Horizontal vs. Vertical Transmission of Fertility Preferences

Andrea Di Miceli*†

Preliminary and Incomplete, please do not cite or distribute without permission
This Version: April 6, 2015

Abstract

I study the cultural transmission of fertility preferences among second generation immigrant women observed in U.S. Censuses from 1910 to 1970. As hypothesized by [Bisin & Verdier, 2001], the transmission of preferences can be "vertical" or "horizontal". Using a unique source documenting the variation in fertility behavior in Europe before and after the first demographic transition (1830-1970), I unpack the influence of parents (measured by source-country fertility at the time of departure from Europe) versus the influence of peers (measured by fertility of the same-age cohorts living in the source country and transmitted by same-age recent immigrants). I find that the transmission mechanism is crucially affected by the number of foreign born immigrant peers living in the same MSA. On one hand, the "vertical" channel of transmission is stronger in places where there are few newly arrived foreign born immigrant couples from the same source countries. On the other hand, fertility choices of second generation women are strongly correlated with marital fertility choices measured over peer cohorts in the source countries whenever they live in MSAs densely populated by recently arrived immigrants.

JEL-Classification: J13, Z10, Z13, N30

*UCLA Anderson School of Management, CCPR, e-mail: andrea.di.miceli.1@anderson.ucla.edu

†I am particularly grateful to Sascha Becker, Bastien Chabé-Ferret, Christian Dippel, Paola Giuliano, Erasmo Papagni and Romain Wacziarg for several comments and suggestions that substantially improved the paper. I also thank seminar participants at UCLA Anderson School of Management. All remaining errors are my own.

1 Introduction

During the last decade, economists have shown growing interest in the effect of culture on outcomes, social norms and traits such as living arrangements, labor force participation, level of trust and fertility decisions just to name a few.[1] However, the mechanism through which culture is transmitted is still a black box, for culture can be very "sticky" or rapidly evolving according to the social norm of interest (see Giavazzi, Petkov & Schiantarelli [2014] for a recent discussion). Following the seminal contribution of Bisin & Verdier [2001], papers documenting the persistent effect of culture have generally not been able to distinguish between the transmission channels through which the persistence in social norms occurs. In fact, Bisin & Verdier [2001] mention two distinct channels: the vertical one, that is, from parents to children, and the oblique-horizontal (from peers, henceforth, I will simply refer to the latter channel as to the horizontal one). In their model parents exhibit "imperfect empathy": their utility function is affected by children's choice to pick up one of the two social norms in the society.[2] Since parents' utility is increasing in their own social norm, they are willing to incur a socialization cost in order to maximize the probability of children acquiring the parental social norm. Moreover, the socialization effort exerted by the parents decreases the larger its social group as the two channels substitute each other.

Previous empirical work studying the transmission of social norms across second generation immigrants has generally taken different approaches to measure the transmission of social norms among second generation immigrants.[3] The most popular of these is the epidemiological approach which I adopt in what follows. According to this strategy, the key explanatory variable capturing vertical transmission of preferences should be measured in the country of origin of the parents (henceforth *source country* in this article) to reflect the prevailing social norm of interest.[4] However, a careful analysis of the previous literature shows that the choice of such variable has generally been limited by data availability. For instance, Alesina & Giuliano [2010], Algan & Cahuc [2010] and Ljunge [2014] use the same World Value Survey's waves to obtain dependent as well as independent variables when using the epidemiological approach. This choice is problematic as it does not take into account that habits and social norms might change across different generations living in the same country, more importantly it does not shed light on the transmission channel through which cultural norms persist.

My approach takes one step forward, focusing on individual fertility decisions of second generation married women living in the U.S. between 1910 and 1970, I perform a "horse race" between the two

[1] See [Guiso, Sapienza & Zingales, 2011] as well as [Guiso, Sapienza & Zingales, 2006] and [Fernández, 2011] for a thorough review on the effect of culture on several outcomes.

[2] In their model Bisin & Verdier [2001] assume that a monoparental family has one child which is born without one of the two existing cultural traits $\{a, b\}$. Parent chooses the optimal socialization effort level having perfect information of how many individuals in the population share its social norm.

[3] Henceforth in this paper, second generation immigrants are defined as U.S. born with at least one foreign born parent. First generation immigrants, i.e. those immigrants that were born outside the U.S., are here called foreign-born immigrants.

[4] See Fernández [2011] for a review of the advantages and drawbacks of this method. Other methodologies include the dummy variable approach (see Giuliano [2007] for an example), others have approximated the social norm simply averaging across migrants' population (Borjas [1995, 1992] Card, DiNardo & Estes [2000] provide examples of such studies).

transmission channels of preferences in Bisin & Verdier [2001] that are: vertical and horizontal (that is from foreign-born peers that migrated from the same source countries). Since fertility in the U.S., and in immigrants' source countries, changed considerably during this time window, there is sufficient longitudinal variation to investigate whether women's fertility choices are influenced by their parents' choices or by their peers from the same source country. The reasoning underlying the "horse race" is the following: in presence of vertical transmission, I expect the number of children of second generation women to be correlated with the marital fertility rates (hereinafter MFR) measured in their source countries at the time of migration of their parents. Therefore, lagged values of marital fertility rates in the source countries s (i.e. $MFR_{s,t-30}$) capture the vertical transmission. Measuring the horizontal channel is more challenging because of the reflection problem ([Manski, 1993]) one cannot plug in the MFR_t computed among peers in the U.S. at the time of the Census. Since, in the particular time window analyzed, fertility rates underwent sharp changes in the source countries the low autocorrelation of MFR enables the inclusion of both $MFR_{s,t}$ and $MFR_{s,t-30}$ in the same model. Therefore, $MFR_{s,t}$ measures the horizontal transmission occurring whenever second generation women socialize with recently arrived foreign-born immigrant peers from the same source country. Indeed, whenever this happens, I expect their observed fertility to be strongly correlated with their peers $MFR_{s,t}$ in the source country, this is what I call the "horizontal" channel of transmission. This strategy is feasible as I have a unique source documenting fertility decisions for almost one hundred years (i.e. from 1880 to 1970) in almost thirty European countries before and after the first fertility transition occurred Coale & Watkins [1986]. Although I cannot fully test for the extent to which women in the sample are exposed to the influence of peers from the same *source countries* over their lifetime, I can observe the fraction of peers living close to them at the time of the Census.[5]

In order to run the horse race, I use a pooled Negative Binomial model. While my results confirm past findings about the effect of cultural norms on family size (Fernández & Fogli [2009, 2006]), I find limited evidence of direct vertical transmission, that is, adding contemporaneous fertility among the set of explanatory variables causes lagged fertility to lose most of its statistical significance.[6] In line with the theoretical results found in Bisin & Verdier [2001], I find that the presence of foreign-born married couples within the same geographical area is strongly correlated with the transmission of fertility norms from the source countries. Specifically, the horizontal channel of transmission clearly wins the horse race in Metropolitan Statistical Areas (henceforth MSAs) populated by a large community of foreign-born peers. Therefore, second generation women living in MSAs that underwent continuous inflows of immigrants ended up having preferences that were closer to their peers in the source country rather than their parents' ones. Since I do not observe where the women in the sample were born and lived before

[5]Although the longitudinal dimension allows me to control for a set of fixed effects which purge my estimation from time-invariant unobservable characteristics, internal validity is still challenged by the potential presence of time-varying unobservables. Previous studies relying on cross-sectional data such as: Fernández & Fogli [2009, 2006], Alesina & Giuliano [2010], Ljunge [2014] are potentially affected by the presence of country-specific unobservables.

[6]The persistent effect of fertility preferences is such that an increase of one child in the *source country* marital fertility rate is associated with an increase by a factor of 1.07 in the number of children a second generation woman had.

filing the Census, I cannot completely rule out that my results are driven by self-selection of immigrants into areas with a high (or low) density of foreign-born immigrants. If this is the case, my estimates are likely to be an upper bound of the horizontally transmitted cultural effect.

The rest of the paper is organized as follows: section 2 reviews the literature about cultural transmission of preferences with a special focus on the studies looking at second generation migrants and using the epidemiological approach. Section 3 describes Coale & Watkins [1986] data on fertility. Moreover, this section explains how individual data on married couples was chosen for Censuses from 1910 to 1970. Section 4 explains the identification strategy adopted together with its advantages and drawbacks with respect to what has been done in the past. Section 5 shows the results of the pooled Negative Binomial estimation and suggests a potential channel through which the transmission of preferences observed in the data occurred. Section 6 concludes.

2 Literature Review

Previous work attempting to identify the role of culture on a set of diverse outcomes has used migrants and, more often, their children.[7] Guiso, Sapienza & Zingales [2004] were among the first to use migrants data to show that, within Italy, variation in the level of social capital had a causal impact on the use of formal credit and checks. However, differences in choices among foreign born migrants might reflect an "endowment effect", that is, they might be partially caused by early life experiences such as growing up in places with different institutional environments. In order to address this criticism, in a series of original articles Fernández & Fogli [2006, 2009] analyzed fertility choices and Labor Force Participation of second generation women in the U.S. Indeed, differently from their parents, migrants' children who were born and raised within the identical institutional environment of a single country, represent the ideal individuals on which is possible to test the persistence of preferences inherited from their parents. In a series of path breaking articles, the authors showed that fertility and labor force participation (henceforth abbreviated with LFP) measured in the 50's in the *source country* explains the variation in preferences for the number of children as well as for LFP's decision of second generation's migrants women.[8] This attempt to single out "cultural" from "environmental" beliefs using a variable measured in the country of origin of the parents is called epidemiological approach and has now been adopted widely in economics.[9] At the same time of Fernández & Fogli [2006, 2009] other articles showed that the heterogeneity in outcomes and choices of second generation's migrants within the U.S. is accounted for by the variation at the parents' country of origin level. Giuliano [2007] for instance, shows that important decisions such as the living arrangements of second generation's migrants in 1970's and 2000's are highly correlated with the ones in place in the country of origin of the parents. Similarly, Alesina

[7]Guiso et al. [2006] define culture as *the set of customary beliefs and values that ethnic, religious, and social groups; transmit fairly unchanged from generation to generation* , for a thorough discussion see Guiso et al. [2011].

