THE URBAN AND REGIONAL DIMENSIONS OF ECONOMIC INEQUALITY IN CANADA, 1996 – 2006

Kenyon C. Bolton
Department of Geography
McGill University
Montreal, Canada

June 2010

A thesis submitted to McGill University
in partial fulfilment of the requirements
of the degree of Master of Arts

ABSTRACT

There is a general consensus that economic inequality increased in Canada between 1996 and 2006. However, few studies examine the multi-dimensional causes of this trend at the sub-national scale. Chakravorty (1996, 2006) and others (Morrill, 2000; Martin, 2001, Drennan, 2005) argue that an understanding of what influences national-scale inequality requires in-depth consideration of urban and regional socioeconomic processes. Using microdata drawn from the most recent 20-percent samples of the *Canadian Census of Population* (1996, 2001, and 2006), this thesis examines the spatial and socioeconomic dimensions of earnings inequality among individuals in Canada's labour force.

The thesis makes two main contributions to the literature. First, it provides a detailed analysis of the key socioeconomic determinants of earnings inequality across Canadian urban areas using regression analysis. The findings provide new evidence that substantial changes have occurred in the contribution of specific factors to inequality since 1996. Second, spatial data analysis points to changes in the geographic distribution of earnings inequality. Between 1996 and 2006, high levels of inequality across Canada's census divisions have increasingly clustered in Alberta and Newfoundland, and results from spatial regression models shed further light on changes in the nature and structure of earnings inequality.

RÉSUMÉ

Il existe un consensus que l'inégalité économique a augmenté au Canada de 1996 à 2006. Toutefois, peu d'études se sont penchées sur les causes multidimensionnelles de cette tendance à l'échelle sous-nationale. Chakravorty (1996, 2006) et d'autres chercheurs (Morrill, 2000; Martin, 2001, Drennan, 2005) soutiennent que la compréhension de l'inégalité à l'échelle nationale requiert une considération en profondeur des processus socio-économiques au niveau urbain et régional. Cette thèse utilise des micro-données détaillées de l'échantillon 20% des plus récents Recensement du Canada (1996, 2001, et 2006) pour examiner les dimensions spatiales et socioéconomiques de l'inégalité des salaires et traitements chez les travailleurs canadienne.

Cette thèse contribue en deux temps à la littérature préexistante. Premièrement, elle fournit une analyse détaillée des plus importants déterminants socio-économiques de l'inégalité des salaires et traitements à travers les régions urbaines en utilisant une analyse de régression. Les conclusions de cette analyse fournissent de nouvelles preuves que des changements considérables se sont produits dans la contribution de facteurs particuliers à l'inégalité depuis 1996. Deuxièmement, les données spatiales indiquent des transformations importantes dans la distribution géographique de l'inégalité des salaires et traitements. Entre 1996 et 2006, des taux élevés d'inégalité de revenu à travers les divisions de recensement au Canada se sont concentrés en Alberta et à Terre-Neuve. Les résultats de modèles de régression spatiale soulignent les transformations autant de la nature que de la structure de l'inégalité des salaires et traitements.

ACKNOWLEDGEMENTS

It has been a great pleasure to work with my supervisor Sébastien Breau on this project. Throughout the last two years he has provided a fresh and resourceful perspective to my thesis, and I cannot thank him enough for his encouragement, patience, and guidance. I also thank my thesis reader, Nancy Ross, as well as Benjamin Forest and Michel Lapointe of the Graduate Affairs Committee, who have provided helpful feedback and support over the course of my Master's program.

I sincerely thank the departmental staff, especially Maria Marcone, Pauline Nesbitt, Nancy Secondo, Joseph Vacirca, and Paula Kestelman. A graduate student could not ask for a better group of people to help in navigating McGill's administrative labyrinth.

The work of the Quebec Inter-university Center for Social Statistics (QICSS) staff, especially Gaël Lejeune and Élise Comoé of the McGill-Concordia Research Data Centre, was integral to the completion of this project, and I offer them many thanks.

I am grateful for funding from the Warren Fellowship in GIS, the John Bradbury Award, and the QICSS grant program.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1
INTRODUCTION

Economic inequality in Canada has risen substantially in recent years (Frenette, Green, and Milligan, 2009). The OECD's (2008) *Growing Unequal* singles out Canada for having one of the highest rates of growth in inequality among member-nations from the mid-1990s onwards. Looking at longer-term trends, Yalnizyan (2007) argues that the country's income gap is at a 30-year high. This stands in stark contrast to the 1980s when Canada was thought by many to have experienced equitably shared economic growth relative to other highly industrialized economies (Beach and Blackburn 1991; Heisz 2007). The contemporary shift in the distribution of Canadian income is one that increasingly reflects income dynamics in the United States, where individuals in the highest percentiles of the income distribution have claimed a growing proportion of total income since the mid-1970s (Saez and Veall, 2005; Osberg, 2008). This is occurring while individuals in the lowest percentiles are experiencing declines in real earnings (Statistics Canada, 2008).

The observed increase in inequality across many industrialized countries has reinvigorated debates as to its possible causes (Atkinson 1997). Social scientists have linked high levels of economic inequality to outcomes such as higher rates of crime and various other social problems (Hipp, 2007; Wilkinson and Pickett, 2009) as well as disparities in health, a subject in which geographers have made important contributions (see, for instance, Ross, Nobrega, and Dunn 2001). In Canada, much of the work on the sources of inequality has been carried out by economists who have focused on national-level income dynamics (MacPhail, 2000a; Moore and Pacey, 2003; Myles 2003; Frenette et. al, 2007, 2009). Little attention has been paid to sub-national level trends and patterns. Yet, as Chakravorty (1996, 2006) and others (Morrill, 2000; Martin, 2001, Drennan, 2005) have argued, an understanding of what influences national-scale inequality requires in-depth consideration of urban and regional socioeconomic processes.

This project seeks to address part of this gap in the literature with the overall objective of advancing our knowledge of the recent spatial and socioeconomic dynamics of earnings inequality in Canada. Much of the observed increase in inequality has taken place in the distribution of workplace earnings (Saez and Veall, 2005), and so this thesis centres its attention upon the earnings of those in the core labour force. Specifically, its analysis is guided by the following three research questions: *(1) How has earnings inequality evolved in recent years across both urban and rural areas in Canada? (2) Are some areas within Canada becoming more unequal than others? (3) What structural, institutional, and spatial factors explain these changes in inequality?*

To address these questions, I use data drawn from the 20-percent samples of the 1996, 2001, and 2006 long-form *Census of Population.*[1] These large samples are not commonly accessed, and each provides place-of-residence and income data that is more precise than what is publicly available.[2] As such, they provide the opportunity for a uniquely detailed analysis of un-aggregated individual earnings at both the urban and regional scales.

The thesis' findings point to the changing structure and nature of earnings inequality in Canada. First, at the urban scale, results from Chapter 4 suggest that the size of a city grows in prominence over the studied period as a predictor of inequality, as many of the largest cities in the country have experienced the greatest rises in earnings inequality levels. The level of a city's economic development, measured by the median earnings of the labour force, loses its mitigating effect on inequality during the period. This is likely a result of lopsided growth in the earnings distribution, as overall wage gains are not trickling down to those who earn less. With regards to demographic changes, the percentage of women in the workforce emerges as a strong dampener of inequality, and the percentage of city's population composed of visible minorities becomes less influential in predicting earnings inequality between 1996 and 2006. While

[1] Access to these samples was obtained through Statistics Canada's Research Data Centres (RDC) Program, which is described in Section 3.2.
[2] Additionally, each census' public-use microdata file (PUMF) reflects only a three-percent sample.

further research is necessary to determine why in both cases, the former may indicate the effect of wage increases for women in recent years, or alternatively the high representation of men in the fastest growing sectors of the economy. The latter may be evidence of less structural wage discrimination faced by visible minorities in the labour market, or on the other hand it may only suggest other factors are becoming more dominant. And, as expected, higher rates of workers employed in manufacturing dependably keep wages more equally distributed.

Second, at the regional scale, findings in Chapter 5 suggest that there have been similarly important changes in the geographic manifestation of regional earnings inequality over the 1996 to 2006 period. Large areas in Western Canada, as well as the Greater Toronto Area and portions of Newfoundland show statistically significant clusters of high rates of growth in earnings inequality between 1996 and 2006. There is also evidence of a growing urban—rural divergence in levels of earnings inequality as growth in urban Gini coefficients tends to outpace that in rural areas. In line with results in Chapter 4, certain socioeconomic determinants of earnings inequality have changed in their influence across all regions in Canada. The level of dispersion in educational attainment emerges as a contributor to earnings inequality by 2006, and the rate of manufacturing employment increases its role as a dampener of inequality. Higher rates of unemployment and higher proportions of senior citizens to workers both consistently contribute to higher inequality. Finally, the population shares of visible minorities and Aboriginal peoples are less certain in predicting inequality once the model controls for urbanization in the spatial error model. The results from the spatial error model prove to be more reliable in this application; however, its comparison with both OLS and spatial lag regression highlights possible confounding variables within the models.

The remainder of the thesis is organized as follows. Chapter 2 provides the theoretical and empirical foundations for approaching the research objectives and developing a study on inequality. Chapter 3 details the nature of the data used and provides the economic context in which changes in the earnings distribution have taken place. Chapters 4 and 5 form the bulk of the thesis, and comprise the analytical studies; Chapter 4 examines inequality and its determinants across urban areas using cross-sectional and panel data regression methods, and Chapter 5 undertakes a broader regional analysis using spatial autocorrelation mapping and spatial regression techniques. Chapter 6 concludes with a summary of the thesis' main findings and discusses possible policy implications.

CHAPTER 2:

LITERATURE REVIEW

2.1 INTRODUCTION

This section begins with a review of the literature on how one might problematize inequality, and highlights some of the debates that have shaped approaches to the study of inequality in geography and other disciplines. Evidence of the renewed interest in inequality within industrialized countries is then presented, and a case is made for its study within the Canadian context. A discussion of the empirical foundations of the models estimated concludes the chapter.

2.2 WHY STUDY INEQUALITY?

Concerns of fairness and social justice usually motivate studies of economic inequality. Within geography, two scholars have been particularly influential in how it may be problematized. First, in *Social justice and the city*, Harvey (1973) transforms the conceptualization of justice beyond the liberal, Rawlsian notion of distributional equity. In the Marxist tradition, he criticizes the latter for omitting the role of production in creating and sustaining inequalities. Rather than simply a feature of liberal markets, he argues that "production *is* distribution and that efficiency *is* equity in distribution" (p. 15). This argument positions the study of modern industrial organization at the centre of social relations and wellbeing, and in effect the study of economic inequalities. In that sense, Bluestone and Harrison's *The deindustrialization of America* (1982) might be considered as a landmark study of late 20th-century changes in economic and social production that favours capital versus society, and the inequalities that ensue.

Second, Smith (1994) builds on Harvey's ideas to argue that distributional outcomes are a necessary focal point for study, especially to catalyze social action. He presents the *justice as equalization* concept, in which any shift towards greater equality implies greater social justice. The concept both avoids an absolute or static definition of social justice and allows for empirical investigation of the causes and outcomes of income inequality.[3] In other words, if anything practical is to be done about inequalities, we have to look at the quantifiable distribution of resources among specific social groups, and how it changes. This argument has a place in the ongoing debate regarding the use of quantitative methods in critical approaches to geographic questions.[4] Ellis (2009), one of the contributors to this debate, points out: “Numbers and the methods we use to estimate them shape how clearly and extensively we can see injustice and inequality; I cannot imagine how human geography can be critical if it does not embrace this fact” (p. 308).

Myles (2003), a sociologist, echoes Ellis’s (2009) sentiment in making the following assessment regarding the study of Canadian economic inequality in his field: “With a few notable exceptions...we have left most of the heavy lifting—both theoretical and empirical—on this topic to the economists” (p. 551). Notwithstanding sociology’s long history of studying inequality, he argues that the field has been engaged in disciplinary boundary battles with literature and the humanities regarding interpretations of (post)structural changes in society, and in doing so lost its footing in practical approaches to real world problems. Economists, on the other hand, have diversified their methods to include consideration of a variety of social factors. Gone, he argues, is the ability of sociologists to pinpoint drivers of economic inequality in Canada. As discussed later in this chapter, geographers have advanced similar arguments (see Dorling and Shaw, 2002).

It follows from the above that one way to problematize inequality is to see evidence of injustice in the distribution of income, and to date there is growing inquiry in the empirical literature regarding the consequences of economic inequality. Geographers have contributed significantly to this literature by examining the relationship between inequality and population health, as well as segregation outcomes and processes. For instance, at the urban scale, Ross et al.

[3] See Lake (1994), a special issue of *Urban geography*, for an in depth treatment of social justice in geography, including the roles of Harvey and Smith in its conceptual development.
[4] See issues 3 and 4 of *The Professional Geographer, vol. 61 (2009)* for a discussion on the use of quantitative methods in critical geography.

(2004) as well as Walks and Maaranen (2008) link income inequality with spatial polarization of population groups within Canadian cities, and greater spatial polarization within U.S. cities has been shown to negatively affect health outcomes (Ross et al., 2001). Similarly, epidemiologists Wilkinson and Pickett (2009) provide a study on the repercussions of high levels of inequality, as well as an example of a quantitative study that incorporates social and scalar contexts.

Figure 2.1: Correlation between health and social problems and income inequality, Wilkinson and Pickett (2009)

At the international scale, Wilkinson and Pickett (2009) find a strong correlation between health and social problems and income inequality (Figure 2.1), yet they are careful in arguing causation, as the magnitude of inequality does not directly lead to social outcomes.[5] Rather, they assert that "It is much more plausible that [inequality] works through all the processes of social status stratification that have been central to the social sciences for so long, including the ways in which so many marks of social position become imprinted on us from early childhood onward" (p. 509).

[5] The authors construct the health and social problem index from 10 measures including infant mortality rates, public trust, homicides, life expectancy, and social mobility.

At the intra-national scale the authors apply a similar approach among U.S. states and counties, and highlight the importance of scale in problematizing inequality. While higher inequality predicts higher levels of social dysfunction at large scales, they are less reliably (statistically) associated at the other end of the spectrum. The authors posit that "perhaps what we are seeing is less of a social comparison between neighbours than of the effects of the extent of social class differentiation in society as a whole" (p. 503). This thought complements the approaches of Smith (1994) and Ellis (2009) in bridging the gap between studying societal scale numerical outcomes and maintaining a contextual awareness of individual experiences. Inequality is a relevant predictor of social dysfunction at societal scales, and yet its causal elements may extend down to the micro-level.

In sum, we can problematize inequality when the modes of production tend to benefit some greatly, in both the nature of the workplace and the magnitude of gains, but not others. The case is strengthened when the latter not only sees little to no improvement in gains over years of steady work, but also is put at increased risk of tangible health and social consequences. In the next sections I discuss how the nature of economic growth and inequality has changed, and why it is a problem to be addressed.

2.3 The current interest in economic inequality

The return to interest in income inequality is largely associated with its accelerating growth in highly industrialized economies (Atkinson, 1997; Martin, 2001; OECD 2008). Discussions of inequality have intensified in academic and policy circles as well as in the public realm following the most recent economic crisis. Even Alan Greenspan, the former head of the U.S. Federal Reserve, acknowledges that if left unchecked, growing economic inequalities "may eventually threaten the stability of democratic capitalism itself" (Grier, 2005).

2.3.1 The "new inequality"

In 1955 Simon Kuznets theorized that industrialization and its resultant economic growth would directly affect income gaps in two stages of rising and then diminishing inequality. First, the transition of workers from the agricultural sector to industrial sectors would widen income gaps within the labour force. The second stage would be characterized by a ‘catching-up’ of the low-wage sector, effectively diminishing income inequality as industrialization broadened (Conceição and Galbraith, 2001, p. 148). This explanation of the Kuznets curve holds for much of the growth experienced by industrialized countries after the Second World War, when the share of total income represented by those in the top 1% of income-earners among OECD countries decreased steadily until the late 1970s, and median earnings grew across the income spectrum (OECD, 2008). However, as Martin (2001) articulates, a “new inequality” has emerged in recent decades (p. 268).

This “new inequality” is driven by a range of structural and institutional shifts within industrialized economies. For instance, Saez and Veall (2005) show that in the past 30 years the top earners in the United States are accumulating a rapidly growing share of the country’s total annual income, with the highest percentile (99^{th}) of wage-earners accounting for 12.5% of total wages in 2000, up from 5% in the late 1970s. Meanwhile, researchers such as Levy and Temin (2007) find that those in the middle and lower echelons of the U.S. income distribution have seen their incomes stagnate, and this is the recent case in Canada as well (Statistics Canada, 2008). While Canada at first lagged in the growth of economic inequalities relative to the U.S., its earnings distribution is beginning to mirror that of its southern neighbour (Saez and Veall, 2005; OECD 2008).

2.3.2 Recent research on inequality in Canada, and the contribution of this thesis

In a recent paper, Lars Osberg (2008, p. 3), a leading figure in the field of economic inequality in Canada, reflects on his early work in the early 1980s: “Back in those days, an oft-repeated gibe in academia was that the study of economic inequality was boring.” In the years of post-war growth in developed countries, there was little interest in income inequality due to its negligible growth in industrialized countries, as mentioned above. Canada persisted in maintaining such a trend into the late 1970s and 1980s, while the U.S. and others saw their levels begin to increase (Heisz, 2007; OECD, 2008). Evidence shows that the country's redistribution mechanisms were able to offset pre-tax and pre-transfer growth in income inequality until 1990 (Frenette, Green, and Milligan, 2007 and 2009).

The increase in Canadian inequality levels since 1990 has sparked new interest in the topic, as evidenced by a wealth of recent research (Myles, et. al. 2000; Picot and Myles, 2005; Heisz, 2007; Statistics Canada, 2008). Yalnizyan (2007) and Osberg (2008) at the Canadian Centre for Policy Alternatives have become public figures in the discussion, bringing their findings to the media, academia, and government alike.[6] The emergent role of Alberta in provincial and national levels of earnings inequality (to be discussed later in this thesis) is also a topic of research by the Parkland Institute (Gibson, 2007), and the Centre for the Study of Living Standards has published on inequality and productivity in the labour force (Sharpe, Arsenault, and Harrison, 2008). While there is an emergent and energetic consensus that the trends of growing inequality are irrefutable, the bulk of the above research remains mainly descriptive, and few studies look at sub-national processes and determinants of inequality.

Myles (2003) pairs his argument that sociology has lost ground in studying Canadian inequality with a list of key academic articles published by economists on the subject. His sentiment could be shared by geographers, as there are few studies on the determinants of inequality at the sub-national scale in Canada with a geographic perspective. In 1997, MacLachlan and Sawada emphasized a need for such research at the inter-urban and regional scales, and still only few responded (Moore and Pacey, 2003; Breau, 2007; Breau and Rigby, 2009;). Meanwhile economists such as Erksoy (1994), Soroka (1999), MacPhail (2000a), Fortin and Schirle (2006), Lu, Morissette, and Schirle (forthcoming) and others have analyzed various determinants of inequality across Canada. Frenette

[6] See the 2009 Walter Gordon Massey Symposium: *Rising inequality in Canada: A problem for public policy:* http://www.publicpolicy.utoronto.ca/NewsandEvents/Lists/Upcoming%20Events/DispForm.aspx?ID=64; or their website: http://www.policyalternatives.ca/.

et al. (2007, 2009) in particular have steadily pursued this subject, as it applies to after-tax family income inequality. Yet, there remains a significant gap in the research when it comes to up-to-date analyses specific to employment earnings in the labour force, and applications to urban and regional levels. As mentioned in Section 2.2, to explain societal-level inequalities requires approaches that are tuned to finer scales, where micro-socioeconomic factors shape the nature of the labour market income distribution.

It is clear that inequality in Canada, especially in recent years, has grown. In this thesis I address the gap in research with various models that analyze determinants of inequality in Canada at the urban and regional scales. Additionally, spatial modelling and autocorrelation measures are used to address geographic concerns of regional spatial context and change over time. And finally, as described in Chapter 3, the datasets of individual-scale microdata used in this analysis are not commonly accessed, and provide more detail and coverage than data such as the Survey of Labour and Income Dynamics or publicly available census microdata (based upon 3% samples), which are commonly used by researchers. The following section outlines the relevant theoretical foundation I use in building these models.

2.4 MODEL FOUNDATIONS

This section provides the conceptual basis for constructing and analyzing models of the socioeconomic determinants of inequality. The independent variables studied in Chapters 4 and 5 are derived from the body of theory presented below. It is important to note that there is no overarching or general theory of earnings distribution (Osberg, 1981). As mentioned above, a multidimensional approach is necessary when studying inequality; micro-scale factors must be conceptualized within macro-scale frameworks (MacPhail 2000a). For instance, while human capital theory seeks to explain the determinants of individual earnings, the distribution of earnings within a population requires consideration of broader socioeconomic and institutional processes. As such, the following discussion indentifies some of the key drivers of earnings inequality, and is categorized into two subsections. The first looks at labour market-specific factors such as labour force skill and demographic makeup, and the second looks at other lines of theory including institutional changes and geographic context.

2.4.1 Labour market demand and supply factors

Changes in both demand and supply factors with respect to labour markets have been found to affect income inequality, and some approaches have focused upon human capital explanations.

On the demand side, skill-biased technological change is seen as driving demand for higher educated workers, even in a time of their increasing supply, especially in the U.S and Canada (Berman, Bound and Machin 1998; Breau, 2007; Levy and Temin, 2007). However, this approach has been characterized as weak by Morris and Western (1999) as it is difficult to proxy the skill-level of workers by simply using their level of educational attainment. Instead, they attribute the growth of the college premium in recent years to the collapse of the high-school premium (p. 633). Drennan (2005) finds that human capital gains in the United States between 1979 and 1999 played a role in spatial divergence of income among metropolitan areas, a reversal from 1969 to 1979. Rigby and Breau (2008) argue that the effect of skill-biased technical change played a role in raising inequality in Southern California during the 1990s, but that its effect was eclipsed in later years by the impact of international trade. Likewise, the declining rate of manufacturing within industrialized countries is associated with the replacement of high-paying, unionized jobs with lower-paying service-sector jobs characterized by higher turnover, adding to the overall effects of structural unemployment (Massey, 1979; Bluestone and Harrison, 1982; Peck, 1992), and evidence in Canada points to the impact of these inter-related factors upon inequality (Soroka, 1999; MacPhail, 2000a; Breau 2007). The decline in manufacturing in industrialized countries has been attributed in part to the rise of foreign competition (Bluestone and Harrison, 1982), and there is a well-developed literature on the effects of ‘globalization’ and trade upon inequality (Rigby and Breau 2008). Breau and Rigby (2009) have also shown that these effects vary across regions, depending on their level of exposure to foreign competition. In this study, I will proxy such effects by focusing upon the rate of manufacturing employment.[7]

On the supply side, changes to the demographic composition of the labour force are also seen as potential sources of increased inequality. Much attention has been given to how immigration may be linked to rising inequality, especially in the US where it has significantly changed the composition of the labour force over the last few decades of the 20th century. Between 1970 and 1996, immigrants increased as a proportion of the population from 4.6% to 9.6%, with much of the change reflected in large urban areas (Morris and Western, 1999). As immigrants generally enter the workforce at lower wages and lower levels of human capital in the U.S., earnings of the native-born may face downward pressures as a result of competition (Borjas et al. 1997; Morrill 2000). In the Canadian context, policy shifts over the past 20 years have greatly increased the influx of immigrants. Picot, Hou, and Coulombe (2007) show that those who have recently immigrated have faced a declining share of market wages in recent years, regardless of the fact that they are on average more highly educated than the native-born workforce. This is reflected by Moore and Pacey's (2003) findings that immigration plays a significant role in the growth of national income inequality, with recent immigrants accounting for 46% of the rise in inequality between 1990 and 1995. They also note that as an increasing share of Canada's visible minorities is represented by immigrants, the two categories will become 'confounding variables' within models, with both contributing to inequality.

