Beyond Household Walls: The Spatial Structure of American Kinship Networks

For Presentation at the Annual Meeting of the Population Association of America, 2014, Boston

Jonathan Daw
Ashton Verdery

March 23, 2014

Abstract. Kinship networks are a social network in which pairs are connected through biological, social, and legal pathways. Using the Panel Study of Income Dynamics, this paper investigates the spatial structure of these networks. We find that spousal and parent/child pairs are the most likely to live in the same Census tract, followed by full/half siblings and grandparent/grandchild pairs. Less obviously, kin pairs vary widely in the average distance between them, and this varies by age and educational attainment. Analyzing the determinants of moving behaviors that influence these relationships, we find: distance moved follows a normal distribution with a thick right tail; persons age 40+ are less likely to move; Hispanics have higher odds of moving; women are more likely to move (but shorter distances); higher-educated persons are less likely to move (but move further); and respondents in more recent periods have been more likely to move (but shorter distances).

Introduction

Although social differences in kin contact, co-residence, support, and exchange have attracted much sociological research, studies of kinship structures and their characteristics have received less attention in demographic research. In this extended abstract, we describe kinship patterns of one characteristic that has received scant scholarly attention: distance between kin. Although contact with kin (Raley 1995) and instrumental support (Mazelis and Mykyta 2011; Sarkisian and Gerstel 2004) have been well-studied, distance to the kin with whom one might have contact or provide support has not. In this paper, we use restricted data from the Panel Study of Income Dynamics to provide descriptive statistics on the spatial patterns of kinship by relationship type, age, and educational attainment.

Data & Methods

Data for this study are drawn from the Panel Study of Income Dynamics Family Information Mapping System (PSID FIMS). The PSID began in 1968 by following a nationally representative, household-based sample of over 18,000 respondents (Panel Study of Income Dynamics). It has been collected since 1968, with data available through 2009. The PSID follows the original, nationally representative sample of U.S. households using a genealogical design – as members of the original households left home, the study followed the new households they formed in addition to the original households. Since most households consist of bio-legal kin, this study tracks the evolution of biological, adoptive, and marital lineages over a 41 year span. The FIMS dataset provides linkage variables delineating parent-child (biological and adoptive) and sibling (distinguishing full-, half-, and step-siblings) ties among observations; marital ties are determined using the primary dataset. An important note is that we define co-resident pairs who bear at least one child together as 'partners'. We treat partners in the same way we treat spouses.

In contrast to other work on kinship which has primarily focused on a small set of ties specific to the question being explored (e.g., child and parent ties), our analysis describes a broad range of different kinship ties. We characterize kinship ties using a modification of previous methods (Batagelj and Mrvar 2006; Verdery et al. 2012; White 1963), incorporating information on biological, adoptive, and marital ties to characterize the full kinship networks. The key intuition is that all bio-legal kinship ties can be defined as a function of three elementary matrices: parent matrices (**P**, a non-reciprocal matrix in which person *j* is person *i*'s parent if $\mathbf{P}_{\mathbf{i,j}}$=1 and =0 otherwise), sibling matrices (**S**, a reciprocal matrix in which *j* is *i*'s sibling if $\mathbf{S}_{\mathbf{i,j}}$ =1), and spousal matrices (**E**, a reciprocal matrix in which *j* is *i*'s spouse if $\mathbf{E}_{\mathbf{i,j}}$=1). For instance, one's grandparent is one's parent's parent, and one's aunt is one's parent's sibling or the spouse of one's parent's sibling (algorithms available upon request). Using these methods, we characterize the following kinship ties: parents and children; full and half siblings; current and former spouses/partners; non-relatives connected through chains of co-residence; grandparents; full and half aunts and uncles; grandchildren; parents- and children-in-law; nieces/nephews; great grandparents and grandchildren; full and half great aunts/uncles and great nieces and nephews; and full and half cousins. For relations that are frequently biological, we distinguish between biological and non-biological kin (connected through step- or adoptive ties) as well. We define spousal matrices to include those to whom one has ever been married or co-resided with in a household with the couple's children.

Because the spatial information provided by the PSID is only available at the tract level, we measure each household's geographic location as the spatial centroid of the Census tract in which they reside. Thus this spatial information is left-truncated, as some kin pairs will live in the same tract but not in the same household, and this spatial information is missing from the dataset. Thus, we assess pairwise geographic proximity in two ways: first, we measure whether a given kin pair resides in the same tract; second, we calculate the distance in miles between each kin pair's Census tract centroids.

Education is measured as the most recent valid response (in 2009) to the question, "What is the highest grade or year of school that (he/she) has completed?" which is recoded into four categories: less than high school, high school, some college, and a 4-year degree or higher. All analyses of education are restricted to those 25 years old or older. *Age* is measured as either one's PSID-calculated age in 2009 or, if one is still alive but a non-respondent in 2009, the last valid reported age plus the differences in years since that report.

We also perform regression analyses on two individual behaviors that influence these patterns: moving from one Census tract to another, and the distance between the centroids of those Census tracts. We estimate the former model using logistic regression as a function of age (estimated polynomially), sex, race/ethnicity, educational attainment, and year among persons aged 25 and older (who are likely to have completed their educations). We estimate the latter model using linear regression as a function of these same demographic characteristics. Both models are estimated using the sandwich estimator (Rogers 1993) to allow for the non-independence of multiple observations of individuals across years and the correlated characteristics of individuals within kinship networks.

Results

Table 1 provides frequencies with which different directed kinship ties are measured in the PSID, at the pair level. Table 2 provides three pieces of spatial information on pairs of kin in the PSID: the probability that they live in the same Census tract; the average distance between them if they do not; and the standard deviation of distance between such kin pairs. A few patterns are immediately clear: First, intuitively, spouses, parents, and children are the kin with whom one is most likely to live in the same

Census tract. However, if one does not, one typically lives far away: the average parent-child distance for those not in the same tract is 745 miles, with a standard deviation of 729. This is a typical distance distribution between kin. Also intuitively, one is far less likely to live in the same tract as one's more distant relatives, such as nieces and nephews, cousins, and great-grandparents. Only a small set of ties have probabilities of living in the same tract of greater than 0.1: parents, children, full and half siblings, spouses, grandparents, and grandchildren. The average distance between oneself and one's kin who do not live in one's Census tract ranges between 500 and 900 miles for all kin types, with comparably sized standard deviations.

However, these patterns vary considerably by age and educational attainment. One is much more likely to live with one's parent when one is age 20 or younger (0.727) than when one is 61 or older (0.109). Similarly, as one grows older one's children are less likely to live in the same tract, though the average distance from one's children does not strongly increase between ages 21 and 61+. Furthermore, the odds that one does not live in the same tract as one's siblings decreases sharply after age 20, and the average distance from one's siblings in whose tract one does not live increases concomitantly. Similar patterns are observed for grandparents. Finally, the probability that one lives in the same tract as one's grandchildren decreases with age, and the average distance to them increases.

These results also vary quite strongly by educational attainment, as shown in Table 4. For instance, 52.8% of persons aged 25 and older with less than a high school education live in the same tract as their parents, but only 14.6% of such persons with a four-year degree or higher does so. However, the average distance between more highly educated persons and their parents is lower than for less highly educated persons. Similar patterns are observed for siblings, grandparents, aunts/uncles, nieces/nephews, and cousins. However, there are three exceptions: spouses, children, and grandchildren. The probability of living in the same tract as one's current or former spouse/partner is proportional to one's educational attainment, and the average distance from one's spouse in whose tract one does not live is negatively related to educational attainment. Those with less than a high school diploma are the least likely group to live in the same tract as their children and have the furthest distance to them when they do not, followed by those with high school diplomas, four-year college degrees, and some college attendance. The pattern for grandchildren is somewhat different: those with four-year college degrees are the least likely to live in the same tract as their grandchildren, followed by those with less than a high school diploma, some college, and high school graduates. However, those with less than a high school diploma live the furthest away from their grandchildren on average when they do not live in the same tract.

However, these patterns in spatial relationships by kinship relationship unfold over time as a series of moves as an individual moves residences and their kin do not, or the absence of such moves. As such individual moving behavior will prove an important input into the determination of spatial relationships between kin pairs. Figure 1 provides one such piece of evidence, showing the distribution of the distance moved among those that moved in any given year (expressed as the natural log). As is clearly evident, the distribution very closely follows a normal distribution up to the value of approximately 4 (about 55 miles). Between this value and 7 (1097 miles), there is a significantly thickened curve compared to that expected from a normal distribution, which then tapers down as it approaches 8 (2980 miles). Intuitively, this is consistent with the view that there are two kinds of moves: local, semi-random ones within a 50 mile radius that conforms closely to the normal curve predicted by an aggregation of random inputs, and a significant subsets of long-distance moves, which do not conform to this process (and are likely the result of moves for school, work, or family reasons).

As shown in Figure 2, the probability of moving is structured by age. As can be seen, up to the age of approximately 40 the probability of moving between Census tracts in any given year is quite high, between 0.2 and .25. After the age of 40, however, the probability of moving appears to move strongly downward. (Although relatively few valid observations are observed above the age of 65, leading to increased noisiness in this range, the data is consistent with a continued downward trend in this probability above this age.)

Figure 3 demonstrates the average distance moved by those who do move is less associated with age. Although the age-specific values vary between approximately 100 and 200, these means do not clearly trend in any clear pattern. Thus, although the probability of moving is related to age, the distance moved among movers does not appear to be strongly related to age.

Table 5 presents descriptive regression results describing these patterns for characteristics that include, but are not limited to, age. In the left-hand side results, logistic regression models are fit to estimate the probability of moving in any given year as a function of age, race, sex, educational attainment, and year (measured categorically as 1986-1979, 1980-1990, 1991-2000, and 2001-2009). The right-hand side of this table depicts the effects of these same variables on distance moved among those who move. These effects were estimated only among those who were ages 25 and over in each year's survey.

Consistent with the descriptive statistics portrayed in Figures 3 and 4, in these models there is a statistically significant association between age parameter of moving, but far less association with distance moved among movers. The linear age parameter has an odds ratio of 1.05 (P=0.001), and the age-squared parameter has an estimated effect of 0.999 (P=0.007). The effect of age-cubed is not statistically significant in this model. Although age and age-squared coefficients do not have statistically significant coefficients on distance moved, this is not true for the age-cubed term, which is statistically significantly associated with distance moved (B=0.001, P=0.046).

Among other demographic variables, race/ethnicity is somewhat associated with the probability of moving, as Hispanics have an odds' ratio of 1.125 (compared to whites; P=0.000), showing evidence of an increased probability of moving between tracts. However, Hispanics who do move do not move statistically significantly further than whites, and Blacks and members of other races/ethnicities have no statistically significant associations with either behavioral characteristic.

In contrast, sex and educational attainment are strongly associated with these behaviors. Compared to men, women have an odds' ratio of 1.085 for moving between tracts, but an average distance moved that is 25 miles shorter. Far stronger are the effects of educational attainment: compared to lower-educated persons, higher-educated persons are statistically significantly less likely to move, but those that do move go further. Specifically, high school graduates, some college attendees, and four-year degree holders have odds' ratios for moving of 0.847, 0.844, and 0.772 compared to those with less than high school degrees, respectively. Furthermore, compared to those who did not graduate high school, higher educated persons move 49, 85, and 154 miles further on average for high school graduates, some college attendees, and four-year degree holders respectively. All of these effects are statistically significant.

Finally, there is some evidence of period effects on these outcomes. Compared to those observed in the 1960s and 1970s, persons observed in the 1990s and 2000s have odds of moving that are 1.16 and 1.37 times higher respectively, which are statistically significant effects. However, there is also evidence that those who move go less far than persons in the 1960s and 1970s on average, as those observed in the 1980s, 1990s, and 2000s move 16, 52, and 38 fewer miles respectively compared to those observed in the earlier period. Thus, in recent years participants in this survey have been more likely to move, but typically over shorter distances, compared to persons observed in earlier periods.

Turning back to age patterns of spatial relationships with kin, we can see that these individual behaviors have consequences for the spatial structure of kinship ties. Figure 4 demonstrates this for parent-child relationships. Between the ages of 19 and 62, there is a steady, semi-linear relationship between increasing age and the average distance between oneself and one's parents. Persons who are 19 and do not live in the same tract as their parents live on average in a tract that is 14 miles away, whereas 62 year olds with living parents who do not live in the same tract live an average of 258 miles away. The highest observed mean distance is for persons who are 58 (358 miles), and the shortest is for persons who are 19.

Conclusion

To understand how kinship functions it is important to study its structure and characteristics. We contribute to this goal by analyzing patterns of spatial relations between kin of different types in the U.S. population, and by age and educational attainment.

This research is subject to a number of limitations. First, our measure of spatial distance between kin is subject to left-truncation, such that the distance between persons in the same tract is unmeasured whether or not they reside in the same household. Second, our measurement of kinship networks is limited by the fact that the families of persons who ‘marry in’ to the focal lineages are unmeasured unless they live with a member of the focal lineage at some point during the survey. Third, to date these calculations are descriptive only, and do not differentiate between substantively different reasons for moving that could distinguish between cycling between apartments in a single metropolitan area and moving to another city or state for school or work, or to take care of one’s sick relatives. Greater understanding of the spatial dynamics of kinship networks will be promoted by the considerations of such processes in future research on this subject.

In summary, the PSID spatial data, combined with their information on the elementary ties of American kinship networks, provides significant and heretofore untapped opportunities to study the spatial dynamics and functions of kinship networks, of which this analysis only scratches the surface. We encourage future researchers to take full advantage of this data to better understand how spatial distance and kinship functions dynamically influence one another.

REFERENCES

Batagelj, V.and A. Mrvar. 2006. "Multiplication of Networks." in *Algebraic Combinatorics and Theoretical Computer Science*. Bled, Slovenia.
Mazelis, J.M.and L. Mykyta. 2011. "Relationship Status and Activated Kin Support: The Role of Need and Norms." *Journal of Marriage and Family* 73(2):430-445.
Panel Study of Income Dynamics, p.u.d., edited by S.R.C. Produced and Distributed by the Institute for Social Research, University of Michigan, Ann Arbor, MI 2011.
Raley, R.K. 1995. "Black-White Differences in Kin Contact and Exchange among Never Married Adults." *Journal of Family Issues* 16(1):77-103.
Rogers, W.H. 1993. "sg17: Regression standard errors in clustered samples." *Stata Technical Bulletin* 13:19-23.
Sarkisian, N.and N. Gerstel. 2004. "Kin support among blacks and whites: Race and family organization." *American Sociological Review* 69(6):812-837.
Verdery, A.M., B. Entwisle, R.R. Rindfuss, and K. Faust. 2012. "Social and Spatial Networks: Kinship Distance and Dwelling Unit Proximity in Rural Thailand." *Social Networks* 34(1):112-127.
White, H. 1963. *An Anatomy of Kinship: Mathematical Models for Structures of Cumulated Roles*. Englewood Cliffs, NJ: Prentice Hall.

TABLES

Table 1: Measured Kinship Ties in the PSID

Relationship	N	%
P	54,841	3.27
C	54,841	3.27
PNB	16,797	1.00
CNB	16,795	1.00
SF	40,276	2.40
SH	37,704	2.25
SNB	9,446	0.56
SP	36,326	2.16
NR	622,909	37.11
GP	27,311	1.63
GPNB	21,871	1.30
AU	30,432	1.81
AUH	21,615	1.29
AUNB	62,289	3.71
GC	27,311	1.63
GCNB	20,980	1.25
CIL	24,202	1.44
NN	30,512	1.82
NNNB	62,363	3.72
SIL	65,282	3.89
NNH	21,584	1.29
PIL	23,650	1.41
GGP	6,994	0.42
GGPNB	3,026	0.18
GAU	7,983	0.48
GAUH	6,038	0.36
GAUNB	18,887	1.13
CO	46,494	2.77
CONB	29,543	1.76
COH	26,324	1.57
GGC	6,998	0.42
GGCNB	3,449	0.21
OIL	30,639	1.83
GNN	7,983	0.48
GNNNB	2,627	0.16
GNNH	6,004	0.36
OR	146,338	8.72

NOTE: 'P' stands for parent, 'C' for child, 'S' for sibling, 'SP' for spouse, 'R' for relative, 'G' for great/grand, 'AU' for aunt/uncle, 'NN' for niece/nephew, and 'O' for other. The modifiers are 'F' for full, 'H' for half, 'NB' for non-biological, 'IL' for in-law. Combining these elements together defines each of these relations.

Table 2: Kinship Spatial Patterns, By Relationship Type

Relationship	P(Same Tract)	Mean Distance	SD Distance
P	0.603	745	729
C	0.603	745	729
PNB	0.297	829	706
CNB	0.297	829	706
SF	0.460	667	721
SH	0.448	752	722
SNB	0.167	838	694
SP	0.713	897	703
NR	0.075	749	696
GP	0.120	703	715
GPNB	0.038	793	706
AU	0.059	636	693
AUH	0.069	678	695
AUNB	0.025	733	692
GC	0.120	703	715
GCNB	0.039	793	708
CIL	0.066	771	720
NN	0.060	634	693
NNNB	0.025	732	691
SIL	0.040	727	710
NNH	0.069	678	695
PIL	0.066	773	722
GGP	0.041	710	691
GGPNB	0.016	792	690
GAU	0.048	505	657
GAUH	0.034	565	649
GAUNB	0.024	642	667
CO	0.035	602	669
CONB	0.016	797	677
COH	0.038	636	683
GGC	0.041	710	691
GGCNB	0.015	800	704
OIL	0.020	677	675
GNN	0.047	505	656
GNNNB	0.019	744	719
GNNH	0.034	564	650
OR	0.024	568	652

Table 3: Kinship Spatial Patterns for Selected Kinship Types, by Age Group

Relationship	Age	P(Same Tract)	Mean Distance	SD Distance
P	0-20	0.727	955	727
	21-40	0.335	453	648
	41-60	0.137	623	677
	61+	0.109	790	658
C	0-20	0.590	995	699
	21-40	0.931	401	658
	41-60	0.594	402	604
	61+	0.219	476	677
SF	0-20	0.622	969	731
	21-40	0.231	364	580
	41-60	0.080	446	622
	61+	0.098	522	609
SH	0-20	0.538	914	717
	21-40	0.228	432	617
	41-60	0.085	514	663
	61+	0.160	498	665
SP	0-20	0.660	988	685
	21-40	0.839	737	735
	41-60	0.733	747	704
	61+	0.689	807	682
GP	0-20	0.140	721	729
	21-40	0.062	644	660
	41-60	0.023	812	766
AU	0-20	0.063	712	711
	21-40	0.049	464	618
	41-60	0.093	411	628
GC	21-40	0.531	21	38
	41-60	0.257	266	513
	61+	0.142	410	614
NN	0-20	0.034	964	700
	21-40	0.120	297	524
	41-60	0.062	426	591
	61+	0.045	508	637
CO	0-20	0.035	694	696
	21-40	0.036	440	584
	41-60	0.051	310	527

NOTE: Some age rows are omitted because insufficient numbers of ties were observed.

Table 4: Kinship Spatial Patterns for Selected Kinship Types, by Educational Attainment

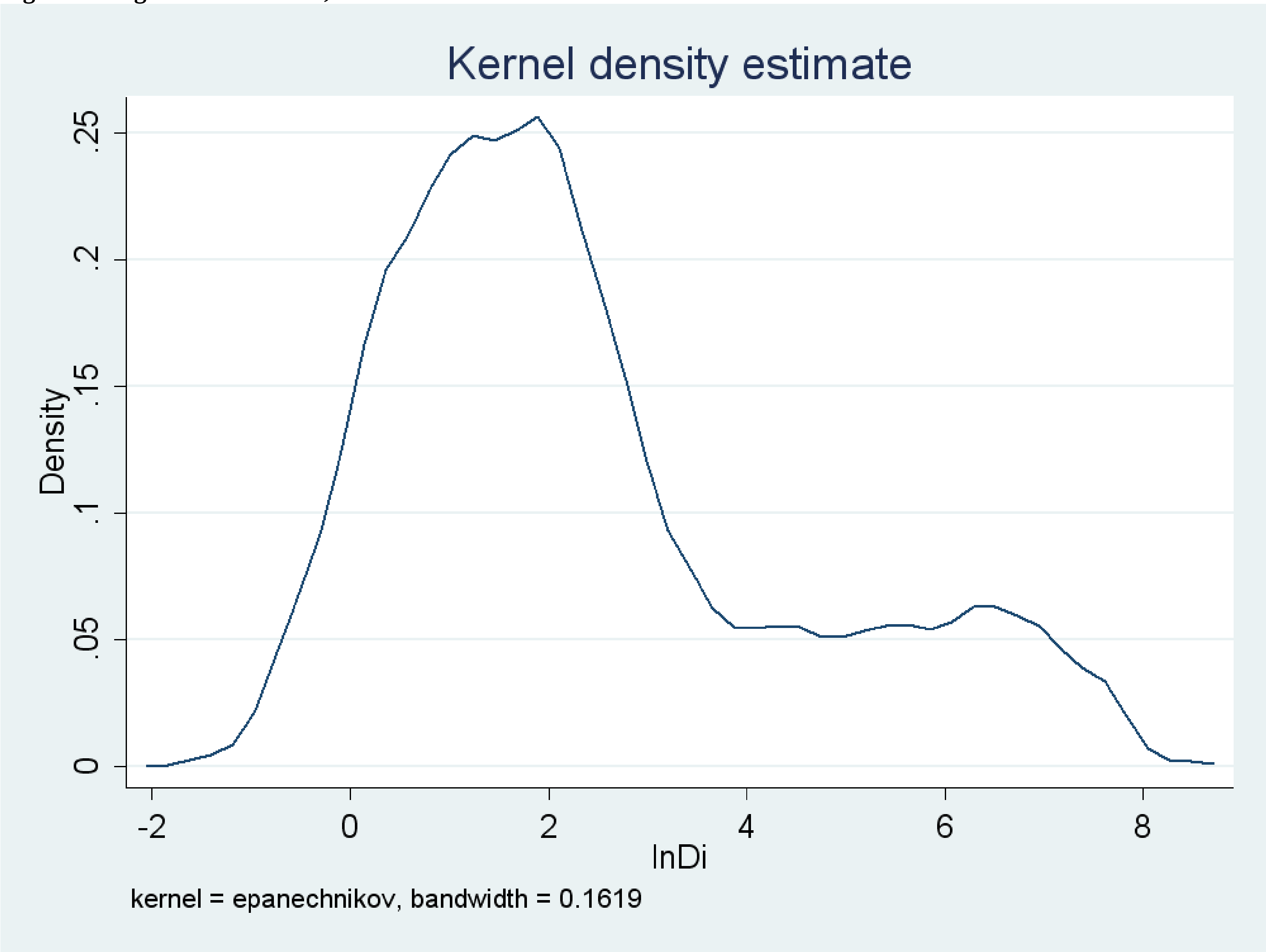
Relationship	Education	P(Same Tract)	Mean Distance	SD Distance
P	<HS	0.528	803	731
	HS	0.304	724	716
	SC	0.279	703	738
	BA+	0.146	639	700
C	<HS	0.362	803	729
	HS	0.541	698	714
	SC	0.612	703	740
	BA+	0.599	665	698
SF	<HS	0.352	713	744
	HS	0.165	626	696
	SC	0.161	615	700
	BA+	0.070	607	671
SH	<HS	0.320	761	722
	HS	0.184	681	696
	SC	0.130	756	742
	BA+	0.069	760	740
SP	<HS	0.544	938	715
	HS	0.646	883	696
	SC	0.681	896	728
	BA+	0.766	834	658
GP	<HS	0.098	692	699
	HS	0.069	733	700
	SC	0.059	693	686
	BA+	0.031	726	663
AU	<HS	0.053	637	672
	HS	0.048	555	663
	SC	0.039	561	651
	BA+	0.028	579	646
GC	<HS	0.088	790	734
	HS	0.126	632	688
	SC	0.111	653	714
	BA+	0.067	629	679
NN	<HS	0.071	670	722
	HS	0.052	604	673
	SC	0.056	619	687
	BA+	0.029	632	675
CO	<HS	0.042	614	662
	HS	0.036	509	614
	SC	0.029	529	605
	BA+	0.019	575	653

Table 5: Regression Predictors of Moving and Distance Moved

	Move		Distance	
	OR	P	B	P
Age	1.049	0.001	7.243	0.134
Age2	0.999	0.007	-0.181	0.084
Age3	1.000	0.420	0.001	0.046
Race				
White (Ref)	--	--	--	--
Black	1.020	0.101	-3.803	0.389
Hispanic	1.125	0.000	-5.810	0.403
Other	1.016	0.669	-19.90	0.057
Sex				
Male (Ref)	--	--	--	--
Female	1.085	0.000	-25.03	0.000
Education				
<HS (Ref)	--	--	--	--
HS	0.847	0.000	49.28	0.000
SC	0.844	0.000	84.61	0.000
>=BA	0.772	0.000	153.6	0.000
Year				
1968-1979 (Ref)	--	--	--	--
1980-1990	0.983	0.246	-15.67	0.004
1991-2000	1.161	0.000	-41.73	0.000
2001-2009	1.373	0.000	-38.37	0.000

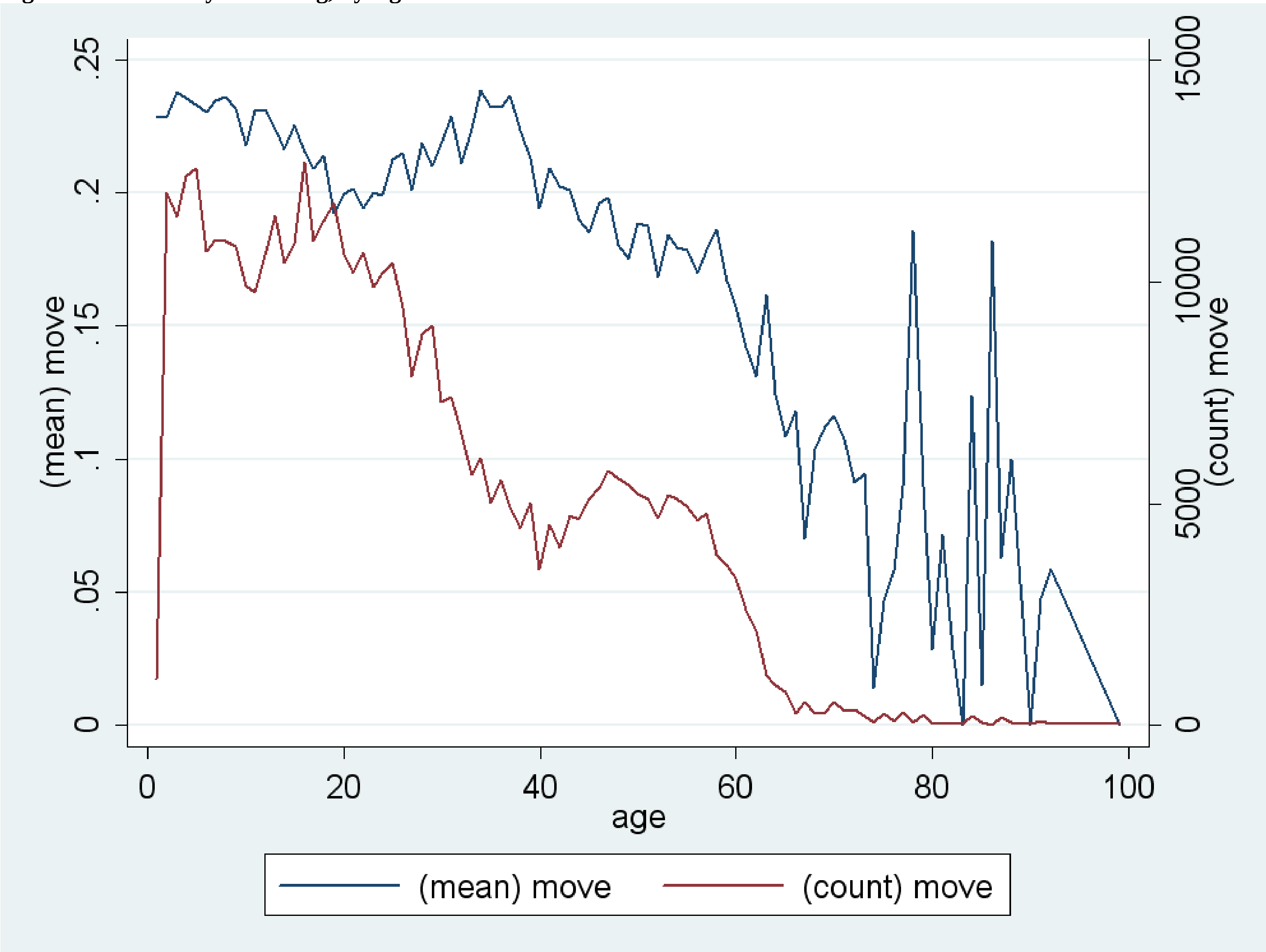
FIGURES

Figure 1: Log Distance Moved, All Years



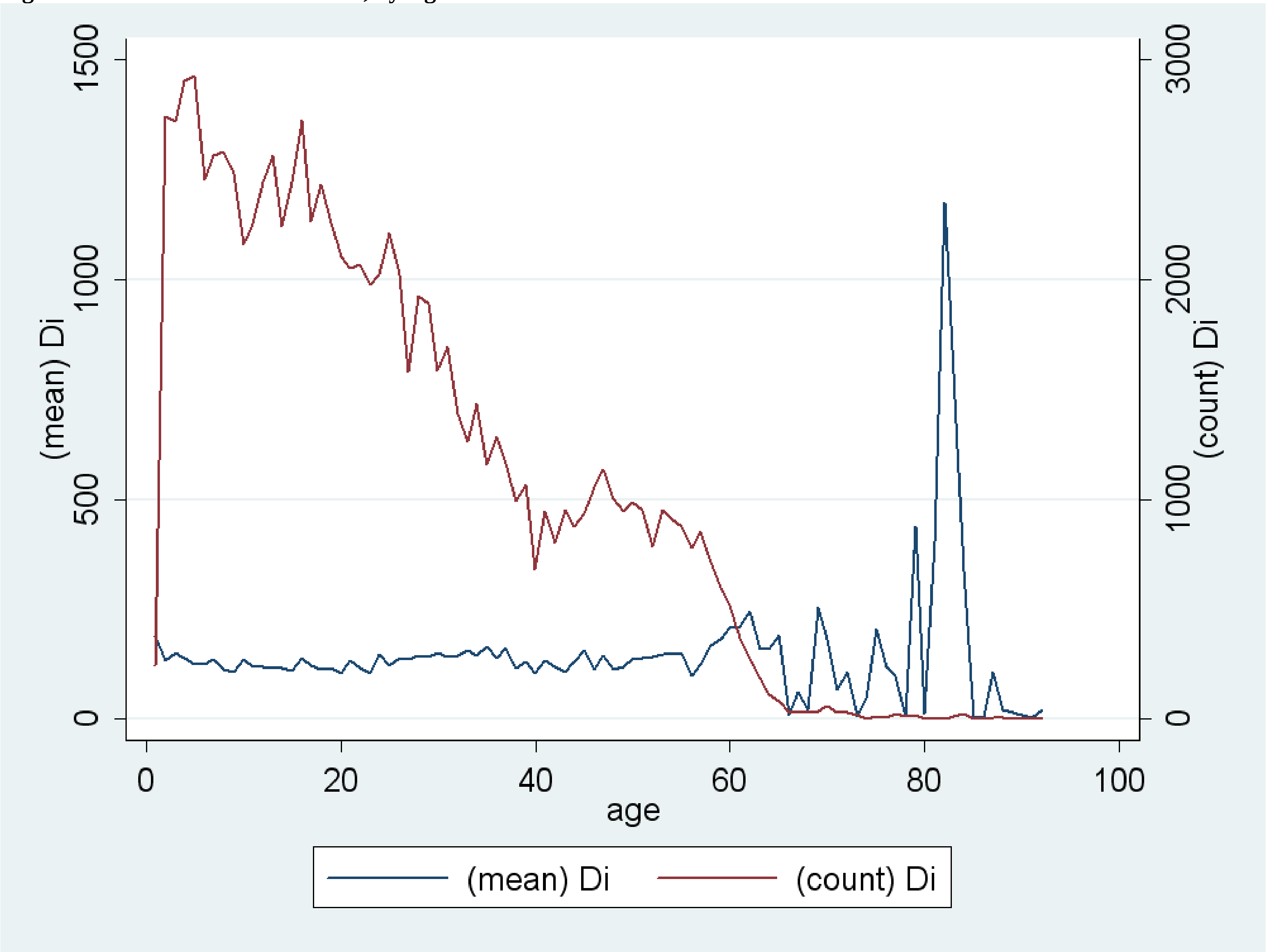
NOTE: lnDI indicates the natural log of the distance between the centroids of Census tracts in consecutive PSID waves among those who moved.

Figure 2: Probability of Moving, by Age



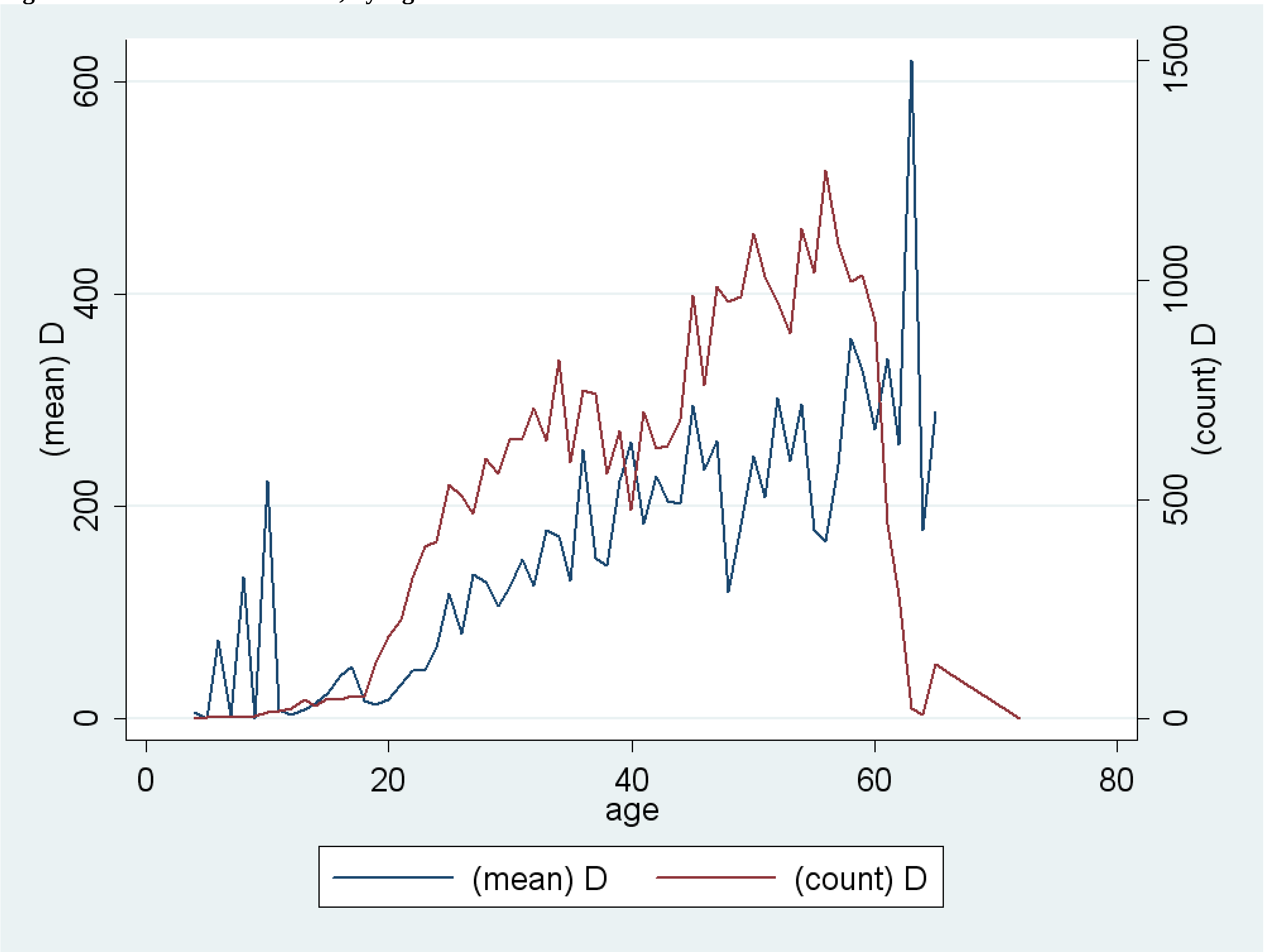
NOTE: "(mean) move" is the probability of moving; "(count) move" is the number of age-specific valid observations in this calculation.

Figure 3: Distance Moved if Moved, by Age



NOTE: "(mean) Di" is the age-specific mean distance moved; "(count) move" is the number of age-specific valid observations in this calculation.

Figure 4: Distance from Parents, by Age



NOTE: “(mean) D” is the age-specific mean distance from parents; “(count) D” is the number of age-specific valid observations in this calculation.