[8]The authors use information on the country of origin of the father to define the *source country* of second generation women observed in the 1970 U.S. Census and in multiple GSS waves.

[9]See Fernández [2011] for an introduction to the epidemiological approach.

& Giuliano [2010] use the beliefs about the family from the World Value Survey as a proxy for second generation's "cultural baggage" inherited from their parents. The authors demonstrate that culture has high explanatory power with respect to women's as well as youth's LFP measured from the CPS data and the American Time Use Survey. Furthermore, the authors also find that the "cultural baggage" variable affects a wide array of choices such as: family size, home production, living arrangements and geographic mobility of second generation migrants.

Since the sample of analysis is always made of a cross-section of individuals, these articles also face some limitations linked to the absence of time dimension. For instance, it is impossible to control for place-of-origin unobservable characteristics that might be driving the results through a spurious correlation. Algan & Cahuc [2010] were able to control for *source country* unobservables by looking at different cohorts of immigrants' descendants over time. In order to study the effect of trust on GDP per-capita growth in a set of countries, the authors estimate values of trust for the beginning of the twentieth century (1910) by looking at GSS answers of second, third and fourth generation U.S. citizens whose parents moved to the U.S. around 1910. Provided that the transmission of trust is vertical (i.e. from parents to child) and that immigrants' descendants are not influenced by shocks occurring in the *source country* after their ancestors left, the trust level should differ over consecutive cohorts of immigrants. The main problem of the paper lies in the fact that the transmission of trust across generations need not be vertical. Different sources of transmission can occur through the interaction with newly arrived immigrants from the same *source country* of their ancestors. Alternatively, higher generations could be assimilating and simply reflect the trust level of the country in which they are living.

As a matter of fact, Bisin & Verdier [2001] show that there are multiple channels through which heterogeneous preferences can persist over time. In their model, the authors hypothesize the existence of two channels of transmission: vertical (i.e. through the parents) and horizontal-oblique (i.e. through peers, teachers etc.) and show that both substitutability and complementarity among the two channels can sustain stationary states in which heterogeneous traits persist in the population. In light of this theoretical result, one cannot be sure that a second or higher generation immigrant will acquire his social trait *exclusively* from his family. In fact, if consecutive generations from the same *source country* have different social traits, socialization among them will increase the probability of acquiring a trait that differs from the one of their parents.

Mostly because of data shortage, studies documenting the persistent effect of cultural norms on preferences and choices could not check for the presence of these two channels. Fernández & Fogli [2009, 2006] for instance, use 1950 female LFP and fertility from a set of *source countries* as epidemiological variables explaining the variation in economic outcomes between women aged thirty to forty years old in 1970. Therefore, 1950 is not an ideal choice as their parents were certainly born at the beginning of the century when values for women LFP and fertility were certainly different and, because of the fertility transition, not highly correlated with the values observed in 1950. Hence, from their studies, it is not clear which transmission channel among the vertical and the horizontal one is driving the correlation. By the same token, many articles applying the epidemiological approach suffer from this problem: Alesina

& Giuliano [2010] for instance, employ the independent variables as well as the key right hand side one from surveys conducted roughly at the same time. Despite being very recent, Ljunge [2014] has this issue as well when studying the intergenerational transmission of trust among the children of immigrants in several European countries. Similarly to Algan & Cahuc [2010], the author uses trust measured in the parents' countries as a key independent variable to estimate the intergenerational transmission of this value. However, both the left and right hand side variables are again measured at the same time using different World Value Survey waves. Finally, Giavazzi et al. [2014] analyze the convergence of a set of values among immigrants up to the fourth generation within the U.S. and finds substantial heterogeneity in this process. Namely, the authors show that persistence is specific to some topics such as religious ones as well as linked to descendants whose ancestors came from specific countries.

3 Data Description

3.1 Fertility Data for European Countries 1880-1970

I use data on marital fertility from the following source: *The decline of fertility in Europe: the revised proceedings of a conference on the Princeton European Fertility Project* [Coale & Watkins, 1986], which to date represents the most complete source of information on European fertility during the nineteenth and early twentieth centuries. The main goal of this study was to date the onset of the fertility transition in every European region. Specifically, for every country s and different years t, this source includes the Marital Fertility Rate. Coale & Watkins [1986] also reports another variable: I^g_{st} which is a ratio of the number of births occurred to married women divided by a hypothetical fertility plateau that would be reached if all women in the population were to adopt the Hutterites' fertility schedule.[10] Throughout the rest of the paper I use MFR_{st} as right hand side variable, I also replicate my analysis using I^g_{st} in the Appendix's section A.3. Data frequency differs by country, France, for instance, has data from 1831 until 1961. Other countries like Romania and Bulgaria have only three data points starting from 1900 and ending in 1956. In general, most of the countries in the sample have at least four different observations divided by a 30 years lag between each other starting from 1880 until 1970.[11]

In light of the criticism moved by Guinnane, Okun & Trussell [1994] to Coale & Watkins [1986], it is important to stress that I do not use the authors estimated date of the fertility transition in this analysis. In fact, I employ their data exclusively to have a measure of the MFR for more than twenty countries in a time window of almost one hundred years. Since the first fertility transition occurred in Europe mostly during the second half of the nineteenth and the first half of the twentieth centuries this implies that, as shown in Table 1, the autocorrelation of MFR_{st} (as well as I^g_{st}) is relatively low when

[10] The Hutterites are an Anabaptist sect that migrated from Europe to the north central regions of the U.S. as well as south central Canada in order to avoid religious persecution. Since any sort of contraception or abortion is strictly forbidden within this sect, their Fertility rate is taken as an upper bound by Coale & Watkins [1986]. Additional details on how the variables are constructed are included in the section A.1 of the Appendix.

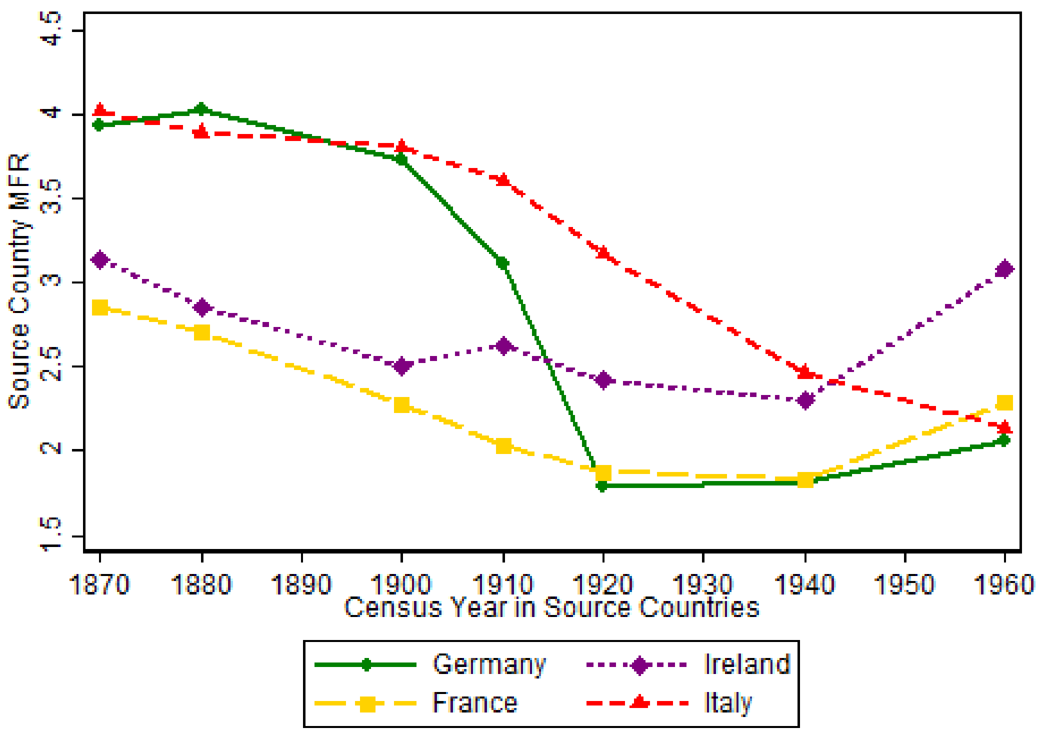
[11] Table 8 in Appendix A.1 shows data availability for the countries in this study.

these variable are opportunely spaced using a 30 years lag.[12]

Table 1: **Autocorrelation of the two variables with a 30 years lag**

	MFR_{st}	I^g_{st}
MFR_{st-30}	0.5536	
I^g_{st-30}		0.7436

Figure 1: **Marital Fertility Rates over time for five countries**



Source: author's calculation using data from Coale & Watkins [1986].

