Immigration and the Canadian Earnings Distribution in the First Half of the 20th Century

Alan G. Green[*]and David A. Green[†]

July, 2015

Abstract

We use newly available micro data samples from the 1911, 1921, 1931 and 1941 Canadian Censuses to investigate the impact of immigration on the Canadian earnings distribution in the first half of the 20th Century. We show that earnings inequality increased dramatically between 1911 and 1941, with most of the change occurring in the 1920s. This coincided with two of largest immigration decades in Canadian history (the 1910s and 1920s) and then the smallest immigration decade (1930s). We find, however, that immigration was not a prime cause of the increase in inequality in these years. The relative lack of effect arose for three reasons. 1) in the laissez-faire immigration policy before WWI, immigrants self-selected to have an occupational distribution that was similar to that of the native born; 2) in the 1920s, when immigration policy brought in a large number of farm labourers, immigrants adjusted geographically and occupationally after arrival to again end up with an occupational distribution similar to that of the native born; 3) general equilibrium adjustments in the economy helped mitigate the effects of occupation-specific immigrant supply shocks.

[*]Department of Economics, Queen's University (dec.)

[†]Vancouver School of Economics, University of British Columbia and International Research Fellow, IFS

The 1910s and 1920s were, respectively, the second and fourth largest immigrant inflow decades in Canadian history in per capita terms. Between 1911 and 1920, an immigrant inflow amounting to approximately 24 percent of Canada's 1911 population arrived in Canada, and between 1921 and 1930, the inflow equalled 14 percent of the 1921 population. In comparison, the largest immigrant decade (in a proportional sense) for the United States since Canadian Confederation in 1867 was the decade from 1900 to 1909 when immigrant inflows totaled 11 percent of the 1900 US population.[1] For Canada, the inflows in the 1910s and 1920s were accompanied by large emigration to the U.S., and we discuss this as part of immigrant adjustment after arrival, but even after accounting for that, immigrants who arrived in the previous decade made up 17 percent of male employment in 1921 and 14 percent in 1931.[2] Those percentages are at least double the size of the Mariel Boatlift inflow to the Miami labour market in the 1980s that has been the focus of considerable attention in discussions of immigration impacts in modern economies (Card 1990), and are two to three times the proportion of the Canadian population in 2006 who had arrived in the previous decade (6.2 percent). Moreover, in 1931, 38 percent of male wage earners in Canada were foreign born. Thus, whatever measures we use, Canada received substantial immigration inflows in the 1910s and 1920s. These large inflow decades were then followed by the lowest immigrant inflow decade on record: the 1930s.

We will show that this large influx and then almost complete cessation of immigration coincided with a substantial change in the Canadian earnings structure. Between 1910 and 1930, the ratio of the 90th to the 10th percentiles of the male annual earnings distribution increased by nearly 50 percent before experiencing a slight decline between 1930 and 1940. Our primary goal in this paper is to investigate the contribution of the large and variable immigrant inflow to the changes in the Canadian earnings structure between 1910 and 1940. That is, we are interested in the counterfactual question of how different the earnings structure would have been in the absence of immigration. We do this, first, using a decomposition technique which shows the direct impact of immigrants on the earnings distribution in the absence of general equilibrium effects. We then estimate the impact of immigration on the earnings of the native born and previously arrived immigrants within the context of a nested CES model of production that allows for broader, general equilibrium effects of immigration. In estimating that model, we address potential endogeneity concerns arising because the immigration inflows were not ex-

---

[1] For both countries, the immigrant numbers are from official statistics and so refer to legal immigrants.

[2] We will provide evidence that many immigrants who ultimately left Canada entered the Canadian labour force for some period, indicating that the ultimate stock of immigrants who settled permanently in Canada understates the impact of the large immigration inflows on the Canadian economy.

ogenous shocks but were distributed across occupations partly in response to demand patterns arising from demand for Canadian exports, technological change, etc.. In particular, we use an ethnic enclave instrument that takes advantage of the fact that new immigrants tend to locate where previous immigrants from their source country are living. Perhaps surprisingly, we find that the immigrant inflows had some effects on low skilled earnings but very limited effects on the overall shape of the earnings distribution. While there are studies of the effect of the set of changes the Canadian economy experienced on average wages in Canada in this period (e.g., Lewis 1975), to the best of our knowledge, there is no prior empirical analysis of the impact of the large pre-WWII immigrant inflows into Canada on its earnings structure.

Following from this first result, our second goal is to examine why such a large inflow would have such apparently small impacts on the earnings distribution. The first main reason is because of immigrant adaptation. As Ran Abramitsky et al (2014) emphasize, in periods with laissez-faire immigration policy, one might expect that it is mainly potential immigrants with skills that match demands in the host country economy who will actually migrate. This would be reflected in occupation distributions of new immigrants being similar to those of the native born and earlier arrivals. This is what Abramisky et al(2014) find for the U.S. in the era of mass migration. For Canada, the federal government had limited ability to control anything, including the total level of immigration, in the period before WWI, and we find that the occupational distribution of new immigrants was similar to that of the native born. This fits with the findings in Green and Green (1993) using data from manifests of ships arriving to Canada in 1912. To the extent that the occupational distribution of recent immigrants was similar to that of the native born and previously arrived immigrants, one would expect immigration to affect the level of earnings but not the shape of the earnings distribution.

Immigrants also made substantial occupational adjustments after arrival. Compared to their statements about intended occupations at the time of landing, immigrants were ultimately much less likely to be employed in agricultural labourer jobs and more likely to be in clerical and general labourer jobs. These post-arrival adjustments were important in all decades but were particularly important in the 1920s - a decade when Canada had adopted a more activist immigration policy that helped generate an inflow dominated by farmers and farm labourers. Immigrants made the adjustment through a variety of channels: re-migration (often to the U.S.); geographic migration within Canada; and switching occupations. We cannot distinguish the relative importance of these channels but we show that approximately two-thirds of immigrants arriving in the late 1920s were working in Canada at the time of the 1931 Census and so

immigrants did not merely pass through Canada without impacting the Canadian labour market. But whatever the form of the adjustment, the net effect was an occupational distribution that looked quite similar to that of the native born within 10 years of an immigrant cohort arriving. Again, this implies a limited impact of immigration on the shape of the earnings distribution. This result fits with arguments in Dustmann and Preston (2012) that one needs to be clear on where immigrants actually end up working in an economy to understand their impact on wages. In addition, our estimates imply general equilibrium adjustments in response to the immigrant supply shocks.

We investigate the ultimate impact of immigration on the earnings structure by working with the estimates of the parameters of our model of production to construct a set of counterfactual exercises. We find that a counterfactual scenario with no immigrants entering Canada in the 1920s implies small effects on both the level and spread of the native-born earnings distribution. In contrast, a policy that blocked only agricultural and general labourers from entering Canada in the 1920s would have increased earnings below the 40th percentile of the native-born distribution by about 10 percent and reduced the increase in inequality in this decade by a third. Focussing on farm labour (the main immigrant group among males), we show that their annual earnings would have fallen by 42 percent over the 1920s if all immigrant arrivals during that decade had stayed in Canada and in their stated intended occupation at time of arrival. Once we allow for the post-arrival adjustments by the immigrants themselves, that impact is reduced to 13 percent, and once we allow for general equilibrium effects, it is further reduced to a 4 percent decline. We conclude, much as earlier discussions that incorporate full general equilibrium effects (e.g. Hatton and Williamson 1998) that immigration inflows within an occupation can have substantial impacts on wages in that occupation if the inflows are viewed in isolation. But once we take account of spill-over effects through patterns of substitutability across occupations and of the fact that immigration inflows are either more balanced at arrival than is typically recognized or become more balanced with time in the host economy, the net impact on earnings inequality is limited.

As part of investigating the impact of immigration on the earnings distribution, we document movements in Canadian earnings inequality in the first half of the Twentieth Century. We see this as a central contribution of the paper. To document movements in the Canadian earnings distribution, we use recently available micro data from the decadal Censuses of 1911, 1921, 1931 and 1941. The key advantage of this data is that it includes consistent questions on earnings across these Censuses. Working with this, we document large increases in inequality in the 1920s

that were previously undocumented and also show new patterns in terms of experience profiles by occupation and intra-occupation changes in inequality. Our results are broader than those in earlier papers that either used aggregate wage data or relate only to a few occupations in the construction and manufacturing trades plus labourers (Bertram and Percy 1979; Mackinnon 1996; Emery and Levitt 2002). We document a decadal pattern in which wage and earnings inequality changed little between 1910/11 and 1920/21, increased sharply between 1920/21 and 1930/31 then declined somewhat between 1930/31 and 1940/41. The net result was a substantial increase in inequality across the period stretching from just before WWI to just before WWII. The main implication from our estimates of immigration impacts is that the large increase in earnings inequality in the 1920s was not primarily due to immigration and must, instead, have reflected other forces relating to technological change, institutional change, or trade.

1 Immigration Policy and Patterns

We begin with an overview of Canadian immigration policy before WWII which will form the basis for our argument that changes in that policy helped determine how immigrants adapted to the Canadian economy. Alan Green (1993), in a discussion of the political economy of immigration policy in Canada from 1896 to 1930, shows that the period can be divided into two periods: a period with a more laissez faire policy before WWI and a period with a more actively selective policy after WWI. He argues that the central theme of Canadian economics policy from Confederation in 1867 to WWI (and, possibly, into the 1920s) was the settlement of the West. In this context, immigration directed to the West was seen as having benefits both in providing a market for the burgeoning manufacturing sector in Central and Eastern Canada and in keeping the West safe from American expansionism. To carry it out, the Departments of Agriculture and the Interior established immigrant agents in the US, Britain and, to a limited extent, Continental Europe whose job was to attract farmers and farm workers who would be willing to settle in the West. Importantly, before 1906 there was no legislation in place to allow the government to refuse immigrants apart from "the diseased, the criminal or vicious, and those likely to become public charges." (Green 1993)[3]

The shift to a selective immigration policy began with revisions to the Immigration Act in 1906, but the real watershed was the new Act introduced in 1910. The key section in the 1910 Act allowed the government to prohibit immigrants from landing if they belonged to "any race

---

[3]The clear exception to this was Asian immigrants who faced out a head tax and other restrictions set out with the Chinese Immigration Act of 1885 and its subsequent revisions.

deemed unsuited to the climate or requirements of Canada," or in any specified class, occupation or character. Further, these prohibitions could be implemented through regulations, without the need to go to Parliament for approval (Kelley and Trebilcock 2010). The government now had almost unlimited power to select exactly the immigrants they wanted. It was a power they did not use to any real extent before WWI but it was used afterwards.(Green 1993)

Green(1994) argues that WWI was also a watershed in the role the West played in Canadian economic growth. Between the closing of the American frontier in 1896 and 1914, the expansion into the unsettled lands of the Prairies induced an unprecedented investment boom. But it was after WWI that Western exports of grains truly began to contribute as a driver of economic growth. In the new phase, too, there was strong demand for labour in the West. Partly this was because there was still land to be filled (Lew 2000), but it was also because the farm export sector needed labour for its highly seasonal production cycle.

