

respect to the economic conditions, it has been argued that economic downturns actually decrease criminal opportunities, as when unemployment is high more people stay at home serving as “guardians” of their property. During high unemployment periods, people who are out and about also are likely to carry less cash and possessions.[23]

Figure 2 plots the three measures of income inequality for Canada over the period 1982 to 2017, and it reveals that income inequality as measured by these median income ratios and Gini coefficient has grown over time though there has been some moderation of the trend since the 2008–2009 recession. The ratio of the median total income of the top 1 percent to bottom 50 percent rose from 15.1 in 1982 to 18.2 by 2017. For the ratio of the top 5 percent to the bottom 50 percent and the top 10 percent to bottom 50 percent, the increases were from 8.5 to 9.1 and from 6.9 to 7.0, respectively. Finally, the Gini coefficient increases from 0.32 in 1982 to 0.35 in 2017. Thus, the evidence from Figs. 1 and 2 suggests that at a national level, falling homicide rates have been accompanied by rising income inequality. However, these results are simply correlations that do not take into account differences across the country in both homicide rates and income inequality.

The relationship between greater inequality and higher crime rates is not automatically evident when one examines the evidence visually at the CMA level. Figure 3 plots homicide rates for Canada’s major CMAs in 2017, and they range from rates of 0 in Saguenay, to a maximum of 5.7 for Thunder Bay. When the ratio of median total income of the top 1 percent to the bottom 50 percent is ranked and plotted for a reduced set of available CMAs (Fig. 4), the range is from a low of 11.2 in Thunder Bay to a high of 20.7 in Calgary. Toronto in 2017 was the most unequal CMA with respect to income distribution in Canada, as measured by the ratio of the median total income

[23] Scheider, Spence, and Mansourian (2012).

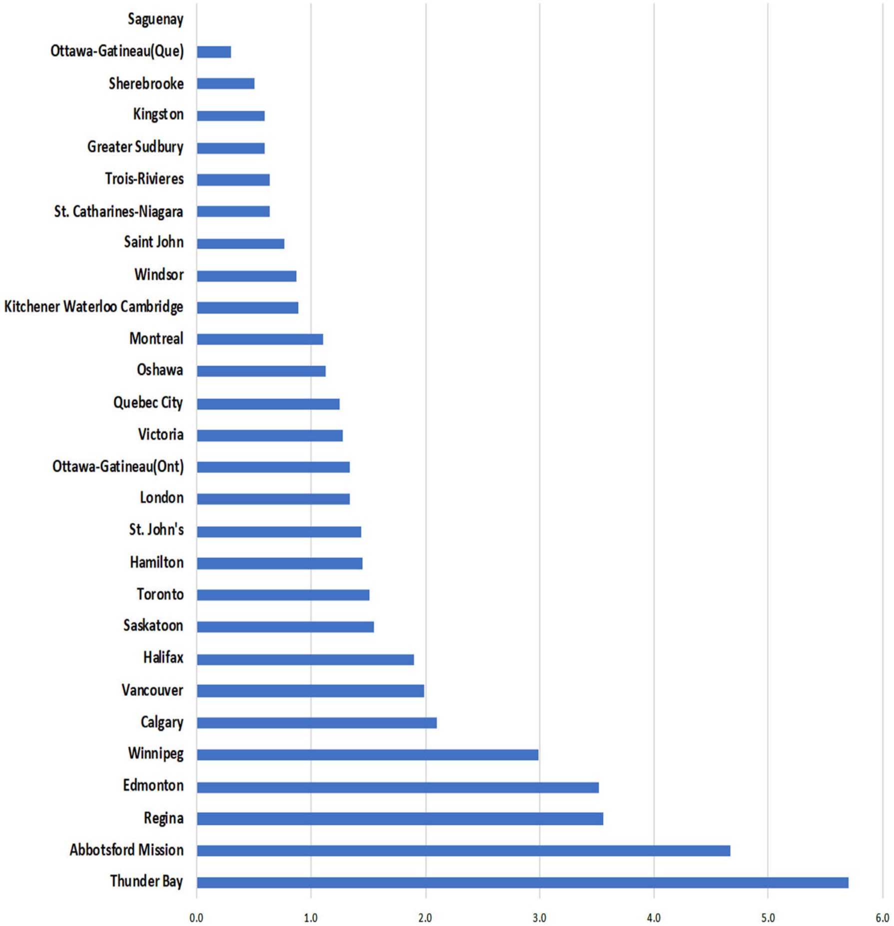

Fig. 3 Homicide rates per 100,000 population in major Canadian CMAs, 2017 Source: Statistics Canada

of the top 1 percent to the bottom 50 percent, and its homicide rate was approximately in the top third of CMAs. Calgary, which was the second most unequal, was also in the top third of CMAs for homicide rates. However, Thunder Bay, which ranked first in homicides, also was the CMA with the second most "equal" income distribution in Fig. 4, while Ottawa–Gatineau (Que) had the second lowest homicide rate and was the most "equal" of then plotted CMAs.

In assessing the broad correlation between inequality and homicide rates, Fig. 5 provides four plots done using a simple nonparametric smoothing technique: LOWESS.[24] Figure 5a to 5c plots the relationship for three years: 2017, 2000 and 1982. The 2017 cross section demonstrates a somewhat positive relationship between the homicide rate and income inequality across Canadian CMAs. The relationship is noticeably flatter in 2000 and more of an inverse u-shaped relationship in 1982. This demonstrates the

24 Locally weighted scatterplot smoothing.

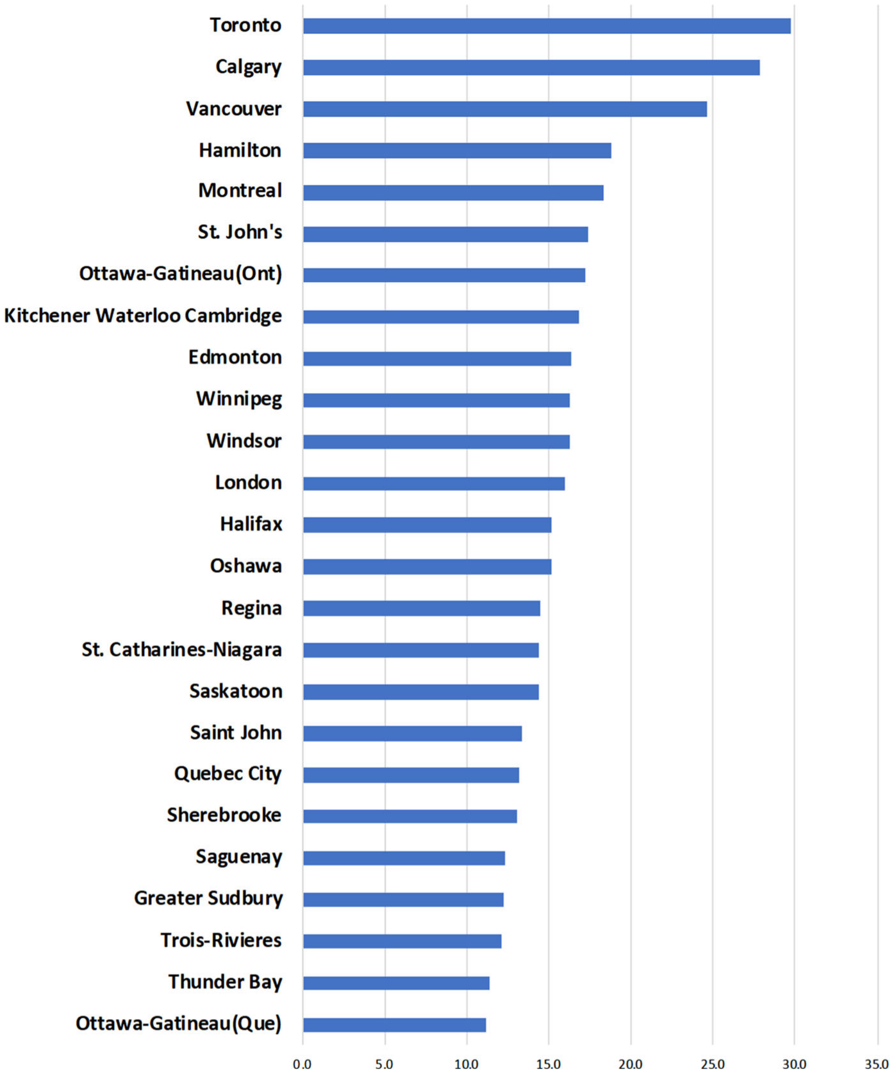

Fig. 4 Income inequality, ratio of median total income of top 1 percent to bottom 50 percent of taxfilers, Major Canadian CMAs in 2017 *Source*: Statistics Canada

obvious pitfall of relying on a single year cross section of data to draw conclusions on the relationship between income inequality and homicide rates. When all of the data from 1982 to 2017 are pooled, as shown in Fig. 5d, the relationship appears to be flat. Again, these simple cross sections do not account for the long-term decline in homicide rates over time as well as for other socioeconomic confounding factors that may be affecting homicide rates such as regional effects.

