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

Alan G. Green*and David A. Green†

Novemer, 2014

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 second and third decades of the 20th century were, respectively, the largest and fourth largest immigrant inflow decades in Canadian history. Between 1910 and 1919, an immigrant inflow amounting to approximately 27% of Canada's 1910 population arrived in Canada, and between 1920 and 1929, the inflow equalled 15% of the 1920 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% 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% of male employment in 1921 and 14% in 1931.[2] Those percentages are at least double the size of the Mariel Boatlift shock 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%). Moreover, in 1931, 38% of male wage earners in Canada were foreign born. Thus, whatever measures we use, Canada received substantial immigration shocks in the 1910s and 1920s. These large inflow decades for Canada (and the US) 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% 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. 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 the potential endogeneity of occupatonal choices by immigrants by using 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

[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.

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 (e.g., Lewis(175)), to the best of our knowledge, there is no prior empirical analysis of the impact of the large immigrant inflows into Canada on its wage structure for any part of the pre-WWII period.

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. To understand that, it is useful to understand the nature of Canadian immigration policy in the first half of the Twentieth Century. Prior to WWI, Canadian policy could best be described as laissez-faire: refusing entry based only on disease, criminality, and racist reasons. Following our earlier work (Green and Green(1993)), we show that this resulted in an immigrant inflow with an occupational distribution that was very similar to that of the native born and previously arrived immigrants. Because of that, the immigration inflow had little impact on the shape of the earnings distribution in the first decade of our sample period (from 1910 to 1920). This is similar to what Abramitsky et al(2014) argue for the US. It suggests an international labour market that reached across the Atlantic such that immigrants were not just blindly arriving in New World countries unaware of the shape of labour demand in those countries. In the 1920s, a dominant component of Canadian immigration policy was an agreement with the main railway companies under which they brought they brought in large numbers of farm labourers destined for the West. We show that these immigrants adjusted geographically (both by moving from Western to Central Canada and by leaving Canada, likely for the U.S.) and occupationally (shifting to other low and semi-skilled occupations). These adjustments spread the immigration shock across the economy in a way the implied a lower impact on the shape of the native born earnings distribution. 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. The comparison between this counterfactual and the actual earnings distribution provides our main conclusion on the ultimate impact of immigration on the earnings structure. In contrast to that counterfactual, 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% 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 approximately 42% 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%, and once we allow for general equilibrium effects, it is further reduced to a 4% 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 complementarity and 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.

Our third goal in the paper is to document movements in Canadian earnings inequality in the first half of the Twentieth Century. To do this, 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 and weeks worked 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. Ours, of course, is not the first depiction of Canadian wage movements. Accessible wage and price series have been developed by several authors (Bertram and Percy(1979), Mackinnon(1996), Emery and Levitt(2002)) and there are wage series in the Historical Statistics of Canada (Urquhart and Buckley(1965)). But these are either aggregated or relate only to a few occupations in the construction and manufacturing trades plus labourers, while our data represents the entire occupational wage structure. Using the Census data, 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.

The remainder of the paper proceeds in 5 sections. In section 1, we describe Canada's immigration policies and immigration patterns, documenting both the size of the inflows and

the extent of the adjustments made by immigrants. In section 2, we describe our data and use it to document the main changes in the earnings distribution in Canada in this period, including a description of changes in age and occupational earnings differentials. In section 3, we present the results of the decomposition of changes in the earnings distribution using the Firpo et al(2012) approach and also show relative immigrant and native born earnings distributions. In section 4, we estimate the impact of the immigration inflows on native born wages and use them to generate counterfactuals. Section 5 concludes.

1 Immigration Policy and Patterns

We begin with an overview of Canadian immigration policy before WWII. As we will see, understanding changes in immigration policy in this period is helpful in interpreting how immigration affected the earnings structure in different decades. 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. This policy was seen as having two main benefits. The first was that a settled west would provide a market for the burgeoning manufacturing sector in Central and Eastern Canada. The second was to keep the West as part of Canada - safe from American aggression.[3] These objectives were to be attained through three inter-related policies, referred to collectively as the National Policy: 1) raising tariffs to support Central and Eastern Canadian infant industries; 2) building a transcontinental railroad; and 3) acquiring Rupert's Land from the Hudson's Bay Company. At the heart of these tripartite policies was immigration, i.e., a policy to fill the empty lands of the West as quickly as possible.

To carry out the latter policy, the Department of Agriculture (and later the Department of the Interior) established immigrant agents in the US, Britain and, to a limited extent, Europe. Their job was to attract not just immigrants to Canada but farmers and farm workers who would be willing to settle in the West. Importantly, while the government had clear preferences for British and American immigrants, its approach before 1906 was welcoming to farmers or farm workers from a much broader range of countries (Kelley and Trebilcock(2010)).[4] Moreover,

[3] Aitken(1969) termed this second goal "Defensive Expansionsim."

[4] Clifford Sifton, the Interior Minister from 1896 to 1905, famously stated in talking about what constituted a good quality immigrant,"I think that the stawart peasant in a sheepskin coat, born to the soil, whose forefathers have been farmers for ten generations, with a stout wife and half-dozen children is good quality." (quoted in Kelley and

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))[5]

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 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 brought with it 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 what was a highly seasonal production cycle. The continued leading and, indeed, expanding role of agriculture in Canada's economic performance distinguished it from its southern neighbour which had progressed more rapidly to focusing on manufacturing (Wright(1990)).

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 apart from the old ones of not being prospective criminals or wards of the state. 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 of any kind 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 (a requirement put in place by Order in Council in 1923) in conjunction with the strategic

Trebilcock(2010), p. 122)

[5]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.

placement of visa offices in preferred as opposed to non-preferred countries. Lew(2000) argues effectively that 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 - except for one factor. In 1925, the Canadian government entered into agreements with the two main railway companies for them to recruit agricultural labour in non-preferred countries. Lew(2000) shows that in the remaining years of the 1920s there was a surge of immigration dominated by immigrants from the non-preferred countries of Europe and headed for the West.[6] The flexibility of Canadian policy allowed the government to find a way to recruit labour targeted directly to farms in the West. [7] Green(1993) argues that this policy is a reflection of the rising economic power of farmers in the West, expressed through farm centered political parties.

Finally, with the onset of the Depression, the Canadian government used its powers to all but shut down immigration.

1.1 Immigration: Flows and Stocks

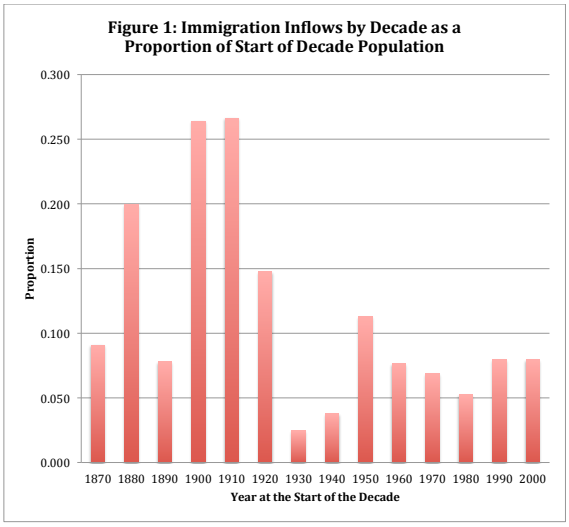
Figure 1 shows the decadal immigration inflows to Canada for each decade since Confederation expressed as a proportion of the start of decade population. Even with the relative lack of immigration during WWI, the decade from 1910 to 1919 was the highest inflow decade on record, with the inflow being an astonishing 26.6% of the 1910 population. The 1920s inflow was smaller at 14.8% 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 it. In contrast, the 1930s were the lowest immigration years on record, with the decadal immigration falling to 2.5% of the 1930 population.

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).[8] 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

[6]The government tried actively to bring in British farmers and farm labourers in the 1920s through the Empire Settlement Agreement and other initiatives. It was the substantive failure of these initiatives that led to the railway agreements (Kelly and Trebilcock(2010)).

[7]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)).

[8]The differences in the estimated flows related to differing assumptions on mortality rates and whether to treat transient entrants as immigrants then emigrants or not to include them in the gross flows at all (see Keyfitz(1950) and McDougall(1961)).



Source: Statistics Canada, CANSIM Tables V742019 (Estimated Population) and V742546 (Immigrant Arrivals).

