Canadian Migration to the US in the late Nineteenth Century: Positive Selection or Positive Opportunity?

Byron Lew* and Bruce Cater

Department of Economics

Trent University

1600 West Bank Drive

Peterborough, ON K9J 7B8

Canada

October 16, 2017

*corresponding author. Email: blew@trentu.ca.

Introduction

North Americans are a highly mobile people. Of those born in the U.S. in the middle of the nineteenth century, many would move outside their state of birth. The Canadian-born were also highly mobile in pursuit of opportunity, particularly if the U.S. is included as a possible destination. Unlike the American-born, the Canadian-born moved across the international border in large numbers through the late nineteenth century right up until the Great Depression. By the early twentieth century a significant share of the Canadian-born—both English- and French-speaking—were living in the U.S.[1]

Evidence of the cross-border mobility comes from the U.S. Census and from records of border crossings (McInnis, 1994; Ramirez, 2001). The U.S. census provides information on the status of movers but nothing about their background. The border crossing data yield some information on background, but nothing about outcomes of migrants in the U.S. To identify both outcomes and selection characteristics, data must be linked from different census years. The 1881 Census of Canada is digitized and contains records for all individuals. We find Canadians living in the U.S. in the IPUMS 1900 US Census 5% sample and match them backwards to the 1881 Census of Canada. We are able to match 3,612 males out of a potential 18,534 listed in the 5% sample.

The two dimensions of mobility we will explore are geographic and occupational. Along the first dimension, geography, we examine the location of origin of Canadian migrants, and link sending to receiving locations. This will help shed light on regional and local disparities of opportunity. Canada at the time was a highly agricultural economy, and by the late nineteenth century unclaimed lands were becoming very scarce. In addition,

---

[1]At its peak in 1900, the Canadian-born living in the U.S. amounted to 22% of those born in Canada.

Canada was highly industrialized, but it was a small country lacking large cities, so labour markets were relatively thin compared to those of the U.S. Canadians looking to exploit particular specialized skills were more likely to find employment matches in larger labour markets. We examine the evidence to gauge the extent to which migrant movement was predominately from rural to large urban centers.

We can also examine migration networks knowing both from where migrants were coming and to where they moved. The choice of a migrant's destination is determined by the opportunities in the destination and the skills of the migrant. Opportunities to acquire skills are partly determined by family background and location. These data allow us to see how background and location affect destination choice.

The second question we explore is identifying the socioeconomic status of immigrants. In particular, we identify the occupation of fathers of immigrants in Canada. Occupation is our measure of socioeconomic status. We can compare fathers of immigrants with the general population to determine if immigrants came from a particular subset of family by occupational status. We are also able to capture intergenerational mobility by comparing occupational status of father and son. We hope to supplement this with information from the North Atlantic Population Project once data on intercensal mobility in Canada has been released.

We are able to sub-categorize by language of origin: French vs. non-French based on identification from the 1881 Census of Canada. Identification by mother tongue is introduced in the 1910 U.S. Census. We will (in the future) compare our results for differences between French and non-French to those reported in the 1910 Census.

Matching Strategy

Census Matching Keys

The full 1881 Census of Canada is available electronically serving as an anchor (Ruggles et al., 2017). We match the Canadian-born in the U.S. 1900 Census 5% sample to the 1881 Census of Canada.[2] The large majority of those reporting an arrival in the U.S. after 1881 should be listed in the 1881 Census.[3]

Our matching strategy follows closely that of Antonie et al. (2014) and Baskerville et al. (2014). We match individuals by surname and firstname. We match only males, as surnames of women marrying between census years will usually change.[4] In addition, we use the following characteristics to identify matches: age, marital status, marriage duration, birthplace, and parent birthplace.

The age of someone reporting age of n_{1900} will be $n_{1900} - 19$ in 1881.[5] We use a range of ± 3 years, $[n_{1900} - 22, n_{1900} - 16]$, to account for the lack of precision in records of age of the period. A window for age uncertainty is often $\pm 20\%$. Since the censuses are about 19 years apart, a 20% window would be ± 3.8 years, so we truncate to 3 rather than round to 4. We miss some matches, but we also limit duplicate matches.

---

[2]The microdata file downloaded January 2017 from `http://www.prdh.umontreal.ca/census/en/main.aspx`.

[3]Not everyone will necessarily be included. Some may be overlooked due to enumeration error. Others could be away at the time of the Census, perhaps working in the U.S. prior to the decision to move permanently.

[4]Married women from Québec who keep their surname from birth, change to their husband's surname in the U.S. Census.

[5]The Census of Canada is as of April 1, 1881 and the U.S. Census is as of June 1, 1900 for a span of 19 years and 2 months. Rounded to years only, for those born in April or May this is a 20-year difference; for all others it is a 19-year difference.

Matches are also identified that are consistent with permissible changes in marital status. A change from married, divorce or widowed in 1881 to single in 1900 is non-permissible. The U.S. Census reports marriage duration. When marriage duration is 19 years or more, and if a spouse is listed on both censuses, then the spouse's name and age must match across censuses.

