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

Matthew Curtis[*]

November 2, 2020

For the latest version click here.

Abstract

When did strongly assortative marriages begin to increase inequality? Many believe spouses matching strongly on ability are a modern development, a consequence of rising female employment and education levels. Surprisingly, I find that marriage in Quebec was highly assortative as far back as the early 19th century. Moreover, this assortment played a key role in social mobility and inequality. It was not mere matching between similar families, but instead depended on the human capital of the women as individuals. Although they were rarely employed, their ability still mattered as much as their husbands' for child outcomes. Even in the deeply conservative society of Quebec 1800–1970, the human capital of women mattered both for marriage and for its outcomes.

Keywords Assortative mating, marriage, intergenerational mobility

JEL codes J12, J62, N31, N32

1   Introduction

Women had limited educational and employment opportunities until the late 20th century. Before then, it is easy to assume that who people married was economically unimportant. If only men held formal employment, who they married would not matter for income inequality. However, this assumption overlooks an important channel through which marriage — and women — mattered: intergenerational mobility.

To determine if marital assortment had economic consequences before the late 20th century, this paper sets out to answer three related questions. First, how did the degree of marital assortment evolve over the long run? Second, did the abilities of women matter, or was assortment just between families? Third, how much did the human capital of mothers matter compared to that of fathers for the outcome of their children?

Using a large new linked dataset of families from Quebec 1800–1970, this paper shows that the degree of marital assortment was surprisingly high and stable from the 1830s through the 1960s. Moreover, this assortment was on the human capital of both men and women. Using remarriages, I show that child outcomes depended equally on the ability of mothers and fathers. Despite deeply conservative gender roles, in Quebec 1800–1970 women mattered as much as men for marriage and mobility.Thus, assortative marriage was economically significant long before the rise of female employment and education in the late 20th century.

The historical data from Quebec have several features that are particularly suited to answering these questions. First, the data are close to a complete population registry. Families in the sample are not selected by cohabitation (like in census records) or by living descendants (like in many genealogical records). Second, complete family structures are reconstructed in the dataset. Unusually, as Québécoise women retained their family name after marriage and are therefore can be linked to their parents. I am able to use within-family variation, combined with the size of the dataset (several million observations), to identify the mechanisms linking assortment and mobility.

Before I consider mobility, I first develop a method to measure the degree of assortment.

The data have two complementary proxies for a woman's human capital: her ability to sign her name and the occupational status of her husband or father. The ability to write is a direct measure of human capital that is consistently available for both men and women. However, it corresponds to a very low level of education. After 1900 almost all people were literate. Occupations are a more detailed measure but are only available for men. Therefore, I use signatures when I need a direct measure and the occupation of a woman's husband or father when I need a detailed measure. To infer the degree of assortment from occupations, I construct a simple model of assortment. In this model, a man is two degrees separated from his father-in-law (husbands are similar to their wives and children are similar to their fathers), whereas he is only one degree separate from his father. The correlation between spouses is thus equal to the correlation between men and their father-in-laws divided by the correlation between men and their fathers. This correlation was surprisingly strong and stable from the 1830s through the 1960s.

Does this very high and stable level of assortment, though, have any economic significance? I show it does by using within-family variation to disentangle the interactions between matching and mobility. First, I use family-fixed effects to determine if spouses matched on individual human capital. I find that literate women typically married higher status husbands than their illiterate sisters. The same is true for brothers. Notably, the returns to literacy are symmetric across gender. When it came to marriage, the human capital of women mattered as much as that of men.

Second, I use remarriages to show that the human capital of women mattered for child outcomes. A mother's literacy was correlated with her children as strongly as that of her husband. However, this would still be possible if the mother did not directly matter as long as marriages were assorted on literacy. To address this empirical challenge, I use remarriages to control for the father, both through fixed effects and by directly comparing half siblings. I find evidence that the ability of mothers mattered as much as the ability of fathers.

The consequences of these findings are twofold. First, they strongly imply that assortment

contributed to the overall distributions of human capital 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, women appear to have had an equal role in this at least. Second, the results have implications for standard estimates of intergenerational mobility. I show that assortative marriage directly influences estimates of rates of mobility even when only men are considered. A change in an estimate of mobility could be entirely driven by a change in 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. The late 20th century saw a rise in female employment rates and the closing of the gender gap in educational attainment (Goldin 2006). This 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 were was little resulting change in assortment or inequality (Eika et al. 2019, Hryshko et al. 2017). This paper provides a longer-run context that suggests an explanation: marriage had long been highly assortative.

Second, this paper 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). More recent work considers the mobility of daughters as well as sons (Chadwick and Solon 2002). In historical settings, this is typically done by comparing daughters to their fathers (Olivetti and Paserman 2015, Craig et al. 2020). This paper goes even further by considering the role of their mothers as well. It finds that mothers mattered and demonstrates that even if they are not observed they can effect the standard estimates of mobility.

The structure of the paper is as follows. First, I provide a brief summary of the historical context. Second, I describe the dataset, how families were linked, and the measures I use of human capital. Third, I construct a measure of the degree of assortment over time. Fourth, I present evidence that this assortment consisted of matches between individuals of similar

human capital. Fifth, I estimate the effect of parental human capital on child outcomes. Sixth, I discuss the broader implications of these findings. Finally, I conclude: assortment mattered for inequality long before the late 20th century because women had always played an important role in marriage and mobility.

2 Historical context

Before its Quiet Revolution of the 1960's, Quebec was much less secular than its North American neighbors.[1] Catholicism asserted significant control over public education and social norms, and deeply conservative beliefs about gender roles were enshrined by law and public policy. For example, while Québécoise women could vote in federal elections after 1918, they could not vote in local elections until 1940 (Tremblay and Roth 2010).[2]

2.1 The legal rights of women

Women in most historical societies faced systematic legal disadvantages; Quebec was no different. 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. Under the Custom of Paris, and unlike in English-speaking legal traditions, married couples formed a legal entity called the *communauté de biens* (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,[3] though the husband alone was expected to manage 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

---

[1]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]Unusually, 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. This was considered by 19th century reformers a concerning oversight that needed to be addressed.

[3]As mentioned below, this greatly aids the linking of vital records.

Lower Canada, introduced in 1866, only clarified the legal disadvantages of women. Only after reforms starting in 1964 were women no longer legally incapable.

In theory, the law did not discriminate when it came to 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.[4] 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, typically an older son (Greer 1985).

2.2   The demographic regime

Was an unequal partnership in marriage the typical experience for women in Quebec? Quebec had a variant of the European marriage pattern, with younger marriages and less frequent celibacy than France (Greer 1997). Quick remarriage upon the death of a spouse was common. While married, a woman typically gave birth to a child roughly every two years until her forties. Most women married, and for most women marriage marked the beginning of many years of pregnancy and childcare.

While this variant fertility pattern was common in European settler colonies, Quebec kept it longer than most. The demographic transition occurred later, 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 family size (Clark et al. 2020).

2.3   Women and the workforce

While the economy of Quebec evolved dramatically from 1800–1970, women consistently had limited opportunities in the formal labor market. In the first half of the 19th century,

---

[4]One of the few advantages women had was that a widow could renounce the debts of the community of property as it was assumed she was not responsible for their accumulation.

women usually contributed labor to household production as well as to domestic tasks (Baillargeon 2014). In urban areas, formal employment was available to unmarried women as domestic servants. Women also found employment as educators, first as nuns and and later as secular teachers. As the economy began to industrialize in the 1840's, unmarried women were also employed by 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). By the late 19th and early 20th century, female dominated occupations emerged such as telephone operators, typists, and secular nurses. However, married women were still expected to be housewives until the 1970's.