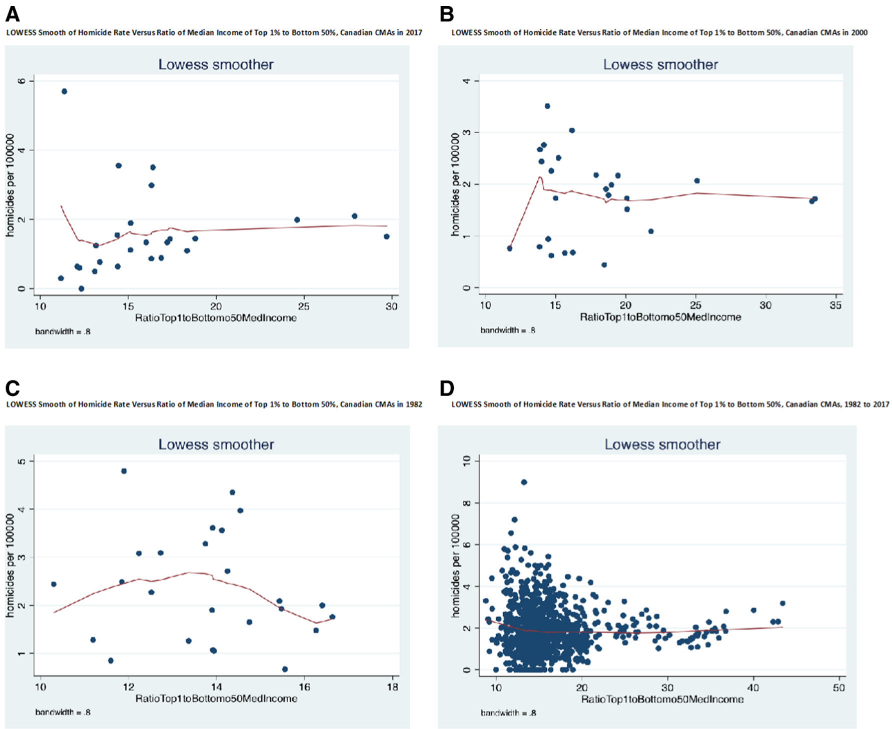

Fig. 5 **a** LOWESS smooth of homicide rate versus ratio of median income of top 1% to bottom 50%, Canadian CMAs in 2017, **b** LOWESS smooth of homicide rate versus ratio of median income of top 1% to bottom 50%, Canadian CMAs in 2000, **c** LOWESS smooth of homicide rate versus ratio of median income of top 1% to bottom 50%, Canadian CMAs in 1982, **d** LOWESS smooth of homicide rate versus ratio of median income of top 1% to bottom 50%, Canadian CMAs, 1982 to 2017

4 Estimation and analysis

4.1 Specification

To control for correlation of additional factors on homicide rates beyond income inequality, we perform regression analysis and do so for both province-level and CMA-level data. We specify the following regression:

$$y_{\mathrm{it}} = \alpha_i + Z_{\mathrm{it}}\delta + X_{\mathrm{it}}\beta + e_{\mathrm{it}}. \tag{1}$$

The regressions have the ln(1 + homicide rate) as the dependent variable y_{it}, while Z denotes the measure of income inequality and α_i is a dummy variable for the province or CMA.[25] The regressions contain two measures of income inequality: (1) the ratio

[25] For the purpose of comparison, we also perform regressions with the homicide rate and ln(homicide rate) as dependent variables. These results are available upon request. We choose report regressions with ln(1 + homicide rate) as the dependent variable as the homicide rate does equal zero on occasions, especially for CMA-level data.

of the median income for the top one percent to the median income of the bottom 50 percent (ratio top 1 to bottom 50) and (2) the Gini coefficient. X is a vector of control variables, which includes: (1) the median income; (2) unemployment rate; (3) the employment level; (4) employment rate; (5) percentage of population below the low-income cut-off; (6) police officers[26] per 100,000 people; (7) minimum wage; (8) percentage of total population who are males aged 15–24 years; (9) new immigrants per 100,000 people; (10) three regional dummy variables with an Ontario dummy excluded; and (11) a time trend. Results are reported for both provinces and CMAs. For the CMA regressions, the provincial measures of the Gini coefficient, percentage males 15–24 of total population and new immigrants per 100,000 are used and assigned to each CMA in a given province, since CMA-level information on these variables is unavailable. For both provincial and CMA data, six specifications are estimated with two specifications including only the ratio top 1 to bottom 50 variable, two specifications including only the Gini coefficient and two specifications including both measures of income inequality.

4.2 Econometric issues

Econometric issues may generate potential biases when estimating Eq. (1), which complicates this empirical relationship. We note that there is a link in the causes of these econometric issues. The first issue involves the possibility that unobserved CMA- or province-specific factors affect both homicide rates and income inequality leading to a spurious correlation between the dependent and independent variables. For example, Fajnzylber, Lederman and Loayza (2002) suggest the potential to either under- or over-report crime rates leads to measurement error. This measurement error becomes an underlying factor as the under- or over-reporting may likely be related to the factors affecting crime rates. Further unobserved heterogeneity may still be present at the CMA or province level. Thus, provincial- or CMA-specific dummy variables are included in Eq. (1) to control for any additional unobserved heterogeneity, such as the proportion of population that is Indigenous.

A second issue is the potential for homicide rates to exhibit state dependence or persistence, whereby the present values are highly correlated with the recent past values. Jurisdictions with high homicide rates in one year are likely to continue to have high homicide rates into the near future. To account for this persistence, we add the lagged homicide rate to the regression specifications. A third issue is the potential reverse causality between homicide rates and the determinants of homicide rates. Endogeneity is more likely to affect certain determinant variables than others. For example, a rising homicide rate can potentially cause an increase in the police presence (police per 100,000 people) for a jurisdiction.[27]

[26] As noted in Curry, Sen and Orlov (2016: 482), many econometric studies of crime have focused on the effects of policing numbers in large part because the variable is readily available.

[27] In general, while one expects an inverse relationship between crime rates and more police resources, higher crime rates in turn can be associated with the demand for more resources. See Di Matteo (2014: 27) for a discussion of this bidirectional causality.

With the addition of lagged homicide rate as a regressor, we modify Eq. (1), so that the underlying model becomes:

$$y_{\mathrm{it}} = \alpha_i + \rho y_{i,t-1} + Z_{\mathrm{it}}\delta + X_{\mathrm{it}}\beta + e_{\mathrm{it}}. \tag{2}$$

The standard approaches to dealing with unobserved heterogeneity in a panel data setting include the random effects, the within/fixed effects and the first-difference estimators. Each of these estimators transforms the dependent and independent variables in the model to remove or handle the unobserved heterogeneity. In particular, Nickell (1981) shows that the least squares dummy variable (LSDV) estimator is inconsistent for $N \to \infty$ with strictly exogenous regressors. Each of these estimators suffers from inducing endogeneity in the form of the transformed lagged dependent variable. For example, if we first-difference the model to remove α_i, then the model to be estimated becomes

$$y_{it} - y_{i,t-1} = \rho\left(y_{i,t-1} - y_{i,t-2}\right) + (Z_{it} - Z_{it-1})\delta + \left(X_{it} - X_{i,t-1}\right)\beta + (e_{it} - e_{i,t-1}) \tag{3}$$

Given the underlying model, the error term in this transformed first-difference model is correlated with $\left(y_{i,t-1} - y_{i,t-2}\right)$, one of the regressors in the transformed model.

To address these econometric issues, we employ a dynamic panel data estimator that allows for the presence of endogenous variables.[28] This estimator is the Blundell–Bond (1998) system GMM estimator. The estimator uses both the levels (Eq. (2)) and the first-difference (Eq. (3)) equations jointly in a system to account for the presence of both lagged dependent variables and unobserved heterogeneity. The system GMM estimator allows for any type of endogeneity, including the induced endogeneity in the first-difference equation and the endogeneity of any regressor in the levels equation, by using appropriate lagged values and lagged first-difference values of the explanatory variables as instruments.[29] Moral-Benito (2013) and Moral-Benito et al. (2019) suggest an alternative maximum likelihood approach for dealing with a dynamic panel data model. Bun et al. (2020) also use GMM estimation to study crime dynamics in Australia. They note that the advantages of GMM versus maximum likelihood is that "it requires much weaker assumptions about initial conditions of the data generating process and avoids full specification of the serial correlation and heteroscedastic properties of the error, or indeed any other distributional assumptions." Further, the GMM allows for the presence of endogenous regressors, which, as noted, is a concern for some of the control variables in the regression. Normally, the time-invariant effects are

[28] An alternative approach uses Bartik-type instruments to handle endogeneity. For example, see Enamorado et al. (2016) and Bouston et al. (2013).

