Inequality and the Distribution of Earnings in Canada at the Turn of the Twentieth Century

by

Oana Ciobanu

A thesis submitted to the

Department of Economics

in partial fulfillment of the

requirements for the degree of

Master of Arts

Queen's University

Kingston, Ontario, Canada

September 2010

Contents

Acknowledgements

Thank you to my supervisor, Ian Keay, for all his help and guidance.

1 Introduction

The study of historical income distribution and inequality during periods of early industrialization has captured the interest of a wide variety of researchers in recent decades, as it adds to our understanding of topics ranging from economic growth, to the role of public redistribution policies, to household decision making. Although the theory and methodology of measuring inequality has developed considerably, data availability and quality severely limit existing studies in both their scope and validity. A new sample of the 1901 Canadian personal census provides groundbreaking insights into within-group distributions and methodological issues associated with measuring inequality with historical micro-data. Instead of seeking to provide a one-size-fits-all approach for measuring and studying inequality, the purpose of this paper is to demonstrate how nation-wide micro-level data can be used to improve on existing studies of historical income distribution. I use within-group distributions to show how mean or median income differentials based on occupations, skill-level, or region are not only incomplete but are also deceptive methods of capturing inequality. Furthermore, I show that local micro-data such as probate samples fair poorly in describing the income inequality of the working-class.

Most papers studying income distribution before 1950 have been preoccupied with tracing movements in inequality over long stretches of time and matching patterns in economic growth with aggregate movements in inequality. Although it is typically believed that rapid industrialization should be accompanied by rising inequality, this

"quest for the Kuznets curve" in the empirical literature has come up with little concrete evidence for or against the existence of such a relationship (Lindert, 2000, 170). The growing consensus today, as expressed by Peter Lindert (2000), seems to be the need to divert attention away from the search for Kuznets curve patterns and focus on episodic movements trying to explain the underlying causes of inequality. In other words, before we can understand the relationship between inequality and economic development we need to study the factors that affect and that are affected by income distribution in its own right.

Until at least the mid-twentieth century income inequality was seen as the embodiment of patterns in factor-prices and factor-shares. Indeed the view originally proposed by Adam Smith of fixed shares of the population for different economic classes, each rewarded by a different factor price persisted in the theoretical explanations of income inequality for decades (Atkinson and Bourguignon, 2000; Bourguignon and Morrisson, 1990). Simon Kuznets in his seminal 1955 paper showed that inequality could change dramatically with the movement of workers from rural to urban regions without any shifts in factor prices whatsoever. In so doing he "ushered in the current era of decompositional inequality accounting" (Lindert, 2000) that began attributing inequality to hosts of other factors including human capital, the accumulation of productive factors, imperfect competition, and life-cycle patterns [1]

Part of this new decompositional wave in the income distribution literature has

[1]See Chapters 1-3 of the Handbook for Income Distribution (2000) for a summary of the theoretical literature on inequality.

been the increased interest in "looking back" to early stages of industrialization and studying historical inequality dynamics. It is a widespread belief that between the 17th and 19th century industrialization and rapid economic growth caused fundamental changes to class structure, employment, and income shares, thereby creating the foundation for the patterns of income distribution found today in much of the industrialized world (Lindert, 2000). Unfortunately, although widespread interest exists, good quality historical income micro-data is extremely scarce. Many attempts have been made by Soltow (1984, 1992), Williamson (1980), Jones (1977, 1980) and Lindert and Williamson (1982, 1983), to name only a few, to recreate historical income distribution estimates for the United States and England. In England the main sources of historical micro-data have been the social tables constructed by then "political arithmeticians" from Gregory King to A.L. Bowley, probate records, a few dozen wage rates, land-rent series, and early partial tax returns. The most comprehensive study to unite these measures was made by Lindert and Williamson (1982, 1983) with their construction of revised social tables by weighting different measures. Yet even they were only able to produce class-average incomes per household and had to use clues from movements in factor-price ratios and compositional changes to deduce real inequality patterns. In the United States, state-level labour force surveys, social surveys, and selected probate records provide the main micro-data for inequality studies and are often supplemented by additional "odds and ends" - indirect evidence in the form of suggestive wage gaps and mortality trends (Lindert, 2000, 186).

Within this literature, discussion of the Canadian experience has been unnecessarily omitted. This can be partly attributed to data limitation, as Canada lacks labour force surveys and social tables of the kind available for the United States and England. However, this has not stopped researchers from investigating income distribution using evidence from a variety of alternative sources including isolated probate samples, and differentials in occupational wages, skill gaps, and regional earnings. The 1901 census, as I will show in this paper, provides superior insights into the income distribution and patterns of inequality found in Canada at the turn of the century. It unearths patterns hidden by aggregate estimates while at the same time allowing for national inequality implications. In many ways, it is also provides noteworthy advantages over foreign data sources with its wealth of information on employment and earnings of the working class. As Baskerville and Sager (1995, 23) argue, "It is clear... that Canadian census takers went far beyond their British and American counterparts in their interest in the 1901 work force." The Canadian census in 1901 asked fourteen questions regarding the characteristics and activities of the work force. In England, only five questions were asked, and in the United States where only two questions were asked it took almost another 50 years for earnings to be included in the personal census. For these reasons, the Canadian census provides an incredibly unique opportunity to study historical income distribution, that for reasons of availability or lack of interest has yet to be utilized in this context.

The paper will be structured in the following way. Section 2 outlines the two

streams of literature that have encompassed the study of historical Canadian income distribution and includes background information on inequality during this period. A methodological discussion can be found in Section 3 where the basic tools for evaluating income distribution and the properties associated with the earnings and occupation data available from the personal census are described. Section 4 presents findings from within-occupation distributions and discusses the position of the chronically under-employed in aggregate distribution measures. National inequality is examined in Section 5 by measuring inequality within urban centres across Canada. Section 6 concludes.

2 Wealth and Inequality in Canada

The period 1896 to 1913 for Canada represents a time of rapid economic growth and fundamental changes to patterns of development, the structure of industry, population growth, and living standards. It also marks the beginning of many economic patterns including regional income differentials and growth patterns of industry, that persisted well into the century. Therefore, describing labour force characteristics for this period including wage rates and income differentials across regions, occupations, and skill-levels, has much value and has, for this reason, been the target of many studies. Those relevant for this paper can be grouped into two categories. The first uses aggregate wage rates and average earnings to track income differentials across regions or skill levels but can say little about their within-group distributions. The second tries to capture inequality directly through data samples on wealth, income, or property but are limited to the region of their sample and have little explanatory power for the majority of wage earners. The Census micro-data presented in this paper, although not without its own shortcomings, can fill in many of the holes left by the current literature to shed better light on the patterns of income distribution in the Canadian labour force at the turn of the century.

In the first group fall Bertram and Percy (1979), McInnis (1968), MacKinnon (1996), Emery and Levitt (2002) and, Coe and Emery (2004), among others. These works primarily focus on reported and estimated wage rates to study growth and regional disparities. One of their biggest advantages is their ability to analyze trends

at the national level. For example, Bertram and Percy (1979) document an annual growth rate of 0.9% in real hourly wage rates for the period 1901-13 using indexes weighted by industry wage rates and cost of living measures. Several papers have shown the existence of persistent regional income disparities going back to as far as 1891 both in nominal and real terms (e.g., McInnis, 1968; Green, 1967; Emery and Levitt, 2002; Williamson, 1965; Coulombe and Lee, 1995). They all find a similar pattern of: (1) High earnings in the West (especially in British Columbia) as well as in central Ontario and Montreal, (2) Lower income levels in the Atlantic provinces, and (3) The lowest earnings in Quebec outside of Montreal. By taking into account real prices and cost of living patterns these differentials become less pronounced but still persist (Emery and Levitt, 2002).

All of these studies depend heavily on point estimates of earnings, often for only a selected number of occupations. The published census results, for example, include only average earnings for fully-employed workers over the age of 16. Another popular source of wage data has been the Wages and Hours of Labour reports published with the Labour Gazette by the Department of Labour between 1907 and 1921.They provide minimum standard wage rates (ie. for fully-qualified employees working standard hours) from a survey of a few dozen occupations to be paid in public sector contracts. Though useful, these rates have been criticized extensively. First, as Bertram and Percy themselves admit, to use these standard rates as a proxy for money income has the potential of skewing estimates considerably during period of excess demand for labour or recessions. In this case, hours of work and employment would be re-

quired to derive earnings or any estimate of income. MacKinnon (1996) argues that because the wages reported by Fair Wage Officers were usually union rates, estimates for unskilled workers such as general or builders' labourers were much higher than the wages received for these occupations in reality. Using hourly wage rates from firm records of the CPR for a selected number of occupations, MacKinnon finds much higher growth of real and nominal wages than previous estimates including those of Bertram and Percy (1979). By picking occupations of different skill level yet common to other industries, she captures working class inequality using skill differentials and finds increased inequality before WWI as growth in wages for machinists outpaced that of general labourers and helpers. MacKinnon's work exemplifies the considerable impact misrepresentation of a single wage rate can have on aggregate trends. Isolating comparison based on point estimates that are easily influenced by selection bias or misreporting opens up analysis to a host of possible errors. This is where distributional analysis becomes very useful as will be shown in Section 4.

In order to assess the real impacts of early industrialization on the Canadian population, including resulting movements in income and wealth, other researchers have been led to use micro-data such as census manuscripts, probate records, and assessment rolls. These studies are able to describe income and wealth distributions and calculate measures of inequality directly instead of relying on point estimates and proxies for inequality such as skill or regional differentials. The evidence they provide shows that micro-data can unearth phenomenon buried by aggregate measures. For

example, according to macroeconomic estimates of Urquhart (1965, 1986) growth of real GNP per capita remained almost constant during the 1890s. However, evidence from micro-data studies of Darroch (1983) for Toronto and Di Matteo and George (1992, 1995) for Wentworth Country show that average wealth levels actually declined during this period. Di Matteo went further to isolate wealth of subgroups and demonstrated that between 1892 and 1902 those foreign born, farmers and Catholics improved their average wealth holdings while natives, non-farmers and non-Catholics fared poorly.

