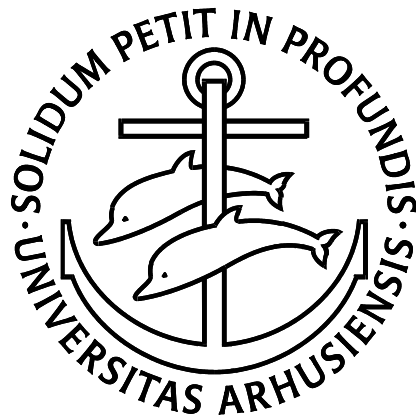# SCHOOL OF ECONOMICS AND MANAGEMENT

## ECONOMICS WORKING PAPER 2006-10

Educational Homogamy:
Preferences or Opportunities?

Helena Skyt Nielsen and Michael Svarer

# UNIVERSITY OF AARHUS

# Educational Homogamy: Preferences or Opportunities?[*]

Helena Skyt Nielsen

School of Economics and Management, University of Aarhus        IZA

Michael Svarer

School of Economics and Management, University of Aarhus        CAM

August 16, 2006

## Abstract

Individuals match on length and type of education. We investigate whether the systematic relationship between educations of partners is explained by opportunities (e.g. low search frictions) or preferences (e.g. complementarities in household production or portfolio optimization). We find that half of the systematic sorting on education is due to low search frictions in marriage markets of the educational institutions. The other half is attributed to complementarities in household production, since income properties of the joint income process show no influence on partner selection.

JEL codes:    J12, J24.

Keywords: positive assortative matching on education, search frictions, hedging, complementarities in household production.

# 1   Introduction

There exists a substantial literature showing that individuals form partnerships with individuals with similar levels of education. The coefficient of correlation between spousal level of education is among the highest between different personal characteristics.[1] In a recent study by Fernandez et al. (2005), the mean correlation is 0.6 based on information from household surveys from 34 countries.

The purpose of this study is to investigate whether the systematic relationship between the educations of the partners is explained by opportunities or preferences.

In relation to the first explanation, educational institutions are presumably very efficient marriage markets. The density of potential partners is rather high (see e.g. Goldin (1992) and Lewis & Oppenheimer (2000)), and search frictions are therefore smaller than in other local marriage markets (see e.g. Gautier et al. (2005) for a model that analyzes the effect of search frictions on marriage market outcomes). That educational institutions function as marriage markets is also rooted in sociology. In Scott (1965) and Blau and Duncan (1967), it is argued that parents place their children in good colleges in order to secure the social position of the family. There is also ample evidence that partnerships form in schools. Laumann et al. (1994) report – based on a US survey conducted in 1992 – that 23% of married couples met their current partner in school. In a Dutch Survey from 1995, 15% reported to have met their current partner in school (Kalmijn & Flap (2001)). In fact, among the different sets of shared settings (neighborhood, family overlap, workplace etc.) the most common place for couples to meet before a partnership is initiated at the same school. In the present analysis, which is based on Danish register-based data, we find that for 20% of the couples the two partners have attended the same educational institution.

It could also be the case that educational homogamy is the outcome of a decision problem solved by rational agents. That is, an individual with a similar level, and perhaps same type, of education might be preferred to an individual with a different level of education. In the following, we focus on two mechanisms for preference-based partnership choice.

First, it might be the case that the mating of different educational groupings occurs as

---

[1] Typically, individuals match positively on individuals traits. This is the case for e.g. income, height, weight, IQ, and parents' characteristics (see e.g. Epstein & Guttman (1984) and Schafer & Keith (1990)). Besides age, education is the trait that has the highest bivariate correlation.

a result of rational behavior of risk averse agents who seek to optimize discounted utility in an environment where future income is uncertain. A number of papers highlight the interdependence between risk sharing and marriage. In their seminal paper, Kotlikof & Spivak (1981) showed that the expected gain that a risk averse agent can expect from the risk sharing elements of marriage formation amounts to 10-20% of his wealth. Since then, Rosenzweig & Stark (1989), Micevska (2002), Chen, Chiang & Leung (2003), and Hess (2004), among others, have investigated related aspects of partnership formation and dissolution in association with the presence of idiosyncratic income risk. The idea is that risk averse agents can benefit from forming marriage with others to insure against unforeseen changes in income. Along the lines of Hess (2004), a good economic match has a high mean income, a low income volatility and an income process that negatively correlates with ones own income process, much like a financial asset portfolio. In the present analysis, we consider matching between individuals with different educations. The income variables are generated as time series means for different educational groupings. As a consequence, the income processes are exogenous to the specific partnership, and we implicitly assume that the agents are able to predict the future income components for different educational groups.

Second, it could also be the case that educational attainment of spouses are complements in the household production function. Becker (1973) argues that positive assortative mating is optimal when traits are complements. According to this argument, a reason why two partners with the same education form a partnership is that they tend to appreciate the same public goods or the same kind of leisure. It is not obvious how to identify to what extent educational traits are complements in the household production function, although it is commonly assumed to be the case (see e.g. Chiappori et al. (2006)). In a recent speed-dating experiment, Fisman et al. (2006) found no evidence of preferences for same field of study. In the present analysis, we attribute the part of the realized partnership formation between individuals that cannot be explained by opportunities (that is, proximity of partners) or by portfolio choices to complementarities in household production.[2]

When it comes to dissolution of relationships, a systematic relationship between divorce rates and education has also been established. In the divorce literature, it is typically

---

[2]There might be more mechanisms that imply systematic matching in education than the ones mentioned here. In the present analysis these will be considered to belong to the residual group and will somewhat crudely be labeled complementarities in household production.

found that level of education (i.e. the sum of years of education of the spouses) is more important than similarity in education among spouses[3] (see e.g. Weiss & Willis (1997), Svarer (2004), and Charles & Stephens (2004)). More detailed information on the level and type of education could be used to make inference about match quality. If the factors related to education that influence partnership formation also influence dissolution risk, there is presumably a systematic difference in match quality which is related to those factors. In order to reconcile the dissolution analysis with the partnership formation analysis in a setup where we assume that decisions about whom to form a partnership with are taken by rational and forward-looking individuals, there must be some disturbances to the partnership that trigger dissolution. We will have this in mind when we present the partnership dissolution analysis.

In this paper, we exploit a rich register data set to disentangle the correspondence between education and marriage market behavior. We have detailed information on individuals' educational attainment, including the exact type of education and where the education was taken. After combining with information on individual income, we investigate to which extent educational portfolios of couples reflect low search frictions, complementarities in household productions and portfolio optimization. From a more general perspective, the analysis allows us to evaluate whether the systematic relationship between the educations of the partners is explained by opportunities (low search frictions) or preferences (complementarities or portfolio optimization).

We find that half of the systematic sorting on education is due to low search frictions in marriage markets of the educational institutions. The other half is attributed to complementarities in household production, since income properties of joint income process show no influence on partner selection. In addition, we find that dissolution risk is not strongly affected by the factors that determine partnership choice. That is, although individuals are more likely to form a partnership with an individual who has attended the same educational institution, the evidence that the distance affects the duration of the relationship is rather weak.

The structure of the paper is as follows: In section 2, we describe the data set. In section 3, we take a closer look at partnership formation. In section 4, we investigate the process of partnership dissolution, and in section 5, we conclude.

---

[3]That is, in terms of lowering the divorce risk, it is better for a low educated individual to marry a person with a high level of education rather than to marrying a person with a similar level of education .

# 2 Data

The data that we use to test our hypotheses come from IDA (Integrated Database for Labour Market Research) created by Statistics Denmark. The information comes from various administrative registers that are merged in Statistics Denmark. The IDA sample used here contains (among other things) information on marriage market conditions for a randomly drawn sub-sample of all individuals born between January 1, 1955 and January 1, 1965. The individuals are followed from 1980 to 1995. The data set enables us to identify individual transitions between different states of the marriage market on an annual basis. In addition, we have information on a number of background characteristics for the individuals as well as for their partners. Information on marriage market status is based on a register that collects information on who is living in all housing units in Denmark. This implies that an individual is registered as either cohabiting or married if they have the same residential information as their partner. We are not able to identify relationships between individuals who do not live together.

Below, we describe in detail how the data is organized and present the most important statistics. First, we explain how we treat different educations. Second, we describe how income measures and distance measures are constructed.

## 2.1 Educational grouping

Throughout the paper, we assume that individuals first decide on where to pursue education after they have finished high school (typical graduation age is 18-19 years), and this is also when they start to search in the marriage market. This assumption implies that educational grouping may be regarded as exogenous in the matching and dissolution analyses. In order to comply with this assumption, we assign the first education an individual attends after high school to the individual for the rest of the sample period. Individuals who change education or drop out are assumed to belong to the educational group they were first assigned to. We impose this restriction to reduce the presence of endogeneity in choice of education. To the extent that individuals are already in a partnership when they start the education, the assumption would be violated, since the decicion regarding education and educational institution might be coordinated with the partner's decision.

