S E D A P

A PROGRAM FOR RESEARCH ON

SOCIAL AND ECONOMIC DIMENSIONS OF AN AGING POPULATION

The Correlation Between Husband's and Wife's Education: Canada, 1971-1996

Lonnie Magee, John Burbidge, Les Robb

SEDAP Research Paper No. 24

For further information about SEDAP and other papers in this series, see our web site:
http://socserv2.mcmaster.ca/sedap

Requests for further information may be addressed to:
Secretary, SEDAP Research Program
Kenneth Taylor Hall, Room 426
McMaster University
Hamilton, Ontario, Canada
L8S 4M4
FAX: 905 521 8232
e-mail: qsep@mcmaster.ca

The Correlation Between Husband's and Wife's Education: Canada, 1971-1996

Lonnie Magee, John Burbidge, Les Robb

SEDAP Research Paper No. 24

July 2000

The Program for Research on Social and Economic Dimensions of an Aging Population (SEDAP) is an interdisciplinary research program centred at McMaster University with participants at the University of British Columbia, Queen's University, Univérsité de Montréal, and the University of Toronto. It has support from the Social Sciences and Humanities Research Council of Canada under the Major Collaborative Research Initiatives Program, and further support from Statistics Canada, the Canadian Institute for Health Information, and participating universities. The SEDAP Research Paper series provides a vehicle for distributing the results of studies undertaken by those associated with the program. Authors take full responsibility for all expressions of opinion.

Note: This paper is cross listed as No. 353 in the McMaster University QSEP Report Series.

The Correlation Between Husband's and Wife's Education: Canada, 1971-1996

Lonnie Magee, John Burbidge, Les Robb
Department of Economics
McMaster University
Hamilton, Ontario, CANADA
L8S 4M4

June, 2000

Abstract

We present a measure of the correlation between the education levels of spouses based on a bivariate ordered probit model. The change in this correlation over time can be measured while controlling for the large changes in the educational attainment levels. The model is estimated with data from 20 Surveys of Consumer Finances in Canada over 1971-1996. Our main findings are a reduction in this correlation among younger couples beginning in the 1980s, and an inverted U-shaped effect of the spouses' age difference on the correlation, with the maximum correlation occurring approximately when the spouses' ages are equal.

JEL classification #: J1, C1, D1.

We thank seminar participants at the University of Guelph for valuable comments. Financial support was provided by the Social Sciences and Humanities Research Council of Canada.

1. Introduction

A person's educational attainment has a large influence on his or her income. It follows that the correlation between the education levels of spouses should have implications for household income inequality. This correlation also is of intrinsic interest. Mare (1991) writes: "... who marries whom [is] a fundamental building block in understanding social structures and social life."

Pencavel (1998) and Mancuso and Pencavel (1999) have used U.S. data to look at trends in "schooling homogamy" among married couples, which is a general term for the tendency for spouses to have the same education level. This also is the emphasis in Mare (1991), who examines the strength of barriers to educational intermarriage, or heterogamy, in the U.S.

We look at the correlation between the education levels of husbands and wives rather than at schooling homogamy. If there is a positive relation between education and income for both spouses, a higher education correlation between spouses should translate into higher household income inequality, all else equal. If the education distribution is similar for husbands and wives, correlation and homogamy have similar interpretations.

To illustrate briefly, suppose there are three education categories, A, B and C. Let (A,B) refer to a couple where the husband has education level A and the wife has B. If all couples were one of (A,A), (B,B) or (C,C), then there would be both perfect schooling homogamy and a correlation of one. The education distributions of husbands and wives would be equal. However, if instead all couples were one of (A,B) or (B,C), there would still be a perfect correlation, but not perfect schooling homogamy, and the education distributions of husbands and wives would differ.

We observe four education categories for persons in married couples in twenty Canadian Survey of Consumer Finances surveys

covering the period 1971-1996. We use the correlation between the two latent standard normal variables in a bivariate probit model estimated by maximum likelihood to measure the correlation between the husband's and wife's ordered categorical education variables. This measure of association dates back to Pearson (1900). Goodman (1984, p.112) provides further references.

We model the ordered probit cutoffs (or thresholds) as functions of time and age. In this way we account for the large changes in education attainments over this period without changing the underlying correlation concept. The correlation itself is specified as a function of the year and the couple's ages. We can then examine the time trend in the correlation while controlling for changes in educational attainment and age distribution.

Other measures of association for ordered categorical variables have been proposed. Goodman and Kruskal (1954,1979) suggest measures that do not rely on a fully specified statistical model. In other work we plan to compare the trends resulting from the Goodman-Kruskal measures to the one used in this paper, for both Canadian and U.S. data.

The statistical model underlying our approach allows us to control for age and year effects within the maximum likelihood framework. A minimum chi-square procedure is used. Coefficients are estimated for each year in the first stage, and then are combined by a GLS regression in the second stage to give smooth time trends.

Section 2 summarizes a few papers on schooling homogamy. Sections 3 and 4 describe the statistical methods used in this paper. Section 5 discusses the data. The results are given in Section 6, followed by the conclusion. Some details of the ML algorithm are given in an appendix.

2. Literature Review

From the vast literature on marriage and assortative mating,

we limit this section to a few recent papers that deal specifically with schooling homogamy.

Mare (1991) uses a “crossings model,” which models the probability that the spouses’ education levels differ at various education levels. Using U.S. Census and Current Population Survey (CPS) data, he finds that “the association between spouses’ schooling increased between the 1930s and the 1970s and was stable or decreased in the 1980s.” He explains this trend as a function of the time gap between schooling and marriage. This gap decreased during 1930-1960 because of an increase in educational attainment and a decrease in age of first marriage, and it increased in the 1970s and 1980s because of an increase in age at marriage. Mare also notes that more highly educated persons are more likely to have the same schooling level.

Qian (1998) also uses U.S. Census and CPS data to examine crossing probabilities, the probability of marriage, and differences between married and cohabiting couples, by age and education. He does not use a summary measure of education correlation or homogamy, but gives many detailed results. For example, since 1980, the reduction in the propensity to marry has been more pronounced among less-educated men and women. Similar to Mare, he finds those who marry at a later age have higher education homogamy. He finds that the relation between age at marriage and the degree of education homogamy is weak for men. Women who marry at older ages tend to have education levels similar to their husbands, while women who marry at a young age tend to be less educated than their husbands, controlling for husbands’ age.

Pencavel (1998) examines cross-tabulations of similar U.S. data. Unlike Mare, he finds an increase in schooling homogamy between 1960 and 1990. He attributes this difference to the availability of data from the 1990 Census. He offers the increase in labour force participation of wives as an explanation for this

increase. This could lead to an increased emphasis by a husband on the wife's earnings potential, and hence her education, in an assortative mating framework.

Mancuso and Pencavel (1999) use similar data and a simulation-based empirical approach based on assortative mating. This method allows them to control for changes in schooling attainment over time. Like Pencavel, they find an increase in schooling homogamy since 1950.

3. Bivariate Ordered Probit Model

Consider the observed pairs (H_i,W_i), where H_i and W_i are ordered categorical variables from the set {1,2,3,4} indicating the education levels of the i^{th} husband and wife. Let z_{Hi} and z_{Wi} be latent variables distributed as bivariate normal with N[0,1] marginal densities. Leaving covariates aside for now, denote the wife's cutoffs as $C_{W,k}$, where $C_{W,k-1} < C_{W,k}$, $k = 1,2,3,4$, and let $C_{W,k} = -\infty$ for $k = 0$ and $C_{W,k} = \infty$ for $k = 4$. Then $W_i = k$ if and only if $C_{W,k-1} < z_{Wi} \leq C_{W,k}$. The husband's indicator H_i is determined in the same way with H replacing W.