Furthermore, since probate records include data on real estate, financial assets, and personal property these studies are not limited in their income measure by only salary and earnings. Their results show a highly unequal society in Canada at the end of the 19th century that experiences a very gradual narrowing of income over time. Darroch (1983) finds that the top 10 percent of the population in Toronto in 1899 amassed 49.9% of total assessed wealth. In the Thunder Bay region which experienced high levels of growth from the wheat boom, the top decile between 1885 and 1906 owned on average 66.7% of real wealth according to Di Matteo (2006). Osberg and Siddiq (1988) study Nova Scotia in 1871 and estimate even higher inequality when the wealth of non-probated decedents are imputed. Their estimates reach 80.9% of total wealth amassing to the top 10% of family units a Gini coefficient of .89 [2]. Darroch (1983) also estimates a Gini value of .598 for his Toronto sample in 1899. By linking probate samples to census records, Di Matteo (2001) is able to analyze in-

[2]See Section 3.1 for a discussion of the Gini coefficient. Higher values are associated with higher levels of inequality

equality using a richer set of information. He regresses average wealth for 38 Ontario counties onto inequality measures taking into account a variety of controls including age profiles, occupation skill level, sex, ethnic origin, proportion of wealth originating in real estate, literacy, and whether the decedent was a farmer. The results indicate a negative statistically significant relationship between inequality and literacy, being aged 30-44, proportion of real estate in wealth, and being a farmer. It is worthwhile to note however that only the proportion of real estate in wealth had an estimated impact higher than a 1% increase to the Gini coefficient.

This impact of land ownership on inequality is one that is consistently observed across samples for this period. Di Matteo (2006, 90) observes that "the rich were truly different from everyone else" not only in the level of wealth they enjoyed but also in its composition. In his sample he finds that the top 10% owned 45 times the value of real estate and 69 times the value of moneys secured by mortgages. In my sample of wage earners from the 1901 federal census I find that land ownership was highly correlated with total earnings. Only 12.5% of the bottom quintile of wage earners reported owning any property. This value increases exponentially across percentiles such that at the top 10% of earners, 71% owned some sort of property. However, growing evidence suggests that age patterns were the decisive link between land ownership and inequality. Ownership of land was highest among those over 40 years of age and generally was positively correlated with age (Di Matteo, 2001). Evaluating access to land, Darroch and Soltow (1992, 169) find that once age groups are ac-

counted for, the "deep class divisions" appearing in land ownership in 1871 largely disappear. Sylvester (2001) also finds that life cycle was the most important factor influencing farm size in 1901. The implication then is that changes to the composition of the population age structure over time will create patterns in income inequality by changing rates of property ownership. Since evidence shows that the Canadian population was aging considerably over the period 1871 to 1921 (Mitchell, 1993) it would be no surprise to see rising inequality as predicted by Di Matteo (1997, 2001, 2006) and Osberrg and Siddiq (1993).

However, these patterns of property holdings are based purely on cross-sectional evidence. Industrialization during this period brought with it fundamental changes to income sources and composition, especially for urban areas. Higher rates of urbanization and increased reliance on wage-labour have been documented characteristics of the working population in Canada beginning in the later half of the 19th century (Emery et al. 2008, Chapters 9-13). Darroch (1983) suspects this, along with the trend in falling commodity prices, explains the fall in inequality found in Toronto between 1861 and 1899. Therefore property as a source of income was very likely falling in the decades surrounding 1901, and must have been doing so at an even greater rate given the aging population.

Where this becomes a problem for studies based on probate samples is that probate records have been widely documented to miss a large proportion of the working class or non-land-owning part of the population. In Di Matteo's (2006, 55) sample from Wentworth Country only 36% of estates were estimated to have been probated

and in the Thunder Bay region, that number falls to 31% (Di Matteo, 1997, 98). The lower limit of wealth under which probates were not required by law was estimated by Darroch (1983, 32) to be so large that it "must simply have excluded many of the working class from assessment altogether." Furthermore, probate records capture the value of estate at death and therefore represent the accumulation of wealth over a lifetime. Therefore, not only is analysis based on probate records isolated to the wealthy and the proportion of the population of higher socioeconomic status but as patterns of property ownership shift over time they become increasingly poorly suited for studying the relationship in Canada between early industrialization and inequality. As I show in Section 5, the inequality among top earners is a poor representation of the distribution of earnings for the rest of the population.

Indeed the potential negative effects of industrialization as they appeared in increased poverty and the conditions of the expanding class of wage earners was a topic of interest even at the turn of the century. As documented by Darroch (1983), sensitivity to increased poverty and division of wealth rose particularly in the latter half of the nineteenth century. Toronto's journals began noting the emergence of extreme poverty in areas with high concentration of industrial production (Masters, 1947) and questions arose nationally of the standards of living and the shares in the expanding economy of the working class. Keyssar (1986, 307) underlines that the widespread nature of unemployment at the turn of the century that a "chronic and pervasive feature of working class life." Baskerville and Sager (1998) expand these findings to show

that one in three workers experienced joblessness and that one in seven working-class families in urban centres had less real income than that required for basic survival. Probate samples have no way of capturing the relative well being of the poor and are therefore poorly suiting for addressing the inequality of the working class.

Studying inequality through census samples of wage earners indeed will miss the impact of non-labour income on inequality but offers considerable advantages over studies based on isolated probate samples. First, the 1901 census is better able to capture the working class that are made up predominantly of wage-earners most impacted by industrialization. Second, previous papers have had limited impact and are ill suited for describing national patterns because of their isolated usage of a single community or region. Di Matteo and George (1996, 47) note that "it is perilous to generalize from 'local' results to 'general' trends or inter-community comparison because comparative studies require a consistent sampling of similar data sources from many communities or regions." In order for the study of historical inequality in Canada to draw attention and interest, it must work to provide "fundamental reinterpretations of Canadian history" (Di Matteo, 1996, 3). This requires a national perspective accompanied by attention to local distributional characteristics.

3 Data and Methodology

The main source of data for this paper is a 5 percent sample of the 1901 Canadian federal personal census created by the Canadian Families Project at the University of Victoria. It includes a rich set of information on personal characteristics (age, sex, nationality, education and household structure) as well as labour market characteristics including annual earnings, chief occupation, and employment status. The sample is nationally representative of national and provincial demographic patterns in published census data (Canadian Families Project, 2003).

3.1 Measuring Inequality

Though many studies have used proxies such as skill, region, or gender differentials to capture patterns in inequality, the importance of micro-data cannot be understated. Point estimates such as mean or median are extremely sensitive to outliers and, more importantly, are a poor representation of average well-being when inequality is high. This stems from both diminishing returns to utility and the nature of aggregate welfare functions that put value on equality and the income of the poor. The field of inequality measurement and the comparison of income distribution has experienced considerable progress in the last decades [3] but criticisms are everywhere being raised as to the proper usage of these (theoretically) extremely informative methods (Atkinson, 1970; Sen, 1976; Szekely and Hilgert, 1999; Deininger and Squire, 1996). "The

[3] See Chapter 2 of Lambert (1989) and Sen and Foster (1997) for a summary.

problem may not be any more that of *how* to measure and to compare income distributions but that of *what* to measure and compare" (Atkinson and Bourguigon, 2000, 4). Assessing data quality including income composition, missing values, and the under-reporting of higher incomes are just some of the necessary steps before measures of inequality are calculated and compared.

One of the most basic tools for comparing income distributions are income shares. They are calculated as the proportion of total income accrued to the top x% of the population. Calculating income shares gives much more information than mean or median estimates since they capture concentration of wealth. However, it is difficult to make general comparisons of inequality across multiple samples and the inequality within quantiles will always be missing. Income shares have been traditionally used for studying historical inequality because micro-level income data is usually only available for top ranked earners, often on account of selective taxation. With a comprehensive sample of individuals from the entire income distribution, I am not limited to these measures.

The traditional graphical representation of income inequality is the Lorenz curve. It ranks income shares by population shares to produce a curve where each point represents the share of income accrued to the xth population percentile. Perfect equality would thus be represented by a straight diagonal (a flat Lorenz curve) and increasing inequality would result in increased concavity. Not only does the Lorenz curve provide a graphical representation of relative income shares but it also allows for robust

comparisons to be made between income distributions through use of Lorenz dominance. If the income distribution of Population-A lorenz-dominates the distribution of Population-B this is akin to having a strictly higher Lorenz curve and means that individuals at every income share are better off with less inequality.

The problem is that Lorenz dominance rarely occurs and therefore cannot be used to rank income distributions whose Lorenz curves cross. Inequality measures have been developed for this purpose. The single most popular measure is the Gini coefficient (index or value), which is easily computed and envisioned as the area below a Lorenz curve. Therefore, a higher Gini index indicates higher levels of inequality. Historical coefficients have been estimated for 1911 for the United Kingdom at .48 (Lindert and Williamson, 1983) based on pre-tax nominal personal income, and for the United States at .83 (Soltow, 1984). However, using just the Gini index to compare income distributions can be severely misleading. Two countries may have the same Gini index but very different distributions. Critiques of the Gini coefficient center on the fact that it is most sensitive to variations in inequality in the middle of the distribution (Champernowne, 1974; Cowell, 1977, 28; Osber, 1984, 16-24).

Generalized entropy levels help alleviate this issue by allowing for differing attitudes with respect to inequality and offer the added ability to decompose overall inequality in an additive way into inequality within subgroups and inequality between groups. Using insights from the entropy concept in information theory Theil (1967) pioneered this approach to income distribution analysis with the Theil Index

and mean logarithmic deviation. Cowell (1977), Cowell and Kuga (1981) and Toyoda (1980) extended these measures to the generalized entropy (GE(a)) family of measures where the parameter a captures sensitivity to changes in particular parts of the distribution. Large and positive values indicate sensitivity to the changes in the distribution of the upper tail of the income distribution while negative values make the GE level sensitive to changes in the lower tail. GE(0) is the mean logarithmic deviation, GE(1) is the Theil index, and GE(2) is half the square of the coefficient of variation.

Making comparisons between income distributions using scalar inequality measures may often lead to contradictory results. Where Lorenz dominance cannot be applied either, it becomes very difficult to categorize certain distributions as 'more equal' and others as 'less equal.' Income distribution analysts are therefore encouraged to make use of multiple approaches and definitions. This, along with making sure to define carefully "inequality of what among whom" (Atkinson and Bourguignon, 2000, 4) is key to decompositional analysis of inequality and the study of patterns in income distribution. As I will show, different distributional measures can give different insights into income differentials and inequality. Using a variety of measures, including the mean, median, Gini coefficient, and generalized entropy levels, allows for a richer, more comprehensive study of historical inequality.

