A Longitudinal Perspective on the French-English Stature Divide

John Cranfield, Kris Inwood, Asher Kirk-Elleker
University of Guelph

In this paper we bring together medical examinations during World War One with household information for 9000 soldiers located in the 1901 census. The analysis of height differentials indicates that Quebec-born and sons of labourers were especially short and that stature for all groups declined more or less continuously during the last third of the 19th century. Being francophone is the principle source of smaller stature in Quebec. A Oaxaca decomposition confirms that only a small share of francophone height penalty arises from personal characteristics and that some significant part of the explanation arises from non-random genetic characteristics and/or specific cultural practices not captured by our data.

I. Introduction

Men born in nineteenth-century Quebec were relatively short by North American standards at the time. The differential in mean stature between those born in Quebec and elsewhere in Canada was sizeable - roughly one inch (Cranfield and Inwood 2007, table 3). The framework of historical anthropometrics provides a way to understand the stature gap (Komlos 1996, 1998; Steckel 1995, 2008). Montreal and Quebec City were notoriously unhealthy. Moreover, Quebec francophones occupied the lower rungs of the socio-economic ladder in nineteenth-century Canada. If the urban environment in Quebec was less healthy, and social inequality was extreme, then childhood disadvantage in diet, disease and workload might have contributed to short stature in Quebec. Admittedly, a first attempt to control for these influences did not appreciably reduce the differential (Cranfield and Inwood 2007, tables 4, 5 and 7). In those estimations a measure of socio-economic status was either unavailable or proxied by the adult occupation of the son rather than the father's occupation, as we would prefer. In this paper we attempt to improve the analysis of the Quebec stature differential.

Improving the socioeconomic and other controls in a multivariate analysis is important because if diet, disease and related influences cannot explain the Quebec effect an alternate explanation of non-random genetic effects attracts attention. At the end of the nineteenth-century much of the Quebec population descended from a small population of migrants from western France who arrived before 1760. Subsequent generations saw some but proportionally not a great deal of immigration into Quebec. It is thus possible that the stature differential reflects non-random genetic characteristics of the founding population of Quebec. At issue is a widely-held assumption in the anthropometric literature that the stature potential of all populations is comparable. The Quebec population may turn out to be an important exception if we cannot explain a significant share of the stature differential with appropriate controls for social and environmental circumstance.

Improved data will help us sharpen the the analysis. Anthropometric analysis is almost always implicitly longitudinal in the sense that adult stature is interpreted as the outcome of net nutrition during childhood. And yet direct evidence of childhood circumstance is seldom available. Many studies use someone's occupation at the time of being measured, ie as an adult, as a proxy for socio-economic status of the household in which they were children. Clearly, the proxy is reliable only to the extent that sons followed their fathers' occupations. We get around this problem by developing direct evidence of intergenerational mobility.

II. Generating a Sample of Linked Records

We generate a sample of linked records using a multi-stage approach beginning with more than 64259 records or roughly 10% of the Canadian Expeditionary Force (CEF) for soldiers whose surname begins with the letter 'B'. Thornton and Olson (2001) argue that B surnames are ethnically representative of the Canadian population and that B is seldom misheard by enumerators or misread by data entry operators. Elimination of duplicate records and records lacking detail leaves us with 53960 records. Removing the foreign-born drops the number to 26990 records for men born in Canada before the 1901 enumeration. We link roughly 8% of these records to the census with an intensive manual process, and then use the linked records to guide the development of an algorithm that expands the linked sample to 28% of the available CEF records.

The manual process begins with a review of all Soundex-equivalent spelling variations to identify the close phonetic matches of the surnames. Thus we merge Berger, Burgher and Burgar with Burger, while Barchard and Brookwire with the same Soundex code are ignored. The pool of phonetically matched surnames is then sorted by first name to permit easy examination of census forenames. Spelling idiosyncrasies and the inconsistent use of nicknames, diminutives, multiple forenames and initials complicate the effectiveness of matching for both first and last names. Potential census name matches are then screened for date of birth and information on next of kin provided by the CEF record in order to identify the appropriate person.

This three-stage method - phonetic rendering to provide a pool of surnames which is narrowed by the use of given names, age and location, and then confirmed by day of birth and next of kin - generates most of our 'manual' links. We obtain a further 10% increase in the number of links with more intensive techniques including a search of the pool of potential surname matches for individuals with the same birth month as the soldier, searching for next of kin directly, or searching very carefully in a precise geographical area. Such techniques compensate to some extent for inconsistencies introduced during data recording and/or processing. This flexibility is helpful because Soundex does not adequately compensate for spelling variation generated by transcription errors (Hershberg et al 1976, 141). The alternate methods are used sparingly and with high standards for supporting evidence in order to accept a link.