The association measure ñ is the correlation between z_{wi} and z_{Hi}. The contribution of observation i to the likelihood function is $\text{Prob}(H_i,W_i) = \text{Prob}(C_{H,Hi-1} < z_H \leq C_{H,Hi}$ and $C_{W,Wi-1} < z_W \leq C_{W,Wi})$. This probability is the integral of the bivariate normal density

$$ö_2(z_H,z_W|ñ) = (2ð)^{-1}(1-ñ^2)^{-1/2}\exp\{-(z_H^2 + z_W^2 - 2ñz_Hz_W)/(2(1-ñ^2))\} \quad (1)$$

integrated over the rectangular area in z_H-z_W space bounded by the four cutoff values $C_{H,Hi-1}$, $C_{H,Hi}$, $C_{W,Wi-1}$ and $C_{W,Wi}$.

The dependence of educational attainment on time and age is captured by parameterizing $C_{W,k}$ and $C_{H,k}$ for couple i as $x_{W,i}^T â_{W,k}$ and $x_{H,i}^T â_{H,k}$ respectively, where the x vectors contain the age of the person in question and its square, etc., and year variables. The dependence of the correlation measure ñ on the spouses' ages and

the year is modeled as $\tilde{n}_i = w_i^T ç$, where w_i is a vector of year and age variables.

The parameters were estimated by maximum likelihood using a GAUSS program available upon request. A hill-climbing algorithm was employed using analytic first derivatives. More details are in the appendix.

4. Minimum Chi Square Estimation

Our sample consists of 428,967 couples observed over the 20 surveys. Sample sizes for individual years range from 15,816 in 1971 to 26,210 in 1990. The estimation algorithm requires repeated bivariate integration and the model involves many parameters, so estimating with the full data set would be difficult and slow. Instead, we split the estimation procedure into two stages. In the first stage, the model was estimated for each survey year separately. The cutoffs and correlation were specified as year-specific polynomial functions of age. In the second stage, the first-stage coefficient estimates were used as dependent variables and their covariance matrix estimates were used in a GLS regression that models the coefficients as functions of time. This step was relatively fast, so most of the model specification searching was done in the second stage.

This is a minimum chi square procedure. In other applications with time series of cross sections, e.g. Burbidge et al. (1997), the first stage involves computing means or medians of the dependent variable, and their estimated variances, for each year/age cell, and no parametric restrictions are imposed until the second stage. We did not use this approach here because there would have been very many cells, many of them empty or nearly so. Some sort of grouping or smoothing was necessary in the first stage. Discrete grouping would create problems with the age/year/cohort aspect.

More specifically, in the first stage the cutoffs are modeled

as

$$C_{W,k}(a_{Wi},t) = \hat{a}_{0Wkt} + \hat{a}_{1Wkt}a_{Wi} + \hat{a}_{2Wkt}a_{Wi}^2 + \hat{a}_{3Wkt}a_{Wi}^3 + \hat{a}_{4Wkt}a_{Wi}^4, \qquad (2)$$

where a_{Wi} is the age of wife i observed in survey year t. Note that wife i in year t is not the same person as wife i in year s, $s \neq t$. Coefficients are estimated for each cutoff (k=1,2,3), each time period (20 values of t), and for the husband as well (with H replacing W).

The first stage model for the correlation is

$$\tilde{n}(a_{Wi},a_{Hi},t) = \tilde{a}_{0t} + \tilde{a}_{1t}\bar{a}_i + \tilde{a}_{2t}\bar{a}_i^2 + \tilde{a}_{3t}\bar{a}_i^3 + \tilde{a}_{4t}d_i + \tilde{a}_{5t}d_i^2 + \tilde{a}_{6t}d_i^3 \qquad (3)$$

where $\bar{a}_i = (a_{Hi} + a_{Wi})/2$ and $d_i = a_{Hi} - a_{Wi}$.

The first stage produces 20 vectors of coefficient estimates, one for each year, of $\{\hat{a}_{0Wkt},\hat{a}_{1Wkt},\hat{a}_{2Wkt},\hat{a}_{3Wkt},\hat{a}_{4Wkt}\}$ for each of the six cutoffs in (2) and 20 vectors $\{\tilde{a}_{0t},\tilde{a}_{1t},\tilde{a}_{2t},\tilde{a}_{3t},\tilde{a}_{4t},\tilde{a}_{5t},\tilde{a}_{6t}\}$ for the correlation in (3). If these were used to plot trends in educational attainment and in the correlation over time, sampling error would cause the plots to be bumpy. We expect that the actual trends are smooth, since they involve characteristics of the stock of married couples, which evolves gradually with time.

The second stage fits time polynomials to these coefficients, resulting in smooth temporal trends. For reasons described in Section 5, four time period dummies are used in place of a single intercept in the cutoff and correlation functions, to account for three changes in the data over the 1971-1996 period.

For each cutoff, the â's in (2) are modeled (suppressing the W,H and j subscripts) as:

$$\hat{a}_{0t} = f(\text{four time period dummies}) + \grave{e}_{01}t + \grave{e}_{02}t^2 + \grave{e}_{03}t^3 + \grave{e}_{04}t^4$$
$$\hat{a}_{1t} = \grave{e}_{10} + \grave{e}_{11}t + \grave{e}_{12}t^2 + \grave{e}_{13}t^3 + \grave{e}_{14}t^4$$

$$\hat{a}_{2t} = \grave{e}_{20} + \grave{e}_{21}t + \grave{e}_{22}t^2 + \grave{e}_{23}t^3 + \grave{e}_{24}t^4 \qquad (4)$$
$$\hat{a}_{3t} = \grave{e}_{30} + \grave{e}_{31}t + \grave{e}_{32}t^2 + \grave{e}_{33}t^3 + \grave{e}_{34}t^4$$
$$\hat{a}_{4t} = \grave{e}_{40} + \grave{e}_{41}t + \grave{e}_{42}t^2 + \grave{e}_{43}t^3 + \grave{e}_{44}t^4$$

This reduces the number of coefficients describing each cutoff from 100 (20 years × five quartic coefficients) to 28 (eight coefficients in the $\hat{a}_{0t}$ equation, five in each of the other four equations).

The $\tilde{a}$ coefficients in (3) are smoothed in a similar way:

$$\tilde{a}_{0t} = g(\text{four time period dummies}) + \grave{u}_{01}t + \grave{u}_{02}t^2 + \grave{u}_{03}t^3 + \grave{u}_{04}t^4$$
$$\tilde{a}_{1t} = \grave{u}_{10} + \grave{u}_{11}t + \grave{u}_{12}t^2 + \grave{u}_{13}t^3 + \grave{u}_{14}t^4$$
$$\tilde{a}_{2t} = \grave{u}_{20} + \grave{u}_{21}t + \grave{u}_{22}t^2 + \grave{u}_{23}t^3 + \grave{u}_{24}t^4 \qquad (5)$$
$$\tilde{a}_{3t} = \grave{u}_{30} + \grave{u}_{31}t + \grave{u}_{32}t^2 + \grave{u}_{33}t^3 + \grave{u}_{34}t^4$$
$$\tilde{a}_{4t} = \grave{u}_{40} + \grave{u}_{41}t; \quad \tilde{a}_{5t} = \grave{u}_{50} + \grave{u}_{51}t \; ; \quad \tilde{a}_{6t} = \grave{u}_{60} + \grave{u}_{61}t.$$

This reduces the number of coefficients summarizing the correlation from 140 (20 years × seven coefficients in (3)) to the 29 coefficients appearing on the right hand side of (5). We found that a linear trend alone is sufficient for modeling the age difference coefficients $\tilde{a}_{4t}$, $\tilde{a}_{5t}$ and $\tilde{a}_{6t}$. Specification tests are discussed in Section 6.1.