2.4   Quebec, North America, and external validity

Overall, how much was Quebec an outlier? In general, it 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 the gap was stable over time (Albouy 2008, Geloso and Lindert 2020). Its deeply conservative society delayed the extension of rights to women, but not indefinitely. Its demographic regime was characterized by large family sizes and a delayed demographic transition, but it was still a variant of the 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. While Quebec was behind its neighbors in levels of human capital and in rights for women, it was not fundamentally different when it came to marriage and mobility. If women and assortment mattered for mobility even in Quebec, they probably mattered in neighboring regions.

3 Data

The IMPQ is a large 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 use data from a period with frequently reported occupations for men, 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 2.1 million unique marriages 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 million other individuals are mentioned besides the main participants, providing additional observations over time for many people besides their own vital events.

3.1 Linked family vital records

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 documents. Second, women kept their family names when they married. This means both that women can be linked to their fathers and that most vital records have four names on which to link (the first names and last names of both the husband and wife or mother and father).

In both of the databases, links were formed using two similar computer-assisted matching systems (Vézina et al. 2013, Dillon et al. 2018). These two procedures differ slightly,[5] but they both use the four names to link records and resort to manual linkage in difficult cases. Manual linkages are not necessarily better than automatic linkages; in some applications they produce both more true matches and more false positives (Abramitzky et al. 2019). However, the fact that the Quebec vital records have four names to match on should increase

---

[5]e.g. the BALSAC database standardizes names using the FONEM phonetic algorithm (Bouchard et al. 1981), whereas the RPQA uses a custom-made name dictionary

the accuracy of matching regardless of the 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 (Dillon et al. 2018). Therefore, records of almost the entire population survive; this will reduce false positive rates in an analogous way to the linking of full count to full count censuses (Abramitzky et al. 2019).

3.2   Measures of human capital

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 they were able and the priest to record if they were not. I code a signature indicator variable as one if the individual signed their marriage record, and zero if they were unable to sign. I omit cases that are either missing or ambiguous. As shown in Figure 1 below, this definition produces a trend that is close to external estimates of literacy. Was a signature really a measure of human capital, that is a productive attribute? The qualitative evidence suggests 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).[6]

The vital records also often state the occupations of men. 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 men with that occupation in Quebec i a 5% sample of the 1901 Canadian Census (Canadian Families Project 2002 and Minnesota Population Center 2019).[7] There are numerous other ways to

---

[6]As for reading, it too was likely associated with economic activity 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.

[7]This requires crosswalking occupations from IPUMS's occupational codes to the original HISCO scheme

rank occupations, as discussed in Appendix A2. However, these imputed earnings are easy to interpret,[8] are at least a proxy for the standard variable of interest in mobility studies (lifetime earnings), and produce similar estimates to the other ranking systems. Therefore, the main results in this paper use the 1901 imputed earnings as the primary measure of occupational status.

3.3   Women's presence in the vital records

Do the vital records accurately record the occupational status and human capital of women? Four extracts of Canadian censuses 1891–1911 and data compiled by Long (1958) for 1920–1960 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).[9]

Figure 1 compares the fraction of individuals signed their first marriage record to the fraction who self-reported the ability to write in the censuses. Unlike the censuses, individuals only appear in the marriage records during a specific time in their lives. To account for this, I reweight the census data to match the age distribution of the vital records. As shown in the figure, my estimated literacy rate closely tracks the rate in the census. Two patterns are particularly notable. First, Quebec went from a very low human capital society to a high human capital society from 1800 to 1920. Second, there was actually a gender gap in favor of women from 1850 to 1920.

In contrast, the vital records do a poor job of recording female occupations. Figure 2 shows the employment rate of women by marital status. Here, I reweight the vital records data to match the age distribution in the census. Compared to the other sources, the vital records underestimate the employment rate of married women and almost entirely omit unmarried women with occupations. One pattern, however, is clear. While unmarried
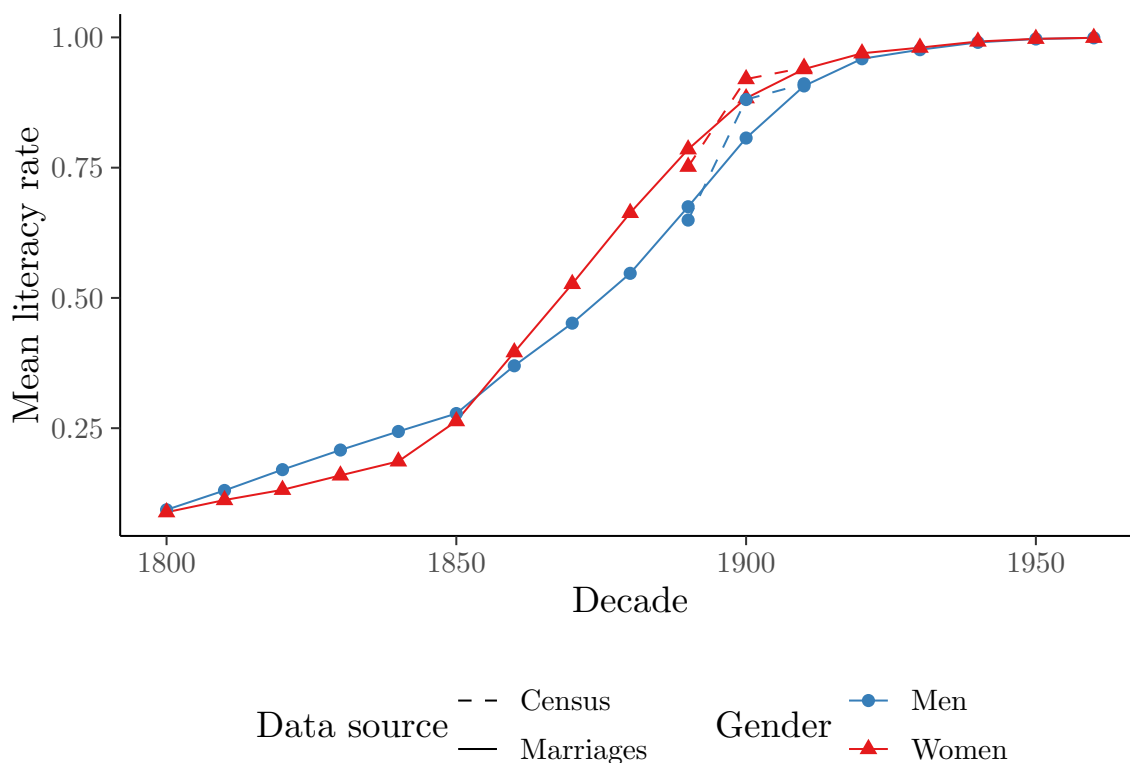

(Zijdeman 2014).

[8]i.e. how much the individual would earn, on average, with their occupation in 1901.

[9]The census extracts are the 100% 1881 sample, the 5% 1891 sample, the 5% 1901 sample and 1901 oversample, and the 5% 1911 sample.

women often worked outside the home, married women did not begin to work in substantial numbers until the 1940's.