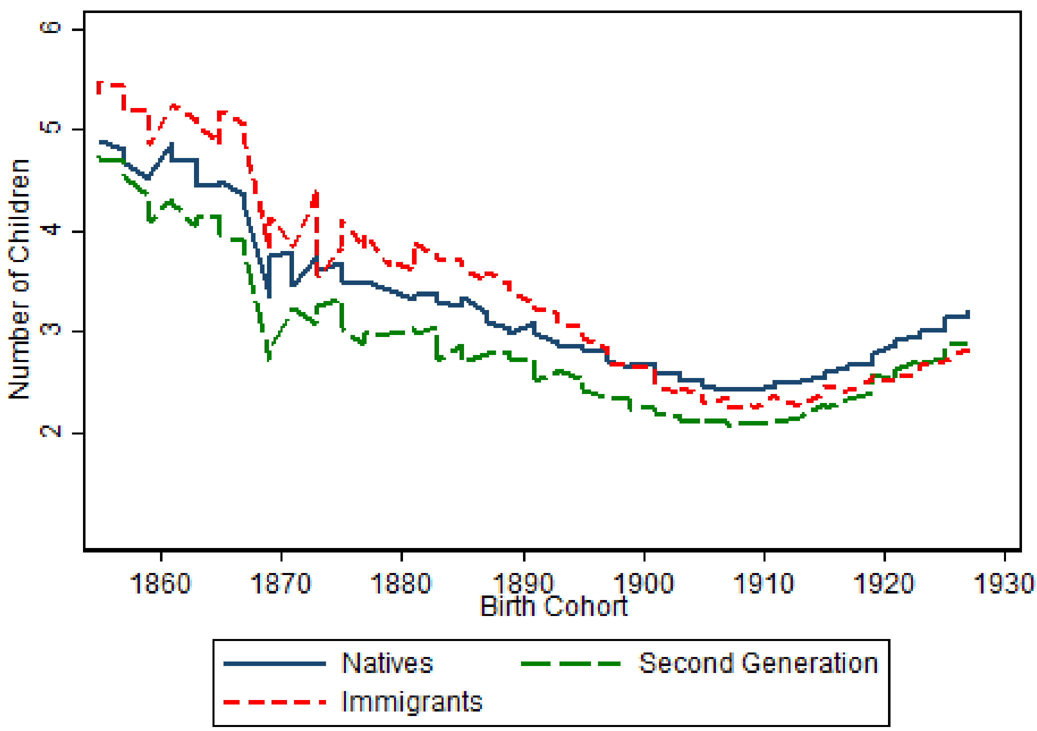
Figure 1 shows the variation in the data for four countries for which the frequency of fertility data is particularly high. As it is evident, fertility levels are sticky when observed over ten years interval, however, once they are opportunely spaced over thirty years intervals, the figure shows more variation. An obvious limitation of using data aggregated at the national level is that I lose the within country heterogeneity dimension. As suggested by Spitzer & Zimran [2013], one should be careful in using national averages when making inference on a heterogeneous population. Indeed, Coale & Watkins [1986] collected data at a finer level than the national one (a pattern visible in Figure 4 in the Appendix). In general, in my study, I am unable to take advantage of the within country variation displayed in this

[12]The choice of 30 years can also be interpreted as a "generational" lag.

source. However, the within country dimension allows me to have fertility data for regions that later became countries such as the Baltic states, Czechoslovakia and Yugoslavia. A within country analysis would require building a matching algorithm that infers the region of origin of the parents based on their last names, a fact that is clearly impossible for women since their last name changes after marriage. Table 7 in the Appendix section A.1.1 replicates one of the main regressions of [Fernández & Fogli, 2009] using [Coale & Watkins, 1986] data showing that results are comparable to the ones she obtained using her dataset.

3.2 Data on Fertility in the U.S. 1910-1970

Figure 2: **Number of Children for Different Birth Cohorts: Natives and Immigrants**



Source: author's calculations selecting women older than 49 in the 1900, 1910, 1940, 1950 and 1970 Censuses.

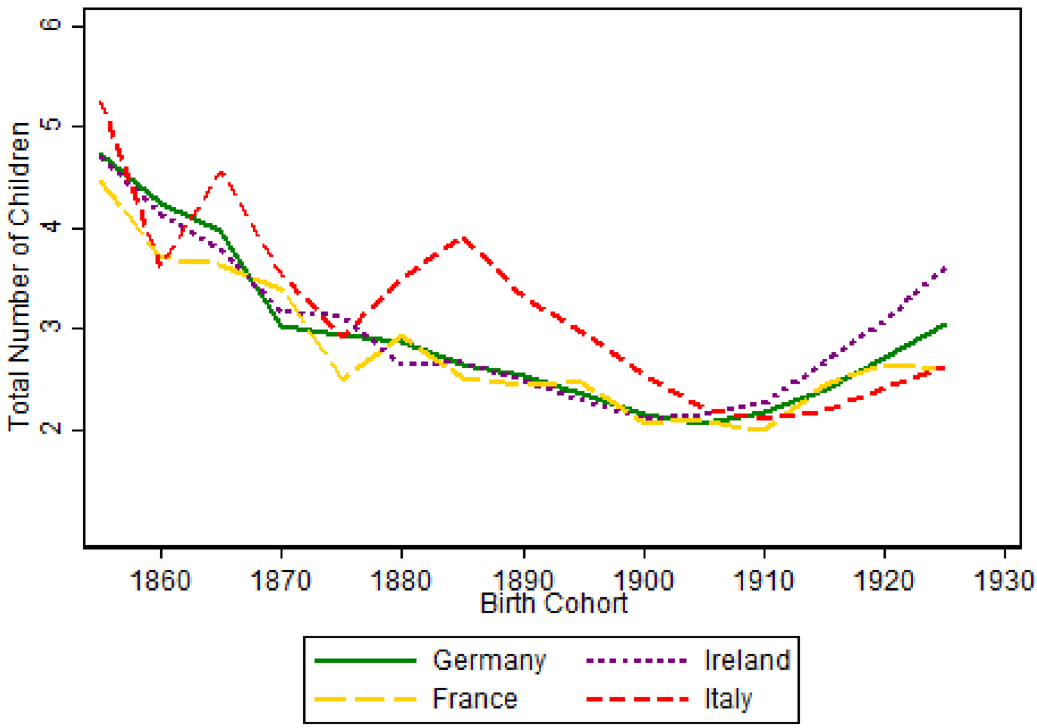
I use individual information on married women born in the U.S. with at least one foreign born parent from the following Censuses: 1910, 1940, 1950 and 1970.[13] I restrict my sample to married women between 20 and 49 years of age as Coale & Watkins [1986] computed their variables using the same age group.[14] The choice of the Censuses is led by the presence of the following variables that are important for the empirical analysis: number of children that a woman had at the time she was filing the Census,

[13]For every year I downloaded the 1% sample from IPUMS, for 1970 the sample used is the 1 % Metro fm2 one.

[14]I understand that the age distribution of the European countries during the years for which the variables were constructed affects their values and might well differ from the age distribution of second generation women in the U.S. observed in the Censuses.

within-state geographical identifier, place of birth of the parents and husband's presence within the household. As I am studying the fertility choices of women in different age groups, I cannot use the 1920 as well as the 1930 Censuses as they only ask the number of children living within the household at the time the Census was filed.[15] I could not use the 1960 Census as it lacks detailed geographical identifiers.

Figure 3: **Number of Children for Second Generation Women from Different Source Countries**



Source: author's calculations selecting women older than 49 in the 1900, 1910, 1940, 1950 and 1970 Censuses.

In Figure 2 I plot the completed fertility for women in different birth cohorts disaggregated by nativity status. Overall, this figure shows that the data on the number of children that second generation women had display a common trend with respect to natives and foreign born immigrants. However, plotting the completed marital fertility rates by *source country* (see Figure 3) shows that there are persistent differences within second generation immigrants. A detail to bear in mind, when looking at Figure 3 is that these data are taken from consecutive cohorts of second generation women whose parents' social and economic background might differ. As the composition of immigrants changed over time, the sample reflects the variation in migrants' source countries over time. Table 2 shows that the sample of second generation women in 1910 is mainly composed by Germans, Irish and English.[16] This is because the

[15]This would imply that for the women in the age group 40-50 years old I would systematically underestimate the number of children they had as some of them might have already moved out of the household.

[16]In Table 2 I select second generation women as having at least one parent foreign born and when both parents are

early comers in the U.S. were mainly from these three countries while at the beginning of the twentieth century immigrants came disproportionately from Eastern and Southern Europe.[17] This pattern can be seen in the following Censuses where the fraction of women whose parents came from countries like Poland, Italy and Russia increases. As the U.S. Census never asked question on religiosity, I am certainly missing this important dimension of heterogeneity by only looking at country of origin of the parents. As a matter of fact, Irish fertility differed according to the religious faith (a fact that is somewhat visible from Figure 4 where the regions nowadays part of Northern Ireland have lower values of I_g in 1900). In order to remove the descendants of Jewish immigrants from the sample I follow Angrist [2002] and look at Census question on the mother tongue (as well as mother tongue of the parents for those in 1910 Census) so that I can remove all native Yiddish and Hebrew speakers.

Table 2: **Sample composition by year: selected countries in % of the total sample**

Countries	**Census year**				
	1910	1940	1950	1970	Total
Scandinavian Countries	6.5	11.7	10.1	7.3	8.6
England	11.8	7.7	5.7	5.3	7.6
Scotland	3.1	2.2	1.9	2.5	2.5
Ireland	22.6	9.2	6.9	5.7	11.3
Italy	0.9	10.1	17.3	23.7	13.6
Austria	2.0	3.7	5.6	4.5	3.9
Czechoslovakia	0.1	2.8	3.1	3.7	2.4
Germany	45.3	24.6	13.8	8.3	22.5
Poland	0.2	8.6	12.4	12.3	8.4
Baltic States	0.0	1.2	1.8	1.8	1.2
Russia	1.0	8.0	10.9	10.6	7.6
# of Second Gen. Women	22,761	13,102	18,713	24,514	79,090

4 Empirical Strategy

4.1 Identification and Challenges to Internal Validity

In this paper I apply the epidemiological approach to study the persistence of cultural heritage on fertility choices of second generation migrant women during the period 1910-1970. As discussed in section 2, this strategy uses a variable measured in the country of origin to capture the effect of belonging to a specific group on a certain outcome. This identification strategy relies on the assumption that there is no omitted variable systematically correlated across different countries, if this is the case, then the epidemiological

foreign born and come from two different countries I assign the woman to belong to the mother's source country see Table 9 in the AppendixA.2 to have a full list of women whose parents were foreign born.