[29] Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998) develop GMM estimation for dynamic panel models. Their estimator uses and instruments for only the first-difference equation and is known as the "difference GMM" estimator. A third option is to use only the levels equation. Depending on assumptions, the difference GMM estimator, the levels GMM estimator or the full system GMM estimator generates consistent estimates. The system GMM estimator assumes variables follow a stationary process. Fajnzylber, Lederman and Loayza (2002) estimate both the levels and system GMM models.

removed through differencing in a panel setting. However, the system GMM estimator also allows for the estimation of coefficients on time invariant covariates by using the levels equation.[30] Our set of control variables does contain three regional dummy variables with Ontario excluded. These regional controls now account for within region correlations.[31]

We also report two specification tests for the validity of the instruments.[32] The first test is the Sargan test of over-identification restrictions. The system GMM is over-identified as there are more instruments, and by extension moment conditions, than regressors. If the population moments are valid, then the equivalent sample moments should be close to zero. For the Sargan test, non-rejection provides support for the validity of the model. The second test examines the serial correlation of the first-difference error term (e_{it}–$e_{i,t-1}$). By construction, this first-difference error term suffers from first-order serial correlation. However, second-order serial correlation indicates the error term in the levels equation, e_{it}, is also serially correlated and invalidates some lagged variable values as instruments. In this case, restricting to higher-order lags is necessary in order to obtain valid instruments. The econometric analysis restricts the instrument set to higher orders, and we report the test statistic for third-order serial correlation if the test rejects no second-order serial correlation. Finally, robust standard errors are reported.[33]

4.3 Results

Table 1 provides the regression results when considering homicide rates across provinces. These results essentially represent an update to Daly, Wilson and Vasdev (2001) by including additional controls, accounting for the econometric issues mentioned in Sect. 4.2, and using a longer interval of provincial data spanning the 1982 to 2017 period.[34] The results indicate that the income inequality measures both have a positive relationship with homicide rates at the provincial level. The coefficient on the Gini coefficient ranges from 3.140 to 4.184 and is always statistically significant. The coefficient on the ratio of the top 1 to bottom 50 variable lies between −0.008

30 For example, Huynh and Petrunia (2010, 2016) use the system GMM to investigate firm growth and demonstrate the initial conditions matter to an entrant's future growth.

31 For example, a further factor of consideration in this analysis is Indigenous population within a jurisdiction. Unfortunately, annual measures of the proportion of population that is Indigenous are unavailable at both the provincial and CMA levels from 1982 to 2017. This issue is partially dealt with by the inclusion of regional dummy variables in the regression. Kripfganz and Schwarz (2019) discuss the inclusion of time-invariant regressions in linear panel data models.

32 The dynamic panel data GMM estimators are over-identified in nature with the number of instruments increasing with the length of the panel (T). The overfitting of endogenous variable in finite samples potentially leads to a bias versus efficiency trade-off. Alvarez and Arellano (2003) and Bun and Kiviet (2006) examine the overfitting bias of panel data GMM estimators. Roodman (2009a, b) discusses approaches for instrument choice in dynamic panel GMM estimators: (i) limit the maximum lag length used as instruments rather than all lags and (ii) collapse instruments into smaller sets through addition. The results do not change substantially when changing the lag length of the instrument set and, thus, are robust to the lag length of the instrument set.

33 All GMM estimation is done using Roodman (2009b) *xtabond2* package in Stata 16. We also cross-validate select results using Kripfganz (2019) *xtdpdgmm* algorithm.

34 Daly et al. (2001) use provincial data over the 1981 to 1996 period.

Table 1 System GMM provincial dynamic panel regressions homicide rate

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathrm{Ln}(1 + \mathrm{Hom\_rate}_{i,t-1})$	0.220***	0.227***	0.226***	0.382***	0.393***	0.404***
	(0.068)	(0.072)	(0.073)	(0.075)	(0.071)	(0.082)
Ratio top 1 to	0.010		0.024***	− 0.008		0.008
Bottom $50_{prov,t}$	(0.008)		(0.007)	(0.007)		(0.006)
$\mathrm{Gini}_{prov,t}$	3.140**	3.930***		4.184**	3.455**	
	(1.570)	(1.338)		(1.711)	(1.454)	
Median $\mathrm{income}_{prov,t}$	0.009	0.019	0.020	0.047	0.032	0.059
	(0.043)	(0.042)	(0.043)	(0.053)	(0.053)	(0.055)
Unemployment	− 0.023*	− 0.025*	− 0.020	− 0.034**	− 0.031**	− 0.032**
$\mathrm{Rate}_{prov,t}$	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)
Employment	− 0.001*	− 0.000	− 0.001**	− 0.001**	− 0.001***	− 0.001***
$\mathrm{Level}_{prov,t}$	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Emp $\mathrm{rate}_{prov,t}$	− 0.013*	− 0.014**	− 0.012	− 0.022***	− 0.021***	− 0.021***
	(0.007)	(0.007)	(0.008)	(0.005)	(0.005)	(0.006)
LOINC $\mathrm{portion}_{prov,t}$	− 0.014	− 0.017**	− 0.001	− 0.018**	− 0.017*	− 0.002
	(0.007)	(0.007)	(0.009)	(0.008)	(0.009)	(0.009)
$\mathrm{MinWage}_{prov,t}$	0.014	0.005	0.018	0.001	0.008	0.010
	(0.023)	(0.023)	(0.024)	(0.018)	(0.017)	(0.018)
Percentage adult	0.757	1.551	− 2.286	0.639	0.218	− 2.694
$\mathrm{Male}_{prov,t}$	(3.081)	(3.227)	(4.269)	(3.201)	(3.233)	(4.224)
NewImm Per	10.882	10.396	9.097	7.344	7.881	4.154

Table 1 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
$100{,}000_{prov,t}$	(8.755)	(8.392)	(9.160)	(8.034)	(7.535)	(8.325)
Police Per	0.006***	0.006***	0.006***			
$100{,}000_{prov,t}$	(0.001)	(0.001)	(0.001)			
Sargan *p* value	0.369	0.394	0.476	0.780	0.836	0.870
AR(1) *p* value	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) *p* value	0.019	0.017	0.020	0.014	0.012	0.007
AR(3) *p* value	0.317	0.324	0.320	0.189	0.236	0.233

Dependent variable: $\mathrm{Ln}(1 + \text{homicide rate}_{prov,t})$

Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables

and 0.024 but is positive in three out of four specifications. Statistically significance occurs in only one of the four specifications that include this variable.

The results change when using CMAs as the cross-sectional unit of observation. Table 2 presents the regression estimates for CMAs over the 2000 to 2017 period.[35] For the CMA ratio of the top 1 to bottom 50 variable, the estimated coefficients fall between − 0.016 and − 0.013 and are always statistically significant. Across specifications, the coefficient on the provincial Gini measure ranges from − 1.526 to − 0.203 and is always statistically insignificant. The results at the CMA level suggest homicide rates and income inequality become negatively correlated after controlling for a range of confounding factors. The contrasting results for the regressions using provincial data and the regressions using CMA data indicate that the level of aggregation is indeed a relevant and important factor.

We highlight the following additional results from Tables 1 and 2. First, homicide rates do appear to have some positive persistence. The coefficient on the lagged homicide rate is positive and statistically significant. This estimated coefficient is between 0.22 and 0.404 for all provincial specifications in Table 1, while the estimated coefficient is between 0.22 and 0.25 for the CMA regressions in Table 2. However, these relatively small (not close to one) estimated coefficients indicate that the persistence in homicide rates is relatively short, suggesting rapid adjustment of homicide rates to changing conditions over time.

Second, with respect to the estimated coefficients of the other variables, the results are fairly consistent across all the specifications within each table. The estimated coefficients on the other control variables do vary quantitatively but do not vary greatly qualitatively within each table. However, there are qualitative and quantitative differences when comparing the regression results for the provincial (Table 1) versus CMA (Table 2) data. Specifically, we highlight the following findings. To start, the estimated coefficient of median income is essentially zero. The unemployment rate appears to have a negative relationship with homicide rates. The coefficients on unemployment rate are statistically significant only in the regressions using provincial data. The estimated coefficient on the employment level is positive and statistically significant for five out of six specifications using the CMA data, while this coefficient is negative and statistically significant in four out of six specifications using the provincial data. The estimated coefficient on the police per 100,000 people variable is always positive and statistically significant for the provincial data. The coefficient on minimum wage is positive and statistically insignificant. The new immigrants per 100,000 people variable always has a positive coefficient and is statistically significant in half the specifications for both tables. The percentage of persons in low-income variable generally has a statistically insignificant negative coefficient for the provincial results in Table 1, while the coefficient is positive and significant for the CMA results in Table 2. Finally, the coefficient on the fraction of males aged 15–24 years in the total population variable is positive in most cases. This coefficient is statistically significant for only the CMA regressions in Table 2.