Manual efforts link 2184 of 3124 records, a proportion comparing well to that of projects (Condran and Seaman 1981; Dillon 2006; Di Matteo 1997). We next use the manual links to guide the development of an automated linkage program for the entire set of 27000 records. We experiment with a variety of multiple-stage blocking algorithms involving as many as nine steps that gradually loosen the criteria in different ways. Unexpectedly we find that a one-step program with strict criteria returns fewer false links and almost as many correct links as the more complicated algorithms. Our preferred program requires day and month of birth to be exact and allows census age to vary two years in both directions. Because birthplace is unavailable from our census data we require people less than ten years old to be in the province of their birth in 1901. This eliminates many false positive links at the cost of insignificant bias towards 'stayers' since interprovincial (largely westward) movement of eastern Canadians was at a low ebb during the 1890s. We require an exact spelling of surnames because the easiest alternative, Soundex-equivalency, permits too many false links. The vagaries of spelling and prevalence of initials in lieu of given names in one or both sources preclude exact matching on forenames. However, not requiring any consistency on forename would lead to a high rate of false positives. After some experimentation we find that requiring at least one matching initial most closely mimics the linkage pattern of the manual process.

Our automated program for linkage, then requires a single pass that matches exact surname, sex, birth day and month, at least one initial of forename, birth year within two years and consistent province for those born 1891 or later. The crafting of these criteria balances the need for a substantial number of linked records with a minimization of false positive links. The automated program generates 7487 links from 26990 records, a 28% linkage rate. The potential for incorrect links may be estimated by running the program on manually-linked records. Only

1.5% of the links predicted by the program are incorrect in the sense of finding someone other than the person identified by the manual process. Combining the manual and automatic links gives us 9000 soldiers for whom family circumstances in 1901 are known.

The military records reasonably represent the Canadian male population since WWI enlistment was widespread among Canadian men (Morton 1994, 2005). Of course military-derived records inevitably over-represent young adult males (Fogel 1986; Lamm 1988). In Table 1 we compare the Canadian male population of military age, all Canadian-born CEF enlistees with B-surnames and records linked to 1901. These data indicate that enlistment drew disproportionately from the West, white-collar occupations, English-origin churches and the unmarried men but in other respects was reasonably representative. Slightly weak representation from Quebec francophones and those of Irish ethnicity contributes to the smaller share of Roman Catholics in the CEF.

The linked records differ slightly again. We are slightly more successful linking the 1890s cohort, farmers, white collar workers and the Ontario-born. Lower linkage rates for the 1870s cohort and general labourers undoubtedly reflect the greater imprecision of descriptions of older men and of occupations with higher rates of illiteracy and poverty (Guest 1987; Emery and McQuillan 1988). Those born in Quebec were more difficult to link because of the homogeneity of the French-Canadian name stock and a pattern of migration to English-speaking communities in which some francophones adopted English name equivalents and many enumerators were unfamiliar with French spelling (Dillon 2006). These differences between the military and the overall male population, and between linked records and the entire army, do not threaten the usefulness of the linked data but they are large enough to suggest the value of incorporating into our analysis some adjustment for demographic categories in order to adjust for compositional effects.

III. Empirical Analysis of Stature

Conventional economic indicators convey a mixed impression of the Canadian economy during the final third of the nineteenth century. The economy was growing, per capita income increased and industrialization created the same patterns of structural change as elsewhere (Green and Urquhart 1987; Urquhart 1986). It is therefore not surprising that Canada experienced the same deterioration of physical well-being visible in other countries during this period (Komlos 1987; Haines 2004; Steckel and Haurin 1997). Stature varied a great deal across the country; short stature in Quebec is the single most striking regional effect (Cranfield and Inwood 2007).

A starting point for most anthropometric research is the recognition that while genetic variation within a population typically produces a normal or near-normal distribution of height, differences between social groups and changes over time may reflect variation in the experience of 'net nutrition', or the cumulative effect of food consumption, work effort and disease exposure (Bielicki 1986; Fogel 1986; Fogel et al 1983; Steckel 1995; Tanner 1981). From this perspective the patterns of physical stature are understood as indirect evidence of physical well-being and its social and economic determinants. Because the patterns of physical growth are determined early in life, childhood circumstance is needed to comprehend adult stature. Nevertheless direct evidence of early life circumstance frequently is difficult to obtain resulting in the use of adult occupation as a proxy for occupation of the head of household during childhood. This is helpful only insofar as there is occupational persistence between generations.

The linked data allow us to compare, in Table 2, the occupations given by WWI enlistees at the time of attestation with the occupation reported in 1901 by the future soldier's household head. Using broad occupational classes we find in panel (a) that at the time of enlistment three-quarters of farmers, one-third of those with white collar occupations and even lower proportions of other groups were following their father's 1901 occupation. Overall more than two-thirds of the soldiers were not in the same occupational class as their fathers in 1901. Some differences may arise from the fact that we observe father and son at different points in their life course. We explore this in a partial way by examining soldiers by age at enlistment in panels (b) and (c). The men reporting white-collar occupations were more likely to follow their fathers as they grew older, and the older unskilled labourers were less likely to do so but in other respects the patterns across age classes are similar. We conclude that life-cycle considerations account for only a small portion of the apparent intergenerational differences. Collectively, these data describe considerable deviation, with the exception of the farming population, from a pattern of occupational persistence even with broadly defined classes.

