The geography of economic mobility in 19th Century Canada

Luiza Antonie (Guelph), Kris Inwood (Guelph), Chris Minns (LSE) and Fraser Summerfield (St. FX)

April 2022

PRELIMINARY: please do not cite.

Abstract: To be written

Geographically extended economies may exhibit considerable regional differences in economic performance (Roses and Wolf 2018). Recent research has demonstrated that these spatial differences in economic performance have implications for inter-regional differences in intergenerational economic mobility. Chetty, Hendren, Kline, and Saez (2014) document how the economic characteristics of different locations are strongly connected with parent-child economic progress in the present day United States. Similar patterns have been identified in contemporary Canada (Corak 2020). But when did these geographical differences in opportunities emerge in the first place, and how persistent are they over time? Looking at long-run patterns in the United States, Connor and Storper (2020) find evidence of changes in mobility in the Midwest, where the composition of industry and employment has seen significant change over the twentieth century. Persistence of low mobility is a feature of the US South, where racial inequality is a dominant feature limiting local labour market opportunities.

Recent research by Corak (2020) lays out the geography of economic mobility in the Canadian landscape during the late 20th/early 21st century. But what patterns of local economic mobility looked like a century earlier are currently unknown, as are whether Canada has experienced a similar mix of persistence and change. The causes of such patterns are also an open question. While Canada may have had a less overt racial divide between regional labour markets than the US, labour market segmentation by language limited economic mobility in the late 19th century (Antonie, Inwood, Minns, Summerfield 2022) and the early 20th century saw notable differences in immigrant outcomes depending on origin country (Inwood Minns & Summerfield, 2016). Deindustrialisation has been a key feature of the Canadian economy in recent decades, with the declining share of manufacturing connected to unfavourable changes in employment opportunities for men (Morisette 2020). Did the rise of Canadian industry a century earlier have the opposite effect and create new opportunities for young men in Southern Ontario and Southern Quebec?

In this project we use linked Census data spanning 1871 to 1901 to provide the first portrait of local intergenerational mobility in Canada circa 1900. We construct multiple measures of economic mobility by 1871 Census sub-district and use these measures to document disparities in intergenerational mobility and their correlates at a local level. Our measures include broad occupation-based measures of mobility as well intergenerational elasticities and measures of absolute and relative rank mobility that draw on more fine-grained evidence of earnings patterns by occupation, age, and region (Connor and Storper 2020; Chetty , Hendren, Kline, and Saez 2014). Our next step will be to match up 1871 Census districts as closely as possible to contemporary data to explore how the landscape of mobility and their causes have changed over time.

Intergenerational Mobility in late 19th Century Canada

Our analysis draws on newly linked Canadian Census data for the period 1871-1901. A full overview of the approach to linkage and the characteristics of the data are available in Antonie, Inwood, Lizotte and Ross (2014) and Antonie, Inwood, Minns, and Summerfield (2021; 2022). Linkage draws on complete count census samples from 1871, 1881, 1891, and 1901. We begin by taking the records of boys aged 0 to 14 and co-resident fathers aged 18 to 80 in the 1871 Census Sample (in practice, fathers are the first adult male named in each household, see Antonie, Inwood, Minns, and Summerfield 2022; Collins and Zimran 2019) in 1871. We then link the records of the boys through to each of the 1881, 1891, and 1901 complete count census files. Linkage based on a machine learning approach, with a set of high-confidence links used for training (see Antonie, Inwood, Lizotte and Ross (2014). Names, year of birth range, and province of birth (?) are used as time-invariant characteristics on which linkage is based. We retain the records of the 1871 boys that we identify uniquely in every subsequent Census to 1901. We are able to double the size of the sample by using co-resident household members, where available, to disambiguate among candidate records for linkage.

Table 1 taken from Antonie, Inwood, Minns, and Summerfield (2022) summarizes the characteristics of the resultant linked sample in comparison to the complete count data from 1871. We present three versions of the linked data in this table: raw summary statistics from the linked sample with disambiguation (col. 2), a weighted version of the full linked sample, where we follow Bailey, Cole, Henderson, and Massey (2020) in constructing linkage weights (col. 3), and the linked sample based on unique links only (col. 4). The linked samples overrepresent younger, anglophone individuals, born Ontario or the Maritimes. Disambiguation does not appear to notably worsen the any bias in linkage – unique links (Col. 4) are only modestly closer on age and family size, while birthplaces are more mixed.

We use two approaches to looked at intergenerational economic outcomes. The first approach, used in recent studies of historical intergenerational occupational mobility, focuses directly on occupational outcomes by classifying occupations into four categories. White-collar work includes professionals, clerical workers, and proprietors. Skilled and semi-skilled work encompasses trades and craft work, factory operatives, and lower status service work. Unskilled includes labourers and farm labour. The final group, farmers, consists of farm owner/operators. To place all occupied men into these categories, we assign occupational strings a 4-digit OCCHISCO code (as prepared by IPUMS for US data), convert this code into HISCO using a crosswalk prepared by Evan Roberts, before placing these into 12 HISCLASS categories (van Leeuwen and Maas 2011). We then follow Perez (2019) in assigning these 12 categories into four broad groupings.[1]

We use the occupational information to compute gross mobility (M) as the share of sons who were in a different occupational class to their father. For comparison purposes, we also present Altham statistics (Altham 1970; Altham and Ferrie 2007) that summarize mobility of sub groupings of the adult male Canadian population. The general form of this statistic $d(\boldsymbol{P},\boldsymbol{Q})$ compares the column-row associations between any two contingency tables $\boldsymbol{P}$ and $\boldsymbol{Q}$ with r rows and s columns using the following formula:

$$d(\boldsymbol{P},\boldsymbol{Q}) = \left[\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{l=1}^{r}\sum_{m=1}^{s}\left|log\left(\frac{p_{ij}p_{lm}q_{im}q_{lj}}{p_{im}p_{lj}q_{ij}q_{lm}}\right)\right|^{2}\right]^{1/2} \qquad (1),$$

When a counterfactual table $\boldsymbol{J}$ with independent rows and columns is used as the comparison, $d(\boldsymbol{P},\boldsymbol{J})$ provides a ranking of mobility in table $\boldsymbol{P}$ against the benchmark of complete occupational mobility.[2]

An alternative approach to estimating mobility is to estimate the extent of intergenerational persistence in earnings. Estimates of intergenerational mobility in earnings in existing literature commonly take three forms. The first of these is the intergenerational earnings elasticity associated with a regression of the following type:

$$lny_{i,01} = \alpha + \beta lny_{i,71} + \varepsilon_i \qquad (2),$$

where for each observation i, β captures the elasticity of son's 1901 earnings $y_{i,01}$ with respect to father's 1871 earnings $y_{i,71}$. Our second measure of intergenerational earnings mobility estimates rank-rank correlations, or relative mobility, between the income rank of sons in 1901 and their fathers in 1871:

$$R_{S,i} = \omega + \gamma R_{F,i} + \varepsilon_i \qquad (3),$$

[1] We assign HISCLASS groups 1 to 5 (higher managers, higher professionals, lower managers, lower professionals, clerical, sales) to white-collar, groups 6, 7, and 9 (foremen, skilled workers, lower skilled workers) as skilled/semiskilled, groups 10 to 12 (lower-skilled farm workers, unskilled, unskilled farm workers) as unskilled, and group 8 (farmers and fishermen) as farming.

