The her in inheritance: assortative marriage and intergenerational mobility in Quebec 1800–1970

Matthew Curtis

September 20, 2020

Rough draft

Abstract

Before it was common for married women to earn incomes, was assortative marriage an important determinant for inequality? Using a new dataset of linked marriages from Quebec 1800–1970, I show that assortment — and women — played a major role in intergenerational mobility and therefore mattered for inequality over the long-run. Using a latent variable model, I construct a measure of marital assortment that shows that matching was as strong in 1800 as it was in the mid 20th century. I then show that this assortment was on the characteristics of individuals, not just between similar families. Finally, I show that the characteristics of both mothers and fathers mattered to a comparable degree in determining the outcomes of children. Assortative marriage mattered because, despite severe legal and economic disadvantages, the abilities of women played a surprisingly equal role in both marriage matching and the passing of economic status to the next generation.

1 Introduction

Before the late 20th century, married women had limited economic roles outside of the household. During this long era, it is easy to assume that marital assortment mattered less for economic inequality; if only men held formal employment, who they married was of little importance for the distribution of income. However, this assumption overlooks an important channel through which marital assortment — and women — could have mattered: intergenerational mobility. As I show in a simple model, if marriages were matches between individuals of similar characteristics, and the ability of mothers mattered for the outcomes of her children, then the degree of assortment would matter for social mobility. Then, if the degree of assortment was high for a long period of time, there would be lower social mobility, contributing to higher overall economic inequality. This paper demonstrates that all these conditions were true in Quebec 1800–1970.

To determine if marital assortment played a significant role in long-run inequality, this paper sets out to answer three related questions. First, how did the degree of marital assortment change over the long run? Second, were marriages formed between individuals of similar economic abilities? Third, did the human capital of mothers matter for the outcome of their children? Using a sample of linked families from Quebec 1800–1970 to answer these questions, this paper finds that marriages were as assortative in 1800 as in the mid 20th century. Moreover, even though they may not have been formally employed, the ability of women determined the husband they matched to and the outcomes of their children. Even in as conservative a society as Quebec 1800–1970, I find that women played a comparable role to that of men in intergenerational mobility. As a consequence, assortative marriage mattered for intergenerational mobility and therefore influenced economic inequality long before the 20th century.

This paper is only possible due to some very unusual features of the dataset. First, as married women kept their family names, they are linked to their parents. Second, the dataset contains a measure of female human capital, albeit a fairly rudimentary one: the ability to sign one's marriage certificate. Third, it contains several million individuals, allowing me to use empirical methods that require large sample sizes such as family-specific fixed effects. Finally, it covers an extensive period of time, allowing me to comment on the longer-run trends. While no historical data source is without any limitations, this rich set of features makes Quebec 1800–1970 a uniquely suitable setting to investigate this set of questions.

To explain the interactions between intergenerational mobility, marriage, and inequality, I first propose a latent variable model. Using this model, I suggest a new method to measure of the degree of assortment. This measure works around the lack of a reported occupation for the woman by comparing the occupational status of her husband both to her father and to his father. This measure shows that the degree of assortment was high from 1800 through the mid 20th century, decreasing somewhat after 1920. Moreover, while the model is simple, it has strong predictions. For assortment to matter for inequality and mobility, marriages must match on ability and both parents must play a role in the outcomes of their children. Using this framework, I set out to test these criteria.

The first test uses family-fixed effects to determine if matching was on family or individual characteristics. I find that a woman who was able to write her name could have expected to marry a higher status husband and have higher status children than her sisters who could not. The same is true for a man and his brothers. Notably, the estimates of the effect of human capital is symmetric across the gender of both parents and children; the ability of women seems to matter as much as men.

The second test tries to determine if the ability of mothers matters directly for their

children. The limited information I have on a mother's human capital is correlated with her children to a comparable degree as the ability of her husband. However, this could also be the case if the mother didn't directly matter but instead provided additional information about the underlying economic status of her husband. To address this empirical challenge, I use the fact that re-marriages were common in Quebec to control for the father, both through fixed effects and by directly comparing half-siblings. I find evidence that the ability of mothers did matter, and that it did to a comparable degree as fathers.

The consequences of these findings are twofold. First, they closely fit the predictions of the simple model. This strongly implies that assortment contributed to the overall distributions of ability and socioeconomic status in the province. Moreover, there was remarkably symmetrical effects across the gender of both the parents and children. Despite severe legal and economic disadvantages, in this at least women appear to have had an equal role. Second, the results have implications for standard estimates of intergenerational mobility. I show that the often overlooked influence of both assortative marriage and the ability of women directly influence the magnitude of father-son intergenerational elasticities and rates of multigenerational mobility. An observed change in father-son or grandfather-grandson mobility could be entirely driven by assortment or the influence of mothers.

This paper thus contributes directly to two literatures. First, the paper provides longer-run context to more contemporary studies of assortative marriage. In many societies, the late 20th century saw a rise in female employment rates and the closing of the gender gap in educational attainment (Goldin 2006). This revolution has led to widespread concern that an increase in both assortative marriage and two-income households would lead to greater inequality. The empirical evidence, however, suggests that it had little effect on ether the degree of assortment or on contemporaneous inequality (Eika et al. 2019, Hryshko et al. 2017). This paper provides a longer-run context that helps explain why the changes in the late 20th century, while so revolutionary elsewhere, did not seem to matter for assortment and inequality. First, inequality was as high, if not higher, in the 19th century as it was in the 1960's. Second, assortment had always mattered for inequality through intergenerational mobility, albeit over a longer time frame through intergenerational mobility.

Second, it adds to our understanding of intergenerational mobility over the long run. Until recently, studies of intergenerational mobility have traditionally overlooked women (Black and Devereux 2011). This oversight has been addressed by considering the mobility of daughters (Chadwick and Solon 2002). This approach typically considers the mobility in the incomes of households, not of individuals, consequently acknowledging the importance of assortative marriage. Recently, it has been extended to historical datasets, an impressive feat given the difficulty of linking married women to their parents (Olivetti and Paserman

2015, Craig et al. 2020). However, an additional challenge in historical studies is the lack of direct measures of household income. Instead, occupational status measures are used. As only men held formal occupations, the association between the occupational status of the husband and the father is used as a proxy for the association between the household incomes. While this is a reasonable way to estimate the association between daughters and fathers, this paper goes farther by considering how mothers and marriage determine the mobility of children of both genders.

The structure of the paper is as follows. First I outline a simple model of intergenerational mobility, marriage, and inequality. Second, I provide a brief summary of the history of the legal and economic situation of women in Quebec. Third, I describe the dataset, how families were linked, and the measures I use of economic status and human capital. Fourth, I present a series of empirical results supporting the paper's main findings. Fifth, I discuss the broader implications of these findings. Finally, I conclude, summarizing how all the evidence fits into a broader picture: assortment mattered for inequality long before the late 20th century because women had always played an important role in marriage and mobility.

2 A simple model of marriage, mobility, and inequality

I develop a model of how marriage and intergenerational mobility contribute to inequality over the long-run. This framework, while simple, motivates the empirical work that follows. The two main findings — that assortment was on individual characteristics and that the ability of mothers mattered for child outcomes — individually answer rather narrow questions; together, as illustrated by the model, they show that assortment could have mattered a great deal for inequality if it were high enough and stable enough over a long period of time. The model also suggests a new method to measure the degree of marital assortment, which I then use to show that assortment was indeed high and stable for most of the period studied.

2.1 Setting up the model

Consider the following simple model. We are interested in a specific measure of human capital or socioeconomic status, call it x. In the intergenerational mobility literature, the measure is typically lifetime earnings, but there are many other potentially interesting variables as well.

Following Solon (1992) and Clark and Cummins (2015), assume we only have an imperfect measure or proxy y for x. The classic example is proxying for lifetime earnings with the

earnings in a specific year, but again there could be many potential reasons why an observed measure is imperfect. This introduces classical measurement error:

$$y_i = x_i + u_i \tag{1}$$

for all individuals i, where u_i is an error term uncorrelated with x_i.

Then assume that the measure of child c, x_c, is inherited depending on the measures the child's father x_f and mother x_m.

$$x_c = \beta_f x_f + \beta_m x_m + e_c \tag{2}$$

where e_c is a random term uncorrelated with the x's. For now, assume that the effect on children is the same regardless of gender. While this seems a strong assumption, it makes the model much more tractable, and I will later provide evidence that it appears reasonable in my context. Moreover, note that u_c and e_c can be allowed to vary by the gender of the child to model the likely scenario where there is a gender gap. In Appendix A1 (under construction), I show some numerical simulations that relax this assumption.