Canada's immigration policy in the 1920s, like that in the US, distinguished between preferred countries (essentially, Britain, the US, and Northern and Western Europe), less preferred countries (Southern and Eastern Europe), and more or less prohibited countries (Asia). Prospective immigrants from preferred countries faced virtually no restrictions. Those from non-preferred countries were only to be admitted if they were farmers or farm labourers or, if they were women, domestic servants. Unlike the US, Canada did not impose numerical quotas to enforce this policy. Instead, Green (1993) argues, control was carried out indirectly by the requirement that immigrants from countries other than Britain had to get a visa before entering Canada in conjunction with the strategic placement of visa offices in preferred as opposed to non-preferred countries. This policy combination on its own would likely have created an immigration inflow much like that in the US in the 1920s - small relative to the pre-war period and mainly from preferred countries (Lew 2000). However, in 1925, the Canadian government entered into agreements with the two main railway companies for them to recruit agricultural labour in non-preferred countries. Byron Lew (2000) shows that in the remaining years of the 1920s there was a substantial composition shift in immigration toward immigrants from the non-preferred countries of Europe and headed for the West. [4] Finally, with the onset of the Depression, the Canadian government used its powers to all but shut down immigration.

_____

[4]In order to tighten the restriction that the immigrants were intended only for agricultural labour, the 1927 renewal of the agreement stipulated that the immigrants brought in by the railway companies had to have assured agricultural employment in advance (Lew(2000)).

1.1 Immigration: Flows and Stocks

Figure 1 shows the decadal immigration inflows to Canada for each decade since Confederation expressed as a proportion of the start of decade population. The decade from 1911 to 1920 was the second highest inflow decade on record, with the inflow being an astonishing 24 percent of the 1911 population. Because we will be interested in the impact on the earnings distribution in 1920, it is worth noting that nearly two-thirds of that inflow occurred in 1911, 1912 and 1913. But even if we removed those three years, the inflow in the rest of the decade made up 8 percent of the 1911 population, which is the same percentage as for the 2000s. The 1920s inflow was smaller at 14 percent of the start of decade population, but this still makes it the fourth largest inflow on record, and much larger than any decade that has followed. In contrast, the 1930s were the lowest immigration years on record, with the decadal immigration falling to 2.5 percent of the 1931 population.

Figure 1: Immigration Inflows by Decade as a Proportion of Start of Decade Population
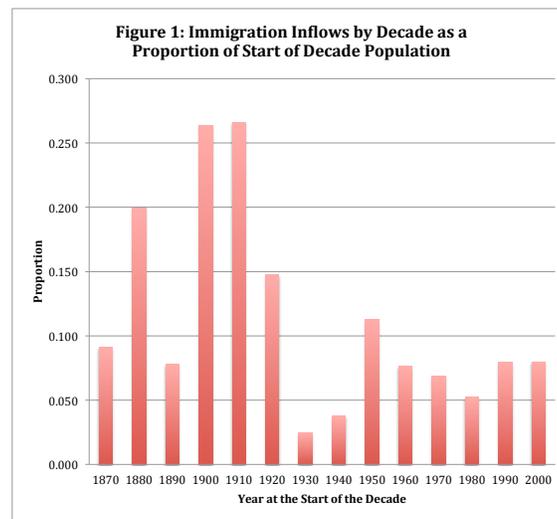
Source: Statistics Canada, CANSIM Tables V742019 (Estimated Population) and V742546 (Immigrant Arrivals).

Throughout the early twentieth century, the Canadian labour force and population were

defined not only by immigration but also by substantial emigration. In the period from 1911 to 1921, emigration has been estimated at between 78% and 85% of the gross inflow for the decade (Urquhart and Buckley(1965), p.22). Importantly, a large portion of the emigrants were Canadian born or immigrants who had arrived in earlier decades and so it is not necessarily the case that 78 to 85% of immigrant arrivals in this decade re-migrated. This is important for us because we are interested in the impact of immigrants on the Canadian earnings distribution and so want to know how many immigrants worked at least a few years in the Canadian labour market. By using Census data on year of arrival, we can obtain the proportion of immigrants who arrived in the years just preceding the Census and who were working in Canada at the time of the Census. Combining that with administrative data on inflows, we find that 37.7% of male immigrants who arrived in the 10 years before 1921 and 63% of those who arrived in the 2 years before 1921 were in Canada at the time of the 1921 Census.[5] The fact that a larger proportion of more recent arrivals were still in Canada at the time of the Census indicates that even if many immigrants eventually left Canada, a substantial proportion worked in and potentially had an impact on the Canadian labour market. Indeed, immigrants who arrived in the preceding decade comprised 17% of employed males age 15 to 64 in 1921 and 14% in 1931.[6] To take account of emigration related effects, we will work with net changes in numbers of both native born and immigrant workers between Censuses. Thus, our numbers will reflect labour force changes net of emigration and our estimates will include the impacts of the temporary migrants who were in the workforce at the time of the Census but later left.

One important question for assessing the impact of immigration on the labour market is whether immigrants actually entered (and stayed) in their intended occupations. Some immigrants would have moved on to the United States, others returned home and still others switched occupations within Canada. To investigate the net impact of these adjustment channels, in Table 1 we present distributions for males across broad occupations from the immigration department flow data and the Census.[7] Specifically, we sum the number of immigrants entering in each

---

[5]In 1931, 43% of immigrant male arrivals in the prior decade were working in Canada. Sources: Flow data: Urquhart and Buckley (1965), p. 24. 1921 Stock data: Census of Canada, 1921, Volume 2, p.61; and Census of Canada, 1931, Volume 6, Table 31.

[6]Source: Canadian Census micro data described in section 2.

[7]We describe the Census data in detail in the next section. The immigration inflow series are from the Parliament of Canada Sessional Papers. Immigrants with no classified occupation make up less than 10% of male immigrants in most years. The non-classified category includes both non-workers and professionals. The occupation categories are from the Sessional Papers data. Assignment of the Census occupations to these categories is somewhat rough. In particular, the "General Labourer" category appears to include semi-skilled workers but some of the occupations we

Table 1: Proportions in Broad Occupations: Native Born and Immigrant Stocks and Flows

	Including Farmers						Excluding Farmers			
	Immigrant Entrants				Native Born		Immigrant 1911-21 Entrants Stock	Native Born	Immigrant 1921-31 Entrants Stock	Native Born
	1911-20		1921-30		1921	1931	1921	1921	1931	1931
	Flow	Stock	Flow	Stock						
Farmers and Farm Labourers	0.38	0.21	0.66	0.32	0.46	0.40	0.14	0.14	0.23	0.11
General Labourers	0.38	0.42	0.14	0.41	0.26	0.29	0.45	0.42	0.47	0.43
Mechanics	0.17	0.18	0.14	0.15	0.15	0.16	0.20	0.24	0.17	0.24
Clerks, Traders, etc	0.06	0.19	0.06	0.12	0.13	0.15	0.21	0.20	0.13	0.23
Total	1	1	1	1	1	1	1	1	1	1

Sources: Flows: Sessional Papers of Canada, various years. Stocks: Census, various years. The left panel includes farmers in the "Farmers and Farm Labourers" numbers to match the inflow data. The right panel is based only on Census data and does not include farmers.

category across the years 1911 to 1920 and then calculate proportions (these are labelled "Flow" in the table). We compare those proportions to the proportions of immigrants in our 1921 Census sample who arrived in the previous 10 years (labelled "Stock" in the table). We repeat this exercise for the 1921-30 period, using the 1931 Census, but do not include numbers for the 1930s because the 1941 Census does not report year of arrival for immigrants. We include occupation distributions for the native born for comparison. Farmers and farm labourers are not separated in the inflow data and in the left panel we also include farmers in the Census data for comparability. The right panel contains Census tabulations of the occupational distributions of immigrants and the native born excluding farmers, which is the sample we will use in our investigations in the remainder of the paper.

The numbers in the table provide our first main results in terms of understanding the impact of immigrants on the earnings distribution. Immigrants themselves made adjustments that mitigated their potential impact, but did so in different ways in the different policy regimes. The 1910's immigrants had an intended occupational distribution that was similar to the native born distribution in 1921, though more concentrated in the General Labourers category and less in Clerical. By the time we reach the 1921 Census, recent immigrants had moved away from farming occupations and toward both General Labourer and Clerical occupations. Moving to the right panel (which does not include farmers), the result was a recent immigrant distribution that is strikingly similar to that of the native born. The implication is that recent immigrants were much

_____

classify as semi-skiilled may belong in the "Mechanics" category in the immigration data. We put all skilled trades in the latter category and combine miners with the General Labourer category. We drop professionals and managers to match the definition of the non-classified category (which we drop) in the immigration data. Nonetheless, we are confident in the Farmers and Farm Labour category that is crucial in our discussion.

less likely to become farmers than the native born but otherwise were in very similar occupations. That similarity arises both because, under a laissez-faire immigration policy, the migration choices made in the source countries resulted in an inflow that matched the occupational patterns in the host country, and because of adjustments after arrival. The point about choices made before migration is made in Abramitsky et al(2014), who show that the skill levels of immigrants to the United States in this era also matched well with those of the native born. In Green and Green (1993) we examine the intended occupations of a micro sample of immigrants arriving by sea in 1912 and also show that immigrants had a balanced mix of intended occupations that matched both the occupational distribution of the native born and the sectoral growth in the Canadian economy. Based on this and other evidence, we argue in that paper that the patterns fit with there being strong informational channels stretching back across the Atlantic.