Changes in the distribution of human capital among individuals also affect the composition of the workforce. While overall increases in educational attainment among workers can have a dampening effect on inequality, as was the case between 1950 and 1990 in the United States, the current wage premium paid to those with university degrees has reversed this correlation (Morrill, 2000). A

[7] MacPhail (2000a) uses manufacturing employment rates to proxy trade, however she does so by isolating those manufacturing industries that are more likely to engage in international trade. The lack of detailed information on trade patterns at the sub-provincial scale limits the capacity to model these effects.

measure of dispersion between those without a high-school education and those with a Bachelor's degree or higher captures this effect, and is expected to be positively correlated with inequality. Chakravorty (1996) and Cloutier (1999) confirm this link between educational attainment and inequality in the United States, and Breau (2007) finds evidence of it in Canada.

Finally the evidence is mixed with regards to the increasing participation rate of women in the workforce. Fortin and Schirle (2007) document the rising presence and relative wages of women in Canada's workforce, citing that during the 1990s women's real earnings increased by 10% while those of men exhibited a decline. They find that between 1982 and 1997 increasing female participation reduced income inequality among families in Canada, and Breau (2007) finds similar results at the inter-provincial level. These findings may reflect evidence from the U.S. that substantial wage gains for women in recent decades greatly reduced inequality among all workers (Kopczuk, Saez, and Song, 2010). However, MacPhail (2000a, citing Richardson 1997) cites evidence from Canada that increasing income inequality is related to a higher a female participation rate, and MacLachlan and Sawada (1997) postulate that women with high-earning partners could be more likely to enter the workforce at higher income brackets and for that reason add to inequality among households. In any case, the role of women in the workforce may indeed be a significant one in determining inequality.

2.4.2 Other structural, institutional, and spatial factors

In addition to the labour market structural changes described above, broader changes in the economies of industrialized countries will also influence income inequality patterns. These can be seen as the result of macroeconomic structural changes, as well as changes in institutional factors.

Unemployment is prominent in the literature as a significant factor that predicts higher inequality in both the United States and Canada (Chakravorty, 1996; MacLachlan and Sawada, 1997; MacPhail, 2000a; Breau 2007). Cyclical unemployment, which reflects the ups and downs of the business cycle, is expected to increase inequality; high unemployment implies weak labour markets in which wages (especially for those in the low-wage sector) are likely to be lower due to excess supply of labour (Chakravorty 1996; MacPhail 2000a). Structural unemployment, which refers to a condition in which workers lack the skills to find employment, do not live where jobs are available, or are unwilling to work for the wage rate available on the market (Osberg and Lin, 2007), may be linked to processes of deindustrialization as discussed above.

The decline in unionized labour also parallels deindustrialization, and is considered to contribute to higher levels of inequality (MacPhail 2000a). Researchers often connect this decline to wider neoliberal and institutional change, such as financial deregulation and increasingly free markets (Galbraith, 2002; Peck and Tickell, 2002; Saez and Veall, 2005; Harvey, 2006; Levy and Temin, 2007). While unionization rates are not directly measured in the census, their magnitude may be approximated to levels of manufacturing employment, and thus the latter accounts for some institutional change.[8]

Geography, of course, also influences the above factors. City size and urbanization have been argued to affect the level of income inequality (Chakravorty, 1996; Korpi, 2008). Larger cities are expected to have a more diversified industrial mix, as well as a higher level of economic development (approximated by the level of median wages and income) (Cloutier, 1997). A debate lies in whether the population size of the city acts as a cause, or at least captures residual causes, of economic inequality after accounting for the industrial mix and development level of the city. Betz (1972) finds that after controlling for the percentage of the population that is non-white in U.S. cities (where increasing city size is closely correlated with the number of non-white individuals living there), as well as industrial mix, city size is not a relevant factor in predicting

[8] While not investigated in this study's analysis, other institutional factors have been attributed to higher levels of inequality in Canada. According to Osberg (2008, p. 30), the increase in total income inequality after 1990 is attributable to the combination of increased interest rates in the early 1990s and a federal deficit crisis, to which the government responded with a rapid reduction of income redistribution. In addition to cutbacks for various government programs, Sharpe, Arsenault, and Harrison (2008) link rising inequality to the Bank of Canada and Finance Canada's policy shift targeting a two percent inflation rate, which, they argue, lead to a prolonged period of stagnant growth from which Canadian workers never fully recovered.

inequality. On the other hand, Chakravorty (1996) argues that it does play a role, and Korpi's (2008) study on urban areas in Sweden finds that labour market size is significantly associated with higher inequality, especially for inequality driven by the top of the income distribution. Within cities, the spatial effect of income inequality is perhaps more immediately provocative, in that the distribution of income is extremely uneven within a relatively small space, and with visible manifestations (Smith, 1994). For instance, Doussard, Peck, and Theodore (2010) take a close look at Chicago's emergent inequalities in its industrial restructuring since the 1980s. In a study that combines aspects of urban analysis, deindustrialization, and economic growth, they make clear that urban labour markets are going through considerable changes which impact their earnings distributions. These areas may well be the nexus of the transformation in the structure of inequality, and MacLachlan and Sawada (1997) highlight the need for more inter- and intra-urban analyses of income inequality within Canada.

In defining causal socioeconomic structural factors at work in driving inequality, MacLachlan and Sawada (1997) also highlight the difficulty in deciphering the contributions of variables, and their roles at various scales. While researchers have addressed this difficulty with the application of models at the inter-metropolitan scale outside of Canada (see Chakravorty, 1996; Cloutier 1997; Morrill, 2000; McCall, 2001; Korpi 2008), similar work in the country has been limited, leaving opportunity for research.

2.5 Conclusion

This chapter presents the theoretical and empirical literature relevant to a study of the determinants of inequality. It begins by focusing on the problematization of inequality, which includes both an incorporation of social justice arguments and an emphasis on the consequences of rising inequalities. The concept of a changing structure of inequality, or a "new" inequality, is then introduced (something to be investigated in greater length in Chapters 4 and 5), and a gap in inequality research within Canada is identified. The last section presents theory that identifies the drivers of inequality which are included in this study's models.

Before diving into the analyses per se, Chapter 3 provides a general discussion of the data used in this thesis, and an overview of general trends of inequality in Canada.

CHAPTER 3

DATA, METHODS, AND DESCRIPTIVE STATISTICS

3.1 INTRODUCTION

One of this study's unique aspects is the nature of its data. Few studies of inequality use the long-form Canadian census datasets described in this chapter (for other examples see Pendakur and Pendakur, 2007; Breau and Rigby, 2009; Frenette et. al, 2009), yet, as detailed in section 3.2, it provides certain advantages from an empirical perspective. Furthermore, this study includes the 2006 census, which provides for an up-to-date analysis.

This chapter begins by discussing the study's data and the steps undertaken to gain access to it, including the necessary ethical considerations. Second, it describes the parameterization of the labour force sample that serves as the foundation for the analytical aspects of the thesis. A discussion of methodological issues regarding the measurement of inequality among individuals follows, and the chapter concludes with a brief look at general trends in inequality at both the urban and national scales.

3.2 DATA ACCESS AND ETHICS

This study's analyses are based on the 20-percent long-form samples of the 1996, 2001, and 2006 versions of the Canadian *Census of Population*. In constructing each sample, Statistics Canada distributes the long form questionnaire (2B) to approximately every fifth dwelling in the country. It relies upon a calibration estimation methodology to weight each individual surveyed so that the sample accurately reflects the full population. Each sample contains roughly six million individual responses with information on demographic, social, and economic characteristics, and it is estimated that the national response rate of the 2006 census was 96.5%.[9]

---

[9] For more detail on how Statistics Canada constructs its census data, see: http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3901&lang=en&db=imdb&adm=8&dis=2.

There are a number of advantages to using this microdata in the study of economic inequality. First, each census sample contains a high number of observations. This helps to reduce sampling error, and increases each sample's representativeness of the population. Second, as Frenette et al. (2007, 2009) point out, compared to other commonly-used datasets, the Census long-form sample has ample coverage. For instance, the Survey of Labour and Income Dynamics (SLID) has a 20% under-response rate relative to the census (in addition to being a much smaller sample). Third, incomes and earnings are not top-coded. MacPhail (2000b) and Moore and Pacey (2003) highlight the limitations of top-coded data and smaller datasets, as their use generally underestimates levels of inequality. And finally, particularly important from a geographical perspective, the long-form sample contains place-of-residence information available down to the census tract level. This is crucial for constructing indices of inequality at the metropolitan and regional scales that are consistent over time.

Statistics Canada deems such microdata to be confidential due to its level of detail. To protect the identity of census respondents the data is housed in a secure facility (i.e. locally at the McGill-Concordia Research Data Centre – RDC), which is itself operated by the Quebec Inter-university Centre for Social Statistics. Gaining access to the RDC requires (1) project approval and acceptance by the *Canadian Initiative for Social Statistics- Access to the RDC Program*, a joint initiative between Statistics Canada, the Social Sciences and Humanities Research Council, and the Canadian Institutes of Health Research, (2) a Government of Canada Security Clearance, and (3) a contract with Statistics Canada in order to become a 'deemed employee' and subject to its research guidelines. Finally, the use of secondary data requires approval by the McGill Research Ethics Board. The work carried out in this thesis conforms to all of the above considerations, and all results presented have been screened by an RDC Statistical Analyst to ensure that no confidential information is released.

3.3 Data development

Chapters 4 and 5 analyze the determinants of inequality at different geographic scales within Canada. Both approaches aggregate subsamples of individual census respondents by geographic area, and the subsamples are parameterized in order to reflect the core labour force for each census. The targeted sample comprises individuals in the active labour force ages 25 to 64, who reported $1,000 or more in both annual earnings and total income for the previous calendar year.[10] 'Earnings', as defined in the Census, reflect only an individual's salary and/or wages including self-employment earnings, and differs from 'market income' in that the latter includes other non-government sources such as investment income. In contrast, 'total income' comprises all sources of income including federal government transfers such as Employment Insurance.[11]

These distinctions are especially important due to the fact that there may be inconsistencies between the 2006 census and earlier ones regarding total income. Frenette et al. (2009) note that in the 2006 census, respondents were each given the option to allow Statistics Canada to link his or her census survey to government tax data rather than filling out the income section in the survey. The authors find that while reported total income may differ from what is documented in tax data, market income inequality measures are consistent between datasets in the previous census. Thus, while individuals may under-report government transfer income (Kapsalis, 2001, cited by Frenette et al., 2009) they may be more likely to accurately report employment income. As such, while analyses of total income are presented in this thesis, the focus is placed mainly on earnings inequality. Earnings are the primary source of income for most people, and measures of inequality based on individual earnings data better reflects labour market processes and outcomes (Moore and Pacey, 2003; Frenette et. al, 2007, 2009).

[10] SAS 9.x is used for the majority of the data management and variable construction tasks.
[11] I use the term 'earnings' throughout this thesis for what is termed 'wages' in the census. See the Census Dictionary for definitions of these terms: http://www12.statcan.ca/census-recensement/2006/ref/dict/azindex-eng.cfm.

Imposing a minimum earnings and total income threshold effectively removes those individuals not actively engaged in the labour force and controls for self-employed individuals who might have netted losses in total income. The chosen age range also captures those more closely attached to the labour force, as often individuals younger than 25 have yet to enter steady employment (Korpi 2008). The sample is weighted by individual person weights provided by Statistics Canada in order to reflect the characteristics of the wider population.

In sum, the variables in this study are derived from a weighted sample of individuals who represent the core active labour force in each census period. In Chapters 4 and 5 I describe in further detail the mechanics of the chosen models, as well as the construction of the spatial units (i.e. geographic areas) used in the analyses.

3.4 MEASURING INEQUALITY

The Gini coefficient acts as the measure of inequality in both analytic chapters. It is commonly found in the literature due to its ease of application and interpretation. It is derived from the Lorenz curve (MacLaughlin and Sawada, 1997), and Cowell (1995, p. 23) generally defines it as "the average difference of all possible pairs of incomes in the population, expressed as a proportion of total income." More formally, it is specified as:

$$Gini = \frac{1}{2n^2\overline{y}}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|y_i - y_j\right|$$

where n is the number of individuals and y is the income of individual i or j. The Gini's possible values lie between 0 and 1, with 0 representing 'perfect' equality and 1 representing 'perfect' inequality (i.e. one individual in the sample receives all the income and the others none).

One of the Gini's major limitations is its sensitivity to income differences at different points in the distribution (Cowell, 1995, p. 23). A transfer of income from a richer to poorer person can have a greater effect on the Gini's value if the two individuals are in the middle rather than either tail of the distribution. Certain entropy indices such as the Theil may be manipulated to emphasize different parts of the distribution. In robustness-check comparisons I found the Gini to maintain sufficient rank-ordering of inequality levels among the studied geographic areas, and thus keep it as my chosen indicator.[12] For each geographic unit studied in Chapters 4 and 5, Gini coefficients of earnings and total income are generated using *ineqdeco.ado* in Stata. This program was initially developed by Jenkins (2001) and provides multiple measures of inequality.

3.5 General inequality trends in Canada, 1996 - 2006

The following section provides an overview of national-level earnings inequality trends from 1996 to 2006. In order to gain a general understanding of a labour force's earnings distribution, it is common to divide it into ranked quintiles. Each contains 20% of the overall number of observations. The bottom quintile contains respondents who earn the least; the 20% - 40% quintile contains those with the next lowest earnings, and so on. Real median earnings over time are calculated for each, providing clear insight into tangible economic trends for different earnings brackets. The 40% - 60% quintile represents the middle earnings bracket, and its median level of earnings also reflects the median level of all observations. Table 3.1 below provides median earnings by quintile for all individuals in the core labour force pooled across the 87 urban areas studied in Chapter 4.[13]

Table 3.1: Median real earnings for the targeted labour force across 87 urban areas in Canada (2002 dollars), by earnings quintile

	1996 Census Median Earnings	2001 Census Median Earnings	2006 Census Median Earnings	1996 – 2001 % Change	2001 – 2006 % Change	1996 – 2006 % Change
Top 20%	67,500.00	71,610.00	75,838.65	6.1%	5.9%	12.4%
60% - 80%	45,000.00	47,058.00	48,865.10	4.6%	3.8%	8.6%
40% - 60%	33,187.50	34,403.49	34,872.59	3.7%	1.4%	5.1%
20% - 40%	22,072.50	23,529.00	23,192.76	6.6%	-1.4%	5.1%
Bottom 20%	8,673.75	10,230.00	9,534.97	17.9%	-6.8%	9.9%
Total of weighted observations	7,884,350	8,701,220	9,297,640	-	-	-

Source: Author's tabulation of the 1996, 2001, and 2006 censuses, with the CPI used as a deflator.
Earnings reported in each census correspond to the total calendar-year earnings of the previous year.
The quintiles are created using Osiro's (2007) Stata module *quantiles.ado*.

[12] See Appendix 3.1 for a correlation table between the Gini coefficient's values for earnings across census divisions in Canada (Chapter 5's datasets for 1996, 2001, and 2006) and eight other common inequality indices constructed in Stata using *ineqdeco.ado* (Jenkins, 2001).
[13] Here I focus on Chapter 4's urban-scale dataset, which represents over 80% of the Canadian labour force. See Appendix 3.3 for these methods applied to Chapter 5's regional-scale dataset.

Between 1996 and 2006, both the top and bottom quintiles of the labour force in Canada reported significant gains in real earnings. The 10-year gain of 9.9% in median earnings for the lowest quintile is substantial as that quintile seems to have benefited from Canada's overall GDP growth in the late 1990s (see Cross 2007a, for trends in GDP). However, these figures may be misleading outside of wider temporal contexts; Yalnizyan (2007) documents similar gains by the poorest working Canadians in the late 1990s, but notes that those constant-dollar figures stand *lower* than they were in the 1970s. Picot and Myles (2005) show that during the late 1990s Canadian families overall saw reductions in unemployment, increased earnings, and a drop in numbers below the low-income cut-off. The authors cite the possible influence of the Child Tax Benefit, which reduced employment disincentives; however, they highlight the plight of those far below the low-income cut-off line, many of whom saw major losses in real income during this period yet may not be included in this study's sample. Furthermore, Sharpe et al.'s (2008) argument that workers never recovered from the fall in labour's share of GDP in the early nineties strengthens Yalnizyan's conclusion that no long-term gains were made in this period. And lastly, the 1990s wage gains seen in the bottom quintile of the core labour force may reflect an "asymmetrical polarization" of job growth, which is recorded by Doussard et al. (2010) in the U.S. In other words, the gains may be due to a rapid growth in numbers of high- and low-paid jobs in the service industry, while middle class job growth is stagnant.

Nevertheless, Table 3.1 shows that the bottom 40% of urban workers saw real losses in median earnings between the 2001 and 2006 censuses, and those in the middle quintile saw their earnings stagnate. This occurred as those in the top quintile continued to make substantial gains. These figures can be further compared to the Statistics Canada (2008) report on earnings trends since 1980. While that study focuses on a narrower slice of Canadian workers (only those working full-time for more than 49 weeks in the year), it also shows impressive gains in median real earnings among the top quintile of earners (6.2%), little growth in real earnings for the middle quintile (2.4%), and losses in real earnings for the bottom quintile (-3.1%) between the 2001 and 2006 censuses.

The share of earnings commanded by each quintile provides additional evidence of growing inequality during this period, and is shown in Table 3.2:

Table 3.2: % Share of total earnings by earnings quintile, for the core labour force

	1996 Census Total Earnings % Share	2001 Census Total Earnings % Share	2006 Census Total Earnings % Share
Top 20%	41.79	43.29	45.80
60% - 80%	24.60	20.92	22.73
40% - 60%	17.52	19.57	16.37
20% - 40%	11.59	11.35	10.76
Bottom 20%	4.51	4.87	4.34
Total	100%	100%	100%

Source: Author's tabulation, earnings are reported for the previous calendar year.
Created with Jenkins' (2006) Stata module *sumdist.ado*.

Divergence in the distribution of overall earnings held by each quintile is apparent here, as all but the top quintile declined in their relative shares between the 1996 and 2006 censuses. Note, however, that the middle and lower quintiles increased their shares of overall earnings in 2001, before falling in 2006. Again, it is clear that earnings were distributed quite differently during the first period (1996 to 2001), relative to the second period (2001 – 2006). I investigate this difference further in the following analytic chapters.

Finally, inequality growth was not uniform across industries. Table 3.3 compares Gini values across industry aggregates during the studied period, and they are listed from highest percentage change in the Gini between 1996 and 2006 to the lowest. [14]

[14] The industrial classifications in this table are based on the SIC 1980 Divisions for 1996 and 2001, and the NAICS 2002 aggregates for 2006. As such, comparing the 2006 inequality values to the other two may not be exact (the most significant difference is that the real estate sector is included with Finance and Insurance in the SIC 1980 aggregates, but is separate in NAICS); however the chosen aggregates are those that show a general consistency in categorization over time.
Also, it should be noted that these values are based upon annual earnings, and not per-hour earnings. Thus, while they accurately reflect inequality by industry, they may not adequately capture inequalities in hours worked, i.e. workers in manufacturing may not be putting in as many hours as some do in the mining and oil sector. See Williams (2007) for more information on how the industry has changed in terms of hours worked, hourly-pay, and overall economic growth.

Table 3.3: Gini values by industry aggregate and percentage change over time

Industry	*Gini of earnings*			*% Change*		
	1996	2001	2006	96 – 01	01 – 06	96 – 06
Mining and oil	0.315	0.341	0.425	8.3%	24.6%	34.9%
Finance and Insurance (and real estate in 2006)	0.393	0.442	0.473	12.5%	7.0%	20.4%
Wholesale trade	0.372	0.384	0.418	3.2%	8.9%	12.4%
Retail trade	0.395	0.400	0.430	1.3%	7.5%	8.9%
Construction	0.374	0.359	0.403	-4.0%	12.3%	7.8%
Manufacturing	0.344	0.353	0.365	2.6%	3.4%	6.1%
Education	0.330	0.328	0.341	-0.6%	4.0%	3.3%
Health care	0.350	0.359	0.360	2.6%	0.3%	2.9%

Based upon the author's tabulation of the 1996, 2001, and 2006 one-in-five sample of the Canadian Census.

As seen in the above table, the aggregated mining and oil industries show the largest increase in earnings inequality across all other industry aggregates, although the finance, insurance, and real estate aggregate exhibit the highest absolute level of earnings inequality in 2006. As expected, service-based industries are characterized by both high absolute levels of inequality and rates of change, and workers in the public sector are more equitably paid. Lastly, manufacturing's generally low levels of inequality are exhibited here, as well as the low rate of change in its Gini coefficient.

3.6 CONCLUSION

This chapter provides an overview of the relevant data and the steps taken to access it. Furthermore, the chapter presents general inequality trends among individuals in the core labour force. It is clear that over a relatively short period, 1996 – 2006, the structure and shape of the earnings distribution has changed, and the following chapters investigate this change further by measuring the contribution of chosen socioeconomic factors to earnings inequality.

CHAPTER 4

CANADIAN ECONOMIC INEQUALITY IN THE URBAN LABOUR FORCE

4.1 INTRODUCTION

This chapter examines the determinants of economic inequality across 87 cities in Canada, for the years 1996, 2001, and 2006. It investigates the influence of key structural, institutional, and geographical predictors of earnings inequality as they are identified in the academic literature. Cross-sectional regression methods are used to highlight changes in their influence on inequality at various temporal points, and panel data regression methods analyze their influences during the total period studied. Its main findings point to major changes that have taken place in the composition of Canadian earnings inequality.

The chapter is organized as follows. First I define the urban areas studied, and present the steps taken to establish consistent geographical boundaries between census cycles. Second, I provide summary statistics of the labour force across the relevant urban areas, and highlight how levels of urban inequality vary by city size. I then define the dependent and independent variables, as well as the model specifications for the cross-sectional and panel data models. A presentation and discussion of the results concludes the chapter in the context of the relevant literature presented in Chapter 2.

4.2 RESEARCH DESIGN AND METHODS

One of the key challenges researchers face when working with data drawn from various census cycles at the urban and regional scales of analysis is maintaining consistency in geographical boundary definitions. Census data are disseminated for a variety of geographical areas, the boundaries of which change frequently over time to reflect annexations, municipal amalgamations or simple name changes (Statistics Canada, 2007). This section discusses the criteria used to build the study sample, and the steps taken in establishing a consistent urban geography among census cycles. The cross-sectional and panel model specifications to be estimated are also presented.

4.2.1 Dataset construction: the geography

Statistics Canada defines cities as either census metropolitan areas (CMAs) or census agglomerations (CAs). According to the 2006 definition, each CMA has a minimum population of 100,000 which includes an urban core population of at least 50,000, and each CA has an urban core population of at least 10,000. Statistics Canada uses adjacent census subdivisions (CSDs) that surround an urban core as building blocks to delineate CMAs and CAs, including only those CSDs which have populations highly integrated with their respective urban core.[15] However, as an urban area's influence expands geographically, Statistics Canada adds CSDs from that urban area each new census. This may be due to population changes, changes in commuting patterns, or administrative changes at the provincial scale.[16] With respect to the latter, unlike most of Canada's census geographical units, CSDs are subject to municipal boundary changes made by provincial and territorial legislations. These conditions present the difficulty of creating consistent urban areas over multiple censuses, especially when including smaller cities.[17] Essentially, the boundaries of CMAs and CAs from one census may cut across their respective CSDs in another census, rendering a comparison of multiple cross-sectional datasets potentially inaccurate.

To address this difficulty, I use ArcMAP 9.3 (a Geographic Information System) to rebuild the 1996 and 2001 cities according to their 2006 CMA/CA boundaries, using 1996 and 2001 CSDs as building blocks. For the 1996 and 2001 censuses, a list of CSDs is generated. These spatial units are categorized according to their corresponding 2006 CMA/CAs boundaries, and then hardcoded into a SAS program that aggregates the individual respondents in each census by

[15] See CMA detailed definition provided by StatCan: http://www12.statcan.ca/english/census06/reference/dictionary/geo009a.cfm. A CMA surrounds an urban core population of 50,000, and a CA surrounds an urban core population of 10,000.
[16] In order to maintain some historical continuity, CSDs are normally retained in the CMA or CA even if commuting patterns shift elsewhere.
[17] Census divisions (CDs) are often used for their stability over time in studies of larger cities (Vinodrai, 2001).

2006 CMA/CA. The SAS program for the 2006 census relies upon CMA/CA boundaries rather than CSD ones. The result is a list of cities for which boundaries remain constant through the 1996, 2001, and 2006 censuses. Individual survey respondents are then grouped by census cycle and city. Those cities in which major boundary conflicts exist between earlier CSDs and their respective 2006 CMA/CAs are excluded from the study, as described in more detail below.