The available educational information gives a complete picture of an individual's educational history. Individuals are grouped according to the educational information. In order to focus on colleges as marriage markets and educational homogamy, we restrict at-

tention to high school graduates. Although this implies a substantial reduction in sample size (see below), it enables us to give a very detailed description of partnership formation and dissolution for individuals in partnerships where both partners have graduated from high school. Generally, the intention is to group individuals into educational groups by the first education they enrolled in after high school, as long as this is not an additional high school education.

All individuals with educational information are then divided into 13 educational groups[4] which differ in level and subject of education. In Table 1, we give an overview of the sample reduction. The representative gross sample consists of 26,048 individuals making up about 20,000 couples. When all relevant individuals have been assigned to one of the educational groups, this results in 2,965 couples.

TABLE 1

SAMPLE SELECTION

|  | Number of observations |
| --- | --- |
| Sample of individuals born 1955-1965 | 26,048 |
| Relationships formed between 1980-1995 | 19,938 |
| - where both partners completed high school | 3,144 |
| | |
| **Partnership formation analysis** | |
| Relationships where information on institution and | |
| municipality of education is available for both partners | 2,965 |
| | |
| **Partnership dissolution analysis** | |
| Relationships formed before 1994 and | |
| where information on institution and | |
| municipality of education is available for both partners | 2,896 |

Now we go through the definition of educational groups in detail. The first group consists of those individuals who do not enroll in an education after high school. The remaining 12 groups then consist of individuals who enroll in one of the following educations: vocational education and training, short-cycle higher education, medium-cycle higher education and long-cycle higher education. Individuals who enroll in vocational education and training are subdivided into two groups where one consists of the mercantile educations, such as sales assistants, and the other consists of both crafts, such as electricians or plumbers, and health or pedagogical-related educations, e.g. orthopedists. The short-cycle higher educations are all grouped together and are subject-wise more

---

[4]The number of educational groups are restricted by the number of individuals in each group. We have tried to form groups that are as homogenous as possible while still having a sufficient number of observations to calculate the desired statistics.

diverse than the other groups. Examples of short-cycle higher educations are real estate agents and various forms of technicians.

Individuals who enroll in medium-cycle higher educations are subdivided into five groups. The first group consists of pedagogical educations, such as nursery teachers and social workers. The second group comprises school teachers at the basic and lower secondary level. The third group consists of educations that lead to jobs in the public health system, e.g. nurses and physical therapists. The fourth group consists of educational subjects within the range of humanities and business, e.g. journalists, librarians, and graduate diplomas in business administration. The final group comprises technical, veterinary, agricultural, and military educations, e.g. engineers.

Finally, individuals who enroll in long-cycle higher educations are divided into four groups: the humanities, the natural and technical sciences, the social sciences, and the medical sciences. These are university educations. In Table 2, the distribution of males and females across educational groups are presented.

From Table 2, we see that women are overrepresented in educational groups that contain social and pedagogical and health care elements, whereas men are more inclined to take an education in natural sciences, technical, veterinary and agricultural sciences, and social sciences.

TABLE 2

EDUCATIONAL GROUPS, NUMBER OF OBSERVATIONS

| Education group | Men | Women |
|---|---|---|
| (1) No further education than high-school | 208 | 181 |
| (2) Vocational - mercantile | 206 | 262 |
| (3) Vocational - health and crafts | 152 | 136 |
| (4) Short-cycle further education | 142 | 280 |
| *Medium-cycle further education* | | |
| (5) Social and pedagogical | 44 | 199 |
| (6) School teacher | 218 | 262 |
| (7) Health care | 39 | 440 |
| (8) Humanities and social sciences | 51 | 64 |
| (9) Technical, veterinary and agricultural sciences | 377 | 81 |
| *Long-cycle further education* | | |
| (10) Humanities | 218 | 332 |
| (11) Natural and technical sciences | 590 | 250 |
| (12) Social sciences | 547 | 331 |
| (13) Medical sciences | 173 | 147 |
| In total | 2965 | 2965 |

In Table 3, we report the ratio of actual to expected frequency of a given educational combination in order to get an overview of the educational combinations of couples. The expected frequency is the number of couples in a given cell had the matching been random

7

by education.[5] For instance, for couples within medical sciences (combination 13-13, see Table 2), the expected frequency would be 8.58 (147*173/2965), whereas the actual frequency is 33 couples; this makes up a ratio of 3.85 (33/8.58). Ratios above 1 are highlighted, and they represent educational combinations that are more common than would be the case under random matching by education.

A pattern of positive assortative matching on education shows up. All couples with the same education (i.e. at the diagonal) are systematically more common.[6] However, there are large differences between the tendency of homogamous marriage. In Appendix A, Table A1, we show the ranking of couples by ratio. The top three couples are social and pedagogical couples, teacher couples and medical science couples. The first ten places on the ranking are occupied by couples with the same educations. Among the couples that do not consist of people with similar educational attainment, the more popular are, as expected: female nurses who mix up with male medical doctors (rank 19).

---

[5]Note that the expected frequencies are calculated based on the numbers in Table 2 and not on the marginal distributions for the whole of Denmark.

[6]Couples on the diagonal make up 22% of all couples, whereas couples with the same length of education amount to 43% of all couples in our sample.

Table 3

Ratios of actual to expected frequency

| | | | | | | | Females | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Males | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 1 | **2.49** | **1.47** | **1.96** | 0.88 | **1.08** | 0.72 | **1.00** | 0.50 | 0.93 | 0.65 | 0.39 | 0.53 | 0.59 |
| 2 | **1.69** | **3.06** | 0.82 | **1.12** | 0.79 | 0.67 | 0.60 | 0.35 | 0.42 | 0.51 | 0.45 | 0.79 | 0.61 |
| 3 | **1.92** | **1.07** | **1.32** | **1.36** | **1.42** | 0.83 | 0.94 | **1.56** | 0.25 | **1.16** | 0.40 | 0.43 | 0.55 |
| 4 | **1.54** | **1.13** | **1.39** | **2.72** | 0.96 | 0.71 | 0.90 | 0.00 | **1.05** | 0.51 | 0.51 | 0.58 | 0.43 |
| 5 | 0.32 | 0.66 | 0.51 | 0.50 | **6.69** | **1.30** | **1.09** | 0.00 | 0.00 | **1.27** | 0.00 | 0.00 | 0.00 |
| 6 | 0.38 | 0.26 | 0.52 | 0.70 | **1.85** | **4.53** | 0.91 | **1.31** | 0.52 | 0.55 | 0.28 | 0.56 | 0.58 |
| 7 | 0.00 | 0.49 | **1.73** | 0.28 | 0.00 | 0.88 | **2.98** | **1.22** | 0.97 | 0.72 | 0.63 | **1.44** | 0.00 |
| 8 | 0.55 | 0.38 | 0.44 | 0.43 | **2.13** | 0.23 | **2.28** | **2.80** | 0.74 | **1.28** | 0.96 | 0.74 | 0.00 |
| 9 | 0.83 | **1.34** | 0.93 | 0.93 | 0.80 | 0.74 | **1.30** | 0.74 | **3.61** | 0.79 | 0.82 | 0.65 | 0.76 |
| 10 | 0.63 | 0.17 | **1.02** | 0.55 | 0.98 | **1.04** | 0.68 | **2.37** | 0.17 | **2.69** | **1.00** | **1.02** | **1.14** |
| 11 | 0.49 | 0.64 | 0.98 | **1.06** | 0.60 | 0.77 | 0.90 | 0.88 | **1.14** | **1.03** | **2.32** | **1.04** | **1.26** |
| 12 | 0.96 | 0.74 | 0.93 | 0.93 | 0.81 | 0.72 | 0.80 | **1.20** | 0.27 | **1.12** | 0.91 | **2.06** | **1.02** |
| 13 | 0.24 | 0.44 | 0.26 | 0.82 | 0.71 | 0.33 | **1.73** | **1.10** | 0.65 | 0.91 | **1.20** | 0.97 | **3.85** |

Note: Bold indicates that the actual frequency is higher than the expected frequency.

The empirical analysis that follows is going to shed more light on the reasons for positive assortative matching and other systematic matching patterns on education. The next two subsections present the variables needed to test whether partnership formation happens due to preferences or opportunities or both. First, we define variables that describe the couples' income processes. Second, we present variables that measure the search costs for matching with a specific educational type.

### 2.1.1 Income measures

To assess whether a given portfolio of two educations fulfills the requirement of being a good economic match, we calculate a number of simple income measures based on the time series variation in incomes between different educational groups. We base the income measures on income information sampled over educational groupings instead of using data from observed partnerships. The latter suffers from potential endogeneity bias (see e.g Hess (2004)). Along the lines of Hess (2004), we present three narrowly defined income measures: income correlation, relative volatilities, and mean difference to describe the income processes in relationships within different educational groupings. In addition, we summarize the characteristics of the income processes of two partners given educations by the standardized return.

The income measures are based on residuals from a Mincer wage regression of log net income on experience and experience squared. The residual purges wages for differences stemming from systematic experience differences between educational groups.