Another complication arises from the CEF standards of physical fitness. The examination of fitness was multi-dimensional of which one element was a preference for a minimum height of 63 inches. Admittedly the requirement was not enforced systematically: "in the scramble for men, many battalions had not been fussy about whom they accepted. Doctors ignored physical defects and passed men as fit because their colonel needed them" (Morton

1994, 60). However men who were less fit, some of them shorter than 63 inches, were rejcted. Soft or 'fuzzy' truncation of this nature is common (Fogel et al 1982). Even an inconsistently applied truncation requires a maximum likelihood truncated regression technique that ignores all observations with height less than the truncation point and estimates the lower tail of the distribution from the observations above 63 inches and a presumption of normality (Komlos 1990; Fogel 1986; Komlos and Kim 2004). In this case the frequency distribution of recorded heights follows a near-normal pattern (Figure 1). Estimation with ordinary least squares generates results very similar to those of the truncated model – as expected given the apparent normality of the data. Nevertheless we report truncated regression results where possible because in principle it is appropriate to recognze the potential for truncation effects.

We report in Table 3 a standard equation that explains height as a function of birth cohort, birth location, occupation using alternate sources of occupational information, and dichotomous variables for age of those less than 21 years of age. Columns on the left are based on the soldier's occupation as reported in CEF enlistment papers while the right column uses occupation of household head in 1901. Surprisingly, given the extent of occupational change from father to son, a comparison of the two columns reveals no important differences in the pattern of estimated co-efficients. From an inferential perspective the assumption of occupational persistence appears to have limited impact on the Quebec effect or anything else. One possible explanation is that occupational designations are sufficiently crude and imprecise indicators of influences on stature that the degradation of information from use of the proxy is of limited marginal consequence.

The next step is to take advantage of additional information that arises from the census linkage. The census reports mother tongue and whether the birthplace was rural or urban. It is possible that the Quebec effect arises entirely or in part from the francophone majority in that province and that those born in urban areas were systematically shorter because the environment was intrinsically less health. An estimation capturing these effects is reported in Table 4. Introducing mother tongue into the estimation markedly reduces the Quebec height disadvantage. It also increases the stature superiority of the Maritime or Atlantic coastal region which had a substantial francophone minority. French-speakers who were largely of francophone descent were systematically shorter. Separate regional regressions (not shown) confirm that the francophone disadvantage was largest in Quebec (-4.9 cm against -4.0 cm in Ontario and -2.8 cm in the Atlantic region). It is clear that much, although not all, of the Quebec height penalty originates with its francophone community and that, while being francophone was disadvantageous outside of Quebec, stature inequality by language/ethnic group was greatest in that province.

The results also confirm that the rural-born were taller as adults although, unexpectedly, the effect is not statistically significant. It is worth noting that enumerators were given no guidelines as to what constituted urban and rural implying that the 'rural' influence captured by this variable may not have been consistently specified (Canadian Families Project 2002 , 23). Nevetheless, in spite of this imprecision, the potentail importance of urban ill-health and the possibility that it varied across the country recommends another estimation that estimates separately on people born in urban and rural environments. This experiment reported in Table 5 leads to a number of useful insights. The birth cohort co-efficients for the urban-born show that the stature decline becomes signficant in the late 1880s (previously this decline was confined to the late 1890s). In contrast, none of the birth cohort co-efficients are statistically significant in the rural equation. The regional effects also diverge for the rural- and urban-born. The Quebec effect is substantial in the urban equation but the rural-born in Quebec were no different than those in Ontario.

The revelation that the Quebec effect arose primarily from urban areas and is associated with French as a mother tongue, and that the height penalty for urban birth was larger among francophones recommends a standard Oaxaca decomposition of stature differences by mother tongue based on the regression reported in Table 6. Here we estimate with ordinary least squares rather than truncated regression to simplify the decomposition (given the robustness of estimated coefficients across models estimated with OLS and maximum likelihood truncation regression models, use of OLS in this instance should not create substantial bias). The coefficients in Table 6 are not of immediate interest. Rather we are interested in how much of the average stature difference between language groups is explained by any tendency for francophones to have been born disproportionately in less healthy areas (urban environment, Quebec), to fathers with lower socio-economic status occupations and in less healthy cohorts (eg late 1890s). The calculations summarized in Table 7 indicated that the 'explained' share of mean stature difference is roughly 10% or 20% depending on whether defferent endowments are weighted by co-efficients from anglophone or francophone equation. The important conclusion, of course, is that even the most sensitive estimation that we can assemble explains only a small share of the francophone stature difference.

IV. Interpretation

The various estimates agree that the various occupational groups tended toward different mean heights. The use of occupation of the soldier himself and of the head of the family in which he was growing up have broadly similar patterns. Farmers and white collar workers were tall. Labourers tended to be roughly 2 centimeters shorter. This pattern reflects a familiar pattern of social inequality (Darroch 1983). Many labourers experienced low wages, frequent underemployment and congested living conditions (Emery and Levitt 2002). The skills and resources of professionals, merchants and other white collar workers brought a more reliable income and healthier living conditions for their families. Those growing up on farms were taller because the relative price of food was lower and the lower-density environment implied a reduced exposure to disease (Fogel 1986). Of course some although by no means all of the farmers were relatively wealthy.