[17] For more details on the Age of Mass Migration and immigrants composition over time see Hatton & Williamson [1998].

approach fails as the key right hand side variable might be capturing a spurious correlation driven by the omitted variable.[18] The main difference between my reduced form identification strategy and previous articles using a similar approach is that I take advantage of the longitudinal dimension of Coale & Watkins [1986] data to purge estimates from time-invariant unobservables. The variation in fertility rates over time is a product of the large time frame considered as well as of the differential timing in which the first fertility transition occurred among European countries. The dependent variable is the number of children ever born observed at the individual (i.e. married woman) level. As shown in Table 2 the sample is composed of more than seventy nine thousand observations. However, the empirical estimation relies on the variation observed at the country-year level. Since the outcome of interest is a discrete nonnegative integer, I estimate a count data model as it is more interesting to understand the effect of the epidemiological variables on having one, two or more children rather than being able to tell what is the effect of the conditional mean. In order to address overdispersion of the dependent variable I run a pooled negative binomial model.[19]. I list the regressors of matrix $\boldsymbol{Z}$ in equation NB2 (1) in more detail in equation (2) and discuss them below.

$$\boldsymbol{Z}'\boldsymbol{\delta} + \epsilon_i = \alpha_s + \boldsymbol{\gamma}_1'\boldsymbol{X}_{it} + \beta_1\left(MFR_{st}\right) + \beta_2\left(MFR_{s(t-30)}\right) + \tau_t + r_m + \epsilon_{ismt} \tag{2}$$

In order to test the cultural transmission of preferences I run a "horse race" between contemporaneous MFR (i.e. MFR_{st} in equation (2)) and lagged MFR (that is $MFR_{s(t-30)}$ in the same equation) measured in parent's *source country*. The subscript t labels the year in which the MFR has been measured in the *source country*. That is, in order to explain fertility choices of women in the sample, I include two observations of the MFR measured with a lag of thirty years. The choice of the thirty years lag reflects the change of fertility norms across two generations: MFR_{st} and $MFR_{s(t-30)}$ represent in equation (2), respectively, the prevailing norm across their peers and parents' cohorts in the *source country* measured at time t and $t-30$. The rationale for doing so is the following: if transmission of fertility preferences occurs from parents to daughters (the so called vertical channel in Bisin & Verdier [2001]) then the $MFR_{s(t-30)}$ will explain today's fertility preferences of second generation women. On the contrary, the presence of *horizontal* transmission of preferences is captured including the MFR_{st} among the explanatory variables. In addition, according to this classification, most papers reviewed in

[18] See Fernández [2011] for additional details on the caveats of using the epidemiological approach to identify the transmission of preferences.

[19] Equation NB2 (1) shows the general expression of the Negative Binomial model:

$$\begin{aligned} f\left(y_i|\boldsymbol{Z}_i u_i\right) &= \frac{e^{-(\boldsymbol{Z}_i u_i)}(\boldsymbol{Z}_i u_i)^{y_i}}{y_i!} \\ \boldsymbol{Z}'\boldsymbol{\delta} + \epsilon_i &= \ln\lambda_i + \ln u_i \\ \mathbb{E}\left(y_i|\boldsymbol{Z}_i,\epsilon_i\right) &= \exp\left(\boldsymbol{Z}_i'\boldsymbol{\delta} + u_i\right) \end{aligned} \qquad \text{NB2 (1)}$$

The Negative Binomial estimation requires assuming that the individual heterogeneity term $\exp^{\epsilon_i} = u_i$ is distributed as a Gamma (with parameters $\alpha = \theta \;\; \beta = \theta$) so that the conditional mean of y_i given $\boldsymbol{Z}_i$ equals to λ_i. See Cameron & Trivedi [2013] for a discussion on the Negative Binomial model. For robustness, in section A.3 of the Appendix I present results when a pooled OLS is used.

section 2 are unable to distinguish among the two channels as the right and left hand side variables are generally measured during the same time window in Alesina & Giuliano [2010], Fernández & Fogli [2006, 2009], Ljunge [2014]. The variation in fertility at the source country level is crucial as the relatively low autocorrelation of MFR allows me to use contemporaneous values of MFR as a proxy for fertility preferences transmitted by peers rather than parents. In fact, an alternative approach that would use the observed MFR among immigrants currently living in the U.S. would suffer from the reflection problem ([Manski, 1993]).

The dependent variable, y_{ismt} is the number of children ever born to woman i whose parents came from country s, living in MSA m and surveyed in Census t. $\boldsymbol{X}_{it}$ is a set of individual characteristics correlated with fertility measured in Census t. Namely, these variables are: age, a set of dummies for husband's age and a dummy for farm status.[20] The choice of using women from consecutive Censuses suffers from the drawback that some questions changed over time. In fact, the 1910 Census did not ask for the years of completed education of respondents, therefore, I cannot control for this important determinant of fertility. The concern here is that the cultural effect might be upward biased as it is capturing the outcome caused by parents' underinvestment in education rather than fertility preferences *per-se*. Past studies analyzing the intergenerational transmission of fertility have taken different stances on whether including LFP and education status. On the one hand, Fernández & Fogli [2009] control for as many variables as possible thus including LFP status and educational attainment to avoid the upward bias discussed above. On the other hand, [Blau, Kahn, Liu & Papps, 2013] omits women's education level and LFP status when analyzing preferences' transmission arguing that fertility preference might be the cause leading to the choice of not investing in education (or entering the labor market). The authors argue that if this is the case, their inclusion among the controls biases downward the estimate of the cultural transmission coefficient. I choose to estimate the model above with and without LFP status, since results are generally identical, I only report models in which LFP status is included. In order to have a proxy for family's income, I create a dummy for high earnings occupation based on the occupation score assigned to the husband in the household. I compute the sex ratio at the MSA level following Angrist [2002]'s aggregation procedure among *source countries* as well as generating the sex ratio for each individual *source country*.[21]

The advantages of the time dimension in the fertility data are manifold as I can control for time-invariant unobservables both at the geographical and country of origin level. In fact, α_s in equation (2) is a *source country* fixed effect, i.e. it equals one for all second generations women whose parents came from country s. In order to control for Census specific FE I add τ_t in my specification[22] In addition, the

[20] I generate husband dummies in 10 years interval, from 25 to 34, then 35 to 44 and so on.

[21] Whenever a woman lived outside an MSA I computed this value in the smallest identifiable geographical area, that are respectively: counties for 1910 Census, state economic areas (SEA) for 1940 and 1950 Censuses and County Groups (CNTYGP97) for 1970, see IPUMS website for additional details on these variables.

[22] Note that, since my sample does not have as many observations for the 1940 and 1950 Censuses, I treat them as a unique Census when adding τ_t. Results with the Census Year FE treating 1940 and 1950 Censuses separately are available upon request and do not change much with respect to those shown in the following sections.

period studied is one of sharp changes in women LFP within the U.S.[23] Hence, I include a FE for each MSA labeled with r_m in equation (2) to control for different labor market opportunities at the MSA level. Furthermore, I also run a more demanding specification that is identical to the one in equation (2) except for the fact that I interact the MSA FE r_m with the Census Year FE τ_t, by this token, I control for unobservables characteristics that change over time at the geographical level. There are, in fact, several factors affecting fertility whose impact might be changing over time such as: infant mortality, female labor market opportunities in the different MSAs.[24] Importantly, the interaction term among fixed effects ($r_m \times \tau_t$) absorbs the fertility within the MSAs in different Census years so that I do not need to add these variable within the set of explanatory variables in equation (2). Lastly, unobserved human capital transmission from parents as well as variation in women's education level represent a major threat to internal validity as I cannot control for them in the specification above.
In section 5.1 I first run the horse race to assess what is the prevailing channel of transmission of fertility preferences among second generation women. Following a short discussion of results, I try to explain what is the underlying mechanism and provide evidence about it in section 5.2.

5 Results

5.1 The Horse Race Contest

In order to show how results and coefficients change when only one of the two $MFRs$ is added, together with the horse race, I initially run equation NB2 (1) including only one of the two epidemiological variables among the right hand side ones. Therefore, the first two columns of every table that follows report results when only the peers' fertility (i.e. MFR_{st}) is included among the regressors. In particular, the first column of every table reports the specification without interacting Census year FEs with MSA ones while in column 2, and more generally in every even column that will follow, I interact the two FEs with each other.[25] By the same token, the ensuing two columns report results of the two specifications having only lagged fertility ($MFR_{s(t-30)}$) as epidemiological variable. Finally, the last two columns (i.e. column 5 and 6 of each table) display results of the horse race. Standard errors are clustered at the parents' *source country* level and reported in parentheses.[26]
The Negative Binomial coefficients reported in Table 3 represent marginal effects that can be given an additive interpretation, that is, a one unit change in the regressor x_j equals $\frac{\partial E(y|x)}{\partial x_j} = \beta_j * exp(x'\beta)$, in other words, holding all other variables constant a one unit increase in MFR_{st} increase the difference in logs of expected children by more than 0.05 (when considering the horse race result). The Incidence

[23] Fogli & Veldkamp [2011] document the transition of female LFP participation in the U.S..
[24] That is, female labor market opportunities (or child mortality) in Chicago in 1910 are not the same as the ones in Chicago in 1970.
[25] The bottom of each table has a list of which FE are included in the regression.
[26] Note that significance tests on the Incidence Rate Ratios are run against the null hypothesis that if the regressor has no effect on the number of children ever born then $\exp^{\hat{\beta}} = 1$.