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

Note: The vital record literacy rate is the average of an indicator variable that is one if a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. The census record literacy rate is the fraction of individuals who were reported as able to write, reweighted to match the age distribution in the vital records. The two sources broadly agree.
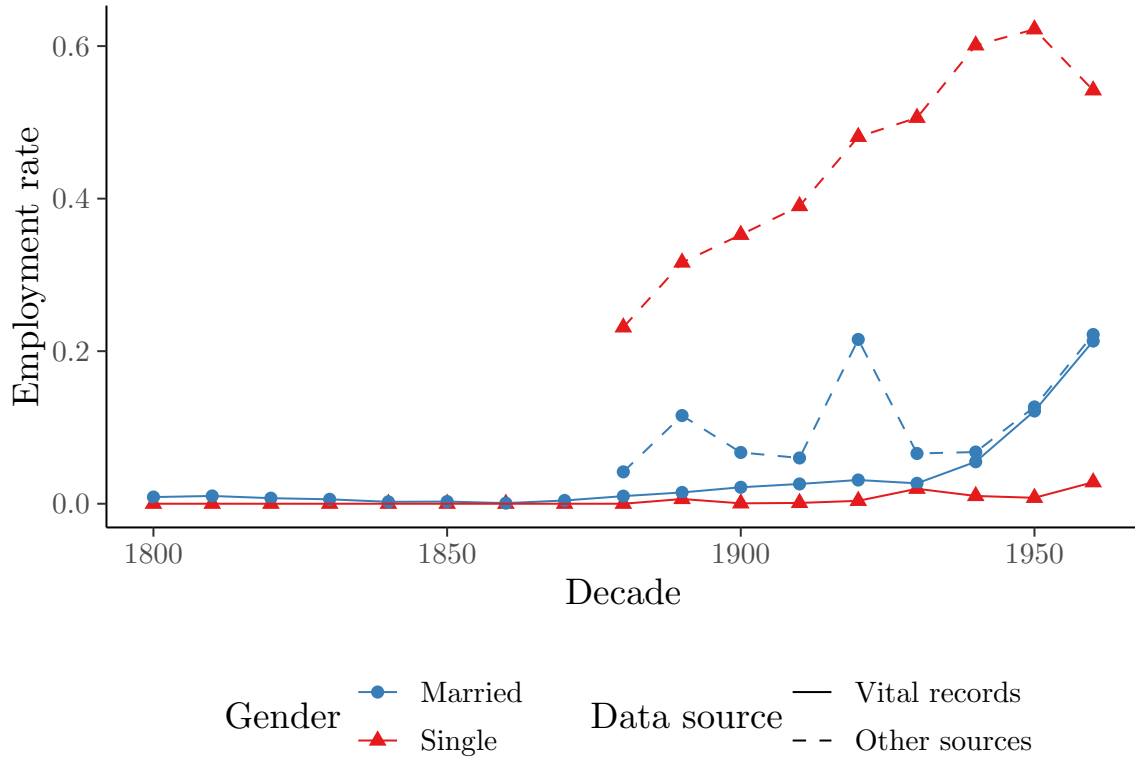
Figure 2: The vital records fail to record female employment
Note: Each woman in the vital records is counted as employed if she ever had a job reported and as not employed otherwise. She is then assigned a year equal to the median of all the years at which she is observed. Then, I compute the average employment rate for each decade, reweighted by age to match the age distribution in the census data. Before 1920, the other sources are census extracts and the rates the fraction of women aged at least 16 with an occupation. After 1920, the other sources are rates compiled by Long (1958), with the married rate calculated as an average of the rates for currently married women and widowed or divorced women weighted by the relative frequencies of the two categories in the censuses.

4 Measuring the degree of marital assortment

Below, I develop a model to illustrate how marriage and intergenerational mobility contribute to inequality over the long-run. This framework, while simple, suggests a new method to measure the degree of marital assortment. Using this method, I show that assortment was surprisingly high and stable over the period 1830–1969.

4.1 A simple model of marriage, mobility, and inequality

We are interested in a specific measure of potential socioeconomic status, call it x. In the intergenerational mobility literature, the typical measure of status is lifetime earnings. The distinction I make here is that I consider *potential* socioeconomic status. For example, while women often were not employed, they still possessed human capital that could, in different circumstances, be used to earn a wage.

Following Solon (1992) and Clark and Cummins (2015), assume only an imperfect measure or proxy y is observed for x. The typical example is proxying for lifetime earnings with the earnings in a specific year. As this model considers potential status, for some individuals no y is observed. For the others, y is observed with classical measurement error:

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

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

Then assume that the status of child c, x_c, is inherited depending on the status of 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 realized status.

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

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

Now note that if variances are equal, I can re-write the assortment correlation equation as a linear relationship:

$$x_i = \gamma x_s + v_i \tag{4}$$

where s is i's spouse and v_i is an uncorrelated error term. If I substitute this into the intergenerational mobility equation, I get:

$$x_c = (\beta_f + \gamma\beta_m)x_f + \gamma\beta_m v_f + e_c \tag{5}$$

This can be estimated with a regression:

$$y_c = \alpha_0 + \alpha_1 y_f + \epsilon_c \tag{6}$$

where $\alpha_1 = (\beta_f + \gamma\beta_m)$. However, as y_f is correlated with u_f the estimate is attenuated down. Specifically, as:

$$y_c = (\beta_f + \gamma\beta_m)y_f - (\beta_f + \gamma\beta_m)u_f + \beta_m v_f + e_c - u_c \tag{7}$$

there is bias of the form:

$$plim \ \hat{\alpha}_1 = (\beta_f + \gamma\beta_m)\frac{\sigma^2_{x_f}}{\sigma^2_{x_f} + \sigma^2_{u_f}} \tag{8}$$

If β_f and γ are greater than zero, then mothers contribute to the observed correlation of x between fathers and sons. Appendix A1 extends this model, demonstrating the effects β_f and γ on inequality.

4.2   Measuring assortment if women are not observed

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

14

during this interval. As the average education level increased, the average education level of individuals who could at least sign 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 it became more common, they are concentrated in lower status occupations such as housekeeping). 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, 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.

Using the model, I can construct an estimate of the degree of assortment by comparing the correlation between father-in-laws and son-in-laws to the correlation between sons and fathers. Letting y_{fl} be the observed status of the father-in-law of i:

$$y_i = \gamma(\beta_f + \gamma\beta_m)y_{fl} - \gamma(\beta_f + \gamma\beta_m)u_{fl} + \gamma\beta_m v_{fl} + \gamma e_i + v_i - u_i \tag{9}$$
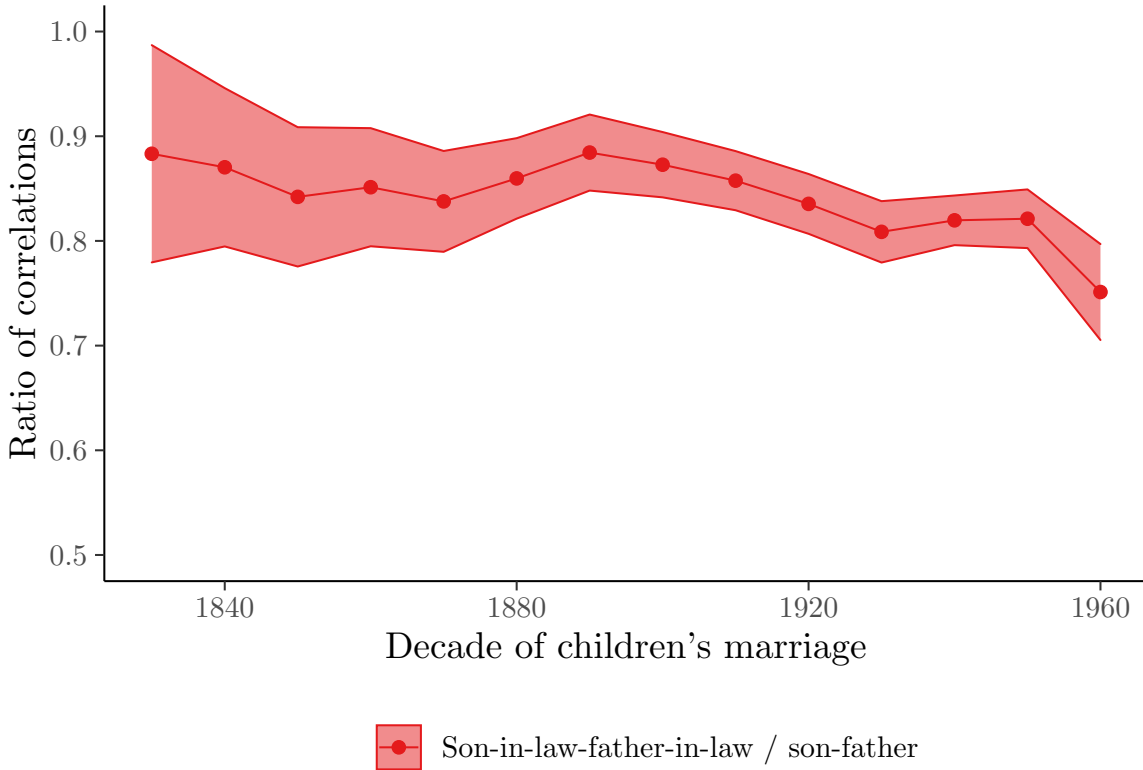
and