Birthplace is used to generate our sample within the U.S. Census sample. The detailed birthplace measure used in the U.S. Census captures state of birth of U.S.-born, but for Canada, province of birth is not generally identified. Birthplace does contain additional information on the Canadian-born that we utilize. For a large majority of records, birthplace distinguishes between "English Canada" and "French Canada." This seems to suggest geography, with "French Canada" presumed to be Québec, and "English Canada" indicating birthplace anywhere other than Québec. But the use of these two categories could also refer to ethnicity.

The instruction to enumerators for those reporting Canada as their birthplace is ambiguous. The enumerator is to: "ask whether the person is of English or French descent."(United States. Department of the Interior. Census Office., 1900, number 143, p. 30) The use of the word "descent" seems to suggest it was intended as a question of ethnicity linked to language spoken. This would be odd in the context of a question on birthplace that would otherwise refer to geography. Some enumerators clearly interpreted it as a question of geography since a small number of records list the specific province of birth. For those that report "English" or "French" Canada, it is therefore not clear whether "French" indicates a birthplace of Québec or French-speaking heritage/ethnicity. It is possible that an individual from a predominately French-speaking community outside Québec but of French ethnicity could be reported as coming from

5

"French" Canada.

Our default interpretation is that birthplace generally refers to geography. If we get a match where geography does not match, we then check ethnicity using "ethnic origin" from the Census of Canada. The field "ethnic origin" contains a detailed set of categories. For this purpose, we treat ethnic origin as a binary variable: French vs. every other ethnicity.[6] Where we match someone whose birthplace is identified as "French" Canada but is not born in Québec, we check for French ethnicity to confirm. Or, equivalently, if we match someone from "English" Canada born in Québec, we also check for non-French ethnicity to confirm.

We assume that when birthplace in the U.S. Census is reported as "French" Canada, it means Québec, while "English" Canada means any other province. For a potential match where the U.S. Census reports a birthplace of "French" Canada but the birthplace in the Census of Canada is not in Québec, if ethnic origin is French, we consider it a potential match. And conversely for a potential match where the U.S. Census reports a birthplace of "English" Canada but the birthplace in the Census of Canada is in Québec, if ethnic origin is not French, we consider it a potential match.

There are unambiguous matches—cases where we find several members of the same family in both censuses—where individuals identified in the Census of Canada of non-French ethnic origin born in Québec are identified with birthplace of French Canada in the U.S. Census. There are also examples of individuals identified in the Census of Canada of French ethnic origin born outside of Québec, but identified with birthplace of French Canada in the U.S. Census.

---

[6]We use codes French, French Canadian, or Acadian as French.

The U.S. Census also reports parents' birthplaces. We use this variable to confirm matches and discriminate among multiple matches. Because there is no variable identifying family relationships in the 1881 Census of Canada, we must make assumptions about family structure in order to identify parents.[7] It is very clear that family relationships are implied by the order in which individuals are listed within households.

The vast majority of Canadians are in households where a large majority share a single surname, so identifying family relationships from ordering by household is generally straightforward. There are 802,973 unique households identified in the 1881 Census of Canada. Of these, 613,909 households have one unique surname, and 146,572 have two unique surnames.[8] These account for almost 95% of households in the census.

Families are typically ordered as follows: head of household, spouse if alive, children, then others. Marital status of parent must be married or widowed. We assume a parent-child relationship when the surnames match and the ages differ by eighteen years or more, and the presumed mother and youngest child's age differ by no more than 47 years.[9] If the mother and youngest child's ages differ by more than 47 years, we then assume that child must be a grandchild.

There are several confounding possibilities. There may be multiple generations in one

---

[7]The benefit for our purpose of generating a family-relationship variable would be tiny relative to the cost. To date, we do not know of any researchers who have created a family relationship variable for the 1881 Census of Canada.

[8]Many of the households with two surnames have the same surname spelled slightly differently across family members. Many of the others reflect the presence of boarders, domestic servants, sometimes their children, and some married women in Québec.

[9]After age 45, the probability of conception falls to zero very quickly. Among Hutterite couples studied by Tietze, where the female partner was 45, 13% of couples were still fertile; but an extrapolated linear trend falls below zero by age 47.

household, in which case the order of individuals varies. There are two standard variants. Children of married children might be listed after all siblings, effectively as grandchildren of the head of household. Or grandchildren may be treated as children of children, in which case a married child is listed consecutively with their spouse and children, followed by the rest of the household head's children. Married women in Québec may be listed immediately after their husband, or if they report their birth surname they may be listed after all individuals with the same surname as the household head. In these cases, the marital status indicator generally makes the relationship clear. Widows and widowers living with their children are usually listed after the household head's family unless they themselves are identified as the head of the household. There may be multiple, unrelated families in one household. In most of those cases, family structure is identified by order as well. If the relationships are not clear, we do not use the information on parent birthplace.

The major problem we encounter is determining if the married female listed right after a married male is the mother of the children listed next. The Census does not report marriage duration. If the children have a different surname, then we consider these as children of a previous husband and do not identify the male household head as a parent.[10] But if there is only one family surname, then there is no direct way of knowing if the female spouse is the mother. The father could have remarried after the death of the previous wife. The only basis for us to make this distinction is the age differences between the female spouse and the children. If the age difference is less than 18 years, we do not consider the wife as a mother. Otherwise, we assume that the wife is the mother.[11]

---

[10]If the family is in Québec and the wife reports her birth surname, then we assume she is the mother.
[11]If age difference is greater than 47 years, we assume she is the grandmother.

