Marital Matching and Earnings

Evidence from the Unmarried Population in Sweden

Robert A. Nakosteen
Olle Westerlund
Michael A. Zimmer

ABSTRACT

Social scientists have devoted substantial research to economic basis for matching of men and women in marriage. A common feature of existing studies is their reliance on samples of married couples. The principal shortcoming of spouse data is that spouses' earnings correlations are contaminated by the partners' behaviors and other events that occur after marriage and affect their earnings. This study addresses that problem by exploiting a longitudinal data file from the Swedish population. By selecting a sample of married couples in a given year, we retreat through the file to years before the marriage. Using data from the spouses' single years, we apply the correlation methodology to their earnings. Evidence from the model supports positive assortative mating.

I. Introduction

It has been 30 years since Becker (1974) first focused on economic dimensions of spouse selection. Although previous research had emphasized the extent of marital sorting based upon physical and social characteristics, Becker argued that marital sorting systematically matches high-earning males with females specializing in nonmarket activities, particularly home production and child rearing.

Empirical testing of Becker's hypothesis has been problematic, primarily due to data deficiencies. Several studies have estimated earnings regressions from samples of

Robert A. Nakosteen is an associate professor in the School of Management, University of Massachusetts in Amherst. Olle Westerlund is an associate Professor in the Economics Department, University of Umeå, Sweden. Michael A. Zimmer is Bristol Myers Squibb Professor of Economics, University of Evansville, Indiana. *We wish to thank two anonymous referees for helpful comments. An earlier version of this paper was presented at the 2002 conference of the Missouri Valley Economic Association. The data used in this article are proprietary and derived from Swedish tax and labor market records. The authors will provide information regarding obtaining the data from Statistics Sweden and The Swedish Labour Market Board.*
[Submitted June 2001; accepted May 2003]
ISSN 022-166XE-ISSN 1548-8004

married couples and obtained correlations between the ensuing earnings residuals. Negative correlation provides evidence that, after controlling for measured factors such as age, education, and work experience, there is negative marital sorting; positive correlation indicates positive sorting. There has been mixed evidence, based upon this approach. Smith (1979), Becker (1981), and Lam (1988) report results supporting positive sorting, while Zimmer (1996) reports negative correlations for whites and positive correlations for blacks. These studies rely on data that report spouses' earnings after marriage, rather than before. That makes it impossible to disentangle the effects of marital sorting from those of the apparent earnings premium of married relative to single males. Previous studies of the marriage premium suggest that high-earning males possess unmeasured traits that are associated with marriage (see, for example, Cornwell and Rupert 1997). Other research, however, suggests a reversal in the chain of cause and effect: marriage affects labor force attachments and productivity of males, inducing them to earn higher remuneration (Korenmann and Neumark 1991; Akerlof 1998). In any case, all studies to date are based on data from existing marriages. They do not address the phenomenon of spousal matching among single persons.

Nakosteen and Zimmer (2001) provide a partial solution to this problem. Using the Panel Study of Income Dynamics (PSID), they record individual earnings before and immediately after marriage, but are able to observe spouses' earnings only after marriage. This arises from the way in which the PSID is constructed, following individuals and families, and reporting spousal characteristics only as spouses join the sample after marrying someone who was previously in the panel. It is assumed that earnings immediately after marriage have not yet been significantly affected by the decision to marry. Immediate post-marital earnings are used as a proxy for premarital earnings. Using this approach, Nakosteen and Zimmer find evidence of positive marital sorting.

The ideal data for testing the Becker hypothesis would allow observation of both spouses in years preceding marriage. The panel data files with which we are familiar are not structured in a way that allows this approach. Once an individual is selected for the panel, his/her earnings are recorded for each year. The spouse of the originally sampled individual will have her/his earnings recorded only after the couple becomes married.

Although the authors are not aware of U.S. data that address this problem, we have access to longitudinal data that comprehensively report Swedish labor market behavior. Because these data cover the entire Swedish labor force for a number of years, it is possible to identify the year in which a couple becomes married and observe labor market outcomes of both individuals in the years *preceding* marriage. Using these data, we develop regression models of earnings that link the records of married couples during their single years before marriage. We then estimate pair wise correlations between spouses' earnings that are not contaminated by the effects of the marriage premium or with external factors that affect both spouses' earnings after marriage.

II. Marital Matching

It is well known that individuals tend to marry persons similar to themselves in physical and social attributes. To some extent, this might result from

routine social interactions: people encounter prospective mates in their schools, churches, and neighborhoods. Economic models of the marriage market move beyond intersections of individual choice sets, ascribing to individuals the intention to maximize the lifetime value of their marriage. According to Becker (1981), individuals of superior productivity tend to marry one another and are compensated for their higher productivity. As a consequence, outcomes in the marriage market tend to produce a matching of traits. In Becker's framework, spouses reinforce one another in traits that are complementary within the family and offset one another in traits that are substitutes.

Other authors have emphasized social outcomes that result from marital matching. Although it is commonly believed that the most important outcome has to do with income distribution, Kremer (1997) argues that sorting has a somewhat greater effect on the persistence of economic status across generations than on inequality. Chadwick and Solon (2002) contribute evidence to that effect. Using a sample of married couples from the PSID, they find a substantial elasticity between incomes of daughters (wives) and the incomes of families in which they were raised. They then report that earnings of the daughters' *husbands* have nearly the same elasticity with respect to the daughters' parents. Chadwick and Solon interpret that result to mean that assortative mating plays an important role in the transmission and persistence of economic status across generations.

Becker (1981) suggested that optimizing behavior in the marriage market would pair high-wage males with low-wage females, since the time of wives at home would naturally complement that of husbands at work. Lam (1988) pointed out that many gains from marriage result from the family's joint consumption of goods purchased in the market. If those gains offset the gains from respective specialization of husbands and wives in market employment and the home, then high-earning spouses tend to match with one another. Resolution of these conflicting hypotheses is a useful empirical issue.

This study addresses marital matching on the basis of measured earnings as well as earnings traits that are not directly measured. To a certain extent it is natural to observe nontrivial correlation in spouses' earnings. Individuals marry persons similar to themselves in education and age. Schooling in turn correlates closely with earnings, and spouses accumulate work experience as they age together. In addition, they share lifestyles and some health outcomes, and they experience the consequences of similar regional labor markets. As a matter of course, then, their earnings are correlated in a manner that does not bear directly on choice-based optimizing behavior at the time their marriages were formed. Recognizing this, most studies attempt to capture unmeasured factors that determine earnings. This approach is used by Van der Klauww (1996), who studies marriage decisions of females. The model assumes that unmarried women confront a marriage decision each period. At the time of her choice, a woman knows her potential husband's actual current earnings as well as information that she uses to forecast his expected future earnings. In Van der Klauww's model, the forecast is generated from a regression equation based on current data.

In a similar spirit, most studies measure marital matching in a multivariate context. Becker (1981) estimates partial correlations of spouses' wages that control for age and schooling, obtaining 0.32 for white couples and 0.24 for blacks. Smith (1979) uses residuals from spouses' wage regressions, where controls include schooling,

experience and region. He obtains residual correlations of 0.098 for whites and 0.035 for blacks.[1] Zimmer (1996) replicates the wage-residual approach on a sample of low-income couples. He obtains correlations of –0.12 for whites, 0.15 for blacks and –0.31 for Hispanics.

A variation of the wage-residual approach uses information from spouses' family backgrounds. Behrman, Birdsall and Deolalikar (1995), using panel data from a sample of marriages in India, exploit family data in a two-step approach. First, they estimate three regression models in which the wife's education, age, and dowry, respectively, depend on characteristics of her husband and his parents. They obtain a set of residuals from each of the three regressions and interpret them as unmeasured components of the *husband's* human capital. Each residual is then entered as a distinct explanatory variable in regressions explaining the husband's labor force participation, hourly wage, and wage growth. Behrman et al. interpret the results to mean unobserved characteristics of the husband, proxied by the first-stage residuals, exert important effects on his (post-marriage) wage and wage growth. The authors argue that those traits, which are observed by potential wives and their parents, command economic value in the marriage market.

Another approach that attempts to construct proxy measures for spouses' attributes is found in papers by Lam and Schoeni (1993, 1994). Using data for married couples from the United States and Brazil, they estimate models of hourly wages for husbands. In addition to standard human capital variables, the models include several measures of family background. Of particular interest are years of schooling of the wife's parents. The estimates indicate a significant relation between husbands' earnings and education of the father-in-law. Lam and Schoeni interpret the result to mean that some unmeasured traits of husbands are proxied by observed characteristics of their wives' fathers, reflecting an underlying process of marital matching.

As we noted in Section I, the principal contribution of this paper is to exploit data that exist prior to marriage. The findings of Behrman et al. (1995) and Lam and Schoeni (1993, 1994) reveal a potential shortcoming of this approach, at least in the context of our data set. Their results indicate that the family background of an individual's prospective spouse is informative with respect to marital matching. Extending a suggestion by Freiden (1974), if parental backgrounds increase the nonmarket productivity of the spouse who specializes in household production or the market productivity of the spouse who earns wage income, then they increase the gains from marriage. Thus, parents' traits are potentially important in the process of marital matching. The Swedish data files contain information on individuals but not their parents. Thus our approach, which has the advantage of revealing pre-marriage data for both spouses, does not address the role of parental and family backgrounds.

III. Data and Sample Construction

The data derive mainly from official registers administered by Statistics Sweden and registers held by the Labor Market Board in Sweden. The

1. When Smith corrects the estimates for nonparticipation of some wives, he obtains correlations equal to 0.035 for whites and 0.216 for blacks.

Statistics Sweden data are from a longitudinal micro data file containing information for the entire Swedish population of working age, encompassing individuals 18–64 years of age. For the purpose of this study, we use data that originate from the population registration system, managed by the National Tax Board. The data include earnings, marital status, and other family characteristics such as place of residence and origin country in the case of immigrants.

We matched information from the Statistics Sweden file with data from a second longitudinal file maintained by the Swedish Labor Market Board. In order to receive benefits, the unemployed must register at the Labor Market Board's Public Employment Service as job seekers. Virtually the entire labor force is covered by unemployment insurance, or is eligible for basic unemployment benefits if they are not insured. Data from the combined files provide a complete record for each individual.

We have access to data for the years 1992 through 1996. The sample consists of individuals who were 26 to 34 years of age as of 1996 and became married or cohabitants in either 1994 or 1995. Thus the lower end of the age interval in the first year of analysis (1992) is 22 years. This provides two distinct cohorts of married couples, whom we observe before and after the year of marriage. For the 1994 cohort, we observe one year in the single state (1992) and two years in the married state (1995, 1996). For the 1995 cohort, we observe single years 1992 and 1993 and married year 1996. Access to the post-marriage years allows us to compare estimates before and after marriage for both cohorts.

Cohabitation is widely practiced in Sweden, and is not only viewed sympathetically, but is given a social status nearly equivalent to marriage. In recognition of this, we include cohabitants in our target population. Official data on unmarried cohabitants include, with minor exceptions, only couples that have a child in common. This means that individuals observed as being single in 1995 and married or cohabitants in 1996 may in fact already have been cohabitants in 1995 or earlier, but did not report that relationship. In order to minimize measurement error, we do not include the year immediately prior to recorded marriage or cohabitation as the final year of being unmarried. Instead, we drop the final unmarried year entirely, and include the observation two years earlier than the recorded marriage as the final pre-marriage year. In order to diminish the possibility that couples could have been cohabitating two or more years prior to having a child, we report results for a subsample of couples that, two years prior to marriage or cohabitation, resided in different parishes, which number 2,552 in Sweden. For example, for a couple recorded as first married or cohabitating in 1996, the couple will be included in the restricted subsample only if they lived in different parishes in all years from 1992 to 1994. In the empirical analysis, we relax the "different parish" rule and perform other sensitivity tests to examine the robustness of the estimates.