Rate Ratio, labeled with **IRR** at the bottom of table 3 is easier to interpret as it gives "exponentiated" coefficients ($\exp^{\hat{\beta}}$) that can be given a multiplicative interpretation. The Incidence Rate Ratio of the first four columns shows that a one child increase in the *source country's* MFR is associated with an increase of the number of children ever born by a factor of 7% (5% for lagged values MFR_{st-30}). When the two variables are horse raced, the lagged measure MFR_{st-30} is marginally significant at the 10% level while the coefficient (and the incidence rate ratio) of the contemporaneous MFR_{st} remains significant and its size decreases only marginally. The main result stemming from table 3 is that the MFR measured among overseas peers is a better proxy for the transmission of preferences than the same variable over the previous generation. In other words, rather than imitating their mothers, second generation women tend to behave more like their peers abroad. While the lack of imitation of mothers' fertility is consistent with previous result in the social biology literature ([Murphy, 1999]), the fact that fertility behavior of second generation women evolved along the same lines of the *source country* of the parents is striking and new.

Table 3: **Horse Race Results**

	Dependent Variable Children Ever Born					
	Current Fertility		Lagged Fertility		Horse Race	
MFR_{st}	0.075***	0.083***			0.058*	0.066**
	(0.025)	(0.025)			(0.030)	(0.027)
MFR_{st-30}			0.067***	0.070**	0.050	0.051*
			(0.025)	(0.028)	(0.031)	(0.029)
Labor force status	-0.255***	-0.257***	-0.255***	-0.257***	-0.255***	-0.258***
	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)
IRR β						
MFR_{st}	1.07***	1.07***			1.06***	1.05**
MFR_{st-30}			1.05***	1.05***	1.04*	1.04*
# Countries	27	27	27	27	27	27
# Observations	79090	79090	79090	79090	79090	79090
Ancestry FE	✓	✓	✓	✓	✓	✓
Census Year FE	✓	×	✓	×	✓	×
MSA FE	✓	×	✓	×	✓	×
MSA*Year FE	×	✓	×	✓	×	✓

The coefficients shown are marginal effects estimated from a Negatve Binomial model, controls include woman's age, age squared, sex ratio among migrants from the same *source country* within the geographical area, dummies for husband's age group and income. * p<.1, ** p<.05, *** p<.01 S.E. in parentheses clustered at the *source country* of the parents level.

In section A.3 of the Appendix I include the results of several robustness checks. I show that results are qualitatively identical if, instead of using a negative binomial model, I simply run pooled OLS keeping the right hand side variables unchanged with respect to the ones in equation (2). Besides, Coale

& Watkins [1986] construct the variable I^g_{st} which is a "noisy" proxy for MFR, so I run again every specification shown above in section A.3 and show that, again, results are qualitatively unchanged.[27] In the following section, I propose a mechanism to explain the findings of this section.

5.2 Mechanism Underlying the Horse Race Result

Table 3 shows that the vertical transmission of preferences alone is unable to explain the persistence of fertility preferences in the sample. Since results of the previous section suggest that second generation women in the U.S. picked up social norms of peers living in their parent's respective *source countries*, I analyze what explains the above result. While the role of second generation peers might be important in amplifying the transmission effect, as previously found in [Fernández & Fogli, 2009], I argue that these cohorts alone are unlikely to know what are the prevailing contemporaneous fertility norms in their *source countries*. For these norms changed considerably after their parents' departure, second generation women (as well as their husbands) need to "learn" what the contemporaneous fertility norms are in their *source countries*. In order to substantiate this claim, I construct the density of second generation immigrants over the total population at the MSA level (I call this variable $AncestryRatio$).[28] Table 4 provides evidence that once this variable is interacted with the fertility rates and added among the regressors, it is not significant in explaining fertility choices of the women in the sample. Moreover, comparing the coefficients of the epidemiological variables in Tables 4 and 3, it is straightforward to notice that results are not sensitive to the inclusion of these variables.

Therefore, I investigate the role of foreign born immigrants as catalysts of the fertility norm measured with MFR_{st} and prevailing in the horse race of Table 3. In order to provide evidence about the prevailing fertility norms among foreign born cohorts that lived in the U.S. during that same time, Table 5 displays results after estimating the model of equation NB2 (1) on a sample of foreign-born married women. In this case, the horse race has a clear winner, lagged values of MFR are never significant in the horse race, this result shows that fertility preferences of foreign born immigrants in the U.S. reflect the ones of their overseas peers. As expected, the explanatory power and significance of the contemporaneous MFRs are higher when the sample is made of foreign born immigrant women rather than second generation ones.

The role of social learning and behavioral change is not new in the analysis of fertility preferences. A recent paper by [Spolaore & Wacziarg, 2014] provides evidence that the fertility decline, occurred during the first demographic transition in Europe, was the result of the diffusion of new social norms and behavioral changes from the innovator (i.e. France) to the countries nearby and, gradually, to the rest of Europe. Because of the time frame considered, alternative channels of transmission such as television,

[27] See Appendix sections A.1 and A.2 for details and limitations of I^g_{st}, an important caveat to keep in mind when looking at the last two columns where both variables are added is that the correlation of I^g over time is 0.74.

[28] Whenever a woman in the sample was not living in a MSA, I computed this figure for the smallest geographical area which were, respectively, counties (in 1910 Census), state economic areas (in 1940 and 1950 Censuses) and county groups (in 1970 Census)

Table 4: **Horse Race and presence of Second generation immigrants**

	Dependent Variable Children Ever Born					
	Current Fertility		Lagged Fertility		Horse Race	
MFR_{st}	0.084***	0.085***			0.067*	0.067**
	(0.029)	(0.028)			(0.036)	(0.033)
MFR_{st-30}			0.066**	0.064**	0.049	0.050*
			(0.033)	(0.030)	(0.033)	(0.030)
$AncestryR * MFR_{st}$	-0.159	-0.018			-0.154	-0.010
	(0.175)	(0.176)			(0.168)	(0.173)
$AncestryR * MFR_{st-30}$			0.020	0.169		
			(0.239)	(0.178)		
Ancestry Ratio	0.931	0.586	0.411	-0.082	0.909*	0.556
	(0.570)	(0.581)	(0.923)	(0.662)	(0.534)	(0.556)
Labor force status	-0.255***	-0.258***	-0.255***	-0.258***	-0.255***	-0.258***
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)
IRR β						
MFR_{st}	1.09***	1.09***			1.07*	1.07**
MFR_{st-30}			1.06**	1.06**	1.05	1.05*
# Countries	27	27	27	27	27	27
# Observations	79090	79090	79090	79090	79090	79090
Ancestry FE	✓	✓	✓	✓	✓	✓
Census Year FE	✓	×	✓	×	✓	×
MSA FE	✓	×	✓	×	✓	×
MSA*Census Year FE	×	✓	×	✓	×	✓

* p<.1, ** p<.05, *** p<.01 see 3 for details on controls added and S.E. computation.

newspapers and the radio are unlikely to play a decisive role in shaping fertility preferences.[29] In order to test whether women living closer to Europe are more likely to "be in touch" with their respective *source countries*, I analyzed whether the distance between the MSAs where the women in the sample lived and European capitals is correlated with the transmission of preferences, finding no significant results.[30] Moreover, Bisin & Verdier [2001] model provides the theoretical foundation of the proposed channel: the authors show that, parents' socialization effort (i.e. the effort to *directly* transmit their social trait) is reduced whenever they perceive their social trait to be widespread in the society. Of course, measuring this channel would require having more detailed data than Censuses' ones. As a matter of fact, I would need to observe women's (as well as their husbands') network of peers since their early life which is not possible with Census data.

In order to measure peer influence on fertility decisions of second generation married women, I computed

[29] [Ferrara, Chong & Duryea, 2012] show that soap operas shaped women's preferences for lower fertility rates in Brazil.
[30] These regressions are available upon requests.

Table 5: **Placebo: Horse Race on Foreign Born immigrants**

	Dependent Variable Children Ever Born					
	Current Fertility		Lagged Fertility		Horse Race	
MFR_{st}	0.136***	0.128***			0.129***	0.121***
	(0.030)	(0.027)			(0.029)	(0.026)
MFR_{st-30}			0.090	0.086*	0.062	0.060
			(0.055)	(0.048)	(0.052)	(0.045)
Labor Force Status	-0.294***	-0.295***	-0.294***	-0.295***	-0.292***	-0.293***
	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)
IRR β						
MFR_{st}	1.12***	1.12***			1.12***	1.12***
MFR_{st-30}			1.07	1.08***	1.05	1.06
# Countries	27	27	27	27	27	27
# Observations	35344	35344	35344	35344	35344	35344
Ancestry FE	✓	✓	✓	✓	✓	✓
Census Year FE	✓	×	✓	×	✓	×
MSA FE	✓	×	✓	×	✓	×
MSA*Census Year FE	×	✓	×	✓	×	✓

* p<.1, ** p<.05, *** p<.01 see Table 3 for details on controls and S.E. computation.

a ratio expressing, for each country of origin, how large the pool of foreign born immigrants was with respect to the one of second generation within the geographical area of residence. This variable, labeled $MarMigRate$, takes values between zero and one. The numerator of the ratio counts, by source country and within the MSA, the number of childbearing age couples with at least one member being born overseas residing. The denominator of the ratio counts, by source country, how many first and second generation couples live within a specific MSA.[31] The caveat to bear in mind here is that I cannot control for selective migration of the women in the sample inside or outside geographical areas with more or less peers from the same *source country*. In Table 6 I augment equation (2) interacting the newly generated variable with the current and lagged values of MFR. The first column of Table 6 shows that MFR_{st} is no longer significant once the interaction with the fraction of migrants married couple is added to the regression, moreover, the interaction term's Incidence Rate Ratio is larger than the one for MFR_{st} in Table 3. The interaction term is not significant in the more demanding specification of column 2 where I interact Census year FEs with MSA's ones.[32]. The significance of the lagged fertility's coefficient in the horse race specifications of columns (3) and (4) implies that there is some residual variation captured by this variable. A possible interpretation for the results in the first four columns of Table 6 is that the

[31]The main geographical areas are MSAs, whenever a woman was not living in a MSA, I computed this figure for the smallest geographical area which were, respectively, counties (in 1910 Census), state economic areas (in 1940 and 1950 Censuses) and county groups (in 1970 Census)

[32]Note that the interaction of $MarMigRate$ with MFR_{st-30} is not significant in explaining fertility preferences, more importantly, the inclusion of this interaction term has no effect on the horse race results (results available upon request)

marriage market matters. Columns (5) and (6) show that dropping the women married to an foreign born husband from the sample changes the horse race result.[33] Additional evidence of the presence of a "learning effect" comes from columns (7) and (8) in the same Table. In fact, in these columns I show that the horse race results of Table 3 is reversed whenever I drop from the sample women living in areas where values of $MarMigRate$ for their respective *source countries* is equal or above 0.5. Despite the fact that I cannot directly observe parents' socialization efforts or selective migration to specific MSAs, the evidence from Table 6 suggests that parents' socialization effort was higher in MSAs where second generation women were less likely to socialize with foreign born peers from the same source country.