In the 1920s, Canada's active immigration policy with its emphasis on agricultural labour had an impact that is evident in the stated intentions of immigrants arriving in the 1920s in Table 1. However, among the 1920s arrivals, the percentage in farming declined from the 66% who claimed they would enter those occupations when they arrived to the 32% who were actually in those occupations in the 1931 Census. As in the 1910s, the distribution shifted away from farming occupations toward the general labourer and clerical categories. But while the Census distribution represents a very substantial shift relative to the inflow, recent immigrants were still quite different from the native born in 1931. In particular, even after dropping farmers, they continued to show a higher share as farm labourers and a lower share in the skilled trades than the native born. This indicates that the more directed policies of the 1920s did have some effect - that it was not simply a case of relabelling immigrants of different types as farmers. Nonetheless, the table shows that in a regime when immigrants were not able to adjust their choices before arrival, they changed occupations to a considerable degree after arrival.

To better understand the nature of the adjustments that immigrants made after arrival, in Table 2, we present more detailed occupational breakdowns from the 1931 Census, with workers categorized as: native born; preferred country immigrants who arrived over 10 years before the Census; preferred immigrants arriving within the last 10 years; and non-preferred source country immigrants in both arrival categories. In addition, we show regional breakdowns between the West, and Central and Eastern Canada. Several points emerge from this table. First, the importance of immigration in the West if evident from the fact that 61 percent of the male workforce in that region in 1931 was foreign born. Second, long-standing immigrants from preferred countries were disproportionately represented in the trades and professional occupations and

under-represented in both the farm labour and general labourer categories. Third, immigrants from non-preferred countries tended to be heavily represented in the general labourer category and under-represented in more skilled occupations. Fourth, about a third of the non-preferred country immigrants in the West in 1931 who arrived in the previous decade were in farming, but even for that group a larger proportion were general labourers. Fifth, 71 percent of the non-preferred immigrants who arrived in the 1920s and who were in the West in 1931 were in lower skilled farm or general labour jobs. Sixth, though it is not shown in the table, among the non-preferred group who arrived in the 1920s, 47 percent were in the West. Thus, a large proportion of them ended up in general labourer jobs in the East. This table reinforces the view that immigration policy in the 1920s was effective in the sense that it generated the intended low-skilled labour supply flow to the West. However, it is also evident that this inflow spilled over into other regions and other occupations. Finally, these results imply that in our examination of immigration impacts later in the paper, we will need to be concerned about endogeneity since the observed occupational distribution, even of the recent migrants, is clearly not just being assigned to specific regions and occupations through policy.

2   Movements in Earnings Inequality

We now turn to characterizing the shifts in the earnings structure that took place over our sample period (1911 to 1941). This represents the first complete characterization of shifts in the Canadian earnings structure for this period and has been made possible by the recent availability of micro samples from the Canadian Censuses

2.1   Data and Basic Patterns

We use microdata samples drawn from the 1911, 1921, 1931 and 1941 Canadian Censuses under the auspices of the Canadian Century Research Initiative (CCRI).[8] The samples correspond to 5 percent, 4 percent, 3 percent and 3 percent of the Census population files for the 1911, 1921, 1931 and 1941 Censuses, respectively. The samples are random samples of dwellings with fewer than 30 inhabitants with a more complicated sampling scheme for larger dwellings. We use the weights provided in all of our estimation to take account of this scheme.

---

[8]We accessed the data in the Queen's University and UBC Regional Data Centres. The interpretation and opinions expressed are those of the authors and do not represent those of Statistics Canada. Further details on the data are provided in the Appendix.

Table 2: Occupation Distributions at Census Dates by Region, 1931

A: West

| | Native Born (percent) | Immigrants | | | |
		Pref <10 yrs (percent)	Pref >10 yrs (percent)	Not Pref <10 yrs (percent)	Not Pref >10 yrs (percent)
Labourers	17.3	23.3	16.0	37.8	41.3
Farm Labour	18.2	28.8	8.1	32.9	8.1
Farm Managers	1.5	1.9	1.2	1.1	0.6
Clerical	8.2	3.7	7.1	0.3	1.2
Managers	5.9	2.6	7.2	0.5	1.6
Professionals	6.7	2.9	6.9	1.2	1.6
Sales	8.3	4.3	7.5	1.5	2.8
Semi-Skilled	11.8	11.3	13.5	13.4	16.5
Service	4.7	7.0	7.3	3.8	12.5
Trades	17.5	14.3	25.4	7.6	13.8
Total	100	100	100	100	100
Proportion in Group	0.39	0.11	0.30	0.09	0.11

B: Central and East

| | Native Born (percent) | Immigrants | | | |
		Pref <10 yrs (percent)	Pref >10 yrs (percent)	Not Pref <10 yrs (percent)	Not Pref >10 yrs (percent)
Labourers	22.7	20.4	14.8	48.4	32.9
Farm Labour	6.5	17.4	3.6	5.5	1.6
Farm Managers	0.9	1.5	0.4	1.1	0.5
Clerical	6.5	5.0	7.5	0.8	1.7
Managers	4.4	2.8	4.5	1.0	2.9
Professionals	5.8	4.0	6.4	1.7	2.8
Sales	8.2	5.2	6.5	2.3	4.8
Semi-Skilled	17.6	14.7	17.7	17.7	22.5
Service	5.2	7.6	8.1	4.6	11.0
Trades	22.2	21.5	30.5	17.0	20.0
Total	100	100	100	100	100
Proportion in Group	0.71	0.07	0.15	0.04	0.04

Sources: Census, various years.

The key variable for our purposes is annual earnings which refers to wage, salary, commission or piece rate earnings from all jobs.[9] Importantly, the question about annual earnings is identical across the Censuses in our sample, asking the respondent to report "total earnings in the preceding 12 months." Using a variable on type of work, we selected only "wage-earners" which, in both the introductions to the 1921 and 1931 Censuses is defined as "a person who works for salary or wages, whether he be the general manager of a bank, railway or manufacturing establishment, or only a day labourer." This definition excludes the self-employed and unpaid family workers (e.g., farmers sons). It is worth emphasizing that farmers working on their own farm would not be included in our sample even though they make up a substantial share of the overall workforce.[10] We do not include them because we are concerned that their income would reflect a combination of the wages we are interested in and returns on their capital.[11] We restrict our sample to males aged 15 to 64.

We convert earnings into 1913 Toronto equivalent dollars using the cost of living indexes in Emery and Levitt(2002). Because our earnings data correspond to 12 month periods spanning half of two consecutive calendar years, we use the average of the listed 1910 and 1911 values for the first Census, and similar averages for the other Censuses. We use the direct numbers for the 13 larger cities investigated in Emery and Levitt(2002) and use the value for the smallest of their cities in a province for workers in smaller towns and rural areas in the province.

Occupation is another important variable in our investigation. The CCRI mapped from Census responses into a joint occupation and industry coding corresponding to the Integrated Public Use Microdata Series (IPUMS) system based on the 1950 US Census codes. We used those codes to form aggregate occupation groups at approximately the 3-digit Standard Occupational Classification group level. We drop occupation categories with fewer than 5 observations in any of our Census samples.The result is 146 consistently defined occupations which we call the "narrow" occupational grouping. At times, we will also work with a "broad" grouping in which we aggregate the narrow occupations into 10 groups: general labourers, farm labourers, farm

---

[9]We present results based on weekly wages in the appendix. Our main conclusions are not affected by switching from annual earnings to weekly wages.

[10]In 1946, among males, farm owners constituted 19 percent of the total workforce, defined as including paid and unpaid workers, the self-employed, and employers, and unpaid family members in agriculture made up another 6.7 percent.

[11]We also exclude other self-employed workers for the same reasons as farmers. It is worth noting, though, that the ratio of wage earners to the complete set of "gainfully occupied" workers rises in this period. Thus, the wage movements we observe will partly reflect selection issues related to shifts in the labour force toward paid employment. As with the rest of the wage inequality literature, we do not pursue those selection issues here.

managers, clerical workers, managers, professionals, sales occupation workers, service workers, semi-skilled workers, and tradesmen.

Finally, we group individuals into five age categories: 15 to 24; 25 to 34; 35 to 44; 45 to 54; and 55 to 64. In our estimations, we work with data aggregated to the age by occupation level since we don't have any observable characteristics within these groups. In those exercises, we make a further cut, dropping all age by occupation cells with fewer than 10 observations in any year.

2.1.1 Changes in Earnings and Wage Distributions

Figure 2 depicts changes in the real annual earnings distribution for each of our three decades (1911-21, 1921-31, 1931-41). The solid line in this figure corresponds to the difference between log earnings in 1921 and log earnings in 1911 at each percentile, and thus roughly shows the proportion differences in the distributions at each percentile. The line with squares shows the same difference between 1931 and 1921, and the line with triangles shows the difference between 1941 and 1931. Importantly, we are comparing percentiles, not specific people, across years; thus, when we report that, say, the 10th percentile declined by 15 percent this does not necessarily mean that a person who was at the 10th percentile in 1911 experienced a decline of 15 percent in his earnings. As a result, the movements depicted in this figure reflect a combination of shifts in real wages within occupations, related changes in rankings of occupations, and changes in the proportion of workers in each occupation. A rising line reflects an increase in inequality between the pair of years as that arises when increases at the top of the distribution are greater than at the bottom.