[2] Probabilities p_{ij} and q_{ij} are shares of first generation in occupation group i whose corresponding second generation is in occupation group j in economies P and Q, respectively. Thus, $p_{ij} = n_{ij}/\sum_i n_{ij}$. The Altham statistic can also be expressed using four-way odds ratios: $d(\boldsymbol{P},\boldsymbol{Q}) = \sqrt{\left[\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{l=1}^{r}\sum_{m=1}^{s}\Theta_{ijlm}{}^{2}\right]}$

where γ is estimated correlation between rank of sons ($R_{S,i}$) and their fathers ($R_{F,i}$). Finally, we generate an estimate of absolute mobility by calculating the estimated income ranks of sons of fathers at the 25th percentile in the income distribution. We follow Corak (2020) and Chetty Hendren Kline and Saez (2014) in estimating absolute mobility from equation (3) by computing the estimated rank of sons of father at the 25th percentile of the income distribution ($R_{S,i} = \omega + \gamma * 25 + \varepsilon_i$).[3]

The lack of detailed earnings information in Census data has presented a major challenge in the estimation of intergenerational earnings mobility in the past. While the 1901 Census of Canada enquired into the income of respondents, no such information was collected in 1871. As a result, we follow much of the recent literature and construct proxy measures for income by occupation. A long literature uses occupational income scores as a measure of economic status (see Inwood, Minns, Summerfield 2019 for a review), and it is an approach with drawbacks and advantages. One main drawback is the lack of granularity that comes when assigning an income score based on occupation. Solutions have been developed to reintroduce plausible individual variation into occupational earnings data (Saavedra and Twiman 2019), but these methods will struggle with occupations like farming where reported incomes are scarce even in benchmark years. A second drawback is that occupational premiums may have changed over time. In this case using information from benchmark years that imposes the pay structure at some future date could bias estimates of intergenerational mobility, in particular if there have been large unobserved changes to the wage structure over time. Occupational income scores may have an advantage in the context of evaluating intergenerational mobility if an occupational score is a better measure of permanent status than the income observed in the year the Census was taken (see Solon 1992). Previous studies have suggested that the use of income scores rather than earnings present less of an issue in this context (see Feigenbaum 2019, Saavedra and Twiman 2019; Inwood, Minns, and Summerfield 2019). In addition to computing intergenerational elasticities, occupational income scores also allow the estimation of rank-based measures of relative and absolute economic mobility (Chetty, Hendren, Kline, and Saez 2014) that permit direct comparisons to both contemporary and historical findings for North America (Corak 2020; Connor and Storper 2020) and rank-based estimates may be less sensitive to mismeasurement of individual earnings that the conventional intergenerational elasticities

We use complete count data from the 1901 Census to construct occupational income scores that can be applied to sample individuals in 1901 and 1871. We have experimented with different

[3] To provide inference for regression predictions at $R_{F,i} = 25$ we redefine this variable as $R_{F,i} - 25$, at which point the constant term provides the value γ and it's 95% confidence interval.

levels of aggregation in this exercise: occupational income scores are created for 432 5-digit HISCO occupations, and in our preferred versions, also computed separately by region and age cohort.[4] We have several approaches to deal with the issue of incomes for farmers. The first is to use mean actual income as reported in the 1901 Census sample. This is somewhat unsatisfactory due to the majority of farmers not reporting their earnings, but it is the simplest approach one can take to the problem. In the next round of revisions we will apply estimates of regional differences in 1871 farm income from Inwood and Irwin (1992) which will provide more fine grained variation in farm earnings within Eastern Canadian provinces.

The regional geography of Canadian occupational mobility, 1871-1901

We begin our analysis by calculating measures of economic mobility by province, region, and demographic group. In all models we follow Bailey et al (2020) and use weights to account for variation in the likelihood of different individuals being part of the linked sample, though for each set of results we display differences between weighted and unweighted estimates for the complete sample. In these tables we show separate results for English and French Canadians, and for sons originating in Ontario, Quebec, or the Maritimes.

The first measure is gross intergenerational occupational mobility (M). M is the share of all cases where sons in 1901 were not in the same occupational group as their fathers had been in 1871, taking on a theoretical maximum value of 1. Table 2 presents calculations of gross mobility in the linked 1871-1901 sample. These figures indicate only modest differences in overall mobility by region, with the highest rates of mobility in Ontario and the Maritimes (0.52) and the lowest rates of mobility in Quebec (0.48). The regional differences are matched by a similar gap between English and French Canadians (0.52 and 0.48). Table 2 also presents Altham statistics by regional and ethnic group. These show more marked differences in mobility between regions and groups, with higher occupational mobility in Ontario and among Anglophones, and the lowest occupational mobilty in Quebec and among Francophones.

* Insert Table 2 here *

Table 3 shows results for estimated intergenerational elasticities. In addition to the seven groups shown in Table 2, we present estimates with four different ways of calculating occupational scores.

---

[4] Three regions are Ontario, Quebec and the Maritimes. Cohorts are 5-year age groups, starting from age 16, with the final group including individuals aged 65-70.

The results in Panel A show intergenerational correlation coefficients between 0.42 and 0.46 when a uniform occupational income score is used in all regions and for all ages. Using more disaggregated occupational income scores in Panels B, C, and D yields lower rates of estimated intergenerational mobility (higher intergenerational elasticities) of 0.46 to 0.54 for the full Canadian sample in columns (1) and (2). Distinctive patterns of geographical variation emerge when we allow occupational income to vary by region in panels C and D. If we focus on our preferred specification in Panel D, where scores are allowed by vary both by region and age cohort, the results show a clear east-west gradient in mobility, with lower mobility for sons originating in Maritime provinces (b=0.47) than in Quebec (b=0.45), and much higher mobility in Ontario (b=0.37) than anywhere else in the country. This specification also yields somewhat higher mobility overall for francophones than English Canadians (b=0.47 versus b=0.53). These elasticity estimates are higher than what has typically been reported in similar models for the US and Canada in both historical and more recent settings. Ward (2021) reports historic elasticity estimates for white Americans in the order of 0.3 before correcting for measurement error; Chetty, Hendren, Kilne and Saez have similar findings (b=0.30 to 0.35), while Corak's (2020) results suggest much higher mobility in Canada (b=0.20).

* Insert Table 3 here *

One conclusion to draw from comparisons of Tables 2 and 3 is that high gross mobility did not necessarily translate into higher intergenerational elasticities when more fine-grained measures of status are used. This pattern is supported by further results in Tables 4 and 5 for rank-based measures of intergenerational mobility. Focusing on our preferred specification in Panel D, the rank-rank regressions in Table 4 show Ontario as the clear regional leader (g=0.28), with Quebec and the Maritimes near equality (g=0.44 and g=0.43). Differences between English and French Canadians are small in this specification (g=0.47 and g=0.46). As with the intergenerational elasticity estimates in Table 3, these findings suggest lower mobility in late 19th Century Canada than in present-day United States. Corak (2020) reports rank-rank elasticities in the order of 0.24, while Chetty, Hendren, Kline, and Saez's figure is about 0.34. Table 5 reports on absolute mobility through predicted son ranks for fathers at the 25th percentile of the occupational income score distribution. Ontario remains the leader in this specification (predicted rank 43rd percentile), with a smaller gap to Quebec (predicted rank 39th percentile), and the Maritime region much further behind (predicted rank 27th percentile). These results are further evidence of lower regional mobility in 19th century Canada than in available comparisons for present-day Canada. Corak (2020) reports the equivalent mean rank of sons lay at the 44th percentile for all of Canada, and quite large differences for Ontario (47-48th percentile), Queberc (46th percentile), and New Brunswick and Nova Scotia (both at 38th percentile). Historical estimates for the US in Connor and Storper suggest a similar wide range in mobility to Canada, with sons of 25th percentile fathers predicted to lay below the 40th percentile in the South, but above the 50th percentile in the West and Midwest.

* Insert Table 4 here *

* Insert Table 5 here *

Note to self – make nice visuals of these for the CNEH talk.

Mapping Mobility in 19th Century Canada

The linked 1871-1901 sample is large enough to allow a more fine-grained exploration of Canadian mobility patterns in the 19th century. We have constructed a subset of the mobility measures described above for divisions that correspond broadly to Canadian Census districts of 1871. The divisions we use are partially aggregated census districts to account for small sample size in some cases.[5] As a result we have xxx divisions for which we have computed mobility measures. While we have computed numerous mobility measures at district level, for the purposes of the current draft we focus primarily on gross mobility (M), rank-rank correlations of relative mobility and the mean rank of sons of fathers at the 25th percentile of the occupational earnings distribution (absolute mobility). These three measures allow for international and intertemporal comparisons, and for further exploration the contrasting outcomes between unranked mobility and the more ordinal outcomes associated with change in rank.