$$y_i = (\beta_f + \gamma\beta_m)y_f - (\beta_f + \gamma\beta_m)u_f + \beta_m v_f + e_i - u_i \tag{10}$$

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

$$\gamma\frac{\sigma_{x_{fl}}^2(\sigma_{x_f}^2 + \sigma_{u_f}^2)}{\sigma_{x_f}^2(\sigma_{x_{fl}}^2 + \sigma_{u_{fl}}^2)} \tag{11}$$

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

4.3 Measuring the degree of marital assortment

Figure 3: Ratio measure of martial assortment using imputed earnings
Note: 95% confidence interval shaded. Standard errors are bootstrapped with 50,000 repetitions. In the simple model described in the paper, this ratio is equal to the correlation in potential socioeconomic status between spouses. The father-in-law and son-in-law correlation measures both assortment and intergenerational mobility with attenuation bias. The father and son correlation measures only intergenerational mobility, again with attenuation bias. The ratio then, under a few assumptions, measures assortment directly. The correlations are using imputed earnings, which are the average earnings for the individual's occupation in Quebec based on a 5% extract of the 1901 Canadian Census. The ratio is computed as 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 only approximately equal to Pearson correlation coefficients.

5   Spouses matched on their human capital

Were marriages matches on the individual 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 an alliance between a husband and his father-in-law with the characteristics of the wife an afterthought at best.[10] In this hypothetical society, there could still be an observed correlation in the human capital of spouses if a woman's human capital is partially determined by her father.

To test if individual characteristics mattered, consider the following fixed effects regression:

$$y_s = \alpha y_i + \phi_F + \beta \mathbf{X_s} + \epsilon_{i,F} \tag{12}$$

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 are the crucial fixed effects that control for family, $\mathbf{X_s}$ is a vector of controls, and $\epsilon_{i,F}$ is a family-clustered error term. To address any time trends, $\mathbf{X_s}$ includes fixed effects for both decade and the order of siblings.[11]

In other words, the regression asks if, compared to their siblings, an individual with higher ability matches with a spouse of higher ability? If so, α will be positive.

As shown in Table 1 Panel A below, a woman who signed her marriage record married a man with higher status than her sisters who did not. Being able to write was associated with an increase in the probability a woman's husband was literate by 30 percentage points, an increase in her husband's imputed earnings by 4%, and an increase in her father-in-law's imputed earnings by 2%. This is evidence that marriage matches took into account individual characteristics. 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

---

[10]Marriage as a way of cementing a commercial alliance was not unknown to the early settlers in Quebec. Indeed, marriage *à la façon du pays* ("after the custom of the country") between an indigenous woman and a French fur trader was commonly practiced to cement commercial relationships (Baillargeon 2014).

[11]As I only have date of birth through 1849, I order siblings by the date of their first marriage.

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 Panel B below, men who were able to write also were at an advantage when finding a wife. Being able to write associated with an increase in the probability that a man's wife was literate by 28 percentage points and an increase in his father-in-law's imputed earnings by 3%. These estimates are remarkably similar to those for women. The returns to human capital for marriage matching appear to be the same regardless of gender. Appendix A3 discusses the robustness of these estimates.

What is the economic significance of marriages matching on the individual characteristics of women? If it were not the case, the woman's family could after all still matter for the outcomes 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

	Signed		Log imp. earnings		Father's log imp. earnings	
Dependent variable: Spouse's characteristic						
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effect of wife's human capital						
Wife signed	0.49***	0.30***	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
Identifying observations		203,284		124,731		108,199
Observations	1,937,871	1,937,871	1,148,769	1,148,769	971,173	971,173
Adjusted R^2	0.60	0.64	0.06	0.38	0.03	0.32
Panel B: Effect of husband's human capital						
Husband signed	0.41***	0.28***			0.11***	0.03***
	(0.00)	(0.00)			(0.00)	(0.00)
Husband's family FE		X				X
Identifying observations		230,364				123,465
Observations	1,928,239	1,928,239			986,398	986,398
Adjusted R^2	0.62	0.64			0.04	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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in 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. Identifying observations are the number of observations from families where at least one child signed and one child did not sign.

6   The ability of mothers mattered for child outcomes

Did the abilities of mothers directly influence the outcomes of their children? Here, I am not interested in the exact mechanism.[12] All that is required for women to play a role in intergenerational mobility is for their abilities to matter as well as those of their husbands.

---

[12]Such as, for example, the distinction between human capital and innate ability (Black et al. 2005).

6.1 Controlling for the father with fixed effects

A mother's literacy is correlated with that of her children even after controlling for that of the father (Table 2). However, this pattern could still be observed if the mother did not directly matter for the outcomes of the children. 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 do not identify the effect of the ability of mothers. Notably, there appears to be no difference between the associations with children of different genders.

To identify a causal effect, ideally we would control for the father but randomize the mother. A less ideal yet possible approach 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 are scarred by whatever event resulted in a second marriage (such as a death or divorce). Assuming the penalty is a constant, it can be controlled for by including fixed effects for the 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:

$$y_c = \alpha y_m + \phi_f + \beta \mathbf{X_c} + \epsilon \tag{13}$$

where y_c is an outcome of a child, y_m is a characteristic of the mother, ϕ_f are the crucial fixed effects that control for the father, and $\beta \mathbf{X_c}$ are controls. To address any time trends, X_c includes fixed effects for decade, the marriage number of the father, and for the order of siblings.

As shown in Table 3 Panel A, even controlling for the father, a mother who could sign her name had children who were 3% more likely to be able to sign their names and had a

3% higher occupational status score.

For comparison, Table 3 Panel B 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. Appendix A4 discusses the robustness of these estimates.

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

	Dependent variable:			
	Signed Daughter	Signed Son	Log imp. earnings Son-in-law	Son
	(1)	(2)	(3)	(4)
Mother signed	0.12***	0.12***	0.06***	0.07***
	(0.001)	(0.001)	(0.002)	(0.002)
Father signed	0.07***	0.09***	0.15***	0.16***
	(0.001)	(0.001)	(0.002)	(0.002)
Observations	1,551,089	1,435,443	938,558	875,264
Adjusted R^2	0.51	0.50	0.06	0.06

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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time.

Table 3: The effect of parental human capital on child outcomes

	Signed Daughter	Signed Son	Log imp. earnings Son-in-law	Son
	Dependent variable:			
	(1)	(2)	(3)	(4)
Panel A: Controlling for father				
Mother signed	0.02***	0.03***	0.01	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Father FEs	X	X	X	X
Identifying observations	18,407	16,058	8,532	7,537
Observations	1,571,362	1,454,557	950,687	886,907
Adjusted R^2	0.68	0.67	0.37	0.41
Panel B: Controlling for mother				
Father signed	0.02***	0.03***	0.03*	0.04**
	(0.01)	(0.01)	(0.02)	(0.02)
Mother FEs	X	X	X	X
Identifying observations	6,488	5,516	2,906	2,385
Observations	1,563,894	1,447,566	946,275	882,625
Adjusted R^2	0.69	0.68	0.37	0.41

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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time. Identifying observations are the number of observations with a parent with one spouse who signed and one who did not.

6.2 Directly comparing half siblings

One downside of the father fixed effects approach is that it relies on observing a measure of the ability of the mother. As shown in Table 3, the effective sample size is quite low: very few parents had two spouses, one of which was literate and one of which was not. Hence, not all results are significant at the 5% level.

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 if they 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, assuming the penalty is a constant, fixed effects can control for it.

I estimate the regression:

$$y_{i,m_i,f} = \alpha Y_{j,m_j,f} \times I(m_i = m_j) + \beta \mathbf{X_{i,j}} + \epsilon_f \tag{14}$$

Where $y_{i,m_i,f}$ is a characteristic of child i with father f and mother m_i, i is less than j, $I(m_i = m_j)$ is an indicator that is one if the children share a mother, $\mathbf{X_{i,j}}$ are control variables, and ϵ_f is an error term. The controls include fixed effects for decade, the order of the siblings, and the marriage number of the father.

The results are shown in Table 4 below. Full siblings are more strongly associated than half siblings. Moreover, the results are very similar regardless of I let the mothers or fathers vary. Appendix A4 discusses the robustness of these estimates.

Table 4: The effect of parental human capital on half vrs. full siblings

	Dependent variable: Younger sibling's characteristic			
	Signed Daughter	Signed Son	Log imp. earnings Son-in-law	Log imp. earnings Son
	(1)	(2)	(3)	(4)
Panel A: Controlling for father				
Older sibling's characteristic	0.36***	0.36***	0.22***	0.26***
	(0.00)	(0.00)	(0.01)	(0.01)
Same mother	−0.01*	−0.01**	−0.30***	−0.30***
	(0.00)	(0.00)	(0.05)	(0.05)
Signed × same mother	0.06***	0.05***	0.05***	0.05***
	(0.00)	(0.00)	(0.01)	(0.01)
Observations	2,050,264	1,853,707	839,388	756,645
Adjusted R^2	0.64	0.63	0.11	0.14
Panel B: Controlling for mother				
Older sibling's characteristic	0.36***	0.33***	0.24***	0.22***
	(0.01)	(0.01)	(0.02)	(0.02)
Same father	−0.04***	−0.04***	−0.16	−0.53***
	(0.01)	(0.01)	(0.10)	(0.11)
Signed × same father	0.07***	0.08***	0.03*	0.09***
	(0.01)	(0.01)	(0.02)	(0.02)
Observations	1,965,701	1,777,710	806,656	727,123
Adjusted R^2	0.64	0.63	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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time.

7 Discussion

7.1 Assortment matters for intergenerational elasticities

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. In the simple model in Section 4.1, the association between fathers and sons is:

$$y_c = (\beta_m + \gamma\beta_f)y_f + \epsilon_c \tag{15}$$

Note that $\beta_m + \gamma\beta_f$ is often the correlation of interest, as it shows how strongly associated sons are with their fathers. However, it should not be interpreted as the direct effect of the father. If the parents matched on individual characteristics, the mother increases the association through the $\gamma\beta_f$ term. Changes in the observed rates of intergenerational mobility, even if women are omitted, could be driven by changes in marriage matching (γ) or in how strongly mothers influence their children β_f.

Table 5 shows the intergenerational elasticity of imputed earnings estimated separately for more and less assorted parents. The less assorted parents are those where only one parents was literate, the more assorted parents those where both parents were either literate or illiterate. The elasticities for the less assorted parents are 0.30 for the and 0.28 for daughters (using their husbands' imputed earnings as a proxy). For the more assorted parents, the elasticities are 0.42 for sons and 0.41 for daughters. As predicted, the more strongly assorted parents have higher estimate rates of intergenerational mobility. Appendix A5 considers the possibility of endogeneity bias in these estimates.

Table 5: 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***	0.42***	0.28***	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	27,278	125,094	30,022	129,928
Adjusted R^2	0.06	0.11	0.05	0.12

Note: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses. Signature variables are indicators that are one if a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in 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 multigenerational mobility

Although they are less common due to data constraints, several studies also consider correlations across more than two generations (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 6 below.

Table 6: Grandfather-grandson intergenerational elasticities

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

Note: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in 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 intergenerational elasticities between grandfathers and grandchildren 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 imputed earnings of the other grandfather and the father, there is a larger coefficient for the maternal grandfathers.

Should we interpret this as maternal grandfathers being more important to the outcomes of grandchildren? The answer is no. To illustrate why, refer back to the model in Section 4.1.

If it is directly related to the mother's true status, a grandfather's observed status will have a coefficient biased upwards as the mother is omitted. Likewise, if it is directly related to the father's true status, it will have a coefficient biased upwards if the father is omitted. Controlling for the father's observed status will reduce the omitted variable bias from the true status of the father much more than it would reduce the omitted variable bias from the mother. In this framework, we'd expect the maternal grandfather to be more strongly correlated with the mother. We would therefore expect it to have a larger coefficient after controlling for the father, which is in fact what we observe in Table 6.

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.

8   Conclusion

In this paper, I construct a simple model of marriage and mobility. It shows that even with no female employment, assortment will increase inequality if the ability of a woman determines whom she marries and the success of her children. To test if this was true in Quebec 1800–1970, I consider millions of families reconstructed from marriage records by the new IMPQ database. Unusually, married women are linked to their fathers; I use this to estimate assortment, finding it surprisingly high and stable over time. Next, I find pairs of sisters where only one was able to write. Even though she likely never earned a wage after she married, I show that the the more educated sister still typically earned a premium when it came to the status of her husband. Moreover, her ability mattered as much as her husband's for the outcomes of their children. As quick remarriage after losing a spouse was the norm, when comparing child outcomes I can hold one parent constant and allow the second to vary. Sharing a mother mattered almost the same as sharing a father. Altogether, I conclude that

assortment had always mattered. It mattered because, despite severe legal and economic disadvantages, women played a major role in mobility and marriage. Overlooking women would leave any story of assortative marriage, intergenerational mobility, and inequality incomplete.

References

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

Albouy, D. (2008). 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, 36ᵉ année n°6*, 1085–1103.

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

Chadwick, L. and G. Solon (2002). 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). 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). 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). 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). 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). 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). Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the US and Britain, 1850–1911. *The Economic Journal 128*(612), F422–F445.

31

Miller, D., N. Shenhav, and M. Grosz (2019). 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). 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). 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). 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 Modeling inequality

The two main findings of this paper — that assortment was on individual characteristics and that the ability of mothers mattered for child outcomes — individually answer somewhat narrow questions. Together, they imply that a high degree of assortment would have contributed to inequality over the long-run. Here, I illustrate the logic behind the claim with the simple model I use to estimate the degree of assortment in Section 4.1.

To summarize inequality in a given generation, consider the variance of potential status:

$$\sigma_{x_c}^2 = \beta_f^2 \sigma_{x_f}^2 + \beta_m^2 \sigma_{x_m}^2 + 2\beta_f \beta_m (\gamma \sigma_{x_f} \sigma_{x_m}) + \sigma_{e_c}^2 \tag{16}$$

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

$$\sigma_{x_c}^2 = \sigma_{x_f}^2 = \sigma_{x_m}^2 \tag{17}$$

Then:

$$\sigma_{x_c}^2 = \frac{\sigma_{e_c}^2}{1 - \beta_f^2 - \beta_m^2 - 2\gamma\beta_f\beta_m} \tag{18}$$

As the error term u_c is assumed to be independent of x_c, the observed inequality is given by:

$$\sigma_{y_c}^2 = \sigma_{x_c}^2 + \sigma_{u_c}^2 \tag{19}$$

Unsurprisingly, the more children take after their parents (i.e. the higher the β_f and β_m), the higher the level of steady state inequality. 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 4).
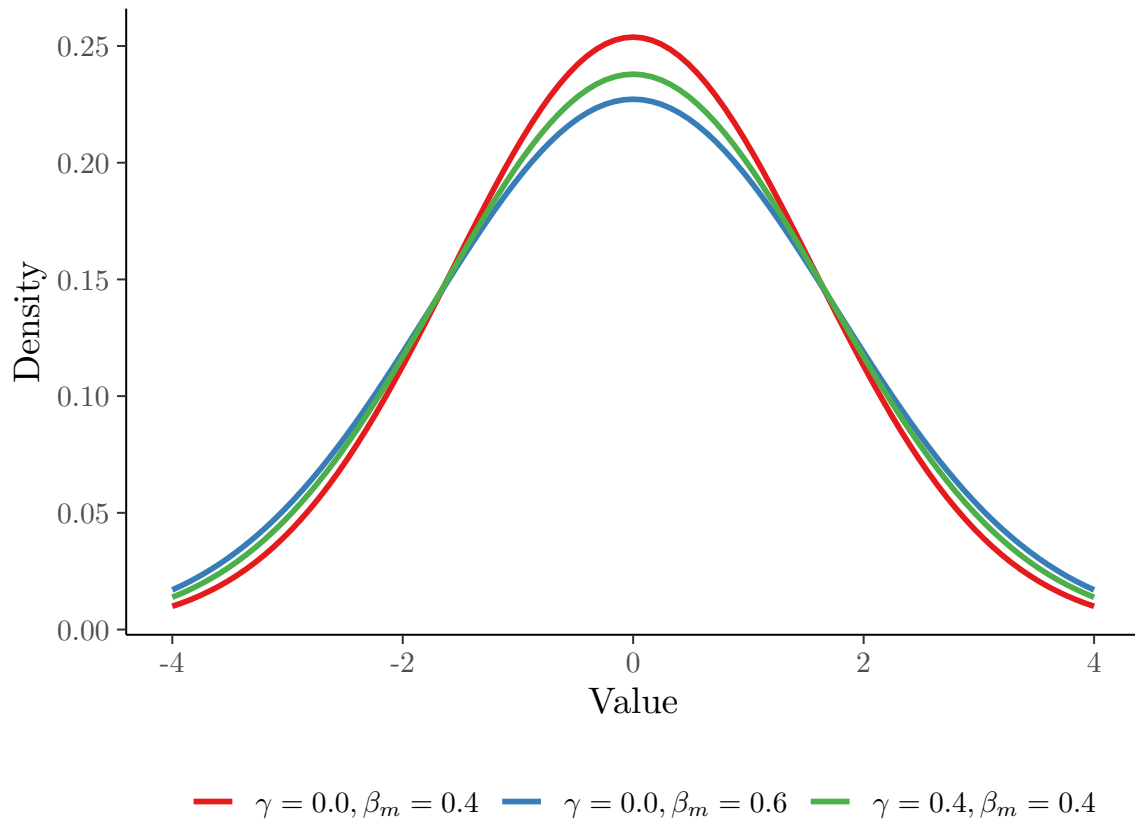
Figure 4: Steady state inequality

Note: Simulated data based on model (see text). γ is the degree of assortment, β_f and β_m are the strength of intergenerational inheritance of potential status from fathers and mothers respectively. 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, increasing γ or β_m increases inequality.

A2: Robustness of estimates of assortment

Figure 5 below computes the ratio measure of assortment used in Figure 3 using a number of different ways to measure human capital. With the exception of literacy, which is not very informative after most people are literate, all the measures give roughly the same picture of the overall trend.
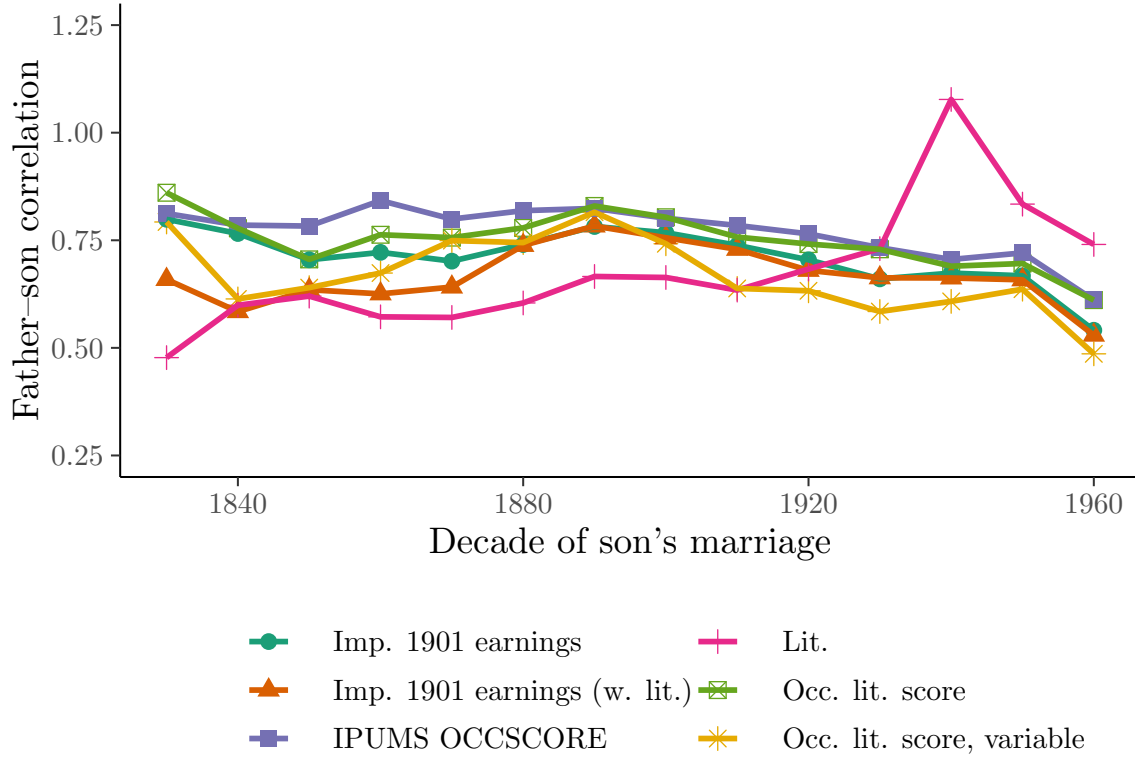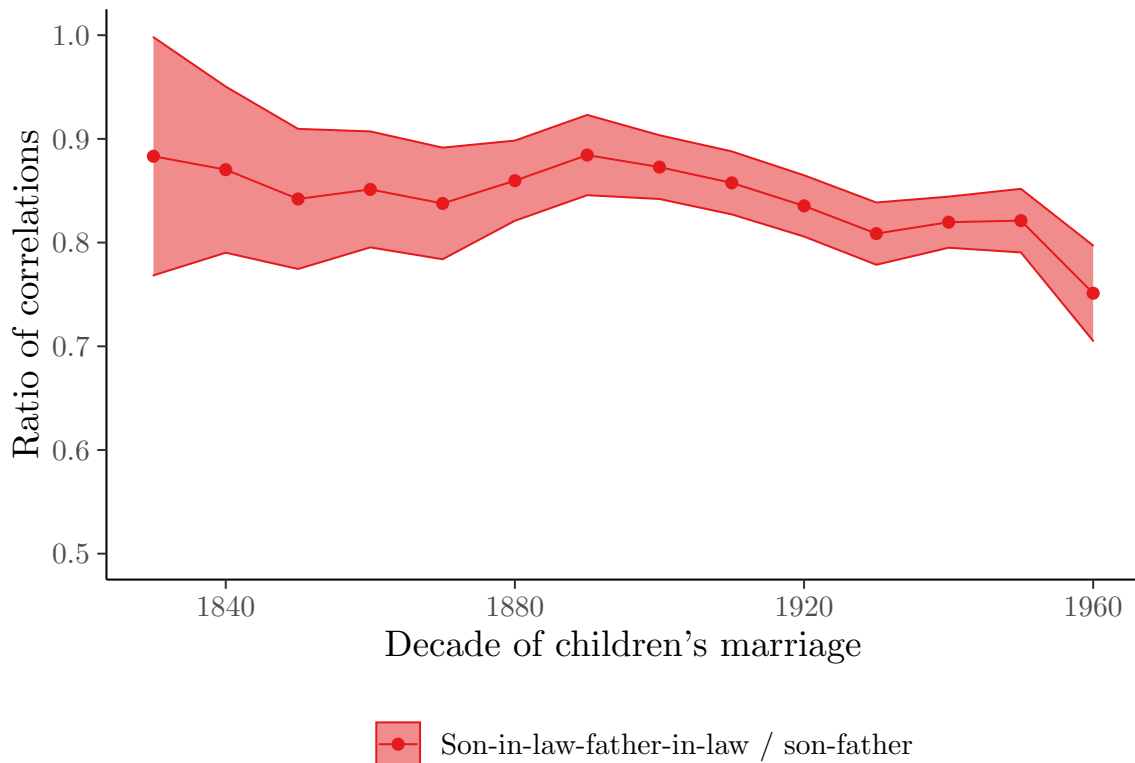
Figure 5: Alternative measures of economic status
Note: 95% confidence intervals shaded. The estimates are ratios of rank-rank regression coefficients, which are equal to the correlation coefficients for the ranked variables assuming no ties in rank (and a reasonable approximation if not) (Chetty et al. 2014). 1901 imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. 1901 imputed earings with signature are the average earnings for a man with that occupation and literacy status. OCCSCORE is the IPUMS imputed earnings which is based on 1950 US Census earnings (Minnesota Population Center 2019). Literate is an indicator variable that is one an individual signed their first marriage certificate and zero if they did not. Occupational literacy scores are the percent of men with that occupation in the 1890's in the vital records who could sign their name. Variable occupational literacy 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 reweighted 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 literacy alone; this is likely because the literacy rate is rising substantially over time and therefore it becomes less informative about socioeconomic status.

Figure 6 below computes the ratio measure used in Figure 3 but first estimates both correlations using an instrumental variable regression. The instrument used is the occupational status for the occupation second closest chronologically to the first marriage. This use of IV regression is a standard approach when there is classical measurement error, and it

requires no assumption about the relative magnitudes of attenuation bias on the numerator and denominator. The resulting measure is very similar.
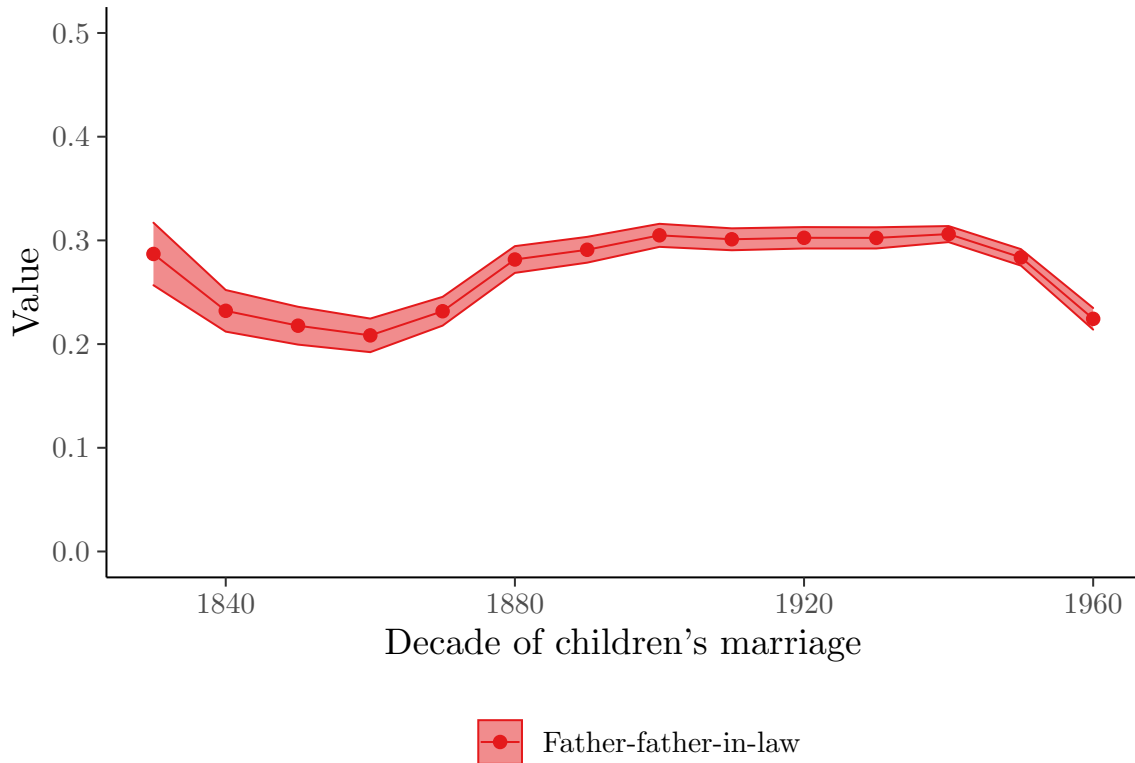

Figure 6: Estimated degree of marital assortment, IV
Note: 95% confidence intervals shaded. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. The estimates are ratios of rank-rank regression coefficients, which are equal to the correlation coefficients for the ranked variables assuming no ties in rank (and a reasonable approximation if not) (Chetty et al. 2014). To reduce attenuation bias, the dependent variable in the regressions are instrumented using a second measure of imputed earnings (using the second closest occupation the the individual's first marriage).

Figure 7 below estimates just the correlation between fathers and fathers-in-law using the same IV strategy mentioned above. This measure is more typically used in the literature, however it is only equivalent to the correlation between wives and husbands if the husband is matching to his father-in-law instead of to his spouse. As I demonstrate in Table 1, husbands and wives are at least partially assorted on ability. If matching was between a husband and

some weighted average of their wife and father-in-law, the true degree of assortment would be between this measure and the ratio measure.

Figure 7: Father-father-in-law correlation, IV
Note: 95% confidence intervals shaded. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. The estimates are rank-rank regression coefficients, which are equal to the correlation coefficients for the ranked variables assuming no ties in rank (and a reasonable approximation if not) (Chetty et al. 2014). To reduce attenuation bias, the dependent variable in the regressions are instrumented using a second measure of imputed earnings (using the second closest occupation the the individual's first marriage).

However, I can also directly test if the matching is between husbands and fathers-and-laws. For fathers-in-law who die before 1849, are the sons-in-law who married before the death different than those married after? As shown in Table 7, there appears to be no difference. In other words, if husbands are matching to their fathers-in-law, they don't seem to mind if their father-in-law is deceased before their marriage.

Table 7: Marriage matching appears not to have been between husbands and fathers-in-law

	Dependent variable:	
	Son-in-law signed	Son-in-law's log imp. earnings
	(1)	(2)
Married after father-in-law's death	−0.01	0.00
	(0.01)	(0.00)
Family FEs	X	X
Observations	83,988	147,641
Adjusted R^2	0.37	0.70

Note: *p<0.10; **p<0.05; ***p<0.01. Family-clustered standard errors in parentheses. Signature variables are indicators that are one if a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time. Deaths are only observed before 1849 and only sons-in-law of men who died before 1849 are included.

A3 Robustness of assortment on individual human capital

The estimates in Table 1 Columns 2 and 4 are identified using family fixed effects. This means that only families where one child signed and one child did not sign are driving the estimation results. The estimated coefficients are an average treatment effect of individuals in these "treated" families being able to sign. However, it is possible that these families have unusual characteristics; a more interesting average treatment effect is, perhaps, that for the entire 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.[13] However, regardless of assumptions

---

[13]The assumptions: 1. There is no selection into treatment within groups. 2. Conditional on observables,

it is still a useful exercise to see if the estimates are robust to reweighting.

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 married children in the family. Missing values are included as an additional category for each indicator variable. As shown in Table 8 below, there is still a positive and significant marriage premium for literacy.

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 8: Marriage selection, reweighting for selection into identification

	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	0.31***	0.20***	0.02***	0.03***	0.04***	0.05***
	(0.00)	(0.01)	(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
Observations	1,937,871	1,937,871	971,173	971,173	1,148,769	1,148,769
Panel B: Effect of husband's human capital						
Wife signed	0.28***	0.35***			0.03***	0.11***
	(0.00)	(0.01)			(0.00)	(0.00)
Re-weighted		X				X
Husband's family FE	X	X			X	X
Decade FE	X	X			X	X
Observations	1,928,239	1,928,239			982,166	982,166

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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order 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 married children in each family. Missing values are included as an additional category for each indicator variable in the logistic regression.

A4 Robustness of estimates of effects of parental human capital

The identifying assumption for the analysis in Table 3 is that the human capital of children of parents who remarry did not change over time faster than those of parents who do not remarry. Table 9 replicates the analysis except it drops children of parents who remarry if they were more than one sibling away from a half-sibling in the order of siblings.[14] This is

---

[14]As elsewhere, I order siblings by date of first marriage as I do not have birth dates for the entire sample.

analogous to restricting the sample to children born on either side of the remarriage, which should limit the importance of differential time trends. The results are very similar.

Table 9: Effect of parents, fixed effects with window

	Signed Daughter	Signed Son	Log imp. earnings Son-in-law	Son
	Dependent variable:			
	(1)	(2)	(3)	(4)
Panel A: Controlling for father				
Mother signed	0.02***	0.03***	0.01	0.03**
	(0.01)	(0.01)	(0.01)	(0.01)
Father FEs	X	X	X	X
Identifying observations	18,303	16,130	8,440	7,438
Observations	1,559,975	1,445,605	944,017	881,588
Adjusted R^2	0.68	0.67	0.37	0.41
Panel B: Controlling for mother				
Father signed	0.02***	0.03***	0.03*	0.04**
	(0.01)	(0.01)	(0.02)	(0.02)
Mother FEs	X	X	X	X
Identifying observations	6,486	5,500	2,799	2,289
Observations	1,522,830	1,411,662	921,530	860,897
Adjusted R^2	0.69	0.68	0.37	0.41

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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time. Children of parents who remarry are only included if they were in a window of one or less away from a half-sibling in the order of siblings.

Table 10 replicates Table 4 except it drops children of parents who remarry if they were more than two siblings away from that a half-sibling in the order of siblings. The logic is the same as above, except I widen the band as the regression compares siblings to full siblings. Again, the results are very similar.

Table 10: Effects of parents, half siblings with window

	Signed (Daughters)	Signed (Sons)	Log imp. earnings (Sons)-in-law	Log imp. earnings (Sons)
	Dependent variable: Younger sibling's characteristic			
	(1)	(2)	(3)	(4)
Panel A: Controlling for father				
Older sibling's characteristic	0.38***	0.36***	0.23***	0.26***
	(0.00)	(0.00)	(0.01)	(0.01)
Same mother	−0.00	−0.01**	−0.22***	−0.30***
	(0.00)	(0.00)	(0.06)	(0.05)
Signed × same mother	0.05***	0.05***	0.04***	0.05***
	(0.00)	(0.00)	(0.01)	(0.01)
Observations	1,953,016	1,853,707	800,284	756,645
Adjusted R^2	0.63	0.63	0.11	0.14
Panel B: Controlling for mother				
Older sibling's characteristic	0.36***	0.33***	0.24***	0.22***
	(0.01)	(0.01)	(0.02)	(0.02)
Same father	−0.04***	−0.04***	−0.14	−0.53***
	(0.01)	(0.01)	(0.11)	(0.11)
Signed × same father	0.07***	0.08***	0.02	0.09***
	(0.01)	(0.01)	(0.02)	(0.02)
Observations	1,946,719	1,777,710	799,342	727,123
Adjusted R^2	0.64	0.63	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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Imputed earnings are the average earnings for men with the individual's occupation in Quebec in the 1901 Canadian Census sample. Fixed effects for decade, marriage number, and sibling order are included in every specification as the signature rate varies significantly across time. Children of parents who remarry are only included if they were in a window of two or less away from the nearest half-sibling in the order of siblings.

A5 Robustness of effect of assortment on intergenerational elasticity

One concern with Table 5 is that families where only one parent was literate were 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. 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.

The gender of children should be, at least at birth, 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 neighborhood or social network. As more of the sisters marry, the remaining sisters will have to be less picky. It is possible that older sisters have a different effect on younger sisters compared to older brothers.[15] However, if it merely changes the human capital of the younger sister, who then matches accordingly, it shouldn't introduce bias.

As shown in Table 11, 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.

---

[15]In fact, this is fairly likely. In preliminary research I have conducted for another project, I find that before 1849, the fraction of older siblings that are male increases the rate of infant mortality for younger sisters.

Table 11: 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	0.37***	
	(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	379,251	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 a signature was recorded, zero if the absence of a signature was recorded, and omitted otherwise. Earning scores are the average earnings for the individual's occupation in Quebec based on the 1901 Canadian Census sample. The fraction is the fraction of the mother's siblings who married before her that were female. Decade fixed effects are included in every specification as the signature rate varies significantly across time.