Looking at the 1911 to 1921, there is slight evidence of an increase in inequality as the percentiles below about the 30th decline by close to 10 percent in real terms while those in the upper end of the distribution increase by just under 10 percent. These movements, however, are dwarfed by the increase in inequality in the 1920s. Below about the 60th percentile, the real difference line increases linearly from a difference of approximately 0.6 log points (or, about a 45 percent decline) near the 5th percentile to approximately zero near the 55th percentile. Above the 60th percentile, the differential continues to increase across percentiles, though at a slower rate. The ratio of the 95th to the 5th percentiles rises by a factor of almost 10 in this decade. Finally, over the 1930s there is a decrease in inequality as the top of the distribution changes very little compared to approximately 20 percent increases through much of the lower half of the distribution. In spite of this, the net effect over the whole period from 1911 to 1941 is a strong

14

Figure 2: Decadal Differences in Log Percentiles, Male Earnings
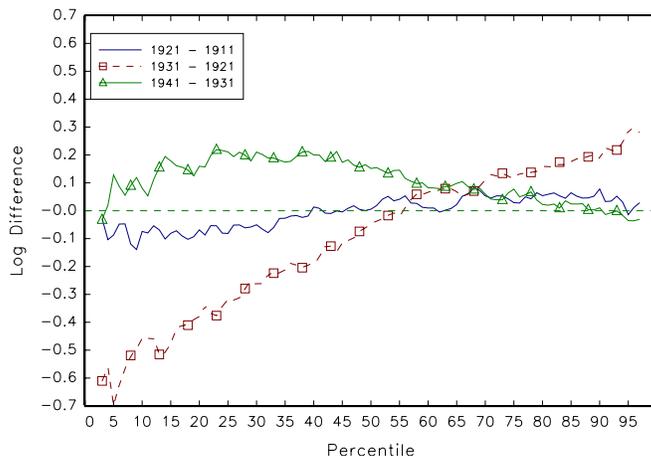
increase in earnings inequality. Overall, the figure indicates that Canada underwent a substantial long-term change in its earnings structure between 1911 and 1941. Roughly speaking, this can be broken down into: a decade with small inequality increases from 1911 to 1921; a decade with a very substantial increase in inequality, driven particularly by large real earnings declines at the bottom of the distribution, from 1921 to 1931; and a mixed period with some reductions in inequality from 1931 to 1941.

The other two main investigations of wage movements in this era are Emery and Levitt(2002)'s study using data from the Labour Gazette and Mackinnon(1996)'s work using data from the Canadian Pacific Railway (CPR). Emery and Levitt(2002) focus on regional wage differentials and do not directly discuss movements in national inequality, while Mackinnon(1996) provides some discussion of wage differentials between skilled trades and labourers. We provide a detailed comparison of results from Census data with the results from both of these other papers in the appendix. Our main conclusion is that the Census data accords well with the Labour Gazette data in terms of long term movements in labourers' wages and wage differentials between the

15

skilled trades and labourers but less well with Mackinnon's CPR data. Given coverage issues with both the CPR and Labour Gazette data, we view the Census data as the most reliable, but we are reassured by the matching of broad patterns between it and the Labour Gazette data.

Regardless of these specific comparisons, it is important to note that all earlier work on wage inequality for Canada has been done in terms of ratios between specific occupations - most often between skilled, blue collar trades or clerks and labourers. Ours is the first presentation of which we are aware of the entire earnings distributions for these years. In our data, general labourers have earnings that place them at approximately the 25th percentile of the overall earnings distribution and the skilled trades and clerical workers have earnings near the 75th percentile. Using the Census micro data, we are able to see whether movements in the skill differential (i.e., what is essentially the 75-25 ratio) are indicative of the overall changes in inequality. An examination of Figure 2 indicates that given the relative similarity of changes in earnings above about the 60th percentile, following the 75th percentile does a good job of capturing movements in the upper tail of the distribution. But the lower tail of the distribution shows a much stronger decline than that observed at the 25th percentile in the 1920s and smaller increases in the 1940s. Thus, focussing on skilled worker versus labourer wage differentials leads to a substantial understatement of the extent of the increase in inequality in earnings over this period.

3   Decomposing the Changes in the Earnings Distribution

We turn now to the question of whether the changes in earnings inequality in this period are linked to the immigrant inflow. The most direct way for the immigrant inflow to affect the earnings distribution is through shifting the composition of the workforce. In the absence of general equilibrium impacts on wages, an immigrant inflow concentrated, for example, in low-skilled occupations would imply a fattening of the left tail of the earnings distribution. In this section, we decompose the decadal changes in Canada's earnings distribution into factors related to immigrant status, age, and occupation.

A change in the earnings distribution can arise through three channels: changes in the distribution of workers across skill related categories (immigrant status, age, and occupation, in our case); changes in the differences in average earnings between those categories; and changes in the dispersion of earnings within the categories. We have already discussed the occupational component of the first channel for immigrants. Table 1 also indicates that the native born occupational composition was relatively constant over this period.

To investigate the between group wage differentials, in Table 3 we present differences in average log wages between groups defined by ten occupation groups and five ten-year wide age groups and the average log wage for a base group (15 to 24 year old general labourers) for each Census year. The wage differentials in 1911 and 1921 were quite similar. In both cases, there were substantial increases in wages between the 15-24 year olds and the 25-34 years olds for all occupations. But age differentials beyond age 24-34 were quite flat, even for white collar occupations like professionals. Virtually all occupations also showed a small decline in average earnings between ages 45-54 and 55-64. The cross-occupation differences at a particular age line up with expectations, with the highest wages for managers and professionals, the lowest for labourers, and semi-skilled and skilled white and blue collar workers falling in between.

Between 1921 and 1931, the earnings structure changed in two ways. First, we see the emergence of more substantial age differentials. For example, the difference between 25-34 year old and 35-44 year old managers was 0.19 log points in 1921 but 0.32 log points in 1931. Second, the earnings differentials among blue collar occupations and between white collar and blue collar occupations increased substantially, particularly . As an example, the earnings differentials between trades occupations and semi-skilled workers and labourers were 0.07 and 0.46 log points, respectively, in 1921. These rose to 0.13 and 0.78 in 1931. In contrast, the differential between clerical workers and managers was virtually unchanged over this period. Finally, between 1931 and 1941, the age differentials either stayed the same or increased slightly while the occupation differentials tended to decrease.

To obtain a measure of changes in within occupation earnings dispersion, we estimated the difference between the 90th and 10th percentiles of the log earnings distribution for each of our 10 broad occupation categories, holding the age dimension constant by looking only at 35-44 year olds. The average of these 90-10 differentials across occupations took values of 1.22 in 1911, 1.23 in 1921, 1.65 in 1931 and 1.63 in 1941. Overall, the earnings calculations indicate that the wage structure (both between and within occupations) changed substantially between 1921 and 1931 but moved only mildly in the decades before and after that.

To judge the relative importance of the composition shifts versus the changes in the wage structure, we employ the ingenious decomposition technique developed in Firpo et al(2009) (hereafter, FFL). At the core of the FFL methodology are influence functions, which show the influence of an observation on the value of a given statistic. Of particular interest is their demonstration that once the influence functions for quantiles are "recentered" around the sample quantile values, one can decompose movements in the quantiles of a distribution in a manner directly

17

Table 3: Differences in Average Log Wages by Age and Occupation

| A: 1911					
Occupation			Age Group		
	15-24	25-34	35-44	45-54	55-64
Labourers	0	0.18	0.19	0.20	0.12
Farm labour	-0.49	-0.22	-0.18	-0.24	-0.32
Farm managers	-0.41	-0.01	0.10	0.18	0.03
Clerical	0.23	0.72	0.78	0.88	0.87
Managers	0.44	0.94	1.11	1.11	0.94
Professionals	0.35	0.82	0.92	1.00	0.92
Sales	0.20	0.72	0.87	0.84	0.86
Semi-skilled	0.17	0.46	0.46	0.47	0.37
Service	-0.09	0.34	0.46	0.42	0.32
Trades	0.27	0.64	0.66	0.64	0.56
B: 1921					
Labourers	0.00	0.29	0.34	0.28	0.19
Farm labour	-0.38	-0.07	-0.01	-0.13	-0.28
Farm managers	-0.39	0.00	0.08	0.09	-0.01
Managers	0.45	1.05	1.24	1.18	1.06
Clerical	0.33	0.81	0.94	0.78	0.87
Professionals	0.43	0.90	1.06	1.03	1.01
Sales	0.31	0.78	0.94	0.89	0.84
Semi-skilled	0.20	0.62	0.63	0.57	0.47
Service	-0.06	0.41	0.55	0.54	0.41
Trades	0.28	0.71	0.80	0.72	0.63
C: 1931					
Labourer	0.00	0.38	0.52	0.52	0.43
Farm labour	-0.28	-0.01	0.01	-0.03	-0.09
Farm managers	-0.03	0.38	0.68	0.64	0.38
Clerical	0.93	1.50	1.69	1.67	1.54
Managers	1.10	1.70	2.02	2.00	1.86
Professionals	1.05	1.59	1.93	1.90	1.80
Sales	0.75	1.44	1.58	1.61	1.47
Semi-skilled	0.55	1.03	1.17	1.11	0.99
Service	0.31	1.02	1.21	1.23	1.00
Trades	0.62	1.14	1.30	1.29	1.11
D: 1941					
Labourer	0.00	0.44	0.51	0.49	0.36
Farm labour	-0.60	-0.24	-0.19	-0.28	-0.44
Farm managers	-0.44	0.05	0.31	0.41	0.29
Clerical	0.37	1.06	1.32	1.44	1.32
Managers	0.54	1.32	1.67	1.80	1.75
Professionals	0.50	1.23	1.56	1.64	1.63
Sales	0.35	1.12	1.34	1.36	1.17
Semi-skilled	0.31	0.90	1.07	1.09	0.99
Service	-0.20	0.74	0.85	0.85	0.77
Trades	0.35	0.96	1.11	1.18	1.06

Sources: Census, various years. The table entries for each year show the difference in average log wages between each occupation by age category and the average log wage for general labourers aged 15 to 24 in that year.

analogous a Oaxaca decomposition. In particular, one can assess the extent to which shifts in the distribution of observable characteristics shift specific quantiles. We use their methodology to assess the impact of shifts in four personal characteristics on the earnings distribution: age (arranged in five, ten-year groups); occupation (in the 146 groups described earlier); whether a person was a recent immigrant (an immigrant who arrived in the 10 years prior to the Census); and whether a person was an established immigrant (an immigrant who arrived more than 10 years before the Census). Native born workers form the omitted category.[12]

In Figure 3a, (in the line labelled "actual difference") we plot the change in each decile of the real earnings distribution between 1911 and 1921. This is a re-plotting of the relevant line in Figure 2 but only showing the decile points. The line labelled "Explained" shows changes in a given decile due to changes in the composition of the labour force along the age, occupation, and immigration dimensions. The third line (labelled, "Residual") equals the difference between the actual change and the change arising purely from changes in the composition in terms of observables. It corresponds to the change that would have happened if the composition were held constant. As such, it captures the impact of changes in the quantile specific coefficients associated with the observables (including the intercept) and reflects changes in returns to observable and unobservable worker characteristics. For the 1911-21 period, shifts in observable characteristics, including immigrant status, explain very little of the overall change in earnings inequality below the 7th decile, However, those composition shifts play an important role at the top of the distribution, explaining 30 percent, 60 percent, and 45 percent of the movements at the 7th, 8th and 9th deciles, respectively. For the 1921-31 changes shown in Figure 3b, composition explains a similarly small proportion of changes in the lower two-thirds of the distribution. In the right tail, composition again explains more but to a lesser extent than in the previous decade (explaining 40 percent, 20 percent and 13 percent of the top three deciles in ascending order).