[35] We also redo the regression analysis for the provincial data over the 2000 to 2017 periods. The provincial results for the shorter time period are similar to those in Table 1. The coefficient estimates on the Gini coefficient variable are larger in magnitude. These results are available from the authors upon request.

Table 2 System GMM CMA dynamic panel regressions homicide rate

	(1)	(2)	(3)	(4)	(5)	(6)
Ln(1 + $Hom\_rate_{i,t-1}$)	0.218***	0.236***	0.219***	0.221***	0.249***	0.222***
	(0.066)	(0.066)	(0.066)	(0.067)	(0.068)	(0.068)
Ratio top 1 to	− 0.013***		− 0.013***	− 0.015***		− 0.016***
Bottom $50_{cma,t}$	(0.005)		(0.005)	(0.005)		(0.005)
$Gini_{prov,t}$	− 0.203	− 0.803		− 0.514	− 1.526	
	(1.657)	(1.841)		(1.509)	(1.649)	
Median $income_{cma,t}$	0.070	0.071	0.068	0.107	0.145	0.103
	(0.101)	(0.102)	(0.104)	(0.127)	(0.134)	(0.129)
Unemployment	− 0.022	− 0.024	− 0.022	− 0.024	− 0.029	− 0.024
$Rate_{cma,t}$	(0.018)	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)
Employment	0.001**	− 0.000	0.001**	0.001***	0.000	0.001***
$Level_{cma,t}$	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	(0.012)	(0.012)	(0.013)	(0.015)	(0.014)	(0.015)
LOINC $portion_{cma,t}$	0.024**	0.023*	0.024**	0.029***	0.033***	0.029***
	(0.012)	(0.013)	(0.012)	(0.011)	(0.012)	(0.011)
$MinWage_{prov,t}$	0.052	0.054	0.052	0.051	0.054	0.050
	(0.036)	(0.034)	(0.036)	(0.037)	(0.036)	(0.037)
Frac adult $male_{prov,t}$	37.825***	35.042***	37.812***	38.247***	35.167***	38.220***
	(9.894)	(9.810)	(9.869)	(9.779)	(9.438)	(9.748)
NewImm Per	21.205**	18.846*	21.163**	22.327**	20.129**	22.238**
$100{,}000_{prov,t}$	(9.711)	(9.713)	(9.648)	(9.086)	(9.105)	(9.038)

Table 2 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Police Per	0.001	0.002	0.001			
$100{,}000_{cma,t}$	(0.001)	(0.001)	(0.001)			
Sargan p value	0.585	0.643	0.598	0.635	0.717	0.647
AR(1) p value	0.001	0.000	0.001	0.001	0.001	0.001
AR(2) p value	0.264	0.243	0.264	0.245	0.219	0.244

Dependent variable: Ln(1 + homicide $rate_{cma,t}$)
Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables

Finally, the Sargan tests fail to reject the null hypothesis in all cases in Tables 1 and 2. This result provides support that the instruments are valid. However, the results for the autocorrelation tests differ across the two tables. We expect to uncover first-order autocorrelation due to construction. Second-order autocorrelation indicates misspecification as t-2 variables become endogenous and are not valid instruments. The tests find first-order autocorrelation as the p-values are low in both cases. The results of second-order autocorrelation differ between the provincial and CMA regressions. Originally, the results using the provincial data indicated the presence of second-order serial correlation. Therefore, we re-estimated the equation for the provincial data by restricting the instrument set to higher orders of the lags, thereby accounting for the possible misspecification problem created by the second-order correlation. Now, these results depend on no third-order autocorrelation for validity of the instruments. The results in Table 1 do not suggest the presence of third-order autocorrelation in the provincial regressions. These test results therefore indicate support for the validity of the restricted higher-order lags of variables as instruments. For CMA regressions in Table 2, second-order autocorrelation does not appear to exist in all specifications. Thus, the evidence suggests the CMA regressions appear to not suffer from any substantial misspecification issues.

4.4 Robustness check: alternative estimation methods

The system GMM estimator is best designed for panel situations with a large cross section (N) and a small and fixed time dimension (T). Flannery and Hankins (2013) discuss the trade-offs between the GMM estimators and fixed effect estimators and note the performance of each estimator depends on the conditions.[36] The ten provinces form a balanced panel of either 32 or 36 observations each depending on the specification. The 26 CMAs comprise an unbalanced panel of between 14 and 18 observations with an average of 17.58. Given the time and cross-sectional dimensions for our data, we consider alternative estimators. These methods include: (1) the LSDV fixed effects model (FE); (2) the bias corrected least squares dummy variable model (LSDVC) of Bruno (2005a), and Bun and Kiviet (2003); (3) the dynamic panel data instrumental variable estimation with defactored regressors and multifactored error structure (IVD) of Norkute et al. (2021); and (4) the split-panel jackknife (SPJ) method of Dhaene and Jochmans (2015).[37] The IVD uses defactored covariates as instruments for endogenous variables. As done for the system GMM estimates previously, the lagged dependent variable, the two inequality measures, and the police variable are treated as endogenous for the IVD method. For the IVD estimates, we use two lags of exogenous variables as instruments and allow up to a maximum of four factors.

The appendix provides results using alternative estimation methods with dynamic panel data. For each method, the specification includes a full set of control variables

[36] An alternative approach suggested by Flannery and Hankins (2013) is to work with long lags as done in Petrunia (2007, 2008).

[37] We implement: the LSDVC method using the *xtlsdvc* of Bruno (2005b); IVD estimation using the *xtivdfreg* command by Kripfganz, and Sarafidis (2021); and the SPJ estimation using the *xtspj* command of Sun and Dhaene (2019).

except the regional dummy variables, since these estimators do not allow for time invariant covariates. As a comparison, these tables also provide the system GMM estimates. Table 8 presents results for the provinces. The coefficient on the lagged homicide rate does fluctuate from close to 0.1 for the FE, LSDVC and SPJ estimators, 0.30 for the system GMM estimator and 0.45 for the IVD estimator. Estimated coefficients on the inequality measures are relatively stable across methods. The coefficient on the ratio of the top 1 to bottom 50 variable is close to zero and insignificant in all cases except for the system GMM estimator. For the Gini variable, estimates fall between 3 and 4.28 and are statistically significant across methods. Table 9 presents estimates for the CMAs. These estimates show less stability. The coefficients on both measures of inequality are always statistically insignificant. These results contrast results in Table 2 where the ratio of the top 1 to bottom 50 variable always has a negative and statistically significant coefficient. Finally, the IVD estimates exhibit a high degree of imprecision with the highest standard errors.

5 Other issues

5.1 Regional analysis

Previous analysis included dummy variables to account for differences across regions. To delve further into possible regional differences in the relationship between homicide rates and income inequality, we conduct a separate analysis for western and non-western provinces. We use these groupings because there are economic and demographic differences not necessarily completely accounted for in the previous analysis and historically, homicide rates in Canada have risen moving from east to west, with the western provinces exhibiting much higher homicide rates. Possible reasons for differences in homicide rates between western and non-western provinces could be unique demographic factors including the proportion of recent immigrants or proportion of Indigenous population. Indeed, the basis for attempting a separate analysis of western Canadian CMAs from those in the remainder of the country is rooted in the distinct historical and cultural development of the Canadian west separately from the east.

Canadian Confederation originally consisted of eastern-based provinces and western Canadian development was essentially a process of rapid settlement and colonization from the east as part of the national policies in the period after 1870. The national policies included a land policy by which the Canadian federal government provided title to 160 acres of land to any settler over 18 years of age after three years of residence subject to minimum use of the land and a 10-dollar fee (Pomfret, 1993: 183; Martin (1973). For the Canadian west, Norrie (1975) documents the dramatic increase in the annual registration of homestead entries, especially during the wheat boom period after 1896 from 1,897 in 1896 to 7,426 in 1900 and then to 44,749 by 1911.

