Genealogies as a Resource for studying Kin Networks in Time and Space

By

Alice Bee Kasakoff
Department of Anthropology
University of South Carolina
Columbia, SC 29206

Email: Kasakoff@sc.edu

June 15, 2005

Prepared for delivery at the IUSSP Conference July 2004

DRAFT- DO NOT CITE

The founder of "modern" kinship studies in anthropology, W.H.R. Rivers, inaugurated the systematic collection of genealogies by cultural anthropologists. For a long time, noone could return from anthropological research without one, Kinship was thought to be the central organizing principle of at least "tribal" societies. Genealogies were thought of as a universalistic template that would facilitate the comparison of social forms. In his classic article, originally published in 1910 "The Genealogical Method of Anthropological Inquiry" (Rivers 1910), he mentioned several uses they could have, particularly among people of "lower culture".   The situation is, of course, quite different today. Now, within anthropology, scholars are deconstructing this method (see, for example, Franklin Relative Matters and the work of Bosquet); it is seen as culturally biased and it is no longer required that an anthropologist return from the field with a genealogy. Within historical demography, by contrast, genealogical research is undergoing something of a revival. [1]

The work I discuss today returns to Rivers' concerns and methods but applies them to much more "modern" people, the native born in American North between 1650 and 1880. Although I am not interested in using genealogies to construct kinship terminologies or marriage systems regulated by clans, as Rivers was,   I agree with his early assessment of their usefulness for studying inheritance, the social obligations between kin, and migration. To Rivers, one advantage of using them was their "concreteness" by which he meant that they provided evidence for patterns of actual behavior which meant that the investigator did not have to rely solely upon people's idealized versions of what should happen in their society. He emphasized that through them one could "formulate laws regulating the lives of people which they have never formulated themselves" (1968:107) This, I think, is still an important advantage to using such materials.   Rivers bragged that "By means of the genealogical method it is possible with no knowledge of the language and with very inferior interpreters, to work out, with

the utmost accuracy systems of kinship so complicated that Europeans who have spent their whole lives among the people have never been able to grasp them." (1968:107)

Today I am going to draw from the research my husband and colleague John W. Adams and I have been doing for over 20 years compiling a genealogical database to document the lives of the descendents of the settlers of the American North, who came from England before 1650. It seems to us that they also have systems of kinship so complicated that they have not yet been accurately grasped, not because they are "exotic" but precisely because of the blindness that any society encourages towards its most intimate obligations: the way in which, as Rivers recognized, the past is distorted by people in the present to advance various ideological agendas. This happens not only among "lower" cultures but in the "higher" as well.

At the beginning of our work, trained in the anthropological tradition of Rivers, we saw the genealogies that had been collected by New Englanders as an important corrective to the fashionable work of family reconstitution. Family reconstitution sought to reconstruct the populations of individual towns and did not include migrants. Since many of the genealogies constructed by New Englanders after the middle of the 19th century sought to follow descendents of the founders of the population wherever they resided across the United States, we thought they would furnish a much more accurate demographic picture than family reconstitution. How they managed to construct genealogies of this great a spatial reach without the aid of web sites, the Church of Latter Day Saints and census microfilms, would be the focus of quite a different project. We do know that they were aided by the newsletters of genealogical societies and by the demographic moment: most of the genealogies we use in this study were compiled at a time when the grandparents of the compilers would have been able to link branches of the family living quite far apart.

Americans historically moved permanently over greater distances than did Europeans at the time – they were, in fact, settling the continent – a difference we can now attest to because we have been comparing data on the life courses of the people we are following with those available in European databases (Kok, Adams and Kasakoff 2000, Egerbladh, Adams and Kasakoff 2005), thus, when we began this project, it seemed important that historical demography of America use materials that included migrants rather than reconstruct the populations of single towns. In most of our work we have treated the genealogical database as though it were a random sample of individuals, useful for charting the demographic characteristics of a "population" which we think it represents: a population which became at the time of the civil war, the Yankees. In fact the map of where the descendents we are studying were living at mid 19th century mirrors almost exactly the counties which voted for Lincoln in the election of 1862.