The overall conclusion from the decomposition exercise is that movements in observable characteristics, including changes related to the inflow of immigrants, were not a major determinant of movements in the real earnings distribution. Most importantly, they explain little of the substantial decline in the lower tail of the real earnings distribution in the 1920s. That decline arises from a combination of the increased earnings differentials between age-occupation groups and the increase in within-group dispersion. In a counterfactual exercise for the 1921-31 decade reported in the appendix, we generate the change in the earnings distribution that would

---

[12]We implemented the decomposition at the ten decile points in Stata using the rifreg ado file found on Nicole Fortin's website.

Figure 3a: FFL Decomposition, 1911−21

Figure 3b: FFL Decomposition, 1921−31
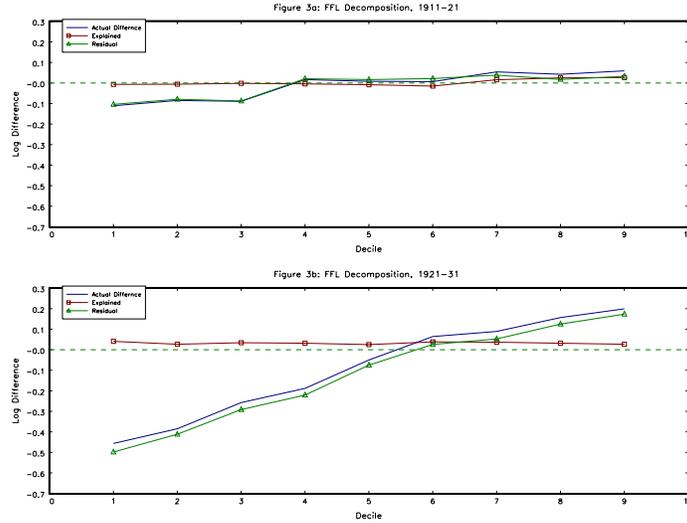
have arisen if there had been no within-group variation. Based on that exercise, within group variation plays a varying but generally important role in inequality changes below the median but essentially no role above the median.

To better understand the role of immigration in these decompositions, in figure 4, we plot differences in percentiles of the immigrant and native born earnings distribution for each Census (immigrant minus native born). The striking point from the figure is that for the 1911, 1921 and 1931 Census years, the immigrant and native born earnings distributions were very similar, with differences typically below 10 percentage points. For 1941, the immigrant earnings distribution is strongly superior to that of the native born. However, since there had been virtually no immigration for a decade, the immigrant workforce was older than the native born in 1941 and this difference likely just reflects returns to experience rather than a true immigrant advantage. Given that immigrant and native born occupational distributions were broadly similar, the similarity of the earnings distributions implies that there was not systematic discrimination (or

20

differentials in productivity) between immigrant and native born workers.[13] It also explains why the decomposition exercise shows little effect from the immigration inflow. However, this is not the final answer on whether the immigration inflow affected the wage structure since this decomposition exercise ignores general equilibrium effects in the form of the immigration inflows altering occupation specific wages over time. To assess the latter potential impacts we will need to use a structural model of wage determination, and we turn to that in the next section.

Figure 4: Immigrant minus Native Born Differences in Log Percentiles, Male Earnings
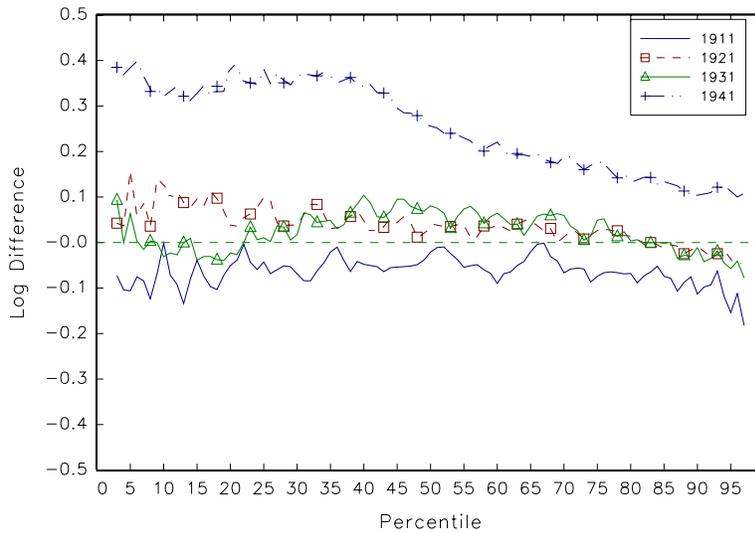
[13]Based on wage data and referring to the report of the US Immigration Commission, Timothy Hatton and Jeffery Williamson (1998) similarly argue that "immigrants received roughly the same wages as natives." In contrast, Alan Green and Mary Mackinnon (2001) find that new immigrants faced a substantial earnings disadvantage relative to the native born that declined only gradually. However, their data is from the 1901 Census, before the real onset of the very large immigration flows that define our period.

4 Estimating the Impact of Immigration on Changes in the Wage Structure

In this section, we examine whether immigration affected earnings inequality in inter-war Canada using techniques from recent papers on the impact of immigration on the native born wage distribution. Given the evidence in section 2 showing potentially endogenous geographic adjustments by immigrants after arrival, we follow approaches that focus on national level wage and employment patterns as opposed to approaches comparing local labour markets.

4.1 Nested CES Model

We address the impact of immigration on the wage structure using a quasi-structural estimation approach based on a nested Constant Elasticity of Substitution (CES) production function. Our specification closely follows Ottaviano and Peri (2010) (hereafter, OP) and D'Amuri et al (2010) (hereafter, DOP). This approach allows for a complete accounting of the impacts of immigration by incorporating the potential effects of an increase in the supply of immigrants in one occupation on the wages for other workers through substitutabilities and complementarities in production. Of course, the CES specification, even in its nested form, imposes significant restrictions on allowable patterns of interactions of use of factors that should be kept in mind. The on-line appendix contains a more detailed account of the estimation approach. Alternatively, readers not interested in estimation and identification issues can skip ahead to the counterfactual exercises in section 4.1.2.

We begin by considering the aggregate production specified as a simple Cobb-Douglas function of a labour aggregate, L_t, and capital, C_t,:

$$Y_t = A_t L_t^\alpha C_t^{(1-\alpha)} \tag{1}$$

where Y_t is aggregate output and A_t is a productivity shifter.[14] We will assume that the economy faces fixed world prices for the output but that C_t is less than perfectly elastically supplied (perhaps because it includes scarce factors such as land in the rural sector, and entrepreneurial

---

[14]In an earlier version of the paper, we allowed, further, for separate production functions in rural, major city, and town sectors. We have also estimated separate specifications in which we dropped the Maritime Provinces (which were recipients of small numbers of immigrants) and for the West alone. The estimated elasticities of substitution were very similar by sector and region, and so we have chosen, instead, to simplify to one aggregate production function.

ability in the other sectors). This implies downward sloping labour demand curves. [15]

Following DOP, we specify the labour aggregate as a CES function of types of skilled labour defined by our 10 broad occupation groups. Given this specification, the labour aggregate in period t can be written as:

$$L_t = (\sum_{k=1}^{10} \theta_{kt} L_{kt}^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where L_{kt} is employment in occupation k, and θ_{kt} are skill-time specific productivity shocks. The parameter σ determines the degree of substitutability among broad occupation groups.

We capture further worker dimensions through a sequence of nestings. First, we treat labour in each broad occupation group as a CES aggregate of employment of the narrow occupations that fit in the broad group. The employment in the narrow occupations is, in turn, a CES aggregate of workers in the five age groups. Within age groups, we divide workers between recent immigrants and a combination of established immigrants and the native born. We implement this division because of the similarity of the established immigrant and native born occupational distribution. The final level of nesting contains the native born and established immigrants separately.

Using the nesting specifications, and again following DOP closely, we can derive a wage equation for a native born worker in broad occupation k, narrow occupation l, and age group j assuming their wage equals their marginal product:

$$\ln w_{Nkljt} = \ln(\frac{\partial F}{\partial L_t}(C_t, L_t)) + \frac{1}{\sigma}\ln L_t + \ln\theta_{kt} - (\frac{1}{\sigma} - \frac{1}{\sigma_o})\ln L_{kt} + \ln\theta_{klt} - (\frac{1}{\sigma_o} - \frac{1}{\rho})\ln L_{klt}$$
$$+ \ln\theta_{kljt} - (\frac{1}{\rho} - \frac{1}{\delta})\ln L_{kljt} + \ln\theta_{ENkljt} - (\frac{1}{\delta} - \frac{1}{\phi})\ln EN_{kljt} + \ln\theta_{Nkljt} - \frac{1}{\phi}\ln N_{kljt} \tag{3}$$

where, L's correspond to employment levels for groups defined by occupation and age, EN_{kljt} is the number of combined established immigrants and native born workers in the cell defined by occupation and age, N_{kljt} is the number of native born workers in that cell, and θ's correspond to productivity shifters. The parameters σ_o, ρ, δ, and ϕ determine, respectively, the degree of substitutability among narrow occupations, age groups, between recent immigrants and all other workers, and between established immigrants and the native born. We can similarly derive the log wage equations for an established immigrant worker in the same cell.