Figures 1 and 2 display gross mobility by adapted district. In figure 1 we array districts in order of increasing mobility. Districts from Ontario Quebec, and the Maritimes are shaded/coloured for identification, and we include confidence intervals n estimates. Figure 1 shows that local mobility varies much more widely than provincial mobility reported in Table 2. The clustering of a number of districts in Nova Scotia at the top end of the mobility Table underpins the high gross mobility rates reported for the region as a whole. The wide range in reported gross mobility is not an artifact of small sample size in some districts – note the almost 10 percentage point (?) difference in gross mobility between Montreal and Toronto, and even larger differences between Toronto and London and Hamilton, two moderate sized industrial cities in Southern Ontario. Figure 2 plots the gross mobility on a map of Canada from the 1871 Census file. On this measure, pockets of high

[5] See the appendix for details on how census districts were merged. All districts used the analysis presented below have at least 30 (???) father-son linked observations.

occupational mobility were in Cape Breton/northern Nova Scotia, southern New Brunswick, and western Ontario, while much of Quebec saw low intergenerational occupational mobility.

*** Insert Figures 1 and 2 here ***

Figures 3 and 4 repeat the previous exercise but with district rank-rank correlations as the main measure of interest. These figures show that once occupations are ranked (and against a relatively fine-grained classification) we get a different sense of 19th century Canadian mobility patterns. Districts with relatively high gross mobility (i.e. a lot of intergenerational occupational turnover) in the Maritimes and Western Ontario often feature among the steepest rank slopes, suggesting less mobility across the income/occupational distribution. Southern Ontario and parts of Quebec feature relatively low rank elasticities between fathers and sons.

*** Insert Figures 3 and 4 here ***

Figures 5 and 6 repeat the previous exercise but with the mean predicted rank of sons of father at the 25th percentile as the outcome used. Figure 3 shows a wide range in son outcomes relative to their fathers as we say in figure 1, but the relative ranking of districts is substantially different. Large and medium sized urban districts offered the best prospects for upward mobility, with Hamilton, London, Ottawa, Montreal, and Toronto at the top if the table. There is some overlap with the leading districts for gross mobility, but the higher position of the largest cities here (Montreal and Toronto) suggests that movement within categories to better paid jobs was significant for sons originating in these districts. The bottom end of the distribution is dominated by districts in Nova Scotia and New Brunswick, including those that had high rates of gross mobility in Table 1. Hence it appears that districts with high gross occupational mobility in 19th century eastern Canada features many more sideways or backwards moves than in Quebec or Ontario. Figure 4 places the son ranks on the map. The regional ranking here is much more clear cut than in Figure 2 for gross mobility. Many (geographically small) Ontario districts saw sons more than 20 percentage points above their 25th percentile fathers. Quebec districts were generally in the intermediate range with some notable high mobility pockets, while Maritime were much lower on the mobility scale, with more than 10 reporting average son ranks no better than that of their father.

*** Insert Figures 5 and 6 here ***

Correlates of 19th Century Mobility

We follow Chetty et al (2014) Corak (2020) and Connor and Storper (2020) in assessing the correlates of the two measures of district-level mobility described above. We use readily available district characteristics from the 1871 Census of Canada that roughly correspond to the features that have been considered in the contemporary literature. For the role or urbanization and population dynamics, correlates are population growth between 1861 and 1871, population density in 1871, and average family size. Labour market characteristics included as correlates are occupational shares for a larger group of occupational classifications than we use to construct gross mobility (professional, industrial, domestic, commerce, agriculture, and other unclassified occupations), and an occupational-based Gini coefficient that measures inequality in occupational earnings within each district. Human capital correlates include current human capital levels in each district through the share able to write, and investment in human capital as the share of school age children enrolled (?) in education. Finally, we also have four characteristics related to migration and ethnicity: the share of population in each district born in province; the share Canadian born, the share Catholic, and the share French speaking.

Figure 7 summarizes bivariate correlation coefficients between the above characteristics and gross mobility M. We find little evidence that district demographics mattered; none of the three characteristics exhibit significant correlation with gross mobility. Occupational structure of the district does have relevance: more agricultural districts show less mobility, while districts with more industrial activity and a higher occupational gini had more mobility. This is likely indiciative of diversity of local opportunity in shaping the occupational paths of sons relative to their fathers. Districts with higher literacy rates were not sources of mobility, but greater school enrolment did promote occupational mobility, while locations with more Canadians, less internal migrants, more Catholics, and more French Canadians saw less occupational mobility.

Insert Figure 7 here

Figure 8 shows the same bivariate correlations but with son rank as the outcome of interest. Some of the patterns observed in Figure 7 continue to hold in this case: more agricultural districts saw less upward mobility, as did those with fewer migrants (internal or international). Upward mobility was positively correlated more strongly with commercial and professional occupations than industry, and also with population density (our close proxy for urbanization), which is where such occupations were more readily available. Schooling of the young remains a strong positive correlate of upward mobility, and we note no notable correlation between religion and language and upward mobility. While French Canadian catholic districts had less intergenerational occupational turnover than elsewhere, there was no notable penalty in terms of sons occupational rank relative to (low status) fathers.

These results can be compared to Conor and Storper's (2020) mobility correlates for the early 20th Century US. Our patterns accord with their in terms of the importance of population density and the presence of migrants in generating greater absolute mobility. We do not share their finding of the strength of manufacturing or industry in creating mobile district – this may in part reflect the timing of the Canadian data, the smaller share of industry in Canada at this time, or differences in how we classify activity into sectors. We also find little evidence that human capital among adults in the district was a driver of absolute mobility, nor that inequality supressed mobility

Insert Figure 8 here

Finally, Figure 9 displays bivariate correlations between district correlations and the rank-rank coefficients on father and son occupational earnings. Far fewer of these correlates are statistically significant – the share born in Canada and in province are positively correlated with rank-rank coeffiicients (indicating less relative mobility), as is the share of adults who could write. It appears that all measures of mobility are positively correlated with the presence of migrants, and interestingly there is no evidence that a more educated local population was associated with greater mobility of any kind.

Insert Figure 9 here

Conclusions

To be written

References

Altham (1970), "The measurement of association of rows and columns for an r x s contingency table." Journal of the Royal Statistical Society, Series B (Methodological) 32(1), pp. 63-73.

Altham and Ferrie (2007), "Comparing contingency table tools for analyzing data from two groups cross-classified by two characteristics." Historical Methods 40(1), pp,. 3-16.

Luiza Antonie, Kris Inwood, Daniel J. Lizotte, and J. Andrew Ross (2014), "Tracking people over time in 19th century Canada for longitudinal analysis." Machine Learning 95, pp. 129-146.

Luiza Antonie, Kris Inwood, Chris Minns, and Fraser Summerfield (2020), "Selection bias encountered in the systenmatic linkage of historical Census records." Social Science History, 44(3) pp. 555-570.

Luiza Antonie, Kris Inwood, Chris Minns, and Fraser Summerfield (2021), "Intergenerational Mobility in a Mid-Atlantic Economy: Canada 1871-1901." LSE Economic History Working Paper 319.

Martha Bailey, Connor Cole, Morgan Henderson, and Catherine Massey (2020), "How well do automated linkage methods perform? Lessons from US historical data." Journal of Economic Literature, 58(4), pp. 997-1044.

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014), "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States." Quarterly Journal of Economics, 129(4), pp. 1553-1623.

Dylan Connor and Michael Storper (2020), "The changing geography of mobility in the United States." PNAS 117(48), pp. 30309-30317.