While age reporting is subject to error, birthplace of a parent appears to be subject to greater error. There are examples of members of the same family reporting different birthplaces for mothers and for fathers. For cases where there are two parents in the household, we consider a match if at least one of two parents' birthplaces match. If there is only one parent, then parent birthplace must match to be considered a match. We only use parent birthplace as a discriminating criterion when comparing matches to individuals living in their parents' household in the 1881 Census. If any of the matches are to individuals not living in their parents' household, then parent birthplace is irrelevant to the match.

When we otherwise have a unique match, but parent birthplace does not match, we make further accommodation for error in the reporting of parent birthplace. We observe among our confirmed matches that a parent's birthplace can be confused with family origin. For example, we observe cases where a child reports their parent's birthplace as Ireland but the parent matched in the 1881 Census reports birthplace as Canada and ethnic origin as Irish. To account for this type of error, when a foreign birthplace is reported for a parent who matches someone born in Canada, we look at the parent's ethnic origin. If that matches, we do not reject a match already made.

We find quite a few families moving together. We have included all males in a family if they are found in both censuses, including a small number of family-member matches not born in Canada.

Algorithm

Our matching algorithm proceeds as follows. We first clean the names to convert to lowercase, eliminate diminutives and adopt a common name for those with variations of spelling. We attempt to anglicize French names where there is a standard link—John for Jean. We drop all records with a first initial only. Where we have an initial and a name—J. Edgar—we drop the initial but keep the name. We then match by blocking. Names generally retain the same first letter when reported on separate occasions and in different documents. This enables us to divide the Census of 1881 into blocks by first-letter of surname. Without doing that, we would quickly create a sample of matches to analyze with huge computer memory requirements.

We first match on age range to speed the process as numeric matching is much faster than string matching. We then match the resultant reduced set on surname and first name. We match both surname and first name on two different string-matching algorithms. We use the Jaro-Winkler edit distance with a threshold of 0.85.[12] We also use the Levenshtein edit distance with a threshold of 0.8 on both the untransformed strings, and on the double-metaphone transformed strings. This yields four scores each for surname and for first name. A match is identified when all eight scores exceed their threshold values.

We then check all these matches manually using the criteria of birthplace, marital status, and parent birthplace. We have checked all unique matches as well as doubles, triples and quadruples. The likelihood of generating a unique match falls with number of duplicates, so for now we have stopped at four.

---

[12]The Jaro-Winkler distance is the minimum number of character transpositions required to convert strings into a match, with 0 indicating a perfect match. Our threshold is $1 -$ Jaro-Winkler distance

Results

Matching success

We start with a sample of 18,534 males 16 years and older, born in Canada and arriving in the U.S. in 1880 or later.[13] We include the year 1880 to allow for heaping on 1880 as a focal point assuming error in reporting year of arrival. We match 9,519 individuals from the 1900 U.S. census sample to the 1881 Census of Canada using names and age only, for just over half of the sample. Of these matches, 2,916 are unique matches to only one person in the 1881 Census of Canada. The rest are multiple matches yielding a large set of possibilities to check. Of these sets of multiple matches, we have 1,287 sets of double matches, 805 triples, 532 quadruples, etc., up to one with 430 matches, for a total of 113,642 potential matches to check.[14]

We have checked 10,033 matches for birthplace, marital status and parent birthplace. We are able to identify 3,138 matches. We also matched an additional 474 individuals—including 67 not born in Canada—from family links who were not otherwise matched, for a total of 3,612 matches. These 3,612 matches include 2,678 individuals where we can link to a father in the Census of Canada.

Two possible sources of bias in our matches results from name matching. Uncommon names are more likely to generate a match and may also be related to a higher-status background (Abramitzky, Boustan, and Eriksson, 2012). As well, name changes across

---

[13]We start with age 16 to remain in our band for error in age reporting.
[14]John McDonald age 36 generated the 430 matches.

censuses for French migrants are more likely if French-speakers choose to anglicize their names once in the U.S.[15] To check for these two effects, unique matches are compared to non-unique matches and to non-matches in both censuses. We look for differences in urbanization, and for differences between French and non-French matches for the 1881 Census of Canada.

Table 1 shows the proportion matched in the 1881 Census of Canada. We have divided our sample into French and non-French on the ethnicity variable. We match 2.25% of the French sub-sample and 6.4% of the non-French. This difference could be due to a lower out-migration rate for French-speakers, but it also could be because of the anglicizing of French names in the U.S. Census. To illustrate this further, we break down all matches into three types: unique and non-unique, as well as matches generated by family links to an individual already matched. Unique matches constitute 5% of matches for French and only 2.4% of matches of non-French. The proportion matched by family relations is very small. The difference is the number of non-unique matches, a larger share for non-French. That does suggest greater name diversity among French migrants. This may well be due to variations in the Anglicized spelling of French names, and is an issue we will have to address (in a future draft).

[Table 1 about here.]