3.2 Earning and Employment from the 1901 Census

The primary drawback of the 1901 census for the purpose of analyzing income patterns is its exclusion of non-wage or non-salary earners. Earnings in this census was taken to mean income received in the form of salary or wages such that it captures "the amount or sum of money which one person employed by another receives for his service, whether the work done be professional, literary or handicraft"[4]. As Baskerville and Sager (1995) note, questions and enumerator's instructions reflected the interest on the part of the Canadian federal government in the emerging urban, industrial class which at the time were primarily wage earners. This group was carefully differentiated from employers, the self-employed and those of independent means who were not encouraged to provide information on earnings or employment. This was done by including additional questions identifying respondents as 'employees', 'employers', retirees, living 'on their own means' and living 'on their own account.'[5] Unfortunately these possible indicators of non-labour earnings were not exclusive (one could be an employee as well as an employer) and were far from comprehensive (being completed in the affirmative or negative for only 10% of the sample with earnings). Therefore they cannot be relied on to completely eliminate non-labour earnings. They do however indicate which occupations and industries are prone to this problem and should not be included in a study of wage earners.

[4]All enumerators' instructions were taken from Fourth Census of Canada, 1901; Instructions to Officers, Ottawa, 1901

[5]Living off one's own means was defined as living off of "incomes, superannuations, annuities, pension, etc." indicating the existence of non-labour earnings. Those living on one's own account referred to those who were "employed in gainful work" or "doing their own work" which was usually the case for farmers.

A big part of what allows me to make this attempt at identifying wage earners is the rich detail available for reported chief occupation. Occupation entries included information on type of work done (labourer, clerk, etc.), the sector or industry where the worker was employed and their "rank in principal calling or occupation." The census's preoccupation with industrial wage earners is seen especially through the emphasis placed on proper occupational disaggregation and description of industry-related trades. Enumerators were given six pages of instructions on how to produce a "fully descriptive [occupational] designation" compared to the third of a page provided in 1881. They were also specifically instructed that "Expressions such as 'manufacturer,' 'merchant' and 'labourer' are inadequate; the particular branch of industry, or trade, or profession or other calling in which the person enumerated is engaged should be given." The Canadian Families Project provides an occupational coding based on the Canadian Classification and Dictionary of Occupations (CCDO) (Ministry of Supply and Services, 1989) which groups occupations into two levels of aggregation based on the type of work done. The occupational groupings in this paper will follow this coding system for the most part. The only exception will be the creation of a category of general labourers compiled with disregard for industry.

With data on occupation, the additional labour market position questions (employer, living on own account, etc.), and relationship to the household head, deducing the likely 'economic position' of each respondent becomes surprisingly comprehensive.

As Baskerville and Sager (1995) also observe, under-reporting in the 1901 census was less common than would be expected. Of all non-institutionalized individuals (not hospital patients or inmates) of sound mind over the age of 14, 52.67% report a chief occupation. Taking into account special characteristics including living on own means, natives, students, and those with some reported income leaves 39.23% with no position. 98% of these are relatives of the household head and 96% are female. Given that enumerators were given specific instructions to disregard women occupied in domestic work, this is not at all surprising. The Appendix lists the main occupation categories with reporting rates and the proportion of missing earnings in each. With this information on economic position I am able to isolate which groups are less sensitive to missing income data.

Without much surprise, farmers and professionals have the highest proportions of missing earnings data and the highest frequencies of entries of living on own means or own account. Because of the large number of farmers in rural areas, this region will be especially prone to non-earnings omissions. However, urban centres pose their own set of problems since non-labour earners including professionals, the self-employed, and those in domestic service occupations were located predominantly in urban centres. The result is poor coverage of income data both in urban and rural regions making analysis between these groups beyond the reach of this study. Comparing coverage rates by occupation across urban centres however shows that regional comparisons are possible. That is why in the first part of this paper I study occupation earnings at

a national level while in the second section I limit my study to urban areas in order to make regional comparisons regardless of occupation.

4 Occupation Earnings

From the drawbacks of this data, I isolate my analysis of occupational earnings distributions to just a handful of occupations that meet the following criteria. First, the majority of income received should be in the form of wage or salary income. For this, I compare proportions of zero-earnings responses making sure there is a response rate of at least 80%, keeping in mind the problematic occupations discussed in the Data and Methodology section. Second, there should be a sufficient number of observations to create robust distributions such that there is little chance that a significant section of the distribution is missing. This census sample is limiting in this sense because only a few occupations satisfy these criteria. These are: machinists, blacksmiths, clerks, and carpenters. In an attempt to capture unskilled workers, I create a category of general labourers by isolating all those identifying themselves as 'general labourer' or 'labourer' or 'common labourer' irrespective of their industry or sector. One might have expected the highest variance in these earnings but on the contrary they have the lowest standard deviation of all occupations listed indicating a viable sample of homogeneous workers.

4.1 Full-time Workers

The sample is isolated to male workers over the age of 18 that report working full-time (at least 10 months of the year) in their principle occupation. This is meant to capture predominant wages for each occupation for fully-qualified workers. This

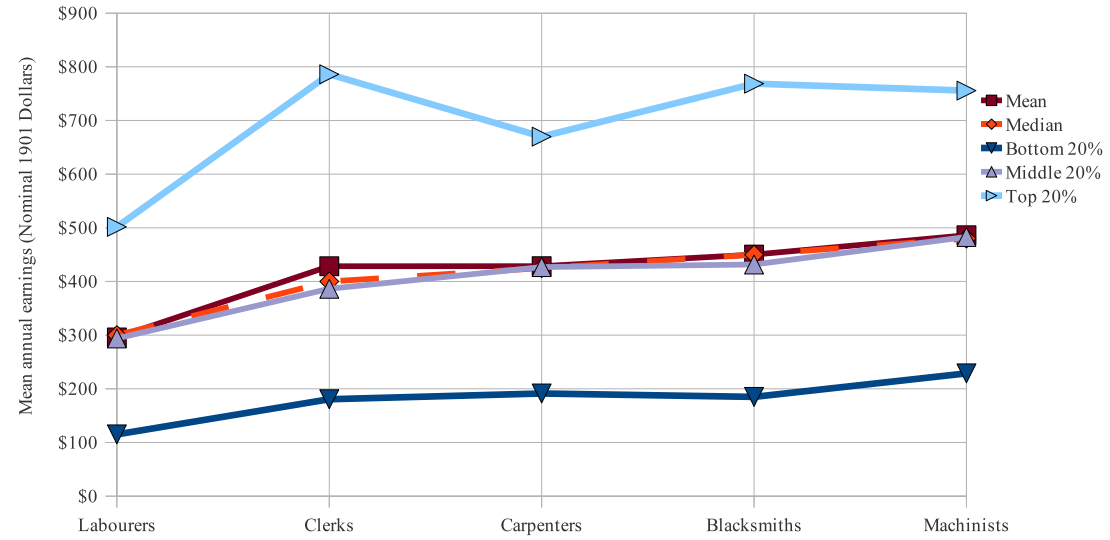
was the goal for wages reported by the Department of Labour and the Census Bulletins, and therefore will allow me to better compare my estimates to those already in common use and diminishes the possible incidence of non-labour income. Capital and land income may still exist but the chances are diminished. In fact, only .04% of total earnings was reported as coming from extra earnings on average among full-time workers in these occupations.

Table 1: Yearly nominal earnings for full-time male workers over 18 years of age, 1901

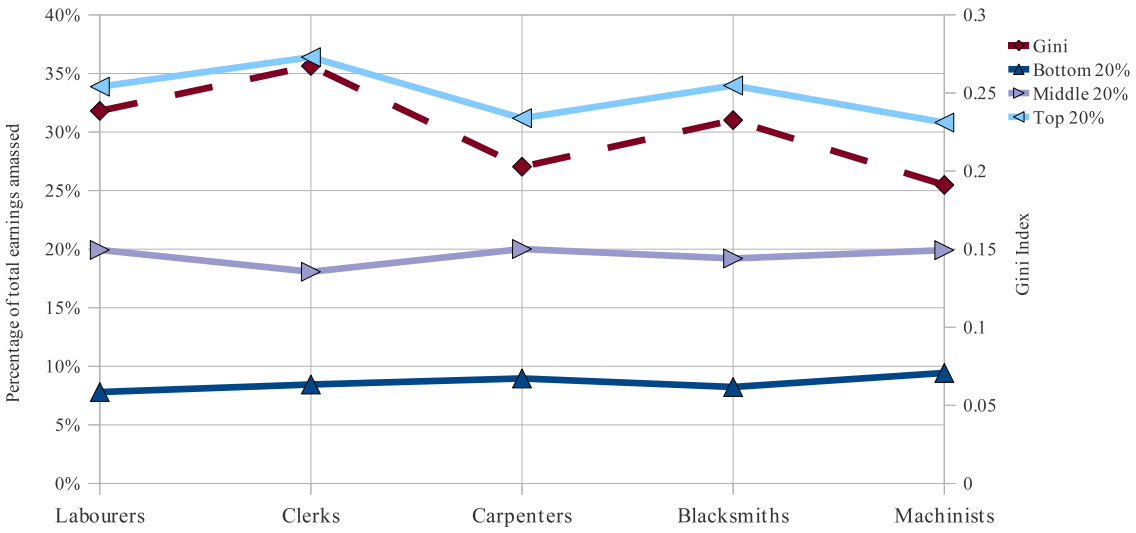
		Labourers	**Clerks**	**Carpenters**	**Blacksmiths**	**Machinists**
	Observations	957	484	667	744	449
	Mean	$295.65	$428.42	$428.45	$450.04	$486.10
	Std. Dev.	148.1	233.44	179.22	251.76	193.84
	Median	$300.00	$400.00	$425.00	$450.00	$480.00
	Gini	0.24	0.27	0.2	0.23	0.19
Mean Earnings	Bottom 20%	$114.98	$180.61	$191.19	$184.97	$228.74
	Middle 20%	$293.79	$386.32	$426.79	$431.56	$482.96
	Top 20%	$501.81	$786.02	$669.89	$768.58	$755.61
Share of Total Income	Bottom 20%	7.80%	8.45%	8.97%	8.23%	9.43%
	Middle 20%	19.94%	18.07%	20.01%	19.20%	19.91%
	Top 20%	33.88%	36.39%	31.18%	33.97%	30.81%