Following Chadwick and Solon (2002), assume that the assortment on the underlying variable can be summarized by:

$$corr(x_f, x_m) = \gamma \tag{3}$$

2.2 Inequality

To summarize inequality in a given generation, consider the variance of the variable of interest:

$$\sigma^2_{x_c} = \beta^2_f \sigma^2_{x_f} + \beta^2_m \sigma^2_{x_m} + 2\beta_f \beta_m (\gamma \sigma_{x_f} \sigma_{x_m}) + \sigma^2_{e_c} \tag{4}$$

Now define a steady-state equilibrium as when there is no change in inequality from generation to generation:

$$\sigma^2_{x_c} = \sigma^2_{x_f} = \sigma^2_{x_m} \tag{5}$$

Then:

$$\sigma^2_{x_c} = \frac{\sigma^2_{e_c}}{1 - \beta^2_f - \beta^2_m - 2\gamma\beta_f\beta_m} \tag{6}$$

In terms of observed inequality, as the error term u_c is assumed to be independent of x_c:

$$\sigma^2_{y_c} = \sigma^2_{x_c} + \sigma^2_{u_c} \tag{7}$$

Unsurprisingly, the more children take after their parents (i.e. the higher the β_f and β_m), the higher steady state inequality will be. Further, if both β_f and β_m are greater than zero, the degree of assortment γ will increase steady state inequality as well. (Figure 2.2).

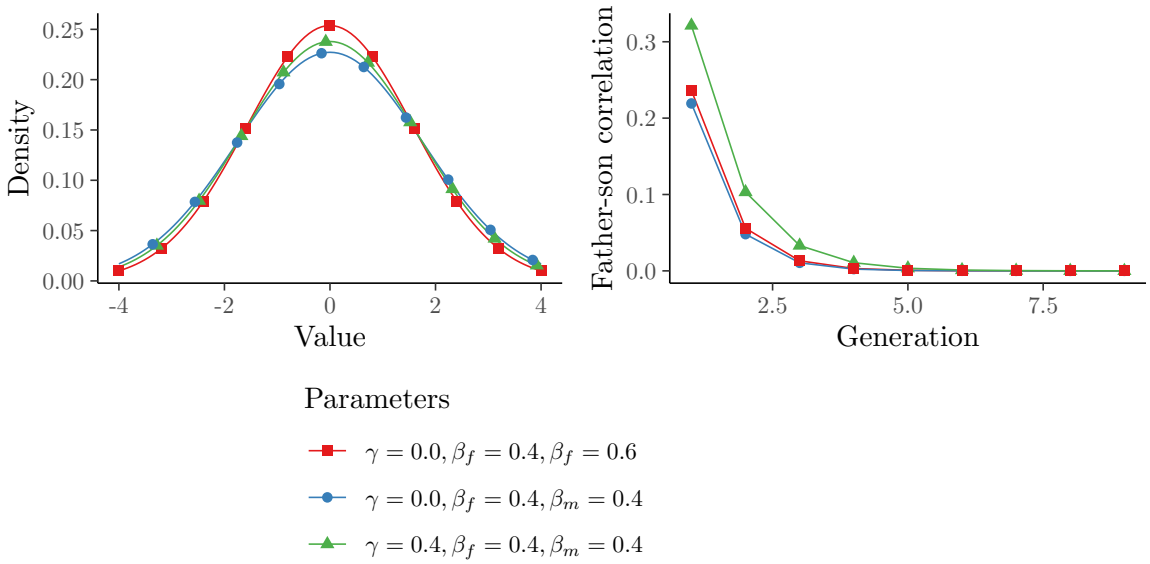


Figure 2.2: Steady state inequality and regression to the mean
Note: See text for the full model. γ is the degree of assortment, β_f and β_m are how associated children are with their fathers and mothers. I assign e_c, the random component of intergenerational mobility, and u_c the classical measurement error term, a variance of one in all simulations. As shown by the simulations, a higher γ and a higher β_m both increase inequality and decrease the speed at which a family line reverts to the mean. As the observed data is measured with error and the size of the resulting bias is inversely related to β_m, increasing β_m when $\gamma = 0$ does slow the observed speed of mean reversion slightly; this would not be true if the plotted data were the true underlying variables.

2.3 Intergenerational mobility

In terms of social mobility, now note that due to the variances being equal:

$$x_i = \gamma x_s + \mu_i \tag{8}$$

where s is i's spouse and μ_i is an uncorrelated error term.

If I substitute this into the intergenerational mobility equation, I get:

$$x_c = (\beta_f + \gamma\beta_m)x_f + \beta_m\mu_f + e_c \tag{9}$$

This can be estimated with a regression:

$$x_c = \alpha_0 + \alpha_1 x_f + \epsilon_c \tag{10}$$

where $\alpha_1 = (\beta_f + \gamma\beta_m)$.

However, if I estimate the regression using the observed variables y_c and y_f, as y_f is correlated with u_f the estimate is attenuated down. Specifically, as:

$$y_c = (\beta_f + \gamma\beta_m)y_f - (\beta_f + \gamma\beta_m)u_f + \beta_m\mu_f + e_c - u_c \tag{11}$$

there is bias of the form:

$$plim\ \hat{\alpha}_1 = (\beta_f + \gamma\beta_m)\frac{\sigma^2_{y_f}}{\sigma^2_{y_f} + \sigma^2_{u_f}} \tag{12}$$

If β_f and γ are greater than zero, then mothers contribute to the observed correlation of y between fathers and sons. While β_m matters for inequality if $\gamma\ = 0$, it does not contribute to the correlation of y between fathers and sons. However, as the attenuation bias depends on σ_{y_f} which is increasing in β_m, the attenuation bias is lower the higher β_m. (Figure 2.2).

2.4 Measuring assortment if women are not directly observed

I can construct an estimate of the degree of assortment by comparing the association between father-in-laws and son-in-laws to the association between sons and fathers. Letting y_{fl} be the value observed for the father-in-law of c:

$$y_c = \gamma(\beta_f + \gamma\beta_m)y_{fl} - \gamma(\beta_f + \gamma\beta_m)u_{fl} + \gamma\beta_m\mu_{fl} + \gamma e_{i,s} + \mu_i - u_i \tag{13}$$

and

$$y_i = (\beta_f + \gamma\beta_m)y_f - (\beta_f + \gamma\beta_m)u_f + \beta_m\mu_f + e_i - u_i \tag{14}$$

regressing y_i on y_{fl} and on y_f, the ratio of the coefficients has the probability limit of:

$$\gamma\frac{\sigma^2_{y_{fl}}(\sigma^2_{y_f} + \sigma^2_{u_f})}{\sigma^2_{y_f}(\sigma^2_{y_{fl}} + \sigma^2_{u_{fl}})} \tag{15}$$

which should be equal to γ if the distribution of y_f is the same as that of y_{fl}.

3 Historical context

Quebec before its Quiet Revolution of the 1960's was a more conservative and less secular society than its North American neighbors.[1] Public education was highly influenced by the Catholic Church, and conservative norms about gender roles informed public policy. For example, while Québécois women could vote in federal elections after 1918, they did could not vote in local elections until 1940 (Tremblay and Roth 2010).[2]

3.1 The legal rights of women

When it came to property rights, women were systematically disadvantaged, albeit in a different way than their Anglophone neighbors. While Quebec was ceded to the British in 1763, laws pertaining to civil matters remained governed by the Custom of Paris, a codified system of customary French law from the Paris region. Under the Custom of Paris, and unlike in English-speaking legal traditions, married couples formed a legal entity called the *communauté de biens*, or community of property, in which both partners theoretically had equal stakes (Greer 1997). As a consequence, both the husband and wife were required to sign legal documents (which, as mentioned below, greatly aids the linking of vital records), though the husband was expected to have full control of the management of the joint property. Worse still, married women were legally considered incapable, being unable to sign contracts or initiate a lawsuit (Baillargeon 2014). The Civil Code of Lower Canada, introduced in 1866, only clarified the legal disadvantages of women. Only after reforms, starting in 1964 when women were no longer legally incapable (Baillargeon 2014).

In terms of inheritance, the community of property was dissolved by giving the surviving spouse their half and dividing the rest equally amongst the children regardless of gender.[3] Perhaps as a consequence of being unable to write children out of a will, parents had little legal recourse to block a match they disapproved of after the children reached a certain age (Greer 1997). However, it gradually became the norm to "gift" property to favored heirs, albeit still subject to some restrictions designed to protect all potential heirs (Greer 1985).

[1] Confusingly, the term quiet revolution is also used for the increase in female labor force participation and educational achievement of women in many countries starting in the 1970's (Goldin 2006).

[2] Oddly, from the first elections in 1792 until 1849, suffrage was only restricted to individuals meeting age and property requirements; a very small number of women who independently owned property could and did vote based on these criteria.

[3] One of the few advantages women had was that wives could renounce the debts of the community of property on her husbands deaths as it was assumed she was not responsible for their accumulation.