The right-hand side coefficients in (4) and (5) are estimated by GLS, using the estimates of the â's and ã's from the first stage as dependent variables and using their estimated variance-covariance matrices to construct the variance-covariance matrices for the error terms that consequently would appear in (4) and (5). Cross-equation contemporaneous correlation exists in the error terms that result from replacing the left-hand-side variables in (4) and (5) by their first-stage estimates.

This model does not have the commonly-used additive year (t), age (a_i) and cohort (as represented by birth year = t-a_i) effects structure (e.g. Deaton and Paxson (1994)). Additive effects imply

that $C(a_i,t)$ in (2), for example, is

$$C(a_i,t) = f_t(t) + f_a(a_i) + f_c(t-a_i). \qquad (6)$$

One way to see that our model is not additive is to substitute the â's from (4) into (2), giving C as a function of a_i and t as in (6). But if coefficients such as $è_{44}$ are nonzero, then terms like $t^4a_i^4$ appear in (6). These can only arise in an additive-effects model through a term like $(t-a_i)^8$ in the cohort effect, but such terms would also imply that terms such as t^8 or $t^6a_i^2$ should also appear in (6), and they do not. We found that additive effects do not accurately fit these data when quartic polynomials are used.

The plots given in Section 6 can be thought of as fixed-age or fixed-year "slices" of the three-dimensional surface C or ñ as a function of a_i and t. No attempt is made to assign trends to year or cohort effects, although we think of them as cohort effects.

5. Data

We use twenty Surveys of Consumer Finances from Canada for 1971, 1973, 1975, 1977, 1979, 1981, 1982, and 1984-1996. We take all married couples with the husband aged 25 to 75 and with information present on both spouses' education and age. This includes couples who are living common law, or "cohabiting", although we cannot distinguish these couples from officially-married couples. No same-sex couples are indicated in these data sets.

The education question changed in 1975 and 1989, and our data switches from census family to individual files in 1981. (For more details on these changes, see Bar-Or *et al.* (1995, pp. 789-791).) To capture any effects these changes might have on the levels of the cutoffs and the correlation, each of the intercept

equations in the second stage has time period dummies for the 1971-1973, 1975-1979, 1981-1988, and 1989-1996 periods as indicated in the first equations of (4) and (5).

The education categories are: “1" for less than high school, “2" for high school, “3" for some post-secondary, and “4" for a university degree. (See Bar-Or *et al.* (1995) for the algorithm converting the education variable for each year to these four-category variables.) The ordered nature of the categories is strong, but not perfect. For example, someone could be ranked “3" who has a diploma from some specialized training without having completed high school.

6. Results

6.1 Specification Tests

The first stage gives a total of 740 estimates: 37 coefficients (30 of them from the six cutoffs times five coefficients per cutoff in (2), plus seven from the correlation in (3)) for each of the 20 survey years. Out of the total of 120 coefficient estimates on the highest-order terms in the age quartics for the 6 cutoffs from each year, 78 are significantly different from zero at the 5% level and 67 are significant at the 1% level. There are a total of 40 highest-order terms in the average-age and age-difference cubics for the correlation from the 20 years, of which 9 and 5 estimates are statistically significant at the 5% and 1% levels respectively. These 40 t-statistics all are less that 3.3 in magnitude. We take this as informal evidence that underspecification of these first stage polynomials is not likely too serious, particularly in the correlations, where these tests would support reducing the order of most of the polynomials.

The second stage, as mentioned in section 4, imposes polynomial smoothing constraints, which reduce the number of

coefficients describing the model. If the first stage model is well-specified, these restrictions can be tested using asymptotically valid chi-square tests of the form:

$$W = (\hat{\hat{\alpha}}_{(2)} - \hat{\hat{\alpha}}_{(1)})^{T}[\text{EstVar}(\hat{\hat{\alpha}}_{(1)}]^{-1}(\hat{\hat{\alpha}}_{(2)} - \hat{\hat{\alpha}}_{(1)}),$$

where $\hat{\hat{\alpha}}_{(1)}$ is a vector of estimates of some of the $\hat{\alpha}$'s from the first stage, $\text{EstVar}(\hat{\hat{\alpha}}_{(1)})$ is the estimated variance-covariance matrix of $\hat{\hat{\alpha}}_{(1)}$, also from the first stage ML, assuming independence across years and across observations within year, and $\hat{\hat{\alpha}}_{(2)}$ is the estimate of the same vector of $\hat{\alpha}$'s from the second stage, as predicted from the estimated equations (4). With appropriately chosen $\hat{\alpha}$ vectors, the number of restrictions is clear.

The second-stage restrictions imposed on the intercepts of the cutoff functions are rejected at the 1% level in 5 of the 6 cases. There are 12 restrictions imposed on each intercept, as the 20 first-stage intercepts, $\hat{\alpha}_{0Wkt}$ or $\hat{\alpha}_{0Hkt}$, for 20 values of t, are modeled by 8 parameters as shown in the first equation of (4). Because we centred the age variable to equal 0 at age 50, these $\hat{\hat{\alpha}}_{0t}$'s can be interpreted as the ordered probit cutoffs for a person aged 50. The 6 test statistics range in value from 19.3 to 50.6, with a 1% critical value of 26.2.

Although the fraction of rejections is high, they are not emphatic rejections given the large sample sizes. Figure 1 plots the stage 2 fitted cutoffs (solid lines) against the stage 1 intercepts (dots) for the 6 cutoffs. They appear to fit very well. The period dummies play an important role for cutoffs 1 and 2 in 1975 and for cutoffs 2 and 3 in 1989, as seen by the jumps in the plots. These appear to derive from changes to the education question that occurred in those two years. Wald test statistics of the null hypothesis of no period effects, which is that the coefficients on these four dummies are equal, range from

83.2 to 528.6 across the six cutoffs. These all are emphatic rejections, testing just three restrictions.

The second stage reduces the total number of coefficients describing the correlations from 140 to 29. A Wald test accepts these 111 restrictions at the 5% level, with a statistic of 132.08, giving a P-value of 8.4%. The four period dummies are not significantly different from each other, with a Wald statistic of 0.34 and P-value of 95.2%. This indicates that the changes in the survey, which had a large effect on the cutoffs as seen in Figure 1 and the tests mentioned earlier, had little or no effect on the correlations.

6.2 The Education Correlation

The top panel of Figure 2 shows the estimated education correlation as a function of the average age of the couple, setting the age difference to zero, for three years. They are taken from the stage two estimates, so that the data from all years have influenced each plot due to the cross-year smoothing in stage two.

This plot shows the main finding of the paper. By the 1990s there had emerged, especially among the younger couples, a decrease in the education correlation. To check that the age effect on the correlation is statistically significant, a Wald test was used to test the three restrictions $\tilde{a}_{1t} = \tilde{a}_{2t} = \tilde{a}_{3t} = 0$ from (5), for each year. The Wald statistics for 1973, 1984 and 1995 are 5.7, 110.1 and 40.3 respectively, with a 5% critical value of 7.8. We reject both null hypotheses that the true 1984 and 1995 plots are horizontal lines, and conclude that the decreased correlation among younger couples, compared to older couples, which emerges by 1995 is statistically significant.

We also tested the null hypothesis that the 1984 and 1995 correlations are equal. This was done separately for ages 30, 40, 50, 60 and 70. The Wald statistics, each testing a single

restriction, are 23.8, 17.7, 8.0, 2.1 and 0.4 respectively, with a 5% critical value of 3.84. This confirms that there has been a statistically significant drop in the correlation among younger couples over this period.

It may be surprising to see a noticeable difference in the correlation between, say, a 40-year-old couple in 1984 and a 51-year-old couple in 1995, since many of the couples in these two populations (not in the samples) are the same. Possible explanations for this difference include: changes in the composition of this group over the 1984-1995 period arising from marriage, divorce, migration, death, or from a change in who gets classified as a married couple, although on this last point we are not aware of any such changes; changes in their education levels (e.g. they go back to school in their forties); and model misspecification.