In order to be included in the analysis a city had to satisfy three conditions: (i) it had to be categorized as a CMA or CA for all three census years; (ii) its total population in 2006 had to be greater than 20,000; and (iii) its constituent 1996 and 2001 CSDs had to correspond with 2006 CMA/CA boundaries. The first and second conditions remove 40 of 144 combined CMAs and CAs in the 2006 census and the third removes an additional 17, leaving 87 observations in total. The first and second conditions were imposed to examine the distinctly and dependably urbanized areas in Canada. The third condition represents a limitation of the data described above, and mainly affects cities in British Columbia, where 11 urban areas ranging in total populations from 22,000 to 91,000 in 2006 experienced major changes to their CSD boundaries after 1996 and are thus excluded.

In sum, 87 cities at three temporal points (1996, 2001, and 2006) act as the observations in this study's analyses. Each observation contains corresponding socioeconomic characteristics aggregated from the census' 20-percent long-form samples. The parameterization of these individual-level samples was described in Chapter 3. Table 4.1 in the following sections provides a provincial breakdown of the cities included, along with selected characteristics of the labour force.

4.2.2 General characteristics of the dataset

As shown in Table 4.1 below, average urban earnings inequality has increased across urban areas in Canada, but at different rates within each province.[18] Cities

[18] This tabulation reflects summary statistics of the 87 observations, and thus is not derived directly from the microdata. This is notable for its level-effect upon the inequality figures and their corresponding percentage changes. For instance, if the national Gini were computed for

within the Maritime Provinces have seen relatively slow growth in their average levels of inequality, and cities in Quebec have experienced only 0.3% average change between 1996 and 2006. Meanwhile, cities in British Columbia, Alberta, Ontario and Newfoundland have seen substantial increases in average inequality. In 2006, Alberta cities have, on average, the highest levels of inequality, and Quebec the lowest. These general provincial comparisons are in keeping with Breau (2007).

Table 4.1: Provincial breakdown of CMAs and CAs included in the study, and selected characteristics.

Province	Number of included cities	Min/max total populations in 2006	Average of urban median earnings in 2006	Average Gini of earnings in 2006	% Change in average Gini of earnings 1996 - 2006
Newfoundland	2	26,250/102,760	31,760	0.397	5.3%
P.E.I.	1	57,720	30,986	0.381	3.0%
Nova Scotia	5	25,800/369,460	29,834	0.389	1.8%
New Brunswick	6	21,060/124,060	30,641	0.386	0.5%
Quebec	22	23,980/3,588,520	33,007	0.354	0.3%
Ontario	28	22,710/5,072,070	38,370	0.377	5.0%
Manitoba	2	47,460/686,040	34,094	0.369	3.7%
Saskatchewan	4	32,230/230,850	34,943	0.372	1.4%
Alberta	9	22,410/1,070,300	43,035	0.418	5.8%
British Columbia	8	22,730/2,097,960	36,211	0.397	6.7%
National	87	Mean: 273,511	35,827	0.379	3.6%

Source: Author’s tabulation

Table 4.2 contains summary statistics of the top and bottom five cities of the studied 87, ranked in order of their 2006 levels of earnings inequality. Excluding Toronto, those cities with the highest levels of inequality are found in the western provinces. The Calgary CMA stands out with a Gini in earnings of 0.491 in 2006, and also with a 19% increase in inequality since 1996 (the largest increase in inequality of all 87 metropolitan areas). All of the top five have above-

individuals in the *target sample* **pooled** across the 87 cities, it would be greater in magnitude and percentage change. However, the above figures accurately reflect relative inequality values and their changes across cities by province.

average levels of median earnings, while the bottom five have low levels of median earnings.

Table 4.2: The top and bottom five CMAs and CAs, ranked in order of their 2006 Ginis

The top and bottom five CMAs and CAs, ranked in order of their 2006 levels of income inequality.		2006 Total population	2006 Labour Force Population	2006 nominal median earnings	2006 Gini in Earnings	1996 – 2006 % Change in the Gini of earnings
1	Calgary, AB	1,070,300	464,980	42,000	0.491	19.1%
2	Toronto, ON	5,072,070	1,961,110	40,227	0.445	14.7%
3	Vancouver, BC	2,097,060	807,630	37,460	0.425	9.2%
4	Grande Prairie, AB	71,450	29,880	44,967	0.425	9.9%
5	Brooks, AB	22,410	8,560	38,332	0.418	5.0%
83	Saint Jean sur Richelieu, QC	86,080	33,950	34,640	0.341	0.7%
84	Rivière du Loup, QC	23,980	9,590	30,956	0.334	-3.8%
85	Saint Georges, QC	30,970	12,510	31,005	0.334	-2.6%
86	Thetford Mines, QC	25,410	9,300	29,150	0.331	-5.9%
87	Sainte Hyacinthe, QC	54,160	21,310	31,641	0.330	-0.1%

Source: Author's tabulation of the 2006 and 1996 censuses

The bottom five cities are all located in Quebec, which highlights the provincial variation in inequality levels. Also notable is the presence of Canada's largest cities, excluding Montreal, in the top five (Montreal ranks 22nd in 2006). Researchers such as Korpi (2008) argue that labour market size (i.e. city size) is positively correlated with economic inequality, and this is apparent here.

The positive city size-inequality relationship is also observed in Table 4.3 which shows that larger cities have, on average, higher levels of earnings inequality, especially in recent years. Furthermore, there is divergence in inequality levels between larger and smaller cities during the period from 1996 to 2006, with larger cities showing rapid growth in inequality levels. While large cities (greater than 500,000 in total population) show an increase of 9.1% in average levels of the Gini, small cities (less than 75,000) show an increase of only 1.9%. Filion (2009) discusses a similar pattern of divergence, but one in economic growth between large and small urban areas in Canada. He foresees two possible scenarios. In the first a 'hinterland revival' occurs, in which the current boom in commodities results in revived economic growth in smaller cities. Alternatively, Canada's few densely-populated urban centres might experience relatively rapid economic growth, resulting in increased polarization between large and small urban areas. He argues that the latter is more plausible due to the advantages of increasing returns found in large cities, and the reliance of the growing service sector on those processes. As evidenced here, Filion's argument regarding an expanding service sector, which is also linked to levels of inequality as discussed in Chapter 2, holds when applied to the divergence of inequality levels.

Table 4.3: Average earnings inequality by census year and city size, and percentage change between 1996 and 2006.[19]

City Population Size	1996 Average Gini of Earnings (# Cities)	2001 Average Gini of Earnings (# Cities)	2006 Average Gini of Earnings (# Cities)	% Change in Avg. Gini of Earnings 1996 – 2006
Large Greater than 500,000 Total Pop.	.375 (9)	.382 (9)	.409 (9)	9.1%
Medium Large 200,000 – 499,999 Total Pop.	.360 (8)	.364 (8)	.383 (8)	6.4%
Medium Small 75,000 – 199,999 Total Pop.	.366 (23)	.365 (23)	.378 (25)	3.3%
Small Less than 75,000 Total Pop.	.365 (47)	.360 (47)	.372 (45)	1.9%

Source: Author's tabulation of the 1996, 2001, and 2006 Censuses.

The above section describes general characteristics of the selected 87 urban areas. The next two sections discuss the nuts and bolts of this chapter's modelling strategy. I specify the cross-sectional and panel data regression models in which the urban areas act as observations, and outline the urban-level variables as well as the theory behind their inclusion.

4.2.3 The cross-sectional model

The OLS regression model applied to each census cycle is specified as:

$$\begin{aligned} INEQ_{it} = \alpha + \quad & \beta_1 LF\_POP_{it} + \beta_2 ECON\_DEV_{it} + \beta_3 UNEMP_{it} + \\ & \beta_4 FEM\_PART_{it} + \beta_5 MANUF_{it} + \beta_6 VIS\_MIN_{it} + \\ & \beta_7 EDUC\_RATIO_{it} + \beta_8 YOUNG\_LF_{it} + \\ & \beta_9 SENIOR\_LF_{it} + PROV_i + \varepsilon_{it} \, . \end{aligned} \tag{4.1}$$

Here, the dependent variable [$INEQ_{it}$] represents economic inequality among individuals in city *i* at census cycle *t*, as measured by the Gini coefficient.

[19] A similar table was constructed for total income Gini averages, with nearly identical rank ordering and percentage changes.

The scalar (α) and error term (ε_{it}) carry the normal assumptions, and the dummy variable vector $PROV_i$ controls for provincial fixed effects (e.g. broad institutional differences across provinces). The potential for heteroskedasticity in the model is addressed with the use of robust standard errors (developed by White, 1980, as cited in Wooldridge, 2006). Heteroskedasticity-robust standard errors can be expressed as:

$$\sqrt{Var(\hat{\beta}_j)} = \frac{\sum_{i=1}^{n} \hat{r}_{ij}^2 \hat{u}_i^2}{SSR_j^2}$$

where $\hat{r}_{ij}$ denotes the i^{th} residual from regressing x_j on all other independent variables, and SSR_j is the sum of squared residuals from this regression (Wooldbridge 2006, p. 274). This procedure does not change the standard OLS coefficient estimates; it only affects the calculation of standard errors (and thus *t*-values).

Equation (4.1) is estimated for two income concepts: one with the dependent variable $INEQ_{it}$ as the Gini of earnings [$GINIEARNINGS_{it}$], and the other with $INEQ_{it}$ as the Gini of total income [$GINITOTINC_{it}$]. A comparison of the two model runs is useful to check the model for robustness and also in understanding the possible effects of government assistance and transfer programs upon inequality.

The independent variables specified in Equation (4.1) are selected in keeping with the relevant literature regarding the determinants of economic inequality presented in Chapter 2. LF_POP_{it} represents the natural log of the number of weighted observations in the labour force. As expected, the size of the labour force is highly correlated with the size of a city's total population (r = 0.999* across all observations and census cycles), and thus points to the role of the total urban population size in predicting inequality. As the labour force and total population size increases, it is expected that inequality will rise in keeping with Table 4.3 and Korpi (2008). $ECON\_DEV_{it}$ represents overall economic development, and is approximated by the median earnings of city *i* at census cycle *t*.[20] This variable is standardized to 2002 dollars using the Consumer Price Index as a deflator (CANSIM table 326-0021).[21] A negative relationship between inequality and a city's level of economic development is expected (Chakravorty 1996).

The remaining variables are all represented as percentages. $UNEMP_{it}$ is the unemployment rate in each city's labour force, and FEM_PART_{it} is the female participation rate. While higher cyclical unemployment is a tested predictor of higher inequality (Johnson, 1995), the effect of increases in numbers of employed women is ambiguous in the literature. As discussed in Chapter 2, women have steadily increased their presence in the labour force in recent decades (Luffman, 2006); however, the evidence of their impact upon inequality is mixed. For instance, MacPhail (2000a) finds that increased levels of female participation increase inequality, while Breau (2007) finds the opposite. Their presence in the labour force has been a dynamic one in recent years in terms of wages and industries, and this variable should capture their current effect. $MANUF_{it}$ represents the percentage of those employed in manufacturing industries.[22] This variable is intended to capture the role of deindustrialization and industrial mix in Canadian cities, as well as the institutional presence of unions which tends to dampen levels of inequality (Soroka 1999; MacPhail 2000a; Breau 2007). VIS_MIN_{it} represents a combined percentage of visible minorities, recent immigrants, and Aboriginal people. Statistics Canada defines visible minorities as respondents who self-report as non-Caucasian, excluding Aboriginal people. While wage discrimination is not equally experienced across minority groups, including Aboriginal people, overall minority access to well-paying jobs is generally limited, impacting their representation in the income distribution

[20] Median income is a better indicator of overall economic development than average income as it is more representative of income levels at the bottom of the scale.
[21] I standardize the *ECON_DEV* variable in order to simplify a comparison of its coefficient across census cycles.
[22]This variable is coded according to 1980 Standard Industrial Classification codes for 1996 and 2001, and 2002 NAICS codes in 2006. This is due to the lack of a continuous coding across the 20% samples; however, as I keep the category deliberately broad, it can be assumed that they are very comparable in this case. Moreover, numbers provided by CANSIM table 282-0012 regarding manufacturing employment relative to overall employment mirrors the relative levels provided in Table 4.4, which describes the variables in this dataset.

(Pendakur and Pendakur 2007). Recent immigrants are those who have settled in Canada within five years prior to responding to the census. Due to shifts in immigration source-countries in the last 15 years, recent immigrants are largely represented by visible minorities, and also may face increased barriers to full employment (Picot and Myles, 2005). Thus, the two segments of the population are correlated. Furthermore the vast majority of recent immigrants live in the major cities, leading to substantial collinearity with the labour force size variable. Because of these conditions I do not include a recent immigrant variable, and instead rely upon the overall visible minority population to capture the effect of recent immigrants upon urban inequality. Furthermore, Aboriginal respondents are included in VIS_MIN_{it} due to their relatively large presence in some western cities, and their associated spatial concentration and income gap relative to non-Aboriginal people (Maxim et al. 2001; Ross et al. 2004). It is expected that VIS_MIN_{it} is positively associated with economic inequality. $EDUC\_RATIO_{it}$ is intended to capture the educational attainment structure of the labour force in each city and census cycle. It represents the ratio of those without a high school degree plus those with a Bachelor's degree or higher to the remaining sample. As the ratio grows, the spread between those with high and low levels of educational attainment increases, which is expected to be positively related to inequality (Chakravorty, 1996; Cloutier, 1997).

Finally, the age dependency variables $YOUNG\_LF_{it}$ and $SENIOR\_LF_{it}$ are the only exceptions to the labour force parameters; they represent the ratios of children (ages 14 or less) and seniors (ages 65 or more) to the labour force in each city. The effect of pensions and government support have been shown by Milligan (2008) to lower poverty rates among seniors, which leads to expectations that overall income inequality across all ages would be negatively correlated with higher percentages of seniors. However, the variables $YOUNG\_LF_{it}$ and $SENIOR\_LF_{it}$ are included to account for the external effects of these populations upon economic inequality experienced by the bulk of the labour force.[23] For

[23] The regression $GINIEARNINGS_{it}$ on $YOUNG\_LF_{it}$ and $SENIOR\_LF_{it}$, results in positive and significant (a = 0.05) coefficients for both variables in 1996 and 2006, after controlling for

example, the unpaid care of either children or the elderly may take from an individual's work activities, possibly lowering the individual's annual earnings. Therefore each is expected to be positively associated with inequality. Table 4.4 below shows the mean values of each variable, by census cycle.

Table 4.4 Mean values for each variable across 87 observations, by census cycle. Standard deviations are in parentheses.

Variable	1996	2001	2006
GINIEARNINGS	0.366 (0.020)	0.364 (0.022)	0.379 (0.027)
GINITOTINC	0.335 (0.020)	0.337 (0.023)	0.346 (0.032)
LF_POP	90,624 (229,809)	100,014 (259,529)	106,869 (275,695)
ECON_DEV[A]	30,929 (3,828)	32,061 (3,854)	32,839 (5,045)
UNEMP	0.060 (0.020)	0.050 (0.023)	0.043 (0.018)
FEM_PART	0.457 (0.021)	0.468 (0.022)	0.479 (0.024)
PER_MANUF	0.166 (0.087)	0.167 (0.088)	0.146 (0.077)
PER_VIS_MIN	0.051 (0.052)	0.062 (0.063)	0.082 (0.077)
EDUC_RATIO	0.387 (0.039)	0.373 (0.044)	0.349 (0.045)
HIGH_EDUC[B]	0.160 (0.050)	0.173 (0.055)	0.189 (0.064)
NO_HIGH[B]	0.227 (0.045)	0.200 (0.046)	0.161 (0.041)
YOUNG_LF	0.580 (0.065)	0.510 (0.058)	0.455 (0.053)
SENIOR_LF	0.341 (0.106)	0.355 (0.109)	0.375 (0.107)

[A] *ECON_DEV* is represented by median real earnings of the *target population*, in 2002 dollars.
[B] These variables are combined in the *EDUC_RATIO* variable, and thus are not individually added to the model. They are provided for descriptive purposes only.

LF_POP_{it} and $PROV_{it}$. In 2001 $SENIOR\_LF_{it}$ loses its significance, which is also seen in this chapter's results (Section 4.3.1). As expected, their effect upon $GINITOTINC_{it}$ is less pronounced as government transfers such as the Child Tax Benefit are included in total income.

4.2.4 The panel data model

Panel data is longitudinal in that it follows the same set of observations through time. While the *Census of Population* does not follow individual respondents as do long-term longitudinal surveys such as the *Survey of Income and Labour Dynamics* and the *Longitudinal Survey of Immigrants to Canada*, census data can be aggregated by geographic areas that remain constant throughout time. As detailed above in Equation 4.1, cross-sectional analysis is useful in comparing the determinants of income inequality across the 87 cities at different points in time (e.g. 1996, 2001, and 2006). However, multiple regression analysis can be applied to a panel dataset to test the strength of associations between the dependent and independent variables across each census cycle. In other words, panel data analysis lets us see which independent variables play major roles in predicting income inequality over the course of the 1996 to 2006 period, and accounts for the fact that the observations are not independently distributed across time, or space in some cases.

According to Hsaio (2003), panel data analysis also improves the efficiency of model estimates because it generally uses a large number of observations, increasing the degrees of freedom in the model and reducing collinearity among independent variables. A comparison of cross-section and panel data analyses offers insight as to both the overall determinants across the period of interest (in the case of panel data), as well as the changing role of each variable at points within the period (in the case of cross-sectional data).

There are several ways to approach panel data, the most straightforward being a pooled-OLS model that includes all observations from all time cycles. In this case dummy variables that control for each time-cycle (minus one for a reference) are included to account for differing distributions among cycles. While pooled-OLS can be useful for general purposes, it neglects to address that observations are not independently distributed through time. For instance, constant unobserved factors at play in Montreal which may affect its level of income inequality in 1996 will likely again affect its level in 2006 (e.g. municipal policies). The serial correlation of these unobserved factors over time is not considered in a pooled-OLS model, and thus possibly violates basic assumptions that the model's errors (ε_{it}) are not correlated with themselves or the model's independent variables. This condition would introduce heterogeneity bias into the model.

Random effects (RE) and fixed effects (FE) panel data models offer alternative approaches that account for the presence of unobserved effects in the model. Recall that in a typical linear regression equation, $y = \alpha + x\beta + e$, the error term e captures variance in the dependent variable y that is unexplained by the scalar α plus the independent variable parameter estimates ($x\beta$). In RE and FE models, the error term is composite, consisting of the unobserved effect and the idiosyncratic error. Thus, a general form panel model can be specified as $y_{it} = \alpha_i + x_{it}\beta + v_{it}$, where the composite error $v_{it} = \mu_i + u_{it}$. Here, μ_i represents the time-constant unobserved effect for the ith observation and u_{it} represents the 'idiosyncratic error'.

The choice between the use of a FE and RE model depends upon the nature of the unobserved effect. If μ_i is assumed to be correlated with the independent variables, the FE model is the optimal choice. However, if μ_i is assumed to be unrelated to the independent variables, a RE model will also give efficient estimates without dropping time-constant variables. The following is an explanation informed by Wooldridge (2006).

The FE model time-demeans the data, or in other words, it subtracts each variable's mean over all time periods from that variable's ith observation at time t, including the composite error term and dependent variable. This effectively removes variables that do not change with time such as geographic characteristics, cancelling out the 'fixed-effects'. Since the unobserved effect μ_i is time-constant, it also is dropped, leaving only the mean-centred idiosyncratic error. The resulting model then resembles standard OLS regression.

Alternatively, the RE model only quasi-demeans the data. A fraction (λ) of a variable's mean over time is subtracted from that variable's ith observation at time t. The fraction's value is dependent upon the number of time periods, and the variances of the unobserved effect μ_i and the idiosyncratic error u_{it}. It is calculated as follows: $\lambda = [\sigma_u^2 / (\sigma_u^2 + T\sigma_\mu^2)]^{1/2}$, where T is the number of time-periods. This allows the model to keep variables that remain constant, including the unobserved effect μ_i which in this case is assumed to be uncorrelated with the independent variables.

The decision to use either the RE or FE model is based upon both theoretical reasoning and how well each model fits the data. In this study's case of geographic fixed effects, it is reasonable to expect that the unobserved effects are correlated with the independent variables, thus making the FE model more attractive. The Hausman (1978) specification test can be used to check the consistency of estimators in the RE model, or in other words whether the RE estimates would converge in probability with parameter values. This procedure tests the null hypothesis that a RE model offers consistent estimators in comparison to the FE model.

Stata 11 includes commands for heteroskedasticity-robust FE and RE modelling as well as the Hausman test. In order to construct the panel dataset, the three cross-sectional samples of 87 observations (cities) each are combined into one matrix. Each observation is marked with an identifying integer corresponding to its census cycle (e.g. 1, 2, or 3 for 1996, 2001, and 2006 respectively), which accounts for the time-series nature of the data. In this case, the panel is balanced as each cycle contains an identical collection of observations (although their characteristics vary between cycles). The panel data model specification for both the FE and RE model regressions is similar to Equation (4.1), and is expressed as:

$$\begin{aligned} INEQ_{it} = {} & \alpha + \beta_1 LF\_POP_{it} + \beta_2 ECON\_DEV_{it} + \beta_3 UNEMP_{it} + \\ & \beta_4 FEM\_PART_{it} + \beta_5 MANUF_{it} + \beta_6 VIS\_MIN_{it} + \beta_7 EDUC\_RATIO_{it} \\ & + \beta_8 YOUNG\_LF_{it} + \beta_9 SENIOR\_LF_{it} + PROV_i + \mu_i + u_{it}. \end{aligned} \tag{4.2}$$

All variable descriptions correspond with those given for Equation (4.1); however, note that in place of the error term (ε_{it}), the uncorrelated effect (μ_i) and idiosyncratic error (u_{it}) are introduced.[24]

[24] I standardize median earnings and median total income to 2002 dollars, as in Equation (4.1). This is particularly important here, due to the differencing in panel data regression.

In sum, Equation (4.2) is applied with both FE and RE methodologies to the data described in Table 4.4, and the Hausman test is used to determine the applicability of the RE model.

4.3 RESULTS AND DISCUSSION

This section begins with a discussion of the results from the cross-sectional analyses, followed by the panel data regression results. While the bulk of the discussion focuses on the models with the Gini of earnings as the dependent variable, I also touch upon model results that have the Gini of total income as the dependent variable.

The findings show that substantial differences exist in the influence of certain variables upon income inequality among the urban labour force between census cycles. Effectively, it appears that the *nature* and *structure* of income inequality within urban labour forces has changed over the 1996 to 2006 period. This is likely due to macroeconomic changes in the business cycle, and more importantly structural changes.

4.3.1 The cross-sectional regression results

Table 4.5 presents the findings from the cross-sectional regression models with the Gini of earnings ($GINI\_EARNINGS_{it}$) as the dependent variable. Table 4.6 contains the regression results that pertain to the cross-sectional models with the Gini of total income ($GINI\_TOTINC_{it}$) as the dependent variable.

There is substantial change in the influence of certain variables between 1996 and 2006. It is important to note that provincial fixed-effects are controlled for in each model. As shown in the descriptive section above, large differences in inequality exist among provinces, and their inclusion in the models as dummy variables captures possible differences in terms of political and economic structures. Throughout this section I focus on the earnings inequality model results in Table 4.5, unless otherwise noted. Table 4.6 is provided for robustness and comparison reasons.

Table 4.5: Cross-section regression results for earnings inequality in1996, 2001, and 2006

Independent variables	Gini (earnings)		
	1996	2001	2006
LF_POP	0.003 (0.108)	0.007** (0.000)	0.009** (0.000)
ECON_DEV	-1.59e-06** (0.029)	-1.90e-06** (0.015)	-9.94e-07 (0.119)
UNEMP	0.104 (0.307)	0.211* (0.092)	0.269** (0.017)
FEM_PART	-0.142 (0.192)	-0.349** (0.011)	-0.428** (0.000)
MANUF	-0.043** (0.042)	-0.042* (0.082)	-0.060** (0.04)
VIS_MIN	0.160** (0.000)	0.092** (0.011)	0.077* (0.057)
EDUC_RATIO	-0.105** (0.039)	0.016 (0.797)	0.038 (0.616)
YOUNG_LF	0.038 (0.221)	-0.000 (0.998)	0.028 (0.556)
SENIOR_LF	0.039** (0.039)	0.019 (0.458)	0.064** (0.013)
CONSTANT	0.445** (0.000)	0.493** (0.000)	0.459** (0.000)
R^2	0.773	0.707	0.778
PROV. DUMMIES	YES	YES	YES
N	87	87	87

One star (*) and two stars (**) indicate coefficients significant at the 0.10 and 0.05 levels, respectively. All models are heteroskedasticity-robust.
p-values are in parentheses.