Calculated summary statistics, found in Table 1, reveal some patterns that would be anticipated but also many that would not have been. Average common labourer earnings are the lowest at $295 per year while machinists and blacksmiths earn the

Figure 1: Ordering Mean Earnings of Selected Occupations, 1901

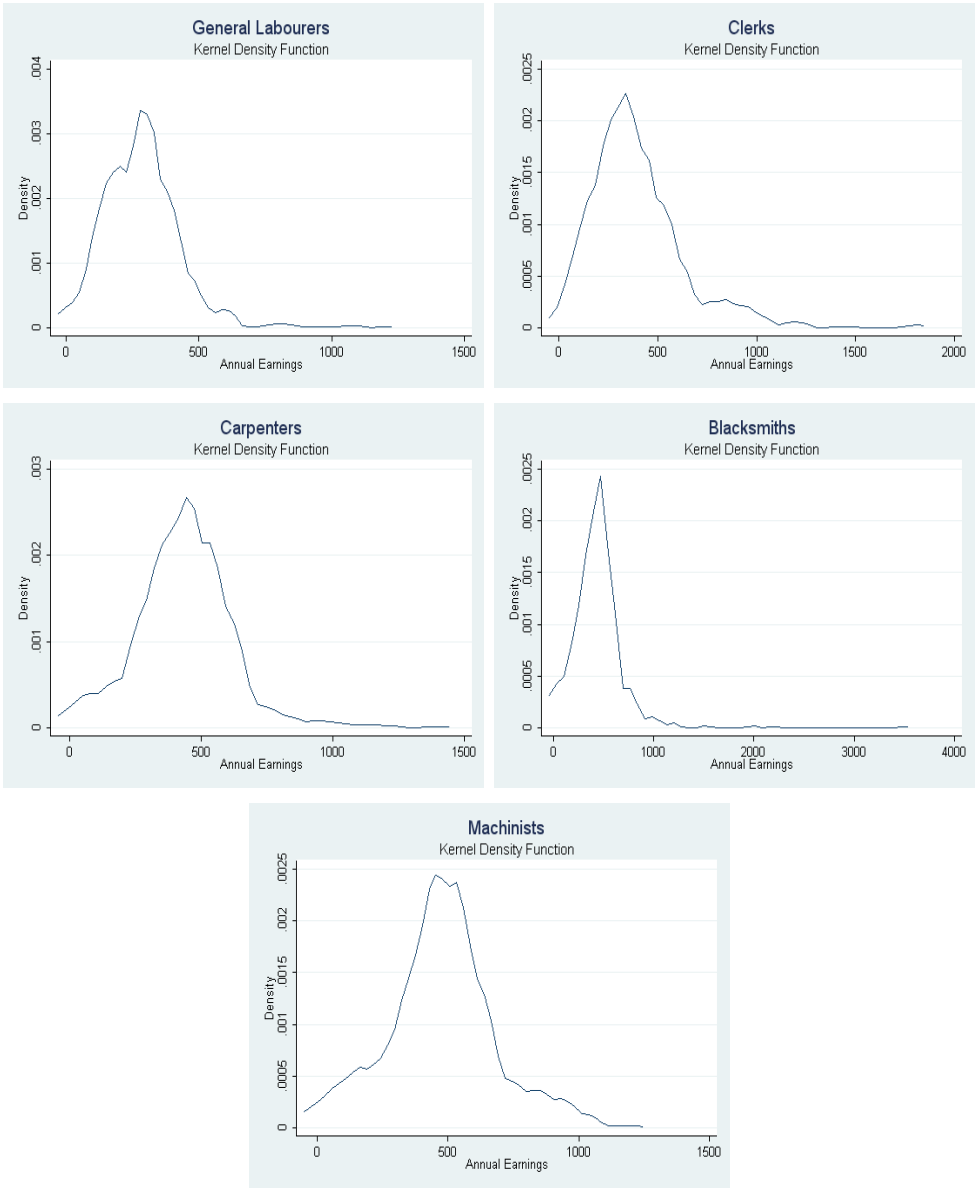


(a) Within-occupation measures of distribution



(b) Percentile shares

Figure 2: Kernel Density functions of Earnings for Selected Occupations



most at $486 and $450 respectively. Clerks and carpenters each earn virtually the same on average at $428.4 per year. These numbers indicate a big discrepancy between the earnings potential of general labourers and skilled occupations. Skills that have become valued from industrialization including those involved in the manufacturing of metal goods and machinery, are being paid well reflective of increased demand. From inspection of patterns in the mean and median, it also appears that clerical labour was a very well-paid occupation compared to common labourers, making annual earnings closer to those enjoyed by carpenters and blacksmiths. A closer examination of within-occupation earnings distributions reveals that these patterns are not as clear cut as one might imagine.

Kernel density functions are presented in Figure 2 and help to illustrate the differing distributional structures within this handful of occupations. All occupations present an earnings distribution that is skewed to the right, with the highest mass of workers found huddled somewhere very slightly to the left of the mean. This is less pronounced for machinists and carpenters. Major differences in distributional structure are found in the size of the upper tail and in the density of mean earners. Blacksmiths as well as clerks (given the low density of the central hump in clerical earnings) have the longest upper tails. In the case of blacksmiths this is an isolated phenomenon since the rest of the distribution is highly concentrated around $500 per year, but the tail extends to over seven times as large as the mean. Clerical earnings are much more spread out, with a larger central mass, and a thicker upper tail, although it only extends three times as large as the mean. The distribution of machin-

ists is even less skewed with a small central mass and short, thick upper and lower tails.

Comparing subgroup distribution by kernel density functions gives a graphical sense of the differing structure of earnings but makes robust comparisons very difficult. This purpose is much better served by a combination of mean earnings by quintile and, in order to analyze inequality, the share of income amassed by each quintile. From Figure 1 it is clear that measures such as the mean, median, and mean earnings of the middle quintile rank occupations in a very similar manner – labourers (making the least) followed by clerks, carpenters, blacksmiths, and machinists (making the most). The only deviation among these central measures is found in the case of clerks and blacksmiths. The middle quintile for these occupations earn less than the mean or median worker. By examining the mean of top earners or their quintile share in total income it is evident that the large upper tails discovered by kernel density functions are pulling up the mean and median. Whether one wants to include these top earners in the calculation of point estimates depends on the purpose of the study. However, making comparisons across occupations or using occupational mean earnings as proxies for movements in inequality will surely produced skewed results that must be carefully analyzed in determining their validity.

The impact of distributional differences is found even between mean and median estimates. Currently available data is often given by wage ranges – a common phenomenon in the Department of Labour's Wages and Hours of Labour. This has lead researchers to typically take the median value as a representative measure of earn-

ings. In this sample, the median underestimated mean earnings by 1-7% for clerks, carpenters, and machinists, while overestimating common labour earnings. These discrepancies can have a major impact on cross-occupational comparisons and in this case artificially shrink skill income gaps.

So far I have concentrated on the earnings of workers from the middle of the distribution and have shown how patterns based on these measures are sensitive to influences from the structure of the entire distribution. These influences are important because they may differ in a non-systematic manner across occupations. Calculating the Gini index shows a range of .19-.27 from machinists (the most equal) to clerks (the most unequal). These relatively high values are proof of how much variation is found in the inequality of occupations that are typically believed to be relatively homogeneous. Unlike the middle of the distribution, upper and lower tails exhibit much more varied levels of inequality across occupations than mean earnings. Across the trades selected here, the bottom fifth earn on average 39%-47% of mean earnings, the middle fifth earn 90-99% of total mean earnings, and the top fifth earn 155-183% of mean earnings. Income shares on the other hand, especially for the middle 20%, vary much less than mean earnings. The bottom quintile amass between 7.8 and 9.4% of total earnings, the middle 18-20% and the top 30-36%. This large and stable income share for the middle quintile hints at the existence of a large and homogeneous middle-class within these occupations.

Mean earnings of top and bottom percentiles can differ quite dramatically. Looking at mean earnings in the bottom 20% we see that relatively poor carpenters actually make more than similarly poor blacksmiths, despite making almost 10% less on average. Top earners vary even more. Highly paid carpenters fair the worst compared to clerks, blacksmiths and machinists which each earn around 50% more than the top 20% common labourers. Identical to the results of Green and Mackinnon (2001) I find that clerks earn on average almost 30% more than general labourers. However, even though average and median earnings of clerks are among the lowest, we can see that top-paid clerk positions are actually very highly paid relative to the rest of the distribution. Many of these clerks earning over $600 per year are in fact working in the banking or financial sector which was one the fastest growing industries in Canada at the turn of the century (Emery et al., 2008, Ch 13). This sectoral differentiation of growth would thus make the earnings of clerks in industries other than bank and finance seem higher than they actually are. Any analysis using clerical earnings at the turn of the century must therefore be careful to differentiate between those working in different industries [6].

4.2 The Chronically Under-employed

There is increasing evidence that the chronically under-employed (at work in their principal occupation for less than 9 months) made up a large proportion of the Canadian work force in 1901. This seemed to be case especially in urban industries where

[6]Ideally I would have liked to look at occupation earnings by region or city but observation numbers limit this exercise. For example, only 153 workers of these five occupations are found West of Ontario.

growing industries produced high rates of unemployment and high turnover rates especially in factory occupations (Baskerville and Sager 1995, 1997). Isolating my sample to male workers over the age of 18, 25.54% of those who reported some work duration reported working under 10 months of the year in their principal trade. What the rates of pay were for these workers is unknown and none of the earnings data available to date include them. Published Census records include only workers employed more than 9 months in their principal occupation in calculating average earnings. Department of Labour wage rates are minimum union rates that apply only to full-time employees fully-qualified in their trade and, as MacKinnon (1996) argues, tended to underestimate the rates of less-skilled trades. In addition, David and Rosemary Gagan (1990) emphasize that actual or potential yearly income is very poorly captured in wage rates when no information on work duration is available. Looking at working-class living standards in Ontario in the late nineteenth century, they find that full employment was much more important than weekly or monthly salaries in determining living standards given the income of only the household head. Baskerville and Sager (1998) demonstrate that the chances of being unemployed did not differ across Canadian cities but rather depended heavily on occupation. This suggests that comparing occupation earnings at the national level is a valid exercise which should be unaffected by local labour market phenomenon.