In most of the work we have done we have tried to ignore the "family-ness" of the sample, but in what follows I will exploit it and ask questions about how being descended from a common ancestor many generations back might influence the lives of their descendents. My goal is to use genealogies to identify turning points in the histories of the families, roads taken and not, as well as general patterns of family

dispersion and link them to economic and demographic circumstances in the lives of their members. I want particularly to show their usefulness at different temporal and spatial scales, particularly scales that are not ordinarily visible to social scientists and to family members themselves. What follows is simply a beginning in this direction.

A Genealogy of Genealogy in Historical Demography

Most of the work using genealogies in historical demography had similar goals to those we had at first and sought to discuss questions of fertility and mortality, not kinship. In the early 20th century information on fertility from genealogies that had been collected since the 1850's when the New England Historic Genealogical Society was founded were codified and tabulated in two papers published in the Journal of the American Statistical Society. They were used to show that the native born Americans had reduced their fertility and urged that American women go back to the ways of their ancestors lest they be engulfed by the more fertile and more recent immigrants. These papers were part of the movement that eventually led the United States to restrict immigration in the 1920's.

In the late 1970's and early 1980's people began to use genealogies to shed light on population growth in the world following World War II.   Hollingsworth used genealogies of the British peerage to generate demographic statistics. Several "Big History" projects began, some of which continue today. Fogel and his students used American "family histories" to create a sample they thought would mirror the American population, but at first they only sampled one line from each family, randomly selecting one sibling to follow in each generation, a method which shows their concern with demography rather than kinship. Later, they included everyone in the families they selected. Other projects began in other locations: The First French Canadians (DesJardins), Chicoutimi (Bouchard), the TRA project for France. Recently genealogies have been used to study migration during the transformation of Europe to an industrialized and urbanized at the end of the 19th century: Pooley and Turnbull in Britain and Paul-Andre Rosental in France.

But it is only recently have scholars begun to exploit the fact that genealogies are by definition about families. Rosental has done the most in this regard, linking migration destinations and chains of relatives, and contrasting ways individuals use their networks and how they developed over space and time. Historical demographers using family reconstitution have also become interested in family issues. The Eurasia project have used their family reconstitutions to examine how the family is involved in adapting to crisis, such as the death of the father (e.g. Oris, When Dad Died) or other life course transistions (Neven   ). And recently there has been interest in how the family negotiates marriage for different siblings (Manfredini   ) and old age care. But these studies, because they are usually based upon single villages, are limited in understanding how kin who lived away from the village might have affected the lives of their relatives there. And, to study the roles of more distant kin within the village, they rely upon the number of people within the village with the same surname, rather than a detailed analysis of genealogical relations.

Campbell and Lee are currently combining the study of genealogies and family reconstitution in their studies of China where genealogies were systematically kept over generations.

In what follows I will be using the genealogical data we have to examine larger spatial and temporal units.

What is a Genealogy ?

A major distinction is between ascending genealogies, often called pedigrees, and descending genealogies. When possible, for demography, descending genealogies should be preferred because ascending genealogies include only individuals who have had children. They will, therefore, be biased towards the fertile members of the society and those who lived longest and had the greatest potential for having children. Lines which have died out will not be remembered or included. To the extent that the migration patterns of people who did not have children differ from those of people who did, ascending genealogies may also bias studies of migration.

Although all genealogies attempt to reconstruct the biological connections between individuals over generations in a population, each set of genealogical materials is unique. Their similarity is deceptive and most of them bear the marks of the records from which they were constructed. For example, compared with genealogies constructed from police or catechial registers, the check-in times in most genealogies are determined by vital events. The "mesh" therefore is relatively coarse grained and there are many periods of the life course that would be of interest from the point of view of kinship when few if any such events occur and thus are not accessible to the genealogist. The two most important are the teenaged years and process of leaving home and the increasingly long period between the birth of the last child and widowhood or death. (Kok, Adams and Kasakoff).