Bill Collins and Ariel Zimran (2019), "The economic assimilation of Irish famine immigrants in the United States." Explorations in Economic History 74, 101302.

Miles Corak (2020), "The Canadian Geography of Intergenerational Income Mobility." Economic Journal, 130(631), pp. 2134-2174.

Inwood, Minns, and Summerfield (2019), "Occupational income scores and immigrant assimilation. Evidence from the Canadian Census." Explorations in Economic History, 72, pp. 114-122.

James Feigenbaum (2018), "Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940." Economic Journal, 128(612), pp. F446-F481.

René Morissette (2020), "The Impact of Manufacturing Decline on Local Labour Markets in Canada." Statistics Canada, Analytical Studies Branch Research Paper Series.

Perez (2019), “Intergenerational occupational mobility across three continents.” Journal of Economic History, 79(2), pp. 383-416.

Joan Roses and Nikolaus Wolf (2018) “Regional economic development in Europe, 1900-2010: a description of the patterns.” CEPR discussion paper 12749.

Saavedra and Twiman (2020) “A machine learning approach to improving occupational income scores.” Explorations in Economic History 75, 101304.

Solon (1992), “Intergenerational income mobility in the United States.” American Economic Review 82(3), pp. 393-408.

van Leeuven and Maas (2011), HISCLASS: A Historical International Social Class Scheme. University Press Leuven.

Table 1: Linked and unlinked sample characteristics, 1871

	(1)	(2)	(3)	(4)
	1871 full count	1871-1901 linked	1871-1901 linked, weighted	Unique links
1871 Age	6.8 (4.3)***	6.6 (4.2)	6.9 (4.3)	6.9 (4.2)***
% hhlds with 5+ children	0.55***	0.53	0.56	0.55***
Born NS	0.10***	0.14	0.11	0.15***
Born NB	0.08***	0.10	0.08	0.12***
Born QC	0.34***	0.21	0.32	0.21
Born ON	0.44***	0.53	0.45	0.49***
Born UK & Ireland	0.02**	0.01	0.01	0.01**
Born Elsewhere	0.04***	0.01	0.03	0.03
Reside NS	0.10***	0.14	0.11	0.16***
Reside NB	0.08***	0.09	0.08	0.12***

Reside QC	0.34***	0.21	0.33	0.21
Reside ON	0.48***	0.54	0.48	0.52***
Head white collar	0.09***	0.08	0.08	0.08*
Head skilled/semi skilled	0.19***	0.16	0.15	0.16**
Head unskilled	0.25***	0.13	0.14	0.14*
Head farm	0.48***	0.64	0.64	0.62***
French Eth.	0.32***	0.18	0.31	0.18
Anglo Eth.	0.60***	0.71	0.61	0.68***
No Female >22 in hhld	0.03***	0.02	0.03	0.02
N	733,355	32,365	32,365	17,256

Sources: 1871 complete count Canadian Census and linked 1871-1901 Canadian Census records.

Notes: See text for sample descriptions. *, **, and *** denote significant differences between each unweighted sample and the linked sample at 90, 95, and 99 percent confidence intervals. Full count sample limited to males aged 0 to 14 in 1871. Unique links refers to all three linkages: 71-81, 81-91 and 91-01. Head occupation "unclassified" are omitted. Children defined as individuals enumerated with the same household id age 0-17, inclusive. Anglophone includes ethnicities reported as English, Welsh, Scottish, Irish and North American. Standard deviations for age in parentheses. Sample size for column 1 is 732,798 for province of residence and 95,176 for head occupations.

Table 2: Occupational Mobility in Canada, 1871-1901

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
M	0.51	0.51	0.48	0.52	0.52	0.48	0.52
Altham Statistic	16.0	16.2	17.8	15.7	16.2	17.4	15.6

Notes: M is gross mobility, the share of sons in 1901 not in the same occupational group as their father in 1871, based on a weighted contingency matrix. Altham statistics calculated with weighted regressions following Modalsli (2015). All Altham statistics are significantly difference from the null hypothesis of independence between son and father outcomes.

Table 3: Canadian Intergenerational Elasticities, 1871-1901

Panel A: Occscores by HISCO only

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
IGE slope	0.446***	0.437***	0.428***	0.440***	0.464***	0.426***	0.427***
	(0.010)	(0.010)	(0.024)	(0.012)	(0.021)	(0.020)	(0.014)
Observations	32,470	32,470	6,018	23,059	7,795	7,036	17,639
R-squared	0.120	0.116	0.110	0.120	0.133	0.118	0.106

Panel B: Occscores by HISCO and 5-yr age group

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
IGE slope	0.482***	0.475***	0.460***	0.485***	0.479***	0.467***	0.476***
	(0.009)	(0.010)	(0.023)	(0.011)	(0.020)	(0.020)	(0.013)
Observations	32,238	32,238	5,984	22,879	7,741	6,978	17,519
R-squared	0.138	0.136	0.126	0.142	0.136	0.141	0.130

Panel C: Occscores by HISCO and 3-regions

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
IGE slope	0.494***	0.459***	0.395***	0.488***	0.455***	0.367***	0.288***
	(0.008)	(0.010)	(0.021)	(0.010)	(0.021)	(0.019)	(0.015)
Observations	32,420	32,420	6,007	23,020	7,777	7,006	17,637
R-squared	0.171	0.150	0.114	0.167	0.126	0.105	0.048

Panel D: Occscores by HISCO, 5yr Age groups and 3-regions

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
IGE slope	0.537***	0.511***	0.473***	0.533***	0.472***	0.449***	0.368***
	(0.009)	(0.010)	(0.022)	(0.011)	(0.020)	(0.020)	(0.015)
Observations	31,548	31,548	5,897	22,315	7,425	6,811	17,312
R-squared	0.175	0.159	0.128	0.177	0.123	0.120	0.077

model: ln(OCC_son)=a+b*ln(OCC_fathrt)
robust SE in parentheses
All columns except 1 use weights
Occ scores nominal

Table 4: Rank-rank correlations for Canada 1871-1901

Panel A: Occscores by HISCO only

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
RR slope	0.330***	0.313***	0.306***	0.320***	0.348***	0.308***	0.295***
	(0.007)	(0.007)	(0.015)	(0.009)	(0.015)	(0.014)	(0.010)
Observations	32,470	32,470	6,018	23,059	7,795	7,036	17,639
R-squared	0.086	0.080	0.080	0.082	0.097	0.084	0.069

Panel B: Occscores by HISCO and 5-yr age group

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
RR slope	0.392***	0.393***	0.407***	0.393***	0.375***	0.416***	0.381***
	(0.006)	(0.006)	(0.014)	(0.008)	(0.014)	(0.013)	(0.009)
Observations	32,238	32,238	5,984	22,879	7,741	6,978	17,519
R-squared	0.130	0.134	0.151	0.131	0.121	0.162	0.120

Panel C: Occscores by HISCO and 3-regions

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
RR slope	0.543***	0.510***	0.473***	0.521***	0.455***	0.425***	0.203***
	(0.006)	(0.007)	(0.016)	(0.008)	(0.020)	(0.014)	(0.013)
Observations	32,420	32,420	6,007	23,020	7,777	7,006	17,637
R-squared	0.209	0.186	0.154	0.194	0.118	0.136	0.021

Panel D: Occscores by HISCO, 5yr Age groups and 3-regions

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
RR slope	0.485***	0.465***	0.460***	0.472***	0.431***	0.436***	0.283***
	(0.006)	(0.007)	(0.015)	(0.008)	(0.017)	(0.014)	(0.011)
Observations	31,548	31,548	5,897	22,315	7,425	6,811	17,312
R-squared	0.195	0.179	0.161	0.187	0.125	0.150	0.056

model:Q_son = w+g*Q_father
robust SE in parentheses
All columns except 1 use weights
Occ scores nominal

Table 5: Son occupational income ranks of fathers at 25th percentile, Canada 1871-1901