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 in the Canadian labour market, at least for a while. 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.[9] 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

[9]In comparison, in 1931, 43% of immigrant male arrivals in the prior decade were working in Canada. The percentage is higher in the 1931 Census likely because of a tightening of US immigration policy. The US attempted to close the Canadian back door by requiring in their 1921 Act that European immigrants entering the US through another Western Hemisphere country had to have lived in the Western Hemisphere for at least a year. In 1922, the requirement was raised to five years.(Lew(2000)). Sources: Flow data: Urquhart(1965), p. 24. 1921 Stock data: Census of Canada, 1921, Volume 2, p.61; and Census of Canada, 1931, Volume 6, Table 31.

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.[10] As a comparison, in 2011, 6% of the labour force was made up of immigrants who had arrived in the previous decade. From figure 1, the 2000s was a high inflow decade by post-war standards but not by pre-war standards. The pre-1930 decades were characterized by very large annual inflows partly offset by 50 to 60 per cent re-migration rates while the post-war decades have been characterized by smaller inflows but approximately 30% re-migration rates by the end of a decade (Aydemir and Robinson(2008)). The net outcome was the same in both periods: Canada historically has been a recipient of immigrant inflows that were large relative to its labour force. Our goal is to study the impact of those inflows on the earnings distribution in the early part of the Twentieth Century. In order to make sure we get estimates of the true impacts, 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.

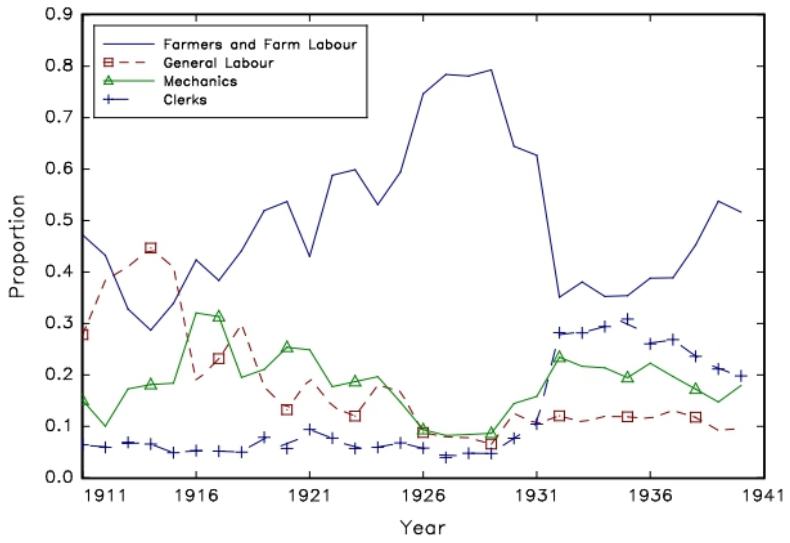
Of course, the potential impact of immigration on the labour market is related to its composition as well as its size. Figure 2 shows proportions of male immigrants with a reported intended occupation for each year from 1911 to 1940.[11] In the years just prior to WWI, agricultural workers made up nearly 50% of the inflow but there was clearly a strong representation from other occupations. Green and Green(1993) examines the intended occupations of a micro sample of immigrants arriving by sea in 1912 and shows that immigrants had a balanced mix of intended occupations that matched the sectoral growth in the Canadian economy. This fits with an era which we have seen had an essentially laissez-faire immigration policy. During the War years, when the inflow consisted mainly of Americans, the importance of agricultural immigrants as a proportion of the inflow rose. Then in the 1920s, as the government began to employ its selective powers, that proportion continued to rise. The onset of the Railway Agreement is clearly visible, as the proportion listing an intended occupation in farming jumped to approximately 80% of the inflow in the late 1920s. Lew(2000) reports that the intended destinations of immigrants within Canada also reflected the impact of the Railway Agreement as 80% of immigrants entering from non-preferred countries in 1925 stated an intention of settling in the West. The proportions in

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

[11] The series are from the Parliament of Canada Sessional Papers for each year. Specifically, they are from what are essentially the departmental annual reports from the Department of the Interior (up to 1917) and then the Department of Immigration. 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 plots do not include the Miners occupation category.

the 1930s were more of a mixed bag but recall that this was a relatively low inflow period.

Figure 2: Annual Immigration, Occupation Proportions



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 US, others returned home and still others switched occupations within Canada. The net result might be an immigrant occupation distribution that bears only limited resemblance to the distribution of their stated intended occupations at time of arrival. To investigate this, in Table 1 we present distributions for males across broad occupations from the immigration department flow data and the Census.[12] Specifically, we sum the number

[12]We describe the Census data in detail in the next section. The occupation categories are from the Sessional Papers inflow 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 classify as semi-skiilled may belong in the "Mechanics" category in the immigration data. We put all skilled trades in the latter

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.

of immigrants entering in each 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 quite 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 did not become farmers to nearly the same extent as the native born but otherwise were in very similar occupations. That similarity arises both because, under a laissez-faire immigration

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.

policy, the migration choices made in the source countries resulted in an inflow that matched, in many ways, 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 US in this era also matched well with those of the native born. As described earlier, in Green and Green(1993), we also find that intended occupational and regional destinations of immigrants match the distributions of the native born quite closely in this period. 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.

As we described earlier, Canada had an active immigration policy in the 1920s, with a strong emphasis on agricultural labour. The resulting concentration in farming occupations in stated intentions of immigrants arriving in the 1920s was evident in Figure 2, and shows up again in Table 1. By 1931, however, the proportion immigrants who arrived in the 1920s in farming had 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, it was still quite different from the 1931 native born distribution. In particular, even after dropping farmers, it 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 made considerable adjustments after arrival. While the result is not as close a match to the occupational distribution of workers already in Canada as was evident in 1921, it still represented a significant step in that direction and, thus, a significant step in the direction of mitigating the expected effects of immigration on the labour market.

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 break the table down between the West, and Central and Eastern Canada. Several points emerge from this table. First, an amazing 61% of the male workforce in the West in 1931 was foreign born. This again highlights the size of the immigration shock. Second, more 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 (immigrants who are likely to have entered under the Railway Agreement) were in farming, but even for that group a larger proportion were general labourers. Fifth, 71% 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% 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 shock in the West. However, it is also evident that this shock spilled over into other regions and other occupations. Based on our earlier discussion, it is possible that some of the adjustment in occupational composition came about through differential tendencies to re-migrate by occupation group. However, there is no way for us to know much of the net change in the occupational distribution after arrival comes through that route versus through movements within Canada. Finally, one implication of the post-arrival adjustments is 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 new 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). As we will discuss, 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).[13] The samples correspond to

[13]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

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

A: West

	Native Born	Immigrants			
		Pref <10 yrs	Pref >10 yrs	Not Pref <10 yrs	Not Pref >10 yrs
Clerical	8.2	3.7	7.1	0.3	1.2
Farm Labour	18.2	28.8	8.1	32.9	8.1
Farm Managers	1.5	1.9	1.2	1.1	0.6
Labourers	17.3	23.3	16.0	37.8	41.3
Managers and Officials	5.9	2.6	7.2	0.5	1.6
Professional	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	Immigrants			
		Pref <10 yrs	Pref >10 yrs	Not Pref <10 yrs	Not Pref >10 yrs
Clerical	6.5	5.0	7.5	0.8	1.7
Farm Labour	6.5	17.4	3.6	5.5	1.6
Farm Managers	0.9	1.5	0.4	1.1	0.5
Labourers	22.7	20.4	14.8	48.4	32.9
Managers and Officials	4.4	2.8	4.5	1.0	2.9
Professional	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.

5%, 4%, 3% and 3% 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.

The key variables for our purposes are annual earnings and weeks worked. Both earnings and weeks worked refer to the twelve month period preceding June 1 of the Census year. Annual earnings refer to wage, salary, commission or piece rate earnings from all jobs in that period. 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.[14] 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.[15] We restrict our sample to males aged 15 to 64.

We construct a weekly wage by dividing annual earnings by annual weeks worked. However, as described in the appendix, weeks worked were elicited differently in 1911 and 1941 compared to 1921 and 1931. This raises issues about the consistency of the weekly wage measure over our whole period and, as a result, we focus mainly on results using annual earnings.[16] We convert earnings and weekly wages 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 project used

provided in the Appendix.