Genealogies also lack systematic economic information, such as occupations and measures of wealth.

Most importantly, genealogies do not have systematic information about the actual contact between relatives. They can be seen as providing information on "potential" relationships, not actual ones. For this reason, especially if one is interested in kin relations, it is important to link the genealogies to other sources. Rosental, for example, used marriage registers to identify witnesses to marriages and discover whether people relied upon relatives for this purpose or not, which he could link to their migration histories. In our case, we have linked the genealogies to the censuses available every 10 years for the United States. This not only provides information on wealth and occupations for the years 1850 through 1870, it also gives some indication of interaction among close kin in the form of households. And, at least every 10 years one gets information about the whereabouts of people that is not dependent upon vital events.

There are also issues of spatial boundedness. In theory genealogies should trace descendents wherever they reside but in practice each genealogy or set of them is a compromise between an ideal of tracing relations of parentage and 'blood" and the available records which means a geographic restriction of some sort. For example, few genealogies cross national boundaries to trace descendents.

There are two other problems worth noting. First, if the researcher relies on genealogies compiled by others, they will not be useful for the study of the entire life course of the people born close to the time of publication. This is because the lives of many of the people have not been traced to their conclusion. This biases statistics on mortality in particular since only those who have died young will have death dates in the genealogy.   Sometimes the temporal reach can be extended by modern genealogical research and linkage to censuses, but careful attention should be paid to distinguishing actual temporal trends from those caused by the falling off of data for the more recent cohorts.

Second, there is the difficulty of following women. Many genealogies make the attempt but few have succeeded. We selected our genealogies because they did include women and their descendents, but in fact, it is the exceptional genealogy that follows descendents of women over several generations. At first we attempted to follow the lives of women as well as men, but we discovered that data on women was biased especially where migration was concerned. When we looked at the distances from which spouses came, there was a curious gender difference: men chose spouses from farther away than did women. At the time there was no reason to believe that the women marrying into the family were any different from those born into the families in their choices of spouses and closer examination showed that the difference stemmed from the fact that information on the origin of spouses was missing from many more of the women's records than the men's. Genealogists would note the husband's name and often his pedigree when he came from the village where the wife lived, and often from nearby villages, but when they came from farther away, they did not. This has led us to study only the men in the genealogies we have and certainly biases the nature of the kinship networks we are able to study in this paper. Other studies have also noted this problem (Rosental 2004).

The Genealogical Sample for the American North

We based our database upon nine published genealogies (Bisbee 1956, Chaffee 1909, Farwell 1929, Faunce 1973, Greely 1905, Holman 1928, Pelton 1892, Shedd 1921, Wellman, J.W. 1918). Each began with a founder who came to what is now Massachusetts during what is known as The Great Migration, that is, before 1650. New England itself grew largely through the natural increase of progenitors like these, as noted by Malthus.   The result is a set of 9 descending genealogies. Although there is abundant information on the women born into the families, several are not followed after marriage, and therefore, we limit our discussion today largely to male kin traced through males. We have linked the genealogical information to the decennial censuses and today we will be taking examples from 1850 and 1860, drawing upon information about wealth

and occupation of the men we have been tracing from the censuses for those years. In our analyses we also use the county level published census data for 1860 to describe the economic landscape of the sending area we are going to analyze.