We obtained earnings data from mandatory income statements reported by employers to the tax authorities. As the dependent variable, we use the logarithm of gross annual income from employment in hundreds of kronor (SEK) where 1 SEK is equivalent to $0.12 at current rates. Earnings for all years are deflated using the Swedish Consumer Price Index and are expressed in constant 1992 SEK. Use of annual earnings rather than hourly wages is necessary because the data do not include information on hours or weeks worked. Generally, compensation for work exceeding 100

SEK must be reported. Compensation includes holiday payment and sick pay paid by the employer, as well as parental leave benefits, which are based on prior earnings. According to law, employers compensate the first 14 days of a period of illness, after which public health insurance provides coverage. The health insurance system was restructured on several occasions during the 1990s. Since 1993, employees receive no compensation for the first day of illness, 75 percent of income for days two and three and 90 percent from the fourth to the ninetieth day. The replacement ratio then declines to 80 percent for the remaining days in a spell of illness.

One issue that arises in connection with the earnings estimations concerns exclusion of nonworkers from the sample. Because we do not observe earnings for individuals who do not work, they must necessarily be excluded in the regression estimations. This raises questions about possible sample selection bias in the remaining sample. One approach is to use Heckman's (1979) procedure to correct the model for potential selection bias. We did not pursue that option, because of two features in the data. First, the high rate of labor force participation by females in Sweden means that the large majority of couples are working during their adult single years, and sample selection bias is not likely to be problematic. For males, the annual sample proportions of workers range between 0.89 and 0.92; for females the range is 0.85 to 0.90. Because our methodology relies on residual *pairs*, in the earnings regressions we restrict the sample to cases in which both individuals worked during their premarriage years. Although that necessitates a further restriction of the sample, again the proportions of dual workers are very high, ranging from 0.78 to 0.83. Second, if we use the Heckman procedure we must find variables in the data that identify workers from nonworkers but do not explain workers' earnings. After perusing the data, we judged that suitable instruments are not available. Consequently, due to the high proportion of workers and an apparent lack of ideal instruments for job status, we chose to estimate the earnings models based on samples of workers and with no adjustments for sample selection.

In spite of the high proportion of workers, sample selection bias might remain a problem. One way to include observations for nonworkers is to exploit the panel nature of the sample and impute those individuals' earnings for years in which earnings are not observed.[2] In the empirical analysis below, we extend the sample by adding person-year observations with imputed earnings. For individuals who were missing earnings during one or more years, we imputed the average of earnings for adjacent years in which they did report earnings. We excluded a small number of couples for whom one spouse or both failed to report earnings for any year from 1992 to 1996.[3]

2. We are indebted to an anonymous referee for this suggestion. The proportion of individuals with imputed earnings does not exceed 6 percent in any year.

3. We imputed incomes if they were available on the same side of the time line relative to the marriage event. For example, in the 1995 marriage cohort, if an individual had zero income in 1993, we imputed only if he/she reported income in 1992 or 1994. On the other hand, if there was no reported income before marriage but income was available for 1996, we did not use 1996 income for imputation. Where appropriate and available, we imputed the average of two years. For example, if the individual reported zero income in 1993, we imputed the average of the adjacent years 1992 and 1994. If only one of those years was available, we used only that year.

IV. Econometric Framework

In order to establish the statistical framework, define y_{mit} as the annual log-earnings of male *i* during year *t*, which by construction of the sample occurs prior to his marriage. His earnings are hypothesized to depend on a vector of explanatory variables x_{mit} describing his human capital and other determinants of earnings:

$$(1) \quad y_{mit} = \beta_m' x_{mit} + \varepsilon_{mit}$$

where β_m is a conformable vector of unknown coefficient parameters and the error term ε possesses a normal distribution with zero mean and constant variance denoted σ_m^2.

In a similar fashion, we specify log earnings for unmarried female *i* during year *t*:

$$(2) \quad y_{fit} = \beta_f' x_{fit} + \varepsilon_{fit}$$

where again the error term is normally distributed with zero mean, in this case with variance σ_f^2. The critical informational content in the data concerns what happens subsequent to year *t*. As described in Section III, by ascertaining eventual marriages we form synthetic couples within the single population, based on a forward look at individuals who later married one another. The available years include $t = 1992$, 1993, and 1996 for couples who married in 1995; and $t = 1992$, 1995, and 1996 for couples who married in 1994. For each unmarried "couple," we estimate Equations 1 and 2 jointly as a seemingly unrelated regressions (SUR) model. In addition to estimates of the coefficient parameters, β_m and β_f, and the disturbance variances σ_m^2 and σ_f^2, the SUR procedure retrieves an estimate of the covariance between the error terms, σ_{mf}. Using this information, we estimate the correlation between ε_m and ε_f:

$$(3) \quad \rho_{mf}(t) = \sigma_{mf}(t) / \sigma_m(t)\, \sigma_f(t),$$

where *t* denotes 1992 and 1993 for couples who married in 1995, and 1992 for couples who married in 1994. For purposes of comparison, we also estimate ρ_{mf} for post-marriage years 1996 (1995 marriages) and 1995 and 1996 (1994 marriages). The latter estimates approximate what has appeared previously in the literature.

The error terms in Equations 1 and 2 capture factors that contribute to measured earnings but are not observed in the data. Our approach assumes that the errors are orthogonal to the included regressors. The implication is that the unmeasured traits cannot be inferred from individuals' observable characteristics. Their importance for spouse selection arises from the fact that many traits not revealed in the data are likely to be visible in the pool of potential marriage partners. Individuals might exhibit good health, whether physical or emotional, or they might exude exceptional confidence or ambition. They might possess favorable attitudes toward work or family connections that portend future success. These traits, which tend to signal strong prospects for individuals' future earnings, might also make them attractive as potential spouses. Even their endowments of physical attractiveness, which are not measured in most data sets but are seen by potential mates, might signal favorable future earnings (Hamermesh and Biddle 1994).

Under the null hypothesis maintaining absence of assortative mating, individuals' unobserved traits would not show significant correlation. Positive correlation, on the other hand, is evidence that individuals with favorable unmeasured traits tend to select

mates whose traits are likewise positive, while negative correlation supports matching of spouse pairs in which persons with positive latent attributes select mates whose traits are below the regression-corrected mean.

V. Results of Estimations

Earnings equations were estimated for six cross sections, as detailed in the preceding sections. The equations contain explanatory variables measuring age, education, region of residence, region of origin for immigrants, disability status, and the regional rate of unemployment. Education is measured by a series of four dummy variables for categories ranging from "less than secondary education" to "university degree or higher." Region of residence is represented by a set of seven dummy variables corresponding to official county groupings used by the Swedish census. Immigrants' origins are represented by two dummy variables that distinguish Nordic immigrants from those originating in other countries.

The sampling plan described in the previous section gives rise to a total of 24 regressions, based on three years for each of the two marriage cohorts and four variations of the sample for each year. For the sake of brevity, we do not present the regression results here.[4] The results are generally consistent with earnings regressions based on other cross section data sets. Earnings increase significantly with education, with particularly strong returns to postsecondary schooling. Regional differences are somewhat pronounced, with a positive advantage for the Stockholm region. Earnings increase significantly with age, presumably a proxy for potential work experience, and immigrants tend to earn significantly less than native Swedes. Consistent with many studies of U.S. data, the R square statistics indicate that the models explain between 23 and 28 percent of measured variation in log earnings.

The earnings regressions are used to estimate residual correlations within spouse pairs, as described in Section IV. The resulting estimates are summarized in Table 1, which encompasses all of the 24 SUR models. The six columns identify individual years for each marriage cohort: 1992, 1993 and 1996 for couples who married in 1995, and 1992, 1995 and 1996 for couples who married in 1994. The four rows correspond to variations of the sample. The full sample is first modified by removing couples who resided in the same parish before marriage, then by removing couples in which one or both spouses reported zero earnings and were assigned imputed earnings (see Section III), and finally by removing both same-parish and imputed-earnings couples.

The entries in each cell are the sample size for the year and sample combination, the within-couple correlation of log earnings, and the within-couple correlation of earnings residuals obtained from the SUR model. Our empirical strategy does not rely on the longitudinal nature of the original data set, aside from the extent to which it was useful in ascertaining eventual marriages. Since we did not construct a panel, there is variation in sample sizes within each row, owing principally to missing data or attrition from the married population.

4. A sample set of regression results, along with a table of variable definitions and descriptive statistics, is available from the authors on request.