With data on both earnings from chief occupation and work duration I estimate implied annual earnings using the full labour force available of both fully-employed

Table 2: Employment and duration reporting rates for adult male workers, 1901

	Labourers	**Clerks**	**Carpenters**	**Blacksmiths**	**Machinists**
Total	5,428	1,882	2,274	1,731	693
Unemployed (no earnings and no duration)	15.46%	9.51%	18.78%	22.13%	6.35%
Employed (with earnings or duration)	84.54%	90.49%	81.22%	77.87%	93.65%
Of those Employed:					
Full time	17.15%	27.95%	31.19%	50.30%	63.02%
Part time	9.17%	1.64%	17.87%	14.32%	12.17%
Unknown duration	73.68%	70.41%	50.95%	35.39%	24.81%

and part-time employees. To do this I calculate average monthly wages given months employed and multiply by twelve. Ideally numbers of weeks or days would have given the best measures of work duration because of the subjective rounding patterns respondents or enumerators must have performed. However I assume that those who would round up and those who would round down would cancel each other out or in the worst case scenario part time workers would work less than the number of months reported.

From Table 2 it is clear that different occupations, even within the same sector, can have dramatically different patterns of employment status. Those that reported no earnings and no work duration can essentially be considered unemployed while reporting some earnings or some work duration would qualify the respondent as employed. Unemployment ranged between 6.35% for Machinists to 22.13% for Black-

Table 3: Implied Annual Earnings including part-time workers, 1901

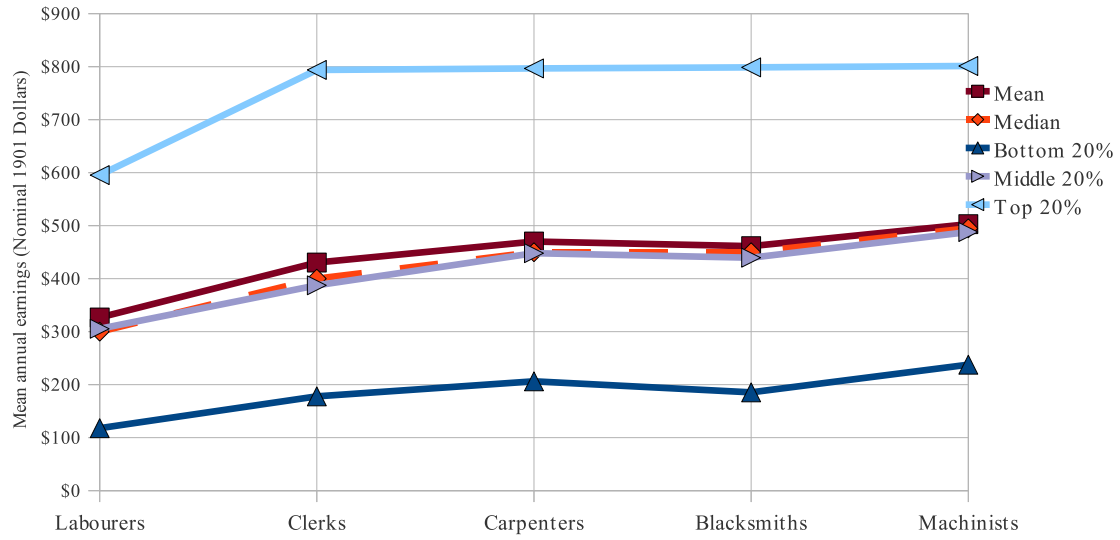
		Labourers	**Clerks**	**Carpenters**	**Blacksmiths**	**Machinists**
	Observations	1208	504	906	871	488
	Mean	$326.74	$430.34	$470.03	$461.31	$502.86
	Std. Dev.	226.33	240.08	265.42	257.45	212.19
	Median	$300.00	$400.00	$450.00	$450.00	$495.00
	Gini	0.27	0.27	0.24	0.24	0.20
Mean Earnings	Bottom 20%	$117.94	$177.99	$206.31	$185.47	$237.71
	Middle 20%	$305.49	$387.42	$448.00	$439.50	$487.97
	Top 20%	$595.38	$793.78	$796.71	$798.64	$801.22
Share of Total Income	Bottom 20%	7.22%	8.29%	8.82%	8.08%	9.49%
	Middle 20%	18.71%	18.04%	19.04%	19.03%	19.29%
	Top 20%	36.33%	36.60%	33.86%	34.59%	31.67%

smiths. Among those employed, blacksmiths and machinists reported work duration much better than labourers or clerks. This might again be a reflection of the need impressed upon enumerators to record information about industry workers. It is impossible to tell the work duration of workers who reported none though it is clear from their earnings that they must have been employed for some period of the year. However there is evidence to suggest that full time employment varied even more than unemployment across occupations and was not a common occurrence. Among workers with a reported work duration, the percentage reporting working over 11 months of the year was only 17.15% for general labourers, 27.95% for clerks, and 31.19% for carpenters. Blacksmiths and machinists were employed for the longest lengths of time but even they only had full time employment rates of 50.30% and 63.02% respectively.

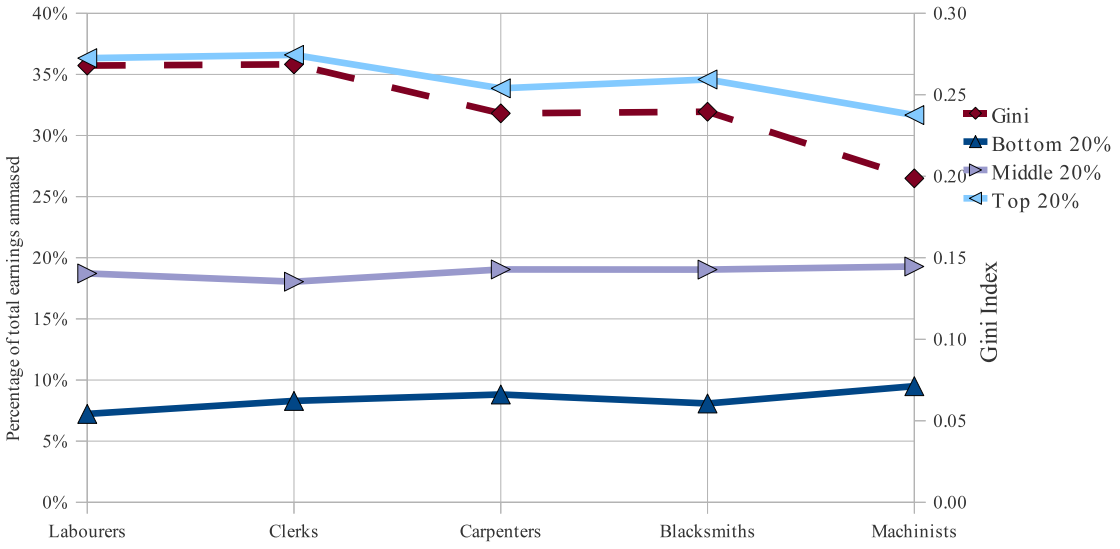
Although the median annual earnings stayed about the same for all occupations, the distribution of earnings shows a clear rightward movement towards higher earnings compared to distributions using only full-time workers. Mean earnings rose for all occupations with the highest increases for labourers, which now appear to have earned about $31 more per year and for carpenters who report more than a $40 increase in average annual wages, or the equivalent of one month's salary. Higher average earnings are also reported for almost all income groups. The shares of total income amassing to the top fifth increase by as much as 4% while lower income brackets make up a smaller proportion of total income though differences are not as considerable.

This exercise shows that the inclusion of the chronically unemployed can dramati-

Figure 3: Ordering Mean Earnings of Selected Occupations with Implied Part-time Earnings



(a) Within-occupation measures of distribution



(b) Percentile shares

cally change cross-occupational income patterns. These results suggest that part time workers earned on average higher monthly rates that those fully employed. This provides direct evidence to the suspicions voiced by Baskerville and Sager (1998) that the chances of being unemployed for some occupations actually rose with higher wages, implying that skilled workers experiencing chronic lay-offs were able to bargain for higher wages with employers. More importantly, it shows that currently used wage rates and mean annual earnings are an inaccurate reflection of the earnings distribution of the Canadian work force at the turn of the century. By disregarding the large proportion of chronically unemployed workers, estimates are biased downwards in a non-constant way across occupations. It appears that occupations with the least permanent employment, like general labourers and factory workers, suffered most from this bias.

From the last two exercises it becomes clear how deceptive point estimates of occupational earnings can be for distributional studies. Top earners can have very different attributes from their colleagues that can impact mean or median estimates considerably. Comparing occupational earnings at different income shares can give markedly different conclusions on how employees at work in separate trades fare relative to one another. Furthermore, chronic unemployment was a widespread phenomenon among the Canadian workforce in the early 20th century and excluding their wages from occupational aggregates distorts comparisons by underestimating the potential earnings of non-permanent occupations.

5 Regional patterns in earnings distribution

Occupational earnings are generally relatively homogeneous and often reflect non-systematic differences across workers such as human capital or premium for working in a particular industry or sector. The distribution of income at the national or regional level on the other hand encompasses many different occupations, skill sets, and local factors that widen the spread of income and increase observed inequality. Assessing the distribution of well-being as captured by earnings in Canada at the turn of the century and its regional patterns must therefore move beyond simple occupational or skill differentials. In this section, I assess regional patterns of earnings distribution by applying decompositional analysis to study the distribution of earnings within cities, emphasizing the importance of sensitivity to tails of the distribution. The point of this exercise is not necessarily to make robust comparisons of regional incomes but rather to explore the methodology of measuring inequality at the national level. Results turn out to be greatly dependent upon within-group distributions, providing further evidence for the importance of decompositional analysis.

I isolate my study to 10 Canadian cities and the combined population of Calgary, Regina, and Edmonton [7] for which price levels are available from Emery and Levitt (2002) to calculate real annual earnings. Urban centers provide a more homogeneous group of workers and minimize the impact of under-reported farming income. In order to compare regional earnings patterns from this data to other sources [Table 4], I present relative differences of earnings in Vancouver, Winnipeg, Montreal and Halifax,

[7] Insufficient population samples for these cities lead me to combine their numbers into a single region which, for the remainder of the paper, will simply be referred to as the Prairies.