Because the population increased exponentially, there were 2088 male descendents of the 9 progenitors in the male line over age 20, known to be alive in 1850, of whom we were able to locate 1651 on the census that year (Adams and Kasakoff 1991).[2]   The population was quite rural: in 1860 only 16 % were living in cities and towns. (which we defined as places with more than 5000 persons in 1860). [3]

Representativeness

A genealogy is not the same as the bounded containers that demographers are used to studying, an administrative unit with clear boundaries. Biological lines inevitably cross those boundaries. Yet because most genealogies are compiled from administrative records and few genealogists systematically trace descendents into other countries, any genealogy reflects some kind of compromise and does not contain all the descendents of particular ancestors.   In fact, any genealogy represents an intersection of biological lineages and closed containers. Because of this double restriction, a genealogy can't represent "the general population". By definition, it must exclude the recent arrivals whose genealogies would have to be traced back to their origins outside the closed container. In fact, a genealogy is the record of the spread of an ethnic group within some kind of bounded unit and the question should be whether the particular set of genealogies chosen "represents" the history of that particular group.

These issues are well illustrated by an attempt to compare the sample we have assembled for the American North with a random sample drawn by IPUMS from the manuscript census, constructed to represent the entire population of the United States. We chose the census year 1850 because that was the first census in the United States to list every individual by name and to include economic characteristics and state of birth.

In constructing the sample for the comparison, we limited ourselves to persons born in the states where the people we were studying had been born and this eliminated the foreign born. Most of the genealogical sample had been in the United States 7 or 8 generations by 1850[4] yet there was no way to distinguish people descended from the earliest migrants from those descended from more recent migrants in the census. It was only in 1870 that the census asked about the birthplaces of parents, and in 1880 the grandparents. Only if an area had not been the destination of more recent immigration, then, could a genealogical sample "represent" the general population. In fact, most of the native born residents of New England in 1850 were, in fact, descended from these earliest settlers. We know that after 1650 people from England settled in the Middle Colonies and the South so there were few new arrivals to New England after what is called "the Great Migration". The more recent arrivals began to come in the 1840's, from Ireland and Germany, and their birthplaces would be marked on the census; their native born children would be too young to affect estimates of wealth or occupation.

It is well known that the United States was settled in waves that proceeded from East to West and that the latitude was a major limit to settlement (Fischer Albion's Seed, Matthews, Steckel). In fact the group we are studying migrated East and remained above the latitude where the variety of corn they had learned to grow from the Indians would prosper. They moved from Massachusetts where the sample was originally drawn, to other New England States, then into New York, Ohio and Pennsylvania, after the revolution in 1776 and from there into the middle west. Their migration effectively stopped at the Mississsippi River in 1850. However, in the newer areas they frequently settled only in the northern parts of states like Ohio and Indiana. As for New York, they settled in the Western areas that were opened to White settlement after the Revolution.

Rather than sample on the basis of where people were living, it seemed wise to take a sample from IPUMS that was like a genealogy because the people were **born** in the states where the people we were following had been born. We ask how well the genealogical sample mirrors people taken from a random sample of the census born in the states where the group we were studying were born, that is, the American North. We could not restrict it to only those parts of the states where New Englanders settled because the census only gave the state of birth, not any finer division. We included all the states in the North where the people in the genealogical sample had been born except for Pennsylvania which we excluded because so few of the people we were studying had been born there – they had settled only a small piece of the state adjacent to New York State - yet it had a very large population in 1850.   In what follows we compared the households we found on the census with the IPUMS households of people born in Northern states. We found only 72% of the men we knew to be alive in 1850 on the census; although differences between those we found and those we did not find might account for some of the differences, we think they result from the different migration history of genealogical sample.

The genealogical sample and the IPUMS sample of native Northerners were surprisingly simple. In terms of wealth and occupation, the genealogical sample was slightly more wealthy than the census sample. The census in 1850 asked only about the value of real property.

Table 1: Comparison of Wealth between genealogical data and IPUMS sample: 1850 census, males 20 years and older

	Genealogical Data	IPUMS Sample
% farm	45 %	41 %
% 0 real property	54 %	61 %
25 % real property	$ 800	$ 500
Median real property	$ 1500	$1000
75 % real property	$2500	$2500

Note: % farm is the proportion who were listed as "farmer" in the census.

