2015

Selection into Mixed Marriages: Evidence from North Carolina, 1894-1906

Matthew T. Gregg
Roger Williams University, mgregg@rwu.edu

Selection into Mixed Marriages: Evidence from North Carolina, 1894-1906*

Matthew T. Gregg†

June 15, 2015

Abstract

Using a novel dataset from agricultural censuses and geological sources, I examine the selection into mixed marriages between whites and Cherokee Indians in North Carolina during the late nineteenth century. I find that the well-documented wealth advantage of Cherokee Indian households containing white husbands is driven mostly by positive selection. Thus, once family fixed effects are controlled for, the observed intermarriage premium is completely eliminated. (JEL D03, N31, O12.)

Keywords: Mixed Marriages, Assimilation, Culture, American Indians

*I acknowledge the financial assistance of the RWU Foundation to Promote Scholarship & Teaching. I am indebted to the Friends of the National Archives: Southeast Region from highlighting the discovery of the Eastern Band censuses in their newsletter.

†Gabelli School of Business, Roger Williams University, 1 Old Ferry Road, Bristol, RI 02809-2921, phone: 401-254-3160, fax: 401-254-3545, e-mail: mgregg@rwu.edu.

1 Introduction

Across time and space, marrying across racial lines is extremely rare (Jr., 2007). When marrying out occurs, most researchers uncover a link between faster assimilation of the minority spouse and intermarriage, though the causal link is disputable (Meng and Gregory, 2005; **?**; Nottmeyer, 2010). Empirical evidence also finds positive assortative matching in mixed marriage markets. The evidence on the role of local marriage market conditions in the decision to marry out is more mixed. This empirical literature is largely based on modern data sources and the focus of the mixed marriages are often between immigrant males and Native females. Inference from historical data is arguably cleaner since the increased rates of co-habitation and other unobservable traits make the drivers of mixed marriages harder to identity.

In this paper, I investigate whether standard economic theories on marriage can be applied to mixed marriages in a racially-divided community during the Jim Crow Era. The community is the Eastern Cherokee Indian Reservation where the miscegenation between Cherokee Indians and whites was not banned. The advantage of using this reservation as a case study in mixed marriages is largely driven by the interesting nature of my constructed data set.[1] The digitization of 1906 Cherokee Indian applications to receive money from the federal government also allows me to create family trees for every husband-wife pair during these years. I can subsequently match the Cherokees who chose to intermarry to their sibling who married a full-blooded Cherokee. I also have micro data on livestock wealth from 1894 to 1906 which allows me to exploit within-occupational variation in outcomes.[2] Thus, I can follow a larger percentage of endogamous and exogamous households over thirteen consecutive years rather than relying on decennial census records.

I characterize my findings into two broad categories. First, like most studies, I find that mixed marriages, especially between a white male and a Cherokee female, are positively related to wealth accumulation and consumption. This relationship is robust to a number of controls and holds when within-district variation is exploited. However, when a controlling for family fixed effects, this positive association disappears, which is consistent with positive assortative matching. I also find that children from exogamous marriages are more likely marry out, which is additional evidence of positive sorting. Second, I find that the fraction of mixed-blooded Cherokees within a district is highly predictive of the decision to intermarry. This evidence is consistent with search models like Adachi (2003) and recent empirical work by .... Taken together, the evidence from mixed marriages on this reservation is consistent with both a Becker-style theory of marriage and local marriage market factors.

These findings support recent research on the causal effects of intermarriage (for a recent survey

[1] This reservation is unique to other Indian reservations in other dimensions. First, the land on this reservation was never allotted so there are no restrictions on alienation based on blood quantum (Beaulieu, 1984). Second, the Cherokees, one of the so-called "civilized tribes," descended from societies with similar institutional and farming traditions as Euro-Americans so there is a built-in bias against finding a positive association between intermarriage and economic outcomes. Another advantages is that mixed marriages only started in the mid to later half of the nineteenth century on this reservation so I will be estimating the role of initial entry into intermarriage market.

[2] On the Eastern Cherokee Reservation, almost 100% of the heads of households identified themselves as farmers in the 1900 Decennial Census.

on identification strategies in the intermarriage literature, see **?**). . Nekby (2010) and Nottmeyer (2010) use panel data techniques to estimate the returns to intermarriage by immigrants in Germany and Sweden, respectively, and ffter controlling for person-specific unobserved heterogeneity, both studies find no post-marriage effects on earnings for foreign-born males who marry natives. This paper is the only study to estimate the role of intermarriage by using sibling comparisons but I find similar results: the wealth gain is mainly attributed to positive assortative matching.

This paper also adds to the discussion on the role of culture in American Indian economic development. To date, the empirical evidence on the role of culture and assimilation on contemporary American Indian economic development is mixed. Cornell and Kalt (1992) and **?** suggest that cultural matches between contemporary self-governance structures and the historical experience of governance lead to economic growth. In these papers, culture only matters through its interaction with institutional quality. Anderson and Lueck (1992), Anderson and Parker (2008), and Trosper (1978) also show that American Indian culture and assimilation do not directly influence agricultural productivity, reservation economic growth and individual profit-maximization behavior, respectively. Each paper uses different samples and different measures of assimilation so comparisons are difficult but the direct role of American Indian culture has been downplayed. Unlike these papers, my study exploits the variation in assimilation at the sub-reservation level and, in doing so, shows that wealth differences, at least on this reservation, existed prior to the twentieth century.