3.2 The demographic regime

Was marriage the typical experience for women? Quebec had an extreme version of the European marriage pattern, with younger marriages and less frequent celibacy than France (Greer 1997). Quick re-marriage upon the death of a spouse was common. Moreover, marriage in Quebec usually meant having a large family. While this variant European marriage pattern was common in European settler colonies, Quebec kept it longer than most. The demographic transition occurred later than in the rest of North America, only reaching substantial numbers of French-speaking Québécois by the 1920's (Vézina et al. 2014). Moreover, from first settlement through at least 1835, there appears to have been no attempt of parents to target a specific number of children.(Clark et al. 2020.

3.3 Rising levels of education

[COMING SOON]

3.4 Women and the workforce

While the economy of Quebec evolved dramatically from 1800–1970, women had limited opportunities in the formal labor market for most of the period. In the first half ot he 19th century, most families engaged in household industry, so the wife would usually contribute labor to the occupation reported by her father or husband as well as to domestic tasks (Baillargeon 2014). In urban areas, formal employment was available to unmarried women as domestic servants. Women could also be educators, first as nuns and and later as secular teachers. As the economy began to industrialize in the 1840's, unmarried women could also find work in factories (typically clothing or tobacco), albeit with substantially lower wages than men. As elsewhere, industrialization also led to the decline of household production and the rise of the male breadwinner household (de Vries 2008). The late 19th century and early 20th century saw the rise of female dominated occupations such as telephone operators, typists, and secular nurses, though married women were still expected not to have a formal occupation until the 1970's.

3.5 Quebec, North America, and external validity

Overall, how much was Quebec an outlier? In general, Quebec was worse-off than the rest of North America but followed the same trends. For example, it had lower wages until the late 20th century but followed a similar trend over time (Albouy 2008, Geloso and Lindert 2020). Its more conservative and less secular society delayed the extension of certain rights

to women, but not indefinitely. Its demographic regime was extreme, with large family sizes and a delayed demographic transition, but it was merely an extreme version of the typical European marriage pattern. The role of women in the labor force evolved roughly the same as the rest of North America (Goldin 2006). I argue, therefore, that my findings from Quebec are likely generalizable to the rest of North America. If women and assortment mattered for mobility and inequality and Quebec, they likely mattered as well in the more liberal neighboring regions.

4 Data

[BRIEF SUMMARY OF SECTION HERE SOON]

4.1 Linked family vital records

The IMPQ is a new database of family reconstitutions from baptism, burial, and marriage records (IMPQ 2020). It integrates two previous databases, the BALSAC database and the RPQA (Project Balsac 2020, PRDH 2020). While it contains data as far back as the founding of the colony, in this paper I consider data from a subset with frequently reported occupations, 1800–1969. While the dataset is still being extended, as of writing it contains 1.4 million unique births, 0.6 million unique deaths, and 4 million unique married individuals from 1800–1969 (though births and deaths are limited to a particular sub-region after 1849). Moreover, in those records a total of 2.7 unique individuals are mentioned besides the main participants, providing additional observations over time for many people besides their own vital events.

Two unusual institutional features of Quebec have resulted in vital records that are particularly easy to link. First, due to the system of community property, both husbands and wives signed their names on all legal family records. Second, women kept their family names when they married. This means that women can be linked to their fathers and that most vital records have four names to match on (the first names and last names of both the husband and wife or mother and father).

As the IMPQ combines two previous databases, it was constructed using two different yet very similar linking methods. For intergenerational linkages (marriages to marriages), the IMPQ uses the SOREP system from the BALSAC database (Vézina et al. 2013). For intragenerational linkages (births and deaths to marriages), the PRHD's matching algorithm is used (Dillon et al. 2018). These two procedures differ slightly (e.g. the BALSAC standardizes names using the FONEM phonetic algorithm (Bouchard et al. 1981) and the PRHD uses a custom-made name dictionary), but they both use the four names to match records and

are linked manually when there are no unique matches. Manual linkages are not necessarily better than automatic linkages; in some applications they produce both higher true matches and false positives (Abramitzky et al. 2019). However, the fact that the Quebec vital records have four names to match on should increase the accuracy of matching regardless of method used. Moreover, the parish records of Quebec have survived remarkably intact as local priests were required to send duplicates of all records to their superiors, so records of very close to the complete population survive; this will reduce false positive rates in an analogous way to full count to full count census linkages (Dillon et al. 2018, Abramitzky et al. 2019).

4.2 Measures of human capital and occupational status

The direct measure of human capital I use in this paper is the presence of a signature on a marriage record. Catholic churches had long required required both the bride and the groom to sign their marriage records if able and often the priest recorded if they complied. I code a signature indicator variable as one if the individual definitely signed their marriage record, and zero if they are recorded as not signing or there is no record of a signature. This is necessary as for some years see no major change in the reported number of participants who signed but a major decrease in the reported number who did not sign. As shown in Figure 1 below, this definition seems to produce a trend that is smooth and close to external estimates. Was a signature really a measure of human capital, that is a productive attribute? The qualitative evidence suggests that it was. Signatures are a proxy for the ability to write, a form of human capital that had always been particularly associated with business activity in Quebec (Greer 1997).[4]

The vital records also often state the occupations of participants or relatives. I assign each individual the occupation listed at their first marriage, or if there is none the one closest chronologically to their first marriage. The occupations are assigned HISCO codes, a classification system designed for comparative studies of historical social mobility (Van Leeuwen et al. 2004). I then assign various occupational status scores to these HISCO codes aggregated to the three digit level. The primary score I construct is the average yearly earnings reported by individuals with that occupation in Quebec in a 5% sample of the 1901 Canadian Census (Canadian Families Project 2002 and Minnesota Population Center 2019). This requires crosswalking occupations from IPUMS's occupational codes to the original HISCO scheme (Zijdeman 2014). There are numerous other ways to rank occupations, as discussed in Appendix A3. However, this ranking is easy to interpret, is at least a proxy for the stan-

[4]As for reading, it too was likely associated with economic in Quebec. As opposed to, say, their majority Protestant neighbors in New England who prioritized literacy education for religious reasons, for Quebec's Catholics reading the Bible was not a religious necessity.

dard variable of interest in mobility studies (lifetime earnings), and produces similar results to the other ranking systems. Therefore, the main results in this paper use the 1901 imputed earnings as the primary measure of occupational status.

4.3 Women's presence in the vital records

Do the vital records record the occupational status and human capital of women? Four extracts of Canadian censuses (the 100% 1881 sample, the 5% 1891 sample, the 5% 1901 sample and 1901 oversample, and the 5% 1911 sample) and data compiled by Clarence D. Long provide external points of comparison (Dillon et al. 2008, Inwood and Jack 2011, Canadian Families Project 2002, Gaffield et al. 2009, Minnesota Population Center 2019, Long 1958, Killingsworth and Heckman 1986).

Figure 1 compares the rate individuals signed their first marriage record to the rate of self-reported ability to write in the censuses. Unlike the censuses, individuals only show up in the marriage records during a specific time in their lives. To account for this, I restrict the census data to married individuals aged 16–35 to approximate the pool of individuals who are getting married for the first time. As shown in the figure, my proxy for the ability to write is somewhat lower than the census rates, but it follows a similar pattern.

In contrast, the vital records do a poor job of recording female occupations. Figure 2 computes the fraction of all individuals with an occupation mentioned in any record who are female by the decade they were first mentioned. Compared to the other sources, the vital records underestimate the number of married women with occupations and almost entirely omit unmarried women with occupations. Here, I restrict the data to ages 16 and above as occupations can be listed at other events besides marriage (e.g. death, marriage of a child, etc.) It appears that vital records omit the occupations of women, especially unmarried women. While the data for occupations are less reliable, some patterns are clear. By the late 19th century, unmarried women often worked outside of the home, but married women did not in substantial numbers until the 1940's (though the data here also disagree by how large the rise was).