However, this process was not particularly democratic and as noted by Velasco (2016: 64): "the displacement of indigenous peoples from their land to make place for the arrival of European settlers and the uneven allocation of land in the hands of corporations like the CPR and the HBC also created the conditions for unequal distribution of wealth in Canada and western Canada particularly." At the same time, it has also been noted that the Canadian west in particular experienced boom and bust economic cycles as a result of settlement and resource development and the initially more dispersed land holding patterns of the farm economy of western Canada was also associated with greater wealth equality (Di Matteo, 2012).

The displacement of indigenous peoples during the process of western settlement has some implications for both crime and inequality in the west given that the population share in western Canada of indigenous origin is much higher than eastern Canada. In 2016, there were 1,673,785 Aboriginal people in Canada, accounting for 4.9% of the total population of whom half lived in western Canada (BC, AB, SK and MB), whereas those same provinces only accounted for one-third of the non-Indigenous population (Statistics Canada, 2017).

In 2015, Indigenous people accounted for 25% of homicide victims, at a rate which was about seven times that of non-Indigenous people (8.77 victims per 100,000 population vs. 1.31). The homicide rate of Indigenous male victims was about seven times that of non-Indigenous males (12.85 vs. 1.87). The homicide rate of Indigenous female victims was six times that of non-Indigenous females (4.8 vs. 0.77) (Canada, Department of Justice, 2017). As well, indigenous peoples have historically experienced lower incomes and higher rates of poverty. Among Indigenous people living in an urban area, about half lived in rented dwellings, compared with 29% of the non-Indigenous population and of those renting in 2016, one in five lived in subsidized housing (Anderson, 2019).

Table 3 presents results for the province data across these two provincial groupings. The coefficient estimates on the income inequality variables (Gini coefficients and ratio top 1 to bottom 50) are positive but are statistically insignificant across groupings and specifications. Although these estimated coefficients are above 2 for the non-western provinces, the results show that a consistent province-level relationship between economic inequality and homicides breaks down once these regional effects are taken onto account. Meanwhile, Table 4 provides estimates for CMAs based on western and non-western province groupings. For western province CMAs, the estimated coefficient for the ratio of the top 1 to bottom 50 inequality variables is negative and statistically significant across all specifications and both tables, while the estimated coefficient on the Gini coefficient is negative and statistically significant in one specification. For the non-western province CMAs, both inequality variables have positive but statistically insignificant estimated coefficients.

These results provide further evidence that results can differ based on combinations of data and levels of aggregation. In explaining these differences, one possible explanation is that there is in fact no consistent positive relationship between inequality and homicide rates. The main reason for unsupportive findings advanced by Wilkinson and Pickett (2006) in the case of the relationship between population health and inequality does not apply here. Whether at the provincial or CMA level, the units used have sufficiently large populations such that we are confident sufficient social and income diversity is present (Table 4).

Table 3 Provincial homicide rate regressions

	Western provinces			Non-western provinces		
Ln(1 + $Hom\_rate_{i,t-1}$)	0.153*	0.165*	0.152*	0.067	0.068	0.062
	(0.087)	(0.087)	(0.087)	(0.073)	(0.073)	(0.074)
Ratio top 1 to	0.012		0.015	0.002		0.011
Bottom $50_{prov,t}$	(0.011)		(0.010)	(0.015)		(0.013)
$Gini_{prov,t}$	0.652	1.299		2.105	2.222	
	(1.342)	(1.234)		(2.003)	(1.618)	
Median $Income_{prov,t}$	− 0.039	− 0.031	− 0.040	− 0.029	− 0.027	− 0.020
	(0.049)	(0.049)	(0.049)	(0.061)	(0.059)	(0.061)
Unemployment	0.021	0.015	0.023	− 0.051***	− 0.051***	− 0.051***
$Rate_{prov,t}$	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)
Employment	− 0.002***	− 0.001***	− 0.002***	− 0.000	− 0.000	− 0.000
$Level_{prov,t}$	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Emp $rate_{prov,t}$	0.010	0.009	0.012	− 0.032***	− 0.033***	− 0.034***
	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)
LOINC $portion_{prov,t}$	0.000	− 0.003	0.003	− 0.009	− 0.009	− 0.002
	(0.012)	(0.012)	(0.010)	(0.014)	(0.014)	(0.013)
$MinWage_{prov,t}$	0.019	0.006	0.021	0.025	0.024	0.023
	(0.027)	(0.025)	(0.027)	(0.035)	(0.034)	(0.036)
Frac 15–24 $male_{prov,t}$	6.014	7.532*	5.793	− 8.876*	− 8.814*	− 11.289**
	(4.625)	(4.479)	(4.616)	(5.239)	(5.202)	(4.744)
NewImm Per	27.769***	27.151***	27.760***	5.051	4.971	4.431
$100,000_{prov,t}$	(6.073)	(6.100)	(6.090)	(7.546)	(7.503)	(7.578)
Police Per	0.002	0.001	0.002	0.005***	0.005***	0.006***
$100,000_{prov,t}$	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N	128	128	128	192	192	192
Sargan *p* value	0.084	0.103	0.099	0.218	0.234	0.260
AR(1) *p* value	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) *p* value	0.414	0.436	0.411	0.095	0.093	0.104

Dependent variable: Ln(1 + homicide $rate_{prov,t}$)
Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables

Table 4 CMA Homicide rate regressions by regions

	Western provinces			Non-western provinces		
Ln(1 + $Hom\_rate_{i,t-1}$)	0.130	0.320**	0.130	0.150**	0.151**	0.150**
	(0.103)	(0.138)	(0.106)	(0.072)	(0.072)	(0.073)
Ratio top 1 to	− 0.027***		− 0.028***	0.006		0.007
Bottom $50_{cma,t}$	(0.006)		(0.006)	(0.007)		(0.007)
$Gini_{prov,t}$	− 1.046	− 2.602**		1.793	2.219	
	(1.300)	(1.065)		(1.928)	(1.943)	
Median $income_{cma,t}$	0.115	− 0.028	0.080	0.234**	0.235**	0.250**
	(0.162)	(0.203)	(0.173)	(0.115)	(0.116)	(0.123)
Unemployment	0.011	0.006	0.010	− 0.045*	− 0.045*	− 0.044*
$Rate_{cma,t}$	(0.029)	(0.027)	(0.030)	(0.024)	(0.024)	(0.024)
Employment	0.003***	− 0.001	0.003***	− 0.000	− 0.000	− 0.000
$Level_{cma,t}$	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Emp $rate_{cma,t}$	0.043***	0.037**	0.044***	− 0.014	− 0.013	− 0.015
	(0.014)	(0.018)	(0.014)	(0.014)	(0.015)	(0.014)
LOINC $portion_{cma,t}$	0.042***	0.040	0.040***	0.019	0.023	0.019
	(0.014)	(0.027)	(0.015)	(0.018)	(0.019)	(0.018)
$MinWage_{prov,t}$	0.060*	0.069***	0.058*	0.038	0.034	0.041
	(0.033)	(0.019)	(0.034)	(0.058)	(0.059)	(0.058)
Frac 15–24 $male_{prov,t}$	36.749***	22.675**	35.547***	6.685	7.150	9.990
	(6.740)	(11.107)	(6.664)	(8.812)	(8.672)	(9.223)
NewImm Per	15.102**	7.846	15.015**	− 0.282	− 0.884	3.616
$100{,}000_{prov,t}$	(5.871)	(7.783)	(5.832)	(11.925)	(11.958)	(10.908)
$100{,}000_{cma,t}$	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N	125	125	125	332	332	332
Sargan *p* value	0.221	0.270	0.244	0.542	0.555	0.559
AR(1) *p* value	0.021	0.020	0.020	0.002	0.002	0.002
AR(2) *p* value	0.484	0.693	0.509	0.281	0.272	0.291

Dependent variable: Ln(1 + homicide $rate_{cma,t}$)

Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables

5.2 Other crime rate measures

Homicides tend to be a small portion of both the overall crime rate and violent crime rate. Further, the previous literature, both theoretical and empirical, suggests a link between economic inequality and overall criminal activity. This link occurs if criminal activity is at least partially about capturing resources as often surmised in the literature. Capturing resources is certainly likely to be a stronger motivation for violent crimes like robbery and non-violent crimes like theft. To further investigate the impact of inequality on criminal activity, we replace the homicide rate with total crime rate, violent crime rate and property crime rate. The total crime rate is defined as all criminal code violations (excluding traffic) per 100,000 persons. Violent crime rate is total violent criminal code violations per 100,000 persons. Finally, property crime rate is total property crime violations per 100,000 persons. Violent crime violations include homicides, assaults, sexual assaults, robberies, abduction and criminal harassment. Property crime violations include breaking and entering, possession of stolen property theft, fraud and arson.[38]

For these alternative measures of crime rates, Table 5 presents the provincial results and Table 6 presents the CMA results. Overall, the estimates indicate that there is more persistence in these crime rate measures when compared to previous estimates looking at the homicide rates. Across both sets of results, the coefficient estimates of the lagged crime rate range from 0.847 to 0.905 in Table 5 and from 0.893 to 0.930 in Table 6. These estimates suggest a high level of persistence with slightly more persistence in CMAs. This persistence at the local level makes sense given that there can be local factors and conditions affecting crime in a particular city or city-regional area that might not affect crime rates in other urban areas.