My study is most closely related to Kuhn and Sweetman (2002)'s analysis of the role of ethnicity on labor market outcomes among Aboriginal Canadians and Gitter and Reagan (2002)'s study of on- and off-reservation Indians. Kuhn and Sweetman (2002) use a cross-section of single- and multiple-origin Aboriginals to show that assimilation variables like ancestral intermarriage and residing off reserves explain more of the employment (and earnings) gap between Aboriginals and non-Aboriginals than traditional labor market controls. They posit that labor market outcomes improve when Aborignals acquire either more skills or cultural traits from ancestors who were part of the majority culture. Gitter and Reagan (2002) also find that living on a reservation has a negative impact on the probability of employment (but no effect on wages), which may reflect a desire to maintain non-market traditions by living in larger ethnic networks. Unlike these studies, my data allow me to estimate the role of assimilation while taking sorting into account.

The rest of the paper is structured in the following way. Section 2 provides a simple theoretical framework to understand the empirical strategies. Section 4 discusses the data used in this paper and section 5 discusses the results. Last, section 7 concludes the paper.

2 Theoretical Considerations

I present a simple model rooted in (Becker, 1973)'s work that fixes the basic ideas behind the decision to intermarry. Suppose that an individual i has a human capital function $H_i = f(x_i, y_i)$ which depends on their ability x_i and the ability of their married partner, y_i. I will assume that

x_i is influenced by the level of assimilation of their parents; therefore, individual's i ability can be written as

$$x_i = x_i^P + v_i$$

where x_i^P is the assimilation level of i's parents and v_i is an idiosyncratic random term. I assume that x_i and y_i are complements in martial output: therefore, $f_{XY} > 0$, which implies that the human capital function increases when similar individuals marry.[3] I also let the return to human capital, denoted as r, depend on whether marriage is mixed: in this case, I assume that $r_W = \mathrm{E}(r|\mathrm{M}{=}1) > r_C = \mathrm{E}(r|\mathrm{M}{=}0)$ where M equals 1 if a Cherokee intermarried, zero otherwise. This assumption is consistent with a long list of studies that finds that intermarried men earn more than men who married within their ethnic group (for a survey, see Furtado and Trejo, 2012).

The last variable of interest is the social cost of intermarriage, C_i. Alienation from the tribe or family would be an example of the social cost of marrying out. I assume this cost is decreasing in x_i^P; thus, individuals from families which previously intermarried will face a lower cost of intermarriage than individuals who are the first in their family to intermarry.

According to historical sources, the cost of intermarriage varied by gender. Historical sources suggest that many white men often married Cherokee women to gain access to the female's property (Perdue, 1999, 83).[4] White males also disrupted the traditional matrilineal clan system of the Cherokees:

> The shamans had always objected to whites marrying Cherokee women, unless the husbands came to live in the Cherokee towns and became one with the tribe and its customs.... They did become Cherokee. but these whites were never pleasing to the shamans. They would not allow the wife's brother to instruct the children; as often as not they moved out of town onto land of their choice and established an independent home. (Ehle, 1989, 35).

Because to these costs, the Cherokee government heavily regulated marriages with white males. White males were charged more for a marriage license, needed to provide evidence of no surviving wife, and were required to provide a letter of good moral character, which needed to be signed by at least ten reputable Cherokees. More importantly, white males forfeited the rights to Cherokee land if the Cherokee wife deserted the residence (Murchison, 1928).

The social costs of marrying a white women *within* the Cherokee community were considerably lower.[5] Traditional gender roles within the household decreased the demand for Cherokee land in

[3] I make a simplifying assumption that marriages are always preferred over remaining single. For a survey on gains from specialization in marriage, see Ribar (2004).

[4] In traditional Cherokee society, improvements to Cherokee land was owned by wife's side of the family. The residence of the married couple was also located near the wife's parents.

[5] The costs of such marriages appeared to be larger in the white communities. When Harriet Gold, a white women from a leading Connecticut family, and Elias Boudinot, a Cherokee, decided to marry, her hometown of Cornwall, Connecticut held a public rally against the marriage, burning effigies of herself and Boudinot. One of the members of the mob was Gold's brother (**?**).

these marriages. The intermarriage laws passed throughout the nineteenth century also did not mention specific laws against marriages with white women.

The decision of an individual is therefore to marry within the tribe or marry out. Given this setup, an individual will marry out (or intermarry) if the net present value of intermarriage is no smaller than the present value of marrying in:

$$r_W f(x_i, y_i) - C(x_i^P) \geq r_C f(x_i, y_i)$$

From this expression, this model predicts positive assortative matching in the intermarriage market. Thus, if a Cherokee decides to marry a white, the Cherokee would have likely descended from a family that previously intermarried. There are two explanations for this result: the human capital (or martial output) function increases when the abilities of a Cherokee and a white individual are similar; and the cost of intermarriage is lower for previously-assimilated families.

This model also predicts that as the cost of intermarriage increases, the degree of positive assortative matching in intermarriages increases.[6] Stated differently, when the social costs of intermarriage increases, the probability of marrying out falls but the degree of positive selection among those marriages will increase. This result will serve as the basis of the empirical strategy as the degree of positive selection should be greater with marriages between white males and Cherokees females.