[14]In 1946, among males, there were 683,000 farm owners, 143,000 paid farm workers, and 245,000 unpaid family workers (Urquhard(1965, p. 63).

[15]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.

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

answers to questions on occupation and industry in the original Census data to map into a joint occupation and industry coding corresponding to the Integrated Public Use Microdata Series (IPUMS) system that is itself based on the 1950 US Census codes. Working with these codes, we formed 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. We will call these 146 occupations 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 managers, clerical workers, managers, professionals, sales occupation workers, service workers, semi-skilled workers, and tradesmen.[17]

Finally, we group individuals into 4 age categories: 15 to 24; 25 to 39; 40 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. We drop all age by occupation cells with fewer than 10 observations in any year.

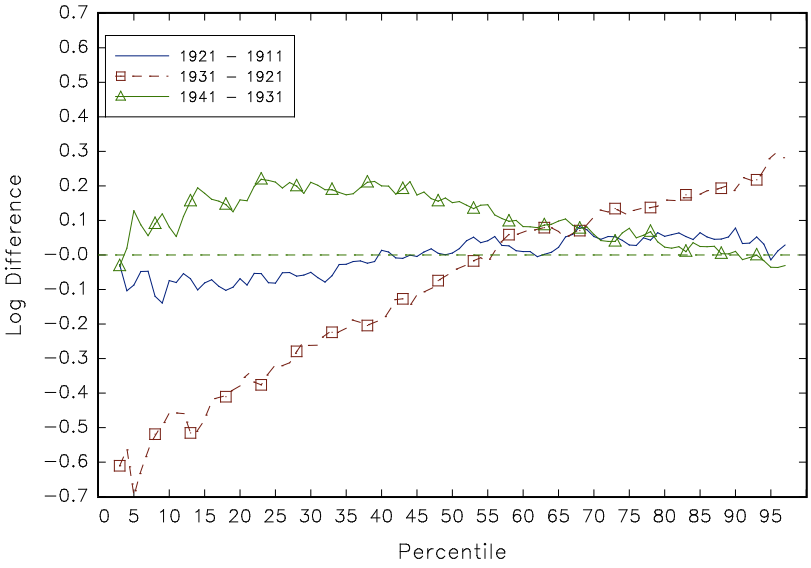
2.1.1 Changes in Earnings and Wage Distributions

Figure 3 contains a depiction of 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 percentage 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. The horizontal dashed line corresponds to zero change between the years. Notice that we are comparing percentiles not specific people across years. Thus, when we report that, say, the 10th percentile declined by 15% this does not necessarily mean that a person who was at the 10th percentile in 1911 experienced a decline of 15% 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. In figures of this type, a line sloping up to the right reflects an increase in inequality between the pair of years because in that case increases at the

[17]Occupation corresponds to jobs held at the time of the Census, while earnings and weeks worked correspond to all jobs in the previous twelve months. Thus, our calculated average earnings for an occupation may not be the true average earnings for people who work only in that occupation. For example, if workers employed in a skilled occupation at the Census date spent part of the year working as labourers then our calculated average earnings would be lower than the average earnings for a person employed solely in that occupation. Comparisons to other data indicate that this is not a concern for the least skilled workers (e.g., labourers) or most skilled (e.g., professors) in Census data but may be important when considering the earnings of tradesmen (Green and Green(2008)).

top of the distribution are greater than at the bottom (or decreases are less).

Figure 3: Decadal Differences in Log Percentiles, Male Earnings

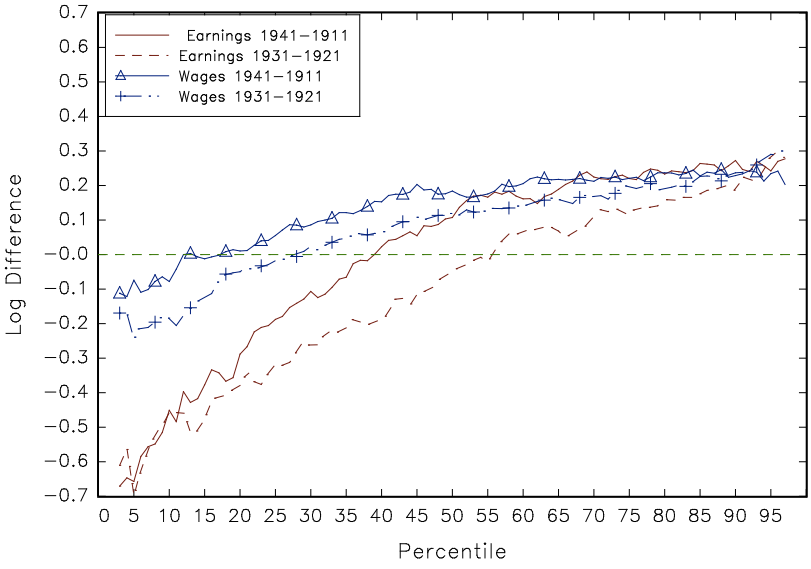


Looking at the line capturing the difference between 1911 and 1921, there is slight evidence of an increase in inequality as the percentiles below about the 30th decline by close to 10% in real terms while those in the upper end of the distribution increase by just under 10%. These movements, however, 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% 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 there is very little in the way of changes at the top of the distribution compared to approximately 20% increases through much of the lower half of the distribution. This pattern is not enough to offset what happened over the 1920s, though, so that the net effect over the whole period from 1911 to 1941 is a strong increase in earnings inequality.

In Figure 4, we show the 1911 - 1941 difference for annual earnings and reproduce the 1921-31 difference. We focus on these differences in order to compare to the weekly wage differences that are presented in the same figure. Recall that these two differences are the ones that are the

most reliable for weekly wages because of changes in questions about weeks of work over time. The 1911-41 difference for earnings shows that there is a strong long-term increase in inequality below the median but a relatively constant increase in earnings levels across percentiles above the median. For weekly wages, the long term pattern has a roughly similar shape to annual earnings, though much more muted and with increases in level across much of the distribution. Whereas with annual earnings there are declines at virtually all percentiles below the median, in weekly wages there are declines only below the 15th percentile. As with annual earnings, the increase in inequality in the weekly wage is driven mainly by movements in the 1920s.

Figure 4: Differences in Log Percentiles, Male Earnings and Weekly Wages



One immediate point of interest is the much more substantial decline in the lower half of the annual earnings distribution than is evident in the weekly wage distribution. Given that weekly wages are constructed as annual earnings divided by weeks of work, the difference must reflect a drop in weeks of work. Indeed average weeks worked per year declined from 44 in 1921 to 42 in 1931. This may reflect selection issues related to the fact that approximately a third of paid employees in the 1921 Census do not report their weeks of work. [18] However, this would not

[18] A decline in weeks of work from 1921 to 1931 may also seem surprising given the labour market problems in 1921. However, while 1921 was a bad labour market year, the labour market in the summer and fall of 1920 was actually strong (exactly when it began to turn bad - either in October or in December - depends on whether the source of information is unions or employers (Dominion Bureau of Statistics (1921)). The 1921 Census asks about

explain the substantial difference between wage and earnings patterns for the 1911-1941 change since there are no such selection issues in either year. In the end, given the changes in weeks worked questions and the possible selection effects in 1921, we conclude that it is preferable to focus on annual earnings, where we have a consistent measure over time.

Overall, the figures in this section provide a portrait of Canada undergoing a substantial long-term change in its earnings structure in the period from just before WWI to just before WWII. 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. 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 Appendix B. Here, we summarize our conclusions. The Labour Gazette was published by the Department of Labour and includes reports from agents across Canada on local labour market conditions including wages and hours of work. The wage data for Labourers from the Gazette is in strong agreement with what we observe in the Census. In that sense, what we show as movements in weekly wages near the 25th percentile (where labourers are located) in Figure 4 fit with what is observed in a different data source. On the other hand, the wage series for Machinists show differences. Both the Census and Gazette data indicate an increase in the trades/labour skills differential over the entire 1911-1941 period but with very different timing since the Gazette data indicate that the main increase in the differential occurred in the 1910s. Mackinnon(1996) argues that the Gazette data for the skilled trades primarily correspond to union contracts and since the unionization rate was only about 11% in this period, do not accurately reflect the main wage movements in the economy. Mackinnon(1996)'s wage data from the records of the CPR show the same pattern of a declining trade/labourer differential in the 1910s and increase in the 1920s as we see in the Census data, but with the 1910s decline being larger than in our data. However, the CPR wage data used in Mackinnon(1996) reflect wages that were under

labour market experiences over the 12 months preceding the start of June Census date. Thus, 7 out of the 12 months are from 1920 and, as a result, the 1920/21 Census data does not represent as bad a labour market as one might expect given the 1921 Census date. The 1930/31 Census data represents a similarly mixed year.

direct government control until into the 1920s and themselves correspond to union agreements. Our main conclusion is that, while the other data show some of the same broad patterns as in the Census, the combination of wage patterns and coverage of the other datasets indicates that Census data is likely to be more reliable.