Table 1 also reports the mean of urbanization by match-type. Here non-match refers to the entire male population. There is no "urban" variable in the 1881 census, but the Census does identify 326 cities, though many are small. About 25% of the population lives in an identified city. Our variable *urban* is an indicator identifying those living in any of these cities. For both French and non-French, the proportion of urban dwellers

---

[15]We note, for example, that married females who report their birth-names on the Canadian Census report their husband's surname in the U.S. Census.

among unique matches and non-matches do not differ statistically. For non-French, *urban* is slightly lower for non-unique matches. The difference is quantitatively relatively small but it is statistically significant. We will consider the impact of urban status on our sample.

For the French sample, there is no difference in *urban* for unique and non-unique matches. But *urban* is lower for non-unique matches compared to non-matches. The sample size for unique French matches is too small to determine conclusively that *urban* differs for unique matches compared to non-matches. Nevertheless, a lower degree of urbanization seems to be a characteristic of migrants themselves rather than bias from name-matching.[16]

A comparison of matches by type for the U.S. Census of 1900 is reported in Table 2. We compare matches to non-matches among the same base sample, those 16 years and older, born in Canada and having arrived in the U.S. in 1880 or later. The U.S. Census includes an *urban* variable. It also reports literacy which may also capture differences in naming patterns.[17] The non-unique matches are more urban than the unique matches and non-matches. If unique names were more likely to indicate higher socioeconomic status, then we would have expected to find more unique matches among the urban. We also tested matches on literacy. Literacy rates are slightly higher for both types of matches compared to non-matches. We suspect this might reflect lower matching rates among French speakers.

[Table 2 about here.]

Occupational status can also be compared. We divide occupations into four aggregate

---

[16]A better definition of *urban* based on population could help sort his out.

[17]We do not divide by French status because we cannot identify status of non-matches.

groupings: white-collar, farmer, skilled, and non-skilled. We match a greater proportion of white-collar workers and farmers, and a lower proportion of skilled workers than the non-matches. Generally unique and non-unique matches are similar, except for farmers for whom the proportion of unique matches is greater than non-unique and non-matches. We match fewer skilled workers as well, though the proportion of non-unique matches lies halfway between unique and non-matches.

Characteristics of Migrants

Geography

Table 3 identifies the origin of migrants by province. The Table shows each province's population and proportion of the country's population, the number of matches per province and their proportion of total matches. For example, Ontario accounts for 45% of the country's population and just under 46% of the matches. Québec, in contrast, accounts for 32% of the population but only 23% of migrants. We cannot be sure if this is undermatching, or evidence of a less mobile population due to language. The maritime provinces all have a greater share of migrants than their share of population. To address the potential for the proportional distribution of migrants by province-of-origin being affected by any bias in our matching against finding migrants from Québec, the last column reports the ratio of migrants to population.[18] This also suggests that the maritime provinces were losing a relatively larger share of their population than was Ontario, and probably Québec as well.

[Table 3 about here.]

We disaggregate the sending regions to determine if certain regions were more likely to see outmigration to the U.S. We disaggregate by district as we don't have enough

---

[18] Assuming that each person represents 20 from the 5% sample.

observations for a full disaggregation by sub-district. Rather than finding a narrow set of locations as the origin for migrants, we find migrants from 189 of 192 districts. We rank these districts in two ways. We first rank them by number of migrants—counting a family as one, shown in Table 4. Montréal City ranks first. We then rank by immigrant share of population in 1881, shown in Table 5. While the district with the largest number of migrants is Montréal, it ranks lower when emigration is expressed as a share of its population. But by either ranking, the Maritime provinces display a strong outmigration to the U.S., particularly Nova Scotia. Districts Annapolis, Yarmouth and Hants rank in the top five by number of outmigrants, and the top three by proportion of outmigrants to population. And Queens in Prince Edward Island, the sixth most populous district, ranks second in numbers after Montréal.

[Table 4 about here.]

[Table 5 about here.]

The U.S. Census reports where Canadians settled, but it does not report on their origin.[19] The connection from province of residence to destination is illustrated in the following three tables, each showing the state of residence for a province or region of residence in the 1881 Census of Canada.[20] The tables report on the proportions of immigrants out of the total of all immigrants in the sample for each province or region. Each table reports separately on French and non-French, and only show destinations accounting for at least 1.5% of the sample of migrants from each province or region. For example, Table 6 shows that 60% of French migrants from the Maritime provinces were living in Massachusetts. For non-French the proportion is 46%.

---

[19]As noted in our method section, it only identifies "English" vs. "French" origin. It does report province of origin, but only for a very few individuals.

[20]The three maritime provinces are amalgamated.

[Table 6 about here.]

[Table 7 about here.]

[Table 8 about here.]

New England was the principal region of settlement for immigrants from the Maritime provinces. Table 6 illustrates that this was true for both French and non-French. New England states were the destination of 87% of French and 67% of non-French Maritime residents. The attraction of Massachusetts was particularly strong for French, accounting for almost 2/3. Distance was quite important. There are relatively few French from the Maritimes found outside New England, with a surprisingly few in New York. But the Midwest did attract some, particularly Wisconsin, Michigan and Minnesota.[21] For non-French migrants from the Maritimes, after New England, New York and the Midwestern states of Michigan, Minnesota and Wisconsin were major destinations. In general, French and non-French Maritime migrants to the U.S. were relatively similar in their choices of destination with French having a greater preference for New England, particularly Massachusetts, and non-French showing a greater preference for New York and Michigan.