When looking at the inequality measures, the coefficients are typically negative for both the provincial and CMA results. However, these coefficients are always statistically insignificant for the provincial results in Table 5. Thus, the positive relationship between crime and economic inequality at the provincial level occurs when using the homicide rate as the measure of criminal activity but disappears when using more broad measures of criminal activity. In the CMA results in Table 6, the estimated coefficient on the ratio top 1 percent to bottom 50 percent variable is always statistically significant, while the coefficients on the Gini variable are statistically insignificant. For CMAs, economic inequality and criminal activity appear to have a negative relationship regardless of how broadly or narrowly defined the criminal activity measure.

6 Concluding discussion

A substantial national- and international-level literature has documented a relationship between greater economic inequality and higher homicide rates. Using Canadian provincial- and CMA-level pooled time-series cross-sectional data, we are only occasionally able to find a positive correlation between inequality and homicide rates. While we find a province-level relationship between greater economic inequality and higher

[38] See Statistics Canada Table 35-10-0177-01 for a complete list.

Table 5 System GMM provincial dynamic panel regressions crime rate

	Total crime rate	Violent crime rate	Property crime rate
Ln(1 + $\text{Crime\_rate}_{i,t-1}$)	0.905***	0.847***	0.889***
	(0.032)	(0.032)	(0.035)
Ratio top 1 to	− 0.002	− 0.002	− 0.003
Bottom $50_{prov,t}$	(0.003)	(0.003)	(0.004)
$\text{Gini}_{prov,t}$	− 0.113	0.109	− 0.334
	(0.405)	(0.406)	(0.498)
Median $\text{income}_{prov,t}$	0.012	0.034**	0.010
	(0.016)	(0.016)	(0.020)
Unemployment	0.005	− 0.011***	0.008*
$\text{Rate}_{prov,t}$	(0.004)	(0.004)	(0.005)
Employment	− 0.000	− 0.000***	− 0.000
$\text{Level}_{prov,t}$	(0.000)	(0.000)	(0.000)
Emp $\text{rate}_{prov,t}$	0.004	− 0.005*	0.005
	(0.003)	(0.003)	(0.004)
LOINC $\text{portion}_{prov,t}$	0.005	0.004	0.003
	(0.003)	(0.003)	(0.004)
$\text{MinWage}_{prov,t}$	0.002	0.008	0.001
	(0.007)	(0.007)	(0.009)
Percentage adult	− 2.999*	− 1.361	− 2.868
$\text{Male}_{prov,t}$	(1.626)	(1.681)	(2.028)
NewImm per	− 3.670***	− 3.110**	− 3.337*
$100{,}000_{prov,t}$	(1.342)	(1.413)	(1.731)
Police Per	− 0.000	− 0.000	− 0.000
$100{,}000_{prov,t}$	(0.000)	(0.000)	(0.000)
N	190	190	190
Sargan *p* value	0.001	0.000	0.010
AR(1) *p* value	0.000	0.000	0.000
AR(2) *p* value	0.064	0.212	0.288

Dependent variable: Ln(1 + Crime $\text{rate}_{prov,t}$)
Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables. Years 1998–2017

homicide rates, we also find that a consistent province-level relationship between economic inequality and homicides breaks down once regional effects are taken onto account. Moreover, when broader definitions of crime are used, there is also not a positive relationship between inequality and crime rates at the provincial level. This result differs from the conclusions reached by Daly et al. (2001) and is partly a function of

Table 6 System GMM CMA dynamic panel regressions crime rate

	Total crime rate	Violent crime rate	Property crime rate
$\text{Ln}(1 + \text{Crime\_rate}_{i,t-1})$	0.930***	0.911***	0.893***
	(0.023)	(0.021)	(0.024)
Ratio top 1 to	− 0.003***	− 0.003***	− 0.004***
Bottom $50_{\text{cma,t}}$	(0.001)	(0.001)	(0.001)
$\text{Gini}_{\text{prov,t}}$	− 0.187	0.429	− 0.405
	(0.311)	(0.346)	(0.423)
Median $\text{income}_{\text{cma,t}}$	− 0.008	− 0.010	− 0.006
	(0.022)	(0.025)	(0.022)
Unemployment	0.004**	0.001	0.005*
$\text{Rate}_{\text{cma,t}}$	(0.002)	(0.003)	(0.003)
Employment	0.000**	0.000***	0.000
$\text{Level}_{\text{cma,t}}$	(0.000)	(0.000)	(0.000)
Emp $\text{rate}_{\text{cma,t}}$	0.001	0.000	0.002
	(0.001)	(0.001)	(0.001)
LOINC $\text{portion}_{\text{cma,t}}$	− 0.000	− 0.002	0.001
	(0.003)	(0.003)	(0.003)
$\text{MinWage}_{\text{prov,t}}$	− 0.010	− 0.012	− 0.006
	(0.006)	(0.009)	(0.007)
Frac adult $\text{male}_{\text{prov,t}}$	− 1.809	− 1.917	− 1.132
	(1.493)	(1.684)	(1.776)
NewImm Per	0.242	− 0.746	0.918
$100{,}000_{\text{prov,t}}$	(2.025)	(2.502)	(2.321)
$100{,}000_{\text{cma,t}}$	(0.000)	(0.000)	(0.000)
N	439	439	439
Sargan *p* value	0.010	0.231	0.013
AR(1) *p* value	0.000	0.000	0.000
AR(2) *p* value	0.461	0.501	0.635

Dependent variable: $\text{Ln}(1 + \text{crime rate}_{\text{cma,t}})$
Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. Median income and employment level are in thousands. Standard errors are corrected for heteroscedasticity and autocorrelation. Regressions also include a time trend and regional dummy variables. Years 1998–2017

a longer period of time-series data available to us as well as of aspects of the regression specification including a much broader range of variables used as confounding factors. The substantially higher homicide rates in western Canada need to be taken into account via regional fixed effects. The implication for policy making may be that the relationship between homicide rates and inequality may be better considered at the local rather than provincial level.

The results for Canadian CMAs generate a different story. Using data from Canadian CMAs over the period 2000 to 2017 allows for newer and a substantially larger number of observations than previous Canadian studies and also controls for many confounding factors. We find a statistically significant negative relationship between homicide rates and income inequality for CMAs in Canada when all the CMAs are combined or when they are broken up into regional units. Furthermore, this result is robust when we also consider the impact of potential econometric issues such as unobserved specific factors, reverse causality or persistence effects. The contrasting results for the provincial and CMA analyses suggest the importance of aggregation issues and the possibility of the Yule–Simpson paradox being at work in the data here.

We thus conclude that much of the literature that has been finding a relationship between greater economic inequality and homicide rates needs to be re-examined within a longer time framework making note of subregional data and regional groupings. Analysis not accounting for these factors is likely to contain estimated correlations that are spurious and calls into question causal explanations linking greater inequality to greater amounts of crime and homicides. Strong empirical support for theory linking economic inequality as a causative factor in homicides or crime requires consistent results across data sets and different levels of data aggregation. The absence of consistent results suggests there is a need for new theoretical work to explain such differences. At minimum, there may need to be more effort expended on rethinking the causal process as to how inequality and poverty affect crime.

Funding This research was not supported by any funding.

Declarations

Conflict of interest Robert Petrunia and Livio Di Matteo have no conflict of interest to declare.

Ethical approval The research did not involve human or animal subjects.

Appendix

See Tables 7, 8 and 9.

