Chapter 2

Before the fall: Child quantity and quality in pre-demographic transition Quebec

2.1 Introduction

Since Becker (1960) economists have posited a trade-off between the average human capital ("quality") and the quantity of children. Growth theorists often emphasize the importance of such trade-off for the transition to modern economic growth (Galor 2005). In these theories, the fall in fertility during the demographic transition and the growth of technology during the industrial revolution are related parts of a positive feedback loop that led to rapid economic growth. However, few estimates of the trade-off exist for pre-demographic transition populations (Clark and Cummins 2018, Klemp and Weisdorf 2018, Tan 2018, Galor and Klemp 2019). If this trade-off was key to modern growth, then was it present in populations on the eve of industrialization?

This paper estimates the quantity-quality trade-off using the IMPQ vital records from

1620–1849.[1] Over 8,000 pairs of twins, an unusually large sample for a historical study, provide a source of exogenous variation in family sizes. Using twins, it finds a trade-off, albeit a small one. One additional birth on average decreases the probability that a sibling signs their marriage record, a proxy for literacy, by 0.6 percentage points.

This estimate is small — seemingly too small for the trade-off to permit rapid increases in aggregate human capital — yet is statistically significant and robust. As the vital records are linked over multiple generations, I show that the effect of quantity on quality does strongly persist between generations. However, as I illustrate with a simple simulation exercise, it is so small that even a large decrease in fertility would result in only moderate gains in human capital.

The structure of the paper is as follows. First, I describe the data. Second, I describe the empirical methods and the baseline results. Third, I demonstrate the results are robust to various threats to identification. Finally, I conclude with a discussion of the broader implications of the finding for the literature on long run growth.

2.2 Data

2.2.1 Sources

While the IMPQ contains marriage records through the 20th century, in this chapter I rely on birth and death records which are only available for the entire province through 1849. The data is near to a complete record of the Catholic population with high quality linking of families together (Bourque 2011, Dillon et al. 2018).[2]

Using the birth records, I am able to identify twins by the exact day of birth.[3] This is an advantage of using vital records instead of census records; many censuses do not record

[1] IMPQ (2020), Project Balsac (2020), and PRDH (2020).

[2] While some Protestants are included, I drop them from the sample as the records are less complete.

[3] I define twins as children born within a day of each other. There are a tiny number of children reported born an implausible amount of time apparant, e.g. one month, but this is likely measurement error.

months of birth, let alone days (Tan 2018). I exclude any family I detect with any triplets or higher order births, as they both are extreme outliers and complicate the empirical strategy. Despite these strict restrictions, I am still able to identify 8,121 pairs of twins (2.7% of children born). Summary statistics, averaged separately for twins and for singletons, are presented in Table 2.1 below.

The main measure of human capital used in this paper is the presence of a signature on a marriage record. A 1678 ordinance required both the bride and the groom to sign their marriage records if able and the priest to record if they complied (Magnuson 1992). However, before 1800 the rate at which individuals were reported unable to sign varies substantially from year to year. Therefore, I define ability to sign as a variable that is 1 if the individual definitely signed and 0 otherwise. See Appendix 1 for further discussion.

Table 2.1: Summary statistics

Variable	Singleton, mean	Twin, mean	Singleton, N	Twin, N
Year of birth	1796	1794	598,384	16,243
Parity	6.04	6.73	598,384	16,243
Signed	0.10	0.08	239,991	2,882
Surv. to 1	0.80	0.53	598,384	16,243
Surv. to 14	0.68	0.44	576,500	15,765
N born	10.99	12.12	598,384	16,243
N surv. 1	8.70	8.97	596,553	16,199
N surv. 14	7.31	7.47	492,046	13,685
Mother's age at birth	30.64	32.22	598,384	16,243
Mother surv. 40	0.9	0.9	598,384	16,243
Mother signed	0.07	0.07	598,384	16,243
Father signed	0.07	0.07	598,215	16,243
Share of sibs signed	0.13	0.11	598,384	16,243

Note: Signed is an indicator which is one if the individual signed their first marriage certificate, zero if they did not or there is no record of a signature, and missing if they did not marry. Surivival to age one is inferred from either a missing death record and a birth more than one year before 1849 or from a death at an age greater than one. Surivival to age fourteen is defined similarly. N represents the count of siblings (and potentially half-siblings) that share a mother.

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 was particularly associated with business activity (Greer 1997). Literacy also allowed young men to become a lay tutor, a frequent stepping stone towards a career as an administrator or notary (Magnuson 1992). Another career choice that required literacy was the Church.[4] Moreover, for the marriages with a known occupation, it does appear that there is a fairly steep signature gradient across different occupations (Table 2.2).

Table 2.2: Occupations by average signature rate, marriage records

HISCO	Occupation	Translation	% of total	Share signed
41025	*Marchand*	Merchant	0.02	0.71
79100	*Tailleur*	Tailor	0.01	0.57
58340	*Soldat*	Soldier	0.02	0.41
77620	*Boulanger*	Baker	0.01	0.38
98135	*Navigateur*	Sailor	0.01	0.34
80110	*Cordonnier*	Shoemaker	0.02	0.34
95410	*Menuisier*	Carpenter	0.05	0.29
76145	*Tanneur*	Tanner	0.01	0.26
83110	*Forgeron*	Blacksmith	0.02	0.25
95135	*Macon*	Mason	0.01	0.21
61110	*Cultivateur*	Farmer	0.50	0.12
98620	*Charretier*	Carter	0.01	0.11
99910	*Journalier*	Worker	0.15	0.09
43220	*Voyageur*	Fur trader	0.01	0.07
62105	*Laboureur*	Laborer	0.04	0.04

Note: All marriages before 1850 with a definite location in Quebec are included. Only men are included in the sample as female occupational titles are rare. The most common title is taken for each HISCO category. % of total is the percent of all males with known occupations that are coded into that HISCO category. *Journalier* is an ambiguous category, as in Quebec it refers to a worker payed by the day regardless of the task or industry.

2.2.2 Comparison to France

Before its demographic transition, Quebec had very high levels of fertility compared to France. In Table 2.3 below, I compare the main dataset to French data from Louis Henry's survey of rural parishes (Henry 1968). The data is partitioned into women born before and after 1748, as the French demographic transition began during the Revolution (Cummins 2013).

Quebec's high fertility, as suggested by the table, was due to a marriage regime where

[4] Though not one I observe in marriage records, so average population-wide literacy is likely underestimated by the signature proxy.

women married younger than their peers in France. Quebec also had relatively low human capital, though the gender gap was also notably quite small. The decrease in literacy from the first to second period, while perhaps surprising, is documented in the historical literature (Greer 1985). The initial colonists were often drawn from urban areas, but their children frequently became *habitants* — Quebec's colonial equivalent of peasants — who neither had easy access to schooling nor a large economic incentive to learn to read and write.

Table 2.3: Comparison of married women who survive to 40, Quebec and France

Period	Country	Age, 1st mrrg	Age, hsbnd	N births	N surv. to 1	Signed	Hsbnd signed
Born 1636–1748	Quebec	22.6	27.5	9.17	7.08	0.10	0.12
	France	24.9	28.9	6.50	5.54	0.09	0.24
Born 1748–1803	Quebec	22.6	26.9	9.31	7.38	0.05	0.05
	France	25.5	29.0	5.05	4.42	0.21	0.44

Note: Sample consists of all women who married, had at least one child, never remarried, and survived to age 40. I only consider women who never remarried as in the Henry data the number of births is per couple, not per woman. I also drop the very few observations where either spouse has a negative age at marriage, as this is presumably due to errors in the records or digitization.

2.3 Methods and results

2.3.1 Identification from twins

There are two types of twins, monozygotic (identical) and dizygotic (fraternal). Monozygotic twins occur at a remarkably consistent rate across societies (around 0.7–0.9 percent of children).[5] The rate of dizygotic twins is more varied and is influenced by several maternal characteristics. The rate increases with a mother's age and previous number of births, and is higher for mothers who previously delivered twins. Some studies find other maternal characteristics associated with higher rates, an endogeneity concern which I address below (Farbmacher et al. 2018). Controlling for maternal age and parity, twins should therefore be

[5] In fact, this rate is observed in all mammals except some species of armadillos (Pison and Couvert 2004).

effectively random. This conditional randomness is key to my identification strategy.

2.3.2 Binned scatter plots

Twins are random conditional on both parity and mother's age at birth. As there are two variables on which to condition, plotting the underlying relationships in two dimensions is not straightforward. Below, I construct a series of four binned scatter plots that show the relationship between these two controls, the number of siblings born, and the average literacy of siblings. In each scatter plot, I hold one control constant using a fixed effects regression on the entire sample and then average the data, separately for twins and singletons, into twenty equal-sized bins over the other control variable.

In the scatter plots holding mother's age constant (Figure 2.1) and holding parity constant (Figure 2.2), there is a higher total number of births at every parity and mother's age for twins. The difference between twins and singletons seems to be the same regardless of parity or mother's age. For the share of siblings who signed, twins have a lower average, though the overall relationship is less clear. All together, these binned scatter plots suggest that twins will have the same effect on family size regardless of what mother's age or parity they occur at. Moreover, if twins are used as an instrument, the first stage is likely to be strong. The reduced form relationship of twinning on sibling literacy appears to be fairly noisy, although it is negative as the theories of the trade-off would predict.

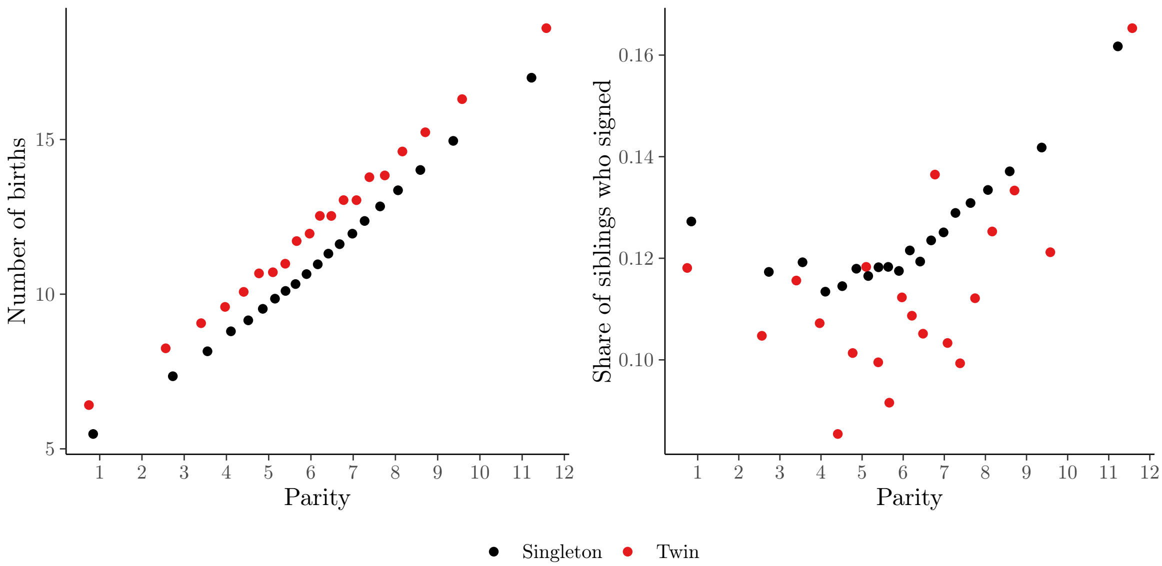

Figure 2.1: Binned scatter plots holding mother's age constant *Note:* All variables are first adjusted by regression on indicator variables for mother's age at birth. Adjusted variables are then computed as the residuals plus the estimated fixed effect for the age closest to the mean. Then, the data are averaged over twenty equal-sized parity bins, seperately for twins and singletons.

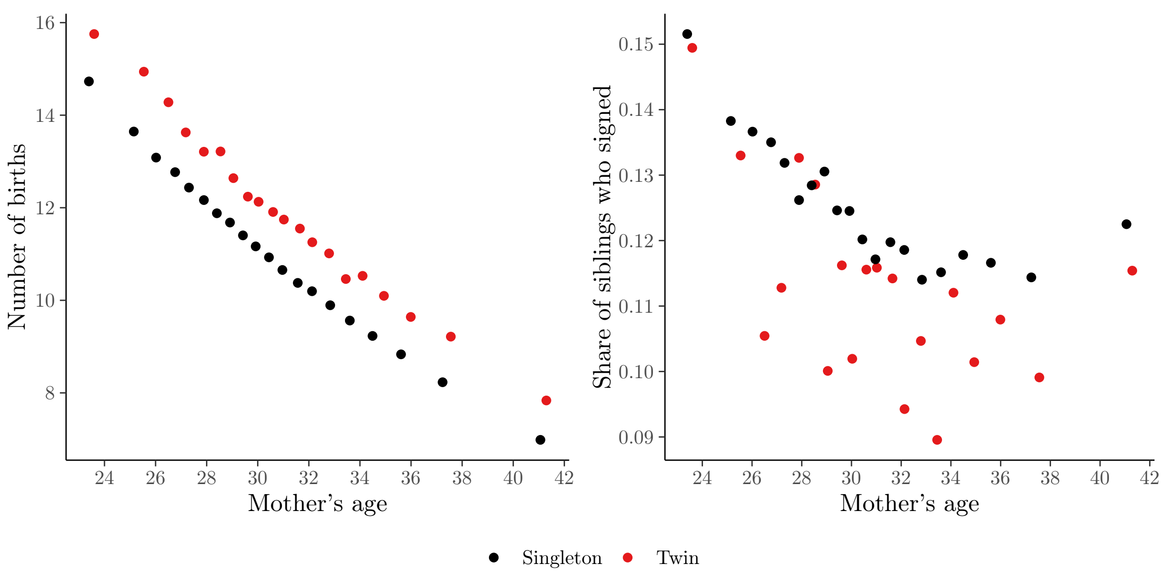

Figure 2.2: Binned scatter plots holding parity constant *Note:* All variables are first adjusted by regression on indicator variables for parity at the birth. Adjusted variables are then computed as the residuals plus the estimated fixed effect for the parity closest to the mean. Then, the data are averaged over twenty equal-sized mother's age bins, seperately for twins and singletons.

2.3.3 Empirical specification

Modern implementations of the twin instrument face different data constraints than those imposed by the historical Quebec data. In populations with parity-dependent control, twin births only increase a family size if they are the last birth planned. This introduces a complex empirical challenge. A twin birth is more likely to occur at a higher parity, but since families with a higher family size target will have a higher average parity, there is the potential for reverse causation. Moreover, the average effect of a twin birth on family size will depend both on parity and on the target family size of the parents. To attempt to address these issues, researchers will often use analysis by parity (c.f. Tan (2018)). This approach uses an instrument for a twin born at a parity n, considers only families with at least n children, and looks only at the effect on children born at parities less than n (Black et al. 2005a, Black et al. 2010, Angrist et al. 2010). As an alternative, the parity-pooled approach uses multiple instruments: a series of indicator variables for a twin birth at different parities (Angrist et al. 2010). This allows for a non-linear effect of twinning on family size depending on the parity of the birth at the cost of increasing potential weak-IV bias.[6]

These approaches are not well-suited to the Quebec data. Fortunately, they are also not necessary. With modern populations, the small average family sizes mean that splitting the sample by parity does not substantially reduce the effective sample size. In the Quebec data, average family sizes are much larger (Table 2.1), dramatically reducing the sample size available for each regression (or increasing the number and decreasing the strength of the instruments in the pooled approach). However, as shown in Clark et al. 2020 and in the binned scatter plots above, parents did not practice parity dependent fertility control and twins did not have a heterogenous effect on family size at different parities. Therefore, I do not use analysis by parity or the parity-pooled approach. Instead, I use a single indicator variable for a twin birth, regardless of parity, as my instrument and look at the effects on all

[6] As the indicator variable for twin at parity n is coded 0 for each birth at parity $m \neq n$, the instruments are weaker as well as more numerous.

other siblings. As twins are only random conditional on parity and mother's age, I control for both using fixed effects. In other words, I include twins and parity as covariates, but not the interaction between the two. This more parsimonious specification requires fewer observed twins at every parity.

The reduced form of the baseline regression is as follows. For each pair of child i and birth j sharing mother m, $i \neq j$, I regress:

$$signature_{i,m} = \alpha \; twin_{j,m} + \gamma_{mother\text{'}s \; age_{j,m}} + \delta_{j,m} + \beta X_{i,j,m} + \epsilon_{i,j,m} \tag{2.1}$$

where $signature_{i,m}$ is if the child definitely signed their marriage record, $twin_{j,m}$ is an indicator for whether birth j is a twin birth, $\gamma_{\text{mother's age}_{j,m}}$ are fixed effects for the mother's age at birth j, $\delta_{j,m}$ are fixed effects for the parity of birth j, $X_{i,j,m}$ is a vector of controls not required for identification, and $\epsilon_{i,j,m}$ is the error term. In most of the regressions, for $X_{i,j,m}$ I only include fixed effects for the year of birth of i to control for trends in literacy over time.

In other words, I compare two mothers, one who gave birth to twins at a given parity and age, and one that gave birth to a singleton at the same parity and age. I then look at each other child of the mothers. Do the children of the twin family have a lower probability of signing than those of the singleton family?

The IV regressions are set up the same way, except various measures of family size are used as endogenous independent variables and the twin indicator is used as an instrument.

2.3.4 Main results

As shown in Table 2.4, one additional birth reduces the probability a child signs their marriage certificate by 0.6 percentage points. This estimate is statistically significant, but the magnitude of the effect is quite small. As shown in Figure 2.3, the trade-off explains little of the observed differences in the average literacy of children of fathers of different occupations.

The number of children surviving to one and fourteen, instrumented by the twin indicator, reduce the probability of signing by 2.2 percentage points and 2.7 percentage points respectively (Table 2.5 and Table 2.6). Comparing first stage Kleibergen-Paap F-stats, the instrument is notably weaker than in the number of births regression (Kleibergen and Paap 2006). This is because, as shown in Table 2.1, twins are less likely to survive childhood.

Table 2.4: Effect of number of siblings born on signature rates

	OLS	IV	1st stage	Reduced form
Number of births	0.001***	−0.006***		
	(0.000)	(0.001)		
Twin birth			1.127***	−0.006***
			(0.013)	(0.002)
N	2,298,174	2,298,174	2,298,174	2,298,174
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
KP F−stat			7,493	

Note: *p<0.10; **p<0.05; ***p<0.01. The dependent variable in each regression is an indicator which is one if the sibling signed their first marriage certificate, zero if they did not or there is no record of a signature, and missing if they did not marry. To observe completed family sizes, the sample is restricted to mothers born before 1810.

Table 2.5: Effect of number of siblings surviving to 1 on signature rates

	OLS	IV	1st stage	Reduced form
Number surv. to 1	−0.006***	−0.022***		
	(0.000)	(0.006)		
Twin birth			0.291***	−0.006***
			(0.015)	(0.002)
N	2,295,644	2,295,644	2,295,644	2,295,644
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
KP F−stat			383	

Note: *p<0.10; **p<0.05; ***p<0.01. The dependent variable in each regression is an indicator which is one if the sibling signed their first marriage certificate, zero if they did not or there is no record of a signature, and missing if they did not marry. Surivival to age one is inferred from either a missing death record and a birth more than one year before 1849 or from a death at an age greater than one. To observe completed family sizes, the sample is restricted to mothers born before 1809.

Table 2.6: **Effect of number of siblings surviving to 14 on signature rates**

	OLS	IV	1st stage	Reduced form
Number surv. to 14	−0.008***	−0.027***		
	(0.000)	(0.009)		
Twin birth			0.181***	−0.005***
			(0.015)	(0.002)
N	2,119,381	2,119,381	2,119,381	2,119,381
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
KP F−stat			139	

Note: *p<0.10; **p<0.05; ***p<0.01. The dependent variable in each regression is an indicator which is one if the sibling signed their first marriage certificate, zero if they did not or there is no record of a signature, and missing if they did not marry. Surivival to age fourteen is inferred from either a missing death record and a birth more than fourteen years before 1849 or from a death at an age greater than fourteen. To observe completed family sizes, the sample is restricted to mothers born before 1786.

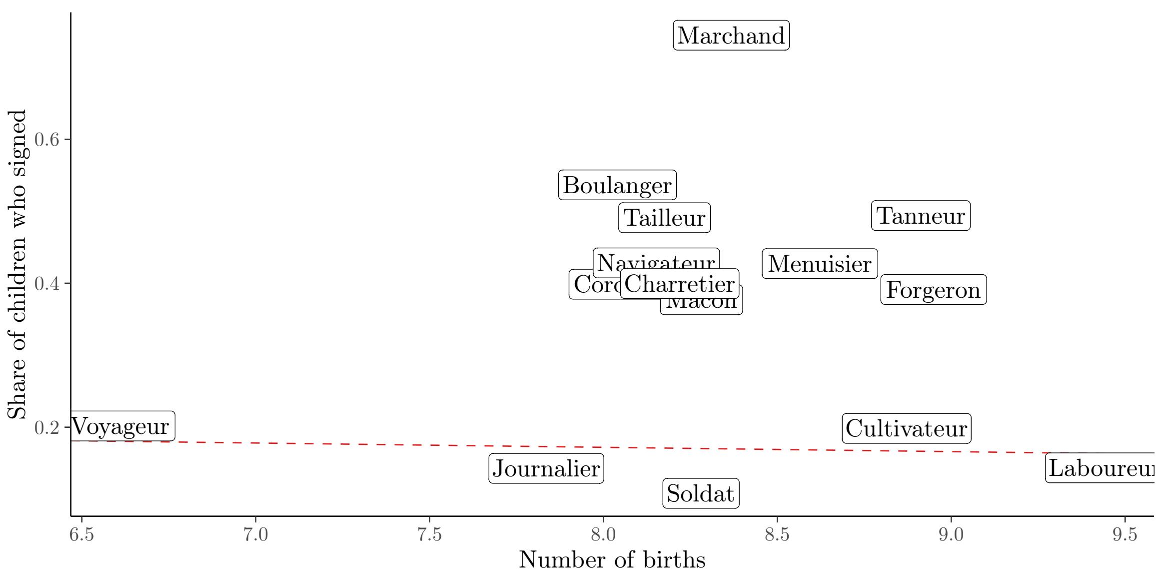

Figure 2.3: Quantity and quality for most common occupations *Note:* This figure plots the average child signature rate and number of births per mother by the occupation of her first husband. The occupations are from Table 2.2 and the sample is restricted to women born before 1849 with at least one married children. To illustrate how the estimated quantity-quality trade-off is small in magnitude, the red dashed line is a line with a slope of −0.006 drawn through the average quantity and quality for this sample.

2.3.5 Intergenerational-effects of the trade-off

Table 2.7 shows the effects on the second generation. Assuming the exclusion restriction holds — that is twins effect the quantity and quality of a second generation only through increasing quantity in the first— then one additional birth increases the average family size of siblings by 0.095 births and decreases the average signature rate of their children by 0.5 percentage points. Accounting for the small increase in fertility, this implies that 92% of the direct effect on signature rates persists to the next generation.

Does this change in human capital have a possible spillover effect beyond the direct descendants? After all, growth theories often assume parents decide between child quantity and quality depending on the rate of technological progress, which in turn depends on aggregate human capital (Galor and Weil 2000). As show in Column 4, for this sample there is no significant interaction between the average human capital in the borough and decade of the twin birth and the size of the trade-off. Note that while this regression provides a causal estimate of the direct effects of a twin birth, the interaction term does not have a causal interpretation. If there was a significant and negative interaction term, the trade-off would in fact be larger in boroughs and decades with relatively higher aggregate human capital. This would not be proof that this steeper trade-off was caused by the higher aggregate human capital. Regardless, I find no evidence of such an interaction. It is also plausible that local aggregate human capital might be of little importance compared to national or even global aggregate human capital. However, without controlling for unrelated time trends with a panel fixed effects regression, it is even less likely that a correlation represents a causal link between aggregate human capital and the magnitude of the trade-off. Overall, while I can't conclusively rule it out, I find no evidence that the trade-off is higher as aggregate human capital increases. Therefore, any increase in human capital from a decrease in fertility would plausibly have no spillovers in the form of a steeper trade-off for subsequent generations.

Together, these estimates allow a simulation of the dynamic effects of reductions in fertility rates. While much simpler than a full structural model of the trade-off, these simulations give a rough illustration on just how small the effects are. Assume that the estimates from Table 2.7 are correct and that the percent of the effect inherited by the next generation is always constant. Permanently decreasing fertility from n_0 to a constant n_1 would have a net effect of $-0.005(n_1 - n_0)\frac{(1-0.92^g)}{(1-0.92)}$ on signature rates after g generations.[7]

The results of this simple simulation (Figure 2.4 below) show that a permanent decrease in fertility to French levels for women born 1636–1748 from Table 2.3 would allow Quebec to close the signature rate gap in 5 generations. As 5 generations is roughly 150 years, this is not a fast convergence. Moreover, the French average of 16.5% is a relatively low bar; five generations from those born in 1748 would be roughly in the 1920's, which as shown in Appendix A1 was a period of near universal literacy.[8]

For the 1748–1803 period, convergence is even slower.

Table 2.7: Estimates of the dynamics of the trade-off

	Sib signed	Sib's N births	Share of sib's children signed	Sibs signed
N births	−0.005***	0.080**	−0.005**	
	(0.002)	(0.037)	(0.002)	
Twin birth				−0.005**
				(0.002)
Lit. rate				0.026***
				(0.006)
Twin birth × Lit. rate				0.025
				(0.028)
N	855,861	855,861	855,861	1,225,871
FE: Boro. of birth				X
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X

Note: *p<0.10; **p<0.05; ***p<0.01. The first three columns consider intergenerational effects. The sample is children of mothers born before 1760 with at least one child of their own. The first dependent variable is the same signature variable as before. The second is the number of births the sibling had. The third is the average signature rate of the sibling's own children. The third column has the same sample and dependent variable as the baseline regression in Table 2.4. Lit. rate is the average signature rate of first marriages in the borough and decade of the potential twin birth.

[7] Note that while is true that the estimates above imply that decreasing fertility from n_0 to a n_1 would also decrease fertility in the next generation by $0.08(n_1 - n_0)$, it is a decrease in reference to n_0 not n_1. Therefore, I instead simply assume an exogenous permanent decrease in fertility to n_1.

[8] And, ironically, when the demographic transition began to reach a substantial share of Quebec's families (Vézina et al. 2014).

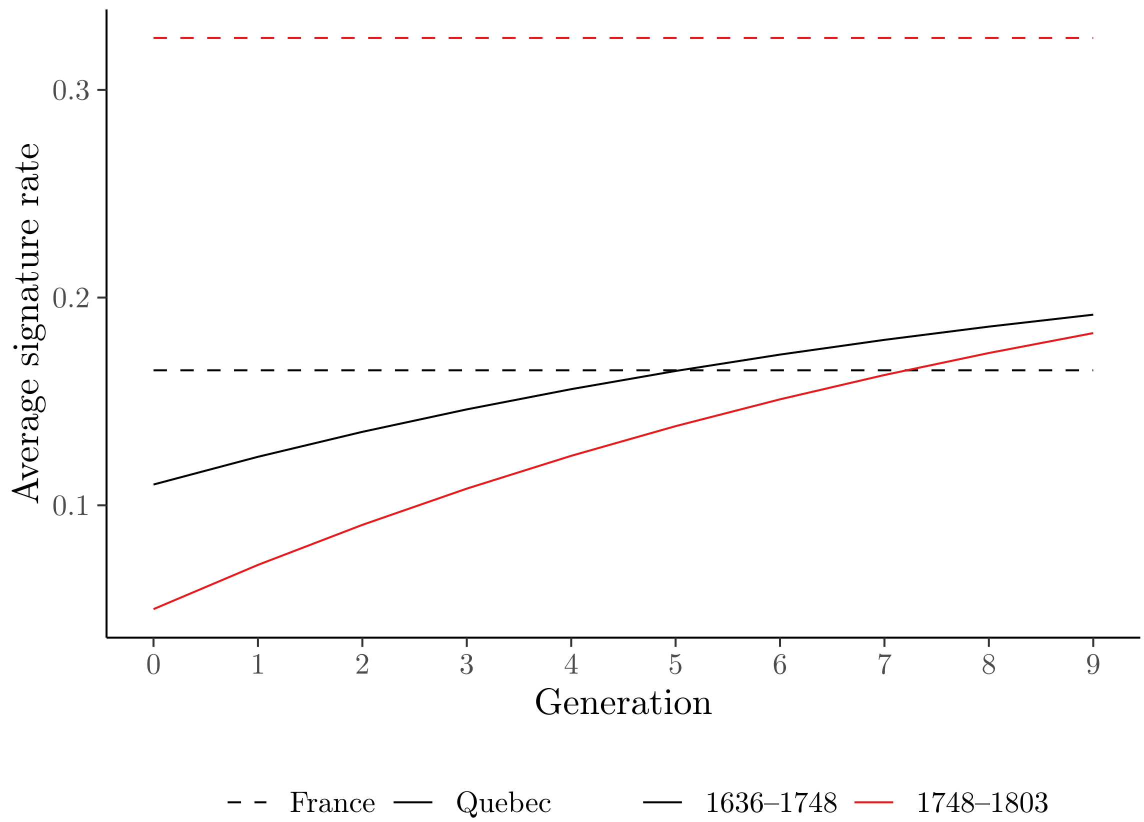

Figure 2.4: Simulation of reducing Quebec fertility to French rates *Note:* The simulation uses the values from Table 2.3 and Table 2.7 to roughly illustrate how the estimated quantity-quality trade-off is small in magnitude. Even with a decrease in births to French levels and an intergenerational persistence of 92% of the effect from one generation to the next, the accumulation of aggregate human capital is very slow.

2.4 Robustness

2.4.1 Alternative specifications

Below, I demonstrate that finding of a small yet statistically significant trade-off is robust to various tests of the empirical framework. The first test omits opposite-sex twins (Table 2.8, Column 2). As monozygotic twins are always same-sex, this restriction increases the share of twins that are monozygotic (Farbmacher et al. 2018). Therefore, any endogeneity bias from the probability of dizygotic twinning is reduced. The estimate is very similar. Moreover, including various parental and sibling characteristics as controls to the baseline specification also does not substantially change the results (Table 2.8, Column 3). These tests suggesting that there is little concern of endogeneity bias from non-random twinning after controlling for parity and mother's age.

Twins are not quite the same as one additional singleton birth. First, twin births are more likely to cause complications in childbirth, potentially leading to maternal death. Dropping mothers who died before age 40 (Table 2.8, Column 4) does not change the result. Second, twins result in atypical spacing of children over time (Black et al. 2005a, Black et al. 2010). If birth order has a different effect on outcomes than child order, then twins have a different effect on the following children than two singleton births. Restricting the sample to just older siblings addresses this concern. Moreover, it also rules out an effect on later siblings from parents altering spacing in response to twins. This restriction does not substantially change the result (Table 2.9, Column 2). Families in Quebec are also quite large, suggesting that siblings sufficiently older or younger than the twin birth might not be affected. Restricting the sample to only siblings within two births of the potential twin birth also does not change the result (Table 2.9, Column 3). Finally, although each mother contributes multiple observations to the regression, clustering the standard errors by mother only modestly increases the standard errors (Table 2.9, Column 4).

Twin infants are typically of below average health (see Table 2.1). While I primarily look at the outcomes of non-twin siblings, it is possible that parents either under-invested in sickly twins or compensated them with additional resources (Rosenzweig and Zhang 2009). To address this concern, I estimate the trade-off using two other instrumental variables.

Table 2.8: Alternative specifications, part 1

	Baseline	Same−sex twins	Extra controls	Mothers surv.
Number of births	−0.006***	−0.005***	−0.005***	−0.005***
	(0.001)	(0.002)	(0.001)	(0.001)
N	2,298,174	2,284,403	2,297,613	2,125,558
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
1st stage KP F−stat	7,493	4,244	7,976	8,680

Note: *p<0.10; **p<0.05; ***p<0.01. The first column is the baseline result from Table 2.4. The second column drops all opposite-sex twins, reducing any bias from non-random dizygotic twinning. The third column adds several control variables (the signature variable for both parents, the gender of the child, and the parity of the child). The fourth column restricts the sample to mothers who were not recorded dying before age 40.

Table 2.9: Alternative specifications, part 2

	Baseline	Older sibs	Sib parity within 2	Clustered SE
Number of births	−0.006***	−0.005***	−0.006**	−0.006**
	(0.001)	(0.002)	(0.002)	(0.002)
N	2,298,174	1,316,500	834,267	2,298,174
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
1st stage KP F−stat	7,493	5,705	2,269	1,145

Note: *p<0.10; **p<0.05; ***p<0.01. The first column is the baseline result from Table 2.4. The second column restricts the sample to only children born before the potential twin birth. The third column restricts the sample to children born within two births of the potential twin birth. The final column clusters the standard errors by mother since each mother has multiple associated observations.

2.4.2 Alternative instruments

The twin instrument might suffer from bias due to the relatively low health endowments of twin children. If parents reallocate resources towards children with higher birth endowments (reinforcement) or away from them (compensation), the IV estimates will be biased. I explore this potential source of bias by using two additional instruments for family size. Both have more concerning potential challenges to identification that the twin instrument. However, these potential challenges are different, and the results from these instruments corroborate those from the twin instrument.

For a second (novel) instrument, I argue that the province-wide infant mortality rate during the year a younger child was born is exogenous to individual family characteristics. I again compare two mothers with the same age and parity, one who gave birth during a year with relatively high infant mortality rate. I then look at other children from both families born in the same year in order to control for both their disease exposure and aggregate trends in literacy rates. For this regression, the spacing of subsequent births will be changed if the instrument has a valid first stage; infant mortality decreases birth spacing. There is also potentially higher risk of maternal mortality in high infant mortality years. Therefore, I restrict the sample to just mothers who do not die before age 40 and older siblings.

As shown in Table 2.10, both twins and children born in high mortality years have on average lower literacy rates than their siblings. As one shock increases and one shock decreases family size but both result in children born with lower endowments, reinforcement or compensation would bias the estimates in different directions. The IV estimates from both measures are in fact quite similar, despite the infant mortality rate being a weaker instrument and proving a less precise estimate (Table 2.11). Assuming both IV's are otherwise valid, this suggests that bias from compensation or reinforcement is not a major concern.

Table 2.10: Effect of twinning and infant mortality rates on signatures

	Self, IMR	Self, twin	Siblings, IMR	Siblings, twin
IMR	-0.779^{***}		0.021^{*}	
	(0.019)		(0.011)	
Twin birth		-0.019^{***}		-0.006^{***}
		(0.005)		(0.002)
N	230,415	259,187	1,215,119	2,295,644
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth			X	X

Note: *p<0.10; **p<0.05; ***p<0.01. The dependent variable in columns 1 and 2 is the signature variable of the children born. The sample is the same as in Table 2.5. The dependent variable in columns 3 and 4 is the signature variable for other children who share a mother. The sample is restricted to both children born before the birth and mothers who were not recorded dying before age 40. IMR is the aggregate infant mortality rate for the entire province of Quebec during the year of the birth.

Table 2.11: Alternative instrument: infant mortality rates

	IV, twin	IV	1st stage	Reduced form
Number surv. to 1	−0.022***	−0.028*		
	(0.006)	(0.015)		
IMR			−0.762***	0.021*
			(0.091)	(0.011)
N	2,295,644	1,215,119	1,215,119	1,215,119
FE: Mother's age	X	X	X	X
FE: Parity	X	X	X	X
FE: Year of sib's birth	X	X	X	X
KP F−stat			70	

Note: *p<0.10; **p<0.05; ***p<0.01. Column 1 is from Table 2.5. In columns 2–4, the sample is restricted to both children born before the birth and mothers who were not recorded dying before age 40. IMR is the aggregate infant mortality rate for the entire province of Quebec during the year of the birth.

For the third instrument, I use the protogenesic interval, the time between the mother's first marriage and first birth (Klemp and Weisdorf 2018, Galor and Klemp 2019). This instrument is a measure of fecundity (i.e. potential fertility) as it captures biological variation in ability to conceive. While this instrument can be used in samples too small to effectively use twin births, it also has some less desirable properties. First, it makes a very strong assumption: that there was no premarital conception. If a couple conceived before marriage, the protogenesic interval will be too small. If premarital conception is more likely for some parents than others, this introduces endogeneity bias. Second, even absent premarital conception, the instrument's validity is uncertain. Various factors influence conception odds, such as maternal health, nutrition, and age. While age can be controlled for, there is still a serious possibility that socioeconomic status is somehow correlated with conception chances. This is why previous studies using the protogenesic interval rely on other control variables for identification.[9]

Despite the concerns related to its exclusion restriction, the protogenesic interval instrument gives a very similar estimate to the twin instrument (Table 2.12). It also would not

[9]For example, Galor and Klemp (2019) use lineage head fixed effects, arguing they control for non-random biological factors. Note that their regressions control for age irrespective of gender, which potentially leaves the instrument invalid. Female age at first marriage almost certainly has a different non-random effect on PI than male age at first marriage. The PI instrument, while potentially much stronger in a small sample, has much more serious endogeneity concerns than the twin instrument.

have bias from reinforcement or compensation, again suggesting that bias from compensation or reinforcement is not a major concern.

Table 2.12: Alternative instrument: protogenesic interval

	IV, twin	IV	1st stage	Reduced form
Number of births	−0.006***	−0.004***		
	(0.001)	(0.001)		
PI			−0.686***	0.003***
			(0.013)	(0.001)
N	2,298,174	221,309	221,309	221,309
FE: Mother's age	X			
FE: Parity	X			
FE: Year of sib's birth	X	X	X	X
KP F−stat			2,710	

Note: *p<0.10; **p<0.05; ***p<0.01. Column 1 is from Table 2.4. Columns 2–4 look at children as the unit of analysis, not child-birth pairs. PI is the protogenesic interval of the mother, the time between her first marriage and her first birth. PI is not arbitrarily trimmed as in Galor and Klemp (2019), leaving values with implausibly short gestation periods. I argue this is better than arbitrarily censoring implausible values, as that only omits premarital conceptions with short PIs and leaves premarital conceptions with longer PIs in the sample, introducing potential sample selection bias.

Together, the estimates from alternative instruments suggest that bias from compensation or reinforcement is not a major concern in this population. It certainly does not imply, however, that in other populations twin instruments would not suffer from compensation or reinforcement. The relatively egalitarian treatment of children with lower or higher birth endowments could be particular to Quebec.

2.5 Discussion

In pre-demographic transition Quebec, the natural experiment of twins provides evidence that there was a small trade-off between family size and the average human capital of children. The estimated trade-off — a decrease in the odds a sibling signed their marriage record of 0.6 percentage points per additional child born — is statistically significant and robust to various tests of potential threats to identification. Using multigenerational linkages, I show that although the effect of an increase in family sizes seems to be strongly inherited, the magnitude of the trade-off is too small to be of major economic significance.