But 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 these ratios between specific occupations - most often between skilled, blue collar trades or clerks and labourers. In fact, both Emery and Levitt(2002) and Mackinnon(1996) present values for wages for these types of workers from the published Census tables. Our 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 1 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 change in the shape of the earnings distribution simply because of adding a disproportionate number of workers in the lower part of the distribution. In this section, we decompose the decadal changes in Canada's earnings distribution into factors related to immigrant status, age, and occupation. There are no measures of educational attainment in the Censuses and so we focus on occupation rather than education as a measure of skill.

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 changes in the proportion of immigrants by occupation. Table 1 also indicates that the native born occupational composition was relatively constant over this period.

In order to investigate the between group wage differentials, in Table **??** we present differences in average log wages between groups defined by 10 occupation groups and 5 10-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 .19 log points in 1921 but .32 log points in 1931. Second, the earnings differentials between occupation groups increased substantially, particularly among blue collar occupations and between white collar and blue collar occupations. As an example, the earnings differential between trades occupations and semi-skilled workers and labourers was .07 and .46 log points, respectively, in 1921. These rose to .13 and .78 in 1931. At the same time, 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, 1.23, 1.65 and 1.63 in 1911, 1921, 1931, and 1941, respectively. Thus, within occupation earnings dispersion was constant between 1911 and 1921, jumped dramatically between 1921 and 1931, and then was relatively constant again between 1931 and 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.

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
Labourer	0	0.18	0.19	0.20	0.12
Clerical	0.23	0.72	0.78	0.88	0.87
Farm Labour	-0.49	-0.22	-0.18	-0.24	-0.32
Farm Managers	-0.41	-0.01	0.10	0.18	0.03
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					
Labourer	0.00	0.29	0.34	0.28	0.19
Clerical	0.33	0.81	0.94	0.78	0.87
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
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
Clerical	0.93	1.50	1.69	1.67	1.54
Farm Labour	-0.28	-0.01	0.01	-0.03	-0.09
Farm Managers	-0.03	0.38	0.68	0.64	0.38
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
Clerical	0.37	1.06	1.32	1.44	1.32
Farm Labour	-0.60	-0.24	-0.19	-0.28	-0.44
Farm Managers	-0.44	0.05	0.31	0.41	0.29
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

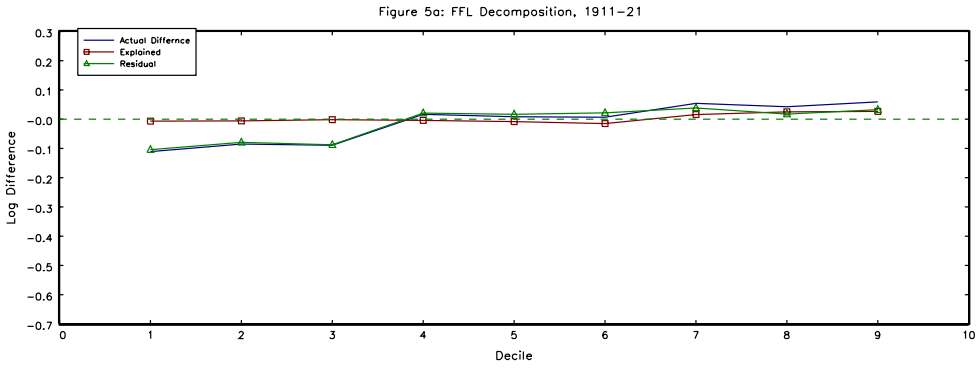
Sources: Census, various years. The table entries for each year show the difference in average log wages between each occupationxage category and the average log wage for general labourers aged 15 to 24 in that year.

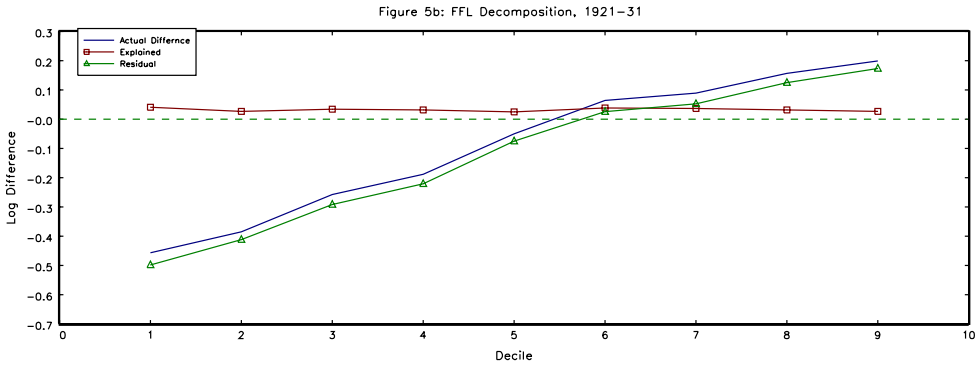
To judge the relative importance of the composition shifts versus the changes in the wage structure, we employ the ingenious decomposition technique developed in Firpo et al(2009) (hereafter, FFL). At the core of the FFL methodology are influence functions, which show the influence of an observation on the value of a given statistic. Of particular interest for us 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 analogous to the way movements in a sample mean can be decomposed in 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 5, 10-year groups); occupation (in the 146 groups described earlier); whether a person was an immigrant who arrived in the 10 years prior to the Census; and whether a person was an immigrant who arrived more than 10 years before the Census (native born workers form the omitted category).[19]

In Figure 5a, (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 3 but only showing the decile points. The line labelled "Explained" shows changes in a given decile that would be predicted as happening purely because of 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 and 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 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%, 60%, and 45% of the movements at the 7th, 8th and 9th deciles, respectively. For the 1921-31 changes shown in Figure 5b, 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%, 20% and 13% 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, are not a major determinant

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



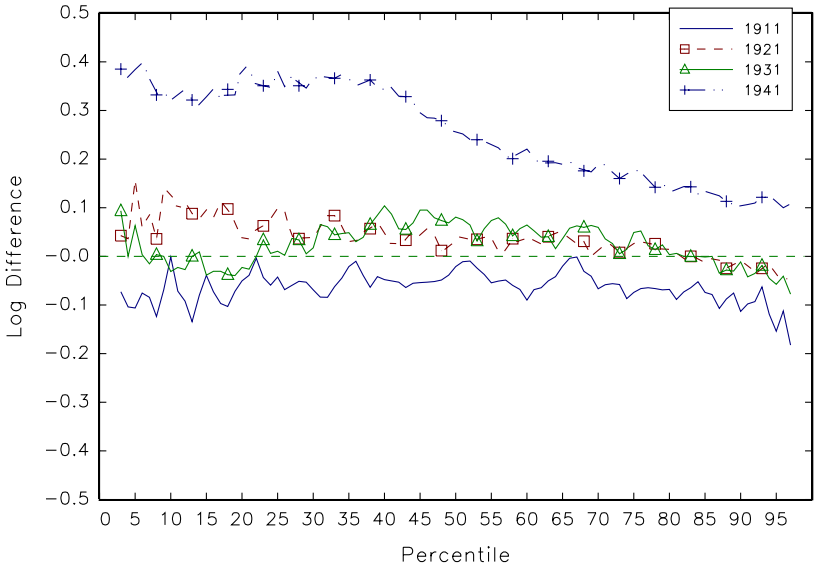


of movements in the real earnings distribution. Most importantly, they explain little of the substantial decline in the lower tail of the real earnings distribution in the 1920s. That decline arises from a combination of the increased earnings differentials between age-occupation groups and the increase in within-group dispersion. As part of the counterfactual exercises in section 4, we generate the change in the earnings distribution that would have arisen if there had been no within-group variation. Based on that exercise, the relative importance of the "between" versus "within" components varies considerably in the lower half of the distribution, with the change in within occupation dispersion (the "within" component") accounting for 80% of the decline in the 5th percentile between 1921 and 1931 but only 20% of the decline in the 10th percentile. Above the median, changes in within group dispersion plays essentially no role in explaining changes in the earnings quantiles over this decade.