We had anticipated that the same logic would apply to those born in rural areas. In fact the co-efficients on rural birth were not significantly different than zero (albeit with a positive sign). Moreover the constant terms arising from estimations on rural and urban populations were very close (Table 5). One explanation is that not every father born into rural Canada was a farmer; even among farmers many were relatively poor. A second line of explanation might be that the urban/rural health gap had already begun to close (Bielicki 1986; Thibault et al 1985, 113). Pelletier et al (1997) argue that urban/rural death rates began to converge after 1870; Harris and Mercier (2002) also suggest that suburban areas were nearly as unhealthy as the cities. Admittedly the precise definition of rural vs urban birth is unclear. Nevertheles, there would appear to be some information value in the 1901 census report of rural vs urban birth since many of the co-efficients reported in Table 6 differ by language group.

Much of the Canadian literature for this period focuses on inequality and poor health and nutrition in the cities (Darroch 1994). Steckel and Haurin (1994, 128) also report that secular changes in stature for the Ohio National Guard arose almost entirely due to changes within the urban population. Nevertheless, the evidence of Table 5 suggests that the health of rural Canadians deteriorated noticeably at the end of the nineteenth century. Dividing the population into rural and urban groups also reveals that, among the urban-born, the Quebec penalty and the penalty for being francophone is even larger than in previous estimates. The joint penalty for being born francophone in a Quebec city is remarkably large. Interestingly, and in stark contrast to Quebec, the height advantage of the Atlantic region disappears entirely if we restrict focus to the cities only.

The regional effects are marked and for the most part conform to expectations. Those from Atlantic Canada were significantly taller, in spited of having lower incomes, because lower population densities and easier access to dietary protein (due to the prevalence of fishing and animal husbandry) made up for any shortfall in wages (Norrie et al 2002, 229). Those born in Western Canada were tallest, despite a higher cost of living, because wages were higher, population densities were low and cities developed later than in the east (Morton 2005; Emery and Levitt 2002; Minns and McKinnon 2007). Davies (2005) suggests that urbanization in British Columbia was informed by knowledge of sanitation that was unavailable at the time of building and organizing the older cities in the east. Low population density, coupled with proximity to significant agricultural production in western Canada also points to food availability as an important factor underlying the western Canadian height premium. The western height differential continued to exist well into the twentieth century (Jette 1980).

Those born in Quebec were noticeably shorter than other Canadians. It is generally thought that public health in Quebec's two major cities were among the worst in North America (Copp 1974; McInnis 2000). Brown and Cook (1974, 128) suggest that industry in Quebec was largely small-scale industrialization which employed a large number of hands at very low wages; many of these jobs were filled by women and children. Yool (2006) presents evidence that food prices were higher in Quebec than in Ontario. The census linkage provides a more nuanced picture. As noted, specifying the language variable reduces the height differential in Quebec roughly from 3 to 1 centimeter. Much of the variation that might otherwise be province of birth is actually due to mother tongue; those whose first language was French were more than three centimeters shorter than average.

The French language coefficient was the largest of all influences investigated in our analysis and is consistent with the report by Ward and Ward (1984) that French speakers lived in conditions which were much poorer than those in Anglophone districts of Montreal. Thornton and Olsen (2001) present evidence suggesting that babies born into Francophone homes were, on average, weaned earlier than those in both Irish Catholic and Anglo-Protestant homes and were thus more susceptible to the range of diarrhoeal diseases which were the most frequent cause of infant mortality. Thornton and Olson also suggest that political discrimination and the resulting unequal distribution of municipal resources exacerbated the health inequalities within Montreal.

Mercier and Boone (2002) suggest that childcare practices contributed to mortality differentials within Ottawa as well. Separate regressions (not reported) imply a substantial francophone height penalty for each of the provinces.

We have not been successful in explaining the stature differential by language group. On the one hand it is clear that a disadvantaged socio-economic status and other measurable characteristics contributed to the francophone height penalty. And yet the decomposition exercise reported in the last table makes clear that 80% or more of the differential remains unexplained.

VI. Conclusion

The use of census-linked military-derived height data to examine social conditions in Canada at the turn of the twentieth century has advanced our understanding of the determinants of stature in nineteenth-century Canada. The linkage itself was quite successful. It provides a large enough group to test the sample as a whole and to break the sample down in a variety of ways that yield a more nuanced picture of certain subpopulations. We have some confidence in the accuracy of the links as they conformed very closely to the painstakingly-generated manual matches. There is, however, always room for improvement. Further refinement of the automated linkage program might produce a higher linkage rate although an even more obvious source of improvement would be a manual review of unresolved links after the program has been run. Linkage has provided a more accurate picture of family income (as revealed through the occupation of the head of household), a revised understanding of the height penalty in Quebec and also led to new questions regarding the relative health benefits of urban and rural upbringing. The penalty for growing up as a francophone accounts for most of the Quebec height disadvantage. Those who grew up speaking French were significantly shorter than those who spoke English outside of Quebec as well – reflecting both the marginalization of French-speakers outside of Quebec and the privileged position of the Anglophone minority inside the province. The failure to explain more than a fifth of the francophone height penalty points toward influences beyond the measure of census and military enlistment records. Everyday cultural practices that were particular to language/ethnic groups (eg the duration of breast-feeding) may have a role, as may some non-random genetic traits of those who descended from Quebec's founding population.