6 Conclusions and Next Steps

The persistent effect of culture on economic outcomes has been widely documented in the economics literature. However, less attention has been devoted to how this effect can be measured and what is the mechanism underlying preferences' transmission. Previous studies have generally assumed the presence of a direct channel of socialization through which second generation children picked up preferences and traits akin to their parents. In this paper, I analyzed the persistence of fertility preference in a time frame in which the outcome of interest was experiencing sharp changes across countries of origin of immigrants to the U.S. The longitudinal variation in fertility norms in these countries allows me to run a horse race from which I find evidence that the "horizontal-oblique" channel of transmission prevails on the vertical one. Interestingly, I find evidence that vertical transmission acts as a substitute to the horizontal one, that is, women living in areas populated by immigrant couples from the same *source country* are more likely to adopt fertility choices similar to them. My findings are in line with the theoretical results of [Bisin & Verdier, 2001]. These results come with some caveats: I am unable to account for women's self-selection into areas more (or less) populated by immigrants. I also lack data on human capital accumulation (for the 1910 Census), i.e. I cannot fully control for the impact of human capital on the fertility decisions.

More research is needed to shed light on the channel of transmission. For instance, it would be interesting to investigate the role that church attendance played on the horizontal vs. vertical transmission of fertility norms across the time frame considered. In addition, IPUMS linked samples might be analyzed in order to test whether self-selection into areas is an issue for the internal results of the paper. In fact, these data would allow observing individuals at different points in time and take advantage of the variation in the density of immigrants' couples to explain the fertility decision of second generation women.

[33]Since only one household member was asked questions about nativity in the 1940 and 1950 Censuses, I have to drop observations from these Censuses in order to be sure not to keep husband that were born abroad.

Table 6: **Horizontal Transmission from Foreign Born Married Couples**

	Dependent Variable Children Ever Born							
	Current Fertility		Horse Race		Marriage Market		Reduced Sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MFR_{st}	0.032	0.039	0.009	0.017	0.032	0.038	0.017	0.017
	(0.029)	(0.040)	(0.033)	(0.032)	(0.023)	(0.024)	(0.017)	(0.015)
MFR_{st-30}			0.055*	0.056**	0.041**	0.040*	0.100***	0.099***
			(0.030)	(0.028)	(0.021)	(0.021)	(0.014)	(0.013)
$MarMigRa * MFR_{st}$	0.080**	0.079	0.078**	0.075*				
	(0.033)	(0.054)	(0.031)	(0.044)				
MarMigrate	-0.113	-0.127	-0.170**	-0.186*				
	(0.072)	(0.118)	(0.074)	(0.096)				
Labor force status	-0.255***	-0.257***	-0.254***	-0.256***	-0.185***	-0.189***	-0.210***	-0.210***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)
IRR β								
MFR_{st}	1.03	1.04	1.00	1.00	1.04	1.03	1.01	1.01
MFR_{st-30}			1.06***	1.05**	1.05**	1.04*	1.10***	1.10***
$MarMigRate * MFR_{st}$	1.08***	1.08	1.11***	1.10***				
# Countries	27	27	27	27	27	27	27	27
# Observations	79090	79090	79090	79090	48154	48154	36628	36628
Ancestry FE	✓	✓	✓	✓	✓	✓	✓	✓
Census Year FE	✓	×	✓	×	✓	×	✓	×
MSA FE	✓	×	✓	×	✓	×	✓	×
MSA*Census Year FE	×	✓	×	✓	×	✓	×	✓

* p<.1, ** p<.05, *** p<.01 see 3 for details on controls and S.E. computation.

A Appendix

A.1 Coale & Watkins [1986] Data

In order to compute marital fertility rates over time I used Coale & Watkins [1986] data. Namely, the authors constructed, for every country, an index (called I^f_{st}) taking values between zero and one. The index expressed how close (or far) total fertility in country s at time t was with respect to an hypothetical plateau. The plateau is constituted by the Hutterites' fertility rate. The total fertility rate index I^f_{st}, computed over all women in reproductive age (i.e. 20 to 49), is composed by the following indices:

$$\overbrace{I^f_{st}}^{\text{Total Fertility Rate Index}} = \underbrace{I^m_{st} * I^g_{st}}_{\text{Marital Fertility Rate Index of Country } s \text{ in Year } t} + (1 - I^m_{st}) * I^h_{st} \tag{A}$$

Where I^f_{st} is the ratio of the actual number of births over the hypothetical number that women would have were they to adopt the Hutterite fertility schedule. I^g_{st} is the ratio of the actual number of births occurring to married women aged twenty to forty nine years old over the hypothetical number that would be observed if the distribution of married women would adopt the Hutterite fertility schedule. Finally, I^m_{st} is a measure of the contribution of marital status to the overall rate of childbearing, this ratio is a weighted average of the proportion of married women in different age groups in the population. I^f_{st} can be written as in equation (A.1) below:

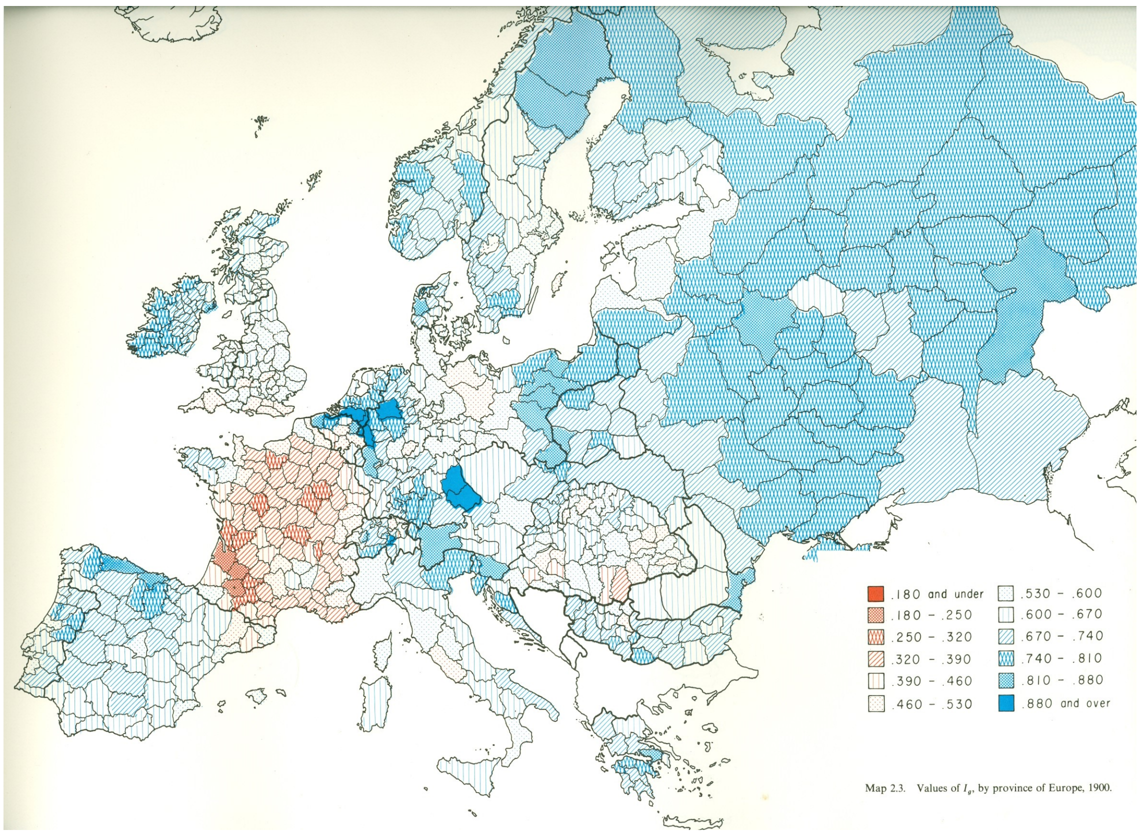
$$I^f_{ct} = \frac{B_{st}}{H^m_{st} \int_{20}^{49} h(a)w(a)_{st} \mathrm{d}a} \tag{A.1}$$

Where B_{st} is the total number of children born by every woman and $\int_{20}^{49} h(a)w(a)_{st}da$ is the plateau of maximum attainable fertility if every woman in age group $w(a)_{st}$ would follow the Hutterites' fertility schedule $h(a)$.

$$\begin{aligned}
\text{where } I^g_{st} &= \frac{B^m_{st}}{\int_{20}^{49} h(a)m(a)_{st}\mathrm{d}a} \quad I^m_{st} = \frac{\int_{20}^{49} h(a)m(a)_{st}\mathrm{d}a}{\int_{20}^{49} h(a)w(a)_{st}\mathrm{d}a} \qquad (1)\\
H^m_{st} &= \int_{20}^{49} h(a)m(a)_{st}\mathrm{d}a \\
B^m_{st} &= \text{\#of births occurred to married women} \\
m(a)_{st} &= \text{\#married women at age } a \text{ in country } s \text{ at time } t \\
h(a) &= \text{Hutterite's yearly fertility schedule}
\end{aligned}$$

In order to compute MFR for country s in year t the authors multiply the MFR's index ($I^m_{st} * I^g_{st}$) with the Hutterites' MFR (that is 10.94 children per woman). Since marriage market and age at marriage in European countries might differ from the one in the U.S., I^g_{st} is a variable measuring the degree to which married women restricted fertility in European countries during the time of analysis. As a matter of fact, I^g_{st} creates a "ranking" among the countries in the sample, from the ones exerting very little fertility restrictions after marriage, i.e. those with a high value of I^g_{st}, to the ones exerting high fertility restrictions during the marriage, that is those displaying low values of I^g_{st}. Figure 4 shows the variation in I^g for many European countries in year 1900. Regions in red are those having lower values of I^g, conversely, regions with a blue scale are those that exert little fertility control after marriage.