Panel A: Occscores by HISCO only

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
mean rank	41.777***	42.569***	41.959***	42.945***	38.313***	43.282***	43.985***
	(0.228)	(0.254)	(0.501)	(0.313)	(0.538)	(0.487)	(0.346)
Observations	32,470	32,470	6,018	23,059	7,795	7,036	17,639

Panel B: Occscores by HISCO and 5-yr age group

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
mean rank	38.549***	38.930***	38.487***	39.150***	36.579***	39.391***	39.597***
	(0.183)	(0.202)	(0.401)	(0.246)	(0.420)	(0.386)	(0.279)
Observations	32,238	32,238	5,984	22,879	7,741	6,978	17,519

Panel C: Occscores by HISCO and 3-regions

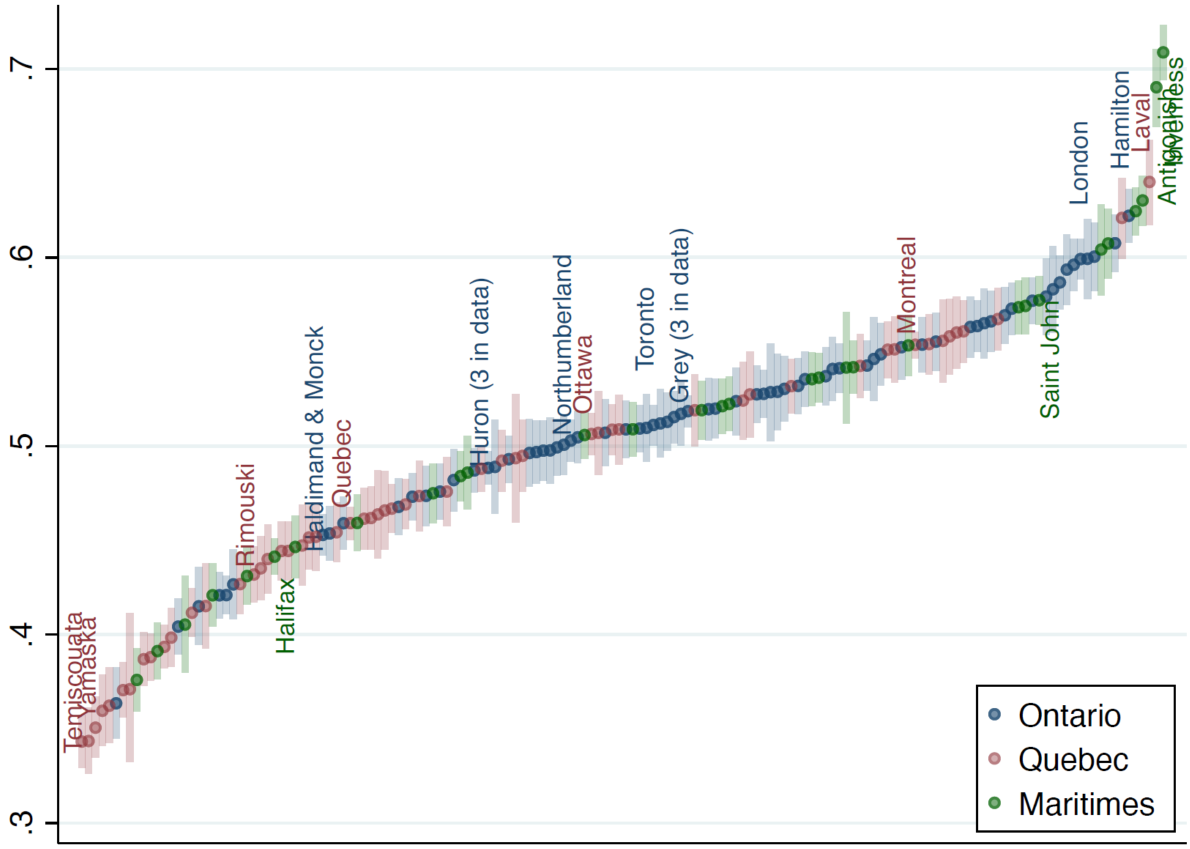
	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
mean rank	40.596***	42.124***	43.728***	41.448***	29.967***	46.299***	52.908***
	(0.179)	(0.198)	(0.364)	(0.248)	(0.396)	(0.356)	(0.408)
Observations	32,420	32,420	6,007	23,020	7,777	7,006	17,637

Panel D: Occscores by HISCO, 5yr Age groups and 3-regions

	(1) Unweighted	(2) Weighted	(3) Franco	(4) Anglo	(5) Mar	(6) Que	(7) Ont
mean rank	35.470***	36.590***	37.794***	35.993***	27.457***	39.433***	43.068***
	(0.170)	(0.190)	(0.356)	(0.233)	(0.336)	(0.349)	(0.346)
Observations	31,548	31,548	5,897	22,315	7,425	6,811	17,312

Fitted values from model in table 2 when father=25
robust SE in parentheses are SEs for the prediction
All columns except 1 use weights
Occ scores nominal

Figure 1: Gross Mobility (M) by 1871 district of origin



Notes:

Figure 2: Gross Mobility (M) by 1871 district of origin

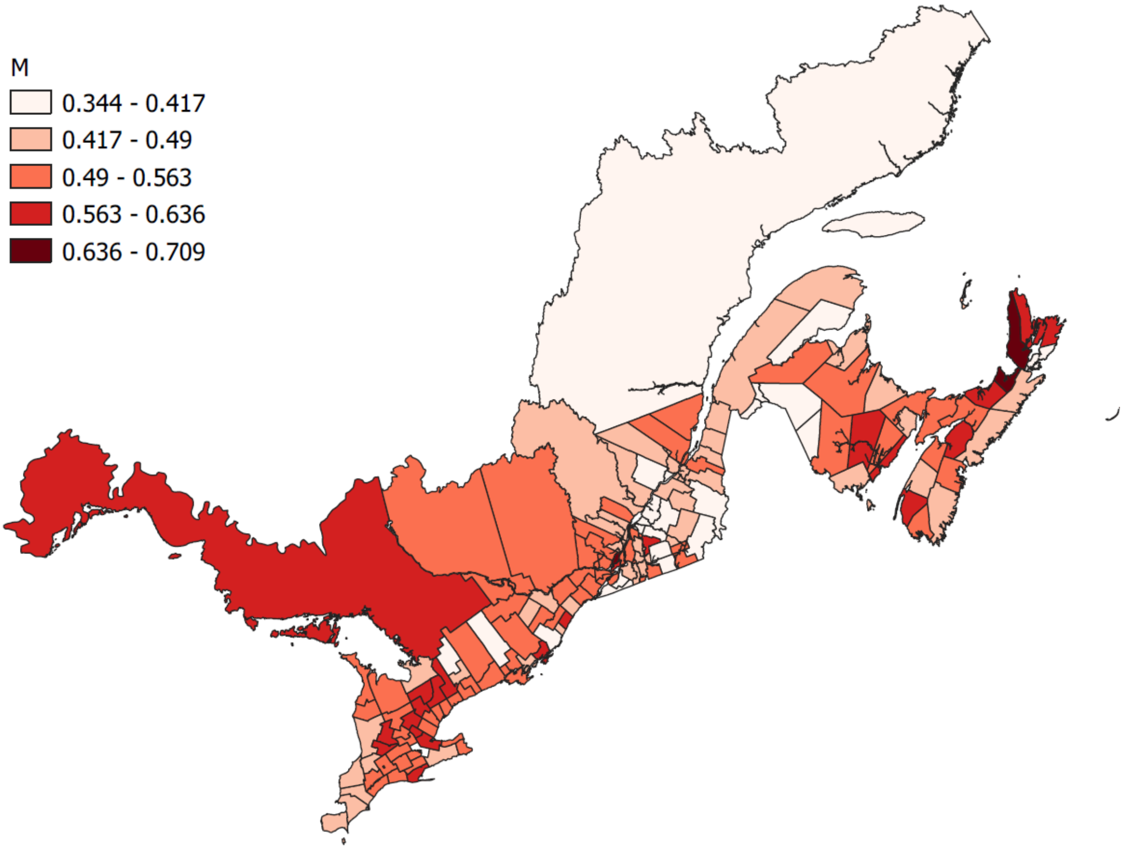
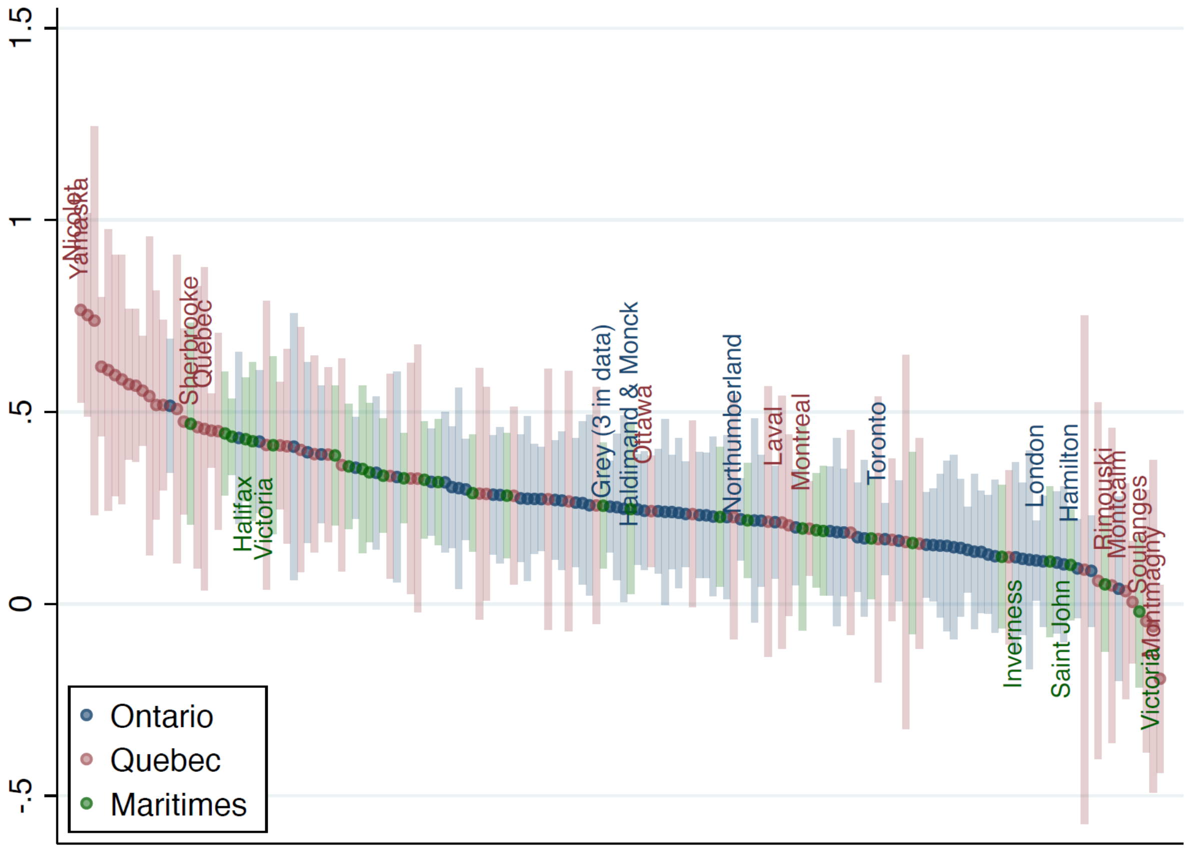


Figure 3: rank-rank correlations by 1871 Census district of origin



Notes:

Figure 4: rank-rank correlations by 1871 Census district of origin

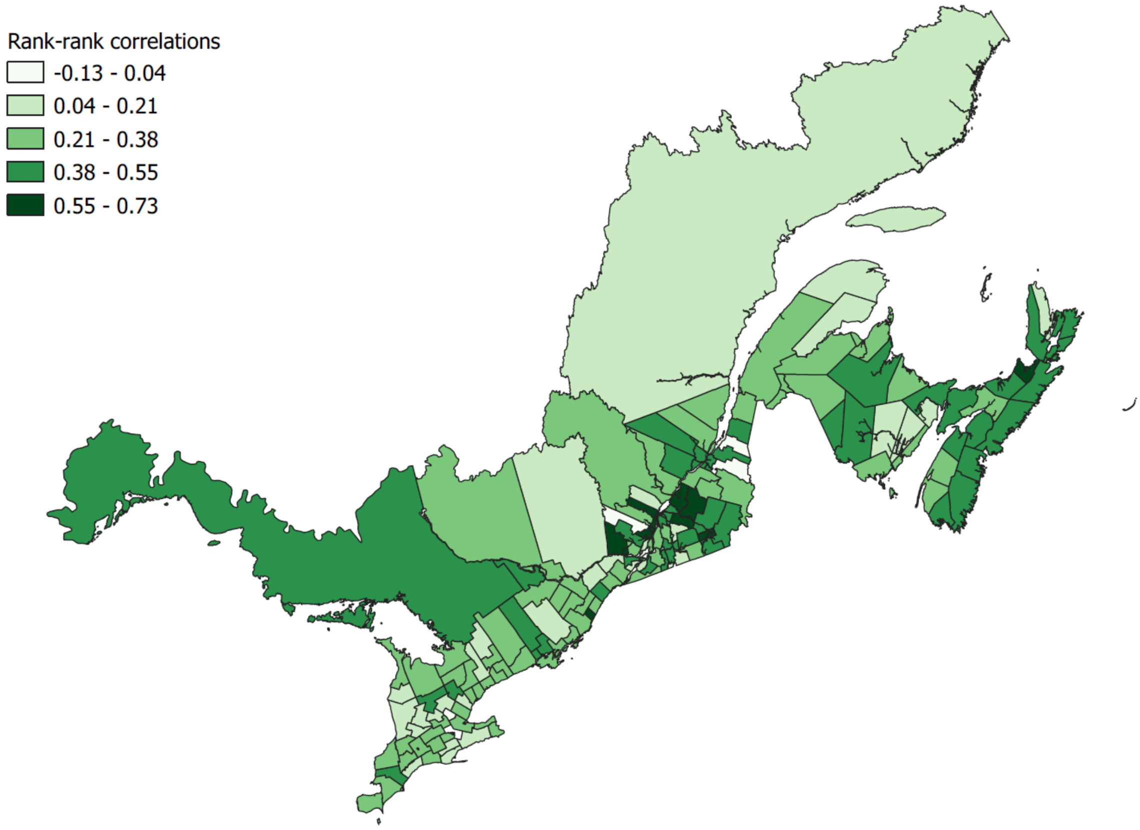


Figure 5: Son rank of 25th percentile fathers by 1871 Census district of origin

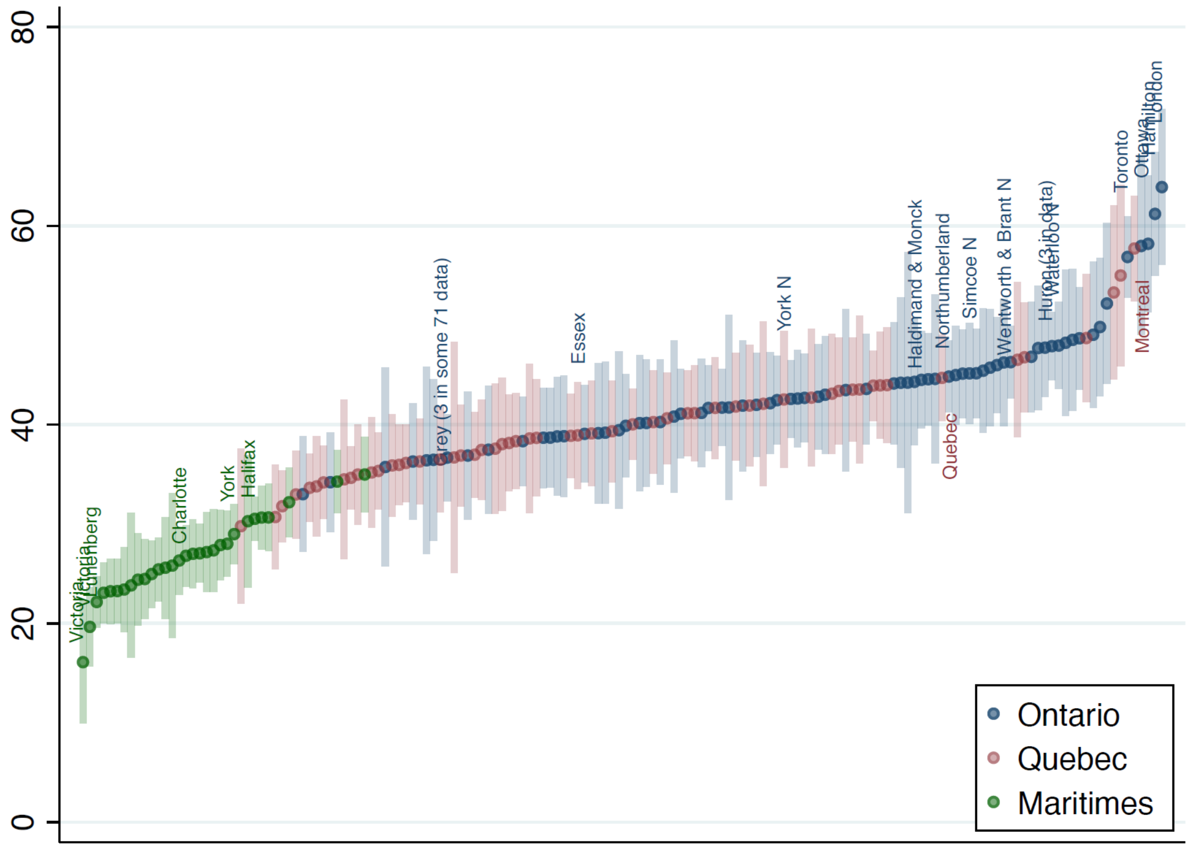


Figure 6: Son rank of 25th percentile fathers by 1871 Census district of origin

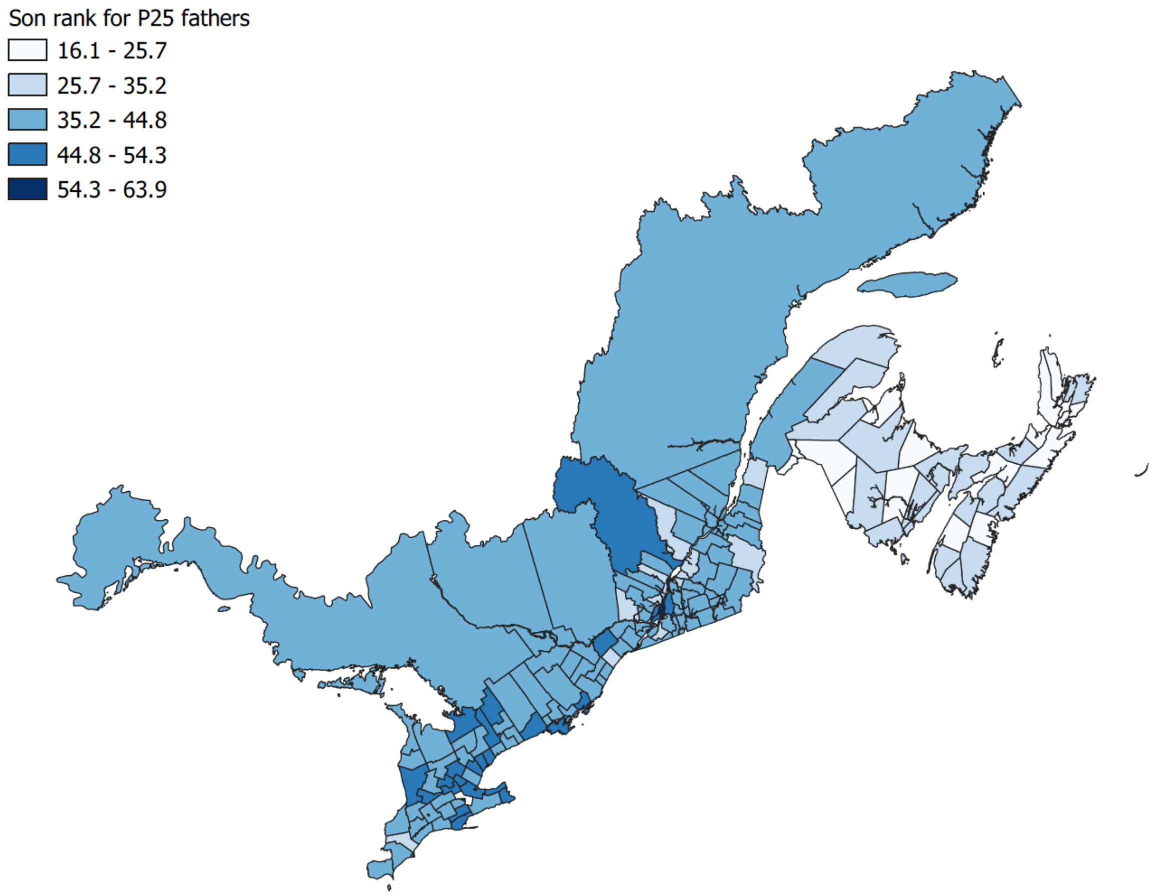
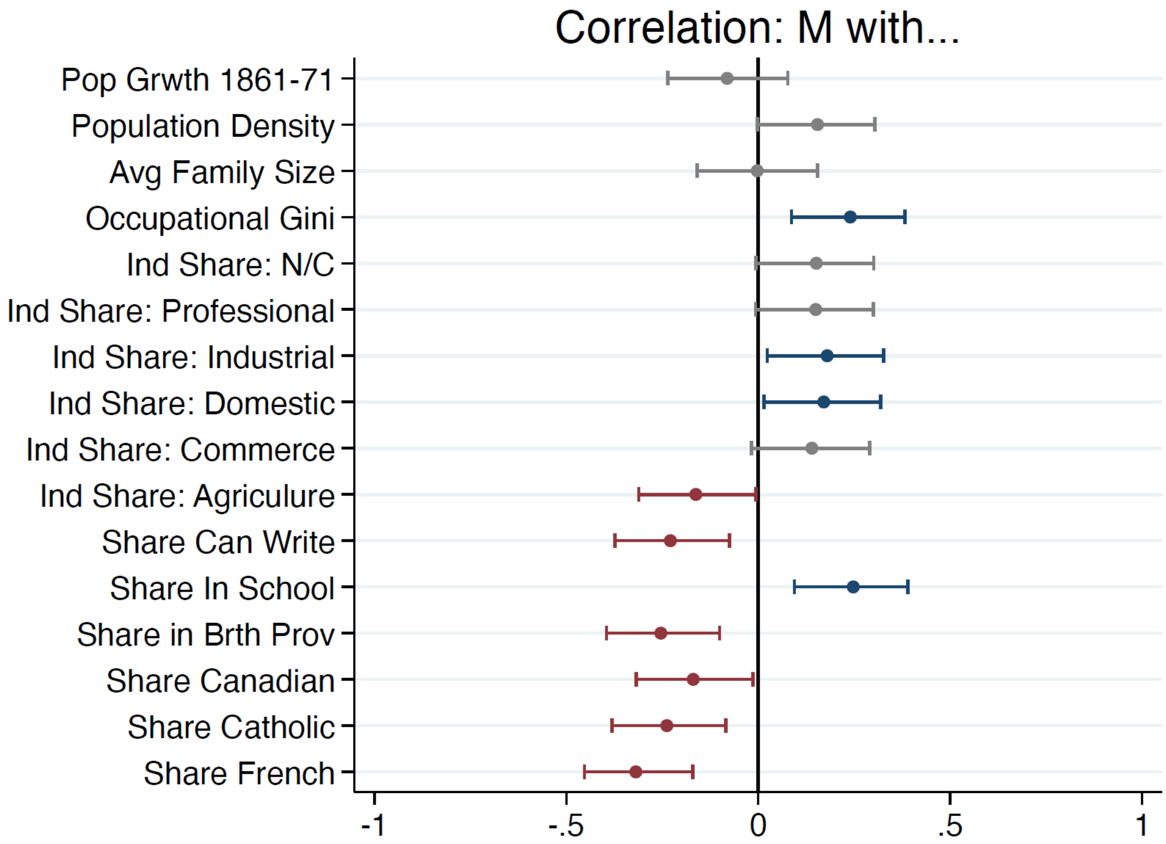
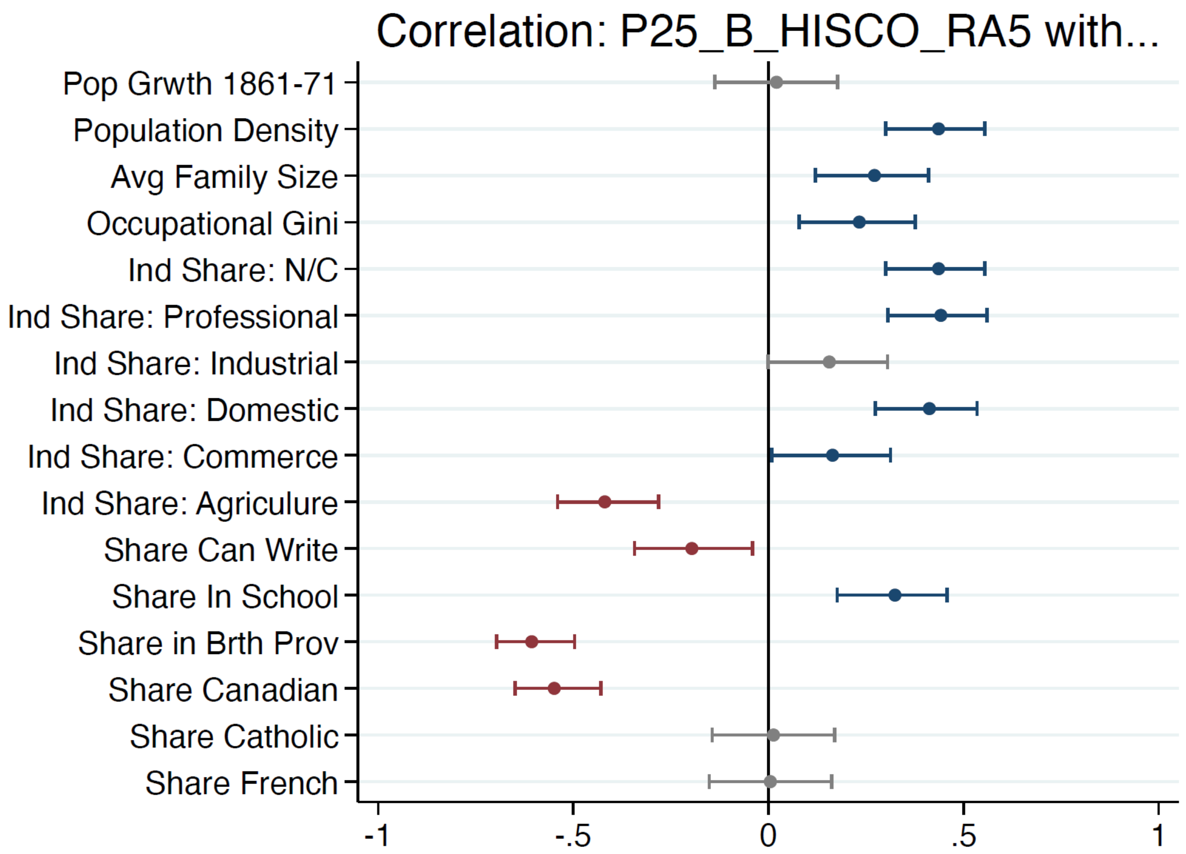