BIBLIOGRAPHY

Tadeusz Bielicki, "Physical Growth as a Measure of the Economic Well-Being of Populations: The Twentieth century", pp. 283-306 in F. Falkner and J.M. Tanner, eds, Human Growth, vol. III (New York: Plenum Press, 1986 2nd ed.).

Robert Craig Brown and Ramsay Cook, Canada 1896-1921: A Nation Transformed. (Toronto: McClelland and Stewart Limited, 1974).

Canadian Families Project, Eric W. Sager, director. The National Sample of the 1901 Census of Canada User's Guide. (University of Victoria, 2002).

Gretchen A. Condran and Jeff Seaman, "Linkage of the 1880-81 Philadelphia Death Register to the 1880 Manuscript Census: A Comparison of Hand- and Machine-record Linkage Techniques." Historical Methods 14 no. 2 (1981), pp. 73-84.

Terry Copp, The Anatomy of Poverty. (Toronto: McClelland and Stewart Limited, 1974).

John Cranfield and Kris Inwood. "The Great Transformation: A Long-Run Perspective on Physical Well-Being in Canada." Economics and Human Biology 5, no. 2 (2007), pp. 204-228.

Adam Crerar, "Ontario and the Great War." pp. 230-271 in David Mackenzie, ed. Canada and the First World War. (Toronto: University of Toronto Press, 2005).

Gordon Darroch, "Occupational Structure, Assessed Wealth and Homeowning during Toronto's Early Industrialization, 1861-1899." Histoire Sociale 16 no. 32, (1983), pp. 381-410.

Gordon Darroch, "Semi-Automated Record Linkage with Surname Samples: A Regional Study of "Case Law" linkage, Ontario 1861-1871." History and Computing 14 (2006), pp. 153-184.

Gordon Darroch and Lee Soltow, Poverty and Inequality in Victorian Ontario (Toronto: University of Toronto Press, 1994).

Megan J. Davies, "Night Soil, Cesspools, and Smelly Hogs on the Streets: Sanitation, Race, and Governance in Early British Columbia." Histoire Sociale 38 no. 75, (2005), pp. 1-35.

Lisa Dillon, "Challenges and Opportunities for Census Record Linkage in the French and English Canadian Context." History and Computing 14 (2006), pp. 185-212.

George Emery and Kevin McQuillan, "A Case Study Approach to Ontario Mortality History: the Example of Ingersoll, 1881-1972." Canadian Studies in Population 15 no. 2 (1988), pp. 135-158.

J.C. Herbert Emery and Clint Levitt, "Cost of Living, Real Wages and Real Incomes in Thirteen Canadian Cities, 1900-1950." Canadian Journal of Economics 35 no. 1 (2002), pp. 115-137.

R.W. Fogel, "Physical Growth as a Measure of the Economic Well-Being of Populations", pp. 263-281 in F. Falkner and J.M. Tanner, eds, Human Growth, vol. III (New York: Plenum Press, 1986 2nd ed.).

R.W. Fogel, Stanley L. Engerman, Roderick Floud, Gerald Friedman, Robert A. Margo, Kenneth Sokoloff, Richard H. Steckel, T. James Trussell, Georgia Villaflor, and Kenneth W. Wachter, "Secular Changes in American and British Stature and Nutrition." Journal of Interdisciplinary History 14 (1983), pp. 445-481.

R.W. Fogel, Stanley L. Engerman and James Trussell, "Exploring the Uses of Data on Height." Social Science History 6 no. 4 (1982), pp. 401-421.

David Gagan and Rosemary Gagan, "Working-Class Standards of Living in Late-Victorian Urban Ontario." Journal of the Canadian Historical Association no. 1 (1990), pp. 171-193.

Granatstein, J.L. and Hitsman, J.M. Broken Promises: A History of Conscription in Canada. (Toronto: Oxford University Press, 1977).

Avery M. Guest. “Notes from the National Panel Study: Linkage and Migration in the Late Nineteenth Century.” Historical Methods 20 no. 2 (1987), pp. 63-77.

Michael R. Haines. “Growing Incomes, Shrinking People - Can Economic Development Be Hazardous To Your Health? Historical Evidence For The United States, England, And The Netherlands In The Nineteenth Century. Social Science History 28 (2004), pp. 249-270.

Richard Harris and Michael Mercier, “How Healthy were the Suburbs?”, Urban History Review 31 no. 6 (2005), pp. 767-798.

Theodore Hershberg, Alan Burstein, and Robert Dockhorn, “Record Linkage.” Historical Methods Newsletter 9 nos. 2 and 3 (1976), pp. 137-163.

Kris Inwood and Jim Irwin, “Land, Income and Regional Inequality: New Estimates of Provincial Incomes and Growth in Canada, 1871-1891.” Acadiensis 31 no. 2 (2002), pp. 157-184.

Maurice Jette, “A Regional Analysis of Anthropometric Measurements in a Canadian Employee Population Study.” Canadian Journal of Public Health 71 (1980), pp. 32-36.

Michael B. Katz, The People of Hamilton, Canada West. (Cambridge: Harvard University Press, 1975).