Table 2: Wealth of farmers and non-farmers in 1850

	Farmers		Non-Farmers	
	Genealogy	IPUMS	Genealogy	IPUMS
% 0 real property	29 %	40 %	69 %	75 %
25 % real property	$800	$600	$600	$400
Median real property	$1500	$1200	$1200	$900
75 % real property	$2500	$2500	$2200	$2000

These differences were probably a result of two differences between the genealogical sample and the census sample: the forbears of the genealogical sample had been in the United States longer than many of the native born in New York State and areas farther West and thus could be expected to have accumulated more property; and, the genealogical sample was more rural than the census sample. In the both the census sample and the genealogical sample, farmers had more real property than did men who were not farmers.   The genealogical sample had slightly more farmers.

This example illustrates the importance of considering migration history before using a genealogical sample as "representative" of a particular bounded spatial unit.   If the genealogies go back several generations, the likelihood is that more recent migrants are excluded. This will have several important effects. There are going to be more relatives that can be traced since the earlier connections cannot be traced for more recent migrants. The descendents of early migrants will be advantaged economically and thus perhaps be more able to retain their kin connections, perhaps because they have taken the good land and established farms that remained stable. The comparative study we made with Northern Sweden illustrates some of the differences. In that study we generated a sample that would be similar to a genealogical sample by selecting couples present in the earliest records, some 100 years before, and tracing their descendents. We compared the descendents of these early couples with the entire population resident in the parish in 1850. Descendents of the founders had more relatives resident in the parish in 1850. There was only one person with no relatives (up to three generations ascending and two descending) while in the general population some 15% had no relatives. We assume that these were young men and women who had come into the parish to work as farm servants before marriage. We don't have direct data on wealth of the founder sample, but we expect that these more established families would be more wealthy. In a study we made

of distances between fathers and their children, the descendents of the founders were more apt to be living in the same location as their fathers than the others.

In most cases, then, genealogies reflect the more established residents of an area. In our case, it is quite surprising that they mirrored the general population as closely as they did.

Macro view I: Distribution of Kin in Space

The distribution of kin at any given time is a residue of a very particularistic history, the previous migration decisions of the family. This map shows the distribution of one of the surnames we are following – the Wellman's - within the area that had been settled prior to 1860: the New England States, New York and the part of northern Pennsylvania where people originally from New England settled. Of course, not all of the places where people in the family lived are reflected in the current locations of descendents. The places occupied at a particular date are a combination of the long established family centers and new places being pioneered whether these are in the cities or the countryside. It will take several years for the places that were recently settled to become family centers, if indeed they do.

MAP A. Distribution of men in the Wellman Family in 1860.

Each family had a unique signature reflecting its specific migration choices. In all of the families the point of first settlement –– in this case Lynn Massachusetts -- remained visible over 200 years later. The opportunity for natural increase was very great among the first to come to America and the families remained in the vicinity of this first place for several generations. This family cut across Southern New Hampshire which we do not find in families whose founders settled elsewhere. The Connecticut River was a thoroughfare and several members of this family went to Maine, again not the case for other families.

But these can also be seen as maps of kin accessibility. Several people have discussed the notion of an area of "familiarity" (see for example Rosental 2004). We settled on the 14-mile radius as a useful measure because it was the distance where people found 80% of their spouses (see Adams and Kasakoff 1984); it also was a distance that could be traveled on horseback easily in a day (plus return). We include children as well as adults in this measure but the children who did not survive to adulthood (age 20) were dropped. Completely separate individuals were rare: usually there were at least two or three males in each location, representing, a man and his young sons. Only 15.5% of the men had fewer than 4 kin in the 14-mile radius. The largest concentration was 76, but such a large concentration was unusual. The 95% mark was at 52. The average number of kin that men over age 20 had in that radius was 17, the median 12. People with few kin within this distance probably had only nuclear families nearby, as far as can be judged by studying only patrilineal kin, but those with many presumably had at least potential ties with people beyond the nuclear family. We have, of course, little direct information about how such kin ties actually functioned in daily life.