The bottom panel of Figure 2 shows the change in these correlations over time holding the average age of the couple fixed, again setting the age difference to zero. This also displays a decrease in the correlation among young couples beginning in the early 1980s. The exact wavy patterns of these plots should not be trusted, but this decrease comes through clearly.

Figure 3 shows the effect of the age difference on the correlation, for a couple with average age 50, for selected years. The age difference in the Figure, defined as husband’s age minus wife’s age, ranges from -6 years to +15 years. This range covers about 97% of the couples in the sample, with roughly half of the remaining 3% at each end. For most years, the correlation is maximized approximately when the spouses’ ages are equal. The 1996 plot shows a lower correlation for couples with older husbands than the 1971 or 1984 plots. The Wald statistic for testing $\grave{u}_{41} = \grave{u}_{51} = \grave{u}_{61} = 0$, is 22.0, which is significant with 3 d.f. This rejects the null that the differences between the three

true plots are constant over time.

The relatively large magnitude of the age difference effect might lead one to wonder about changes in the distribution of the age difference itself over time. In our data, the age difference has decreased slightly and its variance has increased slightly. These changes have not been large enough to have had much of an effect on the correlations. The mode and median age difference in the overall sample is +2 years.

6.3 Trends in the Educational Attainment of Husbands and Wives

A summary of the trends in educational attainment of husbands and wives is not the main purpose of our model. There are other, possibly preferable, ways to examine this issue, for example by nonparametric smoothing over age and time. Nevertheless, this model provides probabilities of someone being in each education group as a function of their sex, age and year. These are shown in Figures 4 and 5. This parametric approach has the virtue of removing the effects of the changes in the survey through the use of the time period dummies, as seen by the jumps in Figure 1. In Figures 4 and 5, the time period effect is held fixed at its estimated value for the 1989-1996 period.

We do not look at these probabilities conditional on the educational attainment of the spouse. This could be done with calculations involving the latent variables and the correlation. However, there may be patterns in the actual conditional probabilities that get wiped out by our bivariate normality assumption. One way to study these in greater detail is with a ‘crossings’ model as in Mare (1991) and Qian (1998).

Figure 4 shows the probabilities as a function of time for husbands and wives aged 35 years. The trends are similar. In 1971, a much higher proportion of husbands than wives had a university degree, but by 1996 this proportion is higher for wives. The probability of a 35-year-old husband having a

university degree in Canada has not risen much since the late 1970s, whereas it increased throughout the sample for wives with no sign of slowing. By 1996, considerably more wives than husbands also are in the greater than high school category. More husbands than wives are in the less than high school and high school groups.

Figure 5 shows these same trends through cross-section plots for 1996. At every age, slightly more husbands than wives have less than high school education, but their age patterns are very similar. Older wives are more likely than older husbands to be in the high school group, but this switches for younger spouses, as mentioned above.

7. Conclusion

We have presented a bivariate probit method for measuring the correlation between spouses' education levels, which are recorded as ordered categories in Canada. The results were used to examine the effects of the year and the spouses' ages on the education correlation during the 1971-1996 period. A two-stage minimum chi square estimation technique was used.

Perhaps surprisingly, we find that the most noticeable trend is a decrease in this correlation among younger couples emerging in the 1980s. This differs from Pencavel (1998) and Mancuso and Pencavel (1999), who use U.S. data and different techniques. One way to explain this finding, following Mare's (1991) reasoning, is to note the increase in the age at which couples are getting married in Canada. As the time gap between the age of completion of one's education and the age of marriage grows, the importance of education as a sorting criterion is reduced. Also, we do not have information on whether these couples are in a first marriage. An increase in the number of people in second-or-more marriages could result in a reduced correlation for the same reason. Mancuso and Pencavel selected only couples in their first

marriage.

The model also allows us to plot age and year patterns in the educational attainment probabilities of husbands and wives, controlling for changes in the education question over the sample period. The trends of husbands and wives are similar in most respects. The growth in the proportion of young wives with a university degree or some post-secondary education is greater than that of young husbands, with the wives' proportion surpassing husbands' by 1996.

Other association measures are available. In future work, we plan to compare the U.S. and Canadian results using similar sample selection criteria and a variety of measures.

One reason for studying this issue is its implications for household income inequality. These implications are much more significant now than say 30 years ago because of the increase in labour force participation of wives. Some empirical work on assortative mating on earnings or wages has been done by Zimmer (1996) and others. (See the references in Zimmer.) An aim of our paper is to present some basic facts on the education aspect of the assortative mating process in Canada to have one more piece of the picture in the determinants of trends in income inequality across households. It appears that there has been a decrease, not an increase, in assortative mating by education among younger couples in Canada. This would not appear to be a cause of any increase in household income inequality in recent decades.

References

Bar-Or, Y., J. Burbidge, L. Magee and A.L. Robb (1995), "The Wage Premium to a University Education in Canada, 1971-1991," *Journal of Labour Economics*, 13(4), 762-794.

Burbidge, J.B., L. Magee and A.L. Robb (1997), "Cohort, Year and Age Effects in Canadian Wage Data," IESOP Research Paper No. 19, McMaster University.

Deaton, A. and C. Paxson (1994), "Saving, Growth and Aging in Taiwan," in D. Wise (ed.), *Studies in the Economics of Aging*, University of Chicago Press, Chicago, 331-357.

Goodman, L.A. (1984), *The Analysis of Cross-Classified Data Having Ordered Categories*, Harvard University Press, Cambridge.

Goodman, L.A., and W.H. Kruskal (1954), "Measures of Association for Cross Classifications," *Journal of the American Statistical Association*, 49, 732-764.

Goodman, L.A., and W.H. Kruskal (1979), *Measures of Association for Cross Classifications*, Springer-Verlag, New York.

Mancuso, D. and J. Pencavel (1999), "Assortative Mating by Schooling and the Timing of Marriage," manuscript.

Mare, R.D. (1991), "Five Decades of Educational Assortative Mating," *American Sociological Review*, 56, 15-32.

Pearson, K. (1900), "Mathematical Contributions to the Theory of Evolution VII: On the Correlation of Characters Not Quantitatively Measurable," *Philosophical Transactions of the Royal Society of London*, 195A, 1-47.

Pencavel, J. (1998), "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands," *American Economic Review*, 88(2), 326-329.

Qian, Z. (1998), "Changes in Assortative Mating: The Impact of Age and Education, 1970-1990," *Demography*, 35(3), 279-292.

Zimmer, M. (1995), "Assortative Mating and Ethnicity in the Low Wage Population: An Examination of Spouses' Earnings," *Applied Economics Letters*, 3, 311-315.

Appendix: Maximum Likelihood Estimation

The log likelihood function is $\ell(\hat{a},\varsigma) = \sum_i \ell_i(\hat{a},\varsigma)$, where

$$\ell_i(\hat{a},\varsigma) = \ell(\hat{a},\varsigma|x_{w,i},x_{H,i},w_i,H_i,W_i) = \ln(\Ö_{2,i}) =$$
$$\ln(\Ö_2(x_{W,i}^T\hat{a}_{W,Wi-1},x_{W,i}^T\hat{a}_{W,Wi};\ x_{H,i}^T\hat{a}_{H,Hi-1},x_{H,i}^T\hat{a}_{H,Hi};\ w_i^T\varsigma)),$$

where $\Ö_2(a,b;c,d;\ñ) = \int_a^b\int_c^d \ö_2(z_H,z_W|\ñ)dz_Hdz_W$ and $\ö_2$ is given in (1). The partial derivatives are:

$$\partial\ell_i/\partial\hat{a}_{W,k} = x_{W,i}\{I[W_i=k](\partial\Ö_{2,i}/\partial b) - I[W_i=k+1](\partial\Ö_{2,i}/\partial a)\}$$

$$= x_{W,i}\ö(x_{W,i}^T\hat{a}_{W,k})\{I[W_i=k][(\Ö((x_{H,i}^T\hat{a}_{H,Hi} - \ñ x_{W,i}^T\hat{a}_{W,k})/(1-\ñ^2)^{1/2}) -$$
$$\Ö((x_{H,i}^T\hat{a}_{H,Hi-1} - \ñ x_{W,i}^T\hat{a}_{W,k})/(1-\ñ^2)^{1/2}))]$$
$$- I[W_i=k+1][(\Ö((x_{H,i}^T\hat{a}_{H,Hi} - \ñ x_{W,i}^T\hat{a}_{W,k})/(1-\ñ^2)^{1/2}) -$$
$$\Ö((x_{H,i}^T\hat{a}_{H,Hi-1} - \ñ x_{W,i}^T\hat{a}_{W,k})/(1-\ñ^2)^{1/2}))]\},$$

where $\ö$ and $\Ö$ are the standard normal pdf and cdf functions. A similar derivative for $\partial\ell_i/\partial\hat{a}_{H,k}$ is obtained by switching W and H in the above expression. The remaining derivative is

$$\partial\ell_i/\partial\varsigma = w_iG_i/\Ö_{2,i},$$

where $G_i = \partial\Ö_{2,i}/\partial\ñ = G(x_{W,i}^T\hat{a}_{W,Wi-1},x_{W,i}^T\hat{a}_{W,Wi};\ x_{H,i}^T\hat{a}_{H,Hi-1},x_{H,i}^T\hat{a}_{H,Hi};\ w_i^T\varsigma))$, with

$$G(a,b;c,d;\ñ) =$$
$$\int_a^b\int_c^d \ö_2(z_H,z_W|\ñ)\{(1-\ñ^2)^{-1}[\ñ + z_Hz_W - \ñ(z_H^2+z_W^2-2\ñ z_Hz_W)/(1-\ñ^2)]\}dz_Hdz_W.$$

Let $\hat{\hat{e}}$ be the entire vector of estimates of the â's and ç, and let $\hat{\hat{L}}$ be the n-by-k derivative matrix with i^{th} row $\ell_i'^T$, where ℓ_i' comprises the $\partial\ell_i/\partial\hat{a}_{W,k}$ and $\partial\ell_i/\partial\varsigma$ vectors described above evaluated at $\hat{\hat{e}}$. The ML estimates were found by hill-climbing using the OPG matrix:

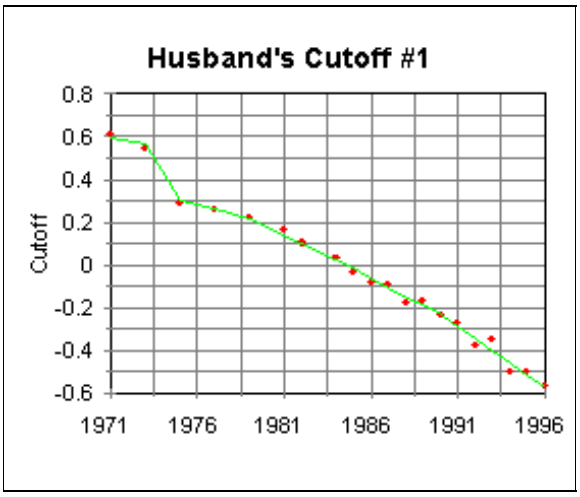
$$\hat{\theta}_{(m+1)} = \hat{\theta}_{(m)} + (\hat{L}^T\hat{L})^{-1}\hat{L}^T\acute{e}, \qquad \text{(A1)}$$

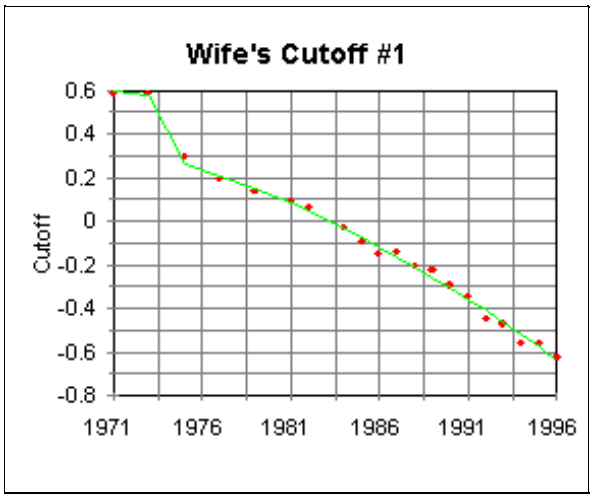
where é is a vector of ones. Upon convergence, using $\acute{e}^T\hat{L}(\hat{L}^T\hat{L})^{-1}\hat{L}^T\acute{e}$ as the convergence criterion, $(\hat{L}^T\hat{L})^{-1}$ was used as the estimator of $\mathrm{Var}(\hat{\theta})$.

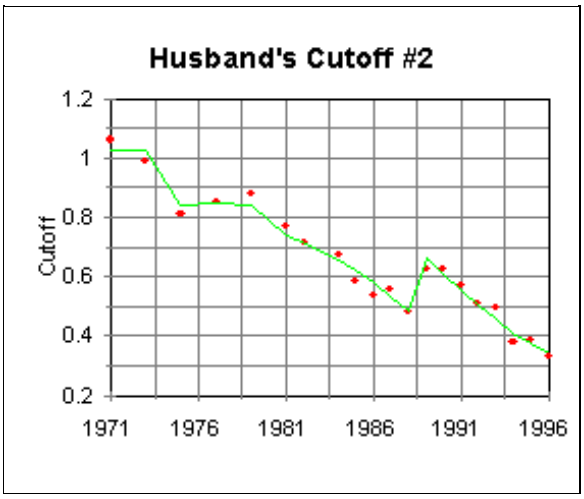
During the hill-climbing iterations, the cutoff functions occasionally crossed. This results in a violation of the inequality between lower and upper bounds necessary for integration. When this occurred, the step size in (A1) was halved until the crossings no longer occurred. If the log likelihood at $\hat{\theta}_{(m+1)}$ was smaller than at $\hat{\theta}_{(m)}$, the step size was halved successively until the log likelihood was higher at the new $\hat{\theta}_{(m+1)}$.

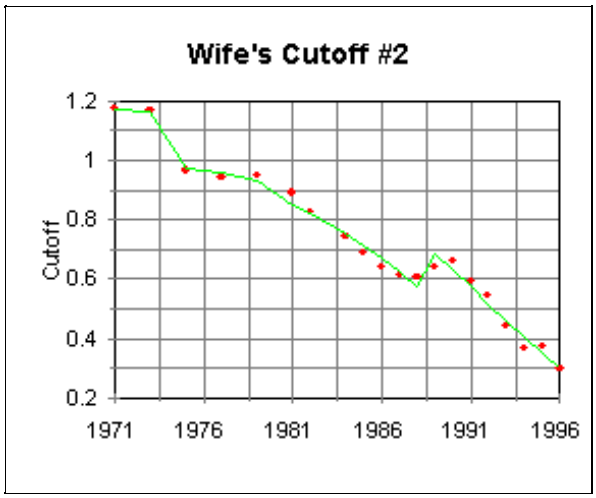
A similar problem occurred with estimates of $\tilde{n}_i$ lying outside the [-1,1] range. This resulted from a small number of observations where the age difference between spouses was large. We handled this by truncating the age difference variable, d_i in (3), to a minimum of -20 years and a maximum of 20 years. This affected only a tiny fraction of the observations.

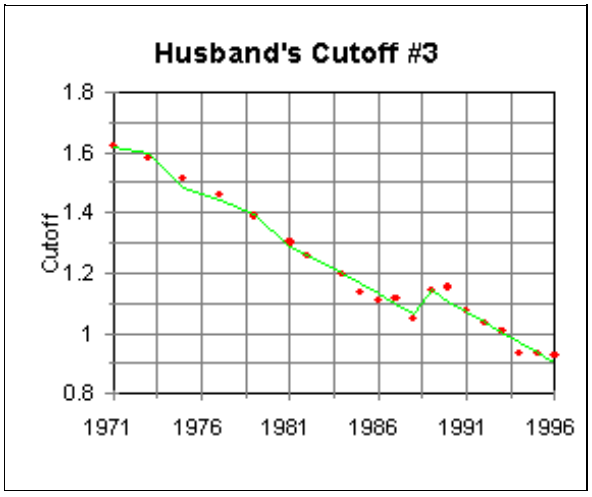
Figure 1. Stage 2 Fits of Stage 1 Intercept Estimates











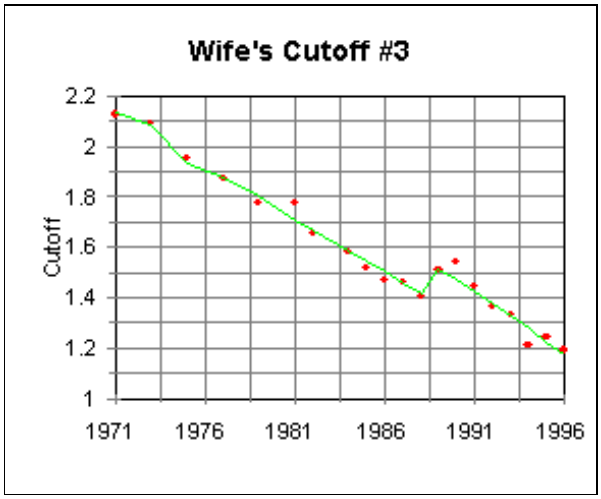
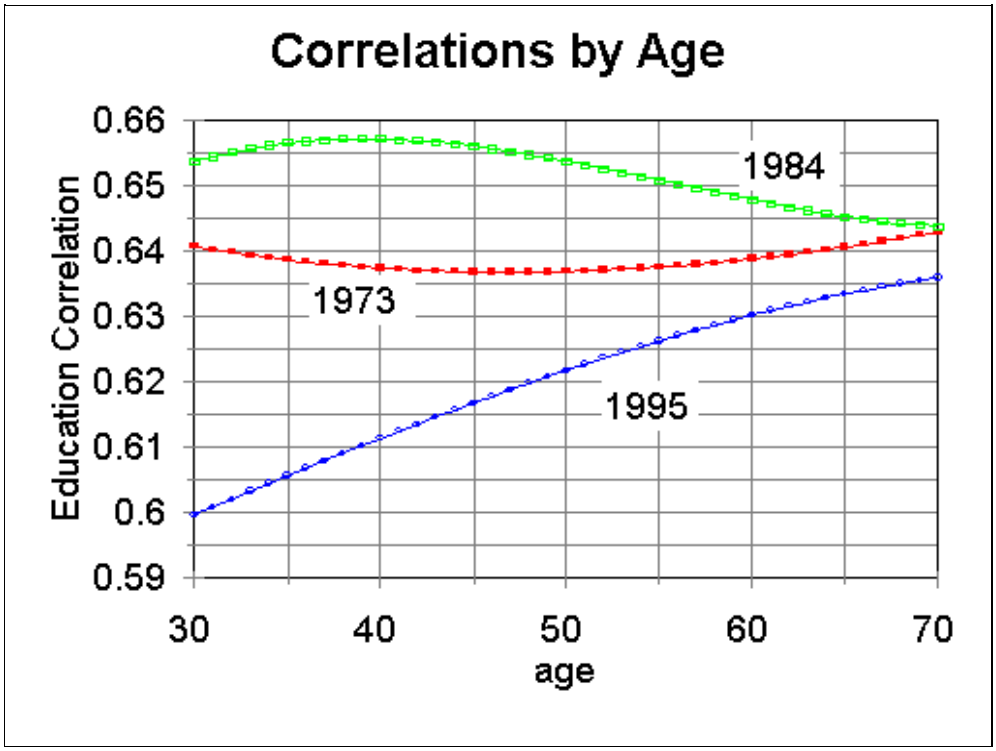