3 Empirical Strategy

The previous framework suggests that the difference in average economic outcomes between a biracial (mixed marriage) household and a mono-ethnic (endogamous marriage) household is driven by the gains from intermarriage and the selection into intermarriage. exogamous and endogamous households is driven by the gains from intermarriage and the selection into intermarriage. These mechanisms can be neatly summarized with an equation commonly adopted in program evaluation (Angrist and Pischke, 2009, 14):

$$\underbrace{E[Y_i|M_i = 1] - E[Y_i|M_i = 0]}_{\text{observed gap}} = \underbrace{E[(Y_{1i} - Y_{0i})|M_i = 1]}_{\text{intermarriage premium}} + \underbrace{E[Y_{0i}|M_i = 1] - E[Y_{0i}|M_i = 0]}_{\text{selection effect}} \tag{1}$$

where Y_i is the outcome of interest (wealth or consumption) and M_i is the intermarriage indicator. The first term on the right-hand side is the average between the outcome of the intermarried (Y_{1i}) and the outcome of the intermarried individuals had they not intermarried (Y_{0i}). This term refers to the intermarriage premium and the leading reasons for this premium include improved linguistic skills, greater adoption of cultural norms and increased social integration from marrying a native

[6] This comparative static result is derived from $\frac{\partial \widehat{x}}{\partial C} = \frac{1}{f_x(r_W - r_C) - C'} > 0$, where $f_x = \frac{\partial f}{\partial x}$ and $\widehat{x}$ is the ability level that makes you indifferent between marrying within the tribe and marrying out.

(Dribe and Lundh, 2008; Gevrek, Gevrek; Duleep and Regets, 1999; Furtado and Theodoropoulos, 2011). The second term is the average difference of the outcome of the intermarried household had they not intermarried and the outcome of individuals in exogamous individuals. The selection bias, according to the theoretical model, is expected to be positive in intermarriages.

The second term is of course the key issue in determining the causal effect of intermarriage. This bias has been shown recently to be the driver of the observed wealth gap. For example, Kantarevic (2004) uses panel-data techniques to control for the selection problem on U.S. immigrants and find no evidence of a intermarriage premium. In addition, Nekby (2010) and Nottmeyer (2010) use panel data techniques to estimate the returns to intermarriage by immigrants in Germany and Sweden, respectively. After controlling for person-specific unobserved heterogeneity, both studies find no post-marriage effects on earnings for foreign-born males who marry natives.

The empirical set-up can be applied to my paper is the following ways. First, I will estimate the observed gap in wealth and consumption between mixed-married households and all-Cherokee households using the following model:

$$y_{i,c,t} = \beta \mathrm{M}_i + \mathbf{X}'_{i,c,t}\Omega + \tau_t + \phi_c + v_{i,c,t} \tag{2}$$

where $\mathrm{y}_{i,c,t}$ is the log value of wealth (and consumption) of household i who resided in township c in year t. M_i is the variable of interest, representing whether the individual i intermarried. In some specifications, I differentiate between a marriage between a Cherokee and a white and a marriage between a Cherokee and a biracial Cherokee. $\mathbf{X}_{i,c,t}$ is a vector that includes a constant and time-varying variables such as the age (and its square) of the head of household, the size of the family, a variable indicating whether a household member attended church, an indicator that is turned on if someone in the household worked on local road construction, a variable indicating whether the household leased any land, and a variable indicating whether the household rented their farm. The township fixed effects, ϕ_c, absorbs the unobservable characteristics that are unique to all households within a township and the year fixed effects, τ_t, are assumed to be constant across all households but vary across years. Finally, the time-varying error term $v_{i,c,t}$ satisfies the usual assumptions.

This empirical strategy simply pools the data and applies OLS to equation (2). Since the township fixed effects are included on the right-hand-side, the estimation of equation (2) will provide the within-township, cross-household relationship of intermarriage on economic outcomes. If the intermarriage effect is fully explained by the observable variables, then this strategy would eliminate any potential bias.

However, these baseline estimates are problematic since, according to the theoretical set-up, parental characteristics, x_i^P, are correlated with M_i. The second empirical strategy attempts to tease out this selection bias by controlling for parental fixed effects. Since x_i^P is constant within the family, I compare the economic outcomes of siblings who and who did not intermarry. Sibling-difference (or family fixed effects) models take into account the fixed unobservable characteristics, x_i^P, that are assumed to be shared by siblings, and, under the above assumption, are uncorrelated with M_i.

How would one expect the sibling-difference estimates of β compare to the OLS estimates in equation (2)? The theoretical model suggests that the OLS baseline estimates are positively biased: therefore, v_i and M_i are positively correlated. Therefore, controlling for x_i^P would reduce the positive bias obtained in the OLS estimate. Additionally, the size of the positive bias will vary by the ethnicity of the husband (or wife). Marrying a white husband should contain a larger positive bias than marrying a white wife.

4 Data

4.1 OIA Censuses

The main data used in this paper are drawn from household-level censuses taken by the Office of Indian Affairs (OIA), later renamed the Bureau of Indian Affairs. From roughly 1875 to 1906, the Interior Department collected annual censuses on Indian reservations and published reservation-level aggregates in the Annual Reports of the Commissioner of Indian Affairs.[7] A typical report contained information on the age and gender distribution on each reservation, the total number of acres broken and cultivated, the amount invested in livestock, and the total farm production. To my knowledge, the Eastern Cherokee Indian Reservation microdata are the only individual censuses to have survived from this period.