These maps indicate a potential kin universe and they also show how the distribution of kin in space reflects settlement history. Of course we know this, but we rarely see it. It begins to pose interesting questions such as whether people in such widely separated clusters ever interacted, the nature of the kin networks within the clusters – in general the smaller the cluster the fewer generations back it went, so that in the largest clusters one wonders whether the second and third cousins did interact even though they were spatially available. Also one would like to have more maps such as these in order to devise a series of predictions linking migration and the availability of relatives. Most modern migrations – the celebrated circulations of Mexicans to the United States, for example, also will produce kin clusters in widely separated areas, but researchers have emphasized choice in their production: processes such as chain migration, local labor recruitment, founder effects and the like. Here we have such clusters which, I believe, resulted simply from the high fertility of the founders who settled particular areas. There were, indeed, by 1850 and 1860 singles who represented remnants, ie all of the family had since moved away, but most of the singles were explorers. It will be important to trace their descendents to see how many of them moved on. As Rivers emphasized, genealogies make it possible for us to see kinship relations that may not have been visible to the individuals at the time.

The migration system was about to change from rural to rural to rural to urban. This would affect these clusters in several ways. The urban clusters might be more stable, there would be less migration out due to jobs and spouses being easier to find. Also there would be reaggregation of distant kin in the cities. This is already visible in our data for 1850.

Macro view II : Differences among Families

One would expect that by the year 1860 when the families had been in America for 7 to 8 generations, the "family" in the sense of shared surname or descent from a common male ancestor, would not be important. Indeed in most of our analyses we do not use the variable family, referring to the 9 surnames we are following and we treat them as though they were a normal sample, not linked to each other. In this section we consider some findings about differences among the families in 1860 and ask how they might have been produced.

We have made a study of what factors led men over age 20 to move between 1860 and 1870. Surprisingly, when we put family in the regression – the nine surnames – they were a significant factor in moving (along with age – younger people moved; occupation – men out of farming were more apt to move; region – men living in the state of New York, the most recently settled, were more apt to move than men in New England; the number of people within 14 miles – the circle of familiarity – kept people from moving as did having a large number of sons, as opposed to daughters). Men in two of the families were less apt to move than the others, the Faunces and the Peltons, but only for the Peltons was the effect statistically significant. Still the family variable overall was significant. In separate regressions by occupation, the family effect was strongest for the

men out of farming and this time both the Peltons and Faunces were significantly less apt to move then men in other families, though the effect of family overall was not significant for the men out of farming alone.

By contrast, family was not a significant determinant of wealth in 1860. Here again the Peltons were deviant and significantly poorer than the other families, but over all the family effect is not significant.   Indeed Peltons had a higher proportion with no wealth at all than the other families (with the exception of one which was very small) and a high proportion of men with no personal property at all. Also, there were strikingly fewer kin within the 14 mile radius, much fewer than in the other families. Since age, economic characteristics of where they were living, occupation (farmer, non farmer and mixed) and kin within the 14 mile radius and size of settlement were in the regression, "Pelton-ness" cannot be explained by any of these.

But both these results point to some underlying family processes that might be causing differences, differences one could not ordinarily detect without genealogical data linking people as far back as 7 or 8 generations.   Where do these differences start ? We could look at sub families as well. When the families are divided into subfamilies in the third generation, the branches were often too small to be useful in the analysis.   So we have divided them at the second generation. In four of the families, there were no such branches, that is, all of the individuals found on the census of 1860 from which we obtained the wealth variables were descended from the same individuals in the second generation, the Peltons among them. But in the other 5 families there were more than one sub branch represented in our 1860 data. Three families had two, one had three and the remaining family had 5 sub branches in our 1860 dataset. The branches are designated according to our genealogical coding system: A represents the founder and the children are AA, AB, AC. So the branch AA is the branch descended from the first child of the founder in the second generation: AC the third child. Since women are given letters, very often there will be missing branches because we do not follow children of women.