Why was the trade-off so low? One explanation for the lack of substantial trade-off in contemporary populations is the availability of public education (Angrist et al. 2010). At one extreme, if education is free, there is no trade-off. At the other extreme, if education is prohibitively expensive, an additional child makes no difference. Before the 1840's, Quebec's formal education system was almost entirely provided by the Catholic Church. Various religious orders and parish schools provided a range of educational services.[10] While tuition was often free, food and board was not, and there was always the opportunity cost in terms of forgone child labor (Magnuson 1992). Perhaps the small trade-off was a product of Quebec's limited public education during the period. Then again, evidence from rural France during a similar period suggests that fertility fell before education rose, which in turn occurred before industrialization (Blanc and Wacziarg 2020).

What, then, do we learn about theories of long-run economic growth? Quebec had already began industrialization by the 1830's, with modern water-powered factories emerging after the Lachine Canal was enlarged in the 1840's and a major railway with the Grand Trunk in the 1860's. (Courville et al. 2006, Bradbury 2003).[11] Montreal emerged as a major industrial center in the second half of the 19th century. Moreover, Quebec's demographic transition occurred substantially later, only reaching substantial numbers of French-speaking Québécois by the 1920's (Vézina et al. 2014).

It appears that the existence of a substantial quantity-quality trade-off is not a necessary condition for industrialization. This does not, however, necessarily contradict the theories that place it at the heart of modern economic growth. These theories allow that regions importing preexisting technologies from an already industrialized area might have different dynamics (Galor and Weil 2000). Perhaps the important trade-off was that faced by families in London or Boston, not Montreal or Quebec City. If so, the so-called "Western

[10] The female Congrégation de Notre-Dame was particularly active. In general, nuns were more focused on education, though there were a few schools run by male orders such as the Recollects (Greer 1997).

[11] The marriage records aren't particularly suited to detecting industrialization because they list only occupations, not industries. The first mechanic shows up in the marriage records in 1818, machine maker in 1832, railway worker in 1855.

Offshoots" (the United States, Canada, Australia and New Zealand), considered to be on the growth frontier alongside Western Europe, should be revised to exclude Quebec (Galor 2005). Alternatively, Quebec could be added alongside France (with its inconveniently early demographic transition) as an example that however elegantly a theory unifies economic growth with demography, the empirical evidence eludes a straightforward grand narrative.

Appendix

A1 Signatures and literacy

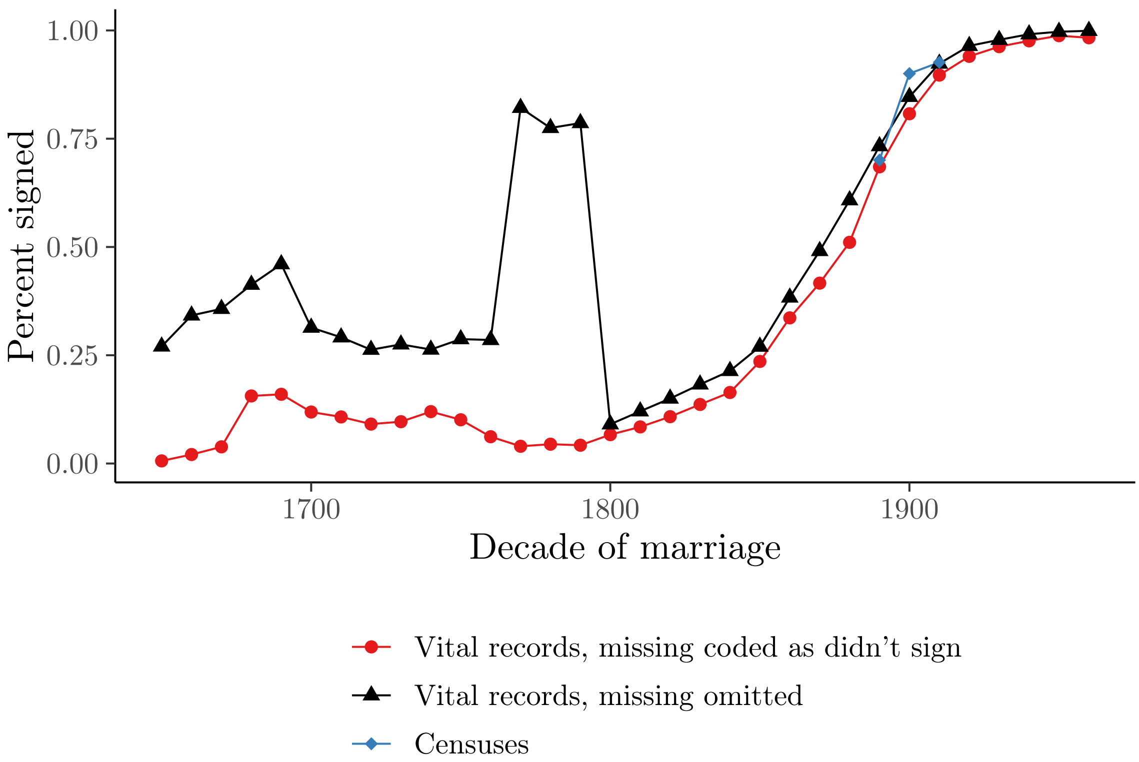

Figure 2.5: Different ways to code the signature variable *Note:* The vital record literacy rate is the average of an indicator variable that is one if a signature was recorded, zero if the absence of a signature was recorded, and either zero or omitted otherwise. The census record literacy rate is the fraction of individuals who were reported as able to write, reweighted to match the age distribution in the vital records.

Figure 2.5 above compares two different ways of coding the signature variable. Four extracts of Canadian censuses 1891–1911 provide external points of comparison.[12] After 1800, coding missing signatures as missing values provides an estimate of literacy closer to that in the censuses. Before 1800, coding the missing signatures as 0 provides a more stable measure over time.

[12]The 5% 1891 sample (Inwood and Jack 2011), the 5% 1901 sample and 1901 oversample (Canadian Families Project 2002), and the 5% 1911 sample (Gaffield et al. 2009). Data provided by the Minnesota Population Center (2019).