There is little evidence suggesting these censuses were imprecisely measured. Census enumerators were selected from members of the tribe and collected information on households living in the same township as themselves. The enumerators often took multiple days to enumerate their townships. For example, in 1903, Sibbald Smith, a Cherokee aged 24, took four days to enumerate 259 individuals within 60 households in Big Cove, the same township where Smith lived. In 1899, John Tahquette took eight days to collect information on 110 households living in two neighboring townships, Big Cove and Yellow Hill. Some enumerators took less time such as Jeff Arneach who took one and a half days to collect information; however, his particular township was occupied by only 74 individuals living in 15 separate households. Thus, the effort by enumerators appears to be uniform across censuses. Also, in most census years, a second person signed off on the accuracy of the census information.

The unit of observation is annual agricultural and population data on Eastern Cherokee households. The OIA data provide information on the degree of animal husbandry of each household by listing the number of heads of six types of livestock: cattle, sheep, pigs, fowls, horses and mules. For each household, the number of literate household members and the number of church members were also collected. These data were typically linked to the name of the head of household; however, in later years, the name, gender and age of each household member were also listed.

The main outcome of interest is the inflation-adjusted value of the heads of livestock. Livestock prices are taken from Carter et al. (2006)'s *Historical Statistics* (Tables Da969, Da971, Da973,

[7] These reservation-level data have been adopted by others, most notably Carlson (1981) who estimated the impact of allotment on American Indian farming.

Da984, Da986, Da1040) and those prices are deflated using the historical BLS price index (Carter et al., 2006, Table Cc1).[8]

The OIA censuses exist from 1894 to 1906 with some gaps between years. For example, there are no data for the year 1896. The data were organized into seven township within the reservation: Yellow Hill, Big Cove, Soco (sometimes called Wolftown), Birdtown, Nantahala, Cherokee County and Graham County. In some years, there were no data collected for certain townships: for example, in 1904, only data in Big Cove, Yellow Hill and Soco exist. In sum, there are 43 separate census books with a total of 2,473 household-by-year observations.

4.2 Eastern Cherokee Applications

The ancestry of each adult listed in the OIA censuses can be recovered by using the 1906-1909 Eastern Band Applications of the U.S. Court of Claims. Detailed familial information on any Cherokee Indian living at this time was collected due to a U.S. Court of Claims decision in favor of three separate civil suits filed against the federal government by both the Eastern and Western Cherokee tribes.[9] In order to receive a share of this award, the U.S. Court of Claims, under the leadership of the court-appointed commissioner Guion Miller, set the following eligibility requirements: (1.) the claimant must be living on May 28, 1906, the date of the award; (2.) the claimant must have been either a member of the Cherokee tribe during the time of one of the treaties listed in the lawsuit, most notably the 1851 Treaty with the Cherokees, or a descendant of a Cherokee who received treaty funds; and (3.) the individual cannot be affiliated with another tribe. From 1906 to 1909, a total of 45,847 applications representing roughly 90,000 individual claimants were collected by Miller. Eventually, 3,436 Cherokees living in North Carolina and 27,384 living in Indian Territory were approved to share in this award.

These applications contain a wealth of personal information on each applicant.[10] Each claimant furnished information on his or her English and Indian name, residence, date and place of birth, married status, name and age of wife or husband, English and Indian names of his or her parents, place of birth of parents, date of death of each parent, and the names of their children, siblings, cousins and grandparents on both their mother's and father's side. To verify authenticity, all applications were made under oath and supported by two witnesses who were acquainted with the claimant.

I use two strategies to determine the race of an ancestor. First, if a spouse of an applicant did not contain any Cherokee ancestors, then the enumerator would list this person by his or her race. Additionally, if a parent or grandparent of an applicant did not contain any Cherokee ancestors,

[8] There are some limitations in the price data. Data on the market price of chickens start in 1909. In place of annual prices, I use the average price per head from 1909-1921.

[9] As a consequence of selling the Cherokee Outlet to the federal government, a Congressional act in 1902 allowed the Cherokees to file suit against the federal government for violations of past treaty stipulations. Three suits, each regarding separate cases in which money was due to the Cherokees, were brought before the U.S. Court of Claims, and in May 1905, the court ruled in favor of the Cherokees. In 1906, Congress finally appropriated over $1 million plus interest which was distributed to the Cherokees on a per capita basis (Archives et al., 1967).

[10] Electronic copies of these applications are available through the military genealogical website, www.fold3.com.

then the enumerator would also typically write "white" beneath the parent's or grandparent's name. For example, Will West Arneach, a Cherokee (in blood quantum terms, $\frac{3}{4}$th Cherokee) born in 1849, had a mother Jenny Arneach (née Reed) whose father, Bill Reed, was white. Even though Bill Reed had a Cherokee name, *Wi-li S-ga-tsi*, the notary added the comment "white" below Bill Reed's name. Therefore, it is commonly straightforward to determine the racial identification of each family member.