Settlement patterns for migrants from Québec are shown in Table 7. The preference for New England is clear for French migrants with 81% of French found in those six states. But distance may have been a factor as New England also accounts for 41% of non-French from Québec. This is a smaller share than from the Maritime provinces, with New York and the Midwestern states accounting for a greater share of non-French migration from Québec. While New England is less of a draw for non-French than for

---

[21]Our sample contains relatively few French Maritimers outside New England and the three other states listed.

French from Québec, Rhode Island and Connecticut have proportionately fewer non-French while Vermont has proportionally more non-French.

Table 8 shows the proportional distribution of migrants from Ontario. This group looks quite different, and of particular note, migration choices of French and non-French are quite similar. The largest proportions of both French and non-French settle in the Midwest, with Michigan having the largest draw, and Wisconsin attracting a relatively large number of French. New England is destination for 31% of French of which half are in Massachusetts and another quarter in Maine. This is a much smaller share than for French from Québec or the Maritimes. Only 10% of English are in New England, the majority of whom are in Massachusetts.

There are some similarities among non-French from Québec and Ontario. For both, Massachusetts, New York and Michigan are large draws. But non-French from Québec and Ontario differ in that New Hampshire and Vermont are attractive to non-French from Québec but draw very few from Ontario. Overall, 30% of non-French from Québec settle in the Midwest and 40% in New England, while for Ontario those shares are 60% and 10% respectively.

Canada was agricultural and also rural with very few large cities. Given the paucity of large urban centres, Canadians may have been migrating to the U.S. to gain advantages that the more urbanized U.S. had to offer. The proportion of urban origins among our matched sample is essentially the same as the overall proportion for Canada. Dividing our sample by language, non-French are about 28% urban, while French are only 20% urban.

Aggregate patterns of changes in urbanization of migrants are shown in Table 9. The Table reports the proportion living in urban areas given their urban/rural status in Canada. It presents three measures of urbanization from the U.S. Census. The definitions of an urban location used are metropolitan area, city and urban area. The city definition is meant to be similar to the measure used to gauge urbanization in Canada. Urban area is the most restrictive as it includes only places with a concentration of at least 50,000 people plus any contiguous urban areas.

[Table 9 about here.]

French migrants were more likely to live in an an urban area regardless of urban or rural origin. While only 42% of those from an urban area ended up in an urban area by the strictest definition, French were more likely than non-French to end up urban across all measures. Non-French chose locations based more on where they originated. Only a minority of those originating in a rural area ended up in an urban area. Approximately 2/3 of those coming from urban areas in Canada tended to move to urban areas in the U.S. Therefore, while French tended to urbanize regardless of origin, non-French chose destinations with the urban characteristics more similar to that of their origin.

Occupations

The second axis of mobility we explore is occupation. We first look at occupations of the fathers of our immigrant sample. Table 10 reports the proportional breakdown for fathers from the 1881 Census of Canada using the four-category occupational grouping. The data are divided into French and non-French. The Table also reports the breakdown for the general population. We note that the differences between the distribution of the general population and the fathers of immigrants over the four occupational categories are not large.

18

[Table 10 about here.]

At the time, Canada was highly agricultural and almost 50% of working males identified as farmers. The proportion is slightly higher for French than non-French. Among the sample of fathers of immigrants, the proportion of farmers is higher for both French and non-French. The differences in the proportions between fathers of immigrants and the general population, however, are the same. Fathers of immigrants were more likely to be farmers and French were more likely to be farmers. There is no interaction effect of note. Fathers of immigrants, both French and non-French, were also over-represented as skilled workers.

Since there was a higher proportion of farmers among fathers of immigrants, then there must have been a lower proportion of other occupational categories. There were fewer unskilled and fewer white-collar among fathers, and again that is true for both French and non-French. The gap between fathers and the general population for white-collar is larger for French than non-French, presumably due to the need for language in white-collar work.

The question of mobility can be addressed by linking occupations of fathers and sons. Table 11 shows the matrix of proportions for matched fathers to son over the four occupational categories, divided into French and non-French. Fathers' occupations are by row and sons' occupations are by column. We now observe some differences between French and non-French. The largest transition for non-French from father to son is farmer-unskilled, followed closely by farmer-skilled. For French, the transition from father-son is largest for farmer-skilled, followed by farmer-unskilled. The difference is large: one-third of all sons of farmers ended up in skilled labour. This is not too sur-

prising perhaps, as work in the textile mills is here classified as skilled. In general, the transition to skilled work for French is greatest, in total accounting for 56% of occupations of sons.

[Table 11 about here.]

The patterns of transition for sons working in unskilled occupations is very similar for French and non-French. The larger differences between French and non-French are the transitions into white-collar occupations or into farming. While a non-trivial share of sons of farmers moved to farming for non-French, for French very few remained in farming. Also white-collar occupations were more common among non-French, though not common for sons of fathers who were unskilled.