Table 4.6: Cross-section regression results for total income inequality in 1996, 2001, 2006

Independent variables	Gini (total income)		
	1996 GINI	2001 GINI	2006 GINI
LF_POP	0.004** (0.033)	0.009** (0.000)	0.012** (0.000)
ECON_DEV	-1.23e-06 (0.101)	-1.28e-06 (0.113)	-7.45e-07 (0.313)
UNEMP	-0.124 (0.248)	0.094 (0.481)	-0.055 (0.641)
FEM_PART	-0.138 (0.209)	-0.376** (0.005)	-0.491** (0.000)
MANUF	-0.044* (0.057)	-0.034 (0.173)	-0.045 (0.127)
VIS_MIN	0.164** (0.000)	0.083** (0.023)	0.076* (0.080)
EDUC_RATIO	-0.091* (0.093)	0.028 (0.644)	0.069 (0.427)
YOUNG_LF	0.035 (0.230)	-0.025 (0.528)	0.022 (0.655)
SENIOR_LF	0.028 (0.124)	0.014 (0.542)	0.060* (0.060)
CONSTANT	0.408** (0.000)	0.463** (0.000)	0.431** (0.000)
R^2	0.781	0.752	0.810
PROV. DUMMIES	YES	YES	YES
N	87	87	87

One star (*) and two stars (**) indicate coefficients significant at the 0.10 and 0.05 levels, respectively. All models are heteroskedasticity-robust.
p-values are in parentheses.

City size, proxied by *LF_POP*, emerges in importance as a predictor of inequality over time for those in the targeted labour force. In 1996 it is not significant (p = 0.108), and by 2006 it is strongly significant. As it is a logged variable, it can be interpreted as a percentage change: a one-percent increase in a city's labour force population in 2006 corresponds to a 0.009 increase in the Gini holding other variables constant. In 2001, a one-percent increase in a city's labour force population corresponds with a 0.007 increase in the Gini, and in 1996 its impact is only 0.003 and not statistically significant. This increase in impact over time points to the growing role of large cities in generating earnings inequality, and inequality as an urban phenomenon. This provides further support of the findings of Soroka (1999), who cites the growth in urban service-sector employment as a driver of inequality for both males and female workers. In addition to containing the managerial offices of national businesses, larger cities are more connected into global financial markets and capital, a condition which greatly benefits the upper levels of the income spectrum, and holds less certain benefits for the lower levels (Skaburskis and Moos, 2008; Wyly 2004). This analysis especially applies to cities in which stock market investment centres are located (e.g. Toronto), or where energy-extraction companies are headquartered (e.g. Calgary), or finally where energy-resources distribution centres are located (e.g. Edmonton) (Cross, 2007b).

The level of economic development as measured by median earnings (or total income on the right side of the results table) also undergoes a notable change in significance. While in 1996 and 2001 it played a significant role in mitigating inequality, in 2006 it is not significant (p = 0.119). Note that the coefficients are extremely small due to the difference in units (dollars versus the Gini). This change in significance might be interpreted, in accordance with the literature, as a shift from the equalizing effect of overall higher incomes to the opposite. This may occur due to "the growth of low-wage tertiary service activities, market failure, and size-related disamenity compensation demands," as summarized by Chakravorty (1996). Additionally, as seen in Tables 3.1 and 3.2, the incomes of the top percentiles have grown at a higher rate than other percentiles, effectively raising median levels of income yet also increasing inequality and thus changing its empirical role.

Unemployment shows substantial change in its influence between 1996 and 2001; however, this result may be more reflective of changes in the business cycle than structural changes. The years 1996 and late 2001-2002 saw low GDP growth rates (Cross, 2007), and cyclical unemployment likely played less of a pronounced role in raising inequality in the labour force during those periods. However, by 2006, unemployment becomes a significant predictor of inequality, with a 1 percentage-point rise in the unemployment rate corresponding with a 0.269 point rise in the Gini of earnings. As some workers saw rises in their earnings during the boom of the mid-2000s, higher rates of urban unemployment translated into downward pressure on wages among those already in the lower earnings quintiles. As such, unemployment may represent the role of lower-paid individuals on overall inequality as they lose more income in times of higher unemployment, similar to findings in Canada during the 1980s by Erksoy (1994). On the other hand, the government effect of Employment Insurance can be seen as a mitigating factor in total income inequality, as Table 4.6 shows no effect of unemployment in any year on total income inequality, which is supported by Johnson (1995).

The participation rate of women in urban labour forces has emerged as a dampener of earnings inequality since 1996. This development breaks from research presented by MacPhail (2000a), which suggests that the female participation rate is positively associated with overall income inequality in the 1980s, largely due to increased inequality among female workers. Soroka (1999) also finds that the effect of service sector employment is greater on female inequality than male, although he presents evidence that female earnings distributions are beginning to reflect male earnings distributions. Evidence presented here from the 2000s, however, suggests that women have an equalizing effect on the overall earnings distribution of individuals, which aligns with findings in Breau (2007). In the 2006 model, a 1 percentage point increase in the female participation rate corresponds with a 0.428 point decrease in the Gini. This finding is likely not so much a result of a rapidly increasing share of women in the workforce since 1996 (the average among the 87 cities has only changed from 45.7% to 47.9%, as seen in Table 4.4, although Picot and Myles (2005, p. 17) note that heads of female lone-parent families saw a major increase in employment rates), but rather evidence of the rapidly increasing level of their earnings relative to men, which corroborates the argument put forth by Fortin and Schirle (2007), as well as findings from the U.S. (Kopczuk, Saez, and Song, 2010). On the other hand, this variable's inverse interpretation raises questions regarding the role of the male participation rate on urban inequality. For instance, of the studied urban areas, those in Alberta have the highest average rate of male representation in the urban labour force (55.5% in 2006), as well as the highest level of average inequality (as seen in Table 4.1). In 2006 men represented 72% of those employed in the country's booming oil and gas industry (Williams, 2007), and here emerges the possibility of economic inequality that is significantly driven by gender-specific industry norms, warranting further study.[25] In sum, this variable's emergence represents an economic structural change as opposed to a reflection of the business cycle, as it is clearly significant in both 2001 and 2006, respectively slow- and high-growth periods.

As expected, higher numbers of workers employed in manufacturing also have a mitigating effect upon earnings inequality within Canadian cities. There is not much change in this variable between 1996 and 2006, although in 2001 its coefficient is only marginally significant. This is testament to the role of manufacturing in maintaining an equalizing effect upon earnings distributions, as blue-collar pay in those industries is generally higher, and there are higher rates of unionization. The deindustrialization literature such as Harrison and Bluestone (1982), as well as the trade competition literature (see Breau and Rigby 2009), points to this effect of manufacturing employment. Declining rates of manufacturing employment in Canada's economy due to a transition to a service-based economy as well as increasing imports of durable goods thus translates to

[25] However, the recent recession has affected earnings of male workers more than female workers due to heavy losses in the manufacturing industries. Again, further study is required.

increasing levels of overall inequality as once well-paid workers integrate the more variably-paid service sector. It is plausible that some former manufacturing workers might take jobs in the booming energy extraction energy, thus further creating unequal regional and overall income distributions.

It is expected and confirmed that the rate of urban visible minority workers (including Aboriginal people) is positively correlated with earnings and total income inequalities. Visible minorities continue to face various forms of discrimination in the labour market, stemming from spatial, racial, or immigration pressures, and thus are less likely to equitably reap economic benefits. This in turn raises levels of inequality. However, what is notable here is this variable's diminishing role over time. In 1996, a one-percentage point increase in the number of visible minorities represented in the labour force translates into a 0.160 point increase in the Gini and is strongly significant. By 2006, that coefficient is only 0.077, and its significance is marginal ($p = 0.057$). This evidence points to at least two possible processes: the focus of Canadian immigration policies upon the recruitment of higher-educated and thus higher-earning immigrants (although immigrant earnings are not climbing in recent years as policymakers had expected; see Picot, Hou, and Coulombe 2007), and the good-news of a fundamental shift away from structural discrimination faced by visible minority native-born and immigrants. On the other hand, its diminishing significance (and its lower contribution to explained variance in 2006) may only suggest that factors other than cultural association are relatively growing in the determination of urban inequality. Further research is necessary to tease out these possibilities, but their presence nevertheless points to a changing structure of inequality.

The remaining three variables, *EDUC_RATIO*, *YOUNG_LF,* and *SENIOR_LF* are less obvious in their effects upon urban inequality. However, the number of a city's seniors (*SENIOR_LF*) relative to the labour force does appear to increase earnings inequality, with an inexplicable loss of significance in 2001. This could be due to the demands of caring for elderly family members, which may take away from an individual's hours at work. Furthermore, there is a negative correlation between median earnings and the percentage of seniors relative to the labour force (r = -0.389* in 1996 and r = -0.577* in 2006). This is evidence that cities with younger working populations are more likely to have greater levels of median earnings, which as discussed earlier, can dampen inequality (up to a point). But, perhaps this correlation points to other exogenous forces, such as increased economic vigour or even greater 'animal spirits' of a younger city, where earnings are more widely shared.

In sum, major changes in the composition of variables associated with labour force economic inequality have occurred between 1996 and 2006. Variables such as city size, economic development, the unemployment rate, the rate of female participation in the labour force, and the percentage of visible minorities have each changed in significance and magnitude over the course of the decade. These changes point to structural changes in the distribution of earnings, and thus the structure of earnings inequality itself. The following section looks at which variables have played major roles during the entire decade.

4.3.2 The panel data regression results

Table 4.7 presents the results from the FE panel data regressions, as well as those from pooled-OLS models which include provincial fixed effects. The left two columns describe the earnings inequality models, and the right two columns apply to total income inequality. Recall that panel data methods allow for the combination of cross-sectional data from each census year, and account for the time-series nature of the data. In the case of the FE, time-constant unobserved heterogeneity is subtracted from the model.

As described in the methods section, both RE and FE models were estimated. However, the Hausman test rejected the null hypothesis that the RE model's coefficients are consistent, and thus confirmed that the FE model is both the most efficient and most consistent model. As such, the RE results are omitted.

Table 4.7 Fixed-Effects Panel Data regression results with either $GINIWAGES_{it}$ or $GINITOTINC_{it}$, as the dependent variable.

	GINIWAGES		*GINITOTINC*	
Independent variables	Pooled OLS	Fixed Effects	Pooled OLS	Fixed Effects
LF_POP	0.006** (0.000)	0.056** (0.014)	0.008** (0.000)	0.039* (0.081)
ECON_DEV	-1.10e-06** (0.002)	-1.38e-06* (0.063)	-5.53e-06 (0.164)	1.53e-07 (0.859)
UNEMP	0.193** (0.002)	0.121 (0.301)	-0.004 (0.957)	0.085 (0.499)
FEM_PART	-0.297** (0.000)	-0.103 (0.368)	-0.314** (0.000)	-0.109 (0.367)
MANUF	-0.046** (0.001)	-0.212** (0.000)	-0.038** (0.013)	-0.150** (0.001)
VIS_MIN	0.115** (0.000)	0.290** (0.000)	0.111** (0.000)	0.268** (0.000)
EDUC_RATIO	-0.004 (0.911)	0.032 (0.637)	0.019 (0.607)	0.065 (0.365)
YOUNG_LF	0.026 (0.236	0.009 (0.817)	0.019 (0.410)	-0.002 (0.954)
SENIOR_LF	0.046** (0.001)	0.036 (0.472)	0.041** 0.004)	0.013 (0.792)
CONSTANT	0.430** (0.000)	-0.146 (0.595)	0.382** (0.000)	-0.043 (0.871)
PROV. DUMMIES	YES	DROPPED	YES	DROPPED
T DUMMIES	YES	-	YES	-
N of Cities	87	87	87	87
ρ	-	0.982	-	0.961

One star (*) and two stars (**) indicate coefficients significant at the 0.10 and 0.05 levels, respectively. All models are heteroskedasticity-robust.
p-values are in parentheses.

In comparing the pooled-OLS and FE models, there is substantial difference in the magnitude and significance of each coefficient, although the signs are generally consistent. As mentioned in section 4.2.4, pooled-OLS does not account for unobserved heterogeneity among the model's residuals. This may insert bias into the model, possibly resulting in one or more Type-I errors, or the erroneous rejection of the null hypothesis that a coefficient is equal to zero. As such, Type-I errors in the pooled-OLS earnings model likely occur for the estimates of *UNEMP*, *FEM_PART*, and *SENIOR_LF*, and render that model's estimates less efficient than in the FE. Wilson and Butler (2007) offer a similar comparison of pooled-OLS and FE models, and show that slope coefficient estimates in the former may be over or under-represented, or even exhibit a sign change when unobserved heterogeneity is not removed.

In the FE model for earnings inequality, four variables significantly predict changes in the Gini during the period studied. The size of the labour force (a proxy to overall city size) is positively related to income inequality, the level of economic development is marginally significant and predicts lower inequality, the percentage of workers in manufacturing is inversely related to income inequality, and the percentage of visible minorities has a positive relationship to inequality. These results support and reflect the explanations provided above for their corresponding coefficients in the cross-sectional models; however, they also represent a broader view of the determinants of inequality during the period from 1996 to 2006.

The decade's most prominent influences upon Canada's urban earnings inequality echo the recent trends found in other industrialized economies. A strong manufacturing base is central to providing more widely shared economic benefits, and the wage differential faced by visible minorities (including those who are native-born, Aboriginal peoples, and the majority of recent immigrants) played a significant role in predicting higher levels of inequality in Canadian urban areas between 1996 and 2006. These influences are present in the determination of total income inequality as well, pointing to the inability of the country's redistribution mechanisms to fully temper rising inequality in recent years.

4.4 CONCLUSION

This chapter uses cross-sectional and panel data regression analyses to shed light on the determinants of earnings and total income distribution for individuals across labour forces in 87 urban areas for the census years 1996, 2001, and 2006. While the panel data results paint a broad picture on the major influences of economic inequality over the course of the decade, including the role of manufacturing in equalizing incomes and wage differentials faced by visible minorities leading to increased inequality, the cross-sectional analyses provide a window into how the nature and structure of economic inequality has changed for Canada's labour force between 1996 and 2006.

Major changes have occurred in the urban determinants of economic inequality in Canada. The rate of female participation in the labour force has emerged as having a significant equalizing effect on income distributions, as female pay rates have climbed substantially in the last decade. Furthermore, the predominance of men in the growing, higher-earning energy extraction industries may account for some of this change as well. With regards to the visible minority variable, while the level of visible minorities in the labour force continues to be positively related to both earnings and total income inequality, its impact has lessened, perhaps a result of immigration policies and reduced wage differentials.

Urban characteristics have also changed in their impact. The role of city size has become a significant contributor to economic inequality, as Canadian income inequality growth becomes increasingly an urban phenomenon. The level of a city's economic development (proxied by its median income) was once inversely related to economic inequality. Now, as income growth is concentrated in the upper income percentiles, median income has lost its significance as an equalizing variable.

CHAPTER 5

ECONOMIC INEQUALITY ACROSS CENSUS DIVISIONS

5.1 INTRODUCTION

The previous chapter analyzed determinants of economic inequality in Canada's urban areas. Here, the scope of analysis is widened to include all census divisions and examine the spatial variation of inequality across regions.

Researchers ranging from Coates, Johnson, and Knox (1977), Drennan (2005), and Excurra, Pascual, and Rapùn (2007) argue that national-scale growth in economic inequality inevitably has a spatial component, whereby regional or metropolitan divergence in overall wage characteristics contributes to national levels of inequality. They argue that in countrywide analyses at the sub-national scale it is *theoretically* necessary to account for the spatial heterogeneity of inequality levels, and that high or low levels may cluster in certain regions. The incorporation of spatial clustering into geographic analysis broadly applies to Tobler's "first law of geography" that "everything is related to everything else, but closer things more so," which Miller (2004, p. 284) places "at the core of spatial autocorrelation statistics."

It is also *methodologically* necessary to account for the spatial clustering of inequality and other variables. A regression analysis that uses contiguous spatial units as observations likely has spatial autocorrelation present in the data, violating the assumption that its error terms are uncorrelated with each other (Anselin, 2005). To visualize this, one may map the residuals of a regular OLS regression: the residuals capture variance left unexplained by the right hand side of the equation, and if they exhibit an embedded spatial structure, autocorrelation is present. This is a concern, as regression estimates may in effect be biased or inefficient.

In this chapter I use spatial autocorrelation and regression techniques to examine the determinants of national labour market inequality across Canada at the census division scale. The chapter begins with a presentation of methods for constructing spatial boundary data that is stable through time. Next, the technical aspects of spatial autocorrelation maps and spatial regressions are discussed. I then present descriptive and analytic maps of inequality across Canada, as well as the spatial lag and spatial error model results. The last section addresses some of the potential social and policy implications of my findings.

5.2 RESEARCH DESIGN AND METHODS

5.2.1 Data and dataset construction: the geography

This analysis uses the same Canadian *Census of Population* 20-percent micro-datasets for 1996, 2001, and 2006 presented in Chapter 3, and the same parameters for the targeted labour force. However, instead of focusing upon labour markets in individual CMAs and CAs, I construct the cross-sectional datasets using 2006 Census Divisions (CDs) as the units of analysis, which provide coverage of the entire country. Statistics Canada publicly provides CD spatial data through the CHASS system at the University of Toronto.

CDs are relatively stable geographic units established by Statistics Canada to facilitate regional planning, and they cover the entirety of the country. Each CD comprises contiguous municipalities (CSDs), and is delineated according to provincial/territory policies, although some provinces/territories defer the responsibility to the federal level. CDs are also categorized into 'types' by Statistics Canada (i.e. County or District), although their designations are not applied in this study.

Statistics Canada indicates minor changes to CD boundaries between 1996 and 2006, and Nunavut's designation as a territory in 1999 resulted in the reassignment of CSDs to that territory. However, the process involved to correct for these inconsistencies is less complex than in the previous chapter because nearly all of the CSD boundary changes that took place remained within CD boundaries. In accordance with the changes noted in the Census Dictionary provided by Statistics Canada,[26] ArcGIS is used to overlay the 2006 CD

[26] Cat. no. 92-566-XWE, "More information on the Census Division:" http://www12.statcan.ca/english/census06/reference/dictionary/geo008a.cfm .

boundaries on top of the 2001 and 1996 boundaries in order to confirm which CDs experienced changes. For example, the Quebec CD "Desjardins" (24 24) was dissolved after 2001, and its component CSDs were annexed to neighbouring CDs by Statistics Canada. Subsequently, I hard-code these CSD changes into the SAS programs that build the 1996 and 2001 census samples, and in effect reassign individuals in those censuses to their corresponding 2006 CD of residence.

This process results in 287 CDs, which constitute the observations in this chapter's analyses.[27] In contrast to the previous chapter on urban inequality, this chapter includes Canada's labour force in both rural and urban areas, and is representative of the entire country. Each CD's respective socioeconomic characteristics are constructed from the core labour force.[28]

To set the scene, Table 5.1 below provides a descriptive summary of labour force earnings inequality (as measured by the Gini coefficient) by census year and various geographic scales (e.g. national, provincial, CD, and urban/rural). The table provides a comprehensive picture of earnings inequality growth across Canada, between 1996 and 2006. The Gini values reflect inequality among census respondents aggregated according to each scale; for instance, Canada's Gini of 0.413 in 2006 represents earnings inequality among 11,911,610 weighted observations (as opposed to an average of 287 CD-level Ginis) for that year.[29] This approach removes any level-effects that may under-represent observed values of inequality. The provided CDs are ranked in order of their 2006 Gini values (as opposed to percentage-change in Gini).

[27] The creation of a new CD (1011) by splitting one original (1010) into two in Newfoundland before 2006 brings the number of actual CDs in Canada currently to 288. However, in order to preserve continuity, I re-merged the new CD (1011) to its original CD (1010). This was necessary due to the change in CSD boundaries within that CD after 1996.

[28] As in the previous chapter, the core labour force is parameterized to include those between the ages of 25 and 64, who earn a minimum of $1,000.

[29] See Appendix 3.2 for the weighted observation counts respective to each labour force sample studied.

Table 5.1 Descriptive summary of labour force earnings inequality across Canada

	Number of CDs	Gini coefficient			Percentage change
		1996	2001	2006	
Canada	287	0.381	0.389	0.413	8.4%
Provinces					
Newfoundland & Labrador	10	0.414	0.413	0.439	6.1%
Prince Edward Island	3	0.397	0.387	0.386	-2.8%
Nova Scotia	18	0.385	0.391	0.397	3.1%
New Brunswick	15	0.391	0.386	0.394	0.7%
Quebec	98	0.367	0.367	0.380	3.7%
Ontario	49	0.375	0.391	0.413	10.1%
Manitoba	23	0.372	0.370	0.383	2.8%
Saskatchewan	18	0.386	0.381	0.391	1.1%
Alberta	19	0.404	0.404	0.451	11.7%
British Columbia	28	0.382	0.384	0.413	8.1%
Yukon	1	0.364	0.358	0.358	-1.7%
Northwest Territories*	2	0.364	0.350	0.366	0.0%
Nunavut*	3	0.418	0.418	0.403	-0.4%
Top 5 Census Divisions in 2006					
1. Division No.9 (AB)	1	0.432	0.390	0.523	21.1%
2. Division No. 6 (AB)	1	0.415	0.423	0.491	18.3%
3. Division No. 7 (NL)	1	0.454	0.433	0.483	6.4%
4. Division No. 9 (NL)	1	0.462	0.447	0.478	3.5%
5. Centre Toronto (ON)	1	0.401	0.437	0.478	19.2%
Bottom 5 Census Divisions in 2006					
283. L'Islet (QC)	1	0.344	0.328	0.315	-8.4%
284. Les Etchemins (QC)	1	0.362	0.327	0.314	-13.2%
285. Acton (QC)	1	0.326	0.317	0.311	-4.6%
286. Montmagny (QC)	1	0.318	0.314	0.309	-2.8%
287. Bellechasse (QC)	1	0.325	0.320	0.304	-6.5%
Urban Census Divisions	63	0.378	0.389	0.416	10.1%
Rural Census Divisions	224	0.384	0.379	0.395	2.9%

Note: Results for the top and bottom five census divisions are sorted according to 2006 Gini coefficients.
Urban and rural classifications refer to census divisions with respectively more or less than 50% of their populations living CMAs according to 2006 population counts (33 CMAs in 2006). CDs for each census year correspond to 2006 boundaries, with exception of Newfoundland.
*Nunavut became a territory constructed from Northwest Territory CDs in 1999; however, these figures reflect 2006 boundaries.

Earnings inequality trends seen here generally correspond to those seen among Canadian urban areas studied in the last chapter. Overall, the country's inequality has increased 8.4% between 1996 and 2006; however, at smaller scales there exists a substantial amount of variation. Alberta and Ontario's levels of inequality have increased by 11.7% and 10.1%, respectively, while P.E.I.'s has decreased. Newfoundland stands out for its relatively high levels of inequality and growth among the Atlantic Provinces. Note that at the CD scale, most of those with the highest Ginis in the country saw dips in inequality in 2001, and major resurgences by 2006. Alberta's Division No. 9, between Edmonton and Calgary, saw over 20% of growth in its Gini after 1996. This phenomenon occurred at the provincial scale as well: only Ontario and Nova Scotia saw steadily increasing Ginis from 1996 and 2006. This is evidence of two very different periods of economic conditions for workers during those 10 years, and relates to the nature of GDP and real earnings growth discussed in Chapter 2 (see also *Earnings and income of Canadians over the past quarter-century;* Statistics Canada 2008). While GDP growth correlated with earnings gains across the income spectrum from 1996 to 2001, only the upper earnings quintile made gains during the second five-year period.