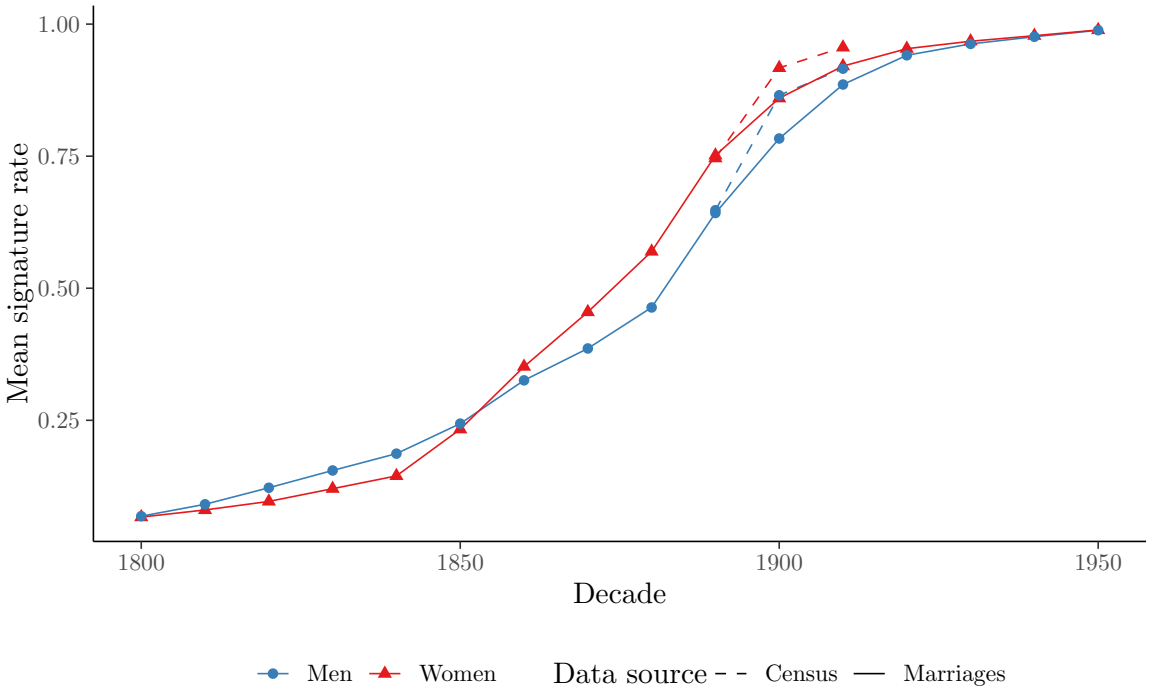


Figure 1: The vital records accurately report the ability to write

Note: The vital record signature variable is the average of an indicator that is one if we observe the individual sign their first marriage certificate, zero otherwise. The census record data is the percent of individuals aged 16–35 (roughly the age range for first marriages) who are reported by enumerators as unable to write. While the data do not exactly agree, the trends and relative positions of men and women are the same. There was actually a gender gap in favor of women from 1850 until the ability to write was near-universal c. 1930. Overall, Quebec went from a very low human capital society to a high human capital society from 1800–1960.

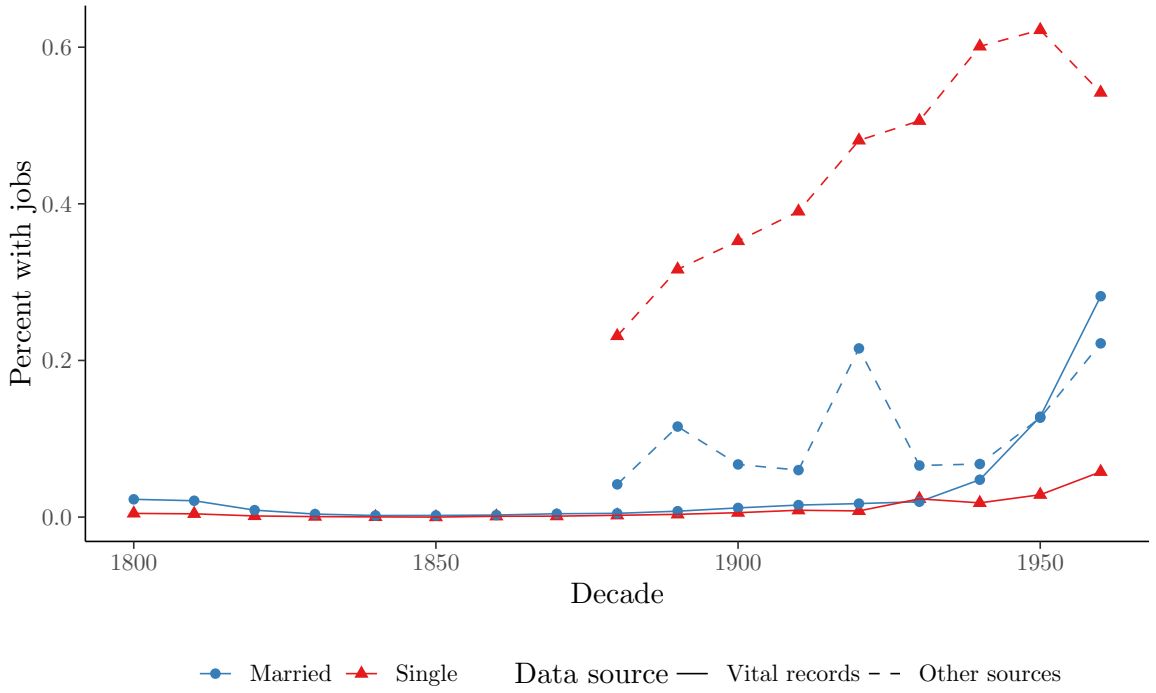


Figure 2: The vital records fail to record female employment
Note: The vital record data are constructed by assigning each individual one if they are observed at any time with a specific occupation and zero otherwise, assigning each individual to a decade based on the median date they are observed at, and then averaging over decades. The other sources data come from two sources. The first are the census samples for 1881, 1891, 1901, and 1911, with which I compute the share of individuals aged 15–64 who have a specific occupation in the given decade. The second are decade averages compiled in Long (1958). Here, I have to make one ad-hoc adjustment; Long reports data by single, currently married, and widowed or divorced. To combine the two categories of individuals who were ever married, I take a weighted average. The weights are the shares in the censuses of married individuals who were currently married versus widowed or divorced, assuming the rates are fairly stable over time. The two different data sources are as close to comparable as possible, yet likely compare different populations for a given decade due to selection into the vital records at certain ages. Still, there are three notable observations. First, the vital records never adequately report employment of unmarried women. Second, before 1930 the marriage records also under-report the employment of married women, though not as badly. Third, starting in 1930 the employment rate of married women began to rapidly rise.

5 The role of women in assortative marriage

[BRIEF SUMMARY OF SECTION HERE SOON]

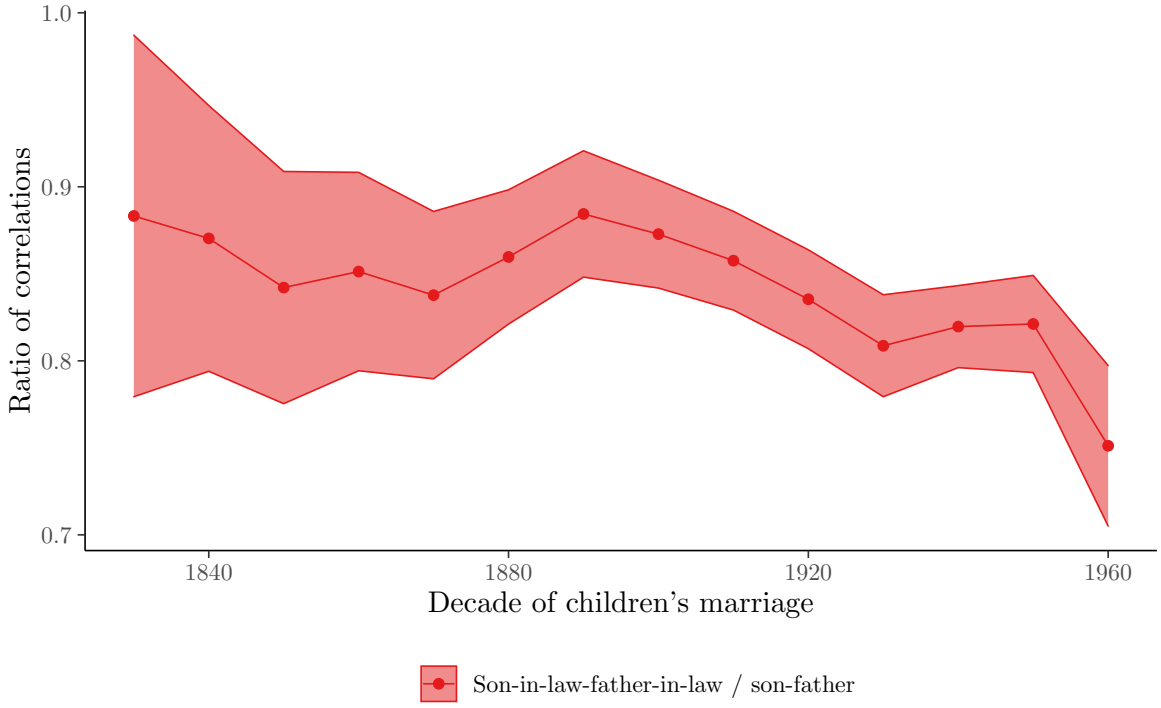
5.1 Measuring the assortment

Typically, assortment is measured by the correlation between the education levels of spouses. However, as shown in Figure 1, the average signature rate changed dramatically during this interval. As the average education level increased, the average human capital of an individual who met the relatively low bar of signing their name would have also increased. Therefore, the relationship between the underlying degree of assortment on human capital and the observed degree of assortment on signature rates is likely changing over time.