Given the very different proportions among each occupation for fathers, Table 12 illustrates the distribution of sons' occupations given their father's occupation. Among non-French, sons of farmers end up in skilled and unskilled occupations with about equal frequency, and they end up in white-collar occupations or farming also with about equal frequency. The latter two categories occupy about 1/3 of farmers' sons. For French, in contrast, skilled and unskilled account for over 90% of occupations of sons of farmers.

[Table 12 about here.]

Sons of unskilled fathers are highly likely to end up in either skilled or unskilled occupations. But among French migrants, sons are much more likely to end up in skilled occupations, with twice the frequency of unskilled. A large share of sons of fathers in white-collar occupations end up in white-collar occupations. While not surprising, the share among French is also fairly large, only 15% lower than non-French.

20

Conclusion

We explored the mobility of Canadians moving to the U.S. in the late nineteenth century along two dimensions: geographic and occupational. In this period Canada was land-constrained and its industrial capacity, though certainly expanding, was limited by the small size of the domestic market. Given the land constraints, we had expected to find a farm-to-city migration concentrated in a few long-settled regions. While we did find this, we also find that Canadian immigrants to the U.S. came from all across the country, that is from Ontario and provinces eastward. Rural regions certainly contributed the majority of emigrants, but the distribution of migrants was not much different than the distribution of urban and rural in Canada at the time.[22]

By region, the Maritimes saw a relatively larger outmigration rate, and Québec's outmigration rate was relatively smaller, though this conclusion may be due to undermatching of French names. It is well established that emigrants from Québec moved to New England almost exclusively. But the same was true of non-French Maritimers as well. Distance and linkages were certainly influential. Non-French speakers from Québec also moved to New England, though they were a little more diversified, moving to New York and the Midwest in greater numbers. Migrants from Ontario moved mostly due South to New York or West into the Midwestern states, with Michigan receiving the most. This was true of both French and non-French, though many of French origin from Ontario did also find their way to New England, albeit at a much smaller rate.

There is occupational evidence to indicate the movement away from the farm was important. Only 8% of non-French migrants consisted of sons following fathers into agri-

---

[22]We intend to look more closely at the distribution of immigrants across sending regions.

culture, and the number among French was less than 2%. Sons of farmers were most likely to end up in unskilled or skilled work, both for French and non-French, though for French there were double the number in skilled compared to non-skilled. There was also evidence of occupational mobility upward. For both French and non-French, more sons of fathers with unskilled occupations ended up in skilled occupations than in unskilled; and again, the ratio was higher among French, at more than double the rate.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review* 102 (5):1832–1856.

Antonie, Luiza, Kris Inwood, Daniel J. Lizotte, and J. Andrew Ross. 2014. "Tracking People Over Time in 19th Century Canada for Longitudinal Analysis." *Machine Learning* 95:129–46.

Baskerville, Peter, Lisa Dillon, Kris Inwood, Evan Roberts, Steven Ruggles, Kevin Schürer, and John Robert Warren. 2014. "Mining Microdata: Economic Opportunity and Spatial Mobility in Britain and the United States, 1850-1881." In *2014 IEEE International Conference on Big Data*.

McInnis, Marvin. 1994. "Immigration and Emigration: Canada in the Late Nineteenth Century." In *Migration and the International Labor Market, 1850–1939*, edited by Timothy J. Hatton and Jeffrey G. Williamson. London: Routledge, 139–55.

Ramirez, Bruno. 2001. *Crossing the 49th Parallel: Migration from Canada to the United States, 1900–1930*. Ithaca: Cornell University Press.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2017. *Integrated Public Use Microdata Series: Version 7.0 [dataset]*. Minneapolis: University of Minnesota.

United States. Department of the Interior. Census Office. 1900. *Twelfth Census of the United States, June 1, 1900. Instructions to Enumerators*. Government Printing Office.

Table 1: Proportion Matched in 1881 Census of Canada, by French Status

Panel A: Proportion Matched				
	French		non-French	
	prop	s.e.	prop	s.e.
All Matches	0.0225	0.000184	0.0643	0.000197

Panel B: Matches, Proportion by Type				
	French		non-French	
	prop	s.e.	prop	s.e.
Unique	0.0539	0.00187	0.0236	0.00048
Family	0.0122	0.00091	0.0029	0.00017

Panel C: Urban by Match Type				
	French		non-French	
	mean	s.e.	mean	s.e.
Unique	0.209	0.01447	0.265	0.00911
Non-unique	0.193	0.00337	0.240	0.00137
Non-match	0.226	0.00052	0.264	0.00037

Note: Panel B reports match from Panel A, broken down with non-unique not reported. French status refers to ethnic origin. See text for definition of French ethnicity and for urban.