Figure 2. Correlations by Age and Year



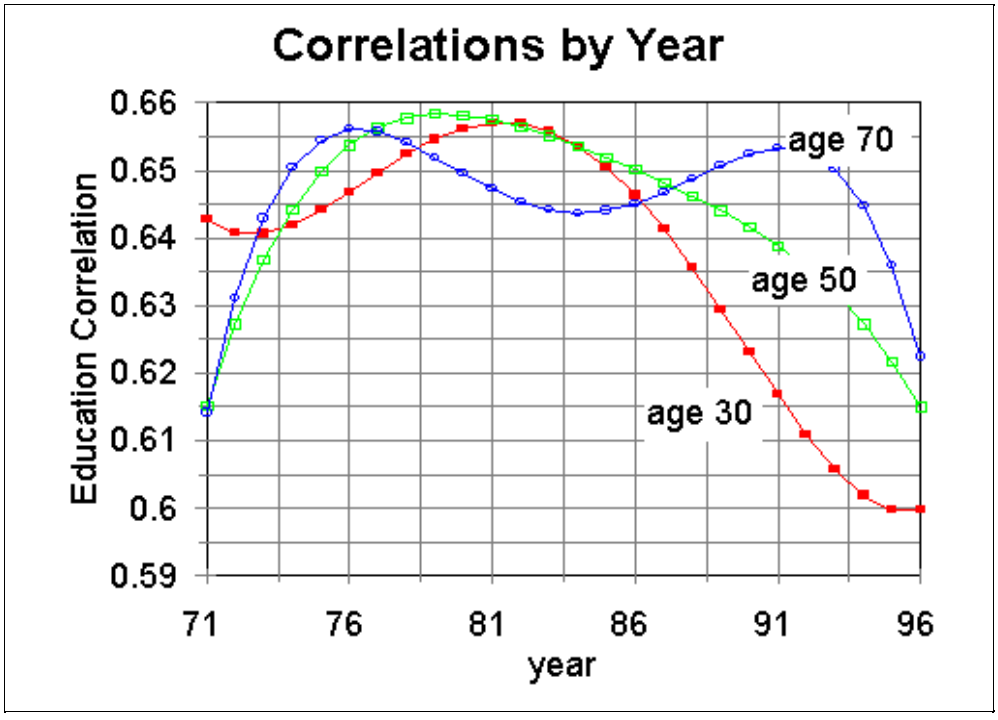


Figure 3. Age Difference Effect on Correlation

Average Age = 50

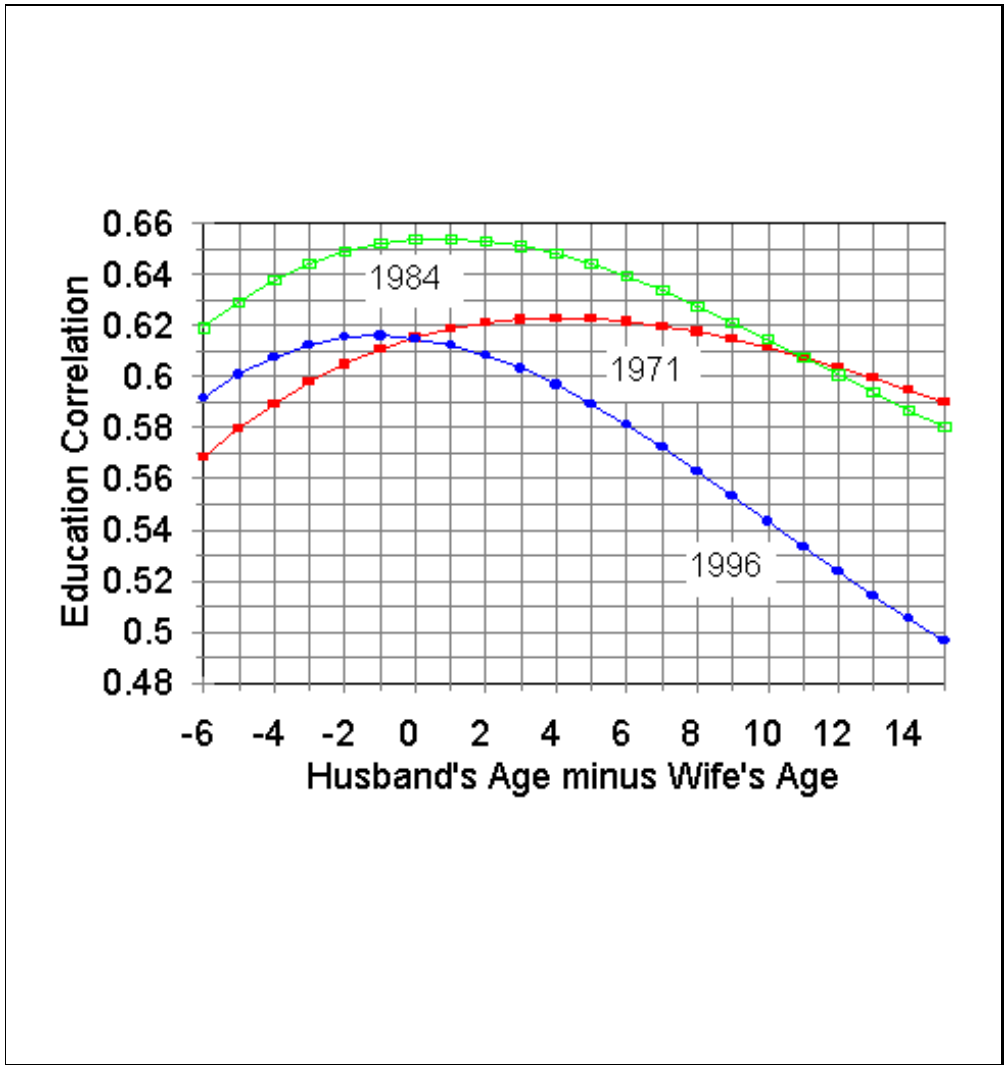
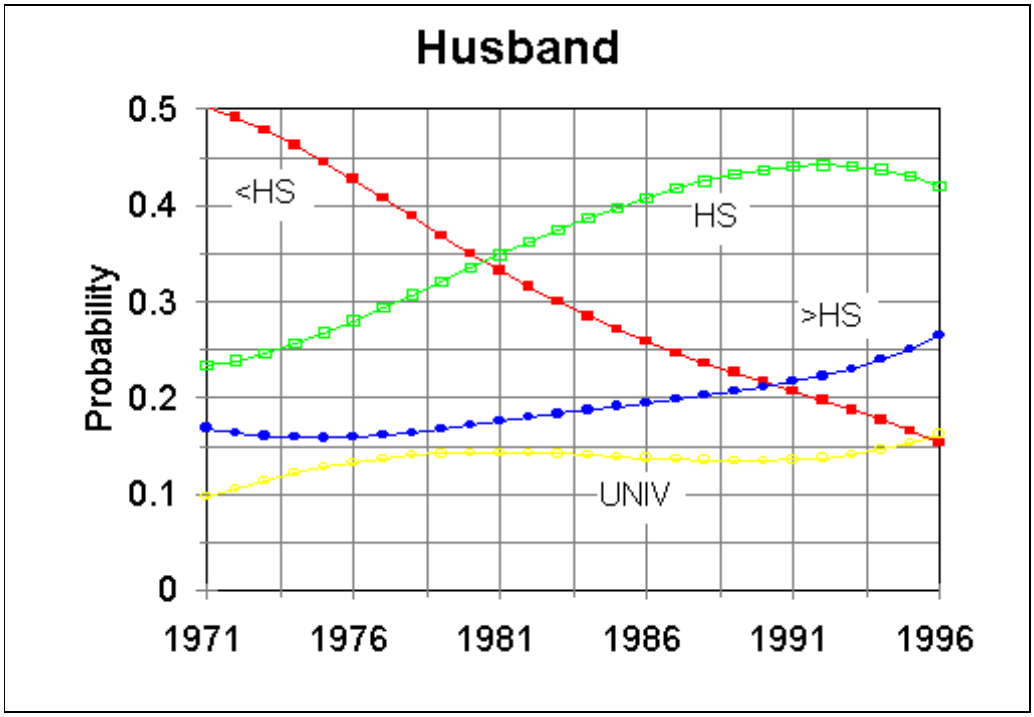


Figure 4. Education Level Probabilities By Year, Age = 35



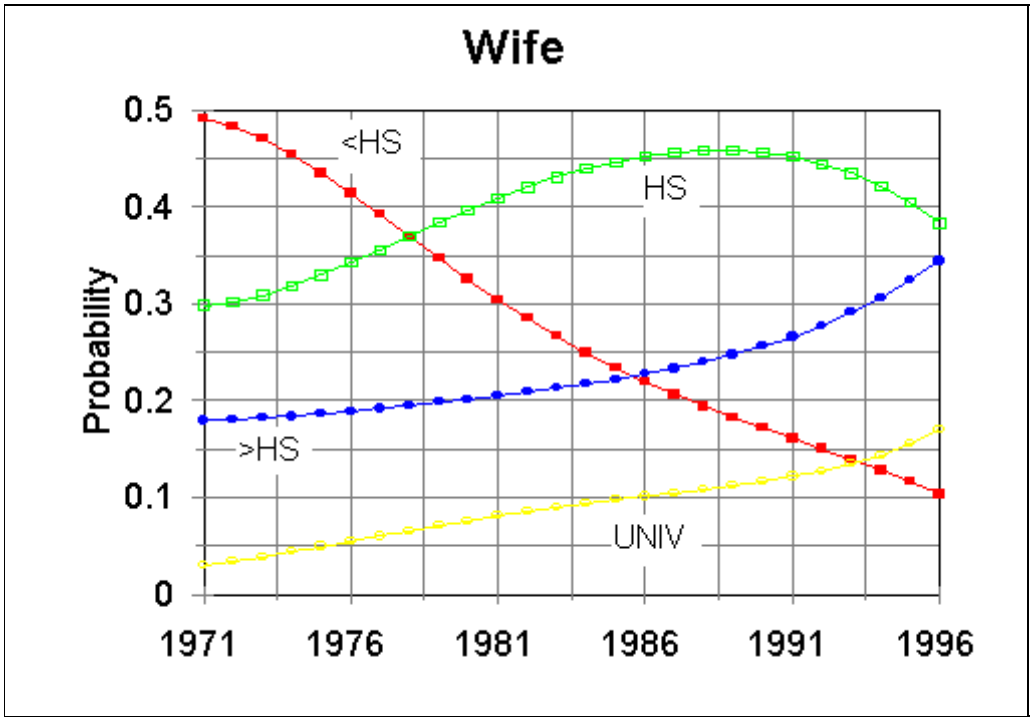
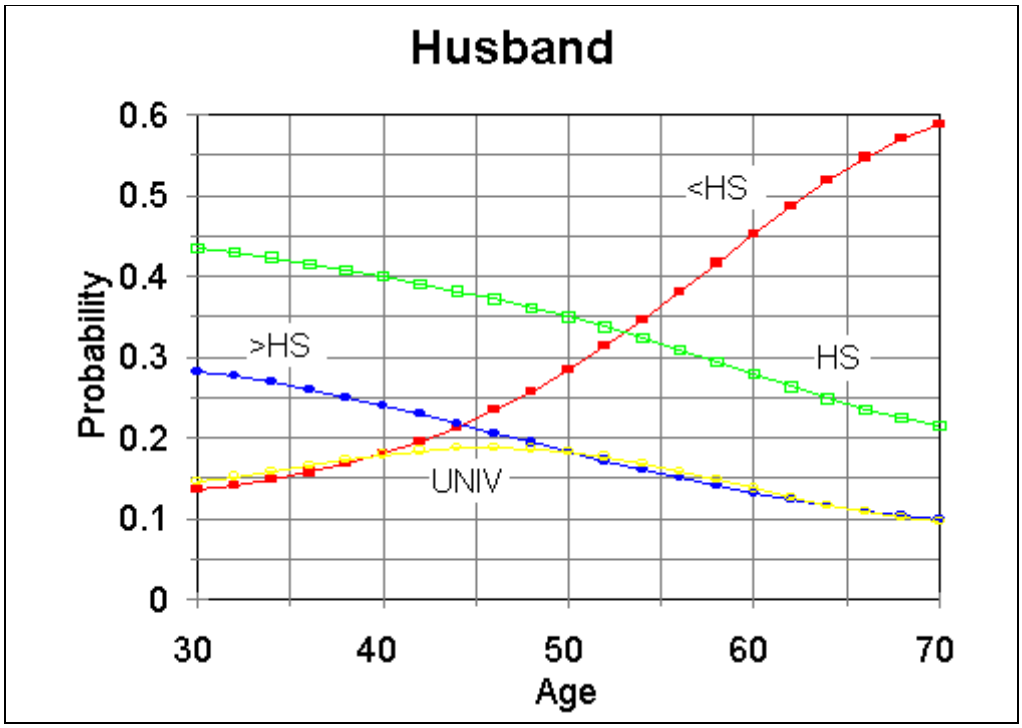
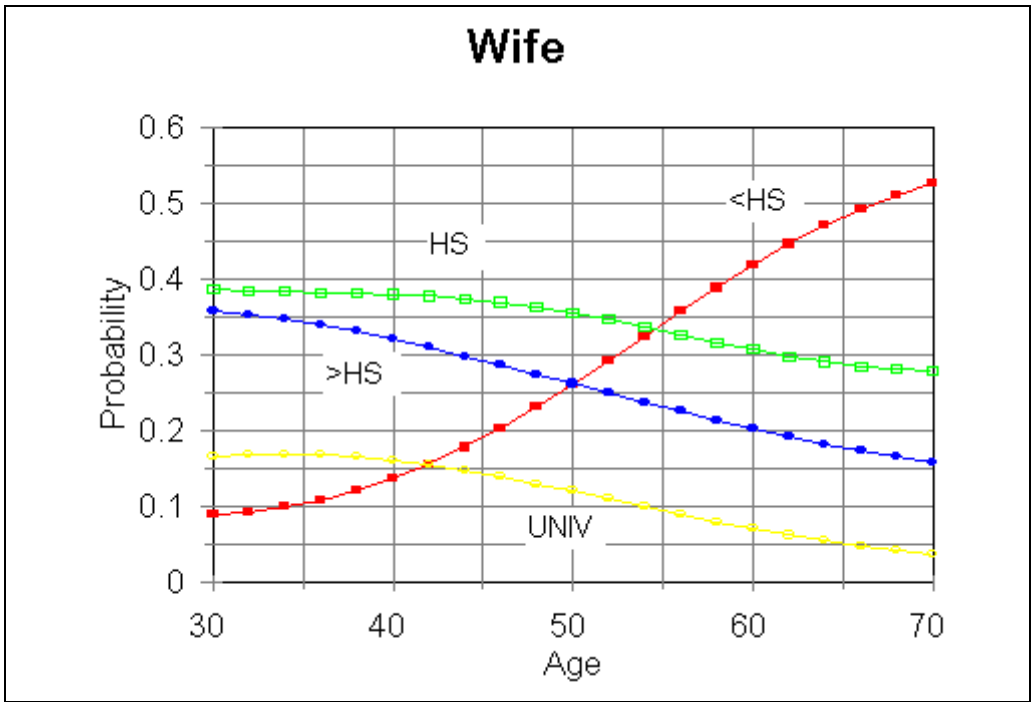


Figure 5. Education Level Probabilities By Age, 1996





SEDAP RESEARCH PAPERS

Number	Title	Author(s)
No. 1:	Population Aging and Its Economic Costs: A Survey of the Issues and Evidence	F.T. Denton B.G. Spencer
No. 2:	How Much Help Is Exchanged in Families? Towards an Understanding of Discrepant Research Findings	C.J. Rosenthal L.O. Stone
No. 3:	Did Tax Flattening Affect RRSP Contributions?	M.R. Veall
No. 4:	Families as Care-Providers Versus Care-Managers? Gender and Type of Care in a Sample of Employed Canadians	C.J. Rosenthal A. Martin-Matthews
No. 5:	Alternatives for Raising Living Standards	W. Scarth
No. 6:	Transitions to Retirement: Determinants of Age of Social Security Take Up	E. Tompa
No. 7:	Health and Individual and Community Characteristics: A Research Protocol	F. Béland S. Birch G. Stoddart
No. 8:	Disability Related Sources of Income and Expenses: An Examination Among the Elderly in Canada	P. Raina S. Dukeshire M. Denton L.W. Chambers A. Scanlan A. Gafni S. French A. Joshi C. Rosenthal
No. 9:	The Impact of Rising 401(k) Pension Coverage on Future Pension Income	W.E. Even D.A. Macpherson
No. 10:	Income Inequality as a Canadian Cohort Ages: An Analysis of the Later Life Course	S.G. Prus
No. 11:	Are Theories of Aging Important? Models and Explanations in Gerontology at the Turn of the Century	V.L. Bengtson C.J. Rice M.L. Johnson
No. 12:	Generational Equity and the Reformulation of Retirement	M.L. Johnson
No. 13:	Long-term Care in Turmoil	M.L. Johnson L. Cullen D. Patsios
No. 14:	The Effects of Population Ageing on the Canadian Health Care System	M.W. Rosenberg

Number	Title	Author(s)
No. 15:	Projections of the Population and Labour Force to 2046: Canada	F.T. Denton C.H. Feaver B.G. Spencer
No. 16:	Projections of the Population and Labour Force to 2046: The Provinces and Territories	F.T. Denton C.H. Feaver B.G. Spencer
No. 17:	Location of Adult Children as an Attraction for Black and White Elderly Migrants in the United States	K.-L. Liaw W.H. Frey J.-P. Lin
No. 18:	The Nature of Support from Adult *Sansei* (Third Generation) Children to Older *Nisei* (Second Generation) Parents in Japanese Canadian Families	K.M. Kobayashi
No. 19:	The Effects of Drug Subsidies on Out-of-Pocket Prescription Drug Expenditures by Seniors: Regional Evidence from Canada	T.F. Crossley P. Grootendorst S. Korkmaz M.R. Veall
No. 20:	Describing Disability among High and Low Income Status Older Adults in Canada	P. Raina M. Wong L.W. Chambers M. Denton A. Gafni
No. 21:	Parental Illness and the Labour Supply of Adult Children	P.T.Léger
No. 22:	Some Demographic Consequences of Revising the Definition of 'Old' to Reflect Future Changes in Life Table Probabilities	F.T. Denton B.G. Spencer
No. 23:	Geographic Dimensions of Aging: The Canadian Experience 1991-1996	E.G. Moore D. McGuinness M.A. Pacey M.W. Rosenberg
No. 24:	The Correlation Between Husband's and Wife's Education: Canada, 1971-1996	L. Magee J. Burbidge L. Robb