The dependent variable in the regressions reported in Table 3 is total property, the sum of real and personal taken from the census for men ages 20 and older. The sample here is rather small (289 men in all) because we are still in the process of gathering the census data. We take the natural log of the property because otherwise men with a great deal would bias the results. Of course wealth increased with age. The role of age squared is to capture the falling off of wealth at the oldest ages. The three "components" are intended to capture important aspects of the economic landscape. They were the most important components from a principal components analysis of county level data from 1860.

When we first did an analysis like this, column 1, we were interested in whether having a large number of kin nearby was related to wealth. Since such clusters were "successful" and represented a history of decisions to remain in an area, one critic reasoned that people in large clusters would be more wealthy than those in small ones. This regression shows that not to be the case: number of kin nearby was not a predictor of wealth. Neither, however, were these clusters economic backwaters, places people were unable to leave.

As Rosental found in France for a later date, people could be economically successful with or without kin.

But it was surprising to us when we decided to see if families varied in wealth once we controlled for all the other variables in the regression. We see here that they did. Moreover the amount of variance explained increased as we divided the families into smaller units.

At first glance, we do not see any startling differences between the poorer branches of these families and the others. However, the analysis raises again some very interesting questions. The fact that occupation drops out as significant when the smaller branches are in the equation may point to a relationship between occupation and family. We can see certain occupations carried on by parents and children and even over more generations: the Chaffees for example were merchants and dentists and jewelers; the Pelton's were sea captains. If such influences existed they might affect family fortunes beyond the point where actual kin helped each other.

Table 3: Regressions on wealth in 1860: No Family vs Families vs Sub Branches

	No Family	Families	Branches
Intercept	-1.89	-1.20	-1.32
Age	.36****	.34****	.35****
Age Squared	-.00****	- .00 ****	-.00****
Occupation	*	*	
Farmer	.30	.25	.17
Non farmer	-.42	-.51	-.46
Mixed	Ref.	Ref.	Ref.
Density Component	-.03	.02	.02
Land Component	.29*	.30*	.26
Rural Manufacturing Component	.10	.08	.06
Family			**
Bisbee		-.31	-.47
Chaffee		.13	
AA			.17
AB			-.02
Coney		-.95	-1.04
Farwell		.44	.32
Faunce		.37	
AE			-.14
AH			.93
Greely		-.22	
AA			- .40
AB			-3.38**
AD			.23
Pelton		-1.04*	-1.05
Shed		-.30	
AB			1.11
AD			.94
AF			-1.63*
AH			.27
AK			- .96
Wellman		Ref.	-.28
AA			
AB			Ref.
Size of settlement			
Less than 5000	.19	.34	.31
Between 5 and 10,000	.03	.11	-.11
More than 10000	Ref.	Ref.	Ref.
Kin within 14 mile radius	.005	-.004	-.003
R Squared	.26	.29	.35

* = p < .05 > .01; ** = p < .01 > .001; *** = p < .001 > .0001; **** = p < .0001

Micro view: Sibling roles

A further advantage of genealogies is that they provide a roster of available kin from which one can evaluate the interaction that actually occurs. They are particular interesting as an indicator of sibling relations.

Here I present some findings about which siblings were most apt to remain in the "same place" as the father as he aged.

Table 4: Distances of sons 15 or older from Fathers by Birth Order: American North 1850

Children	0 km	1 to 9 km	10 to 19 km	20 to 39 km	40 to 59 km	60 to 79 km	80 plus	Total
First Son	65 %	4 %	7 %	5%	2%	1%	16%	445
Middle Son	62 %	4 %	9 %	3%	1%	2%	19%	615
Last Son	77 %	2%	5 %	4%	1%	0	10%	238
Only Son	74 %	5 %	4 %	3%	2%	2%	12%	121
Total	67 %	3 %	7 %	4%	1%	2 %	16 %	1319

Table 5: Fathers by Age, < 60 and 60+ years old: AN

	Fas	Less	Than 60		Fas	Over	60	
Children	0 km	> 80 km	N	Mean age of son	0 km	> 80 km	N	Mean age of son
First Son	75 %	12%	278	22	50 %	23%	167	40
Middle Son	75%	14 %	231	21	51%	23%	284	35
Last Son	89 %	2%	82	19	71%	14%	156	28
Only Son	77 %	10 %	86	22	67%	17%	35	34
Total	77%	11 %	677		57%	20 %	642	

Note: these are all the father son pairs, not just those found on the census. If we limit to census we will be missing several in the same families.