A second strategy is used when the name of an ancestor, commonly a grandparent, is omitted from the applications. The missing information is likely due to the applicant's lack of knowledge of his or her grandparent's name. However, if the missing grandparent is not of Cherokee descent, then the variable of interest would contain measurement error. To resolve this potential issue, I locate the applicant in the 1900 or 1910 Indian schedules in the regular census which lists the degree of Cherokee blood for each individual. For example, if a Cherokee applicant contained only one white grandparent, then this Cherokee would be considered $\frac{1}{4}$th Cherokee in the regular census. By adopting this information, I can deduce the race of a grandparent whose name (and race) was missing in the applications.[11] This strategy was only necessary in a few cases.

Based on this information, I use the following simple indicator variable to capture the ancestry of each household head.

$$\mathrm{M}_i = \begin{cases} 1 & \text{if the } i^{th} \text{ individual married a individual who has at least one white ancestor;} \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Since the racial information of a household head's great grandparent is unknown, this variable equals one if the white ancestor was either a parent or grandparent. In the case that the only white ancestor of a household head was his great grandparent, then this individual would be considered $\frac{7}{8}$th Cherokee (or $\frac{1}{8}$th white). For this analysis, any Cherokee who is at least $\frac{7}{8}$th Cherokee is considered a full-blooded Cherokee. This is not much of a concern here since this reservation was isolated from neighboring white settlements until the mid-nineteenth century.

4.3 Descriptive Statistics

The descriptive statistics of the variables used in this analysis are listed in Table 1. I create an unbalanced panel with successful matches between the OIA censuses and the applications dataset. In sum, 2,473 observations were listed in the OIA census and 2,000 (with 1,656 containing wealth data) were matched to the applications data. This matching rate (81%) did not generate noticeable differences between the outcomes of matched and unmatched households.[12] Table 1 reveals that

[11] This exercise was also used for the entire sample to verify the racial information in the applications. In the rare case where the blood quantum differed between the applications and the Indian schedules, I queried additional genealogical websites like familytreemaker.com and ancestry.com to determine the racial ancestry of a Cherokee head of household.

[12] Differences between the matched and unmatched samples are located in Appendix Table A1.

livestock wealth are noticeable higher on average in households with at least one white ancestor. The number of English users and church members per household also differ by ancestral group. Thus, simple comparisons of livestock wealth across groups may be misleading.

The descriptive statistics highlight the correlation between intermarriage and wealth accumulation. Part of this correlation might be driven by observable differences across households with different racial ancestries. More importantly, those who decided to intermarry and those who did not may differ in unobservable ways. The empirical strategies adopted in this paper attempt to assess whether this correlation holds when these differences are accounted for.

5 Baseline Results

Figure 2 provides an initial view of the relationship livestock accumulation and the cultural ancestry of Cherokee Indians. The effect of having a white man as the head of household is positive, statistically significant, and relatively constant across census years.

Table 2 presents the OLS regressions of equation (2). In each regression, the standard errors are robust to heteroskedasticity.[13] In each regression, the reference group is households with only Cherokee ancestry on both the husband's and wife's side. Column 1 shows that households headed by a white man and a Cherokee female were 30 log points (36 percent) wealthier than households headed by both a Cherokee male and female. The wealth advantage of households with white males increased if the wife was of white descend. In this case, the households comprised of male and female with white ancestry were 34 log points (or 40 percent) wealthier than households without any white ancestry. Last, households with Cherokee males and white females were, on average, less wealthy than households with only Cherokees, though this estimate becomes positive when controls are added.

In column 2, the wealth effect of marrying a white male is qualitatively unchanged when a full set of time-varying household characteristics are included, In this regression, the coefficient on the husband's white ancestry variable is 0.292 (s.e.=0.122) and is significant at 5% level while the coefficient while the coefficient on the wife's ancestry indicator is positive but insignificant. Column 3 shows the results from a regression where township fixed effects are included. Households with a white husband and a Cherokee wife hold a 32 log points advantage in wealth even when locational differences are controlled for. However, the effect of marrying a white female is positive but statistically insignificant at standard levels.

The same exercise is run for determining the effect of marrying out on annual consumption. This measure is likely to be driven more by differences in human capital. In the most parsimonious specification (column 4), households with white males and Cherokee females produced 22 log points more consumption than households without any white ancestry. The role of marrying white females is mixed. In Column 4, the effect of marrying white females was positive and large with respect to consumption levels. Columns 5 and 6 show that the coefficient do not vary qualitatively when

[13] I used other methods such as closeting by surnames and the results are qualitatively similar.

controls are added.

In sum, the positive effect of wealth from intermarriage is the strongest if the husband descend from at least one exogamous marriage. Households with either white wives or wives who descended from an exogamous marriage were not wealthier than households with full-blooded Cherokee wives.

6 Sibling-Difference Results

As mentioned above, the pooled OLS regressions cannot account for unobserved family background effects which might influence the decision to intermarry.

Table **??** shows the results from the sibling-difference regressions which absorb the unobserved fixed characteristics shared by children from the same parents. The individual controls, when included, are listed in the footnote to the table. The first four columns control for family fixed effects by defining the family as siblings from the same parents. Since several Cherokee men had children with multiple women, the last two columns allows for the family fixed effect to be constant among full- and half-siblings. There are 110 observations where deviations of intermarriage from the family mean occurred. When I expand the family to include half-sisters, the sample size increases to 125.