John Komlos, “The height and weight of West Point cadets: dietary change in antebellum America”. Journal of Economic History 47 (1987), pp. 897-927.

-----, “Anomalies in Economic History: Toward a Resolution of the" Antebellum Puzzle". Journal of Economic History 56 (1996), pp. 202-214.

-----, “Shrinking in a Growing Economy? The Mystery of Physical Stature during the Industrial Revolution”. Journal of Economic History 58 (1998), pp. 779-802.

-----, “How to (and How Not to) Analyze Deficient Height Samples.” Historical Methods 37 no. 4 (2004), pp. 160-173.

----- and Joo Han Kim, “Estimating Trends in Historical Heights.” Historical Methods 23, no. 3 (1990), pp. 116-120.

Doron Lamm, “British Soldiers of the First World War: Creation of a Representative Sample.” Historical Social Research 13, n. 4 (1988), pp. 55-98.

Livio Di Matteo, "The Determinants of Wealth and Asset Holding in Nineteenth-Century Canada", Journal of Economic History 57 no. 4 (Dec. 1997), pp. 907-934.

Marvin McInnis, “The Population of Canada in the Nineteenth Century.” pp. 371-432 in Michael R. Haines and Richard H. Steckel eds. A Population History of North America. (Cambridge: Cambridge University Press, 2000).

Michael Mercier and Christopher Boone, “Infant Mortality in Ottawa, Canada, 1901: assessing cultural, economic and environmental factors.” Journal of Historical Geography 28 no. 4 (2002), pp. 486-507.

Chris Minns and Mary MacKinnon, “The Costs of Doing Hard Time: a Penitentiary-based Regional Price Index for Canada, 1883-1923.” Canadian Journal of Economics 40 no. 2 (2007), pp. 528-560.

Desmond Morton, “Supporting Soldiers’ Families: Separation Allowance, Assigned Pay, and the Unexpected.” pp. 194-229 in David Mackenzie, ed. Canada and the First World War. (Toronto: University of Toronto Press, 2005).

Desmond Morton, When Your Number's Up: the Canadian Soldier in the First World War. (Toronto: Vintage Canada, 1994).

K.H. Norrie, Doug Owram and John Charles Herbert Emery. A History of the Canadian Economy. (Scarborough: Thomson/Nelson, 2002).

Francois Pelletier, Jacques Legare and Robert Bourbeau. "Mortality in Quebec During the Nineteenth Century: From the State to the Cities." Population Studies 51, n. 1 (March, 1997), pp. 93-103.

Richard H. Steckel, "Stature and the Standard of Living", Journal of Economic Literature 33 no. 4 (Dec. 1995), pp. 1903-1940.

Richard H. Steckel, "Biological measures of the standard of living, Journal of Economic Perspectives 22 (2008), pp. 129–152.

Richard H. Steckel and Donald R. Haurin, "Health and Nutrition in the American Midwest: Evidence from the Height of National Guardsmen, 1850-1910", pp. 117-128 in John Komlos, ed., Stature, Living Standards and Economic Development (Chicago: University of Chicago Press, 1994).

J.M. Tanner, A History of the Study of Human Growth (Cambridge: Cambridge University Press, 1981).

Harold W. Thibault, Lise LaPalme, Richard Tanguay and Arto Demirjian. "Anthropometric Differences Between Rural and Urban French-Canadian Schoolchildren." Human Biology 57, n. 1 (1985), pp. 113-129.

Patricia Thornton and Sherry Olson. "A Deadly Discrimination Among Montreal Infants, 1860-1900." Continuity and Change 18 (2001), pp. 95-135.

W. Peter Ward and Patricia C. Ward, "Infant Birth Weight and Nutrition in Industrializing Montreal", American Historical Review 89 (1984), pp. 324-345.

James Yool, "An Anthropometric Study of the Standard of Living and Degree of Marginality in Ontario and Quebec, 1870-1900." Unpublished MA Paper, University of Guelph Department of Economics, 2006.

Table 1: Characteristics of the Linked Sample and Canadian Enlistees (% distribution)

	Cdn-born Men	CEF Records	Linked Records
Decade of Birth			
1860s	23.0	0.4	0.1
1870s	10.6	6.7	2.3
1880s	23.1	20.5	18.9
1890s	39.2	71.3	78.1
1900s	4.1	0.8	0.6
Marital Status			
Married	25.1	13.9	11.2
Not married	74.9	84.1	88.8
Occupation			
Farmers	25.3	24.7	27.3
White collar	15.4	21.4	24.7
Skilled	12.6	18.5	17.2
Unskilled	20.6	20.3	17.6
Labourer	12.1	11.1	8.1
Other	13.9	3.9	5.0
Region of Birth			
Atlantic	20.1	18.8	19.0
Quebec	26.7	25.7	22.5
Ontario	47.8	45.1	50.4
West	5.0	10.4	8.1
Religion			
Baptist	6.2	6.8	7.9
Church of England	15.4	18.2	18.4
Methodist	12.6	16.0	19.1
Presbyterian	16.1	17.2	20.6
Roman Catholic	45.0	35.3	28.2
Other	4.4	6.3	5.7
N	48939	26833	8201

Source: Canadian Families Project, 1901 Census Sample; CEF database (see text)