Figure 4: **Values of I^g_{st} when $t = 1900$ across European Regions**



Source: Coale & Watkins [1986]

A.1.1 Robustness of the Fertility Data

In order to show the validity of the data used, I run the baseline OLS regression in Fernández & Fogli [2009] and compare how results vary when substituting the epidemiological variable used by the authors with the one taken from Coale & Watkins [1986].[34] Table **??** replicates the regression in Column 8 of Table 2 in Fernández & Fogli [2009] using the two data sources for the epidemiological variable. Namely, column two of **??** uses the same data as the published paper while column one uses the data adopted to

[34]Fernández & Fogli [2009] use data from the United Nations reporting Total Fertility rates available here.

write this paper. Since there are only fifteen countries for which I have data from both sources I cannot replicate the regression with the same number of observations used in the original paper.[35] Despite these shortcomings and the fact that the size of the coefficient changes when compared to the results in the original paper, results are very similar when I use Coale & Watkins [1986] as a source for the MFR from the source countries. This fact is reassuring and signals that the data, at least for the period in which I have a comparable alternative source, are reliable.

Table 7: **Baseline Regression in Fernández & Fogli [2009] using Different Data Sources**

Source of the Epidemiological Variable		(1) # of children	(2) # of children
Fernández & Fogli [2009]	TFR_{1950}		0.393***
			(0.108)
Coale & Watkins [1986]	MFR_{1950}	0.462***	
		(0.091)	
	# Countries	15	15
	# Observations	4988	4988
	Adj. R-Sq.	0.044	0.043

* p<.1, ** p<.05, *** p<.01 S.E. in parentheses Clustered at the *source country* of the parents level.

A.2 Data on Second Generation Migrants: Additional Details

Table 8 shows the availability of the indices for various countries over time.

[35] Moreover, Fernández & Fogli [2009] dropped the countries that signed the Warsaw Pact of 1955 which are included in this study.

[36] The last observation for France, Ireland, Austria, Yugoslavia, Poland, Switzerland, Hungary, Denmark, Spain, Sweden, Norway and Netherlands is in 1960.

[37] The last observation for Germany is in 1962.

[38] The last observation for England Scotland and Wales is in 1961.

[39] The last observation for Italy is in 1961, the closest observation to 1940 comes from the 1936 Census.

[40] Information about Baltic States comes from Russia's disaggregated data.

[41] The last observation for Greece is in 1961

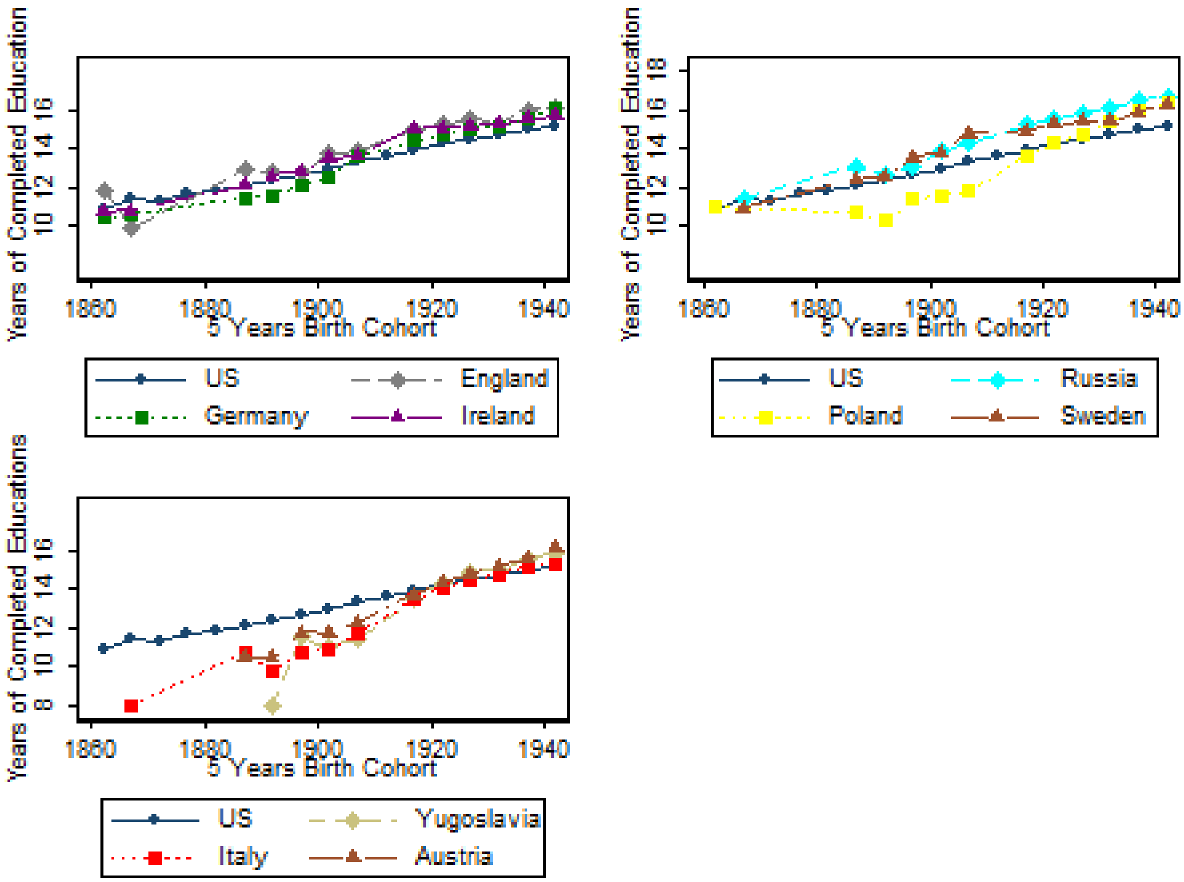
[42] Information about Czechoslovakia before the country was established comes from Austro-Hungarian Empire's Censuses.

[43] The last observation for Romania is in 1956

Table 8: **Data Availability by Year and Country from Coale & Watkins [1986]**

Country	Year						
	1870	1880	1900	1910	1930	1940	1970
France[36]	✓	✓	✓	✓	✓	✓	✓
Germany[37]	✓	✓	✓	✓	✓	✓	✓
Ireland	✓	✓	✓	×	✓	×	✓
England[38]	✓	✓	✓	✓	✓	×	✓
Scotland	✓	✓	✓	✓	✓	×	✓
Wales	✓	✓	✓	✓		×	✓
Italy[39]	✓	✓	✓	✓	✓	✓	✓
Russia	✓	×	✓	×	✓	✓	✓
Baltic States[40]	×	×	✓	×	✓	✓	✓
Norway	×	✓	✓	×	✓	×	✓
Sweden	×	✓	✓	×	✓	×	✓
Finland	✓	✓	×	✓	×	✓	✓
Denmark	✓	✓	✓	✓	✓	×	✓
Austria	×	✓	✓	✓	✓	×	✓
Hungary	×	✓	✓	×	✓	×	✓
Spain	×	✓	✓	✓	×	✓	✓
Portugal	×	✓	✓	✓	×	✓	✓
Belgium	×	✓	✓	✓	×	✓	✓
Netherlands	×	✓	✓	✓	✓	×	✓
Greece[41]	×	×	✓	×	✓	×	✓
Yugoslavia	×	×	×	×	✓	×	✓
Czechoslovakia[42]	×	✓	✓	✓	×	✓	✓
Poland	×	×	✓	×	✓	×	✓
Switzerland	✓	✓	✓	✓	×	✓	✓
Romania[43]	×	×	✓	×	✓	×	✓

Figure 5: **Average Education of Second Generation Married Women**



Source: Author's calculation using 1940, 1950 and 1970 Censuses.

Table 9: **Distribution of Second generation immigrant women across the four Censuses**

	Census year				
	1910	1940	1950	1970	Total
Denmark	237	300	263	325	1,125
Finland	18	137	238	224	617
Norway	721	497	604	624	2,446
Sweden	533	811	859	801	3,004
England	2,690	1,050	1,095	1,370	6,205
Scotland	708	303	367	657	2,035
Wales	327	107	103	75	612
Ireland	5,199	1,262	1,338	1,539	9,338
Belgium	57	54	83	139	333
France	414	194	202	273	1,083
Netherlands	231	189	267	408	1,095
Switzerland	309	159	160	179	807
Greece	0	26	103	538	667
Italy	215	1,388	3,347	6,364	11,314
Portugal	57	94	160	345	656
Spain	30	23	64	185	302
Austria	439	448	1,007	1,048	2,942
Czechoslovakia	21	379	608	978	1,986
Germany	10,381	3,375	2,654	2,219	18,629
Hungary	44	220	486	677	1,427
Poland	0	1,111	2,312	2,816	6,239
Romania	0	57	134	187	378
Yugoslavia	0	85	222	584	891
Estonia	0	1	2	11	14
Latvia	0	6	26	27	59
Lithuania	0	143	289	343	775
Russia	130	683	1,720	1,578	4,111
Total	22,761	13,102	18,713	24,514	79,090

A.3 Robustness Checks

In order to test robustness of results shown in sections 5.1 and 5.2 I take two different approaches. I first estimate a Pooled OLS model rather than the Negative Binomial of equation (2). In addition, I also used I^g_{st} as epidemiological variable instead of MFR. Moreover, I also substitute the fixed effect for the year of birth with a quadratic and linear regressor for age of the woman at the time of the Census, results (available upon request) are unchanged.

