Structuring the Influence of Social Ties in Migration: Genetic and Environmental Variation

Jeffrey S. Napierala[1]
Department of Sociology
University at Albany and the Center for Social and Demographic Analysis[2]

Timothy Gage
Department of Anthropology and Epidemiology
University at Albany and the Center for Social and Demographic Analysis

Migration is generally viewed as a phenomenon rooted in a complex web of social and economic factors, which initiate and perpetuate flows between various interconnected locations. A long history of research in sociology, demography, and economics can be leveraged to support a variety of theories, which operate on different levels of aggregation and were developed to explain migration in different circumstances. Within this vast literature, however, there has been little, if any, attention given to idea that migration may be related to heritable characteristics.

In population genetics migration is seen as a process which alters the frequency of alleles across geographic space, but is not traditionally seen as an attribute that is subject to natural selection. More recent work, however, has explored the pathways that link heritable characteristics to migration in modern populations. These connections have been found to function through personality characteristics linked to specific alleles. In particular, “novelty seeking,” “drive,” and “sociability” characteristics are associated with migration. The “novelty seeking” characteristic has been directly correlated to the DRD4 dopamine receptor 7R+ allele in predicting migration (Chen et al. 1999; Matthews and Butler 2011). Other alleles, such as TaqA1+ have been associated with personality characteristics that have otherwise been linked to migration (see Campbell and Barone 2012). These findings provide a strong foundation for additional research on the heritability of migration.

The sociological literature can be viewed as being in opposition to a genetic perspective on migration. There is a tradition of examining social network connections, which are viewed solely as social products. A well-known framework posits that the strength of social relationships or “ties” is related to both the density of ties between members in a social network and the independence of the resources available through that tie (Granovetter 1973). Stronger ties—often those amongst kin—are dense, but tend to be less valuable on their own since they have more redundant resources, whereas weaker ties, although they are less reliable, tend to provide access to more unique resources.

This framework intersects with a prominent theory of migration in sociology. This theory, sometimes generically referred to as the “network theory of migration,” views the decision to migrate as a product of resources acquired through social network connections. Resources, such as information on employment opportunities in destination communities, reduce the risks and costs of migration, particularly for non-migrants who do not have direct experience with migration. At the aggregate, community-level, migration is predicted to increase rapidly as resources from past and current migrants spread throughout the community via social networks, in turn facilitating migration by additional members (Massey 1990; Massey and Espana 1987). A high prevalence ratio of migrants in the origin community results in the creation of various social institutions, which in turn encourage additional migration. In particular, a “culture of migration” develops that views international migration as a rite of passage and a means to gain status, especially for young men (Massey et al. 1987). Outmigration eventually increases to a point at which “network saturation” is reached, or all the social networks in the community contain numerous links to migrants and their resources. At this stage, migration by additional members is not necessary to compel additional members to migrate (Massey, Goldring and Durand 1994). In general, the sociological perspective views migration as a process that is perpetuated, and in some cases initiated, by social connections to current and past migrants.

Taken together these social network theories predict migration to follow a structure based on the strength of the social connections to current or previous migrants. Recent research has attempted to differentiate between ties with

---

[1] The corresponding author through email: jnapierala@albany.edu. Please do not cite or redistribute without consent.

[2] The Center for Social and Demographic Analysis of the University at Albany provided technical and administrative support for this research through a grant from the National Institute of Child Health and Human Development (R24-HD044943).

kin (or, in some cases, just household members) and other community members. They have found that both play different, but significant roles in predicting migration, although weaker ties may be more strongly correlated with migration (Garip 2008). Stronger ties provide more reliable sources of information and consequently are more associated with migration across longer distances and to more uncertain environments (Garip 2008; Massey and Aysa-Lastra 2011).

From the above we can generate competing hypotheses about how migration will vary across people with different relationships. First, according to the genetic perspective, the degree to which a person is related to another will be predictive of how likely they will be to migrate (or not). In other words, the more genetically related a person is to another, the more likely their experience with migration will predict the other's experience with migration. However, it should be noted that the degrees to which heritable traits appear to influence migration vary with the strength of social and environmental pressure. In instances where social, economic, or other sources of environmental influence are high—even if migration is strongly related to heritable characteristics—genetic influences will appear to be lower or even nonexistent.

In contrast, the sociological perspective suggests that social relationships will largely dictate one's experience with migration. Since weakly connected individuals may have more influence on migration, we might expect that those more strongly connected would be less influential. In other words, that the more socially connected two people are, the less likely they are to have similar experiences with migration. How this hypothesis relates to the genetic hypothesis is somewhat unclear, since the strength of social ties do not directly translate into genetic relationships. However, if we accept that there is a correlation, the association between genetic relatedness and migration would be negative, in contrast with the genetic hypothesis. A more conservative approach would be to assume equal or random social tie strengths across household members, which still provides a contrast.

Data

Both the person- and household-level files of the Mexican Migration Project (MMP)[3] are used. The MMP database was developed specifically to study migration between the US and Mexico and has been used in countless studies to establish the strong relationship between migration and social networks (e.g. Massey 1990; Massey et al. 1987; Massey, Durand and Malone 2002; Massey and Espana 1987; Massey, Goldring and Durand 1994). Each year, starting in 1982 and continuing through 2010, a few communities in Mexico have been selected and randomly sampled. Communities are chosen to represent a wide variety of urbanized areas that have had some experience with international migration. The MMP was released in October of 2011 and contains information on 144,258 individuals in 134 communities throughout Mexico. It uses an "ethnosurvey" method, which collects qualitative and quantitative information from the household head on the migratory experiences of all household members. This includes information on all children of the head—whether they are present or not—along with other household members[4], thus the sample is not biased by the lack of information on current migrants (who are not in the household during data collection), as other cross-sectional surveys.

There are a number of different genetic relationships identifiable in the empirical data. Using multiple variables, separate relatedness groups have sufficient sample sizes to make them suitable for analysis. These are same-sex twins, opposite-sex twins, full siblings, half siblings and unrelated household members. Twins are not explicitly identified in the survey, but are siblings born to the same parents in the same year. These siblings are separated into two groups, same-sex twins and opposite-sex twins; same-sex twins may be monozygotes and hence are more related to each other, and opposite-sex twins are likely to share additional environmental influences over other siblings with the same degree of genetic relatedness. Half siblings are sons/daughters of the household head born during a different union of the household head than the full siblings (or twins) or are stepsons/stepdaughters of the household head if they only had one union. The last group is unrelated household members, who are often the spouse of a son, but also may be the spouse of the daughter, an adopted son/daughter, or someone otherwise unrelated to the household head.

---

[3] The MMP is directed by Jorge Durand and Douglas S. Massey and is a collaborative research project based at the Princeton University and the University of Guadalajara. More information can be found on their website: http://mmp.opr.princeton.edu/.

[4] See "Ethnosurvey V" table A for an example of the exact information collected by the interviewer at http://mmp.opr.princeton.edu/databases/ethnosurvey-en.aspx

Table 1: Migration by Relatedness			
	Non-migrant	Migrant	Total
Unrelated	*73%* 455	*27%* 165	620
Half-Sib	*52%* 156	*48%* 146	302
Full-Sib	*52%* 21324	*48%* 19600	40,924
Opposite-Sex Twin	*57%* 182	*43%* 135	317
Same-Sex Twin	*61%* 294	*39%* 188	482
Total	*53%* 22,413	*47%* 20,236	N=42,649

Counts of individuals in the 5 groups by whether they have migrated are found in Table 1. Since the focus of the analysis is on the variance in migratory behavior within households and between siblings (including unrelated household members) of different degrees of relatedness, households without migrants are dropped from the analysis.[5] The absence of variance[6] in such households will not contribute to the measurement of additive genetic and environmental variance and may make the detection of such variance more difficult. Dropping these households is similar to the use of probands, or index cases, to aid in collecting pedigree data, since we are just selecting out the relatives of those who have migrated. Also, because of the large size of this group, it will make computations much more intensive. In total 84,260 siblings are available after data reorganization; roughly half, or 42,649 have household level variance in migration and are used in the analysis.

There are a number of measures of migratory behavior in the MMP. To simplify the analysis, one dichotomous variable is used that indicates whether a person has migrated internationally to the US, Canada[7] or domestically within Mexico[8]. Although these events are often separated and analyzed differently, particularly domestic and international migration, all migration events are grouped together and are seen as reflective of the same underlying behavior—migration. Future research should consider how different types and measures of migratory behavior are influenced by genetics.

The Measurement of Additive Genetic and Environmental Variation in Migration

Measuring the genetic and environmental variation of migratory behavior involves a number of complexities. To clarify, additive genetic variation refers to the phenotypical variation of individuals due to (the unobserved presence of) transmissible or heritable effects that are assumed to additive (not multiplicative). Environmental variation refers to non-additive variation, which is due to shared social and environmental influences on phenotypes.

The first concern in measuring the heritability of migration is nonlinear effect of age on migration. As seen in Figure 1, the probability of migration is strongly related to age for both males and females. In part, this effect is related to differential social pressures to migrate exerted on individuals depending on their stage in the life course (i.e. Bernard, Bell and Charles-Edwards 2014). Pressure is often high in the late teens and early twenties when people become more independent from their parents' households. This may bring about migration to geographically unite

---

[5] This takes place after the data has been reorganized as described below, so that remaining household groups have at least 1 migrant. The relatively small number migrants interviewed in the US are also not used since households are not sampled randomly, and an outcome of interest—whether a person has migrated to the US—does not vary.

[6] In models with controls there will be variance generated by difference in the observed probability for any individual, which equals 0 if they have not migrated, and the predicted probability of their migration given the controls.

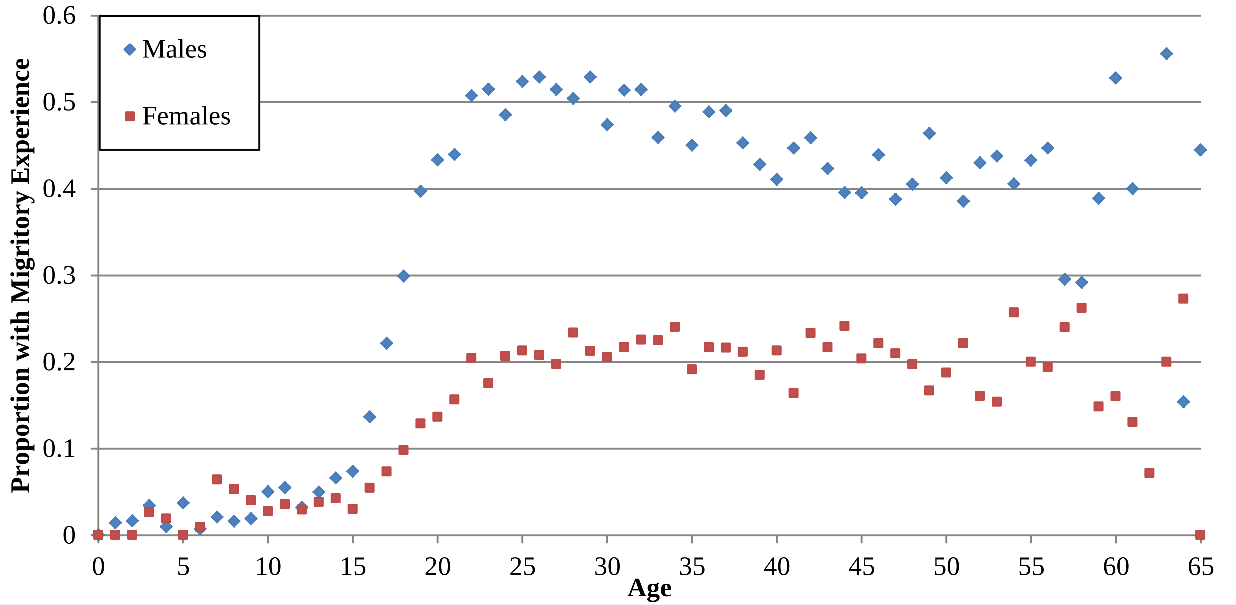
[7] Information on migration to Canada is only available for 25 of the 134 (19%) communities in the sample (data collection on Canadian migration began only recently). When information on Canada is not available, only US migration is used.

[8] Only domestic migrations that cross municipal, state, or provincial boundaries are counted, according to the definition used by the MMP.

with a spouse or to pursue educational opportunities. Particularly for males in more traditional societies, such as Mexico, this stage in the life cycle also coincides with increased participation in the labor force. This may influence them to migrate in order to locate and acquire employment to support themselves and their dependents. Although many of these events are related to social norms and environmental conditions, they are also related age-dependent physiological changes in humans such as sexual maturity and an increased stage of physical and psychological development, which have genetic components.

Age is also related to migration due to the cross-sectional structure of the data. As age increases respondents will have had a longer amount of time to experience influences that encourage them to migrate; the “risk” of migration increases with age. A linear age “fixed effect,” which adjusts the probability of migration according to the observed rate of increase across all ages in the sample, is included in the models below. Only a linear effect is used since the measurement effect is related to the regular, incremental increase of risk due to the passage of time, whereas a non-linear effect is due to differential effects of additive genetic and environmental influences across the life course.

Figure 1: Proportion with Migratory Experience by Age and Sex



Another concern is individuals who have migrated multiple times. Depending on the circumstances of the migration events, they can easily be seen as exhibiting an increased level of migratory behavior. That is to say, these individuals may have an increased genetic predisposition for migration or a stronger source of social or environmental influence. However, given their change in environments during the migration, their subsequent moves may not necessarily be sole products of their original environment since migration will expose them to new environmental influences. The observation that past migration, particularly international migration, is one of the strongest predictors of future migration lends some support to this interpretation (Massey 1986). To control for this source of variation, multiple migratory events for an individual are not differentiated from a single event.

Movement in and out of the surveyed household may also create problems with measuring shared environmental variation. Most children will eventually move out of their parents’ home, and some household members, such as half siblings and unrelated members, may join the household at later points in their lives. Movement in and out of the household will expose members to new, unshared environmental conditions. Since information is not available on when members joined and left the household, the precise measurement of unshared household conditions is impossible, and our results may be subject to increased error.

Members leaving the household may pose less of a problem for measuring shared environmental variation, since the the vast majority of people experience their first migratory event in the mid-to-late teens to mid-20’s. For older members of our sample (age 25+), 37% of eventual migrants had migrated by age 15, 62% by 20, and 82% by age

25. It should also be considered that the formation of new households in close geographic proximity to the parental household (i.e. generally when such events are not counted as migration), are unlikely to fully remove members from social and environmental conditions of the originating household, as members will likely remain in close contact.

People who join the household later in life, however, may have more adverse effects on the results. This is because they may have spent a large part of their lives in a separate environment. There is also the possibility that they were selected into the household based on characteristics related to migration. Care will need to be taken when interpreting results for half sibs and particularly unrelated members.

The next issue concerns the degree to which household members are involved in a "joint" decision to migrate. Young children are a good example since their migratory events are largely determined by their parents or guardians. If, for example, a family relocates permanently, children will be brought along since they are dependent on other members of the household for support and care. In this case, the decision to migrate is not an independent event for dependents and is strongly influenced by the conditions of the household. Nevertheless, since migratory behavior is exhibited, it is counted the same as other migratory events. Individuals under the age of 15 are considered dependents in this analysis and are allowed to share additional environmental sources of variation.

A final concern is differential sources of genetic and environmental influences exerted on males and females. Males will be expected to have increased environmental influences due to the presences of a "culture of migration" and the presence of traditional gender norms in Mexico, which encourage the role of males as the primary income earners. As with dependents, males are also allowed to share additional environmental variation in the empirical model.

Methods

The selection of an appropriate method for computing shared environmental and additive genetic variation is complicated by the structure of the data and the numerous measurement concerns. The model needs to be able to control for a linear effect of age on migration to remove the increasing rate of exposure by age and allow for increased environmental variation for dependents and males. Additionally, given the ambiguities in the data regarding shared environment among household members of different types, it would also be beneficial to examine their shared environmental influences and error components separately.

A simple but flexible mixed models approach developed by Guo and Wang (2002) for complex family behavioral data is adapted for this purpose using the SAS Glimmix Procedure (SAS Institute Inc. 2010). This method satisfies these requirements, and compared to other available techniques (e.g. McArdle and Prescott 2005), it allows for greater utilization of the large MMP dataset by retaining computational efficiency with these additional complications. The main drawback of this method, however, is it does not allow for the direct computation of standard errors for the narrow sense heritability coefficient, h^2, and the shared environmental coefficient, c^2. Given the dearth of literature on the subject, and the limited amount of data suitable for this analysis, tests to establish more precise heritability estimates are left to future research.

The following equations are estimated by using pseudo-maximum likelihood:

$$Y_{ij(t)} = g^{-1}\left[\beta_0 + \beta_2 age_{ij} + u_{0j(t)} + e_{ij(t)}\right] \tag{1}$$

$$Y_{ij(t)} = g^{-1}\left[\beta_0 + \beta_2 age_{ij} + \beta_{3j} sex_{ij} + u_{1j} * males_{ij} + u_{0j(t)} + e_{ij(t)}\right] \tag{2}$$

$$Y_{ij(t)} = g^{-1}\left[\beta_0 + \beta_2 age_{ij} + \beta_3 age15_{ij} + u_{2j} * age15_{ij} + u_{0j(t)} + e_{ij(t)}\right] \tag{3}$$

$$Y_{ij(t)} = g^{-1}\left[\begin{matrix}\beta_0 + \beta_2 age_{ij} + \beta_{3j} males_{ij} + u_{1j} * males_{ij} + \beta_4 age15_{ij} + u_{2j} * age15_{ij} \\ + u_{0j(t)} + e_{ij(t)}\end{matrix}\right] \tag{4}$$

In the above, *i* indexes siblings, *j* indexes families, and *t* indexes the 5 types of genetic relationships. $e_{ij(t)}$ is a residual separated into t parts to capture the unexplained within unit variation in each relationship cluster. The variable "age" is the age, in years, of the individual at the time of survey, and "males" is coded 1 if the individual is male, and 0 if female; the "age15" indicates whether the person is under the age of 15. The u_{1j} and u_{2j} terms capture the additional shared variation due to the "males" and "age15" variables, respectively. A logit link function, g, is used to transform the prediction equation into a linear, unbounded form.

The $u_{0j(t)}$ random effect term captures the shared variation due to each of the t genetic groupings during estimation of the models, and is decomposed as follows:

$$u_{0j(t)} =^{*} u_{j(r)} +^{*} u_{j(h)} +^{*} u_{j(f)} +^{*} u_{j(w)} +^{*} u_{j(s)} \quad [5]$$

The 5 relatedness groups, denoted by *(t)* are: *r* (unrelated person in household), *h* (half brother or sister), *f* (full sibling), *w* (opposite-sex twin), and *s* (same-sex twin). The random effects, "Z" matrix, consists of ones and zeroes and is constructed so that the total sum of shared variance for any given relatedness group is equal to the variance in the particular groups' effect plus all groups of lesser relatedness. In other words, the random effects are constructed so that groups that are more related to each other are nested within all groups that are less related to each other, corresponding with the assumption that more related groups share more variation. For example, the total shared variation for same-sex twins is equal to the sum of all the random effects in (5), whereas the total shared variation for unrelated household members is equal to just $^{*}u_{j(r)}$ term, which is equal to the total shared variation in the household.

Equations (1) through (4) are estimated using the migration indicator as the dependent variable. The first equation (1) is a baseline model, and produces estimates of the random effects after controlling for the linear effect of age. Model (2) includes a "random slope" and fixed effect for males, which captures the variation in the "males" effect within households. Model (3) builds on the baseline model and includes the "age15" variable, and a corresponding random effect. The final model includes both the "age15" and "males" fixed and random effects.

Using these variance terms, intraclass correlations are computed for each of the relatedness groups. The correlation for any relatedness group is equal to the sum of shared variation divided by the sum of the shared variation plus the error, $e_{ij(t)}$, for the relatedness group. The additional random effects in models (2), (3) and (4) capture shared variation in migration due to the environmental influences and are used to compute intraclass correlations specific to sex and dependent status.[9]

The intraclass correlations, $p_{(t)}$ can then be used to compute h^2 and c^2. There are a few approaches that can be used that involve different assumptions concerning shared environmental influences within households (see Guo and Wang 2002). The first method is to assume that the environmental coefficient is constant across all groups with random deviations. The solution can be accomplished through Ordinary Least Squares regression where coefficients of relatedness (and a constant) are regressed on the intraclass correlations. The solution to the regression intercept is equal to the within group correlation when the coefficient of relatedness is equal to 0, or c^2, and the slope (of the coefficient of relatedness) is equal to the within group correlation when the coefficient of relatedness is equal to 1, or h^2.[10]

The second method for computing h^2 and c^2 uses a system of equations—one equation for each relatedness group—which links the h^2 coefficient to the group's coefficient of relatedness, c^2, and $p_{(t)}$:

$$(.693)h^2 + c_{sw}^2 = \rho_{(s)} \quad [6]$$
$$\frac{1}{2}h^2 + c_{sw}^2 = \rho_{(w)} \quad [7]$$
$$\frac{1}{2}h^2 + c_f^2 = \rho_{(f)} \quad [8]$$
$$\frac{1}{4}h^2 + c_h^2 = \rho_{(h)} \quad [9]$$
$$c_r^2 = \rho_{(r)} \quad [10]$$

Equations (6) and (7) have the same, shared environmental correlation term for twins. This restriction is used to solve the system of equations by providing an equal number of equations and variables. The assumption of equal environmental correlation could be made between any two (or more) groups, although equivalence between twin groups is the most justifiable approach.

---

[9] See Guo and Wang (2002) for more details.

[10] Given that h^2 is, by definition, constant across all groups, the residual for each group can be seen as the difference between the average environmental correlation and actual correlation for the particular group.

Given that the exact coefficient of relatedness for same-sex twins cannot be determined due to data limitations, .693 is used as an approximation. This number is reached by comparing the prevalence of opposite-sex twins to same-sex twins in the sample (658 vs. 1070), while assuming that the proportion of dizygotic twins in each grouping is equal (and hence monozygotic twins make up the remainder of the same-sex group). Approximately 38.5% of the twins in the same sex group are monozygotes, and hence the average degree of relatedness among members of the group is: .385*1+(1-.385)*.5 or .693. The estimated proportion of monozygotic twins in this sample is within the range observed by Oleszczuk, Keith and Keith (1999) in various studies conducted around the world, albeit on the lower end.

Results

Solutions for the mixed models depicted in equations 1 through 4 are displayed in Table 2. Although the results are somewhat difficult to interpret in this form, there are a few important observations that can be gleaned concerning both the fixed effects and variance parameters.

First, all the fixed effects are significant and in the expected direction. The effect of "age" and "males" on the log odds of migration is positive, whereas the effect of "age15" is negative and significant. The log odds of migration significantly increase with age, are higher for males, and lower for dependents.

Due to the structure of the random effects, the standard errors for the genetic relatedness groups test whether the variance for a given group is larger than the sum of the variance for all groups that are less related. For the final model, we can conclude that the same-sex twins have more shared variation than opposite-sex twins who in turn do not have more shared variation that full siblings. This is not surprising since they both are related to each other to the same degree and are only separated due to the potential for increased shared environmental influences on twins. Full siblings have significantly more variation than half-sibs and unrelated household members, who have roughly the same amount of shared variation.[11]

Table 2: Solutions for Mixed Models

		Age Fixed Effect Only		Sex Fixed and Random Effects		Age15 Fixed and Random Effects		Age + Sex Effects	
	Variable	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Random Effect Variance	Males			0.481	(.046)			0.697	(.051)
	Age < 15					2.298	(.183)	3.172	(.23)
	Unrelated	0.405	(.092)	0.523	(.107)	0.409	(.09)	0.518	(.104)
	Half Sib	0	-	0	-	0	-	0	-
	Full Sib	0.627	(.095)	0.785	(.11)	0.610	(.093)	0.800	(.107)
	Opposite-Sex Twins	0	-	0.423	(.372)	0	-	0.587	(.375)
	Same-Sex Twins	4.351	(.722)	3.276	(.743)	4.254	(.718)	3.009	(.748)
Error Variance	Unrelated	0.918	(.059)	0.847	(.057)	0.947	(.06)	0.820	(.056)
	Half Sib	0.929	(.082)	0.955	(.088)	0.819	(.075)	0.782	(.078)
	Full Sib	0.868	(.007)	0.803	(.007)	0.821	(.006)	0.745	(.006)
	Opposite-Sex Twins	0.964	(.082)	0.752	(.085)	0.918	(.078)	0.664	(.076)
	Same-Sex Twins	0.412	(.039)	0.401	(.037)	0.404	(.037)	0.391	(.035)
Fixed Effects	Intercept	-1.464	(.036)	-2.209	(.041)	-0.780	(.043)	-1.521	(.048)
	Age	0.051	(.001)	0.058	(.001)	0.032	(.001)	0.037	(.001)
	Males			1.123	(.0233)			1.198	(.024)
	Age < 15					-1.680	(.0643)	-1.88750	(.071)
	N	42645		42645		42645		42645	

---

[11] The SAS mixed procedure sets the lower bound for variance components to 0. Other models (not shown) without the lower bound restriction suggest that these terms are small, negative, and not statistically different from 0.

The random effects for the dichotomous “age15” and “males” variables are large both in absolute and relative terms compared to their standard errors. Thus, we can also conclude that the effects of these variables do vary randomly across households, and environmental variation is higher for males compared to females, and especially for dependents compared to household members aged 15+.

The error variance terms, since they are computed for each relatedness group, measure the amount of non-shared variation in migratory behavior attributable to each group. Genetic theory expects the total amount of error variance to decrease as the degree of relatedness increases, which is evident in the final model when both “age15” and “males” are included in the model.

The intraclass correlations computed from Equation 4 are depicted graphically in Figure 2.[12] A general linear trend in the correlations is evident when plotted against the coefficient of relatedness, suggesting that migratory behavior has additive genetic influences. The unrelated group is somewhat of an exception to the overall trend. This is not so concerning given that they are the least likely to have spent considerable time in the surveyed household. They are often the wife of one of the sons, and it is possible that these wives have migrated (domestically) to live in the surveyed household or that the marriage to a household member is related to their shared migratory experience. Either case could potentially increase their intraclass correlation.

Figure 2: Intraclass Correlations by the Coefficient of Relatedness

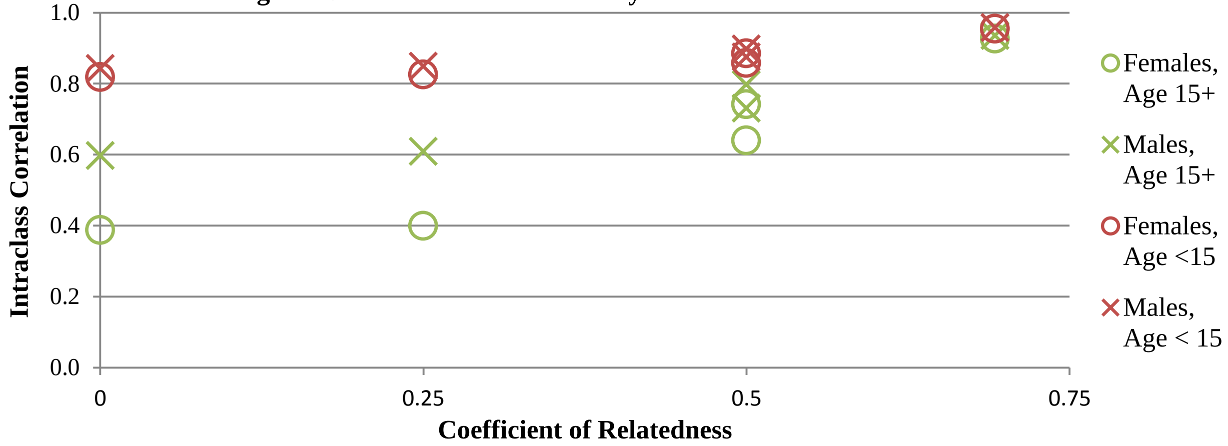


Table 3: Genetic and Environmental Correlations

Computational Assumption	Coefficient	Age & Sex Random Effect			
		Age 15+		Age < 15	
		Female	Male	Female	Male
Method 1: Constant Environment in Entire Household	h²	0.80	0.48	0.18	0.15
	c²	0.31	0.55	0.82	0.82
Method 2: Constant Environment Only Between Twins	h²	0.96	0.72	0.36	0.31
	c² Same-sex Twin	0.26	0.44	0.70	0.74
	c² Opposite-Sex Twin	0.26	0.44	0.70	0.74
	c² Full Sib	0.16	0.37	0.68	0.72
	c² Half Sib	0.16	0.43	0.73	0.77
	c² Unrelated	0.39	0.60	0.82	0.84

Table 3 displays the h^2 and c^2 coefficients calculated according to the two environmental influence assumptions. The assumption of constant variance only between twins produces higher estimates of h^2 than the “constant”

---

[12] The exact values are available in the appendix.

environmental influence assumption. This is due in part to the questionably high intraclass correlations for the unrelated household members; consequently the alternative assumption (of constant environmental influence only between twins) is preferred. The non-significant effect of the Opposite-Sex Twin group supports this limited assumption. Using this method, as expected, females have lower c^2 than males and non-dependents. The highest h^2 coefficient is for non-dependent females (.96) followed by non-dependent males (.72), then non-dependent females (.36) and dependent males (.31).

Conclusions

The goal of this paper is to explore the extent to which migration is related to environmental and additive genetic influences. The results confirm that there is increased environmental variance associated with males, and in particular, dependents. They also suggest that there are strong additive genetic influences on migration in this non-random sample. Individuals that are more related are significantly more likely to be alike in terms of whether or not they have migrated. The exception is unrelated members in the household, who may have increased amounts of shared variation due to their selection into the household. Although this method was not able to provide a precise confidence interval for h^2 or c^2, point estimates indicate that h^2 is highest for non-dependent females, followed by non-dependent males, then dependent females and males.

Theoretically, the results suggest that the heritable characteristics of individuals affect their decision to migrate. Social network theories have been proven to be robust predictors of migration at more aggregate levels, but, as formulated above, appear to lack explanatory power within households compared to the “heritable characteristics” hypothesis. One possibility is that heritable characteristics such as “novelty seeking,” “drive,” and “sociability,” which have been associated with migration, are related to the presence of larger social networks, or perhaps just networks with more “weak” ties that provide valuable information on migration. It also is possible that social networks moderate the relationship between migration and personality characteristics. This idea gains traction if migration is viewed primarily as a response to local social and environmental conditions, such as low resource availability or subordinate social status (see Campbell and Barone 2012). Here, social networks would be seen simply a mechanism that channels individuals in search of a more favorable environment.

Although the above results cannot weigh in on this discussion, additional analyses by the destination of migrants may be useful. International migration is often viewed as a highly social process because of the complexity of long distance moves and the concentration of international migrants in specific communities and geographic regions. Consequently, it would be expected that environmental influences would be very strong, and that heritable characteristics would play a limited role. However, if heritable characteristics affect the ability of individuals to accumulate the resources for international travel or otherwise select individuals into migrating, we would expect the opposite to be true.

Below (Table 4 and Figure 3) are preliminary results, which consider international migration and domestic migration separately using the same techniques discussed above. The two shared environment assumptions used to compute h^2 lead to slightly different interpretations. For international migration, the unrelated household members have much higher intraclass correlations than both half and full siblings. This leads to a lower h^2 point estimate and higher c^2 coefficient than are obtained using the alternative method for international migrants. Using either method, it appears that migration is strongly related to heritable characteristics, certainly more has been previously considered in the sociological literature.

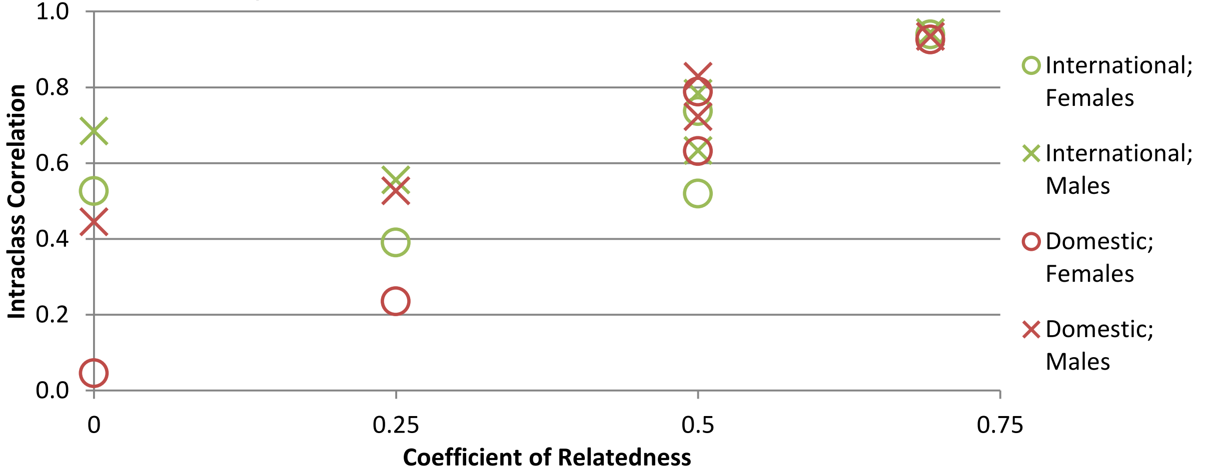
Further research should be done to directly connect migration to specific alleles and personality characteristics. In particular studies that link alleles with both characteristics that are known to affect migration and migration itself would be particularly useful to understanding how migration occurs. Since social networks have been shown to be strongly related to many aspects of migration, further research should consider how characteristics of social networks themselves are related specific alleles.

There are a number of limitations to consider when interpreting these findings. 1. This study examined the most basic measure of migration—whether a person has migrated or not—when there are many aspects of migration that may be more or less heritable. Different measures of migration will provide clarity to the mechanisms through which migration is related to genetic characteristics and evolution. The distance travelled during migration one measure that has been previously explored and found to have significant relationships to DRD4 7+ (Chen et al. 1999; Matthews and Butler 2011). The length of the migration trip and the number of migratory trips is also likely to be related to heritable characteristics associated with general health status and personality characteristics since they will predict the individual's chances for a successful migration. 2. The method used in this study was also unable to calculate exact confidence intervals and 3. Due to limitations of the sample, this study cannot be generalized to the entire Mexican population. We expect that h^2 will be substantially lower in the entire Mexican population then observed in our sample, which only contained households with migrants.

Table 4: Additional Genetic and Environmental Correlations

Computational Assumption	Coefficient	International Migrants, Age 15+		Domestic Migrants, Age 15+	
		Female	Male	Female	Male
Constant Envr. in Household	h^2	0.58	0.36	1.36	0.73
	c^2	0.4	0.58	0	0.41
Constant Envr. Only Between Twins	h^2	1.05	0.83	0.71	0.55
	c^2 Same-sex Twin	0.21	0.37	0.43	0.56
	c^2 Opposite-Sex Twin	0.21	0.37	0.43	0.56
	c^2 Full Sib	-0.01	0.22	0.27	0.45
	c^2 Half Sib	0.13	0.35	0.06	0.39
	c^2 Unrelated	0.53	0.68	0.04	0.45
N		32361		23536	

Figure 3: Intraclass Correlations by the Coefficient of Relatedness



Appendix

Age < 15, by Relatedness

	Age 15+	**Age < 15**	**Total**
Unrelated	*98%*	*2%*	
	608	12	620
Half-Sib	*75%*	*25%*	
	227	75	302
Full-Sib	*88%*	*12%*	
	36113	4811	40,924
Opposite-Sex Twin	*88%*	*12%*	
	279	38	317
Same-sex Twin	*85%*	*15%*	
	412	70	482
Total	*88%*	*12%*	
	37,642	5,007	42,649

Sex, by Relatedness

	Female	**Male**	**Total**
Unrelated	*73%*	*27%*	
	455	165	620
Half-Sib	*52%*	*48%*	
	156	146	302
Full-Sib	*52%*	*48%*	
	21324	19600	40,924
Opposite-Sex Twin	*57%*	*43%*	
	182	135	317
Same-sex Twin	*61%*	*39%*	
	294	188	482
Total	*53%*	*47%*	
	22,413	20,236	42,649

Intraclass Correlations by Relatedness Group

	Age & Sex Random and Fixed Effects Model			
	Females, Age 15+	Males, Age 15+	Females, Age <15	Males, Age < 15
Same-Sex Twin	0.93	0.93	0.95	0.96
Opposite-Sex Twin	0.74	0.80	0.88	0.90
Full Sib	0.64	0.73	0.86	0.87
Half Sib	0.40	0.61	0.83	0.85
Unrelated	0.39	0.60	0.82	0.84

Intraclass Correlations by Relatedness Group					
	Age Fixed Effect	Sex Random Effect		Age Random Effect	
		Females	Males	Age 15+	Age < 15
Same-Sex Twin	0.93	0.93	0.93	0.93	0.95
Opposite-Sex Twin	0.52	0.70	0.75	0.53	0.78
Full Sib	0.54	0.62	0.69	0.55	0.80
Half Sib	0.30	0.35	0.51	0.33	0.77
Unrelated	0.31	0.38	0.54	0.30	0.74

Work Cited

Bernard, Aude, Martin Bell, and Elin Charles-Edwards. 2014. "Life-Course Transitions and the Age Profile of Internal Migration." *Population and Development Review* 40(2):213-39.

Campbell, Benjamin C., and Lindsay Barone. 2012. "Evolutionary Basis of Human Migration." in *Causes and Consequences of Human Migration*, edited by Michael H. Crawford and Benjamin C. Campbell. Cambridge: Cambridge University Press.

Chen, Chuansheng, Michael Burton, Ellen Greenberger, and Julia Dmitrieva. 1999. "Population Migration and the Variation of Dopamine D4 Receptor (DRD4) Allele Frequencies Around the Globe." *Evolution and Human Behavior* 20(5):309-24.

Garip, Filiz. 2008. "Social Capital and Migration: How do Similar Resources Lead to Divergent Outcomes?" *Demography* 45(3):591-617.

Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78(6):1360-80.

Guo, Guang, and Jianmin Wang. 2002. "The mixed or multilevel model for behavior genetic analysis." *Behavior Genetics* 32(1):37-79.

Massey, Douglas S. 1986. "The Settlement Process Among Mexican Migrants to the United States." *American Sociological Review* 51(5):670-84.

—. 1990. "Social structure, household strategies, and the cumulative causation of migration." *Population Index* 56:3-26.

Massey, Douglas S., Rafael Alarcon, Humberto Gonzalez, and Jorge Durand. 1987. *Return to Aztlan: The Social Process of International Migration from Western Mexico*. Berkeley and Los Angeles: University of California Press.

Massey, Douglas S., and María Aysa-Lastra. 2011. "Social Capital and International Migration from Latin America." *International Journal of Population Research*:1-18.

Massey, Douglas S., Jorge Durand, and Nolan J. Malone. 2002. *Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration*. New York: Russell Sage Foundation.

Massey, Douglas S., and Felipe Garcia Espana. 1987. "The social process of international migration." *Science* 237:733-38.

Massey, Douglas S., Luin Goldring, and Jorge Durand. 1994. "Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities." *American Journal of Sociology* 99(6):1492-533.

Matthews, Luke J. , and Paul M. Butler. 2011. "Novelty-seeking DRD4 polymorphisms are associated with human migration distance out-of-Africa after controlling for neutral population gene structure." *American Journal of Physical Anthropology* 145(3):382-89.

McArdle, John J., and Carol A. Prescott. 2005. "Mixed-Effects Variance Components Models for Biometric Family Analyses." *Behavior Genetics* 35(5):631-52.

Oleszczuk, Jaroslaw J, Donald M. Keith, and Louis G. Keith. 1999. "Projections of Population-Based Twinning Rates Through the Year 2100." *The Journal of Reproductive Medicine* 44(11):913-21.

SAS Institute Inc. 2010. *SAS/STAT User's Guide*. Cary, NC: SAS Institute Inc.