Column 1 shows that Cherokee sisters who married men with at least one white ancestor held more wealth than sisters who married full-blooded Cherokee men. The coefficient on husband's with white ancestry is 0.450 (s.e.=0.123) without controls and 0.478 (s.e.=0.126) with controls. When individual controls are added and township and year dummies are included, a household with a husband with white ancestry is 51 log points greater than a household with a husband with only Cherokee ancestry. Thus, the family fixed effects, when compared to column 2 of equation 2, increase the estimated effect of intermarriage by 67%. This suggests that the estimates without family fixed effects are biased downward, which implies that controlling for family fixed effects only strengthens the earlier result.

Column 4 isolates the gains from marrying a white male from the gains from marrying a male with biracial ancestry. Sisters who married a white male are estimated to have held 46 log points more in livestock wealth than sisters who married full-blooded Cherokees. The wealth gain from marrying a male with mixed-white ancestry is slightly larger (52 log points) and highly significant; however, the difference in coefficients among these two measures of intermarriage are statistically indistinguishable.

The final two columns make full- and half-sibling comparisons and the results are qualitatively similar to columns 3 and 4. In each case, the coefficient on the intermarriage indicators is positive, highly significant and larger than the OLS estimates in Table 2. For example, in column 5, the fixed effect estimate of β suggests that the wealth difference between males with and without white ancestry was 40% of a standard deviation of the outcome, when accounting for family fixed effects.

The basic findings from estimating the fixed effect estimate of intermarriage is that unions of white males and Cherokee females earn significantly more than unions of Cherokee males and females. I estimate some additional fixed effect specifications to assess whether functional form errors influence the estimated coefficient on β. I modify the specification in Table **??**, column 3 in four ways: (1.) by omitting potential outliers which are defined as households located in the bottom 12.5% and top 97.5% of the outcome distribution; (2.) by eliminating all observations with a Cook's distance greater than one; (3.) by clustering at the township level which controls for serial correlation within Cherokee towns; and (4.) by introducing a lagged dependent variable as a check against mean reversion (see Arellano and Bond, 1991). Each specification is estimated using OLS except for column 2 which uses weights from an iterative process that down-weights observations with large residuals and Column 4 uses GMM to alleviate the bias from including a lagged dependent variable on the right-hand-side.

Table **??** reveals that the positive coefficient on the husband's white ancestry indicator is relatively stable and highly significant across each specifications. These alternative models continue to reveal that the white ancestry effect is much stronger than the results without the family fixed effects.

In sum, the fixed effects estimates of β are positive, highly significant and stronger than the cross-household comparisons. An interpretation of these results implies that there may have been negative selection into intermarriage by Cherokee families. Another interpretation is that the sibling variation in intermarriage is endogenous in this equation **??**. If a more acculturated sibling is more likely to intermarry, then the sibling-difference estimates could be positively bias. To assess the role of unobserved factors on the positive effect of white ancestry on animal husbandry, I follow Altonji et al. (2005) in estimating the ratio of selection on unobservables to selection on observables needed to overturn the positive effect of each cultural indicator. Intuitively, this approach determines how large the relative effect of the unobservables need to be to explain away the return of white ancestry. A large and positive ratio indicates that the covariance between white ancestry and unobservables needs to be substantially larger than the covariance between white ancestry and the controls to explain away the positive association. A negative ratio on the other hand shows that the effect of white ancestry increases in size when the controls are added which suggests that unobserved skills do not generate an upward bias.

As shown in Tables 2 and **??**, when controlling for time-varying factors and family fixed effects, estimates of the gain from intermarriage become stronger. Therefore, this ratio is negative (equal to -4.07 when compared to a regression without any controls), which implies that if sibling-specific factors could be controlled for, the observed wealth difference by intermarriage would be even greater.

7 Concluding Remarks

In the paper, I use historical data on wealth accumulation to test whether intermarriages influence wealth accumulation among the Eastern Cherokee Indians living in North Carolina at the turn-of-the-twentieth century. I find that cross-household comparisons yield the standard results:

male heads of household that contain at least one white ancestor are substantially wealthier than households without white ancestry. However, once unobservable family background characteristics are controlled for, the fixed effects estimates reveal that the intermarriage effect is reduced.

This paper has disentangled the causal effect of intermarriage on livestock wealth from selection effects into intermarriage. The positive wealth and consumption differences from intermarriage among the Cherokees do appear to be driven by selection into intermarriage. The positive selection is the stronger in marriages between white males and Cherokee females. The simple model presented here suggests that differential costs to intermarriage may be driving the positive selection into those marriages.

References

Adachi, H. (2003, December). A search model of two-sided matching under nontransferable utility. *Journal of Economic Theory 113*(2), 182–198.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1), 151–184.

Anderson, T. L. and D. Lueck (1992, October). Land tenure and agricultural productivity on indian reservations. *Journal of Law and Economics 35*(2), 427–54.

Anderson, T. L. and D. P. Parker (2008). Sovereignty, credible commitments, and economic prosperity on american indian reservations. *Journal of Law and Economics 51*(4), 641–666.

Angrist, J. D. and J.-S. Pischke (2009). Mostly harmless econometrics.

Archives, U. S. N., R. Service, and G. Miller (1967). *Records Relating to Enrollment of Eastern Cherokee.* National Archives, National Archives and Records Service, General Services Administration.

Arellano, M. and S. Bond (1991, April). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *Review of Economic Studies 58*(2), 277–97.

Beaulieu, D. (1984). Curly hair and big feet: physical anthropology and the implementation of land allotment on the white earth chippewa reservation. *The American Indian Quarterly 8*(4), 281–314.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy 81*(4), 813–46.