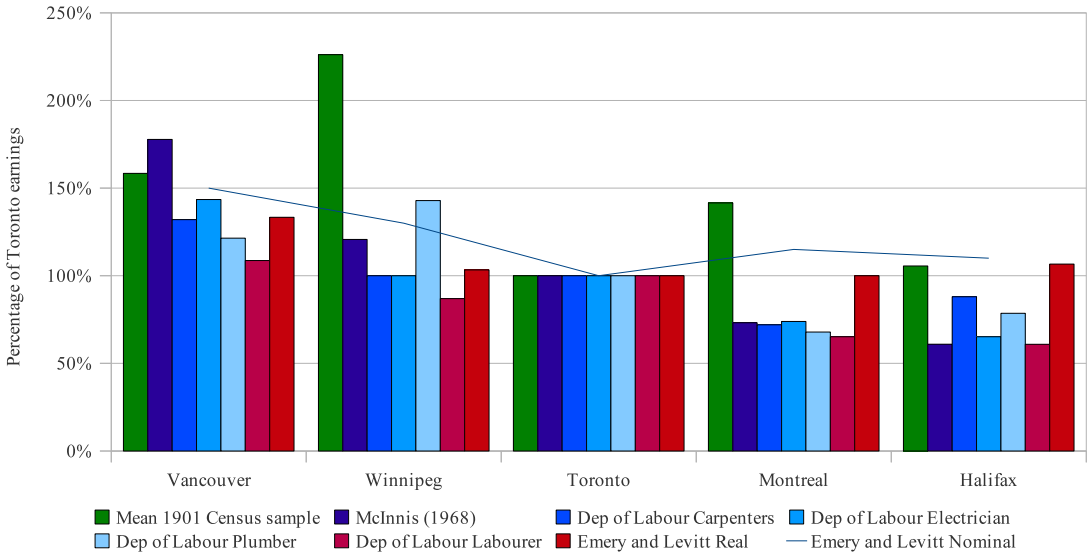
to those in Toronto for adult male workers. Mean adult earnings show similar regional differences to other studies, with the exception of a few anomalies. I find high earnings in Vancouver, Victoria, and western cities, and low relative earnings in Quebec City and the Maritimes, just as other have. However, average earnings in Winnipeg are over 40% higher than the national average, even surpassing cities in British Columbia and the Prairies. At the same time, earnings in Toronto and Hamilton are among the lowest despite being home to some of the greatest documented growths in Canada for this period. These differences can be attributed to a variety of causes including unemployment rates, composition of income sources, and occupational composition, that varied across cities. As discussed in the previous section, currently published wage rates were calculated using only full-time employees and therefore miss a considerable proportion of the working population. Where part-time work was common, such as central Canada (Baskerville and Sager, 1995), it is no surprise to find lower average earnings per adult.

Figure 4 also shows how regional income patterns can differ considerably depending on the data source used. Department of Labour wage rates, for example, amplify earnings in Toronto especially for general labourers. McInnis (1968), using estimated annual participation income per capita for 1910-11, captures much lower income levels in Montreal and much higher levels in the West relative to other Canadian regions. Using these estimates in time-series studies in search of convergence or divergence, as their authors have done, may still be a completely valid and valuable exercise. However, determining regional income differentials at certain points in time can pro-

Table 4: Comparing alternative measures of regional earnings differentials

	Vancouver	**Winnipeg**	**Toronto**	**Montreal**	**Halifax**
Mean adult yearly earnings	$435.65	$621.95	$275.05	$389.27	$290.42
- 1901 federal census sample					
McInnis (1968)	$464	$315	$261	$191	$159
- Estimated annual participation income per capita 1910-11					
MacKinnon (1996) Machinists	$0.29	$0.26	$0.19	$0.19	
MacKinnon (1996) Helpers	$0.17	$0.16	$0.16	$0.13	
MacKinnon (1996) Labourers		$0.13	$0.12	$0.13	
- CPR wage rates (dollars per hour) 1901					
Dep of Labour Carpenters	$0.33	$0.25	$0.25	$0.18	$0.22
Dep of Labour Electrician	$0.33	$0.23	$0.23	$0.17	$0.15
Dep of Labour Plumber	$0.34	$0.40	$0.28	$0.19	$0.22
Dep of Labour Labourer	$0.25	$0.20	$0.23	$0.15	$0.14
- Department of Labour supplements in the Labour Gazette (dollars per hour), 1901					
Emery and Levitt (2002) Real	$0.40	$0.31	$0.30	$0.30	$0.32
Emery and Levitt (2002) Nominal	$0.30	$0.26	$0.20	$0.23	$0.22
- Machinists (dollars per hour), 1901					

Figure 4: Comparing alternative measures of regional earnings differentials



duce vastly different results depending on the characteristics of the underlying data. Occupation wage rates in particular may be problematic even in long-run studies if regional labour markets change relative returns to skills on the count of technological change or industry growth. Regional income differentials should not be determined using wages of single occupations or mean per capita estimates alone. Micro-data from the 1901 census includes a more comprehensive set of workers and can therefore give very different conclusions concerning regional inequality than previous studies.

To assess intra-city earnings distribution, I calculate several inequality measures including the Gini, different income shares, and generalized entropy levels sensitive to different parts of the distribution. I find that the inequality of earnings in Canada was much lower than previous local studies using probate sample have shown, although

its nature differed considerably across regions. The Gini coefficient ranged from .31 for the Prairie cities (the most equal) to .49 for Winnipeg (the most unequal). Other cities of high inequality as captured by the Gini were Victoria at .47 and Montreal at .41. These values are much higher than those for full-time workers within the same occupation as calculated in the last section but are extremely low in comparison to Gini coefficients between .7 and .9 typical of probate-based historical studies. From the point of view of income shares, the top quintile in my sample make up 40%-60% of total earnings while the bottom amass between 4% to 7%. This too is less inequality than the patterns documented by others that report ranges of 70-80% for the top twenty percent and 1-3% for the bottom. It is true that one possible cause for this low measured inequality may be the exclusion of non-labour income earners with typically larger, more diverse incomes. However, if we are willing to assume that under-reporting was not a huge issue in the federal census, interpolating zero-earners as has been done by Osberg and Siddiq (1988) and Jones (1977) in her pioneering paper on the inequality across 13 US colonies, would not be an appropriate modification to this sample. Instead, these measures must be accepted as what they are – a representation of the low inequality levels of workers employed in wage-earning occupations. The fact that inequality measures based on estimates of wealth (including alternate income sources such as property, financial assets, and self-employed income) produce much higher values indicates an unequal distribution of these alternative income sources. Whether they are linked to class differences, age profiles, or other demographic characteristics cannot be deduced from this sample. It is clear however

that by ignoring low-income wage-earners, studies based on probate samples do not fully capture national inequality and very possibly overestimate it.

Compared to the variation in average earnings, the Gini coefficient is considerably stable across these 11 urban centres. Cities with high average earnings, including Victoria, Winnipeg, and Montreal exhibit the highest levels of inequality. This supports the general Kuznets-curve hypothesis that growth is often associated with increased inequality. However, the variation in the Gini index for this sample is low, with a standard deviation of just 0.0667. Furthermore, by selecting alternative inequality measures such as generalized entropy levels it becomes evident that inequality patterns cannot be summarized so easily. The nature and location of inequality in fact differed greatly among these eleven Canadian cities although summary measures like the Gini index and average earnings can hide this.

Table 5 and Figures 5a. and 5b. present multiple inequality measures including three generalized entropy levels, each sensitive to changes in different parts of the earnings distribution, and percentile shares of total earnings. Percentile shares and generalized entropy levels give different information about inequality but will serve the same purpose here - to assess distribution of earnings at different points in the distribution rather than concentrating on overall or aggregate earnings differentials.

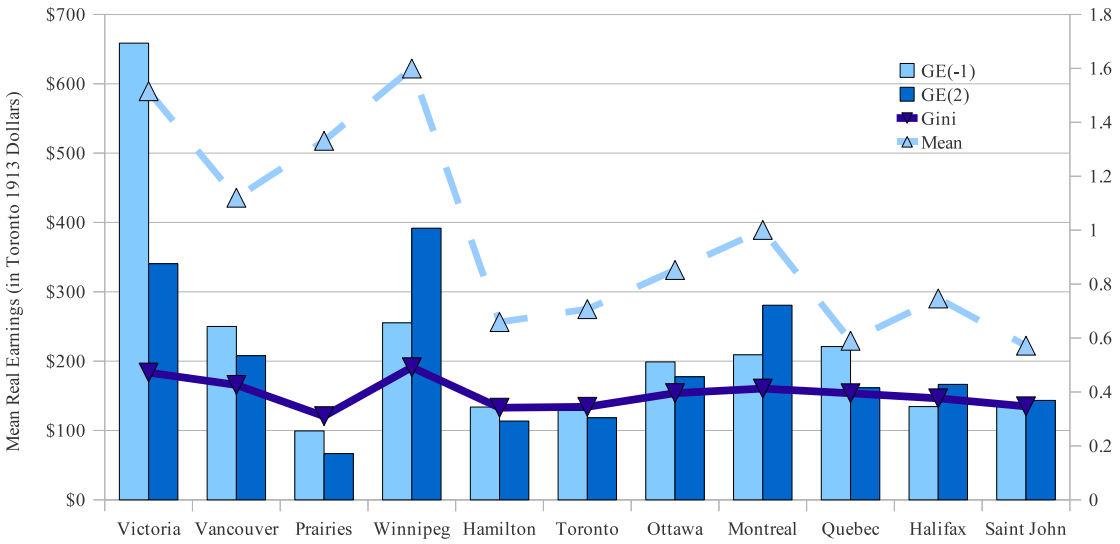
As Figure 5b. shows, the Gini index is extremely sensitive to top earners since it