Based on the residuals, we compute the correlation between partners' income residuals as a pooled time series correlation. More specifically, for a man in educational group $i$ and a woman in educational group $j$, the correlation is defined as the correlation over time between the mean income residuals of men in educational group $i$ and women in educational group $j$. All income measures are per definition the same for any couple with the same educational combination. Similarly, the income gap is defined as $|\overline{y_i} - \overline{y_j}|$ and the variance gap is defined as $max\left(\frac{\sigma_j}{\sigma_i}, \frac{\sigma_i}{\sigma_j}\right)$, where $\overline{y_i}, \overline{y_j}$ are the mean income residuals for groups $i$ and $j$, and $\sigma_j, \sigma_i$ are the standard deviation of the mean income residuals for groups $i$ and $j$.

The standardized return is computed as the sum of the mean residuals for a couple divided by the standard deviation on the sum of residuals. Due to the similarity with the return to a financial portfolio, we denote the standardized return the 'Sharpe' ratio. I.e. for a man in educational group $i$ and a woman in educational group $j$, the Sharpe ratio is the sum of the mean income residual for groups $i$ and $j$ divided by the standard deviation of the sum of the two mean income residuals. This ratio measures the mean income residual per unit of variability, meaning that this measure indicates how good the partnership between two individuals is in terms of generating a certain income level.

In terms of defining a good portfolio of educations in marriage, individuals should seek to form partnership with individuals who have education that gives a high mean, low variance and a negative correlation. Focusing on the summary measure, individuals should seek partners with a high Sharpe ratio. In Table 4, we report the estimated correlation and Sharpe ratios for the 13*13 educational mating possibilities. Looking across the diagonal elements, we see positive correlations for 11 out of 13 educational groupings. Also, the Sharpe ratios are relatively modest (except for couples where both studied medical science after high school and couples where both have mercantile vocational training). This pattern tentatively suggests that finding a good economic match is not the main determinant for observed partnership formation.

### 2.1.2 Distance measures

To investigate how search costs affect marriage market behavior, we include a variety of distance measures between educational groups in the analysis. The basic idea is that educations that are closer to each other, as measured by the physical distance between

10

TABLE 4: CORRELATION (TOP FIGURE) AND SHARPE RATIOS (BOTTOM FIGURE) FOR PAIRS OF EDUCATIONAL GROUPS

| Males | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **0.851** | 0.740 | 0.683 | 0.679 | 0.344 | -0.557 | 0.753 | 0.663 | -0.515 | -0.219 | 0.661 | -0.767 | -0.278 |
|   | **0.126** | 0.430 | -0.279 | -0.107 | 0.158 | -0.212 | 0.267 | 0.048 | -0.302 | -0.571 | 0.085 | 0.077 | 0.576 |
| 2 | 0.928 | **0.886** | 0.616 | 0.811 | 0.213 | -0.230 | 0.834 | 0.708 | -0.495 | -0.169 | 0.784 | -0.415 | -0.304 |
|   | 0.345 | **0.676** | -.069 | 0.190 | 0.589 | 0.287 | 0.527 | 0.312 | 0.104 | -0.009 | 0.391 | 0.584 | 1.181 |
| 3 | 0.665 | 0.665 | **0.419** | 0.462 | -0.028 | -0.248 | 0.716 | 0.366 | -0.461 | -0.152 | 0.704 | -0.130 | -0.135 |
|   | -0.418 | -0.181 | **-0.858** | -0.937 | -0.981 | -1.281 | -0.363 | -0.649 | -1.163 | -2.151 | -0.690 | -1.030 | -0.763 |
| 4 | 0.973 | 0.935 | 0.652 | **0.850** | 0.257 | -0.351 | 0.882 | 0.756 | -0.480 | -0.095 | 0.742 | -0.460 | -0.214 |
|   | -0.084 | 0.159 | -0.477 | **-0.370** | -0.223 | -0.629 | 0.008 | -0.197 | -0.672 | -0.998 | -0.209 | -0.468 | 0.023 |
| 5 | 0.142 | 0.153 | 0.205 | 0.008 | **0.160** | -0.005 | 0.250 | 0.298 | -0.007 | 0.659 | 0.315 | 0.198 | 0.491 |
|   | -0.148 | 0.050 | -0.462 | -0.383 | **-0.214** | -0.386 | -0.070 | -0.222 | -0.426 | -0.498 | -0.223 | -0.268 | -0.074 |
| 6 | 0.267 | 0.265 | 0.001 | 0.147 | 0.276 | **0.461** | 0.155 | 0.282 | 0.161 | -0.350 | 0.396 | 0.218 | -0.562 |
|   | 0.173 | 0.588 | -0.366 | -0.147 | 0.214 | **-0.135** | 0.382 | 0.068 | -0.227 | -0.853 | 0.119 | 0.047 | 1.021 |
| 7 | -0.135 | -0.069 | 0.148 | -0.070 | -0.348 | -0.079 | **0.048** | -0.023 | -0.103 | 0.733 | -0.077 | 0.336 | 0.737 |
|   | -0.651 | -0.439 | -0.875 | -0.923 | -0.857 | -0.929 | **-0.550** | -0.739 | -0.933 | -0.971 | -0.787 | -0.717 | -0.529 |
| 8 | -0.328 | -0.258 | -0.491 | -0.389 | -0.024 | 0.316 | -0.406 | **-0.081** | 0.637 | 0.491 | -0.370 | 0.245 | 0.354 |
|   | -0.115 | 0.246 | -0.741 | -0.551 | -0.200 | -0.380 | 0.030 | **-0.233** | -0.359 | -0.670 | -0.285 | -0.259 | 0.030 |
| 9 | -0.006 | -0.022 | -0.060 | -0.006 | 0.212 | 0.754 | -0.163 | 0.197 | **0.280** | -0.029 | 0.225 | 0.414 | -0.401 |
|   | 0.160 | 0.577 | -0.375 | -0.164 | 0.160 | -0.126 | 0.380 | 0.045 | **-0.219** | -0.593 | 0.089 | 0.018 | 0.664 |
| 10 | -0.277 | -0.190 | -0.530 | -0.234 | 0.165 | 0.860 | -0.428 | -0.023 | 0.680 | **-0.087** | -0.209 | 0.589 | -0.296 |
|    | -0.315 | 0.042 | -0.974 | -0.839 | -0.474 | -0.542 | -0.230 | -0.454 | -0.538 | **-1.348** | -0.572 | -0.458 | -0.288 |
| 11 | 0.205 | 0.275 | 0.066 | 0.325 | 0.035 | 0.681 | 0.155 | 0.507 | 0.214 | 0.284 | **0.215** | 0.578 | 0.067 |
|    | 0.167 | 0.575 | -0.367 | -0.150 | 0.221 | -0.137 | 0.371 | 0.055 | -0.232 | -0.623 | **0.114** | 0.031 | 0.635 |
| 12 | 0.498 | 0.513 | 0.180 | 0.555 | 0.514 | 0.387 | 0.332 | 0.725 | 0.243 | 0.161 | 0.335 | **0.145** | -0.205 |
|    | 0.047 | 0.419 | -0.467 | -0.298 | -0.001 | -0.317 | 0.218 | -0.069 | -0.356 | -0.960 | -0.058 | **-0.145** | 0.475 |
| 13 | 0.259 | 0.452 | -0.017 | 0.094 | 0.440 | 0.136 | 0.173 | 0.296 | 0.375 | 0.473 | -0.139 | 0.130 | **0.031** |
|    | 0.380 | 0.754 | -0.147 | 0.175 | 0.507 | 0.169 | 0.622 | 0.310 | 0.012 | -0.111 | 0.506 | 0.366 | **0.998** |

Females

the institutions at which the particular education is started, should generate more intra-educational matching since it is cheaper to locate a suitable partner.

Three different measures of distance between the partners are used. The first measure is the minimum geographic distance (*minimum distance*) from e.g. a man's actual place of education to a woman in educational group $j$. The second measure is the density (*density*) of women in educational group $j$ in the man's municipality of education in the year he enrolls in the education. The third is a simple indicator of whether or not the spouses have attained their education at the same educational institution (*same institution*).

In calculating the first two distance measures, we use all educational information available, not only on the individuals in the 2,965 couples, but on all high school graduates in the sample. On the basis of this educational information, we determine the minimum distance that you would be required to travel from any municipality in Denmark to find an individual of educational group $i$, assuming that e.g. individuals from educational group $i$ are only to be found at their municipality of education. For a couple where the man has taken his education in a municipality $m$ and the woman is in educational group $i$, the minimum distance is defined as follows: the distance between municipality $m$ and the nearest municipality where women in educational group $i$ are educated. However, the minimum distance to individuals in educational group 1 is assumed to be zero, which means that all individuals are assumed to be close to individuals from this educational group since high schools are spread out over the entire country.

The density is a measure of the concentration of different educational groups in the municipalities. It measures the density of individuals from a specific educational group in a specific municipality. For a man who has taken his education in municipality $m$ and who enrolled in year $t$, and whose partner belongs to educational group $i$, the density is defined as the proportion of women in municipality $m$ from educational group $i$ in year $t$.