Table 2: Comparison of Matches from 1900 Census: Urban, Literacy, and Occupation

	match unique		match non-unique		non-match	
	$\hat{x}$	s.e.	$\hat{x}$	s.e.	$\hat{x}$	s.e.
Panel A: Literacy and Urban Location (mean)						
urban	0.587	0.00881	0.609	0.00622	0.586	0.00522
literacy	0.913	0.00506	0.921	0.00345	0.845	0.00384
Panel B: Occupation (proportion)						
white-collar	0.153	0.00664	0.159	0.00482	0.126	0.00367
farmer	0.101	0.00552	0.073	0.00341	0.076	0.00291
skilled	0.427	0.00912	0.447	0.00656	0.476	0.00552
unskilled	0.319	0.00859	0.321	0.00615	0.322	0.00516

Note: Means of literacy and urban location, and proportion by occupational class over matches—both unique and non-unique—and non-matches. Sample includes only Canadian-born males 16 years and older arriving in the U.S. from 1880 on. See text for definition of occupation.

Table 3: Province of Origin of Matched Migrants

Province			Migrants		
	population	proportion	number	proportion	ratio
British Columbia	48,885	1.1%	5	0.1%	0.2%
Northwest Territories	10,978	0.3%	2	0.1%	0.4%
Manitoba	64,842	1.5%	24	0.7%	0.7%
New Brunswick	321,247	7.5%	332	9.2%	2.0%
Nova Scotia	440,553	10.3%	447	12.4%	2.0%
Prince Edward Island	108,910	2.5%	122	3.4%	2.2%
Ontario	1,924,142	45.0%	1662	45.8%	1.7%
Québec	1,358,222	31.8%	1019	23.4%	1.5%
Canada	4,277,779	1.0	3613	1.0	1.7%

Note: Column *ratio* is the ratio of the number of migrants×20 divided by the provincial population.

Table 4: Sending Districts by Number of Outmigrants, top 35

District Number	Name	Province	count
90	Montréal (Ville/City)	QC	52
2	Queens	PI	47
134	Toronto (City/Ville)	ON	41
16	Annapolis	NS	39
18	Hants	NS	38
14	Yarmouth	NS	38
180	Kent	ON	38
163	Elgin (East/Est)	ON	37
181	Essex	ON	36
33	Westmorland	NB	36
52	Drummond & Arthabaska	QC	35
10	Halifax	NS	34
176	Bruce (South/Sud)	ON	34
1	Prince	PI	32
179	Lambton	ON	32
9	Halifax (City/Ville)	NS	31
97	Ottawa	QC	31
165	Oxford (South/Sud)	ON	29
26	Charlotte	NB	28
139	Simcoe (North/Nord)	ON	28
79	Québec (Ville/City)	QC	27
17	Kings	NS	26
172	Perth (North/Nord)	ON	26
27	Kings	NB	25
25	Saint John	NB	24
31	Carleton	NB	24
175	Huron (North/Nord)	ON	24
59	Shefford	QC	23
155	Grey (East/Est)	ON	23
177	Bruce (North/Nord)	ON	23
174	Huron (Centre)	ON	23
40	Rimouski	QC	22
61	Richelieu	QC	22
173	Huron (South/Sud)	ON	22
133	Ontario (North/Nord)	ON	22

Note: A family of emigrants is treated as one person. Queens in Prince Edward Island was the sixth most populous district in 1881.

Table 5: Sending Districts by Proportion of Outmigration to Population in 1881, top 35

District Number	Name	Province	pop %	raw count ranking
16	Annapolis	NS	3.73	4
14	Yarmouth	NS	3.48	5
18	Hants	NS	3.20	5
13	Shelburne	NS	2.77	37
163	Elgin (East/Est)	ON	2.53	8
8	Guysborough	NS	2.30	37
165	Oxford (South/Sud)	ON	2.28	18
68	St-Jean	QC	2.27	90
116	Kingston (City/Ville)	ON	2.20	66
17	Kings	NS	2.18	22
61	Richelieu	QC	2.14	32
55	Sherbrooke	QC	2.10	101
26	Charlotte	NB	2.10	19
10	Halifax	NS	2.09	12
102	Dundas	ON	2.09	32
23	Albert	NB	2.06	101
64	Iberville	QC	2.05	76
180	Kent	ON	2.01	5
31	Carleton	NB	2.00	25
146	Haldimand	ON	2.00	48
171	Perth (South/Sud)	ON	1.97	37
59	Shefford	QC	1.95	28
119	Prince Edward	ON	1.95	37
27	Kings	NB	1.92	24
2	Queens	PI	1.91	2
54	Compton	QC	1.91	48
58	Bagot	QC	1.87	42
33	Westmorland	NB	1.86	9
162	Waterloo (North/Nord)	ON	1.86	42
173	Huron (South/Sud)	ON	1.84	32
52	Drummond & Arthabaska	QC	1.83	11
1	Prince	PI	1.83	14
28	Queens	NB	1.81	101
177	Bruce (North/Nord)	ON	1.80	28
117	Lennox	ON	1.79	76

Note: A family of emigrants is treated as one person.

Table 6: Destination of Migrants from the Maritime Provinces

	French		Non-French	
	prop	s.e.	prop	s.e.
Massachusetts	0.625	0.0555	0.475	0.0174
Maine	0.154	0.0412	0.117	0.0111
New York	0.013	0.0129	0.051	0.0077
Michigan	0.026	0.0181	0.046	0.0073
California	—		0.037	0.0066
New Hampshire	0.026	0.0180	0.030	0.0060
Washington	—		0.022	0.0050
Minnesota	0.025	0.0174	0.021	0.0050
Wisconsin	0.040	0.0226	0.020	0.0049
Rhode Island	0.027	0.0191	0.017	0.0046
Connecticut	0.013	0.0130	0.017	0.0045
Illinois	—		0.017	0.0044
Vermont	0.027	0.0186	0.015	0.0043

Note: Proportion by Destination by Source Region. Destinations comprising at least 1.5% for either French or Non-French.