Tables 10 and 11 replicate respectively Tables 3 and 6 of the paper using a Pooled OLS model instead, as it is evident, results are mostly unchanged.

Table 10: **Horse Race using Pooled OLS**

	Dependent Variable Children Ever Born					
	Current Fertility		Lagged Fertility		Horse Race	
MFR_{st}	0.194***	0.158***			0.181***	0.119**
	(0.048)	(0.040)			(0.049)	(0.046)
MFR_{st-30}			0.091*	0.143**	0.065	0.111*
			(0.047)	(0.053)	(0.052)	(0.055)
$\log GDP_{st}$	-0.038	-0.159	-0.149	-0.347*	-0.024	-0.217
	(0.134)	(0.152)	(0.174)	(0.179)	(0.141)	(0.158)
LFP	-0.485***	-0.494***	-0.482***	-0.495***	-0.484***	-0.494***
	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.023)
β						
MFR_{st}	0.53***	0.43***			0.18***	0.12**
MFR_{st-30}			0.43*	0.14**	0.06	0.11*
# Countries	27	27	27	27	27	27
# Observations	79090	79090	79090	79090	79090	79090
Adj. R-Sq.	0.197	0.200	0.196	0.200	0.197	0.200
Ancestry FE	✓	✓	✓	✓	✓	✓
Year of Birth	✓	✓	✓	✓	✓	✓
MSA FE	✓	×	✓	×	✓	×
MSA*Census Year FE	×	✓	×	✓	×	✓

* p<.1, ** p<.05, *** p<.01 S.E. in parentheses Clustered at the Country of Origin of the parents level.

Table 11: **Horizontal Transmission from Foreign Born Married Couples Pooled OLS**

	Dependent Variable Children Ever Born					
	Current Fertility		Lagged Fertility		Horse Race	
MFR_{st}	0.107**	0.073			0.077*	0.001
	(0.050)	(0.091)			(0.041)	(0.063)
MFR_{st-30}			0.026	0.096	0.072	0.122**
			(0.064)	(0.081)	(0.049)	(0.052)
$MarMigRa * MFR_{st}$	0.163**	0.157			0.188***	0.211*
	(0.069)	(0.132)			(0.062)	(0.117)
$MarMigRa * MFR_{st-30}$			0.124	0.084		
			(0.079)	(0.098)		
$\log GDP_{st}$	-0.040	-0.169	-0.139	-0.343*	-0.028	-0.221
	(0.135)	(0.151)	(0.172)	(0.181)	(0.141)	(0.160)
LFP	-0.485***	-0.494***	-0.482***	-0.495***	-0.483***	-0.493***
	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)
MarMigRate	-0.157	-0.157	-0.134	-0.039	-0.226	-0.308
	(0.153)	(0.286)	(0.231)	(0.278)	(0.136)	(0.254)
# Countries	27	27	27	27	27	27
# Observations	79090	79090	79090	79090	79090	79090
Adj. R-Sq.	0.197	0.200	0.196	0.200	0.197	0.200
Ancestry FE	✓	✓	✓	✓	✓	✓
Year of Birth	✓	✓	✓	✓	✓	✓
MSA FE	✓	×	✓	×	✓	×
MSA*Census Year FE	×	✓	×	✓	×	✓

* p<.1, ** p<.05, *** p<.01

References

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012a). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, *102*(5), 1832–56.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012b). Have the poor always been less likely to migrate? evidence from inheritance practices during the age of mass migration. NBER Working Papers 18298, National Bureau of Economic Research, Inc.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2012c). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. NBER Working Papers 18011, National Bureau of Economic Research, Inc.

Ager, P. & Brückner, M. (2011). Cultural diversity and economic growth: Evidence from the us during the age of mass migration. Technical report.

Alesina, A., Algan, Y., Cahuc, P., & Giuliano, P. (2010). Family Values and the Regulation of Labor. IZA Discussion Papers 4747, Institute for the Study of Labor (IZA).

Alesina, A. & Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 960–980.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, *114*(4), 1243–1284.

Alesina, A. & Giuliano, P. (2010). The power of the family. *Journal of Economic Growth*, *15*(2), 93–125.

Alesina, A., Harnoss, J., & Rapoport, H. (2013). Birthplace diversity and economic prosperity. NBER Working Papers 18699, National Bureau of Economic Research, Inc.

Algan, Y. & Cahuc, P. (2010). Inherited trust and growth. *The American Economic Review*, 2060–2092.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? evidence from america's second generation. *The Quarterly Journal of Economics*, *117*(3), 997–1038.

Angrist, J. D. & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bandiera, O., Rasul, I., & Viarengo, M. (2013). The Making of Modern America: Migratory Flows in the Age of Mass Migration. *Journal of Development Economics*, *102*(C), 23–47.

Bisin, A. & Verdier, T. (2000). "Beyond The Melting Pot": Cultural Transmission, Marriage, And The Evolution Of Ethnic And Religious Traits. *The Quarterly Journal of Economics, MIT Press*, *115*(3), 955–988.

Bisin, A. & Verdier, T. (2001). The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory*, *97*(2), 298–319.

Blau, F., Kahn, L., Liu, A., & Papps, K. (2013). The transmission of womens fertility, human capital, and work orientation across immigrant generations. *Journal of Population Economics*, *26*(2), 405–435.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics*, *107*(1), 123–50.

Borjas, G. J. (1994). Long-run convergence of ethnic skill differentials: The children and grandchildren of the great migration. *Industrial and Labor Relations Review*, *47*(4), 553–73.

Borjas, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review*, *85*(3), 365–90.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Cameron, A. C. & Trivedi, P. K. (2005). Microeconometrics. *Cambridge Books*.

Cameron, A. C. & Trivedi, P. K. (2013). *Regression analysis of count data*. Number 53. Cambridge university press.

Card, D., DiNardo, J., & Estes, E. (2000). The more things change: Immigrants and the children of immigrants in the 1940s, the 1970s, and the 1990s. In *Issues in the Economics of Immigration* (pp. 227–270). University of Chicago Press.

Coale, A. J. & Watkins, S. C. (1986). *The decline of fertility in Europe: the revised proceedings of a conference on the Princeton European Fertility Project*. Princeton University Press Princeton.

Fernandez, R. (2007). Women, work, and culture. NBER Working Papers 12888, National Bureau of Economic Research, Inc.

Fernández, R. (2011). Chapter 11 - does culture matter? volume 1 of *Handbook of Social Economics* (pp. 481 – 510). North-Holland.

Fernández, R. & Fogli, A. (2006). Fertility: The role of culture and family experience. *Journal of the European Economic Association*, *4*(2-3), 552–561.

Fernández, R. & Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, *1*(1), 146–77.

Ferrara, E. L., Chong, A., & Duryea, S. (2012). Soap Operas and Fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, *4*(4), 1–31.

Ferrie, J. P. (1996). A new sample of males linked from the public use microdata sample of the 1850 us federal census of population to the 1860 us federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *29*(4), 141–156.

Fogli, A. & Veldkamp, L. (2011). Nature or Nurture? Learning and the Geography of Female Labor Force Participation. *Econometrica*, *79*(4), 1103–1138.

Giavazzi, F., Petkov, I., & Schiantarelli, F. (2014). Culture: Persistence and evolution. Technical report, National Bureau of Economic Research.

Giuliano, P. (2007). Living arrangements in western europe: Does cultural origin matter? *Journal of the European Economic Association*, *5*(5), 927–952.

Goldin, C. (1994). The political economy of immigration restriction in the united states, 1890 to 1921. In *The regulated economy: A historical approach to political economy* (pp. 223–258). University of Chicago Press.

Guinnane, T., Okun, B., & Trussell, J. (1994). What do we know about the timing of fertility transitions in europe? *Demography*, *31*(1), 1–20.

Guiso, L., Sapienza, P., & Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, *94*(3), 526–556.

Guiso, L., Sapienza, P., & Zingales, L. (2006). Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives*, *20*(2), 23–48.

Guiso, L., Sapienza, P., & Zingales, L. (2011). Chapter 10 - civic capital as the missing link. volume 1 of *Handbook of Social Economics* (pp. 417 – 480). North-Holland.

Hatton, T. J. & Williamson, J. G. (1998). *The Age of Mass Migration: Causes and Economic Impact*. Number 9780195116519 in OUP Catalogue. Oxford University Press.

Jones, L. E. & Tertilt, M. (2006). An Economic History of Fertility in the U.S.: 1826-1960. NBER Working Papers 12796, National Bureau of Economic Research, Inc.

Ljunge, M. (2014). Trust issues: Evidence on the intergenerational trust transmission among children of immigrants. *Journal of Economic Behavior & Organization*, *106*(0), 175 – 196.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), pp. 531–542.

Mitchell, B. R. (1998). *International Historical Statistics: Europe, 1750-1993*. Macmillan London.

Murphy, M. (1999). Is the relationship between fertility of parents and children really weak? *Biodemography and Social Biology*, *46*(1-2), 122–145.

Solon, G. (1999). Intergenerational mobility in the labor market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics* chapter 29, (pp. 1761–1800). Elsevier.

Spitzer, Y. & Zimran, A. (2013). Self-selection of immigrants on the basis of living standards: Evidence from the stature of italian immigrants at ellis island 1907–1925.

Spolaore, E. & Wacziarg, R. (2014). Fertility and modernity. Technical report, Department of Economics, Tufts University.

Vogl, T. S. (2014). Intergenerational dynamics and the fertility transition.

Wooldridge, J. M. (2002). *Econometric Analysis Cross Section Panel*. MIT press.