As OP point out, inspection of (3) reveals several channels through which an immigrant shock in a specific occupation by age cell can affect the wage of any native born worker. First, the fact

---

[15]This is in contrast to the specification for the manufacturing sector in Chambers and Gordon (1966) and makes our set-up closer to that in Lewis (1975).

that labour of different types is imperfectly substitutable implies that all workers benefit from an increase in the total number of workers in the economy. This is captured in the second term on the right hand side of (3), which is positive. The fourth term on the right hand side captures the effect of increased immigration within the same broad immigration group. This effect will be negative if the elastisticity of substitution among broad groups is less than that among the narrow groups within broad groups (as seems likely). This, and the overall size effect captured in the second right hand side term, are the channels through which immigration in the same broad occupation group but a different narrow occupation group can affect a worker's wage. Similarly, the sixth term allows us to capture the effects of immigration in the same narrow occupation but different age group and the eighth term reflects the direct effect of immigrants entering in the same narrow occupation and age group. The last two quantity terms in the expression relate to previous arrivals and the native born and so will not change with immigration.(Ottaviano and Peri 2010).

4.1.1 Estimating Equations and Results

The parameters in the nested CES model can be obtained sequentially, moving up the levels of aggregation. Thus, following OP, we can obtain an estimate of ϕ by taking the difference between the log wage expressions for native born and established immigrant workers:

$$\ln(\frac{w_{Nkljt}}{w_{Ekljt}}) = \ln(\frac{\theta_{Nkljt}}{\theta_{Ekljt}}) - \frac{1}{\phi}\ln(\frac{N_{kljt}}{E_{kljt}}) \qquad (4)$$

where, E_{kljt} is the number of established immigrants in the occupation by age cell, w_{Ekljt} is their average wage, and θ_{Ekljt} is the productivity shifter associated with those immigrants. This yields an intuitive formulation: holding constant relative differences in productivity (the θ's), comparing relative differences across cells in wages and employment levels between established immigrant and native born workers tells us about the substitutability of the two types of workers. If we see a large relative jump in established immigrant employment but comparatively small changes in their relative wages, for example, we would conclude that established immigrant and native born workers are very substitutable in production - that the established immigration increase affected immigrant and native born wages similarly. Defining cells as narrowly as possible makes it more plausible that there are no productivity differences between the worker types. Thus, established immigrant and native born agricultural labourers of the same age would reasonably be expected to be working with the same technologies and thus to have the same θ's. To isolate this cell-level variation, we include separate occupation, age and time effects. This is

equivalent to assuming that we can write,

$$\ln(\frac{\theta_{Nkljt}}{\theta_{Ekljt}}) = \psi_l + \psi_t + \psi_j + u_{kljt} \tag{5}$$

where the ψ's are fixed effects and u_{kjt} is an independent error term. Identification then relies on the assumption that relative supplies of native born versus established immigrants may respond to longer run factors such as the persistent occupation, age and time effects captured in the included fixed effects, but do not respond to the idiosyncratic disturbances to productivity reflected in u_{kjt}.

We run equation (5) using real annual earnings and employment numbers from the Census data aggregated to the narrow occupation group level. We weight using the square root of the cell size. The dropping of small or empty cells leaves 1534 occupation by age group cells. Standard errors are clustered at the occupation x age group level to allow for flexible forms of serial correlation in the error term. When we move to working with recent immigrants in the higher nesting levels, we will create instruments to address potential endogenous adjustments by the immigrants, but no such instrument is available for the older immigrants and so we only present OLS results for this regression. In Table 4, we present the key estimated coefficients (the negative inverse elasticities) and the implied elasticities from each stage of the CES nesting.

Our estimate of the negative inverse elasticity of substitution is -0.014 and not statistically significant at any conventional level. This falls in the lower part of the range of estimates from DOP for modern day Germany and OP for modern day United States (-0.01 to -0.06). The corresponding elasticity from our estimates is 71, implying that established immigrants and native born workers are almost perfectly substitutable in production, as we predicted.

We next move one step up in the nesting hierarchy to consider substitutability between workers already in Canada (both native born workers and established immigrants) and recent immigrants within the same age-occupation cell. To do this, we aggregate (3) to the established/recent immigrant worker level to obtain a wage expression for the combined native born/established immigrant group. We obtain an expression for the wage for a recent immigrant in the same way and then take the difference between the two wage equations to obtain:

$$\ln(\frac{w_{Ikljt}}{\bar{w}_{ENkljt}}) = \ln(\frac{\theta_{Ikljt}}{\theta_{ENkljt}}) - \frac{1}{\delta}\ln(\frac{I_{kljt}}{EN_{kljt}}) \tag{6}$$

where, w_{Ikljt} is the average wage for recent immigrants, θ_{Ikljt} is the productivity shifter for recent immigrants, I_{kljt} is the number of recent immigrant workers, and $\bar{w}_{ENkljt}$ is the weighted

Table 4: Negative Inverse Elasticities of Substitution from CES Estimation

	OLS	IV	Implied Elasticity	Elasticity Between	No. of Obs	First Stage F stat
$-\frac{1}{\phi}$	-.014 (.018)		71	Est Immigs &Natv Born	1394	
$-\frac{1}{\delta}$	.008 (.018)	.11 (.091)	-	Rcnt Immigs & Est Immigs/Natv Born	1258	12.3
$-\frac{1}{\rho}$	-.073** (.012)	-.045* (.019)	22	Age Groups	1127	552
$-\frac{1}{\sigma_0}$	-.10** (.020)	-.10* (.049)	7	Narrow Occ. Groups	423	35.6
$-\frac{1}{\sigma}$	-.28 (.14)	-.32 (.16)	3	Broad Occ. Groups	30	6.7

Notes: The calculated elasticity is based in the IV estimate in all rows except for the elasticity between established immigrants and the native born, where there is no valid instrument.
Specifications for the first two rows include year, age and occupation fixed effects. The remaining rows include year and occupation fixed effects.
Standard errors in parentheses. Standard errors are clustered at the occupation by age group level in the first three rows, and at the occupation levels in the remaining rows.
*, ** correspond to statistical significance at the 5 and 1 % levels, respectively.

average of established immigrant and native born wages in the occupation-age cell, with the weights being the proportions of established immigrant and native born workers. We again assume that the ratio of productivity terms can be captured by occupation, age and time fixed effects plus an idiosyncratic error term.

There is reason for concern about the endogeneity of the immigration supply changes even after including the fixed effects. In particular, we know that immigrants who were explicitly brought in to work in agriculture in the West, were, by the time we observe them in the next Census, in a variety of occupations and locations. Thus, the set of changes in immigrant employment we observe might partly be due to endogenous responses to within-cell productivity shocks. We use a variant of an instrument originating with David Card (2001) in which we assume that immigrants move to a geographical location based partly on joining an enclave of people from their country of origin these and then, once there, pursue the most common occupations in that location. That is, we assume that when they search for a job in the location they are attracted to for companionship reasons, their probability of finding a job in a given occupation is proportional to the size of that occupation in that location. Moving to a mining town, immigrants are more likely to find mining jobs than if they moved to Toronto.

To construct the instrument, consider an immigrant from source country g, and let the proportion of immigrants from g in location r in year t be given by p_{grt}. Also, let the proportion of all workers in narrow occupation l that is within broad occupation k in location r in period t be given by π_{klrt}. We will predict the number of people from country g in occupation k in year t+1 at the national level as:

$$n10_{klgt+1} = \sum_r \pi_{klrt} * p_{grt} * n10_{gt+1} \tag{7}$$

where $n10_{gt+1}$ is the number of immigrants from country g who arrived in the previous 10 years at the national level. We do this exercise separately by age group with both the π_{klrt} weights and the national level inflows (n_{gt+1}) being age group specific.[16] Finally, we form our instrument for $\ln(\frac{I_{kljt}}{EN_{kljt}})$ as $\ln(\frac{n10_{gt}}{EN_{kljt}})$. The first stage F-stat for this instrument is 12.3, implying that we do not have a weak instrument problem.

The key identifying assumption is that neither the composition of employment by occupation at the start of a period in a location nor the distribution of immigrants from a given country across locations at the start of the period are correlated with the change in productivity within an occupation-age cell at the national level in the ensuing decade. We also require that the general growth in immigration from each country is not correlated with the cell specific productivity shocks. These assumptions would be violated if people from a given source country only immigrated to Canada because there was an increase in demand for a specific occupation in a specific location. For example, if all migrants from Lithuania were miners and they all went to Northern Ontario because there was an ongoing increase in demand for miners across decades in that region, the instrument would be problematic. In that case, the correlation between past concentrations and recent inflows of Lithuanians to the region would partly reflect productivity trends rather than just living situation preferences. The fact that π_{krt} is calculated for all workers in the location, not just for immigrants from the specific source country, makes it more likely that the requirements for the instrument to be valid are met. Both the OLS and IV estimates

---

[16]The regions are defined as follows. We divide Canada into rural areas, towns, and 13 major cities (Victoria, Vancouver, Edmonton, Calgary, Regina, Winnipeg, Hamilton, Toronto, Ottawa, Montreal, Quebec City, St. John, and Halifax). Towns are all locations with populations over 5000 that are not on the list of major cities, and rural areas are all remaining locations. We define a region as either a major city or a town or rural area within a given province (e.g., towns and rural areas in British Columbia are two of our regional groups). We group Prince Edward Island and Nova Scotia together, resulting in 8 provincial groups. Thus, the total number of regions is 13 + (8*2) = 29.

of $-\frac{1}{\delta}$ are positive and insignificant. We interpret this as implying that recent immigrants and previously arrived workers are perfect substitutes.