Carlson, L. A. (1981, April). Land allotment and the decline of american indian farming. *Explorations in Economic History 18*(2), 128–154.

Carter, S. B., S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, and G. Wright (2006). *Historical Statistics of the United States: Millennial Edition.* Cambridge: Cambridge University Press.

Cornell, S. E. and J. P. Kalt (1992). *What can tribes do?: Strategies and institutions in American Indian economic development.* University of California, Los Angeles. American Indian Studies Center. American Indian Studies Center, University of California, Los Angeles Los Angeles, CA.

Dribe, M. and C. Lundh (2008). Intermarriage and immigrant integration in sweden an exploratory analysis. *Acta Sociologica 51*(4), 329–354.

Duleep, H. O. and M. C. Regets (1999). Immigrants and human-capital investment. *American Economic Review*, 186–191.

Ehle, J. (1989). *Trials of tears: The rise and fall of the Cherokee nation.* Anchor.

Furtado, D. and N. Theodoropoulos (2011). Interethnic marriage: a choice between ethnic and educational similarities. *Journal of Population Economics 24*(4), 1257–1279.

Furtado, D. and S. Trejo (2012, February). Interethnic Marriages and their Economic Effects. IZA Discussion Papers 6399, Institute for the Study of Labor (IZA).

Gevrek, Z. E. Interethnic marriages and the labor market integration of immigrants. Working paper, University of Konstanz.

Gitter, R. J. and P. B. Reagan (2002). Reservation wages: An analysis of the effects of reservations on employment of american indian men. *The American Economic Review 92*(4), 1160–1168.

Jr., R. G. F. (2007, Spring). Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th Century. *Journal of Economic Perspectives 21*(2), 71–90.

Kantarevic, J. (2004, May). Interethnic marriages and economic assimilation of migrants. IZA Discussion Papers 1142, Institute for the Study of Labor (IZA).

Kuhn, P. and A. Sweetman (2002). Aboriginals as unwilling immigrants: Contact, assimilation and labour market outcomes. *Journal of Population Economics 15*(2), 331–355.

Meng, X. and R. G. Gregory (2005, January). Intermarriage and the Economic Assimilation of Immigrants. *Journal of Labor Economics 23*(1), 135–176.

Murchison, A. (1928). Intermarried-whites in the cherokee nation between the years 1865 and 1887. *Chronicles of Oklahoma 3*(299-326).

Nekby, L. (2010, November). Inter- and Intra-Marriage Premiums Revisited: It's probably who you are, not who you marry! SULCIS Working Papers 2010:12, Stockholm University Linnaeus Center for Integration Studies - SULCIS.

Nottmeyer, O. (2010, August). Does Intermarriage Pay Off? A Panel Data Analysis. IZA Discussion Papers 5104, Institute for the Study of Labor (IZA).

Perdue, T. (1999). *Cherokee women: Gender and culture change, 1700-1835.* Bison Books.

Ribar, D. C. (2004, January). What Do Social Scientists Know About the Benefits of Marriage? A Review of Quantitative Methodologies. IZA Discussion Papers 998, Institute for the Study of Labor (IZA).

Trosper, R. L. (1978). American indian relative ranching efficiency. *The American Economic Review 68*(4), 503–516.

Table 1: Descriptive Statistics [Matched OIA/Applications Sample]

	Full Sample		White Ancestry		Cherokee Ancestry	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Livestock Wealth ($1900)	106.66	[116.76]	126.47	[136.28]	98.77	[107.05]
Cattle (heads)	2.81	[3.29]	3.34	[3.97]	2.60	[2.95]
Sheep (heads)	1.47	[4.55]	2.27	[5.78]	1.15	[3.91]
Pigs (heads)	5.32	[8.22]	5.50	[8.35]	5.25	[8.16]
Fowls (heads)	26.76	[30.83]	28.15	[36.63]	26.21	[28.17]
Horses (heads)	0.27	[0.67]	0.41	[0.87]	0.22	[0.56]
Mules (heads)	0.06	[0.36]	0.07	[0.34]	0.05	[0.36]
Acres Cultivated	12.28	[11.36]	12.71	[13.28]	12.10	[10.49]
Age of Head	44.40	[14.73]	44.89	[15.15]	44.21	[14.56]
Family Size	5.16	[2.26]	5.08	[2.54]	5.19	[2.14]
Males	2.75	[1.48]	2.85	[1.61]	2.71	[1.43]
Females	2.40	[1.38]	2.23	[1.45]	2.47	[1.35]
English Language Users	1.82	[2.04]	2.95	[2.55]	1.37	[1.58]
Church Members	0.65	[1.02]	0.77	[1.15]	0.61	[0.96]
Observations	1766		546		1220	

Notes: Mean values are weighted by average family size. Livestock and crop values as well as the price deflator are taken Carter et al. (2006), Tables Cc1, Da597, Da770, Da742, Da735, Da678, Da969, Da971, Da984 and Da1040.