The third distance measure is simply an indicator for whether or not the two partners attended the same educational institution.

Descriptive statistics for the distance measures are presented in Table 5 for the partnership formation analysis, and in Table 7 for the partnership dissolution analysis. The distance measures are proxies for search cost of finding a partner with a given level of education since we do not know if a given partnership is formed because the couple met in school. It could also be the case that the reason why we find substantial homogamy in terms of education is that people with similar types of education share the same workplaces after graduation. After school, the workplace is the most likely place to meet a

12

marriage partner according to both Laumann et al. (1994) and Kalmijn & Flap (2001). We address the latter issue in the following section where we perform a multivariate analysis of partnership formation.

# 3   Partnership formation analysis

Below, we investigate how economic conditions and accessibility of partners affect partnership formation. The question we try to answer is: Do positive assortative matching in education persist when we control for proximity of partners and economic factors?

To answer this, we follow the empirical strategy of Dalmia & Lawrence (2001) and Jepsen & Jepsen (2002). They both use conditional logit models to compare actual couples with randomly created couples to see if actual couples are more similar or more different than the random pairings.

The empirical procedure works as follows: In the first step, the relevant explanatory variables are defined. In this application we include: age difference between partners, an indicator for whether they attend the same education, characteristics of the income residuals for the educational grouping of partners, distance measures between partners' educational institutions, an indicator variable for whether the partner attended the same educational institution, and a density measure of the partners' educational group in the local area. In the second step, the randomly created couples are generated. These are created by randomly assigning an individual from the pool of available partners to a given person (we do not construct same sex couples).[7] The final step is to predict the sign of the coefficients based on the level of positive or negative assortative matching. The conditional logit model is:

$$P(Y_i = j) = \frac{e^{x_{ij}\beta}}{\sum e^{x_{ij}\beta}},$$

where $i$ is an individual, $j$ is an alternative, and $x_{ij}$ is the vector of characteristics of the couple created by matching person $i$ with an alternative $j$. Letting the dependent variable take the value 1 for a natural couple and 0 for an artificial couple, we expect to find a negative coefficient to the age difference, since positive assortative matching means that

---

[7]We only construct one artifical match for each couple. Jepsen & Jepsen (2002) also used 1, but state that sensitivity analysis where 3 artificial partners were constructed did not alter their results. In addition, McFadden (1973) shows that the conditional logit model produces consistent parameter estimates when a random subset of nonchosen alternatives is used.

the age difference should be smaller for actual couples than for artificial couples. Likewise, if a couple is more likely to form if the partners share the same educational institution, the coefficient to *same institution* should be positive. We assume in the following that the explanatory variables are exogenous to the partnership formation process.

In Table 5, we present descriptive statistics for the information used in the matching analysis.

TABLE 5

DESCRIPTIVE STATISTICS FOR VARIABLES USED IN MATCHING ANALYSIS

|  | Real couples | | Constructed couples | |
| --- | --- | --- | --- | --- |
|  | Mean | Std. dev. | Mean | Std. Dev |
| Correlation | 0.2608 | 0.3530 | 0.2313 | 0.3613 |
| Sharpe ratio | -0.0694 | 0.5007 | -0.0657 | 0.4815 |
| Age difference (years) | 2.8169 | 2.5666 | 5.3339 | 4.0127 |
| Income gap | 0.2305 | 0.1922 | 0.2447 | 0.1870 |
| Variance gap | 0.4500 | 0.2443 | 0.4577 | 0.2561 |
| Same institution | 0.2037 | 0.4028 | 0.0314 | 0.1743 |
| Density | 0.1382 | 0.1711 | 0.0961 | 0.1209 |
| Minimum distance (km) | 4.2371 | 14.4948 | 6.7450 | 19.1628 |
| # observations | 2,965 | | 2,965 | |

The first columns show the mean and standard deviation for the explanatory variables for the 2965 couples in the sample. Around 20% of the couples attended the same educational institution. This does not necessarily imply that they meet at the time of education, but it strongly suggests that educational institutions do provide facilities for partnership search.[8] For the *minimum distance* variable it is crucial to know that for men (women) 60% (70%) begin an education in one of the major cities in Denmark. Here all 13 educational groupings are available, and as a consequence, for 85% of the couples this variable equals 0.

In Table 6, we present the results from the conditional logit model for partnership formation.

---

[8]Unfortunately, our register-based data set does not allow any identification of where the couples actually meet.

14

TABLE 6

RESULTS FOR PARTNERSHIP FORMATION ANALYSIS

| | Specification | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Age difference (years) | -0.2621* | -0.2578* | -0.2568* | -0.2622* | -0.2623* | -0.2611* | -0.2327* |
| | (0.0117) | (0.0113) | (0.0112) | (0.0117) | (0.0117) | (0.0122) | (0.0128) |
| | [−0.0542] | [−0.0458] | [−0.0493] | [−0.0541] | [−0.0563] | [−0.0573] | [−0.0522] |
| Same educational group | 1.2678* | | | 1.2691* | 1.2770* | 0.6022* | 0.3941* |
| | (0.0989) | | | (0.0989) | (0.1034) | (0.1202) | (0.1281) |
| | [0.2940] | | | [0.2940] | [0.3010] | [0.1400] | [0.0918] |
| **Income measures** | | | | | | | |
| Correlation | | 0.2485* | | | 0.1849* | 0.1062 | -0.0268 |
| | | (0.0906) | | | (0.0936) | (0.0972) | (0.1097) |
| | | [0.0442] | | | [0.0397] | [0.0233] | [−0.0060] |
| Income gap | | -0.5458* | | | 0.1436 | -0.0257 | 0.1480 |
| | | (0.1808) | | | (0.1955) | (0.2070) | (0.2241) |
| | | [−0.0970] | | | [0.0308] | [−0.0056] | [0.0332] |
| Variance gap | | -0.1816 | | | 0.0275 | -0.1560 | -0.1275 |
| | | (0.1396) | | | (0.1451) | (0.1543) | (0.1690) |
| | | [−0.0323] | | | [0.0059] | [−0.0342] | [−0.0286] |
| Sharpe | | | 0.0479 | 0.0700 | | | |
| | | | (0.0865) | (0.0890) | | | |
| | | | [0.0092] | [0.0145] | | | |
| **Distance measures** | | | | | | | |
| Minimum distance (km) | | | | | | -0.0055* | -0.0084* |
| | | | | | | (0.0022) | (0.0025) |
| | | | | | | [−0.0012] | [−0.0019] |
| Same institution | | | | | | 1.9473* | 1.8987* |
| | | | | | | (0.1545) | (0.1668) |
| | | | | | | [0.4517] | [0.4416] |
| Density | | | | | | 0.8263* | 0.7807* |
| | | | | | | (0.2555) | (0.2759) |
| | | | | | | [0.1812] | [0.1752] |
| Number of couples | 2,965 | | | | | | 2,390 |

Note: For each variable we present the estimated coefficient, the standard error (in parentheses), and the marginal effect [in brackets]. * indicates significance at the 5% level. Specification 7 includes only those couples who did not attend the same high school.

In all specifications, we find that a lower age difference between two individuals increases the probability that they form a partnership. This pattern is as expected and in accordance with the marriage formation literature that finds strong positive assortative mating in age.

In the first specification, we add to the age difference an indicator for whether the partners have the same education. We find that when choosing among two otherwise identical partners, the probability of choosing the one with the same education as ones own is 29 percentage points higher. This conforms with the patterns of positive assortative matching on education, which we saw in the previous subsection and which is well-known in the literature. The estimated marginal effect is unaffected by including income related variables (specifications 2-5). However, it is approximately halved when proximity of partners is controlled for, indicating that half of the positive assortative mating on education is explained by low search costs at educational institutions.[9] The other half is attributed to complementarities in household production.

In the second and third specifications, we include only income related variables. We see that the mean income gap has a significant negative effect, which disappears in specification 4 and 5, whereas the variance gaps do not have a statistically significant coefficient. That is, there is no evidence suggesting that individuals choose partners who have educational attainment that makes their expected income streams negatively correlated. In fact, the opposite seems to be the case. The more correlated the income residuals for the educational groups are, the more likely is it that a match is made, which most likely indicates omitted variable bias. Also, the possibility of forming a partnership with a person with whom it is possible to construct a high yield return corrected for variability does not seem to drive partnership formation either. These findings could suggest that either individuals do not pay attention to these considerations when they form partnerships, or that they simply do not have the sufficient information to judge whether a potential partner offers a good hedge and a high variance-corrected return.