Lastly, the growth of urban inequality relative to rural inequality indicates a divergence in Ginis across Canada. Individuals living in highly urban CDs (classified as such when greater than 50% of their populations live in a CMA, according to 2006 population counts) saw an increase in earnings inequality of 10.1% between 1996 and 2006, while those living in rural areas saw an increase of only 2.9%. Furthermore, in 1996 earnings inequality among those living in urban areas was *lower* than for those living in rural areas. Inequality actually decreased in rural areas in 2001, while urban areas sustained steady inequality growth. Both urban and rural areas experienced gains in real average wages between 1996 and 2006 (15.1% and 11.4% respectively). Those gains in wages, however, were much differently distributed in each type of census division.

While the above paragraphs and table offer only descriptive assessments of inequality across Canada, initial impressions point to the effect of industry-specific economic growth in the resource-extraction sector on inequality, as seen in the higher levels of inequality in Alberta and Newfoundland, as well as the geographic variation in urban- and provincial-specific factors. I address these topics in this chapter's analysis.

5.2.2 Exploratory spatial data analysis

Levels and trends of economic inequality in the labour force vary widely across Canada. To emphasize the spatial component of earnings inequality, this chapter includes an array of descriptive and analytic maps created in ArcGIS 9.2 and GeoDa.

Scholars use spatial analysis to study various forms of inequality from the international scale (Ezcurra et al. 2007) to the intra-metropolitan scale (Holt and Lo, 2008). Evident in these analyses is the clustering of socioeconomic variables across political, administrative, and physical boundaries. As such, while descriptive maps can display levels of inequality for general purposes at varying geographical units of analysis (e.g. census tracts, counties, or census divisions), spatial statistical techniques examine whether or not inequality levels are randomly distributed at the global and local scales.

One of the most common techniques, the Moran's *I* statistic, "tests the null hypothesis that there is no underlying pattern, or deviation from randomness" among spatial data points (Rogerson, 2006, p. 235). It is expressed as follows:

$$I = \frac{n\sum_i \sum_j w_{ij} z_i z_j}{(n-1)\sum_i \sum_j w_{ij}}, \qquad (5.1)$$

where n are the number of spatial units, z is a z-score for the variable of interest y $\{ z = (y - \bar{y})/s \}$, and w_{ij} is a weight matrix that represents spatial proximity between units i and j. There are varying methods available for the construction of the weight matrix; however, in this study a binary connectivity matrix is used due to the nature of census division units (*areal polygons*), where $w_{ij} = 1$ if census divisions i and j are contiguous at a minimum of one point (called *queen's* contiguity) and 0 if not. The Moran's *I* is expressed as a value between -1 and 1, with -1 representing negative spatial autocorrelation (rare), 1 representing high spatial autocorrelation, and 0 representing complete randomness.

While the global Moran's *I* offers a broad measure of spatial autocorrelation for the entire sample, in order to map specific clusters of high and low levels of inequality within the sample a local spatial statistic is necessary (Anselin 1995, cited in Rogerson 2006). The local statistic I_i, also known as a

Local Indicator of Spatial Association (LISA), effectively decomposes the global Moran's I among the spatial units and is expressed as follows:

$$I_i = n(y_i - \bar{y})\sum_j w_{ij}(y_j - \bar{y}), \quad (5.2)$$

where y is the variable of interest, w_{ij} is the same weight matrix described above, and $\sum I_i$ is equal to the global Moran's I (Rogerson, 2006). *GeoDa* is used to generate the LISA maps in this chapter, which indicate clusters of high and low inequality that are statistically significant at the 0.05 level.

Each map displays the publicly available CD and CMA boundary shapefiles provided by Statistics Canada. They are in the Lambert Conformal Conic projection based upon the 1983 NAD, and reflect levels of inequality for 1996, 2001, and 2006 according to the 2006 census division boundaries.

5.2.3 The spatial regression models

In light of the above discussion concerning the methodological necessity of controlling for spatial autocorrelation, this section investigates the usefulness of two complementary spatial regression models in this chapter: the spatial lag model (SLM) and the spatial error model (SEM). Rather than relying upon provincial fixed effects ($PROV_{it}$), which capture provincial variation but not inter- and intra-provincial clustering, both models account for potential spatial autocorrelation at the CD scale with the construction of a *queen's* contiguity spatial weight matrix (w_{ijt}) as described in Section *5.2.2*. The SLM does so with the insertion of a spatially lagged dependent variable (e.g. inequality as measured by the Gini) among the model's independent variables. The SEM accounts for spatial autocorrelation captured in the regression's error term (ε_{it}). The application of each model's methodology and specification is discussed in the following section.

Both the SLM and SEM models are estimated in *GeoDa*, a program that offers a straightforward and useful platform for spatial statistics albeit with some limitations. *GeoDa* computes the SLM and SEM using a Maximum Likelihood approach rather than OLS due to the inconsistency of OLS estimates in situations of spatial dependence (see Anselin, 2005; Ward and Gleditsh, 2008). The comparison of one model's fit to another relies upon the value of each model's maximized log-likelihood function, with the higher value representing a better fit (Anselin, 2005, p. 209).

The SLM and SEM incorporate identical dependent and independent variables at the CD scale, with obvious exceptions for their respectively inserted lag and error variables. The SLM's specified form is expressed in the following equation:

$$INEQ_{it} = \alpha + \beta_1 POP\_DEN_{it} + \beta_2 ECON\_DEV_{it} + \beta_3 UNEMP_{it} + \beta_4 FEM\_PART_{it} + \beta_5 MANUF_{it} + \beta_6 VIS\_MIN_{it} + \beta_7 ABOR_{it} + \beta_8 EDUC\_RAT_{it} + \beta_9 YOUNG\_LF_{it} + \beta_{10} SENIOR\_LF_{it} + \rho\sum_{j=1}^{n} w_{ijt} INEQ_{it} + \varepsilon_{it} \; . \tag{5.3}$$

As in Chapter 4, the dependent variable [$INEQ_{it}$] represents earnings inequality as measured by the Gini coefficient but this time among individuals in CD i at census cycle t. The right hand side of the equation contains the independent variables, the scalar (α) and error term (ε_{it}), as well as the spatially lagged dependent variable ($\rho\sum_{j=1}^{n} w_{ijt} INEQ_{it}$). With regards to the latter, w_{ijt} represents the spatial weights matrix at time t. Thus, a positive and significant spatial lag coefficient (ρ) indicates that CDs i and j (at time t) are expected to be spatially autocorrelated with respect to $INEQ_{it}$. In other words, this variable controls for the clustering of high and low inequality values among contiguous CDs.

Table 5.2 Variable means and (standard deviations) for the 1996, 2001, and 2006 census cycles

Variable	1996	2001	2006
GINIEARNINGS	0.376 (0.028)	0.371 (0.028)	0.383 (0.035)
POP_DEN *(people per km^2)*	77.743 (323)	81.291 (340)	85.327 (346)
ECON_DEV[A]	28,164 (5,547)	29,237 (5,336)	30,043 (6,051)
UNEMP	0.085 (0.055)	0.073 (0.057)	0.064 (0.049)
FEM_PART	0.446 (0.028)	0.460 (0.024)	0.473 (0.025
MANUF	0.155 (0.088)	0.157 (0.091)	0.141 (0.082
VIS_MIN	0.0194 (0.038)	0.023 (0.046)	0.028 (0.057)
ABOR	0.048 (0.112)	0.059 (0.122)	0.069 (0.130)
EDUC_RATIO	0.412 (0.051)	0.392 (0.051)	0.370 (0.062)
HIGH_EDUC[B]	0.127 (0.048)	0.137 (0.053)	0.151 (0.060)
NO_HIGH[B]	0.285 (0.071)	0.255 (0.070	0.219 (0.078)
YOUNG_LF	0.671 (0.176)	0.584 (0.168)	0.515 (0.347)
SENIOR_LF	0.368 (0.150)	0.380 (0.144)	0.403 (0.139)
# of Obser. (CDs)	287	287	287

Notes: [A] *ECON_DEV* is represented by median real earnings of the *target population*, in 2002 dollars.
[B] These variables are combined in the *EDUC_RATIO* variable, and thus are not individually added to the model. They are provided for descriptive purposes only.

Table 5.2 provides means and standard deviations of the included variables. The independent variables represent determinants of inequality selected in keeping with the relevant literature presented in Chapter 2, and while most are equal in concept to those presented in Chapter 4, there are two changes. Specifically: (1) population density (POP_DEN_{it}), as measured by population/km^2, replaces the logged labour force population variable (LF_POP_{it}); and (2) a percentage-Aboriginal variable ($ABOR_{it}$) is separate from the formerly combined visible minority variable (VIS_MIN_{it}). The first change is due to the wide variation in CD geographic size. Although two CDs may have similarly sized populations, one may encompass a large rural area while the other a small urban area. Thus, a population density variable more accurately captures highly urbanized areas relative to their suburban and rural surroundings. In accordance with theory presented by Chakravorty (2006), POP_DEN_{it} is expected to be positive, indicating higher inequality in highly urbanized areas. However, if inequality clustering manifests in areas of agglomeration, a spatial lag variable that captures its effect may render the POP_DEN_{it} not significant.

The second change reflects that this analysis covers CDs in rural Canada, many of which have high percentages of Aboriginal residents. While Aboriginal people in urban areas may face structural barriers to employment similar to those faced by visible minorities and recent immigrants (DeVerteuil and Wilson, 2010), their employment contexts are different in rural areas, and thus the variables are separated here.

The SEM contains the same independent variables as described above for the SLM. However, rather than controlling for the embedded spatial dependence of solely the dependent variable as in the case of the latter, the SEM accounts for any spatial dependence in the model's residuals. According to Burt, Barber and Rigby (2009, pp. 569-570), such dependence may arise from "measurement error, or from the influence of spatially autocorrelated variables that are absent from the model and yet have some influence on other variables included in the regression." The SEM's specified form is expressed as:

$$INEQ_{it} = \alpha + \beta_1 POP\_DEN_{it} + \beta_2 ECON\_DEV_{it} + \beta_3 UNEMP_{it} + \beta_4 FEM\_PART_{it} + \beta_5 MANUF_{it} + \beta_6 VIS\_MIN_{it} + \beta_7 ABOR_{it} + \beta_8 EDUC\_RAT_{it} + \beta_9 YOUNG\_LF_{it} + \beta_{10} SENIOR\_LF_{it} + \varepsilon_{it} \ , \tag{5.4}$$

where $\varepsilon_{it} = \lambda \sum_{j=1, i \neq j}^{n} w_{ijt} \varepsilon_{jt} + u_{it}$.

Note that the error term (ε_{it}) includes two components, the first of which includes a coefficient representing spatially autocorrelated errors (λ) and the spatial weight matrix (w_{ijt}). The second component, u_{it}, captures spatially uncorrelated errors at time t.

Before estimating and comparing results from these two spatial regression models, the following section presents recent patterns of earnings inequality across regions.

5.3 MAPPING EARNINGS INEQUALITY

This section provides a discussion of both descriptive and analytic maps of earnings inequality across Canada. LISA maps show where high and low values of inequality are statistically clustered in the country, as well as where inequality levels have risen substantially.

5.3.1 LISA maps of earnings inequality: 1996, 2001, and 2006

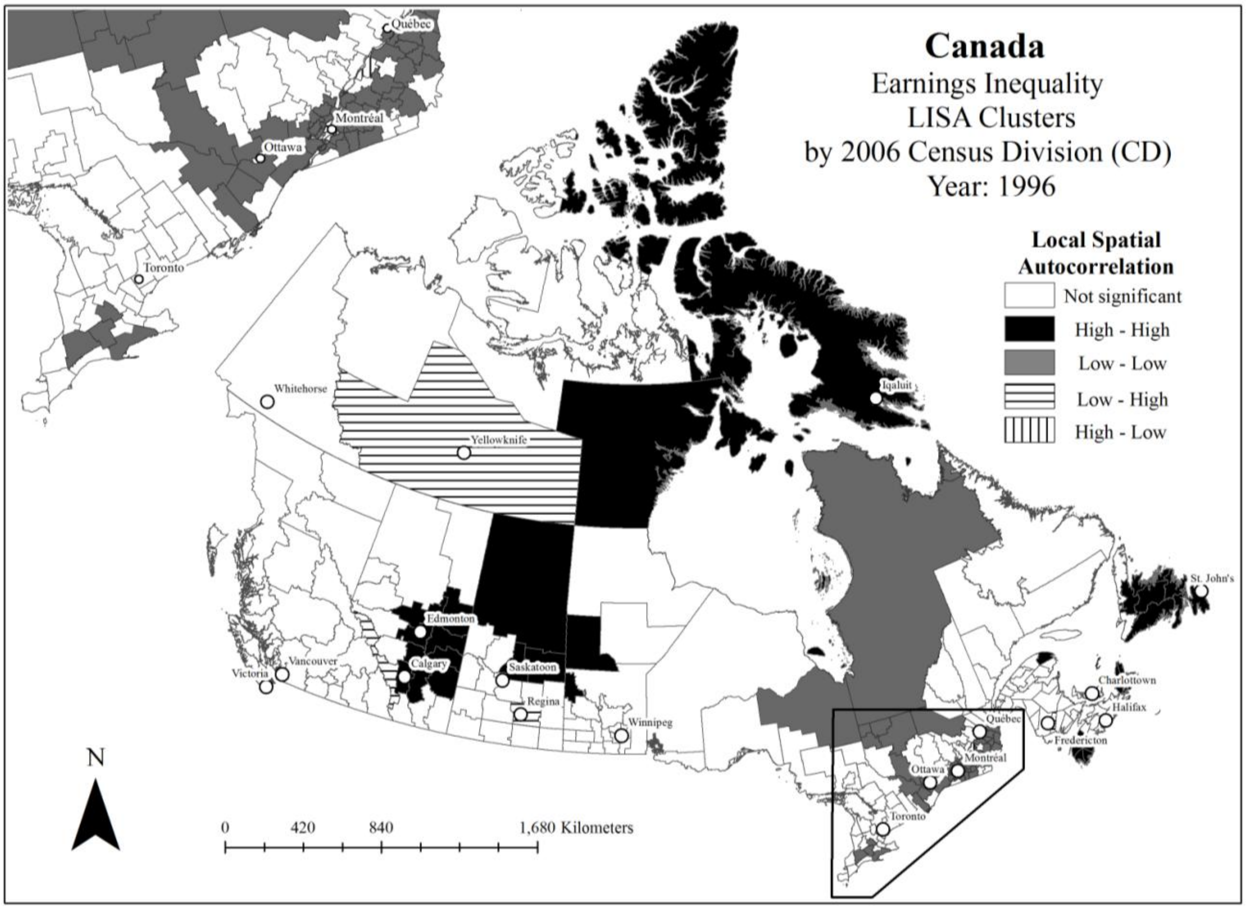
As described in Section 5.2.2, LISA maps display a spatial decomposition of the Global Moran's *I* value into local Moran's *I* statistics, and in this case indicate significant clusters of high or low values of inequality among spatially contiguous CDs. In the maps presented below, CDs within a cluster of high inequality are termed "high-high," as in "a CD with a relatively high level of inequality contiguous to a CD also with relatively high level of inequality." The reverse holds for clusters of low inequality, and a few CDs are "low-high," as in "a CD with a relatively low level of inequality contiguous to a CD with a relatively high level of inequality" and vice versa for "high-low." Table 5.3 breaks down the number of CDs that show significant and non-significant levels of inequality clustering, as well as the Global Moran's *I* value for each census year.

Table 5.3 Global and local spatial autocorrelation for earnings inequality, 1996 – 2006

	Canada						
	Earnings inequality by year	Global Moran's *I* (p-value)	Local indicators of spatial association (LISAs)				
			Not significant	High-high	Low-low	Low-high	High-low
Panel A	1996	0.51 (.0001)	189	34	57	6	1
	2001	0.52 (.0001)	194	32	50	8	3
	2006	0.57 (.0001)	196	31	50	8	2
Panel B	% change 1996-2006	0.34 (.0001)	216	24	33	11	3

As seen in Panel A, column 2, all values of the Global Moran's *I* are positive and statistically significant, a clear indication that positive spatial autocorrelation exists for inequality among CDs. Note that although the Moran's *I* value increases only slightly between 1996 and 2001, in 2006 it jumps to 0.57 from 0.52, indicating that inequality becomes more clustered during the later period. This is also evident in the reduction in both 'high – high' and 'low – low' clustered CDs, as similarly high and low inequality values concentrate spatially over time. For example, if inequality values became more spatially polarized, the resultant overall number of high – high and low – low CDs would decrease and clustering would increase. Panel B indicates that spatial autocorrelation also exists for changes in inequality, which means that clusters of relatively high rates of inequality growth (high-high) and decline (low-low) are found within Canada. The spatial manifestation of these increases in clustering is evident in the following LISA maps.

Figure 5.1 Local spatial autocorrelation map for earnings inequality, 1996



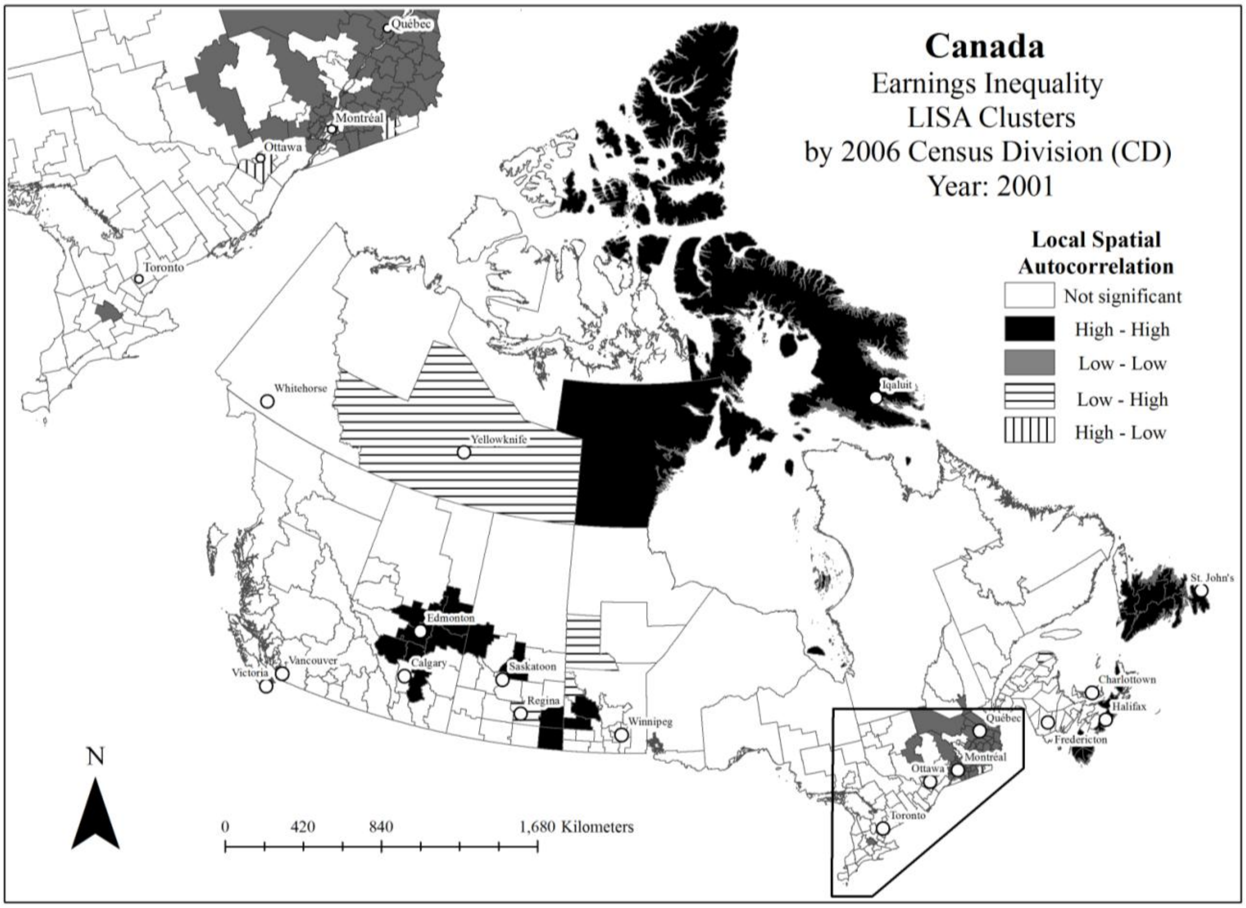
Figures 5.1, 5.2, and 5.3 display clustering of core labour force earnings inequality, respectively for the years 1996, 2001, and 2006. CDs in ‘high – high’ clusters of Gini values are depicted in solid black, and CDs in ‘low – low’ clusters are grey. ‘Low – high’ CDs have horizontal stripes, and ‘high – low’ CDs have vertical stripes. CDs that do not exhibit statistically significant LISA values are hollow. Major cities are also geo-referenced to help situate the reader.

In 1996, clusters of low values of inequality are found primarily in Quebec, as well as in Ontario around Ottawa and to the southwest of Toronto. The cluster of low inequality southwest of Toronto contains much of Ontario’s manufacturing production, including mainstay manufacturing companies such as Electro-Motive Diesels (sold by General Motors in 2005), the world’s largest diesel-electric locomotive producer and General Dynamics Land Systems, a major armoured-vehicle manufacturer. Other heavy manufacturers are located within this cluster including Ford and CAMI Automotive (see Rutherford, 2000). As argued in Chapter 2, areas in which high numbers of workers are in heavy manufacturing industries tend to have lower levels of earnings inequality as those jobs are typically well paid and unionized. Second, Ottawa's surrounding cluster of low inequality reflects its highly government-employed population, and the extension of the cluster into Quebec reflects the province's overall lower levels of earnings inequality.

Significant clustering of high values of inequality in 1996 is visible in Newfoundland, southern and northern Nova Scotia, between Calgary and Edmonton, and to the north and east of Saskatoon. The Baffin and Keewatin CDs in Nunavut (at the time they were in the Northwest Territories) also show high-high values of inequality; however, in this case the large size and small population of both CDs makes 'clustering' seem a bit strong to describe their relationship. These clusters reflect mainly the contemporary economic conditions. First, earnings inequality in Atlantic Canada, especially Newfoundland, is likely due to forces at the bottom and top of the income distribution: (1) the collapse of the groundfish stocks in the early 1990s, which affected earnings of over 20,000 workers in Newfoundland alone (Gien, 2000), and (2) the subsequent rise of the offshore oil and gas industry, which has experienced major growth within the province since the late 1990s. Second, southeast Alberta's cluster of CDs with high inequality reflects the growth of energy extraction industries by 1996, which caused unequal gains in earnings (Gibson, 2007). Gingrich (2009) documents similar economic conditions in Saskatchewan with both mining and energy extraction.

As seen in Figure 5.2, earnings inequality in 2001 continues to cluster in various parts of the country, as it did in 1996. Central Alberta and the region east of Calgary contain clusters of high inequality, as well as Newfoundland, parts of Nova Scotia, north of Saskatoon, and between Regina and Winnipeg. Two CDs in Nunavut continue to exhibit higher relative inequality values as well. Clustering of low inequality values are condensed within Quebec, and the 1996 cluster southwest of Toronto has shrunk to only one CD that registers as statistically significant. Both high and low clusters of inequality are found in pockets around the country in 2001; however, by 2006 Gini values cluster in only three regions.