Table 2: Comparison of CEF Soldier's Occupation to Head's Occupation in 1901

a. All enlistees

		Head's Occupation Code (from census records)					
		Farmer	White collar	Skilled	Unskilled	Labour	Other
Recruit's occupation code (from CEF records)	Farmer	1634	82	122	104	158	20
		0.77	0.04	0.06	0.05	0.07	0.01
	White collar	367	603	357	255	86	67
		0.21	**0.35**	0.21	0.15	0.05	0.04
	Skilled	327	166	335	213	144	30
		0.27	0.14	**0.28**	0.18	0.12	0.02
	Unskilled	335	144	237	346	206	32
		0.26	0.11	0.18	**0.27**	0.16	0.02
	Labour	188	38	89	125	132	9
		0.32	0.07	0.15	0.22	**0.23**	0.02
	Other	75	150	52	47	18	14
		0.21	0.42	0.15	0.13	0.05	**0.04**

b. Enlistees 18-25 years of age

		Head's Occupation Code (from census records)					
		Farmer	White collar	Skilled	Unskilled	Labour	Other
Recruit's occupation code (from CEF records)	Farmer	1387	73	107	83	130	16
		0.77	0.04	0.06	0.05	0.07	0.01
	White collar	237	370	269	190	71	35
		0.20	**0.32**	0.23	0.16	0.06	0.03
	Skilled	185	115	213	157	96	23
		0.23	0.15	**0.27**	0.20	0.12	0.03
	Unskilled	208	99	171	258	141	21
		0.23	0.11	0.19	**0.29**	0.16	0.02
	Labour	124	28	70	94	91	7
		0.30	0.07	0.17	0.23	**0.22**	0.02
	Other	64	119	46	34	12	11
		0.22	0.42	0.16	0.12	0.04	**0.04**

c. Enlistees 26 years of age and older

		Head's Occupation Code (from census records)					
		Farmer	White collar	Skilled	Unskilled	Labour	Other
Recruit's occupation code (from CEF records)	Farmer	239	9	14	19	26	3
		0.77	0.03	0.05	0.06	0.08	0.01
	White collar	127	225	83	61	15	31
		0.23	**0.42**	0.15	0.11	0.03	0.06
	Skilled	140	49	118	52	47	7
		0.34	0.12	**0.29**	0.13	0.11	0.02
	Unskilled	123	43	63	81	62	11
		0.32	0.11	0.16	**0.21**	0.16	0.03
	Labour	60	6	18	28	32	2
		0.41	0.04	0.12	0.19	**0.22**	0.01
	Other	8	21	4	5	4	2

Table 3 - Truncated Regression Results for Linked Soldiers with Alternate Sources for Occupation

		Using soldier's occupation at enlistment			Using father's occupation from 1901 census		
		Coefficient	**z**	**P>z**	**Coefficient**	**z**	**P>z**
Cohort	Born pre-1880	Omitted			Omitted		
	Born 1880-85	-0.4617	-0.68	0.498	-0.2068	-0.26	0.796
	Born 1885-89	-0.9280	-1.54	0.124	-0.8364	-1.17	0.24
	Born 1890-95	-1.1893	-2.04	0.041	-1.0251	-1.48	0.138
	Born 1895+	-2.2411	-3.75	0	-1.9317	-2.77	0.006
Birthplace	Maritimes	0.7671	3.48	0.001	0.5904	2.57	0.01
	Quebec	-3.2379	-13.74	0	-3.3485	-13.52	0
	Ontario	Omitted			Omitted		
	West	1.6903	5.62	0	1.4941	4.74	0
Occupation	Farmer	Omitted			Omitted		
	White collar	0.1931	0.79	0.429	-0.1683	0.63	0.528
	Skilled	-1.3618	-5.05	0	-1.3249	-5.01	0
	Unskilled	-1.3437	-5.03	0	-1.2375	-4.58	0
	Labourer	-2.1233	-5.87	0	-2.1557	-6.7	0
	Other	0.8079	1.97	0.049	-0.7166	1.2	0.228
Age	Under19	-0.5979	-1.79	0.074	-0.4690	-1.39	0.166
	Age 19	0.7114	2.03	0.042	0.7738	2.17	0.03
	Age 20	0.2956	0.95	0.34	0.3806	1.19	0.234
	21 and over	Omitted			Omitted		
	Constant	169.9068	285.95	0	169.8150	248.29	0
	Sigma	6.3937	86.88	0	6.3712	82.5	0
N		7426			6692		

Table 4 - Truncated Regression Results for Linked Records with All Census Variables