A more stable measure of ability is an individual's occupational status. However, in most of the sample married women have no observed occupational status (and in the mid 20th century when they start to in large numbers, they are concentrated in lower status occupations). As an alternative, a measure can be constructed by comparing the correlation of the occupational status of sons-in-law and fathers-in-law to that of sons and fathers. The former are two degrees separate: an intergenerational link from father-in-law to daughter and a marriage link from daughter to son-in-law. The latter has only one degree of separation: an intergenerational link from father to son. In the simple model of intergenerational mobility I develop above, because I assume assortment is simply between the daughter and the son-in-law, the ratio is equivalent to the degree of assortment. Relaxing the model's assumptions, the ratio will still hopefully control for trends in intergenerational mobility and leave only the trends in assortment.[5]

[5]While the model does assume that marriages are matches based solely on the characteristics of the wife and husband, it can be extended to allow the wife's father to play a role in matching. As Appendix A2 shows, under less strict assumptions, bounds for the degree of assortment can still be estimated and they still appear to be fairly high and stable over time.

Figure 3: Marriage assortment on imputed earnings



Note: 95% confidence interval shaded. Standard errors are bootstrapped with 50,000 repetitions. Earning scores are the average earnings for the individual's occupation in Quebec based on a 5% extract of the 1901 Canadian Census. The ratio is the ratio of two rank-rank regression coefficients (Chetty et al. 2014) (As ties are not broken randomly, the rank variables do not have exactly equal variances and therefore the coefficients are not exactly the Pearson correlation coefficients. Chetty et al. also makes this approximation, but it is larger for more granular measures than income such as occupational status.) In the simple model described above, this ratio is a close estimate of the true degree of assortment. More generally, the father-in-law and son-in-law correlation is affected by assortment, by intergenerational mobility, and by attenuation bias. The father and son correlation is affected only by intergenerational mobility and attenuation bias, so the ratio should hopefully capture only assortment.

5.2 Marriages were assortative on individual characteristics

Were marriages matches based on the characteristics of the spouses? One could imagine a society in which this was not the case. For example, perhaps, a society where marriages were negotiated to form a political or business alliance between a husband and his father-in-law, with the characteristics of the wife an afterthought at best. In this hypothetical society, there could still be an observed correlation between a given characteristic of the two spouses if the wife's characteristic is determined at least partially by her father.

To test if individual characteristics mattered conditional on the families, consider the following fixed-effects regression:

$$y_s = \alpha y_i + \phi_F + \epsilon_{i,F} \tag{16}$$

Where y_i is a characteristic of individual i of family F, i_s is a characteristic of spouse s of individual i, ϕ_F is a fixed effect, and $\epsilon_{i,F}$ is a family-clustered error term. In other words, compared to their siblings, does an individual with higher ability match a spouse of higher ability? If so, α will be positive.

As shown in Table 1 below, a woman who signed her marriage record married a man with higher human capital and status than her sisters. Being able to sign her name was associated with an increase in the probability her husband signed by 38%, an increase in her husband's occupational earnings score by 4%, and an increase in her father-in-law's occupational earnings score of 2%. This is evidence that marriage matches took into account individual characteristics and that human capital provided a premium for women seeking a husband. Note that while the family fixed-effect does reduce $\hat{\alpha}$, this does not reveal the degree to which matches are coordinated by families; if matching is only on individual characteristics, the family fixed-effect will still reduce $\hat{\alpha}$ as long as the human capital of sisters is correlated.

What about men and their brothers? As shown in Table 1 below, men who were able to write appear also to have had a premium when matching a spouse. Being able to sign his name was associated with an increase in the probability that his wife signed by 35% and an increase in his father-in-law's occupational earnings score by 3%. These estimates are remarkably similar to the returns to the wife's ability to sign. The returns to human capital in terms of marriage matches appear to be the same regardless of gender.

What is the significance of the result that marriages matches take into account the individual characteristics of women, not just those of her family? Assortative marriage could still matter for inequality, after all, if the woman's family directly influenced the abilities of her children. First, the results show that there was a return to education for women in terms of the economic status of the household she formed at marriage. This was the case even if she did not employ her human capital in a formal occupation. Second, it implies a stronger role for assortative marriage in intergenerational mobility, assuming mothers mattered directly for their children's outcomes. Finally, it at least hints that women may have had some agency over the marriage matching process.

Table 1: Marriage matches were determined by individual characteristics

	Dependent variable: Spouse's characteristic					
	Signed		Log imputed earnings		Father's log imputed earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of wife's human capital						
Wife signed	0.53***	0.38***	0.17***	0.04***	0.07***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Wife's family FE		X		X		X
Decade FE	X	X	X	X		X
Observations	1,994,133	1,994,133	1,161,203	1,161,203	989,747	989,747
Adjusted R^2	0.60	0.63	0.05	0.38	0.03	0.32
Panel B: Effect of husband's human capital						
Husband signed	0.45***	0.35***			0.11***	0.03***
	(0.00)	(0.00)			(0.00)	(0.00)
Husband's family FE		X				X
Decade FE	X	X			X	X
Observations	1,983,328	1,983,328			1,001,536	1,001,536
Adjusted R^2	0.62	0.63			0.03	0.33

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. The sample excludes individuals with one or more unknown parents. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time. Note that, after adding family fixed effects, the effects are close to symmetrical across gender.

Appendices A4 and A5 show that these estimates are robust to two potential concerns. First, although some records are still missing signatures in the mid 20th century when literacy was near universal, the results are robust when restricted to just the 19th century. Second, although families with one literate and one illiterate child are fairly unusual, they results are robust to controlling for selection into identification

6 The role of women in intergenerational mobility

[BRIEF SUMMARY OF SECTION HERE SOON]

6.1 Controlling for the father with fixed effects

Did the abilities of mothers directly influence the outcomes of their children? Here, I am interested in if the ability of the mother directly determines the outcomes of the children;

I am not interested in the exact mechanism by which it does so (such as, for example, the distinction between human capital and innate ability (Black et al. 2005)). All that is required for women to play a role in intergenerational mobility is for their abilities to matter as well as the abilities of their husbands.

The limited measure of ability available for mothers, the presence of a signature at marriage, is in fact correlated with that of her children even after controlling for the signature of the father (Table 2 Panel A). However, this pattern could still be observed if the mother did not directly matter for the outcomes of the children as long as marriages were assortative on signatures. With assortment, if the husband's ability is observed with measurement error, the mother's ability would be correlated with the residual even if its true effect is zero. Therefore, the simple regressions in Table 2 Panel A do not identify the effect of the ability of mothers. One notable pattern here is that mothers seem to matter more for the ability to write and fathers matter more for occupational status. However, there appears to be no strong difference between the association with the outcomes of children of different genders.

To identify the effect, ideally we would control for the father but randomize the mother. A less ideal approach, though one possible with the data, is to consider the case where a father has children from more than one marriage. However, this results in two complications. The first is the chance that the children as scarred by whatever event resulted in a second marriage (such as a death or divorce). This can be controlled for by including fixed effects for which marriage number the children are from. Second, as marriage is assortative on the ability of mothers, the abilities of each wife of the father will be correlated. Therefore, similar to the family fixed effects above, the father fixed effect will absorb part of the effect of the mother's ability.

I regress:

$$y_c = \alpha y_m + \phi_f + \epsilon \tag{17}$$

where y_c is an outcome of a child, y_m is a characteristic of the mother, and ϕ_f is a fixed effect for the father.

As shown in Table 2 Panel B, even controlling for the father, a mother who could sign her name had children who were 2% more likely to be able to sign their names. The evidence for occupational status is not significant at the 5% level, but is at least suggestive of a benefit.

For comparison, as shown in Table 2 Panel C, estimates the effects of the ability of a father controlling for the mother. Notably, the results are very similar to those of the regressions for mothers. Once the correlation between the ability of spouses is accounted for through fixed effects, the effect of parental human capital appears to be symmetrical across gender.