Figure 7: Correlates of District Gross Mobility, 1871-1901



Notes:

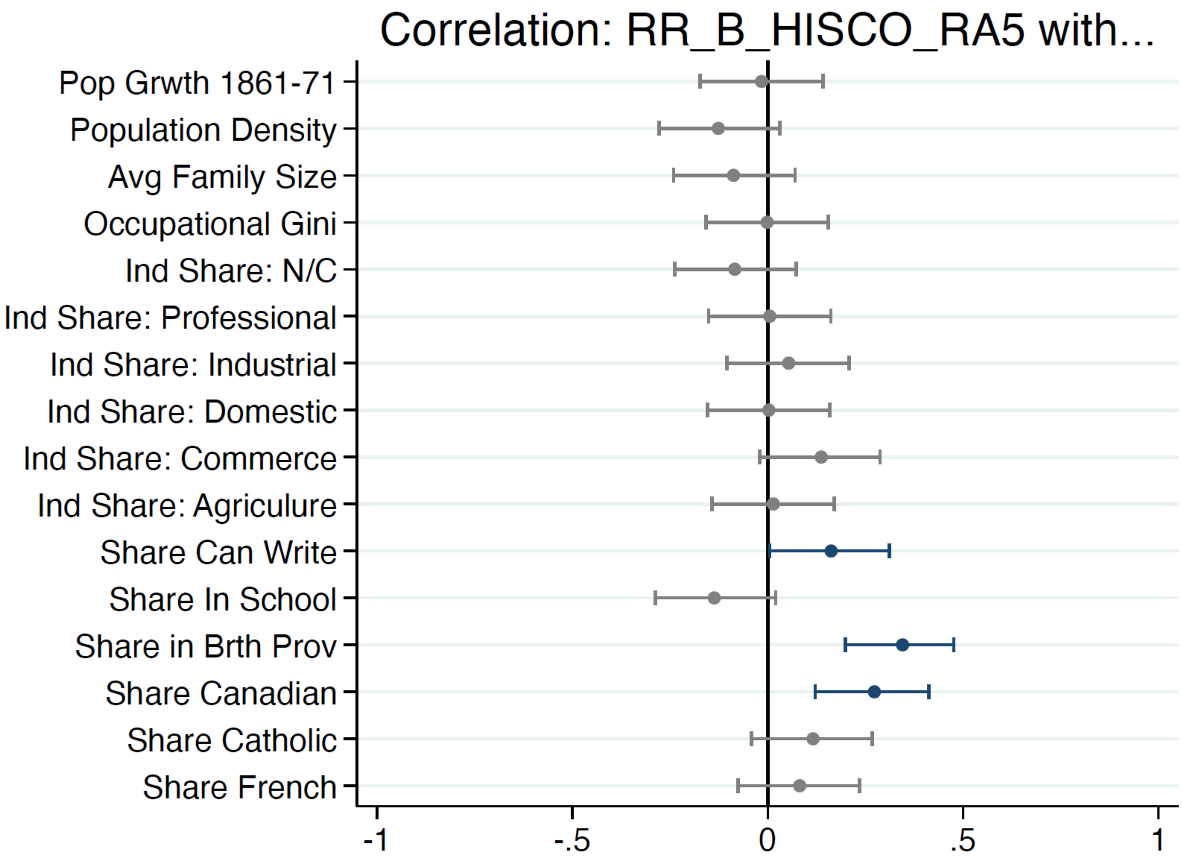
Figure 8: District correlates of son rank for bottom quartile fathers



Notes:

Figure 9: district correlates of rank-rank mobility

/



Notes: