Marriage relationships among households in the mid 19th century Tama, Japan

Nobuyuki Hanaki, Satomi Kurosu

Marriage Relationships Among Households in the mid 19th Century Tama, Japan

Nobuyuki Hanaki

Satomi Kurosu

Marriage Relationships Among Households in the mid 19th Century Tama, Japan

Nobuyuki Hanaki
(GREQAM, Université de la Méditerranée, University of Tsukuba)

Satomi Kurosu
(Reitaku University)

Abstract

This paper studies the formation of marriage relationships between two households in 19th century, Tama, Japan. Previous studies on marriage market or partner selection in the Japanese past tended to rely either on information from a single village in case of statistical analysis, or on collection of oral histories. By using the information from a household register that covers 35 villages, and applying the method of social network analysis, this paper goes beyond the limitation of previous studies. Our empirical results show that there was a tendency for socio-economic homogamy and endogamy (within kinship and within village) among peasants in the mid 19th century Tama, Japan.

Keyword: Marriage, Japan, Network, Household, Household registers

1. Introduction

Marriage is an important social institution that unites two individuals but also binds two families or groups. Marriage, therefore, has been considered a social reproduction (Bourdieu 1976) or strategy for alliance (Levi-Strauss 1976). Marriage was crucial in pre-industrial rural Japan. Given the high mortality of the times, marriage offered the most reliable and least expensive way of achieving both the optimal household size and the optimal gender balance of labor for family farming (Smith 1972; Tsuya and Kurosu 2000).

When one marries (timing), whether one marries or not (celibacy), and to whom one marries (partner selection) are three fundamental questions that have been tackled from various perspectives. Accumulated studies on Japanese peasants show that (1) peasants married early and universal---they married at least once in their life as long as they survived; (2) however, their marriage did not last long due to high mortality and divorce rate; and (3) they often experienced remarriage. In all, regional variations were extremely large (Kurosu, Tsuya and Hamano 1999).

This study deals with the third question of partner selection in marriage, or in our terms, marriage network (formation of relationships via marriage). In the context of historical analysis, at least three approaches to examine partner selection are recognized: socioeconomic background, geographic location, and family alliance.

First, finding a spouse from the same socioeconomic status (SES homogamy, versus to heterogamy) could be seen as an important strategy, especially among landholding farmers to maintain the landholding status and secure social reproduction (Bourdieu 1976). In European agrarian contexts, by finding a spouse from a higher SES origin, an individual increased the chances of advancing to a higher SES, while,

on the contrary, marrying someone with a lower SES potentially increased the risk of socioeconomic mobility downwards (Dribe and Lundh 2008). In Japanese agrarian contexts, homogamy is thought to have been important in securing and assuring an individual living standard as well as an equal economic relationship and exchange between two households. Anthropological studies suggest the importance of having similar pedigrees of households for a good match (Segawa 2006) but demographic studies are scarce on this matter.

Second, finding a spouse from the same village (endogamy, versus to exogamy) was probably considered vital in rural organization in preindustrial Japan. Many anthropological studies point to the fact that partners from the same village were preferred and even some sanctions were practiced against marriage between a woman from the village and outsiders (Segawa 1974). Some studies also indicate that there were customary laws to prohibit (Uematsu 1996: 614) or to require special rituals for exogamous marriage (Yonezaki 1987). Some argue that endogamous marriages fit the best for integrating the functions of two households to cooperate with each other in cases of rituals, festivals, and in exchanging agricultural tools and horses in daily basis (Yagi 1996: 577). It is interesting to note, however, that households with SES at both ends---extremely high SES and extremely low SES, tended to seek marriage partners outside of one's village (i.e. exogamous in order to maintain homogamy) as they could not find any good match within one's villages (Segawa 2006: 58-59).

Some studies examined the geographic distance of marriage migration beyond village boundaries (e.g. Segawa 2006; Kawaguchi 1984). These studies demonstrated that the majority of marriage took place within 3 to 4-*ri* (12-16km)---commuting or a day trip distance. Therefore, the marriage market geographically coincided with the

sphere of daily activities and within neighboring communities, if not within the same village.

Third, family alliance or kinship organization is pivotal in the studies of family sociology and rural sociology in Japan. Numerous studies examine and contrast the two kinship organizations---*dozokudan* (a group of households based on a common ancestor and with hierarchical order based on economic relationship) and *shinzokudan* (a group of individuals related by blood or kinship formed by marriage or adoption[1]). However, how they are related to the formation of marriage networks has not been made clear. Oto (1996: 55) mentions that in order to maintain and strengthen *dozokudan* organization in preindustrial Japan, it was necessary to find marriage and adoption partners within the same *dozokudan*, thus resulting that, in reality, *dozokudan* and *shinzokudan* overlapped. We could assume that these rural kinship organizations, whether it to be *dozokudan* or *shinzokudan*, served to encourage endogamous marriage as well as marriage between relatives.

Previous studies on marriage market or partner selection in the Japanese past had at least two limitations. One, since most longitudinal historical data, that would allow us to examine detailed marriage-related exchanges over generations, come from one village, the study of marriage market tended to be confined to the alliance among households within one village. Two, since many anthropological frameworks on marital customs, which provide us with many important variations in the past, are based on collections of oral histories from pre-war Japan, it is difficult to specify time and space where the customs apply to. The use of 1870 household registers in Tama

---

[1] *The adoption here includes both adoption of children and that of adults. The latter often took place together with marriage with daughters in adopting households (called, mukoyoshi). In this paper, we include the latter "mukoyoshi" in marriage.*

and applying the method of social network analysis allow us to go beyond these limitations.

To study which types of households tend to form marriage relationships between themselves, one needs to study not only the relationships that have been created, but also those that did not materialize. To do so, we construct a pool of potential pairs of household and study which pair actually established a tie. The pool of potential matching would allow us to go beyond existing conceptual frameworks and examine what are the driving factors of households to ally.

Our empirical results show that there was a tendency for socio-economic homogamy and endogamy (both in terms of kinship and locality---within kin members and within village) in the mid 19th century Tama, Japan, but not exclusively so. We also find that households are connected, either directly or indirectly, beyond the village boundaries and socio-economic status. In particular, the network of household demonstrates that about a half of the households from more than a half of the villages in our data formed a large connected component.

The rest of the paper is organized as follows. Section 2 describes the data source, a brief summary of concepts used in network analysis is provided in Section 3. Section 4 illustrates our empirical framework as well as definitions of variables that are used in the analysis. The results are summarized and discussed in Section 5. Section 6 concludes.

2. Data

The data used in this study is constructed from the Commoners' Household Register of Hinojuku, Tama, compiled in 1870. It derives from one of the experimental

household registers conducted in several areas in Japan around 1870, prior to the first nation-wide household register in 1872 (known as *Jinshin-koseki*)(Ishii 1981: 291). The register covers 35 villages that belonged to Hinojuku-group located in Tama, Musashi Province, the areas that correspond to what currently are the cities of Hino and Tama in western Tokyo. In total, there are 2120 households registered.

The household register records the information regarding all the members of the 2120 households who were alive in 1870 (there were 13965 such individuals). The recorded information of individuals includes their age, their relationship to the household head, their presence at the household in 1870, their origin in cases they became a member of the household due to, for example, marriage or adoption.

In this paper, as noted in footnote 1 above, a marriage refers to either a woman moving into a new household as a wife (virilocal marriage) or a man moving into a new household as a husband (uxorilocal marriage). It is important to note that these two forms of post nuptial residence were practiced in early modern Japan where the stem family organization[2] was considered ideal and practiced. That is, in early modern Japan, marriage did not mean a formation of a new household, but rather a succession of household either at the natal household of husband (virilocal) or wife (uxorilocal).

As we will discuss in detail in the next section, we are going to exploit this information regarding the origin of individuals in constructing a network among these 2120 households who are linked through marriage relationships. In addition to these detailed information at individual level, the socio-economic status (SES) of

---

[2] The stem family system in pre-industrial Japan emphasized non-partible inheritance and succession by an eldest child. While the eldest child could have been either male or female, in reality, males were preferred; and thus, the majority of marriage (around 80%) was virilocal than uxorilocal marriage.

households such as household landholdings or whether the household heads were acting as village officials are kept.

In rural Tama where life was centered around agricultural activities, landholding was the basis for political and economic control of the governors (Yasuzawa 1972: 18), and village officials took important roles representing and leading the rural organization. Whether head of a household served as village officers, therefore, indicated social, not always economic, status of the household. Household landholdings are the most frequently used indicator of household SES in the studies of peasants in preindustrial Japan. They are measured as the sum of their agricultural yield as expressed in measure of rice (*koku* or *go*), or where other crops were grown, in rice equivalents (Smith 1977: 30). Although the development of market economy eventually undermined the role of landholding in determining household SES, landholding remains the best indicator of household SES during the period of this study.

With its size and detailed coverage of 35 villages, the household register in Tama provides us with a rare opportunity for statistical examination of various issues including marriage, adoption, and leaving home (Kurosu and Ochiai 1995; Kurosu 1996). In the current study, we exploit the information about marriage relationships. However, because of the nature of the household register at the time, we have to be aware of some limitations of the data. First, since we are constructing our variables based on the information of survivors in 1870, we are not free from the sample selection problem. Therefore, we must assume that the marital behaviors of those survivors are similar to those not present in 1870 because of death. We also construct a separate dataset by restricting the age in order to partially control the mortality effect. Second, when there are a series of exits (a person exits once, comes back, and

exits again), only the most recent reason for exit or entry is recorded in the register. Most of these cases concern those who moved for service and that should not affect our analysis of marriage network. However, we also expect quite a number of divorces and remarriages took place as usual among population in early modern Japan (Kurosu, Tsuya and Hamano 1999). Thus, the new marriage relationship we identify is likely to be a subset of all the marriage relationships that existed. We have tried to overcome possible effects this may have on our result by focusing on marriages that took place in the last 10 years in the data. Of course, if we had information on divorce and re-marriage as well, we could have analyzed not only the formation of new relationship but also voluntary terminations of existing ones. But that is beyond the scope of the current paper.

3. Networks of households

Households are linked, directly and indirectly, through marriages and adoption. In this section, we describe the construction of networks of kinships among 2120 households in our data in year 1870 and discuss their properties.

A network consists of sets of nodes and links. For the network we construct, nodes are households and a link between two nodes exists if there has been at least one individual reported as moved from one node (household) to the other due to marriage or adoption.

Although the household register is a very rich data source, a few limitations should be noted. The major limitation of the data is the lack of information for those who have past away before 1870. This is especially so in trying to construct a kinship networks among households, because the links between households, as they are defined based on individual movements, will be lost with a death of those who created

the link. This limitation prevents us from creating a complete genealogical graph, such as p-graph (White and Jorion 1996) that is often used in analyzing the structure of kinship network in anthropology.

Another limitation is that we have no information regarding households other than those recorded in the household register. For example, there are 1507 individual movements due to marriage or adoption reported between 1861 and 1870, among which 640 (42.5%) were within 35 villages and other 867 (57.5%) were with those households outside. Such a limitation, however, cannot be avoided in any kind of the network analysis that is based on the subgroup of the entire population. Many previous researches have therefore focused on the part of the population they have information about.[3] And, at least, the household register provides the full information regarding those individuals from these 35 villages who were alive in 1870.

The network we construct is, therefore, based on the all marriages and adoptions among 2120 households that are reported to have taken place based on those members of the households who were alive in 1870. Before proceeding to discuss the properties of the so created network of households, we introduce a few terminologies from the network analysis.

Two nodes in the network are directly connected if there is a link between the two. Although two nodes are not directly connected, they can be indirectly connected if there exists a chain of links that connects them. We call such a chain of links, *a path*. For example, consider a network with three nodes, 1, 2, and 3. And assume there is a link between 1 and 2, 2 and 3, but not between 1 and 3. In this case, node 1

---

[3] For example, a seminal work of an evolving social network by Kossinets and Watts (2006) that studies the e-mail exchanges among university students, faculties, and staff members confines their analysis to those exchanges within the university because the authors do not have the information about e-mail exchanges with those inside the university and those who are outside.

and 2 are directly connected (or linked, and so are 2 and 3), but not 1 and 3. But since there exists a path between 1 and 3, they are indirectly connected. Two nodes in the network are connected if there is a path between them. When all the nodes in the network are connected with each other, we say that *the network is connected*.

One can calculate *the distance* between any two nodes in the network by counting the number of links in the path connecting the two. In the above example, distance between nodes 1 and 2 is one, and that between 1 and 3 is two. When there are multiple paths between two nodes, the distance is defined as the length of the *shortest* path between the two. If there is no path between two nodes, they are not connected and the distance between them is infinity.

The average distance between two nodes is not well defined when the network is not connected. In such cases, the network can be decomposed into several connected components. *A connected component* of the network is a subset of nodes that is connected with each other. The size of a connected component is the number of nodes that belong to it. When a network is not connected, instead of defining the average distance of the network, one can calculate the average distance between two nodes that belong to the *largest* connected component.

The number of links a node has, i.e., the number of nodes the node is directly connected to, is called *the degree* of the node. Thus, the average degree of the network is the average number of link a node has in the network.

While the size of and the average distance between two nodes in the largest connected component measure the connectedness and closeness of nodes in the network, they do not capture local properties. It has been demonstrated that many networks, both social and physical, demonstrate dense local connections. *Clustering coefficient* (Watts, 1999) measures the likelihood of nodes that are directly connected

to a node are connected with each other. If nodes in a network are connected randomly, the clustering coefficient will be very low (it approximately becomes average degree/number of nodes), while many networks, not only social networks, exhibit much higher clustering coefficients than that of comparable random networks.

A network that has both the average path length close to and the clustering coefficient that is much larger than those of random graphs of the same size and the average degree is called "small world network." Small world property is found in various networks ranging from social networks such as co-authorship networks (Newman, 2004) to more natural networks such as a network of neurons in bacteria (Watts and Strogatz, 1998).

[Table 1 around here]

Table 1 provides these network statistics for marriage network in Tama. Among 2120 households in 35 villages, 1455 households had at least one link with other households within 35 villages based on marriage or adoption among those who were alive in 1870. The size of the largest connected network was 1010, in other words, close to a half of the entire households in these 35 villages are connected, at least indirectly, through chain of marriages or adoptions. These 1010 households are from 19 different villages. On average, those households, which had at least one kinship relation with other households in these 35 villages, had about 2 links. There was a household that had 9 links. Note that this network demonstrates ten times larger clustering coefficient than, while the average path length of similar magnitude to, that of a random graph of the same size and the average degree. Thus the marriage network in Tama has the properties that are closer to a "small-world network" than

11

being a random graph. These descriptive statistics, especially the high clustering coefficient, suggest that there existed a tendency for endogamy in terms of kinship among households in our data. Also the large connected component that includes 1010 households and spans 19 villages suggests that marriages in Tama in the period under study were not constrained too strongly by village boundaries.

4. Formation of new marriage relationships

Our dependent variable is whether there has been a new marriage relationship between two households between year 1861 and 1870. We do not consider creations of kinship relationships through adoption[4] in this part of analysis. The dependent variable is defined for every pair of households, and takes value 1 if there existed a new marriage relationship between the pair of households, and 0, otherwise. Here, a "new" marriage is not confined to the first marriage of an individual. It could be one's second or higher order marriage. Notice also that we are not going to consider the direction of marriage relationships nor its nature. The reason why we ignore the direction of the movement is because it is reasonable to believe that both households have agreed if a marriage relationship is formed between them. Also note that we are interested in whether a marriage relationship existed or not, and therefore, we do not analyze how many relationships being created, if any. As will be discussed in detail in the result section, there were cases where more than one marriage relationships being established between a pair of households within the 10 years that we consider.

---

[4] As noted in footnote 1 above, in this paper, we consider a *mukoyoshi*, i.e., adopting an adult male who will be married with a daughter of adopting household, as a marriage and not as an adoption.

For the analysis below, we restrict our attentions to those pairs of households in which there are at least one non-married individuals of different sexes in the year 1860 (Non-married individuals (currently not married) are those without identifiable spouses in the data, and include those unmarried, divorced, and widowed).

Therefore, we are excluding, from our analysis, cases in which a household that had no non-married individuals in 1860. In particular, we exclude cases such that a household adopted or gave a birth to a child after 1861 and later received a spouse for this child. This restriction reduces our sample from 2,246,140 possible pairs of households (2120 * 2119/2) to 1,582,405. Among these 1,582,405 pairs of households, 282 pairs (0.018%) are identified as having established a new marriage relationship between 1861 and 1870. [5]

It is possible, because of the high mortality rate for people beyond the age of 55 in the period we are studying (see for example Tsuya and Kurosu 2004), that elderly people were not considered as a candidate for a new marriage[6] To control for possible impacts such consideration may have on our analysis, we have created a second dataset that restrict attention to the pairs of households in which there are at least one non-married individuals of different sexes *who are below the age of 55* in the year 1860.

This additional restriction reduces our sample to 1,345,890. Among these 1,345,890 pairs of households, 268 pairs (0.020%) are identified as having established a new marriage relationship between 1861 and 1870. We report below the results from analyses based on both datasets.

---

[5] If we do not restrict our sample, then among 2,246,140 possible pairs of households, 370 pairs (0.016%) are identified as having created a new marriage relationship.
[6] Although remarriage did take place after this age, the probability of remarriage became extremely small.

Existing research suggests that various household characteristics may affect the possibility of two households in establishing a new marriage relationship. For example, hypotheses related to socio-economic homogamy versus heterogamy, or, at village level, endogamy versus exogamy. To capture these effects, we include in the set of independent variables, the levels as well as similarities of socio-economic status of households, such as their landholdings and *yaku-zuke* (being a village official), and the geographical characteristics of the villages where a pair of households belonged such as being in the same village or the distance between the two villages if a pair of households belonged to two different villages.

Since it is possible for a pair of households in which there are many non-married individuals to establish a marriage relationship with a higher probability, we have controlled for number and the average ages for non-married males and females in a pair of households in year 1860 as well.

One technical issue to be noted is that since the unit of observation is a pair of household, we do not include the individual household characteristics as our control variables. Instead we define a variable such as average landholdings of two households or total numbers and average ages of non-married males and females in the pair of households. [7]

It is also possible for the existing links among households to have an effect on forming another new marriage relationship as suggested by the previous studies on kinship organizations in pre-industrial Japan. In order to examine such effects, we

---

[7] This is a rather technical issue but one can construct a dataset that contains observation for the pair of household twice. For example, one can treat observation for pair 1-2 and pair 2-1 to be different and include individual household characteristics in the control variables. But even if we do so, when we try to estimate the likelihood of a pair of households to engage in a new marriage relationship, it is not possible to separately identify the effect of characteristics of each of the two households, instead what we obtain are the average effects of characteristics of two households.

have included independent variables that are created based on network of household in 1860. The network of household in 1860 is created by deleting the links that are identified by individual movements between 1861 and 1870 from the network of household in 1870 we have discussed in section 3 above.

It should be noted that because of the possible death of some individuals between 1861 and 1870, we do not have the complete information of the individuals who were alive in 1860. And as a result, there may be several links that existed but are missing from our 1860 network. Such missing links can have a large effect on the global measures of network such as the shortest distance (Kossinets 2006); we therefore confine ourselves to locally defined measures in defining our independent variables. In particular, we have constructed dummy variables, *d1* and *d2*, that take value one if a pair of households are at distance one and two in the 1860 network, and zero otherwise.

In addition to characteristics of households, we investigate the effect of village characteristics. In particular, as suggested by previous research on geography of marriage market (e.g. Segawa 2006; Kawaguchi 1984), we would like to investigate the effect of distance between two villages.[8] But because the exact distances between two villages are not obtainable, we approximate them as follows. First, from the *Nihon Rekishi Chimei Taikei* (2002), we identify the contemporary districts corresponding to the 35 villages. We assume the center of each village to be in the proximity of the geographical center of the corresponding contemporary districts, and the distance between two centers to be the distance between two villages.

Table 2 summarizes the description of the variables, and summary statistics of them are presented in Table 3.

---

[8] Ideally, we would like to have geographical distance between all the pair of households. But this is simply not possible to obtain.

[Table 2 and Table 3 around here.]

5. Results

Table 4 presents the result of our regression analyses. We employ the logit model, as our dependent variable, to form a new marriage relationship, is either zero or one. Since we include variables that are defined for pairs of villages while the unit of observations is a pair of households, we have adjusted the standard error for the clustering effect.[9]

[Table 4 about here]

There are four regression results presented in Table 4 depending on the sample and the included control variables. Model 1 and Model 2 are based on the sample without below age of 55 restrictions, while Model 3 and Model 4 are based on the sample with the restrictions. Model 1 and Model 3 include two village level variables, the average and the difference of total yield of two villages (totYield and difVilYields) in the set of control variables, while Model 2 and Model 4 do not.

We are interested in testing homogamy versus heterogamy. This can be seen from the estimated coefficients of the similarities of households. In particular, the similarity of socio-economic status measured by the difference in the landholdings (difWeath) and whether pair of the households both act as village officials (bothByaku). The latter is found to be positive and statistically significant in all the models. The difference in the landholding is found to have a negative and significant

---

[9] Because the unit of analysis is pairs of households, inclusion of village level variables means that these village level variables are repeated for many observations. Since such repeated observations reduce the standard errors of the estimates, one needs to correct for them.

16

effect in all the four regressions. Therefore, it was indeed the case that households with similar socio-economic status were engaging in marriage relationships. This result supports the tendency for socio-economic homogamy. In addition, the average landholdings of the pair of households are found to have a positive and significant effect on the formation of new marriage relationships.

The estimated coefficients of the same village dummy are positive and statistically significant in all the specification. This result is consistent with the previous researches that have shown the tendency for endogamy. In addition, the distances between villages where the pair of households reside are found to have significant negative effect. This finding is consistent with previous findings that most marriage relationships were established in close geographical area, mainly within the same village, or among villages within the distance of daily activities. While the difference in the total yields of two villages is not statistically significant, the total yield in the two villages is found to have a negative and significant effect (in Model 1 and Model 3).

Our results present a strong evidence of marriages within kinship in 19th century Tama. Namely, pairs of households that are directly connected via marriage or adoption are more likely to establish a new marriage relationship than those that are not related. The estimated coefficient of $d1$ (i.e., households directly connected via marriage/adoption) is positive and significant in all the four regressions reported in Table 4. In addition, two households that share a common kin, that is, those with $d2=1$, are found to be more likely to establish a new marriage relationship than those that are not related. The estimated result shows that two households that are directly connected are 0.5% more likely, while those with network distance two are 0.1% more likely, to establish a new relationship than those with network distance 3 or

above. The weak tendency for triadic closure, i.e., those pair of households in network distance 2 is more likely to form a new marriage relationship, partly explain the relatively high clustering coefficient found in 1870 marriage network discussed in section 3. Of course, it should be noted that, as shown by the size of the largest connected component being close to a half of all the households in the data, there are marriage relationships between two households that are in the network distance three or further.

The estimated coefficients of *d1* being positive and significant in all the models means that, there are cases in which a pair of household established multiple marriage relationships within the period we study. As an interesting example of such a pair of households, we have found a pair of households that has established four marriage relationships between 1853 and 1870. Let us discuss this case in detail to illustrate marriages within a kinship. These two households are from the villages of Horinouchi (Household A headed by Takesuke) and Bessho (Household B headed by Chuzaemon). The sizes of landholding in both households were larger than the average---household A had around 22.8 *koku* while household B had 9.0 *koku*. While the head of household B was the head of the village, the head of household A was not.

[Figure 1 around here]

The four recorded marriage relationships are as follows: (1) Around 1853 (judging from the age of the first surviving child): "Iku", an older sister of "Takesuke" from household A, married to the head of household B "Chuzaemon". (2) In March 1860, a daughter of Takesuke married to a son of Chuzaemon and Iku. (3) In November 1860, Chuzaemon's sister married to Takesuke (second marriage for

Takesuke). (4) In November 1870, a daughter of Chuzaemon and Iku married to a son of Takesuke (and his previous wife). These marriages are summarized in Figure 1.

The second and the fourth marriage relationships between these two households are what anthropologists call "cross-cousin marriage" where a son and a daughter of a set of siblings get married. It has been noted that more extensive, stable, and permanent trans-generation ties among allies can be developed by a widely instituted system of cross-cousin marriage (see, for example, Levi-Strauss 1969).

5. Summary and Conclusion

This paper studies the formation of marriage relationships between two households in 19th century, Tama, Japan. Previous studies on marriage market or partner selection in the Japanese past had at least two limitations. One, since most longitudinal historical data, that would allow us to examine detailed marriage-related exchanges over generations, come from one village, the study of marriage relationship tended to be confined to the alliance among households within one village. Two, since many anthropological frameworks on marital customs, which provide us with many important variations in the past, are based on collections of oral histories from pre-war Japan, it is difficult to specify time and space where the customs apply to. By using the household registers in Tama 1870 and applying the method of social network analysis, we have tried to go beyond these limitations.

In particular, one needs to study not only the relationships that have been formed but also those that did not materialize in order to study which types of households tend to form marriage relationships between themselves. For this aim, we have constructed a pool of potential pairs of household and studied which pair

actually established a tie. The pool of potential matching would allow us to go beyond existing conceptual frameworks and examine what are the driving factors of households to ally.

Our empirical results show that there was a tendency for socio-economic homogamy and endogamy (both within kinship as well as village) in the mid 19th century Tama, Japan.

Applying the method of social network analysis as well as statistical analysis to the rich historical data provide us with a very rich research possibility. The results of our study offer a new insight to the understanding of marriage and kinship organizations in early modern Japan. It will be of great interest to apply the same methodology to study formation of marriage networks of other regions of Japan or other countries (for example, using the Korean data explored in Kim, 2009) to investigate the effect of natural and economic environments on formation of such networks.

Other potential area of interest includes investigating the nature of marriage relationships from the perspective of risk sharing in the face of negative shocks. There is a large literature in the field of development economics on the risk sharing among households (see, for example, Fafchamp 2008, for a survey). The use of historical data covering over 100 years including years of famines (such as one studied in Tsuya and Kurosu 2000 and 2004) would be of great interest. We will leave this as a future research.

Acknowledgement

Reference:

Bourdieu, P. and Passeron, J. C. (1977). *Reproduction in Education, Society, and Culture*. London: Sage.

Dribe, M. and Lundh, C. (2008) . "Partner selection and intergenerational social mobility: the impact of age, social origin and place of birth on intergenerational social mobility in 19th century Sweden." Paper presented at the IUSSP seminar on Social Mobility and Demographic Behavior: A Long Term Perspective. UCLA, 11-13 December.

Fafchamp, M. (2008) "Risk sharing between households" forthcoming in *The Handbook of Social Economics* edited by Benhabib, J., Bisin, A., and Jackson, M.O., Elsevier Press

Kawaguchi, H. (1984) "Kinsei Hiryokoku Chiiki no Tukonken ni tuite (geographical marriage market in early modern Japan)" *Rekishi-chirigaku* 124: 17-28. (in Japanese)

Kim, K-t. (2009) "Differing patterns of marriage between a city and villages in 18th century Korea: The case of Taegu Area", *The History of the Family* 14 (1): 69-87

Kossinets, G. (2006) "Effect of Mission Data in Social Networks," *Social Networks* 28: 247-268

Kossinets, G. and Watts, D. J. (2006) "Empirical Analysis of an Evolving Social Network," *Science* 311 (5757), 88-90

Kurosu, S., Tsuya, N. O. and Hamano, K. (1999) "Regional Differentials in the Patterns of First Marriage in the Latter Half of Tokugawa Japan." *Keio Economic Studies* 36(1): 13-38.

Kurosu, S. (1996) "Leaving Home in a Stem Family System: Departures of Heirs and Non-Heirs in Pre-Industrial Japan," *The History of the Family*, Vol. 1(3): 329-352.

Kurosu, S. and Ochiai, E. (1995) "Adoption as an Heirship Strategy under Demographic Constraints: A Case from Nineteenth-Century Japan." *Journal of Family History* 20(3), 261-288.

Levi-Strauss, C. (1969) *The Elementary Structures of Kinship*. Boston: Beacon Press.

Newman, M.E.J. (2004) "Coauthorship Networks and Patterns of Scientific Collaboration," *Proceedings of the National Academy of Sciences of the United States of America*, 101, 5200– 5205.

Nihon Rekishi Chimei Taikei (2002) *Tokyo-to no Chimei* (vol. 13). Heibonsha. (in Japanese)

Oto, O, (1996) *Kinsei nomin to ie, mura, kokka* (Peasants in early modern Japan and Ie, village and state) Kobundo. (in Japanese)

Segawa, K. (2006) *Konin Oboegaki* (Note of Marriage). Kodansha. (in Japanese)

_____ (1974) *Wakamo to musume wo meguru minzoku* (Folk custom surrounding young people) Miraisha. (in Japanese)

Smith, T. C. (1977) *Nakahara: Family Farming and Population in a Japanese Village, 1717-1830*. Stanford: Stanford University Press.

Tsuya, N. O. and Kurosu, S. (2004) "Mortality and Household in Two Ou Villages, 1716-1870," pp.253-292 in Bengtsson, T. and Campbell, C. and Lee J., et al. *Life under Pressure: Mortality and Living Standard in Europe and Asia, 1700-1900.* Cambridge, Massachusetts: MIT Press.

_____(2000) "Economic and Household Covariates of First Marriage in Early Modern Japan: Evidence from Two Northeastern Villages, 1716-1870," pp.131-157 in Caplone, C. and Neven, M. (eds), *Family Structure, Demography and Population: A Comparison of Societies in Asia and Europe.* Liege: Laboratoire de Demographie de l'Universite de Liege.

Uematsu, A. (1996) "Tukonken (marriage sphere)" p.614 in *Jiten Kazoku*. Kobundo. (in Japanese)

Yasuzawa, S. (1972) *Kinsei Sonraku Keisei no Kisokozo* (A fundamental structure of early modern villages). Tokyo: Yoshikawa Kobunkan. (in Japanese)

Yagi, T. (1996) "Sonnaikon (marriage within a village)" p.577 in *Jiten Kazoku*. Kobundo. (in Japanese)

Yonezaki, K. (1987) "Kinsei-koki no konin-girei-kanko to murashakai (marriage custom and village in the latter half of Kinsei). *Kanto kinseishi kenkyhu*. No.22. (in Japanese)

Watts, D. J. (1999). "Networks, dynamics, and the small-world phenomenon." *American Journal of Sociology* 105, 493-527.

Watts, D. J. and Strogatz, S. H. (1998). "Collective dynamics of 'small-world' networks." *Nature* 393(6684): 440-442.

White, D. R. and Jorion, P. (1996) "Kinship networks and discrete structure theory: Applications and implications," *Social Networks* 18, 267-314.
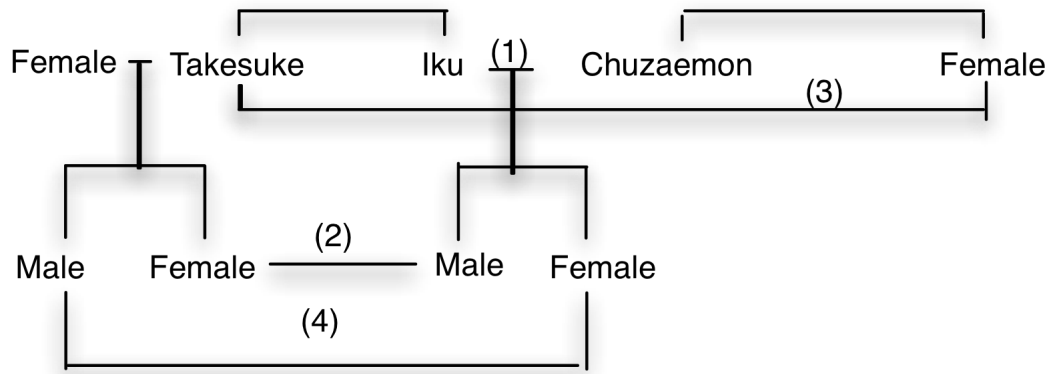
Figure 1: An example of marriage within kinship.

Table 1: Properties of the kinship network in 1870, Tama

Size of network	2120
Number of linked nodes	1455
Size of the largest connected components (LCC)	1010
Number of connected components	140
Average degree among linked nodes	2.09
Max degree	9
Average Path Length in the LCC (for a random network)	12.67 (9.38)
Clustering Coefficient (for a random network)	0.029 (0.0014)