Figure 5.2 Local spatial autocorrelation map for earnings inequality, 2001

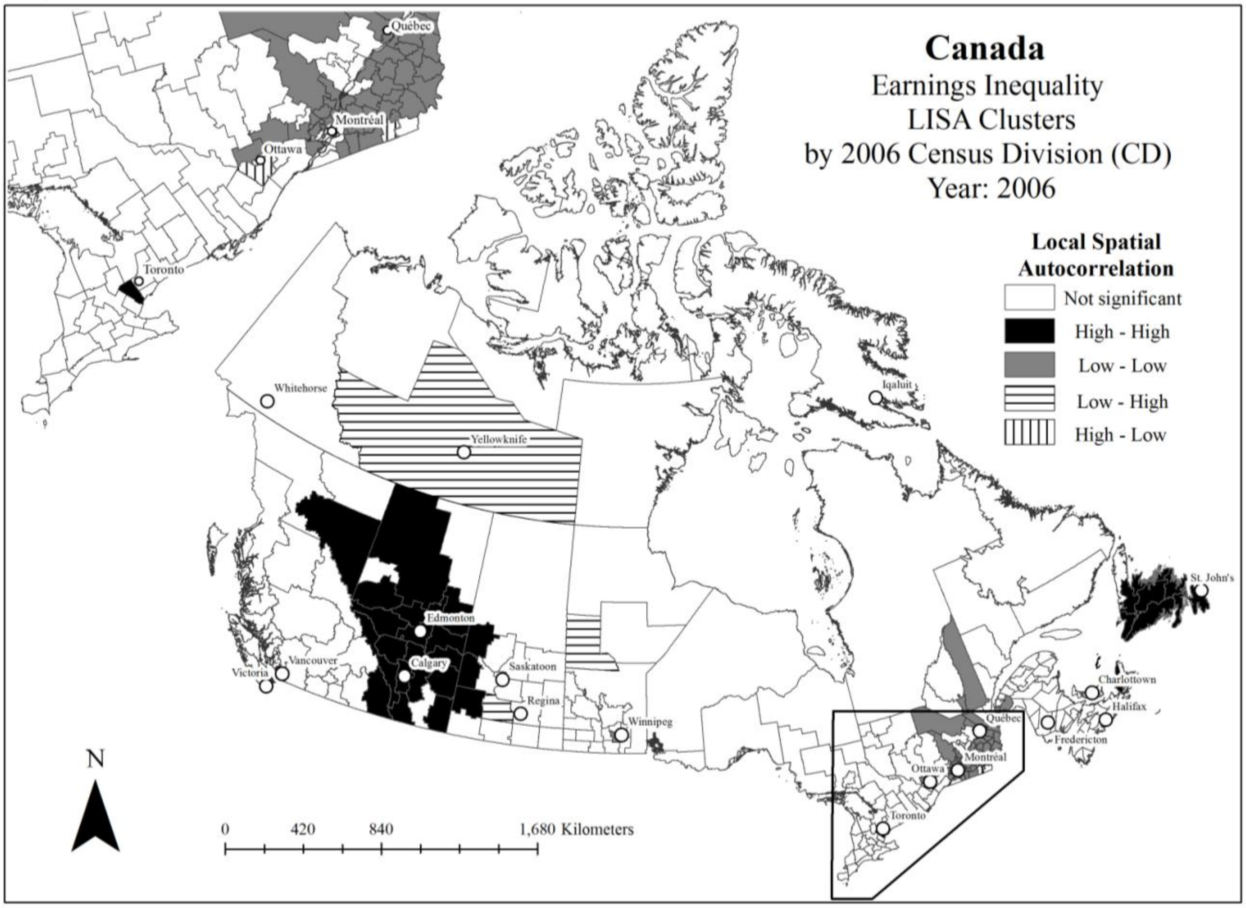


In 2006 (see Figure 5.3) high levels of earnings inequality largely cluster in Alberta, with some overstep of provincial lines, as well as in Newfoundland. Low inequality clustering is reduced to southern Quebec, although Montreal's immediate region is not statistically significant. Nova Scotia's pockets of high inequality are no longer significant, possibly reflecting the rebounding fisheries industry after the late 1990s.[30] The country's largest cluster of earnings inequality overlaps considerably with the Western Canada Sedimentary Region, which holds

[30] See "Fisheries and Oceans Canada" http://www.mar.dfo-mpo.gc.ca/pande/ecn/ns/e/ns7-e.asp. Retrieved January 19th 2010.

the coal, oil, and gas resources at the centre of Alberta's recent energy extraction boom.

Figure 5.3 Local spatial autocorrelation map for earnings inequality, 2006



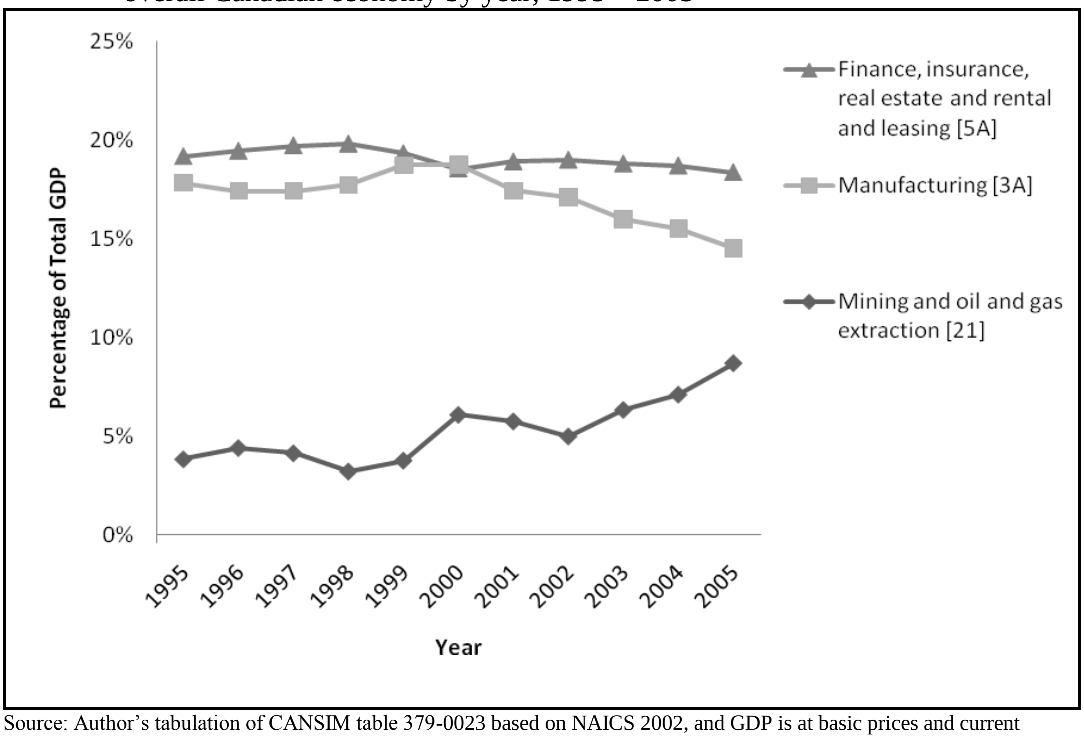
Since 2002, Alberta has experienced "the strongest period of economic growth ever recorded by any province in Canada's history," causing a major labour shortage, higher wages, an oil and gas industry that employs one in every five of the province's workers, and booming constructing and housing industries (Cross and Bowlby, 2006, p. 3.1). The economic boom has created opportunity for many workers in the province, as well as immigrants from elsewhere in Canada. However, in addition to environmental concerns, the economic boom has led to greater earnings inequalities, with many individuals not benefiting from its growth:

> Where incomes for middle class Albertans rose, it was mostly due to working more hours, not higher wages. For those Albertans at the bottom, their incomes are falling; social assistance rates are at 50% of the levels of the 1980s, and minimum wages are not keeping up with inflation and are still amongst the lowest in the country (Gibson, 2007, p. v).

Figure 5.3 above visually supports Gibson's statement, and is reminiscent of Table 3.3, which shows that earnings inequality within the aggregated mining, oil, and gas extraction industries increased 34.9% between 1996 and 2006.

In order to gain a better understanding of the economic role of the oil and gas extraction industries, I show the proportion they represent in the Canadian economy, in current (unadjusted) dollars year-by-year in Figure 5.4. This method does not require adjusting for inflation, as proportions are compared across time rather than dollar amounts, and the use of current dollars more accurately captures the steep price increase of oil in recent years.

Figure 5.4 Representation of the three largest NAICS 2002 aggregate categories in the overall Canadian economy by year, 1995 – 2005



Source: Author's tabulation of CANSIM table 379-0023 based on NAICS 2002, and GDP is at basic prices and current dollars.

In 2005, the aggregated mining, oil, and gas extraction industries formed the third largest share of the Canadian economy (8.7%), behind finance, insurance, real estate, rental and leasing (18.4%), and manufacturing (14.5%). While manufacturing's share of GDP in the Canadian economy grew in the late nineties, after 2000 it fell from 18.8% to 14.5% of GDP. Meanwhile the mining, oil and gas extraction sector has grown in its share of GDP from 3.9% in 1995 to 8.7%, and the timing of its growth parallels the post-2002 Alberta oil boom (Cross and Bowlby, 2006; Williams, 2007). No other sector, including those comprised of highly-skilled technicians and creative workers, grew nearly as much in its share of total GDP in current dollars. However, Cross (2008) notes that since mainly surging prices have driven the resource sector's boom, its output and job-growth have not seen similar growth. He also links a surge in prices to growth in the sector's level of investment, making it "the leading factor in our stock market and direct investment flows into Canada" before the global economic turmoil in 2008 (p. 3.1). This factor may also play a role in the observed increase in inequality among the country's urban financial centres, which is considered in the next section.

In sum, this section's LISA maps show statistically significant clusters of earnings inequality for each of the 1996, 2001, and 2006 censuses. Between 1996 and 2001, clustering did not substantially increase, although its local pattern shifted somewhat. However, by 2006 there is a clear increase in levels of clustering, and two major poles of high inequality emerge in Canada: along the Western Canada Sedimentary Basin, and within Newfoundland.

Figure 5.4 shows that between the mid-1990s and 2000, Canada's manufacturing industries (which employ a large number of workers) maintained a steady share of the national economy. It follows that earnings growth in that period would have been more widely distributed (which is seen in Tables 3.1 and 3.2, as well as in Chapter 4's analysis). However, after 2000 manufacturing lost ground and the composition of the economy shifted. The mining, oil and gas-extraction sector doubled its contribution to overall current-dollar GDP, and became the third largest share in the economy in 2005. This gain occurred without much change in the sector's employment levels (Cross, 2008). Furthermore, it accompanied large increases in within-industry earnings inequality and increased inequality-clustering in its core regions.

5.3.2 Mapping changes in earnings inequality: 1996 – 2006

The poles of earnings inequality described above contain high Ginis relative to the rest of the country for each cross-sectional period; however, the next two maps show where Ginis have changed most between 1996 and 2006.

Figure 5.5 Earnings inequality percentage changes, 1996 – 2006

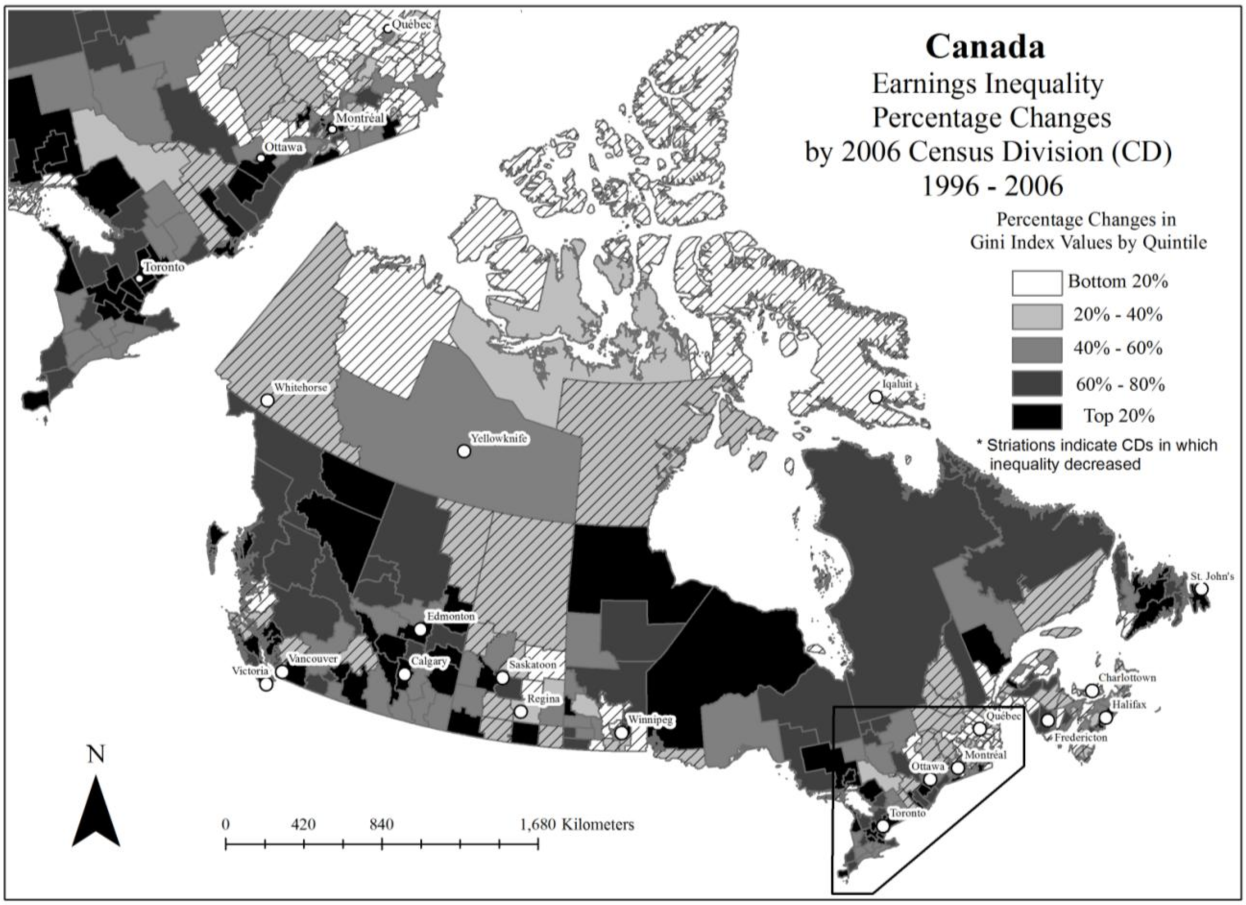
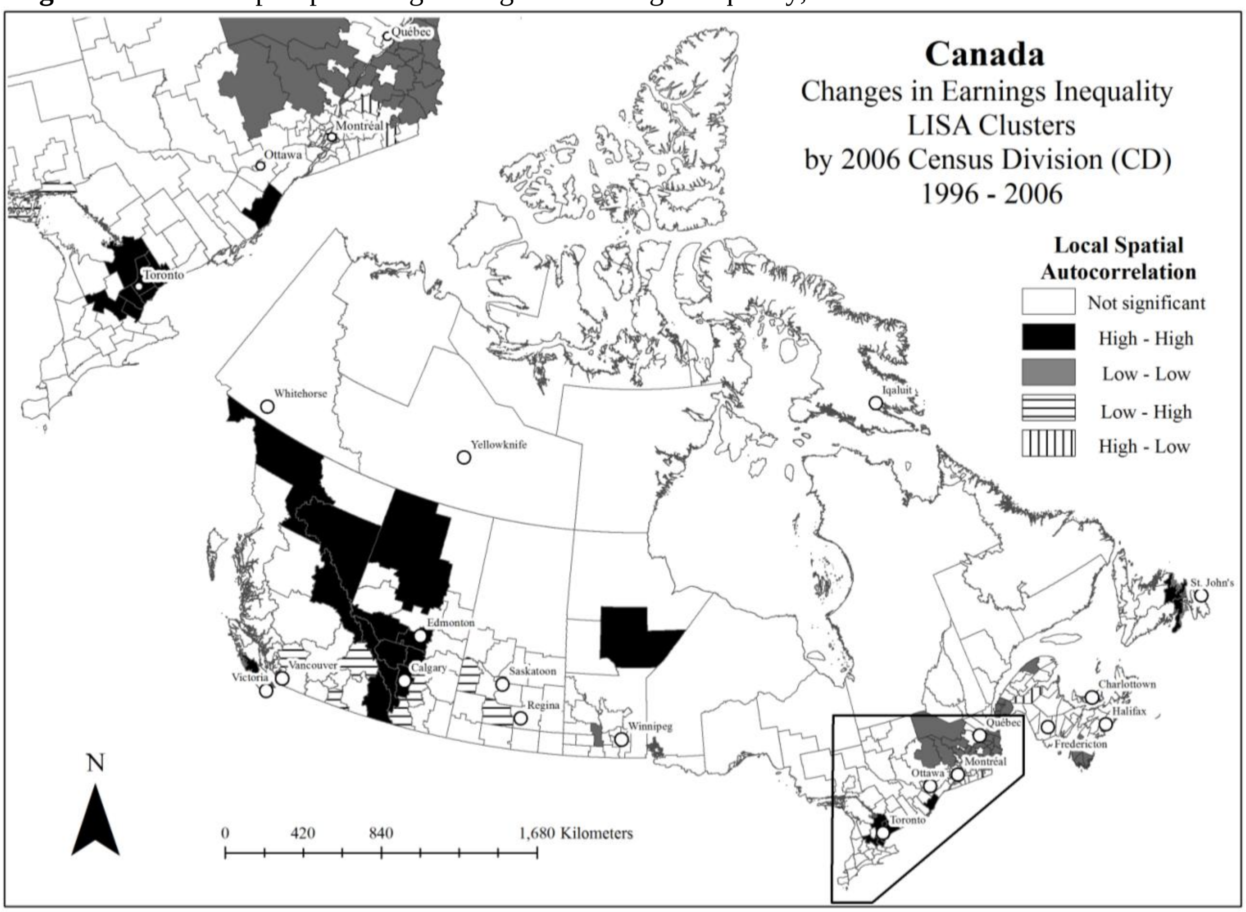


Figure 5.5 shows Gini percentage changes between 1996 and 2006, by CD. The changes range between -13.4% in Les Etchemins, Quebec (just east of Quebec City) to 26.0% in Division No. 15, Alberta (a long CD bordering British Columbia, west of Calgary). In the map they are organized by quintile with each

quintile containing about 20% of the total number of CDs with the lowest 20% including negative values. Striations indicate those CDs in which Ginis decreased.

It is apparent that much of the labour force across Canada experienced rising levels of earnings inequality between 1996 and 2006, however, there is considerable spatial variation in where those increases occurred. Increases are visible among urban areas across the country, as well as much of British Columbia, Alberta, northern Manitoba, Ontario, northern Quebec, and Newfoundland and Labrador. CDs that experienced decreases in inequality are dispersed throughout the country, but are mainly located in parts of Manitoba, Quebec, and the remaining Atlantic Provinces. The LISA map in Figure 5.6 highlights clustered CDs that show statistically significant changes in inequality.

Figure 5.6 LISA map of percentage changes in earnings inequality, 1996 – 2006



Clusters of relatively high increases in earnings inequality appear in Alberta, again overlapping with much of the Western Canada Sedimentary Basin, and crossing provincial lines into British Columbia. A few CDs surrounding the Alberta cluster have significantly low levels of inequality compared to their neighbours, as evidenced by the 'Low – High' horizontal striations, including the mostly rural CD Columbia-Shuswap Regional District, BC, which lowered its overall inequality between 1996 and 2006 as seen in Figure 5.5. It is important to note that high increases in earnings inequality clustered significantly in the Greater Toronto Region and south of Ottawa, even though these centres are outside of the immediate areas surrounding the energy-extraction poles of inequality mentioned earlier. Toronto's concentration of corporate offices, financial markets, and its 'global city' status clearly play into what Chakravorty (2006, p. 176) calls the "increasing returns of urban and metropolitan regions" and their role in raising levels of national inequality. I explore this idea further in *Section 5.5*, but first *Section 5.4* investigates the presence of spatial autocorrelation among the residuals of OLS regression applied to Equations 5.3 and 5.4.

5.4 Spatial autocorrelation of residuals in OLS regression

As discussed in Sections 5.1 and 5.2.2, the presence of spatial autocorrelation among the residuals of OLS regression violates a central assumption of the model, likely inserting bias. To test for such a condition, one may map the regression's residuals and apply the Moran's *I* statistic (see Appendix 5.1). When OLS regression is applied to Equations 5.3 and 5.4 (effectively removing respectively the lag and spatial error variables), strong spatial dependence among residuals is indicated by the Moran's *I* for all census years. However, if the model contains provincial dummies, as in the previous chapter, the residuals for the 1996 and 2001 model estimates do not show significant spatial dependence. This is an indication that provincial level unobserved heterogeneity (i.e. political, social, and economic differences) accounts for much of the regional variation in inequality for census years 1996 and 2001. Here we see the importance of political boundaries, and OLS regressions with provincial dummies for the 1996 and 2001 datasets offer a suitable method for assessing the effects of the selected determinants of inequality. However, this changes in 2006, as spatial autocorrelation is present and significant among the residuals (even with controls for provincial fixed effects), and Appendix 5.1 displays a LISA map in which a cluster of CDs with significantly high residuals exists in Alberta. As such, *new patterns of spatial dependence emerge in 2006, which are processes not explained by provincial differences, that insert bias into the OLS model.*

In light of this finding, it is necessary to apply spatial regression to each dataset in order to maintain methodological consistency across census years. Furthermore, Ward and Gleditsch (2008, p. 64) raise the point that "the regional dummy variable assumes that all observations within every region are homogeneous and interconnected, whereas the spatially lagged *y* model [or the spatial error model] allows the degree of similarity to be estimated." Thus, while provincial variation explains a large portion of regional variation in earnings inequality across Canada for all census years, by 2006 intra-provincial clustering of high and low levels of inequality emerge and render the regional-dummy approach less desirable.

5.5 Spatial regression results and discussion

This section presents a discussion of the results for the spatial regression models, which use earnings inequality as the dependent variable. In extension to the urban findings presented in Chapter 4, they corroborate that both the *nature* and *structure* of earnings inequality have changed between 1996 and 2006. These findings reflect national-scale processes, as well as changes in the spatial structure of inequality.

Table 5.4 Spatial error model (SEM) regression results with earnings inequality as the dependent variable.

Independent variables	Spatial error model (SEM)		
	1996	2001	2006
POP_DEN	3.787e-06 (0.317)	7.376e-06* (0.075)	1.26e-05** (0.008)
ECON_DEV	-7.014e-07 (0.167)	-7.453e-07 (0.103)	4.34e-05 (0.322)
UNEMP	0.216** (0.000)	0.111** (0.003)	0.204** (0.000)
FEM_PART	-0.105** (0.044)	-0.082 (0.191)	-0.102 (0.163)
MANUF	-0.043** (0.009)	-0.065** (0.000)	-0.089** (0.000)
VIS_MIN	0.022 (0.603)	0.072* (0.071)	3.09e-005 (0.999)
ABOR	0.033 (0.105)	0.056** (0.010)	-0.044* (0.091)
EDUC_RATIO	0.061** (0.033)	0.060** (0.043)	0.160** (0.000)
YOUNG_LF	-0.015 (0.177)	-0.012 (0.352)	0.020 (0.184)
SENIOR_LF	0.047** (0.000)	0.033** (0.008)	0.061** (0.000)
ρ (*spatial lag*)	-	-	-
λ (*spatial error*)	0.846** (0.000)	0.708** (0.000)	0.885** (0.000)
CONSTANT	0.401** (0.000)	0.398** (0.000)	0.329** (0.000)
Likelihood Ratio Test	61.097** (0.000)	41.774** (0.000)	97.109** (0.000)
Log likelihood	784.46	756.55	707.31
Breusch-Pagan Test	12.125 (0.277)	**20.880** (0.022)**	2.824 (0.985)
PROV. DUMMIES	NO	NO	NO
N	287	287	287

Note: p-values are in parentheses, and two stars (**) marks those coefficients significant at the 0.05 level, and one star (*) marks those coefficients marginally significant at the 0.10 level.

Table 5.4 provides the SEM regression results in three columns, with each column referring respectively to variable coefficients for the 1996, 2001, and 2006 datasets. The discussion focuses on the SEM results; however, the SLM results can be found in Appendix 5.2. At the bottom of the table, the Likelihood Ratio Test compares the null model specification (the OLS model) to the respective spatial regression model's specification, and is based upon the χ^2 distribution with 1 degree of freedom (see Anselin, 2005). One may interpret it as a redundant measure of significance to the *z*-test for the spatial error (λ) coefficient. The log likelihood function tests each model's fit relative to the others, and the Breusch-Pagan test is significant (and highlighted in bold) when the model exhibits heteroskedastiticity, or whether the estimated variance of residuals is dependent on the independent variables. The latter presents a complication as GeoDa does not offer a heteroskedasticity-robust regression option, and as such caution is needed in interpreting the 2001 SEM results.[31]

The following sections analyze the results presented in Table 5.4. The first compares the spatial error and spatial lag model results, and argues that the SEM is both a theoretically and empirically superior tool in this case. The second portion discusses the independent variable results with a focus on the SEM.