		Coefficient	z	P>z
Density	Rural	0.3109	1.35	0.176
	Urban	Omitted		
Cohort	Born pre-1880	Omitted		
	Born 1880-85	-0.2352	-0.3	0.766
	Born 1885-89	-0.9074	-1.29	0.197
	Born 1890-95	-0.8854	-1.3	0.194
	Born 1895+	-1.5991	-2.32	0.02
Birthplace	Maritimes	0.8667	3.7	0
	Quebec	-0.7546	-2.33	0.02
	Ontario	Omitted		
	West	1.4861	4.73	0
Head's Occupation	Farmer	Omitted		
	White collar	-0.2017	-0.67	0.503
	Skilled	-1.2509	-4.26	0
	Unskilled	-1.1899	-4.08	0
	Labourer	-2.0542	-6.24	0
	Other	-0.5380	-0.87	0.385
Mother Tongue	English	Omitted		
	French	-3.9256	-11.45	0
	Other	-0.9107	-1.59	0.111
Age	Under19	-0.9319	-2.75	0.006
	Age 19	0.4172	1.16	0.244
	Age 20	0.1647	0.52	0.606
	21 and over	Omitted		
	Constant	169.6577	240.63	0
	Sigma	6.2702	82.08	0
N	6494			

Table 5 Truncated Regression Results with Urban/Rural Split

		Rural Only			Urban Only		
		Coefficient	**z**	**P>z**	**Coefficient**	**z**	**P>z**
Cohort	Born pre-1880	Omitted			Omitted		
	Born 1880-85	0.4736	0.48	0.63	-1.2387	-0.93	0.352
	Born 1885-89	-0.0282	-0.03	0.974	-2.0827	-1.78	0.075
	Born 1890-95	-0.3408	-0.4	0.689	-1.6406	-1.45	0.148
	Born 1895+	-0.9435	-1.11	0.269	-2.5898	-2.19	0.028
Birthplace	Atlantic	1.1618	4.29	0	0.0505	0.11	0.915
	Quebec	-0.5325	-1.33	0.185	-0.9511	-1.73	0.084
	Ontario						
	West	1.4725	3.66	0	1.5637	3.09	0.002
Head's Occupation	Farmer	Omitted			Omitted		
	White collar	-1.1999	-2.91	0.004	1.6106	2.03	0.042
	Skilled	-1.2079	-3.2	0.001	0.0574	0.07	0.942
	Unskilled	-1.1947	-3.35	0.001	0.0556	0.07	0.945
	Labourer	-1.8759	-4.95	0	-1.0329	-1.16	0.246
	Other	-0.4752	-0.48	0.634	0.7875	0.75	0.455
Mother Tongue	English	Omitted			Omitted		
	French	-3.7811	-9.54	0	-4.6752	-6.73	0
	Other	-0.0860	-0.13	0.893	-3.1595	-2.55	0.011
Age	Under 19	-0.9285	-2.21	0.027	-0.9124	-1.5	0.134
	Age 19	0.2976	0.66	0.508	0.7089	1.13	0.26
	Age 20	-0.1103	-0.29	0.775	0.6840	1.18	0.239
	21 and Over	Omitted			Omitted		
	Constant	169.3265	201.56	0	169.2422	128.26	0
	Sigma	6.1335	67.42	0	6.4802	47.1	0
N		4183			2311		

Table 6: OLS Regression Results, by Mother Tongue (no adjustment for truncation)

		English mother tongue			**French mother tongue**		
		Coefficient	**t**	**P>t**	**Coefficient**	**t**	**P>t**
Density	Rural	0.2572	1.21	0.225	0.1282	0.28	0.776
	Urban	Omitted			Omitted		
Cohort	Born pre-1880	Omitted			Omitted		
	Born 1880-85	-0.2886	-0.38	0.705	1.0170	0.59	0.554
	Born 1885-89	-0.7329	-1.08	0.278	-0.3727	-0.24	0.809
	Born 1890-95	-0.5934	-0.9	0.367	-0.8782	-0.59	0.557
	Born 1895+	-1.5566	-2.33	0.02	-0.6889	-0.47	0.642
Birthplace	Atlantic	0.7803	3.67	0	1.8236	2.15	0.031
	Quebec	-0.2805	-0.88	0.38	-0.7982	-1.15	0.25
	Ontario	Omitted					
	West	1.3582	4.78	0	2.6762	1.58	0.115
Head's	Farmer	Omitted			Omitted		
Occupation	White collar	-0.1508	-0.53	0.595	-1.3172	-2.28	0.023
	Skilled	-1.0568	-3.82	0	-2.0263	-3.83	0
	Unskilled	-1.2402	-4.58	0	-1.7044	-2.87	0.004
	Labourer	-1.6067	-5.18	0	-2.0837	-3.96	0
	Other	-0.7739	-1.32	0.187	0.1187	0.1	0.918
Age	Under 19	-0.6917	-2.27	0.023	-0.3894	-0.48	0.631
	Age 19	0.5072	1.52	0.127	0.3760	0.46	0.649
	Age 20	0.3875	1.28	0.201	-0.1874	-0.31	0.757
	21 and Over	Omitted			Omitted		
	Constant	169.6792	250.77	0	166.3124	99.41	0
N		5232			1376		

Table 7: Oaxaca decomposition of mother tongue height differential

Using French mother tongue coefficients

	Estimate	z	P>z
Difference	4.0634	21.84	0
Endowments/explained	0.9001	2.09	0.037
Coefficients/unexplained	3.1633	12.03	0

Using English mother tongue coefficients

	Estimate	z	P>z
Difference	4.0634	21.84	0
Endowments/explained	0.3968	1.62	0.105
Coefficients/unexplained	3.6666	12.03	0

Figure 1 – Height Distribution for Linked CEF Recruits