Table 2: Description of the variables

Variable	Description
Lij	=1 if there exists a new marriage relationship between a pair of household during 1861-1870
	=0 otherwise
numMale	Total number of non-married males in the pair of households in 1860
numFemale	Total number of non-married females in the pair of households in 1860
ageMale	Average age of non-married males in the pair of households in 1860
ageFemale	Average age of non-married females in the pair of households in 1860
aveWealth	Average wealth measured by landholding of the pair of households (in koku)
difWealth	Absolute difference in the landholding between the pair of households (in koku)
bothByaku	=1 if both households are village officials (yaku duke)
	=0 otherwise
d1	=1 if the pair of households are directly connected in 1860 network
	=0 otherwise
d2	=1 if the pair of households are indirectly connected with distance two in 1860 network
	=0 otherwise
sameVil	=1 if the pair of households are from the same village
	=0 otherwise
distVil	Approximate distance between two villages (in km)
totYield	Average total yields for a pair of villages (rice + other crops, in koku)
difVilYield	Absolute difference in the total yields between two villages

Table 3: Summary statistics of the variables

Variable	No. of Obs.	Mean	Std. Dev.	Min	Max
Lij	1507494	0.0002	0.013	0	1
	(1277799)	(0.0002)	(0.014)	(0)	(1)
numMale	1507494	3.08	1.523	1	15
	(1277799)	(3.02)	(1.51)	(1)	(15)
numFemale	1507494	2.14	1.187	1	11
	(1277799)	(2.14)	(1.157)	1	11
ageMale	1340733	18.49	10.90	1	75
	(1277799)	(15.52)	(8.59)	1	54
ageFemale	1503894	14.05	12.55	1	78
	(1277799)	(11.03)	(8.44)	1	54
aveLand	1331588	5.10	5.642	0.0	109.7
	(1122414)	(5.36)	(5.799)	(0.0)	(109.7)
difLand	1331588	6.14	9.541	0.0	133.1
	(1122414)	(6.32)	(9.761)	(0.0)	(133.1)
bothByaku	1507494	0.0027	0.052	0	1
	(1277799)	(0.0032)	(0.056)	(0)	(1)
d1	1507494	0.0005	0.023	0	1
	(1277799)	(0.0006)	(0.024)	(0)	(1)
d2	1507494	0.0006	0.024	0	1
	(1277799)	(0.0007)	(0.026)	(0)	(1)
sameVil	1507494	0.0359	0.186	0	1
	(1277799)	(0.0360)	(0.186)	(0)	(1)
distVil	1507494	4.65	2.479	0.0	12.8
	(1277799)	(4.66)	(2.478)	(0.0)	(12.8)
totYield	1507494	292.08	137.16	0.0	868.49
	(1277799)	(290.81)	(137.34)	(0.0)	(868.49)
difVilYield	1507494	196.25	191.99	0.0	868.49
	(1277799)	(196.26)	(192.40)	(0.0)	(868.49)

For the sample restricted to people below 55 in 1860 are in the parentheses

Table 4: Results of Logistic Regression

| | Dependent Variable Lij | | | |
| | All | | Only Below 55 in 1860 | |
Variables	Model 1	Model 2	Model 3	Model 4
numMale	0.113***	0.124***	0.138***	0.149***
	(0.035)	(0.035)	(0.037)	(0.037)
numFemale	−0.051	−0.054	−0.036	−0.031
	(0.058)	(0.057)	(0.062)	(0.062)
ageMale	0.003	0.002	0.015**	0.015**
	(0.006)	(0.006)	(0.007)	(0.007)
ageFemale	0.011**	0.086*	0.017**	0.015**
	(0.005)	(0.005)	(0.007)	(0.007)
aveWealth	0.076***	0.071***	0.075***	0.069***
	(0.010)	(0.010)	(0.011)	(0.011)
difWealth	-0.068***	-0.065***	-0.066***	-0.032***
	(0.013)	(0.014)	(0.013)	(0.014)
bothByaku	1.197**	1.300**	1.016*	1.100*
	(0.565)	(0.560)	(0.607)	(0.608)
d1	3.810***	3.845***	3.866***	3.896***
	(0.271)	(0.265)	(0.269)	(0.265)
	[0.004]	[0.005]	[0.005]	[0.005]
d2	2.586***	2.552***	2.647***	2.610***
	(0.474)	(0.473)	(0.477)	(0.476)
	[0.001]	[0.001]	[0.001]	[0.001]
SameVil	1.287***	1.066***	1.324***	1.109***
	(0.249)	(0.243)	(0.259)	(0.249)
DistVil	−0.290***	−0.308***	−0.278***	−0.295***
	(0.060)	(0.060)	(0.062)	(0.061)
totYield	−0.0017***		−0.0017***	
	(0.0005)		(0.0006)	
difVilYields	0.0002		0.0002	
	(0.0006)		(0.0006)	
Constant	−7.843***	−8.208***	−8.225***	−8.612***
	(0.322)	(0.296)	(0.341)	(0.332)
N	1176522	1176522	1122414	1122414
Pseudo R^2	0.125	0.122	0.124	0.121
Log-likelihood	−1905.66	−1912.12	−1815.24	−1820.89

Standard errors are in the parentheses.
Marginal effects are in the square brackets.
*** Significant at 1% level.
** Significant at 5% level.
* Significant at 10% level.