In specification 6, we include variables capturing the proximity of partners. We find that the effect from income correlation becomes insignificant at conventional levels of significance. So the finding that individuals with higher positively correlated income

---

[9]We also estimated a version of the model where we exclude couples that worked at the same workplace the year before marriage, and then the marginal effect is further reduced to 11 percentage points. Hence, a small part of assortative matching on education is due to the fact that people with the same educations are more likely to meet at their workplace than others.

processes are more likely to marry seems to follow from their closer proximity while undertaking education. This conjecture is consistent with the finding that the included proximity variables all have a significant influence on partnership formation and that the effects are working in the expected direction. That is, an individual is more likely to form partnership with a person who, after high-school, attends the same educational institution (e.g. the same university or the same business school). The marginal effect of the indicator variable is rather high, indicating that the probability for a match is raised by 45 percentage points if two persons share educational institution.[10] Also, we find that the higher density of partners with a given education increases the likelihood that an individual forms partnership with a person with this educational attainment, and finally, we also find that the smaller the minimum distance from e.g. a man's educational institution to a given woman's educational institution, the more likely is it that they form a partnership. A ten kilometer extra distance reduces the probability of marriage by 1.2 percentage points.

Around 20% of the couples also attended the same high school. If they already started to date at this time and subsequently coordinated their choices of where to pursue further education, the distance measures might be endogenous to the partnership formation process (at least for those who continued the relationship after they started further education). To address this issue, we exclude all couples who attended the same high school in specification 7. If endogeneity were a major concern, we would expect that the cofficients to the distance variables were biased away from zero. For the reduced sample it turns out that the minimum distance effect actually becomes stronger. Hence, this does not suggest that this variable is upward-biased in specification 6. The density and same institution variables show a somewhat smaller effect in specification 7, but the change is very modest. Consequently, we do not think that the inclusion of individuals who attended the same high school is a major problem, and we rely on specification 6 for the main conclusion of this part of the paper.[11]

In Table A2 in Appendix A, we present results for a conditional logit model with indicator variables for educational cross terms of couples. First, we estimate the conditional logit model with 13*13 cell indicators, and then we test down using a 5% significance

---

[10]Strictly speaking, they do not have to meet each other at the institution since there are no time limitations to when they enrolled and graduated

[11]We also tried to exclude couples who formed the relationship before they entered college (roughly 500 couples). In this case, the main conclusions are unchanged although the variance gap becomes significant.

level. This procedure leaves us with 17 indicators for education cells. The pattern shows positive assortative matching on education with a few exceptions. However, there is some variation in the size of the estimated effects; marginal effects range from 13 to 57 percentage points. As before, we find that including income related variables leaves the coefficients literally unaffected. The only exceptions are couples of females with vocational mercantile education (2) and males with vocational health and crafts education (3) or further education in the social sciences (12), which goes down after including income variables. However, when we control for proximity of partners, we see that more than half of the indicators referring to diagonal cells go down and become insignificant, while for a few of the frequent off-diagonal couples, the marginal effect goes up, e.g. for couples of nurses and medical doctors (7-13), it goes up from 10 to 17 percentage points. The conclusion from the conditional logits with indicators for education cells generally confirms the above. However, it allows us to name some couples who seem to appreciate the same public goods, that is, couples with a vocational mercantile education (2-2), couples with further education in the humanities (10-10), couples with further education in medical sciences (13-13) and nurses and medical doctors (7-13).

All in all, the partnership formation analysis suggests that search costs are indeed important for partnership formation and that individuals apparently search in marriage markets that are close to the place where they attend school. A main conclusion from this matching analysis is that we find a clear pattern of assortative matching on education, and half of that stems from low search costs for partners at the educational institutions. Another important conclusion is that we find no evidence suggesting that economic conditions are important.

# 4    Partnership dissolution analysis[12]

The next step is to analyze what effects the main variables: economic conditions and proximity of partners, have on the duration of relationships. In order to investigate this, we estimate a duration model where the random variable is the time spent in a given relationship. More specifically, we sample all partnerships that are formed during the period of observation and follow them until they dissolve or the sampling period ends. In the latter case, the observations are treated as right censored. This procedure gives a flow

---

[12]Since we focus on both cohabiting relationships and formal marriages, a split-up will henceforth be denoted a dissolution.

sample of partnership. The duration model is specified as a mixed proportional hazard model. That is, it is a product of a function of time spent in the relationship (the baseline hazard), observed time-varying characteristics, $x$, and unobserved characteristics, $v$ :

$$h(t|x_t, v) = \lambda(t) \cdot \varphi(x_t, v), \tag{1}$$

where $\lambda(t)$ is the baseline hazard and $\varphi(x_t, v)$ is the scaling function specified as $exp(x_t'\beta + v)$.

Since we only observe the transitions on the marriage market on a yearly basis, we specify a model for grouped duration data (see e.g. Kiefer (1990)). The marriage duration $T$ is observed to lie in one of $K$ intervals, with the $k$'th interval being $(t_{k-1}; t_k]$ and the convention $t_0 = 0$ for $k = 1, ..., 15$. The probability that the duration $T$ for an individual with explanatory variables $x_t$ is greater than $t_k$ given that the duration is greater than $t_{k-1}$ is given by:

$$\begin{aligned} P(T > t_k | T > t_{k-1}, x_k, v) &= \exp\left[ - \int_{t_{k-1}}^{t_k} h(t|x_t, v)dt \right] \\ &= \exp\left[ - \exp\left[ x_k\beta + v \right] \cdot \Lambda_k \right] \end{aligned} \tag{2}$$

where $\Lambda_{i,k} = \int_{t_{k-1}}^{t_k} \lambda_i(t)dt$. The interval-specific survivor expression (2) is henceforth denoted $\alpha_k$. The probability of observing an exit out of marriage in interval $k$, conditional on survival until $T > t_{k-1}$, is consequently $1 - \alpha_k$. If we do not specify a functional form for the baseline hazard, the $\Lambda_{i,k}$s are just parameters to be estimated.

The individual contribution to the likelihood function is then

$$\mathcal{L} = \int (1 - \alpha_k)^j \alpha_k^{1-j} \prod_{l=1}^{k-1} \alpha_l g(v)dv, \tag{3}$$

where $g(v)$ is the probability density function of the unobservables and where $j = 1$ if the marriage is not right censored and 0 otherwise. Uncompleted durations therefore only contribute with the survivor probabilities. $g(v)$ is assumed to follow a discrete distribution with two points of support.

In addition to information on educational attainment and the associated income measure as well as distance measures, we use a range of other explanatory variables in this part of the analysis.[13] We have three time-varying indicator variables for the presence

---

[13]The choice of explanatory variables is decided partly by what is available in the data set and partly by what is typically used as explanatory variables in the empirical divorce literature (see e.g. Becker et al. (1979) and Svarer (2004)).

of children. These are *first child, second child* and *third+ children.* Since formal marriages in general are more stable, we also distinguish between cohabiting relationships and marriages by the indicator *married.* A variable indicating the order of relationship the individual currently occupies is measured by the *relationship number* variable. This takes the value 1, if it is the first relationship in which the unit of observation is registered. Subsequent relationships with different partners raise this number. We also include variables measuring the age of the partners in the couple and the age difference. The variable, *sickness*, is an indicator variable taking the value 1 if the individual receives sickness benefits for at least 13 weeks during the year. We also distinguish between individuals living in the Copenhagen metropolitan area and individuals living in the provinces by the indicator variable *province.* We include each individual's annual degree of unemployment. This variable is defined as the number of hours of unemployment divided by the number of potential supplied working hours.

In Table 7, we present descriptive statistics for the variables used in the dissolution equation.

TABLE 7

DESCRIPTIVE STATISTICS FOR DISSOLUTION MODEL AT TIME OF PARTNERSHIP FORMATION

| | Mean | Std. Dev. |
|---|---|---|
| Same educational group | 0.2082 | 0.4061 |
| Sharpe | -0.0663 | 0.4988 |
| Income gap | 0.2288 | 0.1904 |
| Variance gap | 0.4503 | 0.2440 |
| Correlation | 0.2648 | 0.3558 |
| Minimum distance (km) | 4.2103 | 14.5468 |
| Density | 0.1395 | 0.1733 |
| Same institution | 0.2017 | 0.4013 |
| Married | 0.0970 | 0.2961 |
| Relationship number | 1.3222 | 0.6143 |
| Living outside Copenhagen | 0.4858 | 0.4999 |
| **Children (measured at final year of observation)** | | |
| First | 0.5007 | 0.5001 |
| Second | 0.2724 | 0.4453 |
| Third + | 0.0466 | 0.2109 |
| **Age** | | |
| Female between 15-20 | 0.6095 | 0.4880 |
| Female between 21-25 | 0.3111 | 0.4630 |
| Female between 26-30 | 0.0687 | 0.2530 |
| Male between 15-20 | 0.4492 | 0.4975 |
| Male between 21-25 | 0.3809 | 0.4857 |
| Male between 26-30 | 0.1395 | 0.3465 |
| Female more than 4 years older | 0.0297 | 0.1698 |
| Male more than 4 years older | 0.1775 | 0.3821 |
| **Sickness and unemployment** | | |
| Sickness, female | 0.0508 | 0.2195 |
| Sickness, male | 0.0404 | 0.1969 |
| Unemployment degree, female | 0.0679 | 0.1742 |
| Unemployment degree, male | 0.0672 | 0.1753 |
| Number of observations | 16,169 | |
| Number of partnerships | 2,896 | |
| Mean duration | 6.5832 | |
| Dissolutions | 0.3809 | |

Note that the income measures are derived from the educational belongings of the couples. They do not represent the realized income during partnerships. In this sense, the income measures are exogenous to the dissolution process.

In Table 8, the results from the dissolution model are presented. The first specification contains the three income measures, the second specification uses the standardized return, and the third excludes the density and same institution variables from specification 1.[14]

---

[14]We have estimated various versions of the model including different explanatory variables and in sequential order. The main results are robust to these different configurations. We therefore only present the model with all explanatory variables included.

TABLE 8

|  | 1 |  | 2 |  | 3 |  |
|---|---|---|---|---|---|---|
|  | Coeff. | S.E. | Coeff. | S.E. | Coeff. | S.E |
| Same educational group | -0.0453 | 0.0977 | -0.1304 | 0.0934 | 0.0143 | 0.0847 |
| Correlation | -0.0110 | 0.0894 |  |  | -0.0079 | 0.0894 |
| Income gap | 0.6236** | 0.1698 |  |  | 0.6201** | 0.1703 |
| Variance gap | -0.1803 | 0.1331 |  |  | -0.1615 | 0.1321 |
| Sharpe |  |  | -0.3744** | 0.0621 |  |  |
| Minimum distance (km) | -0.0404 | 0.0260 | -0.0380 | 0.0254 | -0.0460* | 0.0256 |
| Same institution | 0.1011 | 0.0901 | 0.0625 | 0.0901 |  |  |
| Density | 0.1262 | 0.2001 | 0.1477 | 0.2018 |  |  |
| Married | -1.4408** | 0.1238 | -1.4266** | 0.1238 | -1.4367** | 0.1238 |
| Relationship number | 0.1811** | 0.0648 | 0.1993** | 0.0633 | 0.1839** | 0.0645 |
| Living outside Copenhagen | -0.1904** | 0.0644 | -0.1856** | 0.0646 | -0.1796** | 0.0641 |
| **Children** |  |  |  |  |  |  |
| First | -0.8078** | 0.1337 | -0.8358** | 0.1338 | -0.8221** | 0.1338 |
| Second | 0.4066** | 0.2004 | 0.4543** | 0.2005 | 0.4112** | 0.2002 |
| Third + | 0.3676 | 0.4213 | 0.3862 | 0.4218 | 0.3701 | 0.4218 |
| **Age** |  |  |  |  |  |  |
| Female between 15-20 | 0.0465 | 0.2393 | 0.0777 | 0.2388 | 0.0097 | 0.2391 |
| Female between 21-25 | -0.0697 | 0.2133 | -0.0556 | 0.2126 | -0.0964 | 0.2130 |
| Female between 26-30 | -0.2001 | 0.1970 | -0.2069 | 0.1959 | -0.2236 | 0.1969 |
| Male between 15-20 | 0.6229** | 0.2054 | 0.6579** | 0.2065 | 0.6513** | 0.2055 |
| Male between 21-25 | 0.2070 | 0.1711 | 0.2359 | 0.1715 | 0.2225 | 0.1711 |
| Male between 26-30 | -0.0115 | 0.1476 | -0.0021 | 0.1480 | -0.0071 | 0.1476 |
| Female more than 4 years older | 0.4760** | 0.2014 | 0.5129** | 0.2021 | 0.4490** | 0.2010 |
| Male more than 4 years older | 0.3999** | 0.1063 | 0.4159** | 0.1070 | 0.4022** | 0.1064 |
| **Sickness and unemployment** |  |  |  |  |  |  |
| Sickness, female | -0.1028 | 0.1283 | -0.1065 | 0.1287 | -0.0875 | 0.1284 |
| Sickness, male | -0.1673 | 0.1586 | -0.1724 | 0.1595 | -0.1589 | 0.1586 |
| Unemployment degree, female | 0.0163 | 0.1643 | -0.0263 | 0.1651 | 0.0154 | 0.1644 |
| Unemployment degree, male | 0.5462** | 0.1614 | 0.4832** | 0.1621 | 0.5456** | 0.1617 |
| Number of observations | 16,169 |  | 16,169 |  | 16,169 |  |
| Log-likelihood | -3,377 |  | -3,366 |  | -3378 |  |

Note: * (**) denotes significance at the 10 (5) % level.

The variables of main interest in Table 8 are the economic conditions variables and the proximity between partners' educational institutions. We saw in Table 6 that proximity of educational institutions had significant influence on partnership formation. We attributed this to lower search costs. The finding that individuals are more likely to form partnership with individuals who are more easily accessible could result in lower quality partnerships if individuals were less choosy in the local marriage market. The results presented suggest that this effect is rather weak and that it is dominated by the effect of a high arrival of offers from partners in the local marriage market. In specification 3, we show that the minimum distance variable has a marginal negative effect on the dissolution risk. This suggests that couples who began education after high school at institutions that were far away from each other have a lower risk of dissolution. This effect is only significant at the 8% level though. The other proximity variables are far from significant. In addition,

couples from the same educational groups do not have higher or lower risk of divorce than couples who started at different types of education.[15] In terms of the very pronounced results from the partnership formation analysis it is somewhat surprising that we do not find that these relationships are more stable. The result does not, however, contradict previous studies in the divorce literature. Here, it is typically found that level of education (i.e. the sum of years of education between spouses) is more important than similarity in education among spouses (see e.g. Weiss & Willis (1997), Svarer (2004), and Charles & Stephens (2004)).

The estimates for the economic conditions variables are in accordance with expectations. Couples who have a higher sharpe ratio are more likely to have longer relationships. This suggests that financially better partnerships are more likely to last. In essence, this is consistent with the divorce literature that typically finds that couples who are economically well off have a lower divorce risk (see e.g. Weiss & Willis (1997), Böheim & Ermish (2001), and Svarer (2004)). The coefficient for income correlation is far from significant and does not support the hypothesis that partnerships based on economic hedging are more likely to endure.[16] Together with the observation that larger income gaps are associated with a higher dissolution risk this suggests that it is a high common income level that is good for the stability of relationships.

Assuming that the length of a relationship carries some information on the quality of the relationship, why do individuals then match with individuals who have the same level of education when it does not prolong the duration of the partnership?

We do find that the further away the educational institutions are, the longer the relationships last. This association is rather weak and also not robust to inclusion of additional distance measures (see specification 1 and 2 in Table 8). Also, individuals do not match based on economic return criteria. Although educational groupings that generate a higher expected standardized return (as measured by the Sharpe ratio) are more stable, this does not have any influence on partnership formation. It is hard to

---

[15]In Table A3 in the appendix, we estimate a version of the dissolution model where we include dummy variables for the educational groupings that were significant in the partnership formation analysis. Non of these turn out to have significant coefficients either.

[16]Hess (2004) found that couples with more positively correlated income processes were more likely to split up. We also estimated a version of the model where we use realized income correlations between the two partners. For the complete sample (including individuals with less than high school as well) we find that more correlated income processes are associated with higher dissolution risk. In that sense the data set corroborates Hess (2004).

reconcile these findings with a model that addresses partnership choice as an outcome of utility maximizing behavior by rational agents. In terms of modeling dissolution processes, it is useful to allow for some ex-post heterogeneity in the quality of the partnership. This can be done either along the lines of Becker et al. (1977), who argue that it is deviation between expected and realized utilities that trigger divorce, or as approached in Brien et al. (2006) where it is assumed that match quality is an experienced good. In the latter model, the match quality is a random variable, and the realization observed in the first period is only a noisy signal of what future draws from the distribution of match quality bring. In relation to these theories of divorce, there is nothing in our analysis that suggests that couples that are similar in terms of educational attainment are more or less likely to survive a shock to the relationship or that they are more able to form expectation on the future match quality of their relationship.

In a recent contribution to the divorce literature, Hess (2004) incorporates love into a model of partnership choice and dissolution. He argues that if love and economic conditions are substitutes (and additive separable) in the utility function, then couples who form a bad match, perhaps due to a low Sharpe ratio (formulated in the terms used in the present paper), or have a high positive income correlation might still find it optimal to form a partnership if love is sufficiently high. When time moves along and if love is temporary (which Hess (2004) heroically concludes) then the latter types of couples fall short of value to their partnership, and it becomes optimal to dissolve the partnership. The findings presented in Table 6 and Table 8 are to some extent consistent with Hess's (2004) findings. That is, couples formed between educational groupings that generate a low Sharpe ratio are not less likely to form, but less likely to endure. This suggests that these partnerships have a lot of initial love. Unfortunately for them, love is temporary and the payoff to the relationship decreases with the length of the relationship.[17]

The effects of the other explanatory variables in Table 8 and A3 are in accordance with previous results found in the divorce literature (see Svarer (2004, 2005) for a separate discussion of the these effects).

# 5   Concluding remarks

In this paper, we take a closer look at the observation that individuals tend to match on length and type of education. We investigate whether the systematic relationship between

---

[17]The same line of argumentation can be made for the income gap variable.

the educations of the partners is explained by opportunities, e.g. low search frictions, or preferences, e. g. complementarities in household production or portfolio optimization. We find that around half of the systematic sorting on education is due to low search frictions in marriage markets of the educational institutions. The other half is attributed to complementarities in household production since income properties of joint income process show no influence on partner selection.

When it comes to explaining dissolution risk we do not find that couples who share the same education are more or less likely to split up than couples who have different types or levels of education. We do, however, find that individuals who form partnership with individuals from an educational group that generates a high standardized economic return are more likely to have a long-lived relationship. Interestingly, these types of partnership are not more likely to form than couples generated by random matching.

For future research it could be fruitful to develop a theoretical justification of the empirical results presented in this paper. Hess (2004) provides a model that is consistent with some of the findings we present, but it does not address the issue of proximity of partners which constitutes one of the main mechanisms for partnership formation.

# References

[1] BECKER, G (1973). "A Theory of Marriage: Part I", *Journal of Political Economy*, 82(4), 813-846.

[2] BECKER, G., E.M. LANDES AND R. T. MICHAEL (1977). "An Economic Analysis of Marital Instability", *Journal of Political Economy*, 85(6), 1141-1187.

[3] BLAU, P. M. AND O. D. DUNCAN (1967). **The American Occupational Structure**. John Wiley & Sons.

[4] BRIEN, M, L. LILLARD AND S. STERN (2006) "Cohabitation, Marriage, and Divorce in a Model of Match Quality", *International Economic Review* 47(2). 451-494.

[5] BÖHEIM R. AND J. ERMISCH (2001). "Partnership Dissolution in the UK - the Role of Economic Circumstances" *Oxford Bulletin of Economics and Statistics* 63(2): 197-208.

[6] CHARLES, K.K. & M. STEPHENS, JR. (2004). "Job Displacement, Disability, and Divorce", *Journal of Labor Economics*, 22(2), 489-521.

[7] CHEN, K-P., S-H. CHIANG AND S.F. LEUNG (2003). "Migration, Family, and Risk Diversification", *Journal of Labor Economics*, 21(2), 353-379.

[8] CHIAPPORI, P.A., M. IYIGUN & Y. WEISS (2006). "Investment in Schooling and the Marriage Market", Manuscript.

[9] DALMIA, S. AND P.G. LAWRENCE (2001). "An Empirical Analysis of Assortative Mating in India and the U.S.", *International Advances of Economic Research*, 7 (4), 443-458.

[10] EPSTEIN, E. AND R. GUTTMAN (1984). "Mate Selection in Man: Evidence, Theory and Outcome", *Social Biology*, 31(4), 243-278.

[11] FERNANDEZ, R, N. GUNER AND J. KNOWLES (2005). "Love or Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality", *Quarterly Journal of Economics*, 120(1), 273-244.

[12] FISMAN, R., S. S. IYENGAR, E. KAMENICA, I. SIMONSON.(2006). "Gender Differences in Mate Selection: Evidence from a Speed-Dating Experiment " *Quarterly Journal of Economics*, 121(2), 673-697.

[13] GAUTIER, P.A., M. SVARER AND C.N. TEULINGS. (2005). "Marriage and the City", Working Paper 2005-01, Department of Economics, University of Aarhus.

[14] GOLDIN, C. (1992). "The Meaning of College in the Lives of American Women: The Past One-Hundred Years", NBER Working Paper No. 4099.

[15] HESS, G. (2004). "Marriage and Consumption Insurance: What's Love Got to Do with it?", *Journal of Political Economy*, 112(2), 290-318.

[16] JEPSEN, L.K. AND C.A. JEPSEN (2002). "An Empirical Analysis of the Matching Patterns of Same-Sex and Opposite-Sex Couples", *Demography*, 39(3), 435-453.

[17] KALMIJN, M. AND H. FLAP (2001). "Assortative Meeting and Mating: Unintended Consequences of Organized Settings for Partner Choices", *Social Forces*, 79(4), 1289-1312.

[18] KOTLIKOF, L. AND A. SPIVAK (1981). "The Family as an Incomplete Annuity Market", *Journal of Political Economy*, 89, 372-391.

[19] KIEFER, N.M. (1990). "Econometric Methods for Grouped Duration Data". In: **Panel Data and Labour Market Studies**, North-Holland.

[20] LAUMANN, E., J. GAGON, R. MICHAEL AND S. MICHAELS (1994). *The Social Organization of Sexuality: Sexual Practices in the US.* Chicago: University of Chicago Press.

[21] LEWIS, S.K. AND V.K. OPPENHEIMER (2001). "Educational Assortative Mating across Marriage Markets: Non-Hispanic Whites in the US", *Demography*, 37(1), 29-40.

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

[23] MCFADDEN. D. (1973). "Conditional Logit Analysis of Qualitative Choice Behavior" in *Frontiers in Economics*, edited by P. Zaremba, New York: Academic Press.

[24] MICEVSKA, M. (2002). "Marriage, Uncertainty and Risk Aversion in Russia", Draft, Center for Development Research, University of Bonn.

[25] NAKOSTEEN, R.A. AND M.A. ZIMMER (2001). "Spouse Selection and Earnings: Evidence of Marital Sorting", *Economic Inquiry*, 39(2), 201-213.

[26] ROSE, E. (2001). "Marriage and Assortative Mating: How Have the Patterns Changed?", Draft, University of Washington.

[27] ROSENZWEIG, M. AND O. STARK (1989). "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India", *Journal of Political Economy*, 97, 905-926.

[28] SCHAFER, R.B. AND P.M KEITH (1990). "Matching by Weigth in Married Couples: A Life Cycle Perspective", *Journal of Social Psychology*, 130(5), 657-664.

[29] SCOTT, J.F. (1965), "The American College Sorority: The Role in Class and Ethnic Endogamy", *American Sociological Review* 30(3): 514-527.

[30] SVARER, M. (2004). "Is Your Love in Vain? Another Look at Premarital Cohabitation and Divorce", *Journal of Human Resources,* 39(2), 523-536.

[31] SVARER, M. (2005). "Two Tests of Divorce Behaviour on Danish Marriage Market Data", *Nationaløkonomisk Tidsskrift (Danish Economic Journal),* 143(3), 416-432.

[32] WEISS, Y. (1997). "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce", **Handbook of Population and Family Economics**, Elsevier Science B.V.

[33] WEISS, Y. AND R.J. WILLIS (1997). "Match Quality, New Information, and Marital Dissolution", *Journal of Labor Economics*, 15(1), S293-S329.

## Appendix

THE 20 EDUCATIONAL COMBINATIONS WITH THE HIGHEST RATIO OF ACTUAL TO EXPECTED
FREQUENCY

| Rank | Educational group - females | Educational group - males | Ratio |
|---|---|---|---|
| 1 | 5 | 5 | 6.69 |
| 2 | 6 | 6 | 4.53 |
| 3 | 13 | 13 | 3.85 |
| 4 | 9 | 9 | 3.61 |
| 5 | 2 | 2 | 3.06 |
| 6 | 7 | 7 | 2.98 |
| 7 | 8 | 8 | 2.80 |
| 8 | 4 | 4 | 2.72 |
| 9 | 10 | 10 | 2.69 |
| 10 | 1 | 1 | 2.49 |
| 11 | 8 | 10 | 2.37 |
| 12 | 11 | 11 | 2.32 |
| 13 | 7 | 8 | 2.28 |
| 14 | 5 | 8 | 2.13 |
| 15 | 12 | 12 | 2.06 |
| 16 | 3 | 1 | 1.96 |
| 17 | 1 | 3 | 1.92 |
| 18 | 5 | 6 | 1.85 |
| 19 | 7 | 13 | 1.73 |
| 20 | 3 | 7 | 1.73 |

## TABLE A2: RESULTS FROM CONDITIONAL LOGIT MODEL

| Educational groups (female-male) | Specification | | | |
|---|---|---|---|---|
| | 1' | 4' | 5' | 6' |
| 1-1 | 0.9727* | 0.9733* | 0.8789* | 0.5433 |
| | (0.3925) | (0.3925) | (0.3970) | (0.4147) |
| | [ 0.2306] | [0.2308 ] | [0.2099] | [0.1287] |
| 1-3 | 0.8073 | 0.8071 | 0.7634 | 0.7205 |
| | (0.4410) | (0.4414) | (0.4470) | (0.4517) |
| | [0.1895] | [0.1895] | [0.1812] | [ 0.1729] |
| 2-1 | 0.7582 | 0.7650 | 0.6848 | 0.4149 |
| | (0.4021) | (0.4031) | (0.4045) | (0.4413) |
| | [0.1771] | [0.1789] | [0.1615] | [0.0971] |
| 2-2 | 1.5770* | 1.5841* | 1.4704* | 0.8927* |
| | (0.3899) | (0.3913) | (0.4004) | (0.4284) |
| | 0.3744 | 0.3760 | 0.3515 | 0.2156 |
| 2-3 | 0.8523 | 0.8564 | 0.8096 | 0.8335 |
| | (0.4384) | (0.4388) | (0.4551) | (0.4515) |
| | 0.2007 | 0.2018 | 0.1927 | 0.2011 |
| 2-9 | 0.7224* | 0.7323* | 0.7649* | 0.6817* |
| | (0.3004) | (0.3037) | (0.3019) | (0.3066) |
| | 0.1681 | 0.1706 | 0.1813 | 0.1630 |
| 2-12 | 0.3537 | 0.3617 | 0.3454 | 0.3142 |
| | (0.2623) | (0.2647) | (0.2670) | (0.2716) |
| | 0.0787 | 0.0806 | 0.0784 | 0.0726 |
| 4-4 | 0.6525 | 0.6500 | 0.5386 | 0.2071 |
| | (0.3517) | (0.3520) | (0.3620) | (0.3960) |
| | [0.1509] | [0.1503] | [0.1253] | [0.0472] |
| 6-6 | 2.7795* | 2.7778* | 2.756* | 1.7490* |
| | (0.4310) | (0.4308) | (0.4318) | (0.4897) |
| | [0.5789] | [0.5785] | [0.5696] | [0.4098] |
| 7-7 | 1.7334* | 1.7245* | 1.7480* | 1.0480 |
| | (0.6953) | (0.6961) | (0.7090) | (0.8193) |
| | [0.4081] | [0.4063] | [0.4106] | [0.2540] |
| 7-13 | 0.8438* | 0.8533* | 0.8482* | 1.1236* |
| | (0.3126) | (0.3154) | (0.3185) | (0.3256) |
| | [0.1983] | [0.2008] | [0.2021] | [0.2721] |
| 9-9 | 1.2104* | 1.2092* | 1.1955* | 0.0697 |
| | (0.3891) | (0.3891) | (0.3908) | (0.4723) |
| | 0.2891 | 0.2889 | 0.2876 | 0.0156 |
| 10-8 | 0.9911 | 0.9907 | 0.9564 | 1.3991 |
| | (0.6863) | (0.6859) | (0.6922) | (0.7197) |
| | 0.2354 | 0.2354 | 0.2294 | 0.3362 |
| 10-10 | 1.3093* | 1.2946* | 1.3558* | 0.6405 |
| | (0.3207) | (0.3271) | (0.3315) | (0.3827) |
| | [0.3126] | [ 0.3092] | [0.3253] | [0.1527] |
| 11-11 | 0.9765* | 0.9800* | 0.9991* | 0.1726 |
| | (0.2104) | (0.2109) | (0.2133) | (0.2463) |
| | [0.2308] | [0.2317] | [0.2391] | [0.0391] |
| 12-12 | 1.0503* | 1.0494* | 1.0855* | 0.5720* |
| | (0.2094) | (0.2095) | (0.2152) | (0.2406) |
| | [0.2490] | [0.2488] | [0.2603] | [0.1354] |
| 13-13 | 1.8759* | 1.8915* | 1.8638* | 1.7259* |
| | (0.4588) | (0.4642) | (0.46289) | (0.5271) |
| | [0.4369] | [0.4400] | [0.4334] | [0.4049] |

Note: For each variable we present the estimated coefficient, the standard error (in parentheses), and the marginal effect [in brackets].

* indicates significance at the 5 % level

TABLE A2:RESULTS FROM CONDITIONAL LOGIT MODEL. CONTINUED

| | Specification | | | |
|---|---|---|---|---|
| | 1' | 4' | 5' | 6' |
| Age difference (years) | -0.2663* | -0.2662* | -0.2664* | -0.2659* |
| | (0.0119) | (0.0119) | (0.0119) | (0.0124) |
| | [−0.0557] | [−0.0557] | [−0.0573] | [−0.0587] |
| | | | | |
| **Income measures** | | | | |
| Correlation | | | 0.1574 | 0.0699 |
| | | | (0.1002) | (0.1041) |
| | | | [0.0338] | [0.0154] |
| Income gap | | | -0.0240 | -0.0738 |
| | | | (0.2213) | (0.2308) |
| | | | [−0.0052] | [−0.0163] |
| Variance gap | | | 0.0870 | -0.1058 |
| | | | (0.1503) | (0.1602) |
| | | | [0.0187] | [−0.0234] |
| Sharpe | | -0.0216 | | |
| | | (0.0956) | | |
| | | [−0.0045] | | |
| **Distance measures** | | | | |
| Minimum distance (km) | | | | -0.0062* |
| | | | | (0.0023) |
| | | | | [−0.0014] |
| Same institution | | | | 2.0330* |
| | | | | (0.1583) |
| | | | | [0.4682] |
| Density | | | | 0.6914* |
| | | | | (0.2695) |
| | | | | [0.1527] |
| Number of "real" couples | 2,965 | 2,965 | 2,965 | 2,965 |

Note: For each variable we present the estimated coefficient, the standard error (in parentheses), and the marginal effect [in brackets].

* indicates significance at the 5 % level.

| | 1 | |
|---|---|---|
| | Coeff. | S.E. |
| 1-1 | -0.6784 | 0.3492 |
| 1-3 | 0.0065 | 0.3305 |
| 2-1 | -0.2970 | 0.3940 |
| 2-2 | -0.4269 | 0.2786 |
| 2-3 | 0.3269 | 0.4326 |
| 2-9 | -0.3142 | 0.3711 |
| 2-12 | 0.1759 | 0.2825 |
| 4-4 | -0.1074 | 0.2820 |
| 6-6 | 0.1480 | 0.2207 |
| 7-7 | 0.0794 | 0.4474 |
| 7-13 | -0.3068 | 0.3786 |
| 9-9 | -0.0663 | 0.3047 |
| 10-8 | 0.7158 | 0.5392 |
| 10-10 | 0.0228 | 0.2438 |
| 11-11 | -0.1302 | 0.1946 |
| 12-12 | -0.0352 | 0.1655 |
| 13-13 | 0.1260 | 0.3790 |
| Sharpe | -0.3195* | 0.0737 |
| Minimum distance (km) | -0.0376 | 0.0257 |
| Same institution | 0.0364 | 0.0907 |
| Density | 0.0597 | 0.2254 |
| Married | -1.4368* | 0.1248 |
| Relationship number | 0.2027* | 0.0642 |
| Living outside Copenhagen | -0.1800* | 0.0653 |
| **Children** | | |
| First | -0.8596* | 0.1356 |
| Second | 0.4709* | 0.2014 |
| Third + | 0.4211 | 0.4227 |
| **Age** | | |
| Female between 15-20 | 0.1131 | 0.2397 |
| Female between 21-25 | -0.0267 | 0.2130 |
| Female between 26-30 | -0.1732 | 0.1972 |
| Male between 15-20 | 0.6606* | 0.2074 |
| Male between 21-25 | 0.2327 | 0.1723 |
| Male between 26-30 | -0.0088 | 0.1487 |
| Female more than 4 years older | 0.4954* | 0.2033 |
| Male more than 4 years older | 0.4067* | 0.1070 |
| **Sickness and unemployment** | | |
| Sickness, female | -0.0976 | 0.1291 |
| Sickness, male | -0.1639 | 0.1602 |
| Unemployment degree, female | -0.0247 | 0.1659 |
| Unemployment degree, male | 0.4524* | 0.1640 |
| Number of observations | 16,169 | |
| Log-likelihood | 3,359 | |

Note: * denotes significance at the 5 % level

**Economics Working Paper**

2005-20    Philipp Festerling: Cartel Prosecution and Leniency Programs: Corporate versus individual Leniency.

2005-21    Knud Jørgen Munk: Tax-tariff reform with costs of tax administration.

2005-22    Knud Jørgen Munk and Bo Sandemann Rasmussen: On the Determinants of Optimal Border Taxes for a Small Open Economy.

2005-23    Knud Jørgen Munk: Assessment of the Introduction of Road Pricing Using a Computable General Equilibrium Model.

2006-1     Niels Haldrup and Andreu Sansó: A Note on the Vogelsang Test for Additive Outliers.

2006-2     Charlotte Christiansen, Juanna Schröter Joensen and Helena Skyt Nielsen: The Risk-Return Trade-Off in Human Capital Investment.

2006-3     Gunnar Bårdsen and Niels Haldrup: A Gaussian IV estimator of cointegrating relations.

2006-4     Svend Hylleberg: Seasonal Adjustment.

2006-5     Kristin J. Kleinjans and Jinkook Lee: The link between individual expectations and savings: Do nursing home expectations matter?

2006-6     Jakob Roland Munch, Michael Rosholm and Michael Svarer: Home Ownership, Job Duration, and Wages.

2006-7     Francesco Busato and Enrico Marchetti: Skills, sunspots and cycles.

2006-8     Peter Sandholt Jensen and Allan H. Würtz: On determining the importance of a regressor with small and undersized samples.

2006-9     Michael Svarer: Working Late: Do Workplace Sex Ratios Affect Partnership Formation and Dissolution?

2006-10    Helena Skyt Nielsen and Michael Svarer: Educational Homogamy: Preferences or Opportunities?