Table 5: Regional Mean Earnings and Inequality

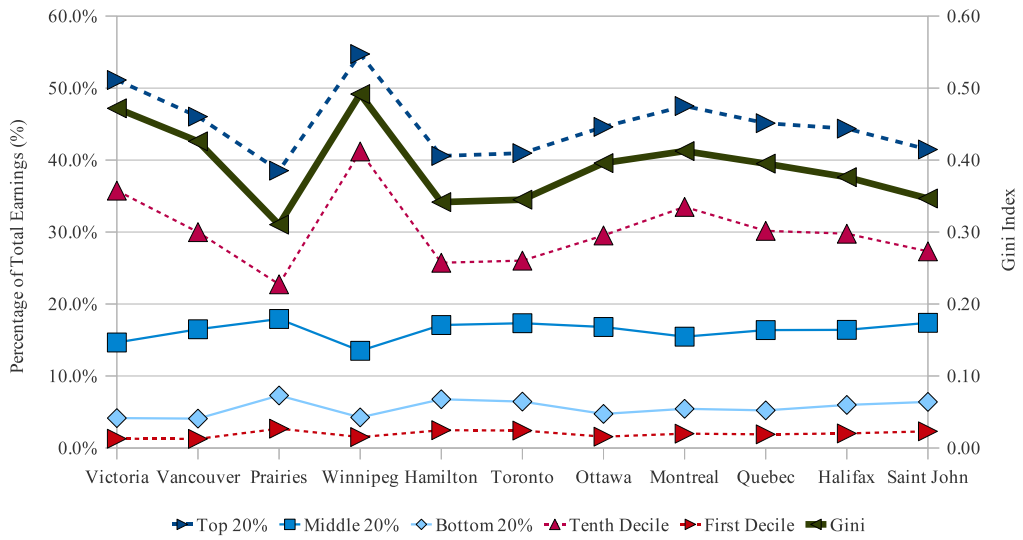
Adult real total earnings by major Canadian city, 1901

	Victoria	**Vancouver**	**Prairies**	**Winnipeg**	**Hamilton**	**Toronto**	**Ottawa**	**Montreal**	**Quebec**	**Halifax**	**Saint John**
Observations	347	459	287	727	798	2994	817	4474	893	506	502
Mean	$589.32	$435.65	$517.83	$621.95	$256.33	$275.05	$331.46	$389.27	$229.41	$290.42	$222.33
Median	$429.84	$363.36	$450.60	$405.50	$215.28	$236.52	$272.90	$297.20	$183.00	$240.12	$184.70
Gini	0.47	0.43	0.31	0.49	0.34	0.34	0.4	0.41	0.4	0.38	0.35
GE(-1)	1.69	0.64	0.26	0.66	0.34	0.33	0.51	0.54	0.57	0.35	0.35
GE(0)	0.43	0.36	0.18	0.43	0.21	0.22	0.3	0.31	0.29	0.26	0.23
GE(1)	0.45	0.34	0.16	0.51	0.21	0.21	0.29	0.36	0.29	0.27	0.24
GE(2)	0.88	0.53	0.17	1.01	0.29	0.3	0.46	0.72	0.42	0.43	0.37
Share held by the:											
Top 20%	51.1%	46.0%	38.5%	54.7%	40.6%	41.0%	44.6%	47.5%	45.1%	44.4%	41.5%
Middle 20%	14.7%	16.5%	17.9%	13.5%	17.1%	17.3%	16.8%	15.5%	16.4%	16.4%	17.4%
Bottom 20%	4.2%	4.1%	7.3%	4.2%	6.8%	6.4%	4.7%	5.4%	5.2%	6.0%	6.4%
Tenth Decile	35.8%	30.0%	22.7%	41.2%	25.7%	26.0%	29.5%	33.5%	30.2%	29.8%	27.3%
Ninth Decile	15.4%	16.0%	15.8%	13.5%	14.8%	14.9%	15.1%	14.0%	15.0%	14.6%	14.1%
Eight Decile	11.9%	13.0%	12.3%	10.3%	12.5%	12.2%	12.2%	11.0%	12.0%	11.4%	11.4%
Seventh Decile	9.7%	10.7%	10.1%	8.0%	10.9%	10.9%	10.3%	9.5%	10.2%	9.7%	10.5%
Sixth Decile	8.2%	9.2%	9.5%	7.1%	9.3%	9.4%	9.1%	8.2%	8.4%	9.0%	9.4%
Fifth Decile	6.5%	7.3%	8.4%	6.4%	7.8%	7.9%	7.7%	7.2%	7.9%	7.5%	8.0%
Fourth Decile	4.8%	5.5%	7.2%	5.1%	6.7%	6.8%	6.3%	6.0%	6.4%	6.9%	7.2%
Third Decile	3.7%	4.2%	6.7%	4.0%	5.4%	5.4%	5.0%	5.1%	4.7%	5.3%	5.6%
Second Decile	2.9%	2.8%	4.6%	2.7%	4.3%	4.0%	3.2%	3.5%	3.3%	4.0%	4.1%
First Decile	1.3%	1.3%	2.7%	1.5%	2.5%	2.4%	1.6%	2.0%	1.9%	2.0%	2.3%

Figure 5: Regional differentials of Inequality



(a) Regional Mean Earnings and Inequality



(b) Regional Percentile Shares of Real Earnings

matches the pattern of inequality as portrayed by the share of total income of the top 20 percent almost exactly[8]. Although the income shares of the top percentiles vary considerably across cities, middle and lower percentiles earn a much more constant proportion of total earnings. Cities such as Winnipeg, Victoria and Montreal have the most earnings amassing to top earners while cities in the Prairies, Hamilton and Toronto have the lowest. Middle and lower quintiles accordingly make up less of total earnings in the former case and a larger proportion in the latter. However, the way in which the rest of this income is divided up among lower percentiles differs across cities. The top ten percent of earners in Victoria for example amass over 51% of total income, over 5% more than those in Vancouver. Despite this, bottom deciles in Victoria earn the same and in some cases more than those in Vancouver. In comparison, although the gap in income share held by the top decile between Vancouver and Toronto is almost the same as that between Vancouver and Victoria, bottom percentiles in Toronto amass consistently more than one percent more income than those in Vancouver. This suggests that Victoria, in contrast to other cities, had a fairly well-to-do lower class with high earning relative to the rest of the city population[9]. The same thing can be found for Winnipeg by comparing patterns in income shares with those in Montreal and Vancouver.

[8]The scales of the two measures are very different so cannot be used to make level comparisons across regions but work well to rank cities in the same order.

[9]I restrain from calculating mean earnings of percentiles here since the objective is to examine within-city inequality. There are many factors including industrial and occupational makeup that can impact mean levels but have nothing to do with how total earnings are distributed and therefore would potentially make mean earnings by quantiles a deceptive method of comparison.

This shows that inequality is not the same within different parts of the earnings distribution for different samples. Percentile shares, however, are a limited method of assessing exactly how this inequality is distributed. In part this is because they ignore the inequality within each percentile. Breaking up quintiles and deciles into more and more percentiles serves to illuminate this issue but only partly. The result would be akin to comparing Lorenz functions and, as has been explained, it is hardly ever easy to find Lorenz dominance among similar subgroups[10]. The alternative is to use ordinal measures of inequality, but as I have shown, the Gini coefficient by itself is not enough for understanding inequality within multiple samples.

The advantage of using micro-data is that we are not confined to use only percentile shares and instead can calculate many different inequality measures. General entropy levels have the advantage over the Gini index of being able to look at different parts of the earnings distribution, but also capture within-quantile inequality which income shares do not.

Generalized entropy levels are calculated with four sensitivity parameters; -1 (which places most weight on dispersion in the lower tail), 0 (which is similar to the Gini index in that it is most sensitive to the middle of the distribution), 1, and 2 (which is most sensitive to inequality in the upper tail). Cities with low Gini coefficients (the Prairies, Hamilton, Toronto and to some extent Halifax and Saint John)

[10]In this sample, neither Lorenz nor Generalized Lorenz dominance was found between any of the cities. This means that there were no two produced Lorenz curves that did not cross at least once. This is an example of one case, as are many, where Lorenz curves cannot produce informative conclusions despite their theoretical superiority in analyzing inequality.

have GE levels that are very low and that do not differ considerably with different sensitivity parameters. Although average earnings differ quite dramatically across these cities, it appears that its distribution within them is very equal. The same cannot be said for the rest of Canada. Once we start looking at cities with higher overall levels of inequality as captured by the Gini index, GE(0), and GE(1) we find that the source of this inequality begins to vary across samples. In other words, the *distribution of inequality* itself differs. In Winnipeg and Montreal respectively, GE(2) is 65% and 75% bigger than GE(-1), indicating that more inequality is coming from the dispersion of earnings among top earners than in from those in the lower tail of the earnings distribution. For Vancouver, Ottawa, and Quebec where the Gini coefficient shows roughly the same level of inequality as in Montreal, more inequality is found among the poor. Victoria is the biggest anomaly because the calculated GE(-1) which is most sensitive to changes in the bottom tail of the earnings distribution is significantly higher than any of the other generalized entropy levels for that city or any other in this sample. Also of note is that high mean earnings are not always accompanied by inequality in top earners as might be expected. This was the case for Winnipeg and Montreal but not at all for cities in the West. Victoria, Vancouver and Prairie urban centres exhibited more inequality when weight is placed on dispersion in the lower tail, despite having some of the largest average annual earnings.

With this discussion, the picture of regional earnings differentials in Canada during the turn of the century has become more complete and richer in detail. Where until

now average regional wages and income per capita may have been the primary tools used to study national inequality trends, using micro-data and multiple inequality measures shows the variability in earnings distribution within cities as well as among them. Some cities may have high average earnings because of highly unequal upper tails, such as Winnipeg and Montreal, while others such as Vancouver and cities in the Prairies were relatively homogeneous in their earning patterns. By showing that an important source of inequality is also the lower half of the income distribution this also demonstrates the potential of probate samples to misrepresent inequality patterns.

6 Appendix

Table 6: Price Levels of 13 Canadian Cities, 1901

Source: Emery and Levitt (2002)

	Urban Place	**Prices**
1	Victoria	79.6
2	Vancouver	75.7
3	Edmonton	75.1
4	Calgary	84.7
5	Regina	90.5
6	Winnipeg	81.1
7	Hamilton	59.8
8	Toronto	65.7
9	Ottawa, Ottawa East	65.6
10	Montreal	74.3
11	Quebec City	61
12	Halifax	66.7
13	Saint John	59.2

Table 7: Occupation Categories and CCDO Codes

For Male Workers over 18 Years Old

		Number Reported	Percentage with Earnings	
White collar				
[1]	11	1576	73.2%	Managerial, administrative, financial management, government, and related
[2]	21	124	62.1%	Scientists, architects, and related professionals
[3]	23	345	42.3%	Law and social institutions
[4]	24	329	2.4%	Students
[5]	25	530	47.9%	Occupations in religion
[6]	27	409	78.2%	Teaching professions
[7]	31	587	32.5%	Occupations in medicine and health
[8]	33	236	61.9%	Occupations in the arts and writin
Pink collar, retail, and other service				
[9]	41	2495	89.1%	Clerical and bookkeeping occupations
[10]	51	3154	47.3%	Commerce and sales occupations
[11]	61	2440	67.6%	Service occupations
Skilled blue-collar				
[12]	8186	8749	79.0%	Primary and secondary processing and manufacturing
[13]	87-88	4849	78.6%	Construction
[14]	91-93	3194	84.8%	Transportation
[15]	95	1077	85.8%	Communications (printing and telegraph)
Unskilled labour				
[16]	71	30377	12.7%	Agricultural
[17]	73	1376	60.3%	Fishing, hunting and trapping
[18]	75	751	81.2%	Logging and forestry
[19]	77	1749	76.8%	Mining, and oil and gas production
General labour				
[20]	991	6076	83.1%	General labour
Unclassifiable				
[21]	992-998	7426	7.1%	

References

. *4th Census of Canada, 1901. Ottawa: S.E. Dawson, 1902.*

. various years. *Department of Labour, Labour Gazette.* (Ottawa, King's Printer).