These statistics show that the youngest son was most apt to remain behind to care for his father and that the differences between siblings increased as the father aged. In earlier

work we were able to show that the youngest son delayed his marriage compared to the other siblings. But here we must be careful to ask how this intersected with the settlement history. Given the family pattern of migration in our materials, where the family migrated to a new area when the older children were in their teens and could be a source of labor to clear land, the older siblings would have been ready to marry earlier perhaps at a time when land was cheaper. This example indicates the way in which family lives intertwine with economic cycles.

Incidentally, studies of household composition also show that more distant relatives were rarely utilized for care of the aged. Of course we are looking at a period of high fertility. As fertility dropped, how did this change ?

Genealogies allow researchers to study how the sibling position in one generation affects the descendents.

Conclusions

I doubt whether the people in these families were aware of any of the findings in this paper and thus I return to Rivers' delight in being able to discover laws that the people themselves might never be conscious they are following. It is not so much to show that humans are subject to forces beyond their control, but rather to indicate the limits to that control in an age when we so often think we can accomplish anything.

These examples illustrate, also, I think, the opportunities genealogies create for studying kinship not only at the micro scale but at a larger scale, one not often available to researchers.The challenge now becomes that of linking macro and micro scales. For example, the study of the consequences of sibling roles over generations. This, of course, cannot be done without considering the economic and social changes that would impact the family. We can ask how families help or hinder taking up those new opportunities. We would all like to know how to equip our children for success in the long term.

A further challenge is the incorporation of women and their descendents into these studies. In a study we are just completing comparing distances of women and men from their kin in a population where they are equally documented, women were residing farther from kin and their kin networks had quite a different shape due to the fact that it was the women who moved at marriage.

Genealogies provide coverage of longer time spans and greater spatial extent and, in kinship, of relatives no longer recognized, thus the edges of the significant family relations, seeing beyond the "official" but still affected by the "official" records. They thus provide information about the set of potential relatives which people might make use of when needed. So as in Rivers' time, they still constitute a template which can be compared with other measures of interaction. The problem for historical demography of kin networks is the lack of such other measures. The ingenuity comes in trying to find

them. But there is some information from the genealogy itself, esp if it is coupled with information about residence, e g from a census.

Genealogies allow us to ask: when does the role of kin end ? They can also be used to see links between kin and demographic outcomes. They provide a macro history not a microhistory due to their reliance on vital events to provide information on migration. Each genealogy is a record of past migration. They teach us that kinship relations are both the result of past migrations as much as they might shape future migrations.

■ ■

References and maps to follow

ENDNOTES

[2] The database contains individual level data on some 36,000 individuals over 10 to 12 generations: all of the individuals born to men in these families up to 1860 along with several people born later so that all of the children of men born before 1840 were included (18,067 individuals) as well as the women who married into these families (about 8,000 individuals.

[3] The database contains individual level data on some 36,000 individuals over 10 to 12 generations: all of the individuals born to men in these families up to 1860 along with several people born later so that all of the children of men born before 1840 were included (18,067 individuals) as well as the women who married into these families (about 8,000 individuals)

[4] 45 % of the entire genealogical sample found on the 1850 census was in the 7th generation and 32 % in the 8th. Among men age 20 or older, 30 % were in the 6th generation and 55% in the 7th.   There were interesting differences among the families: due to timing of arrival in America and birth of children, as well as extinction of some lines, some families were effectively one generation behind others.