Next, we aggregate the wage equations to the occupation-age level. We can write an expression for $\bar{w}_{kljt}$, the weighted average of immigrant and native born wages in the sector-occupation-age cell, with the weights being the proportions of immigrant and native born workers. If we choose an arbitrary age group, j = 1, as the base group, we can take the difference between the mean wage of any other group and the base group to arrive at:

$$\ln(\frac{\bar{w}_{kljt}}{\bar{w}_{kl1t}}) = \ln(\frac{\theta_{kljt}}{\theta_{kl1t}}) - \frac{1}{\rho}\ln(\frac{L_{kljt}}{L_{kl1t}}) \tag{8}$$

Because we want to isolate variation within occupation-time cells to make it more likely that the relative θ values can be considered to be equal, we run this regression in differences including full sets of separate occupation, age and time dummies. This allows for separate trends for each age and occupation group. We instrument for the employment ratio using the log of the ratio of our $n10_{kljt}$ variable for each age group relative to the base group value for $n10_{kljt}$. The resulting first stage is strong (with an F-stat of 552). The omitted category is workers aged 25 to 34 and we drop that group from the estimation. Our IV estimate of $-\frac{1}{\rho}$ is -0.045 and is statistically significant at the 5% level. The implied corresponding elasticity of substitution is 22, which is large but considerably below the substitutability among immigrant and native born workers within age by occupation cells.

Aggregating up to the narrow occupation group level, we arrive at,

$$\ln \bar{w}_{klt} = \ln(\frac{\partial F}{\partial L_t}(K_t, L_t)) + \frac{1}{\sigma}\ln L_t + \ln \theta_{kt} - (\frac{1}{\sigma} - \frac{1}{\sigma_o})\ln L_{kt} + \ln \theta_{klt} - \frac{1}{\sigma_o}\ln L_{klt} \tag{9}$$

where $\bar{w}_{klt}$ is the average wage in the occupation cell in period t. If we assume that the first three terms on the right hand side of (9) can be captured by occupation and time effects then we can estimate (9) including those effects, using within occupation over time variation. As at the third level, we run the regression in differences and include occupation and time dummies. We instrument for the change in $\ln L_{klt}$ using the log of the $n10_{kljt}$ variable aggregated up to the narrow occupation level. The first stage F statistic is again large at 35.6. The resulting estimate of $-\frac{1}{\sigma_o}$ is -.10 and is statistically significant at the 5 percent level. The implied elasticity of substitution among narrowly defined occupations within a broad occupation group is 7.

Finally, we aggregate up to the broad occupation group level and estimate the analogous equation to (9) using the same methods. Again, the instrument is based on an aggregated version

of $n10_{kljt}$. The IV estimate for $-\frac{1}{\sigma}$ is -0.32, which is not statistically significant at the 5 percent level. In this case, we do have a weak instrument problem, with a first stage F-statistic of 6.7. However, based on the combination of this estimate, the OLS estimate, and other IV estimates reported in the appendix, we are confident that the estimated inverse elasticity is larger than that for the previous stages. In the simulation exercises in the next section, we use the IV point estimate, -0.32, which corresponds to an elasticity of substitution for broad occupation groups of 3.

4.1.2 Counterfactual Exercises

The estimated elasticities are of limited intuitive value on their own but can be used with the equation for native born wages, (3), to investigate the impact of immigration on native born wages. Our approach to quantifying that impact is to focus on the 1921-31 decade, constructing counterfactual wage distributions corresponding to: 1) no immigration between 1921 and 1931; and 2) no general or agricultural labourer immigration in that decade.[17] To construct the first counterfactual, consider generating a fitted wage for 1931 using equation (3) but with the number of recent arrival immigrants subtracted from each relevant labour measure on the right hand side. Importantly, this counterfactual is based on the number of new immigrants observed in the 1931 Census and so already nets out any effect that emigration would have in reducing the impact of initial immigration inflows. Since the productivity terms (the θ's and $\ln \frac{F}{L_t}(K_t, L_t)$) are assumed not to change with immigration, we can construct the counterfactual wage in 1931 as,

$$\ln w^c_{Nkljt} = \ln w_{Nkljt} - \frac{1}{\sigma} \ln I_t + (\frac{1}{\sigma} - \frac{1}{\sigma_o}) \ln I_{kt} + (\frac{1}{\sigma_o} - \frac{1}{\rho}) \ln I_{klt} + (\frac{1}{\rho} - \frac{1}{\delta}) \ln I_{kljt} \qquad (10)$$

We then construct a complete counterfactual distribution by creating a set of "phantom" people for each occupation by age cell. That is, if a given cell has n_{klj} native born workers in it in 1931 then we generate n_{klj} phantom people all with the fitted wage, $\ln w^c_{Nklj}$. Doing this for each cell, we obtain a dataset that has the same size and distribution of workers across cells as the true 1931 data but with fitted wages that reflect the estimated elasticities and the counterfactual quantities.
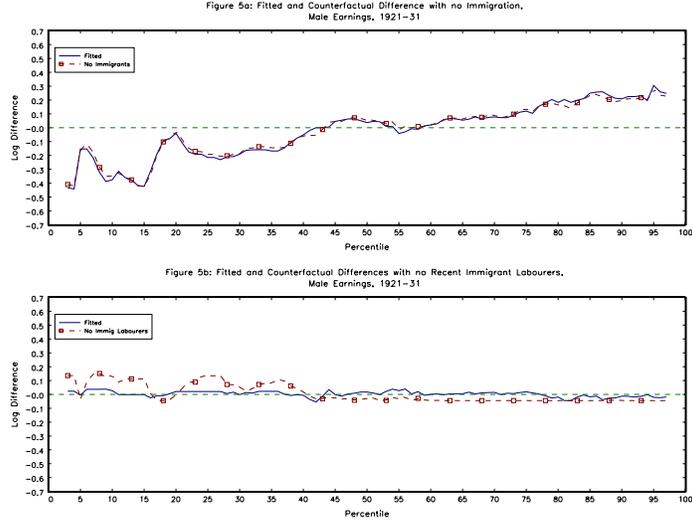
To capture the impact of not having immigration, we need to compare this counterfactual distribution to the actual distribution. This requires forming versions of the true distributions

---

[17]In the appendix, we also present a third counterfactual with no immigration of professionals in that decade as a point of comparison for 2).

that eliminate within occupation by age cell wage variation, just as the counterfactual exercise does. We do this for both 1921 and 1931 by again constructing a set of phantom workers for each cell but, in this case, assigning them the actual average wage for the relevant cell in the given year. We compare the result to the true change in the distribution in the appendix. We find that removing the within-cell variation does not substantially alter the picture of increased earnings inequality in the 1920s apart, although the extent of the decline in earnings in the lower tail of the distribution is reduced. In Figure 5a, we plot the difference in log percentiles from these fitted distributions along with the difference between the log percentiles of the 1931 counterfactual distribution with no recent arrival immigrants and the 1921 fitted distribution. The result is a small implied immigration impact across the distribution. This can be seen more clearly in Figure 6, where we plot the difference in log percentiles between the counterfactual distribution and the fitted 1931 distribution. This difference provides our key estimate of the net impact of immigration on the earnings structure in this period. Because the number of immigrants we remove from the 1931 labour force to construct the counterfactual come from the Census, they reflect the adjustments made by immigrants within Canada. Given our earlier arguments that the occupational distribution of recent immigrants is similar to that of the workers already in Canada, taking them away from the distribution implies small changes in the distribution. But the occupational distributions of recent arrivals versus established workers are not identical and the small impact partly reflects general equilibrium adjustments.

In Figure 5b, we plot differences corresponding to the counterfactual in which we eliminate only recent immigrant labourers. The implied effects of the labourers alone are substantial. Below the 40th percentile, removing the labourer inflow would have resulted in an average increase in native born real earnings of nearly 10 percentage points. Again, this can be seen clearly in the plots relative to the fitted 1931 distribution in Figure 6. There, one can also see the spillover impact of removing immigrant labourers on wages in other occupations. In particular, real earnings in occupations with above median wages would have fallen by approximately 5 percent. This highlights the arguments of OP that one needs to look beyond narrow occupation groups to measure the full impact of immigration. It also implies that the impact of low skilled immigration alone was to substantially increase inequality. The growth in the difference between the log 90th and 10th percentiles in the 1920s would have been reduced by nearly a third if no immigrant labourers had been admitted. The growth in the 50-10 differential would have been reduced by over 50 percent.

We can use different predicted impacts on farm labourers under various scenarios as a means

Figure 5a: Fitted and Counterfactual Difference with no Immigration,
Male Earnings, 1921-31

Figure 5b: Fitted and Counterfactual Differences with no Recent Immigrant Labourers,
Male Earnings, 1921-31

of highlighting the extent of the adjustment channels to an immigrant inflow. In the Appendix, we present estimates of the main specification from Borjas(2003) in which we regress decadal changes in native born wages on the proportional increase in farm labour that is due to recent immigration. [18] This specification can be interpretated as yielding estimates of the elasticity of demand if immigrants and native born workers were perfect substitutes and each occupation represents a separate labour market. That is, the estimated effects can be seen as reflecting immigrant effects on earnings when there are no further general equilibrium effects. Using the narrow occupation groupings in order to match with our more structural estimation, the

---

[18]That is, we estimate the regression, $\Delta \ln(w_{Njkt}) = \alpha_{0t} + \alpha_1 \frac{\Delta L_{Ijkt}}{L_{Njkt-1} + L_{Ijkt-1}} + \Delta u_{jkt}$, where $\Delta \ln(w_{Njkt})$ is the change in native born log earnings in occupation j and age group k, ΔL_{Ijkt} corresponds to the number of recent immigrants who were in that occupation and age category in the time t (e.g., 1931) Census, and $L_{Njkt-1} + L_{Ijkt-1}$ is the total number of workers (native born and immigrant) who were working in that occupation by age cell in the t-1 Census (1921 in this example). Note that the regression is based on Census data and thus the immigrant employment numbers are after immigrant adjustments.