To better understand the role of immigration in these decompositions, in figure 6, 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 at most points being under 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 at that point and this difference likely just reflects returns to experience rather than a true immigrant advantage. Given that, as we have seen, immigrant and native born occupational distributions were broadly similar, the similarity of the earnings distributions implies that there was not systematic discrimination (or differentials in productivity) between immigrants and the native born in the workforce.[20] It also accounts for 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 shocks altering occupation specific wages over time. To assess the latter potential impacts we will need to introduce more economics, and we turn to that in the next section.

Figure 6: Immigrant minus Native Born Differences in Log Percentiles, Male Earnings



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

In this section, we examine the question of whether immigration affected earnings inequality in inter-war Canada using techniques from recent papers on the impact of immigration on the

[20]Based on wage data and referring to the report of the US Immigration Commission, Hatton and Williamson(1998) similarly argue that "immigrants received roughly the same wages as natives." In contrast, Green and 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.

native born wage distribution. Given the evidence in section 2 showing that, for example, immigrants listing an intention of working on a farm in the West ended up as general labourers in Ontario, we believe it is more appropriate in our context to 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 follows closely those in Ottaviano and Peri(2010) (hereafter, OP) and D'Amuri et al(2010) (hereafter, DOP). The advantage of this approach is that it allows for a complete accounting of the impacts of immigration by incorporating the possibility that an increase in the supply of immigrants in one occupation can have impacts on 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. However, more flexible alternatives are not feasible in our situation with a limited number of years of data and a large number of skill groups.

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

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

where Y_t is aggregate output and A_t is a productivity shifter.[21] We will assume that the economy faces fixed world prices for the output but that production also involves scarce factors (land in the rural sector, and entrepreneurial ability in the other sectors), implying downward sloping labour demand curves. [22]

Following DOP, we specify the labour aggregate as a CES function of types of skilled (or unskilled) labour. In their model, these skills are defined by education level while in ours they are the 10 broad occupations used in the investigations in the previous sections. Given this

[21] 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.

[22] 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(1976).

specification, the labour aggregator 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 will express the labour in each broad occupation group as a CES aggregate of employment in each of the narrow occupations within the broad group (e.g., for different specific sales occupations within the broad Sales group), that is,

$$L_{kt} = (\sum_{l=1}^{n_k} \theta_{klt} L_{klt}^{\frac{\sigma_o-1}{\sigma_o}})^{\frac{\sigma_o}{\sigma_o-1}} \tag{3}$$

where, n_k is the number of narrow occupations in broad group k, θ_{klt} are narrow occupation specific productivity shocks, and σ_o is the elasticity of substitution among narrow occupations within the same group. Working with the occupational employment in two levels in this way allows for the possibility that, for example, two types of sales occupations are more substitutable than sales workers and trades workers.

The skill specific labour can itself be written as a CES aggregate of employment for each of the 4 possible age groups described earlier:

$$L_{klt} = (\sum_{j=1}^{4} \theta_{kljt} L_{kljt}^{\frac{\rho-1}{\rho}})^{\frac{\rho}{\rho-1}} \tag{4}$$

where L_{kljt} is employment in narrow occupation l within broad occupation group k, age group j, and θ_{kljt} is the associated productivity shock. Again, ρ is a parameter capturing substitutability; in this case among age groups within the same narrow occupation.

Finally, within each type of labour defined by occupation and age group, we can observe workers broken down into new immigrants (who arrived in Canada in the decade preceding a given Census year, t), old immigrants (who arrived over ten years before t), and the native born. As we saw in section 2, Canada experienced inflows in immigration before 1921 that were both very large relative to its existing labour force and had an occupational distribution that was similar to that of the native born. The latter similarity increased with time such that a decade after arrival, the immigrant occupational distribution was very close to that of the native born in the three Censuses where we observe year of arrival (1911, 1921, and 1931). In reduced form estimates not presented here, we also observe that it is new immigrants who have the main

impact on the wages of the native born. Based on this, we implement a specification in which we assume that old immigrants and the native born are close substitutes while new immigrants are a potentially more dissimilar factor. In terms of the nested CES specification, this implies two further levels of nesting: one at which a combination of old immigrants and the native born are treated as a separate factor relative to new immigrants, and another which is concerned with the aggregation of old immigrants and the native born. Thus, we will first write,

$$L_{kljt} = (\theta_{Ikljt} I_{kljt}^{\frac{\delta-1}{\delta}} + \theta_{ONkljt} ON_{kljt}^{\frac{\delta-1}{\delta}})^{\frac{\delta}{\delta-1}} \tag{5}$$

where, I_{kljt} corresponds to employment of new immigrant workers, ON_{kljt} corresponds to employment of a combination of old immigrants and the native born, and δ is a parameter governing substitutability between new immigrants and old immigrant/native born workers in the same occupation-age group.

At the final level of nesting we have,

$$ON_{kljt} = (\theta_{Okljt} O_{kljt}^{\frac{\phi-1}{\phi}} + \theta_{Nkljt} N_{kljt}^{\frac{\phi-1}{\phi}})^{\frac{\phi}{\phi-1}} \tag{6}$$

where, O_{kljt} represents employment of old immigrant workers, N_{kljt} represents employment of the native born, and ϕ is, again, the relevant substitutability parameter.

Using equations (**??**) through (**??**), and again following DOP closely, we can derive a wage equation for a native born worker in a given occupation-age cell, assuming their wage equals their marginal product:

$$\begin{aligned}\ln w_{Nkljt} = \ln(\frac{\partial F}{\partial L_t}(K_t, L_t)) + \frac{1}{\sigma}\ln L_t + \ln\theta_{kt} - (\frac{1}{\sigma} - \frac{1}{\sigma_o})\ln L_{kt} + \ln\theta_{klt} - (\frac{1}{\sigma_o} - \frac{1}{\rho})\ln L_{klt} \\ + \ln\theta_{kljt} - (\frac{1}{\rho} - \frac{1}{\delta})\ln L_{kljt} + \ln\theta_{ONkljt} - (\frac{1}{\delta} - \frac{1}{\phi})\ln ON_{kljt} + \ln\theta_{Nkljt} - \frac{1}{\phi}\ln N_{kljt}\end{aligned} \tag{7}$$

We can similarly derive the log wage equations for an old immigrant worker in the same cell.

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 previously arrived immigrant workers:

$$\ln(\frac{w_{Nkljt}}{w_{Okljt}}) = \ln(\frac{\theta_{Nkljt}}{\theta_{Okljt}}) - \frac{1}{\phi}\ln(\frac{N_{kljt}}{O_{kljt}}) \tag{8}$$

This yields a simple intuitive formulation: holding constant relative differences in productivity, comparing relative differences across cells in wages and employment levels between old immigrant and native born workers tells us about the substitutability of the two types of workers. If we see a large relative jump in old immigrant employment but comparatively small changes in their relative wages, for example, we would conclude that old immigrant and native born workers are very substitutable in production - that the old-immigration shock 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, old 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,

$$\ln(\frac{\theta_{Nkljt}}{\theta_{Okljt}}) = \psi_l + \psi_t + \psi_j + u_{kljt} \tag{9}$$

where the ψ's are fixed effects and u_{kjt} is an independent error term.[23] Identification then relies on the assumption that relative supplies of native born versus previously arrived 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}. [24]

We run equation (**??**) using real annual earnings and employment numbers from the Census data aggregated to the narrow occupation group level. We drop cells with fewer than 10 observations and weight using the square root of the cell size. The dropping of small or empty cells leaves 1534 occupation by age group cells. In all of our estimations, 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 new 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. The results correspond to

[23]Because the narrow occupations are nested in the broad occupations, the inclusion of the complete set of narrow occupation fixed effects would make the inclusion of broad group fixed effects redundant.

[24]Ottaviano and Peri(2008) provide an extended discussion of the form for $\frac{\theta_{Nkjt}}{\theta_{Okjt}}$, arguing for modelling it as time invariant since general technological shifts over time are captured in higher levels in the nesting. However, they also note that it is common to allow for the possibility that there is time variation and we follow that tradition by including time effects. It is worth noting that even they must really be assuming some time variation in this ratio because without it there would be no error term in the regression.

dependent variables constructed using annual earnings. The parallel set of results using weekly wages are very similar and we do not present them for brevity.