Table 2: The effect of parental human capital on child outcomes, fixed effects

	Dependent variable:			
	Signed Daughter	Signed Son	Husband's log earnings score Daughter	Log earning score Son
	(1)	(2)	(3)	(4)
Panel A: Both parent's signatures				
Mother signed	0.11***	0.10***	0.05***	0.05***
	(0.001)	(0.001)	(0.002)	(0.002)
Father signed	0.06***	0.08***	0.13***	0.14***
	(0.001)	(0.001)	(0.002)	(0.002)
Decade FEs	X	X	X	X
Observations	1,772,249	1,658,043	1,047,263	986,679
Adjusted R^2	0.53	0.51	0.06	0.06
Panel B: Controlling for father				
Mother signed	0.02***	0.02***	0.01	0.02*
	(0.005)	(0.01)	(0.01)	(0.01)
Father FEs	X	X	X	X
Marriage number FEs	X	X	X	X
Decade FEs	X	X	X	X
Observations	1,772,249	1,658,043	1,047,263	951,557
Adjusted R^2	0.69	0.68	0.38	0.42
Panel B: Controlling for mother				
Father signed	0.02***	0.02***	0.02	0.04**
	(0.01)	(0.01)	(0.02)	(0.02)
Mother FEs	X	X	X	X
Marriage number FEs	X	X	X	X
Decade FEs	X	X	X	X
Observations	1,772,249	1,658,043	1,047,263	951,557
Adjusted R^2	0.70	0.68	0.38	0.42

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. The sample excludes individuals with one or more unknown parents. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.

6.2 Directly comparing half-siblings

One downside of the father fixed effects approach is that it relies on observing measure of the ability of the mother. Fortunately, there is another test using parents with more than one marriage that only relies on the characteristics of the children. Consider a pair of children who could be either half-siblings or full siblings. If they share both a mother and a father and the abilities of mothers matter directly, their outcomes should be more correlated than that of a pair that share only a father. Again, there is a concern that the event resulted in a second marriage could have harmed the children of the first marriage; again, fixed effects can control for it.

I estimate the regression:

$$y_{i,m_1,f} = \beta Y_{j,m_2,f} \times I(m_1 = m_2) + \epsilon_f \tag{18}$$

Where $y_{i,m_x,f}$ is a characteristic of child i with father f and mother m_x, $i > j$, $I(m_1 = m_2)$ is an indicator that is one if the children share a mother, and ϵ_f is an error term. The non-interacted indicator $I(m_1 = m_2)$ ought to capture any effect of the even that ended the first marriage; as $i > j$ the child on the left-hand side of the equation is always from a previous marriage. As there are sometimes more than two marriages, I also add an additional fixed effect for the father's marriage number at the time of the older child's birth.

The results are shown in Table 3 below. Children who share both parents are more strongly associated with each other than half siblings. The effects might be very slightly stronger for sharing a mother on sisters and for sharing a father on brothers and very slightly larger for fathers than for mothers, but overall they are very similar regardless of the gender of the children or the parents.

Note that the correlation between half sisters is 89% of the correlation of full sisters for those that share a mother and 85% for those that share a father. For half brothers, the percentages are 89% and 81% respectively. These percentages are very closed to the observed degree of assortment in Figure 3. This is not a coincidence, rather what the simple model would predict. [SIMPLE SUMMARY OF MATH INTUITION HERE SOON].

Table 3: The effect of parental human capital on child outcomes, half versus full siblings

	Dependent variable: Younger sibling's characteristic			
	Signed (Daughters)	Signed (Sons)	Husband's log earnings score (Daughters)	Log earning score (Sons)
	(1)	(2)	(3)	(4)
Panel A: Controlling for father				
Older sibling's characteristic	0.34***	0.32***	0.22***	0.26***
	(0.00)	(0.00)	(0.01)	(0.01)
Same mother	−0.00	−0.01**	−0.30***	−0.30***
	(0.00)	(0.00)	(0.05)	(0.05)
Signed × same mother	0.05***	0.04***	0.05***	0.05***
	(0.00)	(0.00)	(0.01)	(0.01)
Order FEs	X	X	X	X
Decade FEs	X	X	X	X
Observations	2,112,006	1,982,293	839,388	756,645
Adjusted R^2	0.60	0.62	0.11	0.14
Panel B: Controlling for mother				
Older sibling's characteristic	0.33***	0.30***	0.24***	0.22***
	(0.01)	(0.01)	(0.02)	(0.02)
Same father	−0.03***	−0.03***	−0.16	−0.53***
	(0.01)	(0.01)	(0.10)	(0.11)
Signed × same father	0.06***	0.07***	0.03*	0.09***
	(0.01)	(0.01)	(0.02)	(0.02)
Order FEs	X	X	X	X
Decade FEs	X	X	X	X
Observations	2,024,585	1,897,372	806,656	727,123
Adjusted R^2	0.60	0.62	0.11	0.14

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. The sample excludes individuals with one or more unknown parents. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.

7 Discussion

7.1 Assortment matters for estimates of father-son correlations

If women directly matter for the outcomes of their children and marriages are assortative, the correlation between characteristics of fathers and sons will be partially determined by the mother. As shown in Section 2 Equation 9, the observed correlation between the characteristic of fathers and sons in the simple model is $(\beta_m + \alpha\beta_f)$:

$$y_c = (\beta_m + \gamma\beta_f)y_f + \epsilon \tag{19}$$

Note that this is often the correlation of interest, as it correctly shows how strongly associated the son is to the father. However, it should not be interpreted as the direct effect of the father; the mother plays a role through β_f and, if assortment is matched on individual characteristics, through γ. Also, it demonstrates that changes in the observed rates of intergenerational mobility, even if women are omitted, could be driven by changes in marriage matching or in how strongly mothers influence their children.

Note that this also provides an additional way to test if mothers directly influenced the outcomes of their children. If $\beta_f > 0$, then an increase in γ will increase the observed father-son correlation.

Table 4 shows the intergenerational elasticity of the log earnings score between fathers and sons and fathers and sons-in-law estimated separately for more and less assorted marriages of the fathers. The less assorted parents are those where only one parents signed, the more assorted parents those where both spouses either signed or did not sign. The elasticities for the less assorted parents are 0.30 for the son's earnings score and 0.29 for the daughter's husband's earnings score. For the more assorted parents, the elasticities are 0.41 for both the son's earnings score and the daughter's husband's earning score. Appendix A5 considers the possibility of endogeneity bias in these estimates.

Table 4: Father-son intergenerational elasticities, more and less assorted marriages

	Dependent variable:			
	Son's log earnings score		Daughter's husband's log earnings score	
	(1)	(2)	(3)	(4)
Father's log earnings score	0.30***	0.41***	0.29***	0.41***
	(0.01)	(0.00)	(0.01)	(0.00)
Parent's differ on signature	X		X	
Parents same on signature		X		X
Decade FE	X	X	X	X
Observations	28,677	135,839	31,442	140,216
Adjusted R^2	0.06	0.11	0.06	0.12

Note: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.

7.2 Assortment matters for estimates of multigenerational mobility

Studies that link more than two generations are less common, due to data constraints (Clark 2014, Olivetti et al. 2018, Solon 2018, Long and Ferrie 2018). I am able to estimate multigenerational mobility with the Quebec data, as shown in Table 5 below.

Table 5: Grandfather-grandson intergenerational elasticities

	Dependent variable:			
	Child's status measure			
	(1)	(2)	(3)	(4)
Panel A: Log earnings score (male)				
Father's log earnings score	0.42***			0.36***
	(0.00)			(0.00)
Maternal grandfather's log earnings score		0.30***		0.13***
		(0.00)		(0.00)
Paternal grandfather's log earnings score			0.30***	0.10***
			(0.00)	(0.00)
Observations	214,856	214,856	214,856	214,856
Adjusted R^2	0.12	0.04	0.04	0.13
Panel B: Husband's log earnings score (female)				
Father's log earnings score	0.37***			0.31***
	(0.00)			(0.00)
Maternal grandfather's log earnings score		0.28***		0.14***
		(0.00)		(0.00)
Paternal grandfather's log earnings score			0.27***	0.10***
			(0.00)	(0.00)
Observations	219,646	219,646	219,646	219,646
Adjusted R^2	0.09	0.04	0.03	0.11

Note: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.

Note that when estimated separately, the elasticities seem to be the same regardless of if the grandfathers are maternal or paternal. When, however, the partial elasticities are estimated controlling for the log earning scores of the other grandfathers and the fathers, the maternal grandfathers have a higher elasticity.

Should we interpret this as maternal grandfathers being more important to the outcomes of grandchildren? The answer is no. To illustrate why, return to the model in Section 2. If we are interested in estimating β_f, the estimate $\hat{\beta}_m$ is biased upwards from omitting y_m and downwards from attenuation. Modeling the bias quickly becomes complicated as more relatives are added to either the model equations or the regression. However, it is helpful to think in terms of y_m and the classical measurement error, u_f.

If directly related to the mother's status, relative's status will have a coefficient biased upwards. If a relative directly determines the father's true status, their status will be, after conditioning on the observed father's status, negatively correlated with u_f. This biases their coefficient upwards; in other words, they appear more important than they are because they are correlated with some of the missing information about the father.