Table 7 Canadian CMAs and coverage years

City	Number	Coverage years
Calgary	18	2000–2017
Edmonton	18	2000–2017
Greater Sudbury	18	2000–2017
Halifax	18	2000–2017
Hamilton	18	2000–2017
Kitchener–Waterloo–Cambridge	18	2000–2017
London	18	2000–2017
Montreal	18	2000–2017
Oshawa	18	2000–2017
Ottawa–Gatineau(Ont)	18	2000–2017
Ottawa–Gatineau(Que)	16	2002–2017
Quebec City	14	2003–2016
Regina	18	2000–2017
Saguenay	16	2002–2017
Saint John	18	2000–2017
Saskatoon	18	2000–2017
Sherebrooke	18	2000–2017
St. Catharines–Niagara	18	2000–2017
St. John's	18	2000–2017
Thunder Bay	18	2000–2017
Toronto	18	2000–2017
Trois-Rivieres	16	2002–2017
Vancouver	18	2000–2017
Victoria	17	2000–2016
Windsor	18	2000–2017
Winnipeg	18	2000–2017
Total observations	457	
Average observations per CMA	17.58	

Table 8 Provincial dynamic panel regressions homicide rate

	Sys-GMM	FE	LSDVC	SPJ	IVD-1
$\text{Ln}(1 + \text{Hom\_rate}_{i,t-1})$	0.30***	0.08	0.12**	0.08	0.45
	(0.06)	(0.06)	(0.06)	(0.06)	(0.67)
Ratio top 1 to	0.02***	− 0.00	− 0.00	− 0.00	− 0.02
Bottom $50_{prov,t}$	(0.00)	(0.01)	(0.01)	(0.01)	(0.07)
$\text{Gini}_{prov,t}$	4.07**	4.28***	4.20**	4.28***	0.07
	(1.70)	(1.40)	(1.63)	(1.39)	(9.05)
Median $\text{income}_{prov,t}$	− 0.01	0.05	0.05	0.05	0.10
	(0.04)	(0.05)	(0.06)	(0.05)	(0.26)
Unemployment	− 0.02*	0.02	0.02	0.02	− 0.01
$\text{Rate}_{prov,t}$	(0.01)	(0.02)	(0.02)	(0.02)	(0.16)
Employment	− 0.00***	− 0.00**	− 0.00*	− 0.00**	0.00
$\text{Level}_{prov,t}$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Emp $\text{rate}_{prov,t}$	− 0.00	− 0.01	− 0.01	− 0.01	− 0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.11)
LOINC $\text{portion}_{prov,t}$	− 0.01	− 0.01	− 0.01	− 0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
$\text{MinWage}_{prov,t}$	0.04	− 0.00	− 0.00	− 0.00	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.08)
Percentage adult	1.76	4.48	4.21	4.48	0.00
$\text{Male}_{prov,t}$	(3.27)	(3.40)	(4.19)	(3.38)	(.)
NewImm Per	16.20*	8.10	7.38	8.10	0.00
$100{,}000_{prov,t}$	(9.36)	(6.18)	(7.86)	(6.13)	(.)
Police Per	0.01***	0.00	0.00	0.00	− 0.01
$100{,}000_{prov,t}$	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
N	320	320	320	320	320

Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. IVD includes two lags of exogenous variables as instruments and allows up to a maximum of four factors. The number of factors in X equals 1, and the number of factors in U equals 3. The standard error for the LSDVC is bootstrapped

pt

Table 9 CMA dynamic panel regressions homicide rate

	Sys-GMM	FE	LSDVC	SPJ	IVD
$\text{Ln}(1 + \text{Hom\_rate}_{i,t-1})$	0.26***	0.03	0.09*	0.02	0.32**
	(0.06)	(0.05)	(0.05)	(0.05)	(0.16)
Ratio top 1 to	− 0.00	0.01	0.01	0.00	0.05
Bottom $50_{\text{cma,t}}$	(0.00)	(0.01)	(0.01)	(0.01)	(0.05)
$\text{Gini}_{\text{prov,t}}$	− 2.17	1.90	2.00	1.47	− 3.64
	(1.72)	(2.14)	(2.09)	(2.11)	(4.72)
Median $\text{income}_{\text{cma,t}}$	0.06	− 0.02	− 0.03	0.21	− 0.55
	(0.08)	(0.13)	(0.14)	(0.21)	(0.36)
Unemployment	− 0.02	− 0.01	− 0.01	− 0.02	− 0.00
$\text{Rate}_{\text{cma,t}}$	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)
Employment	− 0.00	− 0.00	− 0.00	− 0.00	− 0.01
$\text{Level}_{\text{cma,t}}$	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Emp $\text{rate}_{\text{cma,t}}$	0.02	− 0.01	− 0.01	− 0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
LOINC $\text{portion}_{\text{cma,t}}$	0.03**	− 0.03	− 0.03	− 0.00	− 0.03
	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)
$\text{MinWage}_{\text{prov,t}}$	0.01	0.04	0.04	0.06*	− 0.05
	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)
Frac adult $\text{male}_{\text{prov,t}}$	28.82***	18.90*	17.17	19.14*	5.93
	(7.12)	(10.69)	(11.05)	(10.41)	(26.68)
NewImm Per	11.16	10.59	9.25	5.62	24.89
$100{,}000_{\text{prov,t}}$	(7.04)	(10.00)	(12.19)	(10.39)	(18.00)
Police Per	0.00**	− 0.00	− 0.00	− 0.00	− 0.01
$100{,}000_{\text{cma,t}}$	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
N	457	457	457	457	413

Dependent variable: $\text{Ln}(1 + \text{homicide rate}_{\text{cma,t}})$
Standard errors are in parenthesis with *, ** and *** indicating 10 percent, 5 percent and 1 percent levels of significance, respectively. IVD includes two lags of exogenous variables as instruments and allows up to a maximum of four factors. The number of factors in X equals 1, and the number of factors in U equals 1. The standard error for the LSDVC is bootstrapped

References

Alvarez J, Arellano M (2003) The time series and cross-sectional asymptotics of dynamic panel data estimators. Econometrica 71:1121–1159

Anderson T (2019) Results from the 2016 census: housing, income and residential dissimilarity among indigenous people in Canadian cities. Statistics Canada, 75–006-X. https://www150.statcan.gc.ca/n1/pub/75-006-x/2019001/article/00018-eng.htm

Arellano M, Bond S (1991) Some test of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 58:277–298

Arellano M, Bover O (1995) Another look at the instrumental variable estimation of error-components models. J Econ 68:29–51

Baier J (2014) "Does inequality cause crime? evidence from a Latin American Panel," mimeo, LUP Student Papers, url: http://lup.lub.lu.se/student-papers/record/4497601
Bailey WC (1984) Poverty, inequality, and city homicide rates. Criminology 22(4):531–550
Becker G (1968) Crime and punishment: an economic approach. J Polit Econ 76:169–217
Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. J Econ 87(1):115–143
Boustan L, Ferreira F, Winkler H, Zolt EM (2013) The effect of rising income inequality on taxation and public expenditures: evidence from US municipalities and school districts, 1970–2000. Rev Econ Stat 95(4):1291–1302
Bruno GSF (2005a) Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. Econ Lett 87(3):361–366
Bruno GSF (2005b) Estimation and inference in dynamic unbalanced panel-data models with a small number of individuals. Stata J 5(4):473–500
Bun MJG, Kiviet JF (2003) On the diminishing returns of higher order terms in asymptotic expansions of bias. Econ Lett 79:145–152
Bun MJG, Kiviet JF (2006) The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. J Econ 132:409–444
Bun MJG, Kelaher R, Sarafidis V, Weatherburn D (2020) Crime, deterrence and punishment revisited. Empir Econ 59:2303–2333
Canada, Department of Justice (2017) Indigenous overrepresentation in the criminal justice system, January. https://www.justice.gc.ca/eng/rp-pr/jr/jf-pf/2017/docs/jan02.pdf
Coccia M (2017) A theory of general causes of violent crime: homicides, income inequality and deficiencies of the heat hypothesis and of the model of CLASH. Aggress Violent Behav 37:190–200
Cohen JE (1986) An uncertainty principle in demography and the unisex issue. Am Stat 40(1):32–39
Connor P, Sarafidis V, Zyphur MJ, Keltner D, Chen S (2019) Income inequality and white-on-black racial bias in the United States: evidence from project implicit and Google Trends. Psychol Sci 30(2):205–222
Curry PA, Sen A, Orlov G (2016) Crime, apprehension and clearance rates: panel data evidence from Canadian provinces. Can J Econ 49(2):481–514
Daly M (2016) Killing the competition: economic inequality and homicide. Routledge
Daly M, Wilson M, Vasdev S (2001) Income inequality and homicide rates in Canada and the United States. Can J Criminol 43:219–236
Davies JB, Di Matteo L (2021) Long run Canadian wealth inequality in international context. Rev Income Wealth 67(1):134–164. https://doi.org/10.1111/roiw.12453
Davis GA (2004) Possible aggregation biases in road safety research and a mechanism approach to accident modeling. Accid Anal Prev 36:1119–1127
Deaton A (2003) Health, inequality, and economic development. J Econ Literature 41(1):113–158
Dhaene G, Jochmans K (2015) Split-panel jackknife estimation of fixed-effect models. Rev Econ Stud 82:991–1030
Di Matteo L (2003) The income elasticity of health care spending a comparison of parametric and nonparametric approaches. Eur J Health Econ 4(1):20–29
Di Matteo L (2012) Land and inequality in Canada, 1870–1930. Scand Econ Hist Rev 60(3):309–334
Enamorado T, López-Calva LF, Rodríguez-Castelán C, Winkler H (2016) Income inequality and violent crime: evidence from Mexico's drug war. J Dev Econ 120:128–143
Fajnzylber P, Lederman D, Loayza N (2002) Inequality and violent crime. J Law Econ 45:1–40
Flannery MJ, Hankins KW (2013) Estimating dynamic panel models in corporate finance. J Corp Finan 19:1–19
Frank RH (2012) The Darwin economy: liberty, competition and the common good. Princeton University Press, New Jersey
Frank RH (2013) Falling behind: how rising inequality harms the middle class. University of California Press, California
Frank RH, Cook PJ (1995) The winner-take-all society. The Free Press, New York
Freund C, Oliver S (2016) The origins of the superrich: the billionaire characteristics database, Working Paper Series, 16–1, February, Peterson Institute for International Economics
Getzen TE (2000) Health care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures. J Health Econ 19(2):259–270
Getzen TE (2006) Aggregation and the measurement of health care costs. Health Serv Res 41(5):1939–1954