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		71	Old Immigs	1394	
	(.018)			&Natv Born		
$-\frac{1}{\delta}$	.008	.11	-	New Immigs	1258	12.3
	(.018)	(.091)		& Old Immigs/Natv Born		
$-\frac{1}{\rho}$	-.073**	-.045*	22	Age Groups	1127	552
	(.012)	(.019)				
$-\frac{1}{\sigma_0}$	-.10**	-.10*	7	Narrow	423	35.6
	(.020)	(.049)		Occ. Groups		
$-\frac{1}{\sigma}$	-.28	-.32	3	Broad	30	6.7
	(.14)	(.16)		Occ. Groups		

Notes: The calculated elasticity is based in the IV estimate in all rows except for the elasticity between old 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.

Our estimate of the negative inverse elasticity of substitution is -.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 US (their estimates lie between -0.01 and -0.06). [25] The corresponding elasticity from our estimates is 72, implying that previously arrived 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 and immigrants who arrived at least 10 years before a given Census) and newly arrived immigrants within the same age-occupation cell. To do this, we aggregate (**??**) to the previous/new immigrant worker level. Thus, for the combination

[25] Both DOP and OP use education groups rather than occupation groups, with 3 education groups and 8 experience groups. Thus, the total number of skill cells is similar to ours when we use broad occupation groups.

of previously arrived immigrants and the native born the relevant equation is:

$$\ln \bar{w}_{ONkljt} = \ln(\frac{\partial F}{\partial L_t}(K_t, L_t)) + \frac{1}{\sigma}\ln L_t + \ln\theta_{kt} - (\frac{1}{\sigma} - \frac{1}{\sigma_o})\ln L_{kt} + \ln\theta_{klt} - (\frac{1}{\sigma_o} - \frac{1}{\rho})\ln L_{klt} + \ln\theta_{kljt} - (\frac{1}{\rho} - \frac{1}{\delta})\ln L_{kljt} + \ln\theta_{ONkljt} - \frac{1}{\delta}\ln ON_{kljt} \quad (10)$$

where $\bar{w}_{ONkljt}$ is the weighted average of old immigrant and native born wages in the occupation-age cell, with the weights being the proportions of old immigrant and native born workers. A similar equation can be derived for new immigrant workers. Taking the difference between these two equations then yields,

$$\ln(\frac{w_{Ikljt}}{\bar{w}_{ONkljt}}) = \ln(\frac{\theta_{Ikljt}}{\theta_{ONkljt}}) - \frac{1}{\delta}\ln(\frac{I_{kljt}}{ON_{kljt}}) \quad (11)$$

As with (**??**), we will assume that the ratio of productivity terms can be captured by occupation, age and time fixed effects plus an idiosyncratic error term.

While it is common to approach this type of estimation treating immigration as an exogenous supply shock within a given cell, it is also common to be concerned about potential remaining endogeneity. In particular, in our context, 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 reasonably be expected to be partly due to endogenous responses to within-cell productivity shocks. One approach taken to this problem by a number of authors, starting with Card(2001), is to instrument for the change in immigrant supplies using a variable based on the notion that immigrants will move to locations within the host country where there are enclaves of people from their country of origin, regardless of the local economic circumstances. We use a variant on this instrument in which we assume that immigrants move to a geographical location based partly on these enclave considerations 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, say.

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:

$$n10_{klgt+1} = \sum_r \pi_{klrt} * p_{grt} * n10_{gt+1} \tag{12}$$

where $n10_{gt+1}$ is the number of immigrants from country g who arrived in the previous 10 years at the national level. Note that this exercise is done separately by age group with both the π_{klrt} weights and the national level inflows (n_{gt+1}) being age group specific.[26] Finally, we form our instrument for $\ln(\frac{I_{kljt}}{ON_{kljt}})$ as $\ln(\frac{n10_{gt}}{ON_{kljt}})$. 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. So, 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 then the correlation between past concentrations and recent inflows of Lithuanians to the region would partly reflect productivity trends rather than just living situation preferences and the instrument would be problematic. Note that the issue has to do with trends in productivity not levels (i.e., not just the fact that miners do well in Canada) because we include occupation fixed effects. 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 of $-\frac{1}{\delta}$ are positive and insignificant. We interpret this as implying that new immigrants and previously arrived workers are perfect substitutes.

Next, we aggregate the wage equations to the occupation-age level. As with the previous exercise, we can write an expression (which we omit for brevity) for $\bar{w}_{kljt}$, the weighted average of immigrant and native born wages in the sector-occupation-age cell, with the weights being

[26]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.

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:

$$\ln(\frac{\bar{w}_{kljt}}{\bar{w}_{kl1t}}) = \ln(\frac{\theta_{kljt}}{\theta_{kl1t}}) - \frac{1}{\rho}\ln(\frac{L_{kljt}}{L_{kl1t}}) \tag{13}$$

Since we want to isolate variation within occupation-time cells in order 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 39 and we drop that group from the estimation. Our IV estimate of $-\frac{1}{\rho}$ is -.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 x occupation cells.

Next, we aggregate up to the narrow occupation group level to arrive at,

$$\ln \bar{w_{klt}} = \ln(\frac{\partial F}{\partial L_t}(K_t, L_t)) + \frac{1}{\sigma}\ln L_t + \ln\theta_{kt} - (\frac{1}{\sigma} - \frac{1}{\sigma_o})\ln L_{kt} + \ln\theta_{klt} - \frac{1}{\sigma_o}\ln L_{klt} \tag{14}$$

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 (**??**) can be captured by occupation and time effects then we can estimate (**??**) 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% 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 (**??**) using the same methods. Again, the instrument is based on an aggregated version of $n10_{kljt}$. In this case, we do have a weak instrument problem, with a first stage F-statistic of 6.7. The IV estimate for $-\frac{1}{\sigma}$ is -.32, which is not statistically significant at the 5% level. We also implemented a specification in which we instrumented for the change in the log total employment level using the change in immigration based on Borjas(2003)'s interpretation

of immigration as a supply shock. That results in a first stage F stat of 89 and a (statistically significant) estimate for $-\frac{1}{\sigma}$ of -.44. We do not report this estimate in the table because we do not believe in that instrumenting strategy. However, from the combination of those estimates and the OLS and IV estimates reported in the table, 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, -.32, which corresponds to an elasticity of substitution for broad occupation groups of 3.

4.1.2 Counterfactual Exercises

The elasticities estimated to this point are of limited intuitive value on their own but can by used with the equation for native born wages, (**??**), to investigate the impact of immigration on native born wages. As OP point out, inspection of (**??**) reveals several channels through which an immigrant shock in a specific occupation by age cell can affect the wage of any native born worker defined by his broad occupation group, his narrow occupation within that group, and his age. First, the fact 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 (**??**), 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 won't change with immigration.(Ottaviano and Peri(2010)).

Our approach to quantifying the impact of immigration on native born wages is to construct a set of counterfactual wage distributions. In particular, we focus on the 1921-31 decade and construct counterfactuals corresponding to: 1) no immigration between 1921 and 1931 (in order to match the most common counterfactual considered in the previous literature); 2) no general or agricultural labourer immigration in that decade (in order to examine the inequality implications of the dominant, low skilled component of the inflow); and 3) no immigration of professionals

in that decade (as a point of comparison for 2)). To construct the first counterfactual, consider generating a fitted wage for 1931 using equation (**??**) but with the number of recent arrival immigrants subtracted from each relevant labour measure on the right hand side. That is, form a new version of L_{kt} as L_{kt}-I_{kt} (where I_{kt} is the number of recent immigrants in broad occupation, k), a new version of L_{klt} as L_{klt}-I_{klt}, and so on. 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,