In this framework, we'd expect the maternal grandfather to be more positively correlated with y_m and less negatively correlated with u_f (after conditioning on the father) than the paternal grandfather. We'd therefore expect it to have a larger coefficient after controlling for the father, which is in fact what we observe in Table 5 below.

This exercise demonstrates how caution must be taken in interpreting intergenerational correlations without accounting for the role of women. It would, at first glance, seem plausible to have found evidence that maternal grandfathers mattered more for the outcomes of children than paternal grandfathers. However, it is merely an artifact of measurement error and omitted variable bias.

7.3 Were women active participants in matching and assortment?

[COMING SOON (Basic idea is that I can't say for sure, but it is an interesting quesiton.)]

8 Conclusion

[COMING SOON]

References

Abramitzky, R., L. P. Boustan, K. Eriksson, J. Feigenbaum, and S. Pérez (2019, May). Automated Linking of Historical Data. Technical Report w25825, National Bureau of Economic Research, Cambridge, MA.

Albouy, D. (2008, November). The wage gap between Francophones and Anglophones: A Canadian perspective, 1970–2000. *Canadian Journal of Economics/Revue canadienne d'économique 41*(4), 1211–1238.

Baillargeon, D. (2014). *A Brief History of Women in Quebec*. Studies in Childhood and Family in Canada. Waterloo, Ontario, Canada: Wilfrid Laurier Univeristy Press.

Black, S. E. and P. J. Devereux (2011). Recent Developments in Intergenerational Mobility. In *Handbook of Labor Economics*, Volume 4, pp. 1487–1541. Elsevier.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *The American Economic Review 95*(1), 13.

Bouchard, G., P. Brard, and Y. Lavoie (1981). FONEM : Un code de transcription phonétique pour la reconstitution automatique des familles saguenayennes. *Population, 36^{e} année n°6*, 1085–1103.

Canadian Families Project (2002). *Canadian Families Project.* Victoria, Canada: University of Victoria.

Chadwick, L. and G. Solon (2002, February). Intergenerational Income Mobility Among Daughters. *American Economic Review 92*(1), 335–344.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Clark, G. (2014). *The Son Also Rises: Surnames and the History of Social Mobility* (First ed.). The Princeton Economic History of the Western World. Princeton: Princeton University Press.

Clark, G. and N. Cummins (2015, February). Intergenerational Wealth Mobility in England, 1858-2012: Surnames and Social Mobility. *The Economic Journal 125*(582), 61–85.

Clark, G., N. Cummins, and M. Curtis (2020, August). Twins Support the Absence of Parity-Dependent Fertility Control in Pretransition Populations. *Demography 57*(4), 1571–1595.

Craig, J., K. Eriksson, and G. T. Niemesh (2020). Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1910. pp. 24.

de Vries, J. (2008). *The Industrious Revolution: Consumer Behavior and the Household Economy, 1650 to the Present*. New York: Cambridge University Press.

Dillon, L., M. Amorevieta-Gentil, M. Caron, C. Lewis, A. Guay-Giroux, B. Desjardins, and A. Gagnon (2018, January). The Programme de recherche en démographie historique: Past, present and future developments in family reconstitution. *The History of the Family 23*(1), 20–53.

Dillon, L., North Atlantic Population Project, and Minnesota Population Center. (2008). National Sample of the 1881 Census of Canada (version 2.0).

Eika, L., M. Mogstad, and Z. Basit (2019). Educational Assortative Mating and Household Income Inequality. *Journal of Political Economy 127*(6), 2795–2835.

Gaffield, C., P. Baskerville, S. Cadigan, M. St-Hilaire, C. Bellavance, F. Normand, G. Darroch, C. Amhrein, L. Tepperman, C. Jones, and E. Sager (2009). *National Sample of the 1911 Census of Canada [Dataset]*. University of Alberta [distributor].

Geloso, V. and P. Lindert (2020, September). Relative costs of living, for richer and poorer, 1688–1914. *Cliometrica 14*(3), 417–442.

Goldin, C. (2006). The Quiet Revolution That Transformed Women's Employment, Education, and Family. *96*(2), 43.

Greer, A. (1985). *Peasant, Lord, and Merchant: Rural Society in Three Quebec Parishes 1740–1840*. Toronto: University of Toronto Press.

Greer, A. (1997). *The People of New France*. Toronto: University of Toronto Press.

Hryshko, D., C. Juhn, and K. McCue (2017, October). Trends in earnings inequality and earnings instability among U.S. couples: How important is assortative matching? *Labour Economics 48*, 168–182.

IMPQ (2020). *Infrastructure Intégrée Des Microdonnées Historiques de La Population Du Québec*. Trois-Rivières, Québec: Centre Interuniversitaires d'Études Québécoises.

Inwood, K. and C. Jack (2011). *National Sample of the 1891 Census of Canada*. Guelph, Canada: University of Guelph.

Killingsworth, M. R. and J. J. Heckman (1986). Female Labor Supply: A Survey. In O. C. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 1.

Long, C. D. (1958). *The Labor Force under Changing Income and Employment.* Princeton, NJ: Princeton University Press.

Long, J. and J. Ferrie (2018, July). Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the US and Britain, 1850–1911. *The Economic Journal 128*(612), F422–F445.

Miller, D., N. Shenhav, and M. Grosz (2019, August). Selection into Identification in Fixed Effects Models, with Application to Head Start. Technical Report w26174, National Bureau of Economic Research, Cambridge, MA.

Minnesota Population Center (2019). *Integrated Public Use Microdata Series, International: Version 7.2 [Dataset]*. Minneapolis: IPUMS.

Olivetti, C. and M. D. Paserman (2015). In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850–1940. *The American Economic Review 105*(8), 2695–2724.

Olivetti, C., M. D. Paserman, and L. Salisbury (2018). Three-generation mobility in the United States, 1850–1940: The role of maternal and paternal grandparents. *Explorations in Economic History 70*, 73–90.

PRDH (2020). *Programme de Recherche En Démographie Historique. Registre de La Population Du Québec Ancien (RPQA) [Data Set]*. Montréal, Canada: Département de Démographie, Université de Montréal.

Project Balsac (2020). *BALSAC [Data Set]*. Québec, Canada: Université du Québec à Chicoutimi.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, 393–408.

Solon, G. (2018, July). What Do We Know So Far about Multigenerational Mobility? *The Economic Journal 128*(612), F340–F352.

Song, X., C. G. Massey, K. A. Rolf, J. P. Ferrie, J. L. Rothbaum, and Y. Xie (2020, January). Long-term decline in intergenerational mobility in the United States since the 1850s. *Proceedings of the National Academy of Sciences 117*(1), 251–258.

Tremblay, M. and K. Roth (2010). *Quebec Women and Legislative Representation*. Vancouver: UBC Press.

Van Leeuwen, M. H., I. Maas, and A. Miles (2004). Creating a historical international standard classification of occupations an exercise in multinational interdisciplinary cooperation. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 37*(4), 186–197.

Vézina, H., D. Gauvreau, and A. Gagnon (2014, April). Socioeconomic fertility differentials in a late transition setting: A micro-level analysis of the Saguenay region in Quebec. *Demographic Research 30*, 1097–1128.

Vézina, H., M. St-Hilaire, and J.-S. Bournival (2013). The linkage of micro census data and vital records: An assessment of results on Quebec historical censuses (1852-1911). pp. 25.

Zijdeman, R. (2014). *Oclack - Occupational CLAssifications and CrosswalKs [Github Repository].*

Appendix

A1 Modelling gender-specific intergenerational inheritance

[COMING SOON]

A2 Bounding assortment with different marriage matching assumptions

[COMING SOON]