5.5.1 Choosing the SEM over the SLM

As discussed in Section 5.4, there is no significant spatial autocorrelation of residuals in OLS regressions of the 1996 and 2001 datasets when provincial dummies are included. In the 2006 OLS estimation, however, the Moran's *I* of residuals is positive and significant (see Appendix 5.1), which indicates the presence of spatial autocorrelation unaccounted for by the model. This begs the following question regarding the strengthening of spatial processes between census divisions: is inequality observed in one census division directly influenced by inequality in a neighbouring census division, or is it influenced by observed and unobserved determinants of inequality in neighbouring census divisions? As

[31]The Halton census division (which contains Oakville) in Ontario represents the largest residual in the SEM model, and its omission from the model reduces the problem of heteroskedasticity in 2001. See Appendix 5.3 for model runs that omit this observation.

inequality is theorized as a phenomenon dependent upon the socioeconomic factors presented in Chapter 2, as well as unobservable factors, the latter case is more plausible. Thus the desirable model should address spatial autocorrelation that is present not only within the dependent variable (inequality), but also the independent variables. Based on this reasoning, the SEM is the better model with regards to measuring the determinants of inequality.

In the SEM results, the spatial error term's positive and strongly significant coefficients (λ) in the 1996, 2001, and 2006 estimations, as well as the corresponding likelihood ratio tests, indicate that spatial effects are present among both the independent and dependent variables.[32] Such effects may include exogenous physical, political, economic, and social factors. Conversely, the spatial lag term's coefficient only applies to spatial autocorrelation of the earnings inequality variable. The SLM, exhibited in Appendix 5.2, does not control for how provincial or other differences may affect the independent variables. As evidenced by the spatial lag coefficient's (ρ) non-significance in 1996 and strong significance in 2006 (again, also seen in the likelihood ratio tests), spatial clustering of the dependent variable becomes apparent by 2006. This corresponds with the pocket of significant residual clustering that appears within Alberta in 2006, detailed in Section 5.4 and Appendix 5.1.

There are three important points to take from the above. First, as seen in the emergent significance and increasing magnitude of ρ in the SLM between 1996 and 2006, earnings inequality increasingly clusters in that time. Although the SLMs are heteroskedastic, the likelihood ratio test confirms such a conclusion. The second important point is that the SEM accounts for other spatial factors at work, and indicates a growing presence of overall spatial effects with the increasing value of the λ coefficient. Lastly, since the SEM accounts for significant yet implicit spatial processes that are present in each cycle, it is

[32] In the SLMs and SEMs, provincial dummies are unreliable. For a discussion of why provincial variables are not appropriate in MLE spatial regression, see: http://geodacenter.asu.edu/openspace/2003-December/000109.html.

theoretically the superior model.[33] Furthermore, the SEM displays less heteroskedasticity than the SLM, and better fits the data, making it empirically superior.

5.5.2 The SEM results and discussion

Unless indicated, the following discussion pertains to the SEM results presented in Table 5.4, starting with the variables that show statistical signficance. The focus is on the 1996 and 2006 SEM results since the 2001 model is heteroskedastic. However, the partial coefficient values in the 2001 model are plausible as they generally fall in line with the other models.

The population density variable *POP_DEN*, as measured by number of people per km^2, is not significant in 1996, and by 2006 is strongly significant in 2006 (p = 0.008). The small coefficient in 2006 may seem negligible (an increase of one person per km^2 corresponds with a rise in the Gini by 0.0000126 percentage points); however, it is meaningful in cases of large cities which have population densities of 1,000+ people per km^2. A rise of 1,000 people per km^2 corresponds with a 0.013 rise in the Gini, or nearly half of the latter's standard deviation (s = 0.035). This effect in 2006 corresponds with arguments introduced earlier in the chapter by Korpi (2008) and Chakravorty (2006) as well as Chapter 4's results, and points to the emerging role of large cities in the national level of inequality.

There are negative and significant correlations in the CD-scale sample between *ECON_DEV* and *INEQ* in 1996 (r = -0.50*) and in 2006 (r = -0.13*);[34] however, economic development (*ECON_DEV*), as measured by median wages, is not significant in predicting earnings inequality in the SEMs when controlling for the other independent variables. This finding diverges somewhat from the urban results, in which median earnings have an overall negative impact on

---

[33] Anselin (2005) offers a useful spatial regression discussion and flowchart on p. 198, which both focus on the choice between the SEM and the SLM. His discussion corroborates the argument made in this section.

[34] Note the substantial drop in the correlation coefficient. This alludes to the diminishing negative relationship between median earnings and inequality, as economic development no longer means equitable earnings growth.

inequality in 1996 and 2001, and are not significant in 2006, but points to the role of high incomes in raising inequality: the *ECON_DEV* variable does not dampen inequality as expected once other variables are included in the model. It is likely that higher median earnings are not representative of wider equality as they once were (Chakravorty, 1996). Such a transition was marked in Chapter 4's urban analysis post-2001, and perhaps such is the case pre-1996 when including rural areas into the analysis as I do here. Essentially, this finding is evidence that economic development no longer translates into equitable earnings growth for the core labour force.[35]

Unemployment (*UNEMP*) is a dependable predictor of higher earnings inequality in all census cycles across regions. This is in-line with the literature regarding macroeconomic conditions, whereby increased unemployment represents a labour surplus and thus lower wages (MacPhail, 2000a). This variable likely accounts for inequality driven by the lower echelons of the earnings distribution, as low-paid workers earn significantly less in CDs of higher unemployment.[36] In 2006, an increase of one percentage point of unemployment effects a 0.20 point rise in the Gini. This is comparable to the 1996 value; however, in 2001 the coefficient is only 0.11, indicating that unemployment did not have as much influence in determining inequality in that year.

The percentage of the core labour force employed in manufacturing (*MANUF*) is also a consistently significant explanatory variable. A one-percentage point increase in the rate of manufacturing in 2006 corresponds with a 0.089 point decrease in the Gini. Notable here is the rise in its coefficient since 1996, from 0.043 to 0.089. This highlights an increasing role of manufacturing in dampening inequality, as well as the service-sector and resources-sector's increasing role in raising inequality. These results corroborate findings in Chapter 4 among urban areas as well as the cited deindustrialization literature; however, at

[35] Furthermore, if an average earnings variable replaces the median earnings variable, the former's coefficient increases in magnitude between the 2001 and 2006 model estimations, and is positive and strongly significant in all three. Average earnings are more sensitive to the upper end of the distribution due to its positive skew, and thus pick up on the upper end's influence on raising inequality.

[36] The correlation between unemployment and median wages in 2006, among CDs, is $r = -0.61*$.

the national level among CDs manufacturing plays an even greater role in dampening inequality.

Unexpectedly, the percentage of the labour force composed of visible minorities (*VIS_MIN*) does not significantly contribute to inequality in any SEM. However, it is strongly and positively significant in all cycles of the SLM. I attribute this inconsistency to the spatial patterning of visible minorities in Canada, and that the spatial lag variable in the SLM does not capture such patterning. For instance, in 2006 there are two major clusters of CDs with high percentages of visible minority residents in the country: one in the Vancouver area and the other in the Toronto area. The 2006 Global Moran's *I* value for clustering of *VIS_MIN* is 0.50*, indicating significant spatial autocorrelation. Both clusters are within provincial boundaries and separate from the spatial patterning of OLS residuals that emerges within Alberta presented in Section 5.4. Thus, the spatial error term captures the spatial structure of *VIS_MIN*, likely suppressing it in the SEM.[37] In light of this circumstance, I conclude that a higher presence of visible minorities in a CD's core labour force corresponds with higher earnings inequality, largely due to structural obstacles to full employment faced by visible minorities and recent immigrants (Moore and Pacey, 2003). There is evidence of its diminishing impact between 1996 and 2006, as seen in the coefficient's drop from 0.193 to 0.150 in the SLM, which corresponds with a similar shift across urban areas as presented in Chapter 4. However, the localized nature of visible minority populations within major cities in Canada makes a clear relationship difficult to discern in the national context. Chapter 4's detailed look at urban labour market structures is more enlightening in this regard.

Also somewhat of a puzzle is the *ABOR* variable's change of sign in 2006. In the 2001 model, higher percentages of Aboriginal people significantly predict higher levels of inequality; however, in 2006 the variable is a marginally significant dampener of inequality. As shown in Appendix 5.3, the 2006 variable is not significant with the omission of the highest outlier, which might infer that

[37] In an OLS regression of the 2006 model, the visible minority variable is also suppressed by provincial dummies.

the variable's coefficient estimate is biased in the model. However, the change of sign may reflect the overall higher earnings of people in northern Canada (and the north's lower levels of inequality as seen in Table 5.1), where Aboriginal populations have a higher presence in the labour force. Perhaps the existence of a spatial collinearity between relatively lower-inequality regions and higher percentages of Aboriginal people, or higher earnings of Aboriginal people themselves who work and live in those regions may contribute to this variable's change of sign in 2006.[38]

The spread of educational attainment among workers in the core labour force (*EDUC_RATIO*) is a dependably significant contributor to higher rates of earnings inequality. Recall that the variable is simply a proportion of those without a high school degree, plus those with a Bachelor's or higher, to the remaining labour force: an increased value indicates a higher degree of educational attainment inequality. Its impact upon earnings inequality grows substantially between 1996 and 2006. In 1996, a one-percentage point increase in *EDUC_RATIO* corresponds with a 0.06-point rise in the Gini. By 2006, the coefficient increases to 0.16. This corresponds to findings in the United States regarding an increasing education premium during the 1980s (Cloutier, 1997), and more recently in Canada (Breau, 2007). Furthermore, this finding differs markedly from results in Chapter 4, indicating the greater impact of educational inequality on earnings inequality at the national scale, especially in recent years. The comparison leads to questions regarding how the education premium differs in urban vs. rural areas. For instance, Cross and Bulby (2006, p. 3.11) point out that "rural Alberta has one of the highest rates of high school drop-outs in the country at about 25%, presumably spurred by the promise of attractive pay for relatively unskilled work. However, this leaves youths ill-prepared to deal with the consequences of a slump in the industry." It follows that areas of high

[38] Census division 18 in Saskatchewan, which comprises the northern half of the province, is an example of where inequality decreased between 1996 and 2006, real median earnings increased 14%, and the percentage of residents who self-identify as Aboriginal is over 60%.

earnings inequality in Alberta correspond with areas of high levels of educational inequality, and that both are linked closely to the oil boom in this case

Similar to findings in Chapter 4, a larger percentage of a CD's population that is composed of seniors relative to the labour force increases earnings inequality. This variable captures the age structure of the geographical area, and the effect that older family members may directly have on the earnings of working individuals.

5.6 CONCLUSION

This chapter explores the changing regional distribution of earnings inequality in Canada's labour force between 1996 and 2006, as well as the links between these spatial patterns and possible determinants of inequality at the CD-scale. Its approach is novel in that the spatial heterogeneity of inequality and other variables is statistically tested, mapped, and controlled for in regression modelling.

This chapter's results build on findings in Chapter 4 for urban areas, and strengthen the conclusion that the spatial structure of earnings inequality among individuals in Canada has changed. Its most important findings lead to a number of conclusions. To begin, high levels of urbanization are emerging as major determinants of national levels of inequality, and an overall divergence in inequality has occurred between rural and dense urban regions. However, when mapped, it is clear that the major poles of inequality increasingly centre upon regions most dependent on the energy extraction industries. Intra-provincial spatial processes are at work by 2006, signalling that returns to industry-specific agglomeration are becoming increasingly regionalized within the country. Alberta's unprecedented oil boom has drastically changed its regional economy as well as contributed to national growth in inequality. While this boom has provided much opportunity for some and great provincial and national economic expansion, the findings here indicate that this expansion has not lowered inequality rates as growth did in previous years. I provide evidence that the boom is primarily benefiting those in the top earnings brackets, to the point at which an overall economic development indicator, median earnings, is no longer a predictor of lower inequality rates. Additionally, as in the urban analysis, manufacturing increases its impact in mitigating inequality across Canada. As the labour force shifts from manufacturing to the service and resource sectors, the former's role as a dependable wage-equalizer grows. Lastly, educational dispersion emerges as a significant and positive predictor of inequality.

CHAPTER 6

CONCLUSION

This thesis sought to address the following research questions: *(1) How has earnings inequality evolved in recent years across both urban and rural areas in Canada? (2) Are some areas within Canada becoming more unequal than others? (3) What structural, institutional, and spatial factors explain these changes in inequality?*

In addressing these questions, this study has made a number of contributions to the literature on inequality in Canada. First, the samples used in the analysis draw upon the 20-percent long-form samples of the *Census of Population* for the years 1996, 2001, and 2006, which provide a high level of earnings and geographical detail, as well as widespread coverage across the country. Their use allows for a better picture of earnings inequality trends across Canada in comparison to the use of publicly available microdata. Second, the study is one of few in geography that analyzes Canada's inequality at the sub-national scale. It investigates the determinants of inequality at both the urban and regional scales, and offers evidence regarding changes in the structure of earnings inequality in recent years. Third, it provides insight into the spatial manifestation of inequality in Canada, i.e. where inequality has statistically clustered within the country, and how those clusters have changed. Inequality has risen substantially in Canada in recent years, and this thesis sheds light on why and how this is occurring.

Indeed, Canada's earnings distribution is beginning to reflect that of its southern neighbour. Between 1996 and 2006, the top earners in Canada steadily increased their shares of total earnings, and especially after 2001 those in the middle and bottom of the distribution saw their real wages stagnate, if not decline, year-by-year. However, the comparison is not straightforward. In partial answer to the first and second research questions, my findings show that while very urbanized areas in Canada sustain significantly high growth rates in inequality levels as in the United States (Chakravorty, 1996, 2006), high rates of inequality in Canada are also geographically clustering in areas where energy-extraction industries are highly represented. This phenomenon certainly requires more investigation, as it appears that Canada's resource boom is not only "fuelling" the national economy through exports and financial investments (Williams, 2007), but is also playing a large role in fuelling economic inequality.

Meanwhile, the structure of inequality has changed as seen in the shifting contributions of its socioeconomic determinants across both urban areas and regions. For example, as Canada's manufacturing industries slowly contract relative to the resource and service sectors, the economy is paring down one of its major earnings equalizers. Furthermore, the lack of equitable growth in earnings is further evidenced by the fact that higher levels of median earnings no longer dampen inequality rates; top earners are effectively skewing the distribution sufficiently enough to render that factor insignificant. Some changes, however, are not taking place across both scales. Within Canada's urban areas, the female participation rate has emerged as a significant dampener of inequality. While this is not apparent yet at the broader regional scale, the dispersion of workers' educational attainment has emerged as a significant contributor to inequality across regions and not cities.

These findings have an array of policy implications at both the sub-national and national scales. They might offer a direct response to Dorling and Shaw's (2002) assessment that contemporary studies of human geography are less inclined or less adept in addressing social questions within policy frameworks. The empirical approach of this thesis is designed to pinpoint predictors of earnings inequality in a manner that provides hard evidence for policy decision making.

For instance, I examine inequality among individuals as dependent upon a number of place-based socioeconomic factors. As Partridge and Rickman (2006) point out, place-based policies may be more effective in targeting poverty versus person-based policies. A similar argument may be made in terms of targeting inequality, as it appears that local economic, geographic, and demographic structures impact local levels of individual earnings inequality. For example, a wider dispersion in levels of educational attainment is positively associated with higher levels of inequality across Canada, yet less so in urban areas (as a comparison of Chapters 4 and 5 suggests). As such, a person-based solution such as provincial or national grants for post-secondary education may not be as effective in lowering inequality as grants targeted to certain areas. Think of areas such as rural Alberta where a relatively high number of high-school students drop out without graduating, presumably lured by highly-paid unskilled employment (Cross and Bowlby, 2007).

Furthermore, another policy implication concerns the role of manufacturing in the Canadian economy. It is clear that higher manufacturing employment translates into lower levels of inequality, and yet its presence in Canada is declining. This topic can be connected to the deindustrialization literature starting with Bluestone and Harrison (1982) and others, and its association with inequality is often reflected in more recent literature concerned with neoliberalization (Bluestone and Harrison, 2000; Peck and Tickell, 2002; Levy and Temin, 2007). While the neoliberalization literature is primarily involved in addressing broad questions of complex changes in governance and institutions, manufacturing's role in the provision of high numbers of well-paying jobs is one area where policymakers might focus their efforts in combating inequality.

In researching this thesis I encountered particular limitations which could be addressed in further research. The first was the number of urban areas removed from Chapter 4's analysis, most of which were in British Columbia. This was due to drastic changes in their CSD boundaries, which left no option but to remove them from the sample. This might be addressed with finer-scaled data, which is not available at the RDC. Second, there is no continuity variable in the RDC's 2006 census sample which links recent industry categories to those before 1997 (e.g. NAICS 1997/2002 to SIC 1980). This poses problems for direct comparisons between censuses without going through each industrial individual code for each sample, a task for which time did not permit. Finally, Frenette et al. (2009) point out a major challenge for future studies of income inequality and Canadian income in general: the new option for census respondents to link their tax records (which are likely more accurate) to their census questionnaire. While the authors surmise the impact upon the earnings variable as negligible, other income variables may not be temporally consistent due to the change in survey methodology.

Finally, the findings of this thesis lead to many possibilities for future research. Intra-metropolitan analyses, while challenging in their application of mixed methodologies (Doussard et al., 2010), are needed to further parse the local-level drivers of inequality. Further analysis of the specific dynamics of inequality in rural and northern regions is warranted, and decomposition analyses such as Lu's et al. (forthcoming), but by geography and industry, would provide more insight as to the processes driving changes in inequality levels. Fundamentally, the trends, patterns and changes in inequality investigated in this thesis raise questions concerning the future of Canada's economic and societal wellbeing. What is required for Canada's labour market to once again generate equitably shared earnings? Will Canada suffer the outcomes associated with high levels of inequality? Where will those outcomes be concentrated?

REFERENCES

Anselin, L. (2005). *GeoDa: A workbook*. Urbana: Spatial Analysis Laboratory, University of Illinois.

Betz, M. (1972). The City as a System Generating Income Inequality. *Social Forces, 51*(2), 192-198.

Berman, E., Bound, J., and Machin, S. (1998). Implications of skills-biased technological change: international evidence. *Quarterly Journal of Economics, 112*(4): 1245 – 1279.

Blackburn, M. L., and D. E. Bloom. (1991). *The Distribution of Family Income: Measuring and Explaining Changes in the 1980s for Canada and the United States*. Cambridge, MA: National Bureau of Economic Research.

Bluestone, B., and Harrison, B. (1982). *The deindustrialization of America*. New York: Basic Books.

Bluestone, B., and Harrison, B. (2000). *Growing prosperity: The battle for growth with equity in the twenty-first century*. Boston: Houghton Mifflin Company.

Borjas, G. J., Freeman, R.B., & Katz, L.F. (1997). How much do immigration and trade affect labor market outcomes? *Brookings Papers on Economic Activity, 1997*(1), 1-90.

Breau, S. (2007). Income inequality across Canadian provinces in an era of globalization: explaining recent trends. *The Canadian Geographer - Le Geographe Canadien, 51*(1), 72-90.

Breau, S., and Rigby, D. (2009). International trade and wage inequality in Canada. *Journal of Economic Geography, 10*(1), 55 - 86.

Burt, J. E., Barber, G.M., and Rigby, D.L. (2009). *Elementary Statistics for Geographers* (3rd ed.). New York: The Guilford Press.

Carter, P. L. (2009). Geography, race, and quantification. *The Professional Geographer, 61*(4), 465 - 480.

Chakravorty, S. (1996). Urban inequality revisited - The determinants of income distribution in US metropolitan areas. *Urban Affairs Review, 31*(6), 759 - 777.

Chakravorty, S. (2006). *Fragments of inequality: social, spatial, and evolutionary analyses of income distribution*. New York: Routledge.

Cloutier, N. R. (1997). Metropolitan income inequality during the 1980s: the impact of urban development, industrial mix, and family structure. *Journal of Regional Science, 37*(3), 459-478.

Coates, B., Johnston, R., & Knox, P. (1977). *Geography and Inequality*. Oxford: Oxford University Press.

Conceição, P. a. G., J.K. (2001). Toward a new Kuznets hypothesis: theory and evidence on growth and inequality. In J. K. Galbraith, and M. Berner (Ed.), *Inequality and Industrial Change: A Global View*. New York: : Cambridge University Press.

Cowell, F. A. (1995). *Measuring Inequality* (2nd ed.). London: Prentice Hall/Harvester Wheatsheaf.

Cross, P. (2007a). GDP and employment growth. *Perspectives on labour and income, March*, 14 - 23. Statistics Canada - Cat. no. 75-001-XIE.

Cross, P. (2007b). Year-end review: westward ho! *Canadian Economic Observer, April*, 3.1 – 3.19. Ottawa: Statistic Canada – Cat. No. 11-010-XIB.

Cross, P. (2008). The role of natural resources in Canada's economy. *The Canadian Economic Observer, Nov,* 3.1 – 3.10. Ottawa: Statistics Canada Cat. No. 11-010-XIB.

Cross, P., and Bowlby, G. (2006). The Alberta economic juggernaut: The boom on the rose. *The Canadian Economic Observer, Sept.*, 3.1 - 3.12. Statistics Canada - Cat. no. 11-010.

DeVerteuil, G. and Wilson, K. (2010). Reconciling indigenous need with the urban welfare state? Evidence of culturally-appropriate services and spaces for Aboriginals in Winnipeg, Canada. *Geoforum (in press).*

Doussard, M., Peck, J., Theodore, N. (2010). After deindustrialization: Uneven growth and economic inequality in "postindustrial" Chicago. *Economic Geography, 85*(2), 183 - 207.

Dorling, D., and Shaw, M. (2002). Geographies of the agenda: public policy, the discipline and its (re)turns. *Progress in Human Geography 25*(5): 629 – 646.

Drennan, M. P. (2005). Possible sources of wage divergence among metropolitan areas of the United States. *Urban Studies, 42*(9), 1609-1620.

Ellis, M. (2009). Vital statistics. *The Professional Geographer, 61*(3), 301 - 309.

Erksoy, S. (1994). The effects of higher unemployment on the distribution of income in Canada: 1981 - 1987. *Canadian Public Policy - Analyse de Politiques, 20*(3), 218-328.

Ezcurra, R., Pascual, P., and Rapùn, M. . (2007). The spatial distributions of income inequality in the European Union. *Environment and Planning A, 39*, 869-890.

Filion, P. (2009). Growth and decline in the Canadian urban system: the impact of emerging economic, policy, and demographic trends. *GeoJournal*. DOI: 10.1007/s10708-009-9275-8.

Fortin, N., and Schirle, T. (2006). Gender Dimensions of Changes in Earnings Inequality in Canada. In D. A. G. a. J. R. Kesselman (Ed.), *Dimensions of Inequality in Canada*. Vancouver: UBC Press.

Frenette, M., Green, D.A., and Milligan, K. (2007). The tale of the tails: Canadian income inequality in the 1980s and 1990s. *Canadian Journal of Economics - Revue Canadienne d'Economique, 40*(3), 734-764.

Frenette, M., Green, D.A., and Milligan, K. (2009). Taxes, transfers, and Canadian income inequality. *Canadian Public Policy - Analyse de Politiques, 35*(4), 389 - 411.

Galbraith, J. K. (2002). A perfect crime: inequality in the age of globalization. *Daelalus, 131*, 11-25.

Gibson, D. (2007). *The spoils of the boom: incomes, profits, and poverty in Alberta*. Edmonton: Parkland Institute.

Gien, L. T. (2000). Land and sea connection: The East Coast fishery closure, unemployment and health. *Canadian Journal of Public Health, 91*(2), 121-124.

Gingrich, P. (2009). *Boom and Bust: the growing income gap in Saskatchewan*. Regina: Canadian Centre for Policy Alternatives.

Grier, P. (2005, June 14). Rich-poor gap gaining attention. *The Christian Science Monitor,* p. Online edition: www.csmonitor.com.

Harvey, D. (1973). *Social Justice and the City*. London: Edward Arnold.

Harvey, D. (2006). *Spaces of Global Capitalism: Towards a Theory of Uneven Geographical Development*. London and New York: Verso.

Heisz, A. (2007). *Income inequality and redistribution in Canada: 1974 - 2004*. Ottawa: Statistics Canada: Analytical Studies Branch Research Paper Series.

Hipp, J.R. (2007). Income inequality, race, and place: does the distribution of race and class within neighbourhoods affect crime rates? *Criminology 45*(3), p. 665 – 697.