Harris G, Vermaak C (2015) Economic inequality as a source of interpersonal violence: evidence from Sub-Saharan Africa and South Africa. South African J Econ Manag Sci 18(1):45–87
Heisz A (2016) Trends in Income Inequality in Canada and Elsewhere in D A Green, W Craig Riddell and France St-Hilaire (eds), Income Inequality: The Canadian Story, Montreal: Institute for Research on Public Policy (IRPP) 77–102
Hicks DL, Hicks JH (2014) Jealous of the Joneses: conspicuous consumption, inequality, and crime. Oxf Econ Pap 66(4):1090–1120
Huynh KP, Petrunia RJ (2010) Age effects, leverage and firm growth. J Econ Dyn Control 34(5):1003–1013
Huynh KP, Petrunia RJ (2016) Post-entry struggle for life and pre-exit shadow of death from a financial perspective. Int J Econ Bus 23(1):1–18
Jackson A (2015) The return of the gilded age: consequences, causes and solutions Broadbent Institute. The Harry Kitchen Lecture in Public Policy, April 8, 2015
James A, Aadland D (2011) The curse of natural resources: an empirical investigation of U.S. counties. Resour Energy Econ 33(2):440–453
Kelly M (2000) Inequality and crime. Rev Econ Stat 82(4):530–539
Kennedy LW, Silverman RA, Forde DR (1991) Homicide in Urban Canada: testing the impact of economic inequality and social disorganization. Can J Sociol 16(4):397–410
Kripfganz S, Sarafidis V (2021) Instrumental variable estimation of large-T panel data models with common factors Stata J forthcoming
Kripfganz S, Schwarz C (2019) Estimation of linear dynamic panel data models with time-invariant regressors. J Appl Economet 34(4):526–546
Kripfganz S (2019) Generalized method of moments estimation of linear dynamic panel data models Proceedings of the 2019 London Stata Conference
Levitt SD (1999a) The exaggerated role of changing age structure in explaining aggregate crime changes. Criminology 37(August):537–599
Levitt SD (1999b) The changing relationship between income and crime victimization. Econ Policy Rev 5(3):87–98
Levitt SD (2004) Understanding why crime fell in the 1990s: four factors that explain the decline and six that do not. J Econ Perspect 18(1):163–190
Lobmayer P, Wilkinson R (2000) Income, inequality and mortality in 14 developed countries. Sociol Health Illn 22(4):401–414
Lynch J, Davey Smith G, Hillemeier M, Shaw M, Raghunthan T, Kaplan G (2001) Income inequality, the psychosocial environment, and health: comparisons of wealthy nations. Lancet 358(August):194–200
Lynch J, Smith GD, Harper SA, Hillemeier M, Ross N, Kaplan GA, Wolfson M (2004) Is income inequality a determinant of population health? part 1a systematic review. Milbank Q 82(1):5–99
Macdonald D (2014) Outrageous fortune: documenting Canada's wealth gap Canadian Centre for Policy Alternatives. April
Martin C (1973) Dominion lands policy, The Carleton Library No. 69, Toronto: McClelland and Stewart
Di Matteo L (2014) Police and crime rates in Canada: a comparison of resources and outcomes, Fraser Institute
Moral-Benito E (2013) Likelihood-based estimation of dynamic panels with predetermined regressors. J Bus Econ Stat 31(4):451–472
Moral-Benito E, Allison P, Williams R (2019) Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. Appl Econ 51(20):2221–2232
Nickell SJ (1981) Biases in dynamic models with fixed effects. Econometrica 49:1417–1426
Norkute M, Sarafidis V, Yamagata T, Cui G (2021) Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure. J Econo 220:416–446
Norrie K (1975) The rate of settlement of the Canadian prairies, 1870–1911. J Econ Hist 35(2):410–427
Osberg L (2018) The age of increasing inequality. James Lorimer & Company Ltd, Toronto
Ouimet M (2012) The effect of economic development, income inequality, and excess infant mortality on the homicide rate for 165 countries in 2010
Oxfam (2015) Wealth: having it all and wanting more Oxfam issue briefing. January 2015
Pearl J (2011) Simpson's paradox: an Anatomy UCLA Department of Statistics Papers, 2011–10–25
Pearl J (2018) The book of why. The new science of cause and effect. New York. Basic Books
Petrunia R (2007) Persistence of initial debt in the long-term employment dynamics of new firms. Can J Econ 40:861–880

Petrunia R (2008) Does Gibrat's Law Hold? Evidence from Canadian retail and manufacturing firms. Small Bus Econ, Econ 30:201–214
Pomfret R (1993) The economic development of Canada, 2nd edn. Nelson, Scarborough
Roberts A, Willits (2015) Income inequality and homicide in the United States: consistency across different income inequality measures and disaggregated homicide types. Homicide Stud 19(1):28–57
Roodman D (2009a) A note on the theme of too many instruments. Oxford Bull Econ Stat 71:135–158
Roodman D (2009b) How to do Xtabond2: an introduction to difference and system GMM in Stata. Stand Genomic Sci 9:86–136
Rowlingson K (2011) Does income inequality cause health and social problems? Joseph Rowntree Foundation. www.jrf.org.uk
Saez E, Veall M (2005) The evolution of high incomes in Northern America: lessons from Canadian evidence. Am Econ Rev 95(3):831–849
Scheider M, Spence DL, Mansourian J (2012) The relationship between economic conditions, policing and crime trends. community oriented police services, US Department of Justice
Simpson EH (1951) The interpretation of interaction in contingency tables. J Roy Stat Soc 13(2):238–241
Statistics Canada (1973) 1971 Census of Canada: Families: Families by Size and Type. Statistics Canada, Ottawa
Statistics Canada (1977) The Labour Force. Statistics Canada, Ottawa
Statistics Canada (1982) Income Distribution by Size. Statistics Canada, Ottawa
Statistics Canada (2017) Aboriginal peoples in Canada: Key results from the 2016 Census. https://www150.statcan.gc.ca/n1/daily-quotidien/171025/dq171025a-eng.htm
Subramanian SV, Kawachi I (2004) Income inequality and health: what we have learned so far. Epidemiologic Rev 26(1):78–91
Sun Y, Dhaene G (2019) xtspj: a command for split-panel jackknife estimation. Stata J 19(2):335–374
Szwarcwald CL, Bastos FI, Viacava F, de Andrade CLT (1999) Income inequality and homicide rates in Rio de Janeiro, Brazil. Am J Public Health 89(6):845–850
Veblen T (1899/1994) The theory of the leisure class: an economic study in the evolution of institutions New York; Dover Publications
Velasco G (2016) Natural resources, state formation and the institutions of settler capitalism: the case of Western Canada, 1850–1914. PhD Thesis, LSE. http://etheses.lse.ac.uk/3437/
Wilkinson R, Pickett K (2006) Income inequality and population health: a review and explanation of the evidence. Soc Sci Med 62(7):1768–1784
Wilkinson R, Pickett K (2009) The spirit level: why more equal societies almost always do better. Penguin, London
Wolff EN (2016) Household wealth trends in the United States, 1962 to 2013: what happened over the great recession? Russell Sage Found J Soc Sci 2(6):24–43
Wolff EN (2010) Recent trends in household wealth in the United States: rising Debt and the Middle Class Squeeze-an Update to 2007 Levy economics institute. Working Paper No 589. March
Wolfson MG, Lynch KJ, Ross N, Backlund E (1999) Relation between income inequality and mortality: empirical demonstration. BMJ 319:953–955
Yalnizyan A (2010) The Rise of Canada's Richest 1%." Canadian center for policy alternatives December
Yule GU (1903) Notes on the theory of association of attributes in statistic. Biometrika. 2(2):121–134