A3 Alternative measures of economic status

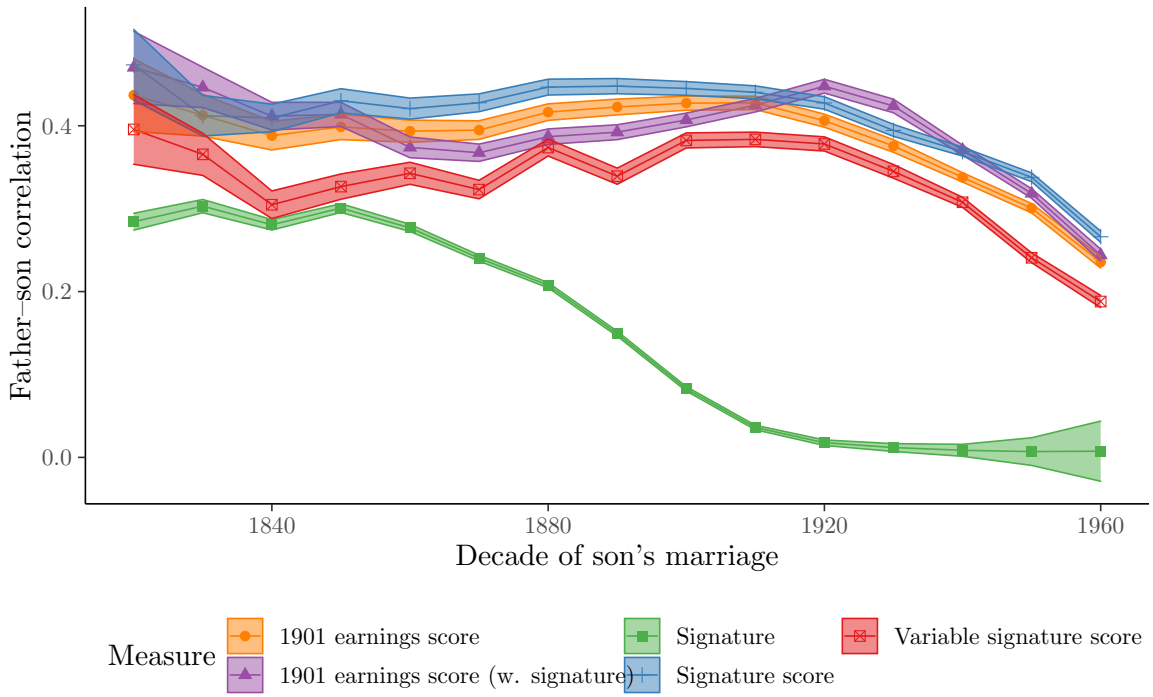


Figure 4: Alternative measures of economic status
Note: 95% confidence intervals shaded. The estimates are rank-rank regression coefficients (Chetty et al. 2014). The 1901 earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. 1901 earnings scores, signature are the average earnings for an individual with that occupation who was reported as able to write if the individual signed their first marriage certificate and not able to write otherwise. Signature scores are the percent of men with that occupation in the 1890's in the vital records who could sign their name. Variable signature scores are computed for each decade using the method in Song et al. (2020); for each occupational category and decade, the score is the sum of the percentile rank of each educational group (signed and not signed) weighted by the share of the occupation in that category. This is essentially a re-weighted version of the average signature rate in that group that accounts for the varying rate of signatures over time. Note that while the levels differ, all of the measures display similar trends over time except for just the signature indicators; this is likely because the literacy rate is rising substantially over time and therefore it becomes less informative about socioeconomic status.

A4 Robustness: Controlling for selection into identification

The estimates in Table 1 Columns 2 and 4 are using off family fixed effects. This means that only families with one child who signed and one child who did not sign are the families which are driving the estimation results. The estimated coefficients are an average treatment effect of individuals in the treated families being able to sign. It is possible that these families

have unique characteristics that make them more likely to be a treated family. A more interesting average treatment effect is, perhaps, for individuals in the whole population.

One method of estimating this population-wide effect is to estimate the effect separately for each treated family and use a weighted average of the effects (Miller et al. 2019). The weights are inverse propensity scores, estimated from a logistic regression of an indicator for being treated regressed on observed family characteristics using the entire sample and normalized to sum to one. For this to be a true average treatment effect, the method does come at the cost of several fairly strict assumptions.[6] However, regardless of assumptions it is still a useful exercise to see if the estimates are robust to re-weighting.

Here, I estimate the propensity scores using indicator variables for the parent's signatures, the mother's decade of first marriage, the mother's borough of first marriage, the denomination of the parish where the mother first got married, and the number of marriages of children observed. Missing values are included as an additional indicator variable category in the logistic regression.

As shown in Table 6 below, there is still a positive and significant premium for the ability of individual to write. The effects on the ability of the spouse to sign increases from 0.38 to 0.58 for women and 0.35 to 0.56 for men; the estimates for occupational earnings scores remain unchanged. The symmetry of the effects across genders is also robust to the re-weighting.

[6]The assumptions: 1. There is no selection into treatment within groups. 2. Conditional on observables, there is no selection into treatment between groups based on heterogenous effects. 3. The logistic regression is the correct functional form. 4. There is a non-zero probability of treatment for every value of observable.

Table 6: Female human capital and marriage matches, selection into identification

	Dependent variable: Spouse's characteristic					
	Signed		Log imputed earnings		Father's log imputed earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of wife's human capital						
Wife signed	0.38***	0.58***	0.04***	0.03***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Re-weighted		X		X		X
Wife's family FE	X	X	X	X	X	X
Decade FE	X	X	X	X	X	X
Observations	1,994,132	1,994,132	1,161,203	1,161,203	984,272	984,272
Panel B: Effect of husband's human capital						
Wife signed	0.35***	0.56***			0.03***	0.03***
	(0.00)	(0.00)			(0.00)	(0.00)
Re-weighted		X				X
Husband's family FE	X	X			X	X
Decade FE	X	X			X	X
Observations	1,983,328	1,983,328			1,001,536	1,001,536

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. The sample excludes individuals with one or more unknown parents. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time. Re-weighted estimates are constructed by estimating the effect separately for each family and then taking the weighted average of the effects. The weights are inverse propensity score weights constructed by running a logistic regression of an indicator for if a family had at least one child who signed and one who did not on indicator variables for the parent's signatures, the mother's decade of first marriage, the mother's borough of first marriage, the decade, and the number of marriages of children observed. Missing values are included as an additional indicator variable category in the logistic regression.

A5 Robustness: Restricting regression samples to the 19th century

As mentioned above, education levels in Quebec increased dramatically in the 19th century. By c. 1900, almost everyone signed their marriage certificates. While signature measure does a good job of tracking the trends in self-reported ability to write, as shown in Figure 1, it also is of smaller magnitude. There is a chance the signature measure includes some false negatives, and therefore after c. 1900, the proportion of those without a signature who can actually write may be high. Table 7 shows that the estimates are robust to restricting the sample to the 19th century.

Table 7: Human capital and marriage matches, 19th century

	Dependent variable: Spouse's characteristic					
	Signed		Log imputed earnings		Father's log imputed earnings	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of wife's human capital						
Wife signed	0.45***	0.27***	0.17***	0.04***	0.07***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Wife's family FE		X		X		X
Decade FE	X	X	X	X		X
Observations	575,874	575,874	352,079	352,079	252,440	252,440
Adjusted R^2	0.31	0.43	0.05	0.42	0.04	0.30
Panel B: Effect of husband's human capital						
Husband signed	0.42***	0.29***			0.12***	0.04***
	(0.00)	(0.00)			(0.00)	(0.00)
Husband's family FE		X				X
Decade FE	X	X			X	X
Observations	570,989	570,989			261,482	261,482
Adjusted R^2	0.38	0.46			0.05	0.31

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. The sample is restricted to the 19th century and excludes individuals with one or more unknown parents. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.

A6 Addressing Table 4's endogeneity concerns

One concern with 4 is that parents who do not match on signatures might be selected on some omitted factor that decreases intergenerational mobility. One way to overcome this endogeneity is to find a variable that changes the degree of assortment of the parents' marriage and only matters for the outcome of the children through the degree of assortment. One plausible variable that meets these criteria is the fraction of the mother's older siblings who are female.

The gender of children should be as good as random, especially as there is no evidence of parity-dependent fertility control (Clark et al. 2020). Why should this matter for assortment? One could imagine a scenario where a set of sisters has multiple potential suitors of similar characteristics in their geographic proximity or social network. As more of the sisters marry, the remaining sisters will have to be less picky. What about the exclusion restriction? While

it is possible that aunts are more beneficial than uncles to children, given the symmetry across gender of the previous results it would be surprising. Finally, it is possible that older sisters have a different effect on younger siblings than older brothers; here this only matters if it changes the younger siblings in a way that does not change marriage matching but does change the outcome of their children. Unfortunately, I do not observe ages in most of the sample, I instead use the percentage of the mother's siblings who got married before her who are female.

As shown in Table 8, the fraction female decreases the association between the signature rates of spouses and decreases the intergenerational elasticity between fathers and sons. This is exactly what we'd expect if the mother directly mattered for the outcomes of children.

Table 8: Father-son intergenerational elasticities, more and less assorted marriages

	Dependent variable:	
	Mother signed	Log earnings score
	(1)	(2)
Fraction female (married before mother)	0.01^{***}	0.13^{***}
	(0.002)	(0.03)
Father signed	0.38^{***}	
	(0.002)	
Father signed X fraction	-0.01^{***}	
	(0.003)	
Father's log earning score		0.41^{***}
		(0.004)
Father's log earning score X fraction		-0.02^{***}
		(0.005)
Sibling marriage order FEs	X	X
Decade FEs	X	X
Observations	390,145	390,145
Adjusted R^2	0.55	0.13

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. Signature variables are indicators that are one if we observe the individual sign their first marriage certificate, zero otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. Decade fixed-effects are included in every specification as the signature rate varies significantly across time.