Figure 6: Differences in Log Percentiles Between
Counterfactual and Fitted Distributions, Male Earnings, 1931
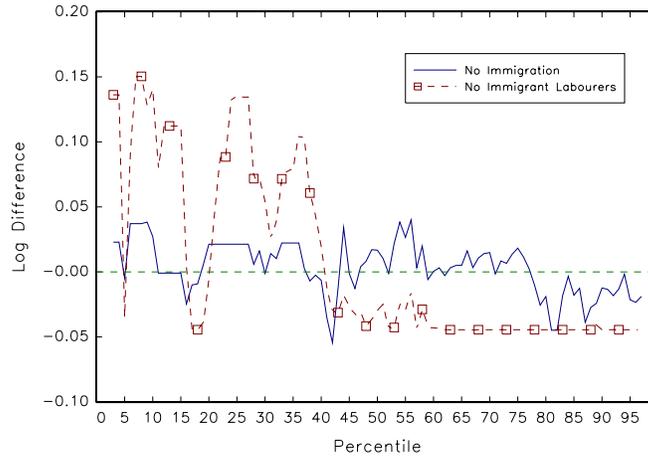
estimated coefficient on the proportion of workers who arrived in the previous decade is -0.35.
If all the immigrants who arrived in the 1920s had remained in Canada and in their intended
occupation, the inflow of new immigrants over the 1920s would have implied a 158 percent
increase in the farm workforce and thus a 42 percent decline in worker earnings in that sector.[19]
However, the actual number of farm workers in 1931 who were immigrants who had arrived in
the previous decade (i.e., the number after accounting for adjustments through re-migration and
occupational switching) amounted to a 39 percent increase in the farm workforce, and thus would
imply a 13 percent decline in farm worker earnings in the absence of other general equilibrium
effects. Finally, making use of our production function parameter estimates and the full set of

---

[19]For this calculation, we adjust the total inflow for the fact that the inflow statistics include farmers as well as
farm labourers. To do this, we multiply the inflow in the "Farmers and Farmer Labourers" category by the proportion
of the sum of farmers and farm labourers who were farm labourers among immigrants who had arrived in the previous
decade in the 1931 Census (0.65). For this calculation and the ones that follow, we use the number of males employed
as farmer labourers in the 1921 Census as the base.

immigration inflows, our first counterfactual implies that the total immigration inflow implied a 4% drop in farm worker earnings. The reduction from a 13 percent to a 4 percent effect represents the impact from taking account of general equilibrium adjustments in the economy.

The counterfactual with no immigration implies that immigration had only limited impacts on the location of the native born earnings distribution. This fits with findings in Pope and Withers (1994) for Australia and Greasley and Oxley (2004) for Australia. Taylor and Williamson (1997) and Hatton and Williamson (1998) generate much larger wage impacts from immigration using CGE models with capital held fixed, but those impacts are reduced by as much as 75 percent once capital is allowed to follow labour. Green and Sparks (1999) find evidence of such endogenous capital inflows for Canada before WWI using a cointegration methodology. A finding of immigration having small impacts on native born wages also fits with the literature examining recent immigration (e.g., Ottaviano and Peri (2012)). Christian Dustmann and Ian Preston (2011) argue that the type of method we use can understate immigration impacts because apparently skilled immigrants based on their reported education actually compete for jobs with less skilled native born workers. Because we match immigrants to the native born in terms of the occupations they actually work in and because immigrant and native born occupational distributions are very similar, this is unlikely to be problem here. We conclude that large inflows in specific occupations can have substantial impacts on the wages in that occupation and in closely substitutable occupations but those impacts can be mitigated by positive spill-over effects due to the arrival of a wider set of immigrants. In our case, the effects of those other immigrants implies near zero net effects of immigration even in low skilled occupations.

5 Conclusion

Using newly available micro Census data, we have shown that the Canadian earnings structure underwent a substantial increase in inequality between 1911 and 1941, with most of the increase occurring between 1921 and 1931. Given that the immigrant inflow in the 1920s was particularly low skilled, it seems plausible that there was a connection between immigration and the increase in inequality in that decade. Our conclusion, however, is that immigration played a minor role as a driver of the changes in the earnings structure. We show that this happened largely because of adjustments made by immigrants themselves. In the 1910s, when there was little policy control over immigration, the people who chose to immigrate to Canada were those with skills that matched the occupational demand in the economy. This is evident in the fact that their occupational distribution was similar to the native born at the time they entered Canada. In

contrast, in the 1920s, the government actively shaped immigration, with the result that the inflow was dominated by farmers and farm labourers. But by the end of that decade, recent immigrants had shifted occupations such that their occupational distribution looked similar to that of the native born. In both decades, with the immigrant occupational distribution being similar to that of the native born, we would expect the immigrant inflow to affect the level of the earnings distribution much more than its shape. In addition to these adjustments, there were general equilibrium effects of immigration on wages that also served to spread out the effects of the mass immigration. Our estimates imply that without any adjustment, immigration would have reduced the wage for farm labour by 42 percent in the 1920s but with all the adjustment channels, the ultimate impact was a reduction of only 4 percent. Of this 38 percent reduction in the wage effect, three-quarters was due to geographic and occupational adjustments made by the immigrants themselves and one quarter was because of general equilibrium price adjustments. Thus, the driving force for the substantial increase in inequality in Canada in this era must have been something other than immigration.

6 References

Abramitsky, Ran, Leah Platt Boustan, and Katherine Eriksson. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122, no. 3 (2014): 467-717.

Anderson, Edward. "Globalisation and Wage Inequalities." *European Review of Economic History* 5 (2001): 91-118.

Bertram, Gordon W. and Michael B. Percy. "Real Wage Trends in Canada 1900-26: Some Provisional Estimates." *Canadian Journal of Economics* 12, no. 2 (1979): 299-312.

Borjas, George J.. "The Labor Demand Curve Really is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market." *Quarterly Journal of Economics* 118, no. 4 (2003): 1335-74.

Card, David. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43, no. 2 (1990): 245-57.

Card, David. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19, no. 1 (2001): 22-64.

Chambers, Edward J. and Donald S. Gordon. "Primary Products and Economic Growth, An Empirical Measurement." *Journal of Political Economy* 74, no. 4 (1966): 315-32.

D'Amuri, Francesco, Gianmarco I.P. Ottaviano, and Giovanni Peri. "The Labor Market Impact of Immigration in Western Germany in the 1990s." *European Economic Review* 54 (2010): 550-70.

Dustmann, Christian and Ian Preston. "Estimating the Effect of Immigration on Wages," CREAM Working Paper No. 21/11, London, UK, 2011.

Emery, J.C.Herbert 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): 115-137.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. "Unconditional Quantile Regressions." *Econometrica* 77, no. 2 (2009): 953-73.

Goldin, Claudia And Lawrence F. Katz. "Decreasing (and Then Increasing) Inequality in America." In F. *The Causes and Consequences of Increasing Inequality* edited by Finis Welch,

37-82 Chicago: The University of Chicago Press, 2001.

Goldin, Claudia and Robert Margo. "The Great Compression: The Wage Structure in the United States at Mid-Century." *Quarterly Journal of Economics* 107, no. 1 (1992): 1-34.

Greasley, David, and Les Oxley. âĂIJGlobalization and Real Wages in New Zealand 1873-1913.âĂİ *Explorations in Economic History* 41 (2004): 26-47.

Green, Alan G.. "The Political Economy of Immigrant Selection in Canada, 1896-1930." Department of Economics working paper, Queen's University, Kingston, Ontario, Canada, 1993.

Green, Alan G.. "International Migration and the Evolution of Prairie Labour Markets in Canada, 1900-1930." In *Migration and the International Labor Market, 1850-1939* edited by Timothy J. Hatton and Jeffrey G. Williamson, 156-74 London: Routledge, 1994

Green, Alan G.. "A Comparison of Canadian and US Immigration Policy in the Twentieth Century." In *Diminishing Returns: The Economics of Canada's Immigration Policy* edited by Donald J.DeVoretz, 31-66 Toronto: C.D. Howe Institute, 1995.

Green, Alan G. and David A. Green. ""Balanced Growth and the Geographic Distribution of European Immigrant Arrivals to Canada,1900-14." *Explorations in Economic History* 30 (1993): 31-59.

Green, Alan G. and David A. Green. "Canada's Wage Structure in the First Half of the Twentieth Century (with Comparisons to the United States and Great Britain)". UBC Economics Working Paper, Vancouver, Canada, 2008.

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

Green, Alan G. and Gordon Sparks. "Population Growth and the Dynamics of Canadian Development: A Multivariate Time Series Approach." *Explorations in Economic History* 36 (1999): 56-71.

Hatton, Timmothy J. and Jeffrey G. Williamson. *The Age of Mass Migration.* New York: Oxford University Press, 1998.

Kelley, Ninette and Michael Trebilcock. *The Making of the Mosaic: A History of Canadian Immigration Policy*, Second Edition. Toronto: University of Toronto Press, 2010.

Lew, Byron. "European Immigration to Canada During the 1920s: The Impact of US Quotas and Canadian Restrictions," Department of Economics working paper, Queen's University, Kingston, Ontario, Canada, 2000.

Lewis, Frank. "The Canadian Wheat Boom and Per Capita Income: New Estimates." *Journal of Political Economy* 83, no.6 (1975): 1249-57.

Mackinnon, Mary(1996). "New Evidence on Canadian Wage Rates, 1900-1930." *Canadian Journal of Economics* 29, no. 1 (1996): 114-131.

Ottaviano, Gianmarco I.P. and Giovanni Peri (2012). "Rethinking the Effects of Immigration on Wages." *Journal of the European Economic Association* 10, no. 1 (2012): 152-97.

Pope, David H., and Glenn Withers (1994). "Wage effects of immigration in late-nineteenth century Australia." In *Migration and the International Labor Market, 1850-1939* edited by Timothy J. Hatton and Jeffrey G. Williamson, 235-57 London: Routledge, 1994.

Taylor, Alan M. and Jeffrey G. Williamson. âĂIJConvergence in the Age of Mass Migration,âĂİ *European* Review of Economic History 1 (1997): 27-63.

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

Williamson, Jeffrey G.. "The Evolution of Global Labor Markets Since 1830: Background Evidence and Hypotheses," *Explorations in Economic History* 32 (1995): 141-96.

Williamson, Jeffrey G.. "Inequality and Schooling Responses to Globalization Forces: Lessons from History," in *Migration, Trade and Development.* Dallas: Dallas Federal Reserve Bank, 2007.