Table 7: Destination of Migrants from Québec

	French		Non-French	
	prop	s.e.	prop	s.e.
Massachusetts	0.400	0.173	0.154	0.0243
New Hampshire	0.136	0.0122	0.073	0.0175
Maine	0.089	0.0100	0.031	0.0115
Rhode Island	0.088	0.0101	0.005	0.0046
Connecticut	0.067	0.0088	0.009	0.0064
Michigan	0.043	0.0072	0.080	0.0181
New York	0.036	0.0066	0.145	0.0238
Vermont	0.034	0.0065	0.138	0.0235
Minnesota	0.025	0.0054	0.048	0.0142
Wisconsin	0.024	0.0055	0.055	0.0154
Illinois	0.017	0.0045	0.065	0.0162
Montana	0.013	0.0040	0.019	0.0092
North Dakota	0.012	0.0037	0.022	0.0097
Washington	0.005	0.0025	0.018	0.0089
Ohio	0.003	0.0018	0.018	0.0091
California	—		0.018	0.0091
Maryland	—		0.018	0.0090
New Jersey	—		0.018	0.0090

Note: Proportion by destination. Destinations comprising at least 1.5% for either French or Non-French.

Table 8: Destination of Migrants from Ontario

	French		Non-French	
	prop	s.e.	prop	s.e.
Michigan	0.286	0.0477	0.268	0.0112
New York	0.121	0.0345	0.142	0.0089
Illinois	0.063	0.0251	0.095	0.0073
North Dakota	0.010	0.0104	0.075	0.0066
Massachusetts	0.167	0.0395	0.064	0.0062
Minnesota	—		0.058	0.0059
Wisconsin	0.089	0.0300	0.039	0.0049
Ohio	0.011	0.0112	0.037	0.0048
California	—		0.026	0.0040
Washington	0.021	0.0149	0.025	0.0040
Maine	0.064	0.0254	0.017	0.0033
Pennsylvania	—		0.017	0.0032
Montana	0.022	0.0153	0.014	0.0030
Iowa	0.022	0.0158	0.011	0.0026
New Hampshire	0.033	0.0188	0.007	0.0021
Vermont	0.034	0.0193	0.005	0.0019

Note: Proportion by destination. Destinations comprising at least 1.5% for either French or Non-French.

Table 9: Proportion Urban in U.S. Given Rural/Urban Status in Canada. Three Definitions of Urban in U.S.

	Metro Area	City	Urban Area
Panel A: non-French			
rural	0.44	0.50	0.36
urban	0.63	0.67	0.55
Panel B: French			
rural	0.60	0.70	0.42
urban	0.69	0.75	0.51

Note: Metropolitan area, urban area and city are defined in U.S. Census. See text for definitions. 28% of non-French and 20% of French resided in urban places in Canada in 1881.

Table 10: Proportion by Occupation: Fathers of Matched Immigrants and Population

	non-French		French	
	prop	s.e.	prop	s.e.
Panel A: Fathers of Migrants				
white-collar	0.097	0.0071	0.065	0.0102
farmer	0.528	0.0121	0.555	0.0206
skilled	0.213	0.0099	0.180	0.0159
unskilled	0.163	0.0089	0.200	0.0166
Panel B: Population				
white-collar	0.116	0.0004	0.098	0.0005
farmer	0.480	0.0006	0.507	0.0009
skilled	0.201	0.0005	0.167	0.0007
unskilled	0.203	0.0005	0.227	0.0008

Table 11: Proportion by Occupation, by Father-Son Match (%)

occupation father	occupation son			
	white collar	farmer	skilled	unskilled
Panel A: non-French				
white-collar	3.9	0.5	2.8	2.3
farmer	8.5	7.8	17.2	19.2
skilled	4.3	1.4	10.3	5.8
unskilled	1.9	0.9	6.8	6.4
Panel B: French				
white-collar	1.7	0.0	3.7	1.2
farmer	3.2	1.8	34.1	16.5
skilled	2.6	0.5	10.3	5.2
unskilled	0.6	0.5	12.2	5.9

Note: $n = 1,818$. Not all fathers or sons report an occupation.

Table 12: Distribution of Sons' Occupation Given Fathers' Occupation (%)

occupation father	occupation son			
	white collar	farmer	skilled	unskilled
Panel A: non-French				
white-collar	0.41	0.05	0.29	0.24
farmer	0.16	0.15	0.33	0.36
skilled	0.20	0.06	0.47	0.27
unskilled	0.12	0.06	0.42	0.40
Panel B: French				
white-collar	0.26		0.56	0.19
farmer	0.06	0.03	0.61	0.30
skilled	0.14	0.02	0.55	0.28
unskilled	0.03	0.02	0.64	0.31

Note: $n = 1,818$. Not all fathers or sons report an occupation. Rows may not sum to 1.0 due to rounding.