Holt, J. B., and Lo, C.P. (2008). The geography of mortality in the Atlanta metropolitan area. *Computers, Environment, and Urban Systems, 32*(149 - 164).

Hsaio, C. (2003). *Analysis of Panel Data*. Cambridge, UK: Cambridge University.

Jenkins, S. P. (1995). Accounting for Inequality Trends - Decomposition Analyses for the UK, 1971 - 1986. *Economica, 62*(245), 29 - 63.

Jenkins, S. P. (2001). Ineqdeco: Inequality indices and decomposition by subgroups. (Version 1.6).

Jenkins, S. P. (2006). Sumdist.ado: Stata module to calculate summary statistics for income distributions: Boston College Dept. of Economics.

Johnson, S. (1995). More evidence on the effect of higher unemployment on the Canadian size distribution of income. *Canadian Public Policy - Analyse de Politiques, 21*(4), 423 - 428.

Kopczuk, W., Saez, E., Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data from 1937. *Quarterly Journal of Economics 125*(1): 91 – 128.

Korpi, M. (2008). Does size of lobal labour markets affect wage inequality? A rank-size rule of income distribution. *Journal of Economic Geography, 8*, 211-237.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy, 99*(3), 483 - 499.

Kwan, M., and Schwanen, T. (2009). Quantitative revolution 2: The critical (re)turn. *The Professional Geographer, 61*(3), 283 - 291.

Lake, R. W. (1994). *Urban Geography, 15*(7).

Levy, F., and Temin, P. (2007). Inequality and institutions in 20th century America, *MIT Department of Economics Working Paper No. 07-17.*

Lu, Y., Morissette, R., and Schirle, T. (forthcoming). The growth in family earnings inequality in Canada, 1980-2005. *The Review of Income and Wealth*.

Luffman, J. (2006). The core-age labour force. *Perspectives on labour and income, Sept.*, 5 - 11. Statistics Canada - Cat. no. 75-001-XIE.

MacDonald, R. (2007). *Canadian and U.S. real income growth pre and post 2000: A reversal of fortunes.* Ottawa: Statistics Canada. CAT # 11F0027MIE – N. 048.

MacLachlan, I., and R. Sawada. (1997). Measures of income inequality and social polarization in Canadian Metropolitan Areas. *The Canadian Geographer - Le Geographe Canadien, 41*(4), 377-397.

MacPhail, F. (2000a). What caused earnings inequality to increase in Canada during the 1980s? *Cambridge Journal of Economics, 24*, 153-175.

MacPhail, F. (2000b). Are estimates of earnings inequality sensitive to measurement choices? A case study of Canada in the 1980s. *Applied Economics, 32*, 845-860.

Martin, R. (2001). Editorial: the geographer as a social critic- getting indignant about income inequality. *Transactions of the Institute of British Geographers, 26*, 267 - 272.

Massey, D. (1979). In What Sense a Regional Problem? *Regional Studies, 13*(233-243).

Maxim, P. S., White, J.P., Beavon, D., and Whitehead, P.C. (2001). Dispersion and polerization of income among Aboriginal and Non-Aboriginal Canadians. *Canadian Review of Sociology, 38*(4), 465 - 476.

McCall, L. (2001). Sources of racial wage inequality in metropolitan labor markets: racial, ethnic, and gender differences. *American Sociological Review, 66*(4), 520-541.

Miller, H. J. (2004). Tobler's first law and spatial analysis. *Annals of the Association of American Geographers, 94*(2), 284 - 289.

Milligan, K. (2008). The evolution of elderly poverty in Canada. *Canadian Public Policy - Analyse de Politiques, 34*(Special Supplement on Private Pensions and Income Security in Old Age: An Uncertain Future), 79 - 94.

Moore, E. G., and Pacey, M.A. (2003). Changing income inequality and immigration in Canada, 1980 - 1995. *Canadian Public Policy - Analyse de Politiques*, *29*(1), 33-52.

Morrill, R. (2000). Geographic variation in change in income inequality among US states, 1970 - 1990. *Annals of Regional Science, 34*(1), 109-130.

Morris, M., and Western, B. (1999). Inequality in earnings at the close of the twentieth century. *Annual Review of Sociology*, *25*, 623 - 657.

Myles, J. (2003). Where have all the sociologists gone? Explaining economic inequality. *The canadian Journal of Sociology, 28*(4), 551-559.

Novak, M. (2000). Defining social justice. *First Things 108*, 11-13.

OECD. (2008). *Growing Unequal? Income distribution and poverty in OECD countries*. Paris: Organisation for Economic Co-operation and Development.

Osberg, L., and Lin, Z. (2000). How much of Canada's unemployment is structural? *Canadian Public Policy - Analyse de Politiques, 26*(Supplement), S141 - S157.

Osberg, L. (2008). *A quarter century of economic inequality in Canada: 1981 - 2006*. Ottawa: Canadian Centre for Policy Alternatives.

Osirio, R. G. (2007). Quantiles.ado: Stata module to categorize by quantiles: Boston College Dept. of Economics.

Partridge, M. and Rickman, D. (2006). *The geography of American poverty: is there a need for place-based policies?* Kalamazoo: W.E. Upjohn Institute for Employment Research.

Peck, J. (1992). Labor and agglomeration: Control and flexibility in local labor markets. *Economic Geography, 68*(4), 325-347.

Peck, J., and Tickell, A. (2002). Neoliberlizing Space. *Antipode, 34*(3), 380-404.

Pendakur, K., and Pendakur, R. (2007). Minority earnings disparity across the distribution. *Canadian Public Policy - Analyse de Politiques, 33*(1), 41 - 61.

Picot, G., and Myles, J. (2005). Income inequality and low income in Canada: an international perspective. Ottawa: Statistics Canada: 11F0019MIE No. 240.

Picot, G., Hou, F., & Coulombe, S. (2007). Chronic Low Income and Low-income Dynamics Among Recent Immigrants. In N. Analytical Studies Branch Research Paper Series (Ed.). Ottawa: Business and Labour Market Analysis Division.

Rigby, D., and Breau, S. (2008). Impacts of trade on wage inequality in Los Angeles: Analysis using matched employer-employee data. *Annals of the Association of American Geographers, 98*(4), 920-940.

Rogerson, P. A. (2006). *Statistical Methods for Geography* (2nd ed.). London: Sage Publications.

Ross, N., K. Nobrega, J. Dunn. (2001). Income segregation, income inequality and mortality in North American metropolitan areas. *GeoJournal*, *53*, 117-124.

Ross, N. A., C. Houle, J.R. Dunn, and M. Aye. (2004). Dimensions and dynamics of residential segregation by income in urban Canada, 1991 - 1996. *The Canadian Geographer - Le Geographe Canadien, 48*(4), 433-445.

Rutherford, T. (2000). Re-embedding, Japanese investment and the restructuring buyer-supplied relations in the Canadian automotive components industry during the 1990s. *Regional Studies 34*(8), 739-751.

Saez, E., and M.R. Veall. (2005). The evolution of high incomes in Northern America: Lessons from Canadian evidence. *American Economic Review, 95*(3), 831-849.

Sharpe, A., Arsenault, J-F., Harrison, P. (2008). *Why have real wages lagged labour productivity growth in Canada*. Ottawa: Centre for the Study of Living Standards.

Skaburskis, A., & Moos, M. (2008). The redistribution of residential property values in Montreal, Toronto, and Vancouver: examining neoclassical and Marxist views on changing investment patterns. *Environment and Planning A, 40,* 905 - 927.

Smith, D. M. (1994). *Geography and Social Justice*. Oxford: Blackwell.

Soroka, L. A. (1999). Male-female urban income distributions in Canada: The service sector in a dependency model. *Urban Studies, 36*(3), 563 - 574.

Statistics Canada. (2007). Standard Geographical Classification, Vol. 1: The Classification. In Cat. No. 12-571-XWE. Ottawa.

Statistics Canada. (2008). Earnings and Incomes of Canadians Over the Past Quarter Century, 2006 Census. In Cat. no. 97-563-X. Ottawa.

Stock, J. H., and Watson, M.W. (2008). Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica, 76*(1), 155-174.

Vinodrai, T. (2001). A tale of three cities: The dynamics of manufacturing in Toronto, Montreal, and Vancouver, 1976 - 1997. Ottawa: Statistics Canada: 11F0019 No. 177.

Wade, R. H. (2004). Is globalization reducing poverty and inequality? *World Development, 32*, 567 - 589.

Walks, R. A., & Maaranen, R. (2008). Gentrification, social mix, and social polarization: testing the linkages in large canadian cities. *Urban Geography, 29*(4), 296 - 396.

Wallace, I. (2002). *A geography of the Canadian economy.* Cambridge: Oxford University Press.

Ward, M. D., and Gleditsch, K.S. (2008). *Spatial Regression Models*. Thousand Oaks, CA: Sage.

Wilkinson, R. G., and Pickett, K.E. (2009). Income inequality and social dysfunction. *Annual Review of Sociology, 35*, 493-511.

Williams, C. (2007). Fuelling the economy. *Perspectives on labour and income, 8*(5), 16 - 24. Statistics Canada Cat. no. 75-001-XIE.

Wilson, S.V. and Butler, D.M. (2007). A lot more to do: the sensitivity of time-series cross-section analyses to simple alternative specifications. *Political Analysis 15*(2), 101-123.

Wooldridge, J. M. (2006). *Introductory Econometrics: A Modern Approach* (3rd ed.). Mason, OH: Thompson/South-Western.

Wyly, E. K., & Hammel, D. J. (2004). Gentrification, segregation, and discrimination in the American urban system. *Environment and Planning A, 36*, 1215 - 1241.

Wyly, E. K., Martin, D., Mendez, P., and Holloway, S. (2010). Transnational tense: immigration and inequality in American housing markets. *Journal of Ethnic and Migration Studies, 36*(2), 187 - 208.

Yalnizyan, A. (2007). *The rich and the rest of us: The changing face of Canada's growing gap*. Ottawa: Canadian Centre for Policy Alternatives.

Yalnizyan, A. (2009). *Opening Remarks: Income Inequality and the Pursuit of Prosperity*. Paper presented at the Walter Gordon Massey Symposium March 10, 2009.

APPENDICES

APPENDIX 3.1 PEARSON CORRELATIONS BETWEEN GINI OF EARNINGS AND OTHER EARNINGS INEQUALITY INDICES

Pearson Correlations between various earnings inequality indices, and the Gini of earnings

Inequality index	Correlation with Gini Index		
	1996	2001	2006
GE(-1)	0.65*	0.70*	0.77*
Mean Log Deviation [GE(0) or MLE)	0.93*	0.94*	0.97*
Theil [GE(1)]	0.95*	0.89*	0.84*
GE(2)	0.49*	0.31*	0.37*
Coefficient of Variation (CV)	0.56*	0.41*	0.54*
A(0.5)	0.98*	0.98*	0.97*
A(1)	0.94*	0.94*	0.97*
A(2)	0.68*	0.72*	0.79*

Note: Author’s tabulation of inequality in the 1996, 2001, and 2006 censuses.
Indices calculating using the *ineqdeco* program by Jenkins (2001).
The CV is equal to 2 * $(GE(2))^{0.5}$ (Cowell, 1995).
A star (*) marks those correlation coefficients significant at the 0.05 level.

This appendix shows the correlation between various indices that measure inequality of earnings and the Gini of earnings across Canada. There is a significant positive relationship between all indices and the Gini, which is expected. As detailed in Cowell (1995), the CV and the GE(2) are sensitive to disparities in earnings in the higher reaches of the earnings distribution, the Theil is sensitive to the middle-upper portion, and the MLE and Gini are more sensitive to disparities in the middle of the distribution. Thus, based on the above table, the impact of gains made by top earners is evident in comparing the GE(2) and the Gini: the decline in their correlations over the years corresponds with the fact that the GE(2) is picking up on rises in inequality due to increases in earnings of those at the top, while the Gini is sluggish due to lower contributions to inequality made by shifts in the middle of the distribution.

APPENDIX 3.2 MANUFACTURING INDUSTRIES GROUPED BY COMPETITIVE PROCESSES

Listed here is the breakdown of the manufacturing sectors into five categories according to the OECD's (1992) grouping of industries by their competitive processes. In most cases, the NAICS 2002 3-digit aggregates are sufficient, but in a few the aggregate had to be broken into its 4-digit components. This list was compiled from Statistics Canada's CANSIM table 379-0027. Note that code NAICS code 3399, "Other manufacturing," is not included for its ambiguity, and that NAICS code 3331, "Agricultural machinery..." is not included due to its lack of reliable data back to 1996.

Category 1: Resource intensive

Food manufacturing [311]

Beverage and tobacco product manufacturing [312]

Wood product manufacturing [321]

Paper manufacturing [322]

Petroleum and coal products manufacturing [324]

Non-metallic mineral product manufacturing [327]

Category 2: Labour intensive

Textile, clothing and leather product manufacturing [31X](2)

Primary and fabricated metal products manufacturing [33A](2)

Furniture and related product manufacturing [337]

Jewellery and silverware manufacturing [33991]

Category 3: Scale-intensive

Printing and related support activities [323]

Basic chemical manufacturing [3251]

Resin, synthetic rubber, and artificial and synthetic fibres and filaments manufacturing [3252]

Pesticide, fertilizer and other agricultural chemical manufacturing [3253]

Miscellaneous chemical product manufacturing [325A](2,3)

Plastics and rubber products manufacturing [326]

Motor vehicle manufacturing [3361]

Motor vehicle body and trailer manufacturing [3362]

Motor vehicle parts manufacturing [3363]

Railroad rolling stock manufacturing [3365]

Ship and boat building [3366]

Other transportation equipment manufacturing [3369]

Sporting and athletic goods, toy and games 3399X](2)

Category 4: Specialized supplier

~~Agricultural, construction and mining machinery manufacturing [3331]~~

Industrial machinery manufacturing [3332]

Commercial and service industry machinery manufacturing [3333]

Ventilation, heating, air-conditioning and commercial refrigeration equipment manufacturing [3334]

Metalworking machinery manufacturing [3335]

Other general-purpose machinery manufacturing [3339]

Electrical equipment, appliance and component manufacturing [335]

Category 5: Science-based

Pharmaceutical and medicine manufacturing [3254]

Engine, turbine and power transmission equipment manufacturing [3336]

Computer and electronic product manufacturing [334]

Aerospace product and parts manufacturing [3364]

Medical equipment and supplies manufacturing [3391]

APPENDIX 3.3 GENERAL CHARACTERISTICS OF THE CANADA-WIDE EARNINGS DISTRIBUTION

Median real earnings for the *target sample* across 287 census divisions in Canada (2002 dollars), by earnings quintile

	1996 Census Median Earnings	2001 Census Median Earnings	2006 Census Median Earnings	1996 – 2001 % Change	2001 – 2006 % Change	1996 – 2006 % Change
Top 20%	66,075.75	69,564	73,798.328	5.3%	6.1%	11.7%
60% - 80%	44,437.5	46,035	47,684	3.6%	3.6%	7.3%
40% - 60%	31,572	32,736	33,790.531	3.7%	3.2%	7.0%
20% - 40%	20,250	22,396.54	22,268.428	10.6%	-0.6%	10.0%
Bottom 20%	7,875	9,207	9,160.8301	16.9%	-0.5%	16.3%
Total of weighted observations	10,270,795	11,202,670	11,911,605	-	-	-

Source: Author's tabulation of the 1996, 2001, and 2006 censuses, with the CPI used as a deflator.
Earnings reported in each census correspond to the total calendar-year earnings of the previous year.
Although in 2006 there are 288 CDs, in order to maintain continuity I merge 1011 back into 1010.

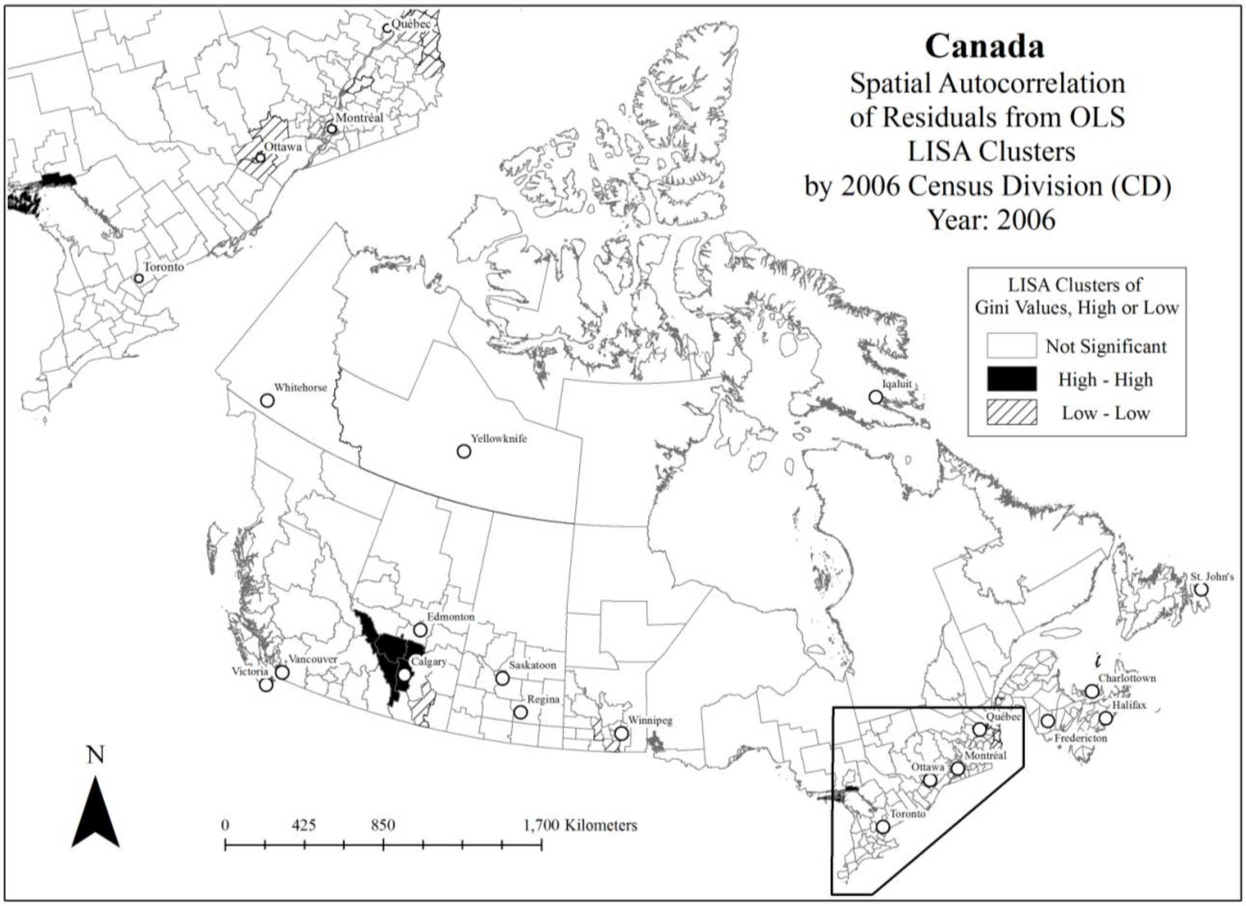
% Share of total earnings by earnings quintile across 287 census divisions, for the *target sample*

	1996 Census Total Earnings % Share	2001 Census Total Earnings % Share	2006 Census Total Earnings % Share
Top 20%	41.9%	43.0%	45.5%
60% - 80%	24.6%	24.2%	23.2%
40% - 60%	17.8%	17.0%	16.3%
20% - 40%	11.3%	11.4%	10.7%
Bottom 20%	4.3%	4.6%	4.3%
Total	100%	100%	100%

Source: Author's tabulation, earnings are reported for the previous calendar year.

APPENDIX 5.1 MAPPING THE RESIDUALS OF OLS REGRESSION

Spatial autocorrelation of residuals from OLS regression (including provincial fixed effects), 2006



Note: This map differs from those in the text in its legend: High-low and Low-high inequality CDs are not shown here.

As seen in this appendix's LISA map of residual clusters, spatial autocorrelation becomes a factor in the 2006 run of the OLS regression with provincial dummies. Although low, the Moran's *I* of 0.10 is significant at the 0.05 level. In identical runs of the 1996 and 2001 cycles, no significant clusters emerge.

There is one main cluster of high-high values in Alberta, west of Calgary, and two pockets of low-low clusters each surrounding Ottawa and Quebec City respectively.

Essentially, OLS regression does not account for the spatial patterning that occurs among inequality values in 2006.

APPENDIX 5.2 SPATIAL LAG MODEL REGRESSION RESULTS

Spatial lag model (SLM) regression results with earnings inequality as the dependent variable.

Independent variables	Spatial lag model (SLM)		
	1996	2001	2006
POP_DEN	-6.99e-06* (0.099)	6.929e-07 (0.871)	1.806e-06 (0.743)
ECON_DEV	-1.01e-06** (0.017)	-6.067e-06 (0.125)	7.795e-07* (0.062)
UNEMP	0.215** (0.000)	0.167** (0.000)	0.293** (0.000)
FEM_PART	0.011 (0.833)	0.032 (0.585)	-0.031 (0.661)
MANUF	-0.090** (0.000)	-0.093** (0.000)	-0.170** (0.000)
VIS_MIN	0.193** (0.000)	0.169** (0.000)	0.150** (0.000)
ABOR	0.033* (0.075)	0.032* (0.084)	-0.068** (0.006)
EDUC_RATIO	0.048* (0.098)	0.052* (0.079)	0.095** (0.014)
YOUNG_LF	0.005 (0.655)	0.000 (0.950)	0.042** (0.011)
SENIOR_LF	0.037** (0.000)	0.034** (0.002)	0.060** (0.000)
ρ (*spatial lag*)	0.010 (0.696)	0.073** (0.007)	0.123** (0.000)
λ (*spatial error*)	-	-	-
CONSTANT	0.347** (0.000)	0.310** (0.000)	0.250** (0.000)
Likelihood Ratio Test	0.145 (0.703)	6.856** (0.009)	13.185** (0.000)
Log likelihood	752.98	739.09	665.37
Breusch-Pagan Test	**61.588** (0.000)**	**21.917** (0.004)**	**36.953** (0.000)**
PROV. DUMMIES	NO	NO	NO
N	287	287	287

Note: p-values are in parentheses; two stars (**) mark those coefficients significant at the 0.05 level and one star (*) marks those coefficients marginally significant at the 0.10 level.

APPENDIX 5.5 THE SEM RESULTS, OMITTING 1 OUTLIER

The SEM results, omitting the largest outlier

Independent variables	Spatial Error model (SEM)		
	1996	2001	2006
POP_DEN	2.72e-06 (0.466)	4.04e-07 (0.144)	1.152e-05 (0.014)
ECON_DEV	-1.06e-06** (0.039)	-1.187e-06** (0.010)	1.400e-07 (0.754)
UNEMP	0.207** (0.000)	0.098** (0.008)	0.191** (0.000)
FEM_PART	-0.115** (0.026)	-0.096 (0.115)	-0.109 (0.133)
MANUF	-0.041** (0.011)	-0.063** (0.001)	-0.086** (0.001)
VIS_MIN	0.193 (0.203)	0.107** (0.008)	0.028 (0.466)
ABOR	0.028 (0.153)	0.052** (0.016)	-0.041 (0.107)
EDUC_RATIO	0.049* (0.084)	0.046 (0.118)	0.143** (0.000)
YOUNG_LF	-0.013 (0.205)	-0.011 (0.383)	0.189 (0.199)
SENIOR_LF	0.043** (0.001)	0.028** (0.025)	0.060** (0.000)
ρ (*spatial lag*)	-	-	-
λ (*spatial error*)	0.853** (0.000)	0.727** (0.000)	0.886** (0.000)
CONSTANT	0.420** (0.000)	0.424** (0.000)	0.349** (0.000)
Likelihood Ratio Test	63.188 (0.000)	47.169 (0.000)	100.750** (0.000)
Log likelihood	785.681	761.030	707.971
Breusch-Pagan Test	14.784 (0.140)	11.381 (0.329)	3.694 (0.960)
PROV. DUMMIES	NO	NO	NO
N	286	286	286

Note: p-values are in parentheses; two stars (**) mark those coefficients significant at the 0.05 level and one star (*) marks those coefficients marginally significant at the 0.10 level.