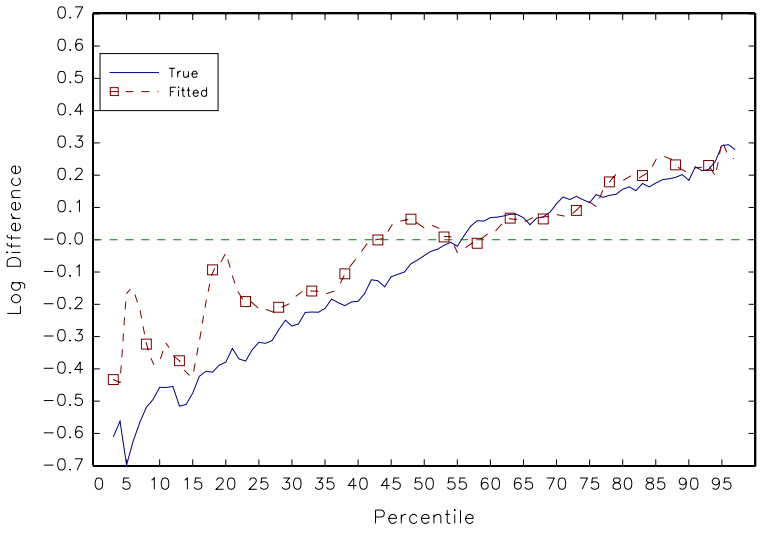
$$\ln w^c_{Nkljt} = \ln w_{Nkljt} - \frac{1}{\sigma} \ln I_t + (\frac{1}{\sigma} - \frac{1}{\sigma_o}) \ln I_{kt} + (\frac{1}{\sigma_o} - \frac{1}{\rho}) \ln I_{klt} + (\frac{1}{\rho} - \frac{1}{\delta}) \ln I_{kljt} \qquad (15)$$

We then construct a complete counterfactual distribution by creating a set of "phantom" people for each occupation x 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 a version of the true distributions that eliminates within occupation by age cell wage variation, just as the counterfactual exercise does. For 1931, we do this by again constructing a set of phantom workers for each cell but in this case we assign them the actual average wage for the relevant cell in 1931. We do this for both 1921 and 1931 then calculate the percentiles of these generated datasets and take the log difference. In Figure 7, we plot the resulting difference line (which we label the "fitted" difference) along with the differences in log percentile between these years from the actual dataset (replotted from Figure 1 and labelled here as "True"). The result shows that removing the within-cell variation does not substantially alter the picture of increased earnings inequality in the 1920s. In terms of the decomposition discussion in the previous section, the fitted line shows the effects of changes in the wage structure that arise only because of changes in between group wage differentials. The fact that it lies above the actual line below the median implies that both between group differential changes and changes in dispersion within groups explain the increase in inequality in the lower part of the distribution. On the other hand, above the median, the fitted and actual lines are coincident, implying that the earnings changes in that

part of the distribution arise only because of within group dispersion.

Figure 7: Actual and Fitted Differences in Log Percentiles, Male Earnings, 1921−31



In Figure 8a, we re-plot the fitted line from the previous figure 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 9, 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 new 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 new arrivals versus settled workers are not identical and the small impact partly reflects general equilibrium adjustments. We return to a measure of the relative importance of those various channels for one group of workers below. In Figure 8b, we plot differences corresponding to the other two counterfactuals: eliminating only recent immigrant labourers and only recent immigrant professionals. 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 be seen clearly in the plots relative to the true 1931 distribution in Figure 9. There, one can also see the spillover impact of

Figure 8a: Fitted and Counterfactual Difference with no Immigration, Male Earnings, 1921–31

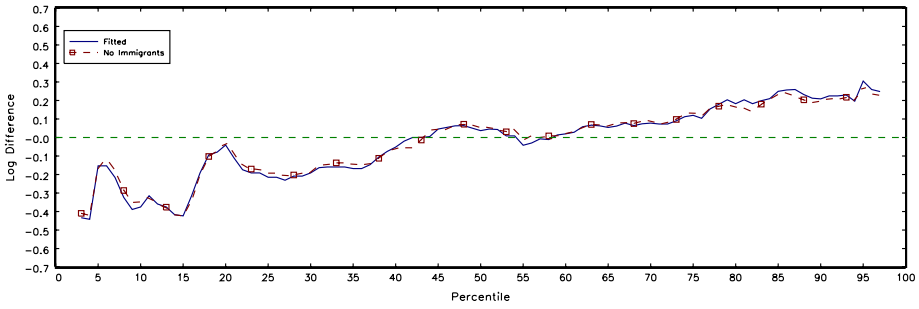
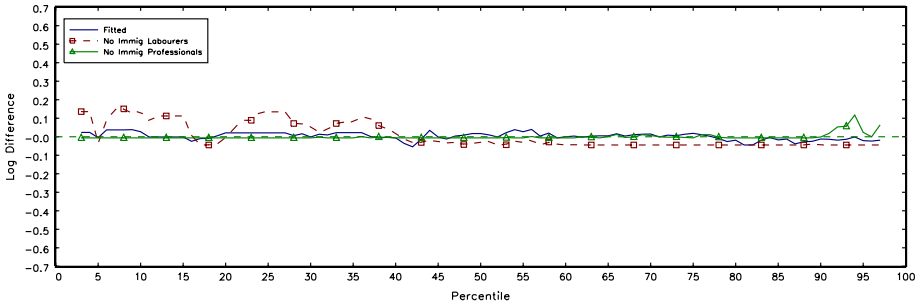


Figure 8b: Fitted and Counterfactual Differences, Male Earnings, 1921–31

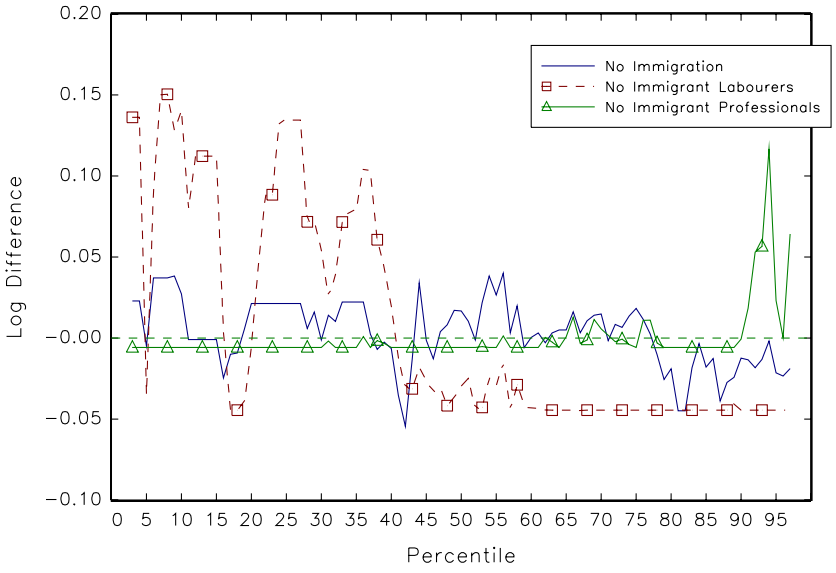


removing immigrant labourers on wages in other occupations. In particular, real earnings in occupations with above median wages would have fallen by approximately 5%. This highlights the arguments in OP that one needs to look beyond narrow occupation groups to measure the full impact of immigration. 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%. The other counterfactual shows the impact of not admitting recent professional occupation immigrants. Removing that inflow would have had substantial positive effects on wages above the 90th percentile but only small negative effects elsewhere in the distribution.

We can use different predicted impacts on farm labourers under various scenarios as a means of highlighting the extent of the adjustment channels we have discussed. 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. [27] This specification can be given an interpretation as showing estimates of the

[27] That is, we estimate the regression, $\Delta \ln(w_{Njkt}) = \alpha_{0t} + \alpha_1 \frac{\Delta E_{Ijkt}}{E_{Njkt-1}+E_{Ijkt-1}} + \Delta u_{jkt}$, where $\Delta \ln(w_{Njkt})$ is the change in native born log earnings in occupation j and age group k, ΔE_{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 $E_{Njkt-1} + E_{Ijkt-1}$ 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 9: Differences in Log Percentiles Between Counterfactual and Fitted Distributions, Male Earnings, 1931



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 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% increase in the farm workforce and thus a 42% decline in worker earnings in that sector.[28] 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 an 39% increase in the farm workforce, and thus would imply an 13% 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 immigration inflows, our first counterfactual implies that the total immigration inflow implied a 4% drop in farm worker earnings. This last reduction (from the 13% effect to the 4% effect) represents the impact from taking account of general equilibrium adjustments in the economy.

[28]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.