Atkinson, A. B., & Bourguignon, F. (eds). 2000a. *Handbook of Income Distribution.* Handbook of Income Distribution, vol. 1, no. 1. Elsevier.

Atkinson, A. B., & Bourguignon, F. 2000b. Income Distribution and Economics. *Chap. 1, pages 1–53 of: Handbook of Income Distribution.* North-Holland.

Atkinson, A.B. 1970. On the Measurement of Inequality. *Journal of Economic Theory*, **2**, 244–263.

Baskerville, P. A., & Sager, Eric W. 1997. Unemployment, Living Standards, and the Working-Class Family in Urban Canada in 1901. *The History of the Family*, **2**(3), 229–254.

Baskerville, Peter A., & Sager, Eric W. 1995. Finding the Work Force in the 1901 Census of Canada. *Histoire sociale-Social History XXVII*, **no. 56**, 521–40.

Baskerville, Peter A., & Sager, Eric W. 1998. *Unwilling Idlers: The Urban Unemployed and Their Families in Late Victorian Canada.* University of Toronto Press.

Beach, C., & Davidson, R. 1983. Distribution-Free Statistical Inference with Lorenz Curves and Income Shares. *The Review of Economic Studies*, **50**(4), 723–735.

Bertram, Gordon W., & Percy, Michael B. 1979. Real Wage Trends in Canada 1900-26: Some Provisional Estimates. *The Canadian Journal of Economics*, **12**(2), 299–312.

Bourguignon, F., & Morrisson, C. 1990. Income distribution, development and foreign trade: a cross-sectional analysis. *European Economic Review*, **34**, 1113–1132.

Champernowne, D. 1974. A Comparison of Measures of Inequality of Income Distribution. *The Economic Journal*, **84**, 788–816.

Coe, Patrick J., & Emery, J. C. Herbert. 2004. Disintegrating Canadian Labour Market? The Extent of the Market Then and Now. *The Canadian Journal of Economics*, **37**(4), 879–897.

Coulombe, Serge, & Lee, Frank C. 1995. Convergence across Canadian provinces, 1961 to 1991. *Canadian Journal of Economics*, **28**, 886–898.

Cowell, F., & Jenkins, S. 1995. 1995. How Much Inequality Can be Explained: A Methodology and An Application to the USA. *Economic Journal*, **102 (2)**, 421–430.

Cowell, F. A. 1977. *Measuring Inequality*. Oxford, Philip Allan Publishers.

Cowell, F.A., & Kuga, K. 1981. Additivity and the entropy concept: An axiomatic approach to inequality measure. *Journal of Economic Theory*, **25**, 131–143.

Darroch, A. Gordon. 1983. Early Industrialization and Inequality in Toronto, 1861-1899. *Labour*, **11**, 31–61.

Darroch, Gordon A., & Soltow, Lee. 1992. Inequality in landed wealth in nineteenth-century Ontario: structure and access. *The Canadian Review of Sociology and Anthropology*, **29**(2), 167–189.

Davidson, R. 2009. Reliable Inference for the Gini Index. *Journal of Econometrics*, **150**, 30–40.

Deininger, K., & Squire, L. 1996. Measuring Income Inequality: A New Data Base. *World Bank Economic Review*, **10(3)**, 565–91.

Di Matteo, Livio. 1997. The Determinants of Wealth and Asset Hodling in Nineteenth-Century Canada: Evidence from Microdata. *The Journal of Eonomic History*, **57**(4), 907–934.

Di Matteo, Livio. 2001. Patterns and Determinats of Wealth Inequality in Late-Nineteenth-Century Ontario. *Social Science History*, **25**(3), 347–380.

Di Matteo, Livio. 2004. Boom and Bust, 1885-1920: Regional Wealth Evidence from Probate Records. *Australian Economic History Review*, **44**(1), 52–78.

Di Matteo, Livio. 2006. Welath and Inequality on Ontario"s Northwestern Frontier: Evidence from Probate. *Histoire sociale-Social History*, **XXXVIII**(75), 79–104.

Di Matteo, Livio, & George, Peter. 1992. Canadian Wealth Inequality in the Late Nineteenth Century: A Study of Wentworth County, Ontario, 1872-1902. *Canadian Historical Review*, **LXXIII**(4), 453–83.

Di Matteo, Livio, & George, Peter. 1998. Patterns and Determinants of Wealth among Probated Decendents in Wentworth County, Ontario, 1872-1902. *Social History*, **XXXI**(61), 1–34.

Di Matteo, Livio, & George, Peter J. 1996. Quantitative Methods, Historical Micro-data and the Interpretation of Canadian Economic History. *Journal of Canadian Studies*, **31**(2), 45–62.

Emery, J. C. Herbert, & Levitt, Clint. 2002. Cost of Living, Real Wages and Real Incomes in Thirteen Canadian Cities, 1900-1950. *The Canadian Journal of Economics*, **35**(1), 115–137.

Emery, J. C. Herbert, Norrie, Kenneth, & Owram, Douglas. 2008. *A History of the Canadian Economy*. 4 edn. Thomson Neilson.

Gagan, David, & Gagan, Rosemary. 1990. Working-class Standards of Living in late-Victorian Urban Ontario: A review of the miscellaneous evidence on the quality of material life. *Journal of the Canadian Historical Association*, **1**, 171–194.

Green, Alan, & MacKinnon, Mary. 2001. The Slow Assimilation of British Immigrants in Canada: Evidence from Montreal and Toronto, 1901. *Explorations in Economic History*, **38**(3), 315–338.

Green, Alan G. 1967. Regional Aspects of Canad's Econmic Growth, 1890-1929. *The Canadian Journal of Economics and Political Science*, **33**(2), 232–245.

Jones, A. H. 1977. *American Colonial Welath: Dcument and Methods, three volumes*. Arno Pres, New York.

Jones, A. H. 1980. *Wealth of a Nation to be*. Columbia University Press, New York.

Keyssar, Alexandar. 1986. *Out of Work: The First Century of Unemployment in Massachusetts*. Cambridge University Press New York.

Lindert, P. H. 2000. Three Centuries of Inequality in Britain and America. *Chap. 3, pages 167–210 of: Handbook of Income Distribution*. North-Holland.

Lindert, P. H., & Williamson, J. G. 1982. Revising Britain's Social Tables, 1688-1812. *Explorations in Economic History*, **19(4)**, 385–408.

Lindert, P. H., & Williamson, J. G. 1983. Reinterpreting Britain's Social Tables, 1688-1913. *Explorations in Economic History*, **20(1)**, 94–109.

Mackinnon, Mary. 1996. New Evidence on Canadian Wage Rates, 1900-1930. *The Canadian Journal of Economics*, **29**(1), 114–131.

Masters, D.C. 1947. *The Rise of Toronto, 1850-1890.* Toronto: Macmillan.

McInnis, Marvin. 1968. The Trend of Regional Income Differentials in Canada. *The Canadian Journal of Economics*, **1**(2), 440–470.

Mitchell, B. R. 1993. *International Historical Statistics: The Americas 1750-1988.* New York, Stockton Press.

Osberg, L. 1984. *Economic Inequality in the United States.* New York, M. E. Sharpe, Inc.

Osberg, Lars, & Siddiq, Fazley. 1988. The Inequality of Wealth in Britain's North American Colonies: The Importance of the Relatively Poor. *Review of Income and Wealth*, **34**, 143–163.

Sen, A. 1976. Poverty: an ordinal approach to measurement. *Econometrica*, **44**, 21913.

Sen, Amartya. 1997. *On Economic Inequality.* Oxford, Clarendon Press. Expanded edition with annexe by James Foster and Amartya Sen.

Siddiq, F. K. 1988. The Size Distribution of Probate Wealth Holdings in Nova Scotia in the Late Nineteenth Century. *Acadiensis*, **18**, 136–147.

Soltow, Lee. 1984. Wealth inequality in the United States in 1798 and 1860. *Review of Economics and Statistics*, **66**, 444–451.

Soltow, Lee. 1992. Inequality in the Standard of Living in the United States. *In:* Gallman, R.E., & Wallis, J.J. (eds), *American Economic Growth and Standard of Living before the Civil War.* University of Chicago Press, Chicago.

Sylvester, Kenneth M. 2001. Household Composition and Canada's Rural Capitalism: The Extent of Rural Labor Markets in 1901. *Journal of Family History*, **26**(2), 289–309.

Szekely, M., & Hilgert, M. 1999. *What's Behind the Inequality We Measure?* Inter-American Development Bank, Research Department Working Paper 409.

Theil, H. 1967. *Economics and Information Theory.* Amsterdam, Noth Holland.

Toyoda, T. 1980. Decomposability of inequality measures. *Economic Studies Quarterly*, **31**(12), 207–246.

Urquhart, Malcolm C. 1986. New Estimates of Gross National Product Canada, 1870-1926: Some implications for Canadian Development. *Pages 9–94 of:* Engerman, Stanley L., & Robert E. Gallman, R.E. (eds), *Long Term FActors in American Economic Growth*. University of Chicago Press, Chicago.

Urquhart, Malcolm C., & Buckley, Kenneth A.H. 1965. *Historical Statistics of Canada*. Toronto: Macmillan.

Williamson, J. G. 1980. Earnings Inequality in Nineteenth-Century Britain. *Journal of Economic History*, **40**, 457–476.

Williamson, Jeffrey G. 1965. Regional Inequality and the Process of National Development: a Discription of Patterns. *Economic Development and Cultural Change*, **13**, 3–85.