Table 2: Baseline OLS Results

	logged Wealth			logged Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Husband with White Ancestry	0.306***	0.292**	0.318**	0.224***	0.169*	0.221**
	(0.099)	(0.122)	(0.125)	(0.074)	(0.100)	(0.099)
Wife with White Ancestry	-0.314**	0.069	-0.063	-0.075	0.061	0.009
	(0.150)	(0.177)	(0.184)	(0.073)	(0.110)	(0.112)
Both Husband and Wife with White Ancestry	0.347*	-0.083	0.081	0.257**	-0.086	-0.104
	(0.189)	(0.242)	(0.246)	(0.117)	(0.180)	(0.179)
F-statistic on controls		20.41	15.02		5.54	4.26
Controls	No	Yes	Yes	No	Yes	Yes
Township Fixed Effects	No	No	Yes	No	No	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,571	1,175	1,175	1,379	1,017	1,017
R-squared	0.047	0.236	0.259	0.087	0.267	0.301

Notes: The Huber-White standard errors are in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels.

Table 3: The Intermarriage Effects of Marrying White Men: Family Fixed Effects Regressions

	Panel A: Parental Fixed Pairs					
	logged Wealth			logged Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Husband with White Ancestry	-0.153	-0.035	0.069	0.085	-0.097	0.088
	(0.157)	(0.339)	(0.343)	(0.137)	(0.356)	(0.127)
Within-Family Variation in M_i	138	138	138	138	138	138
Observations	1,466	1,165	1,165	1,138	1,008	1,277
R-squared	0.535	0.593	0.600	0.558	0.581	0.522
	Panel B: Surname Fixed Effects					
	logged Wealth			logged Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Husband with White Ancestry	0.057	-0.152	-0.077	0.227	0.003	0.354
	(0.115)	(0.177)	(0.186)	(0.143)	(0.212)	(0.221)
Within-Surname Variation in M_i	399	399	399	399	399	399
Observations	1,466	1,165	1,165	1,138	1,008	898
R-squared	0.440	0.557	0.562	0.492	0.501	0.568
Individual Controls	No	Yes	Yes	No	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township Fixed Effects	No	No	Yes	No	No	Yes

Notes: The Huber-White standard errors are in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels.

Table 4: The Intermarriage Effects of Marrying White Women: Family Fixed Effects Regression

	Panel A: Brother Pairs					
	logged Wealth			logged Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Wife with White Ancestry	0.120	3.117	3.404	0.040	-4.583	0.110
	(0.292)	(5.278)	(5.228)	(0.274)	(5.379)	(0.188)
Within-Family Variation in M_i	42	42	42	42	42	42
Observations	1,465	1,164	1,164	1,137	1,007	1,276
R-squared	0.619	0.646	0.651	0.644	0.636	0.615
	Panel B: Surname Pairs					
	logged Wealth			logged Consumption		
	(1)	(2)	(3)	(4)	(5)	(6)
Wife with White Ancestry	-0.292**	-0.450***	-0.467***	0.106	-0.044	-0.084
	(0.136)	(0.169)	(0.170)	(0.099)	(0.132)	(0.131)
Within-Surname Variation in M_i	110	110	110	110	125	125
Observations	1,571	1,175	1,175	1,219	1,017	1,017
R-squared	0.396	0.533	0.541	0.487	0.499	0.509
Individual Controls	No	Yes	Yes	No	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Township Fixed Effects	No	No	Yes	No	No	Yes

Notes: The Huber-White standard errors are in parenthesis. *, **, and *** represent significance at the 10%, 5%, and 1% levels.

Table A1: Differences between Matched and Unmatched Samples

	All Households			Excluding "Blank" Observations		
	Unmatched	Matched	*p-value*	Unmatched	Matched	*p-value*
Livestock Wealth ($1900)	98.05	105.8	0.0522*	139.4	124.7	0.0745*
	(117.7)	(115.3)		(119.5)	(115.8)	
Pigs (heads)	4.175	5.173	0.0192**	5.975	6.098	0.6157
	(6.459)	(8.007)		(7.052)	(8.397)	
Fowls (heads)	28.23	26.23	0.4659			
	(48.17)	(30.47)				
Horses (heads)	0.307	0.278	0.8085			
	(0.836)	(0.679)				
Mules (heads)	0.0785	0.0542	0.3311			
	(0.458)	(0.340)				
Crop Value ($1900)	43.32	49.48	0.0071***	56.95	55.85	0.9108
	(64.22)	(88.11)		(70.98)	(94.21)	
Wheat (bu.)	3.995	4.826	0.0113**	5.166	5.560	0.2375
	(13.19)	(15.47)		(15.01)	(16.64)	
Corn (bu.)	104.8	92.80	0.8632			
	(159.9)	(133.7)				
Potatoes (bu.)	17.46	25.74	0.1956			
	(66.34)	(155.6)				
Oats (bu.)	2.962	2.640	0.2255			
	(10.56)	(10.57)				
Acres Cultivated	11.75	12.01	0.0739*	14.67	13.22	0.5415
	(17.35)	(11.12)		(19.33)	(11.41)	
Crop Output per Acre ($)	3.946	4.546	0.1699			
	(5.458)	(5.978)				
Family Size	5.415	5.076	0.0010***	5.711	5.234	0.1268
	(2.860)	(2.257)		(2.978)	(2.256)	
English Language Users	2.430	1.811	0.0392**	2.606	1.936	0.0751*
	(2.787)	(2.017)		(2.991)	(2.075)	
Church Members	0.820	0.646	0.1342			
	(1.172)	(1.005)				
Observations	473	2000		301	1656	

Notes: The standard deviation is listed in parentheses below the mean. The p-value is from a standard two-sided comparison of means test. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: See Table 1.