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% 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)). Methodologically, Dustmann and Preston(2011) argue that the type of method we use understates immigration impacts to the extent that immigrants have a certain set of skills on paper but actually end up competing for jobs with less skilled native born workers. In our case, though, where we match immigrants to the native born in terms of the occupations they actually work in and where immigrant and native born occupational distributions are very similar, this problem is unlikely to arise.

Impacts on inequality, however, may differ from impacts on the average wage level. Abad(2013) finds substantial impacts of immigration on inequality in Latin America in the nineteenth century, which does not fit with our counterfactual result that total immigration had small impacts on inequality. However, Abad(2013) measures inequality using the ratio of land rents to the unskilled wage, and we do find in our second counterfactual that the substantial unskilled labour inflow in this period reduced earnings in the lower part of the distribution. We conclude that large inflows in specific occupations can have substantial impacts on the wages in that occupation and in closely substitutable occupations but that can be mitigated by positive spill-over effects due 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

We return, at the end, to the question we posed at the outset: what was the impact of the very large inflows of immigrants to Canada in the 1910s and 1920s on the Canadian earnings structure? 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. This arose for a number of reasons. The first main reason is because of immigrant and potential immigrant adaptation. As Abramitsky et al(2014) emphasize, in periods with laissez-faire immigration policy, one might expect that it is mainly potential immigrants with skills that are likely to match demands in the host country labour market 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 new 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 adjustments in terms of their occupational distribution 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 Canadian policy resulted in an inflow that was predominantly composed of 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 have shown 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 affect was to engineer 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.

Because of these adjustments, we show that a standard decomposition of the changes in the earnings distribution reveals limited impacts from the immigrant inflow on the shape of

the distribution. However, such decomposition exercises are based on an assumption of there being no general equilibrium effects, and immigration might well be expected to affect the equilibrium prices in the economy even when the decomposition exercise doesn't reveal direct effects. To investigate this possibility, we implement a quasi-structural estimation that delivers estimates of parameters of an aggregate production function in which immigrants and native born workers are allowed to enter as separate factors. We address potential endogeneity concerns related to immigrant adjustments after arrival by using a variant on the well-known ethnic enclave instrument from the modern immigration literature. Using the resulting estimates in counterfactual exercises, we show that the isolated impact of the general and farm labourer immigration in the 1920s was to lower annual earnings for low earnings workers by over 10% while raising the earnings for higher earnings workers by 5%. But, as we have discussed, the immigrants who arrived in the 1920s did not work only in labouring occupations. Once we take account of their occupational diversity, the spillover effects from immigration into other occupations reduces both the negative effect of immigration on lower skilled native born workers and the positive effects on higher skilled workers. Thus, the third way in which the large immigration inflow translated into small net impacts on the shape of the earnings distribution was through general equilibrium responses within the economy.

The first half of the Twentieth Century were formative years for the Canadian economy. Open land in the West was filled and the mixed resource-manufacturing structure of the Canadian economy was established. Understanding how the economy developed in that period and the role played by key driving forces provides important insights into what the Canadian economy has become. It is hard to deny that large scale immigration was important for the development of a burgeoning economy with untapped resources. But we find that whatever impact immigration had, it did not occur through changing the earnings structure. The driving force for the substantial increase in inequality in Canada in this era must be found elsewhere.

6 References

Abad, L.A.(2013). "Persistent Inequality? Trade, Factor Endowments, and Inequality in Latin America," Journal of Economic History, Vol. 73, no.1, pp. 38-78.

Abramitsky, R., L. Platt Boustan, and K. Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," Journal of Political Economy, Vol. 122, No. 3, pp. 467-717.

Anderson, E. (2001). "Globalisation and Wage Inequalities," European Review of Economic History, Vol. 5, pp. 91-118.

Aydemir, A. and C. Robinson (2008). "Global Labour Markets, Return and Onward Migration," Canadian Journal of Economics, Vol. 41, no. 4, pp. 1285-1311.

Bertram, G.W. and M.B. Percy(1979). "Real Wage Trends in Canada 1900-26: Some Pro-

visional Estimates," Canadian Journal of Economics, 12(2), May, pp. 299-312.

Borjas, G.J. (2003). "The Labor Demand Curve Really is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," Quarterly Journal of Economics, 1335-74.

Card, D. (1990). "The Impact of the Mariel Boatlift on the Miami Labor Market," Industrial and Labor Relations Re Vol. 43, No. 2, pp. 245-57.

Card, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," Journal of Labor Economics, Vol. 19, No. 1, pp. 22-64.

Chambers, E.J. and D.S. Gordon(1966). "Primary Products and Economic Growth, An Empirical Measurement," Journal of Political Economy, Vol. 74, no. 4, pp. 315-32.

D'Amuri, F., G.I.P. Ottaviano, and G. Peri (2010). "The Labor Market Impact of Immigration in Western Germany in the 1990s,"' European Economic Review, No. 54, ppl. 550-70.

Dustmann, C. and I. Preston(2011). "Estimating the Effect of Immigration on Wages," CREAM, WP No. 21/11.

Emery, J.C.H and C. Levitt (2002). "Cost of Living, Real Wages, and Real Incomes in Thirteen Canadian Cities, 1900-1950." Canadian Journal of Economics, 35(1), February, 115-137.

Firpo, S., N.M. Fortin, and T. Lemieux (2009). "Unconditional Quantile Regressions." Econometrica, Vol. 77. May, pp. 953-73.

Goldin, C. And L.F. Katz(2001). "Decreasing (and Then Increasing) Inequality in America," in F. Welch, ed., The Causes and Consequences of Increasing Inequality. Chicago: The University of Chicago Press, pp. 37-82.

Goldin, C. and R. Margo (1992). "The Great Compression: The Wage Structure in the United States at Mid-Century," Quarterly Journal of Economics, 107, pp. 1-34.

Greasley, D. and L. Oxley(2004). âĂIJGlobalization and Real Wages in New Zealand 1873-1913,âĂİ Explorations in Economic History, Vol. 41, pp. 26-47.

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

Green, A.G.. (1994). "International Migration and the Evolution of Prairie Labour Markets in Canada, 1900-1930," in T.J. Hatton and J.G. Williamson (eds.), Migration and the International Labor Market, 1850-1939, London: Routledge, pp. 156-74.

Green, A.G. (1995). "A Comparison of Canadian and US Immigration Policy in the Twentieth Century," in Diminishing Returns: The Economics of Canada's Immigration Policy, D.J.DeVoretz, ed.. C.D. Howe Institute, 1995, pp.31-66.

Green, A.G. and D.A. Green(1993). ""Balanced Growth and the Geographic Distribution of European Immigrant Arrivals to Canada,1900-14," Explorations in Economic History, No. 30, pp.31-59.

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

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

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

Hatton, T.J. and J.G. Williamson (1998). The Age of Mass Migration, New York: Oxford University Press.

Kelley, N. and M. Trebilcock(2010). *The Making of the Mosaic: A History of Canadian Immigration Policy*, Second Edition, Toronto: University of Toronto Press.

Keyfitz, N. (1950). "The Growth of Canadian Population," Population Studies, vol. 4, no. 1, pp. 47-63.

Lew, Byron(2000). "European Immigration to Canada During the 1920s: The Impact of US Quotas and Canadian Restrictions," Queen's University.

Lewis, F.(1975). "The Canadian Wheat Boom and Per Capita Income: New Estimates," Journal of Political Economy, vol. 83, pp. 1249-57.

Mackinnon, Mary(1996). "New Evidence on Canadian Wage Rates, 1900-1930." Canadian Journal of Economics, 29(1), February, pp.114-131.

McDougall, D.M.(1961). "Immigration into Canada, 1851-1920," Canadian Journal of Economics and Political Science, vol. 27, no. 2, pp. 162-175.

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

Pope, D.H., and G. Withers (1994). "Wage effects of immigration in late-nineteenth century Australia," in: Williamson, J.G., Hatton, T.J. (Eds.), Migration and the International Labor Market, 1850âĂŞ1939.

Taylor, A.M. and J.G. Williamson âĂIJConvergence in the Age of Mass Migration,âĂİ European Review of Economic History, Vol. 1, pp. 27-63.

Urquhart, M.C. and K.A.H. Buckley(1965). Historical Statistics of Canada. Toronto: Macmillan.

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

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

Wright, G. (1990). "The Origins of American Indusrial Success, 1879-1940," American Economic Review, vol. 80, No. 4, pp. 651-688.