Table 1
Estimated Correlation of log Earnings and Earnings Residuals

	Married 1994			Married 1995		
	Pre-Marriage 1992	Post-Marriage 1995	Post-Marriage 1996	Pre-Marriage 1992	Pre-Marriage 1993	Post-Marriage 1996
Full sample	n = 19,920	n = 19,131	n = 18,482	n = 19,930	n = 19,986	n = 19,388
	0.203	0.193	0.177	0.187	0.147	0.152
	0.174	0.087	0.061	0.091	0.037	0.055
Full sample minus same parish	n = 5,104	n = 5,504	n = 5,354	n = 5,689	n = 5,809	n = 5,590
	0.162	0.182	0.184	0.171	0.165	0.173
	0.147	0.079	0.075	0.085	0.044	0.061
Full sample minus imputed earnings	n = 17,105	n = 17,235	n = 17,926	n = 19,144	n = 19,448	n = 18,665
	0.199	0.176	0.159	0.177	0.136	0.132
	0.163	0.080	0.052	0.080	0.034	0.042
Full sample minus same parish and imputed earnings	n = 4,826	n = 4,804	n = 5,128	n = 5,378	n = 5,611	n = 5,314
	0.162	0.167	0.174	0.170	0.156	0.157
	0.139	0.082	0.070	0.082	0.043	0.052

*The second entry is the correlation between spouses' log earnings. The third entry is the correlation between residuals from SUR estimation of Equations 1 and 2. All estimates are significantly greater than zero at the 0.01 level of significance or lower.

For the cohort of 1995 marriages, the set of estimates extends from three years before marriage to one year after marriage. For the cohort of 1994 marriages, estimates are for two years prior to marriage and two years after marriage. Estimates of the raw earnings correlations are within a range of 0.13 to 0.20, and in all cases are significantly greater than zero. The estimates appear to be reasonably robust with respect to the four variations of the sample, indicating in particular that evidence of positive earnings correlation before marriage is not an artifact of sample composition based on imputed earnings or same-parish residence.

Inclusion of years before and after marriage permits inspection of the correlations relative to the timing of marriages. In the majority of cases, they appear to diminish modestly after marriage. Since previous studies have necessarily been based on samples of existing marriages, the indication here is that correlations estimated in those studies are present to at least an equivalent extent before the marriages are formed.

As we noted in Section I, the process of mate selection guarantees some degree of earnings correlation. Because individuals tend to choose spouses similar to themselves in age and education, among other attributes, and because age and education are correlated with earnings, there is an extent to which earnings correlation occurs naturally as a consequence of spouse matching based on schooling and age. On the latter point, the present sample is not an exception. Estimated correlations in the age of spouses are 0.391 and 0.376 for the complete sample of 1994 and 1995 marriages, respectively. Schooling attainments are also positively correlated. Because the schooling variable is categorical in ascending order of attainment, we use a rank measure of association for ordered contingency tables (Kendall and Stuart 1961) to arrive at respective estimates of 0.308 and 0.327 for 1994 and 1995 marriages.

These correlations, while informative in their own right, invite a further question: Do earnings correlations persist in a multivariate context that controls for measured spouse traits? That is the purpose of the residual correlations, which are presented as the third entry in each cell of Table 1. All of the estimates are significant at the 0.01 level or lower. The residual correlations are necessarily smaller than the raw earnings correlations, since they reflect the portions of spouses' earnings that remain after adjusting for measured characteristics. They range in relative magnitude from 25 to 90 percent of the raw earnings correlations.

In every case, the residual correlations decline in magnitude after marriage, similar to the pattern in the raw earnings correlations. This appears to be consistent with the phenomenon of specialization in labor supply after marriage. Spouses who matched originally based in part on earnings attributes, both measured and unmeasured, might have a tendency to assume specialized household roles after marriage in a manner that attenuates the correlation between their earnings. It is worthwhile to ask whether the attenuation continues for an extended time after marriage, induced by family events such as the birth of children. The range of data available to us does not permit estimation for additional years after marriage. Consequently, we have only a limited number of estimated correlations following marriage. The evidence that we do have suggests that within-couple correlations reported in previous research appear to have manifested themselves before marriage to an extent that is at least as great as that indicated by common post-marriage estimates. This suggests that marital matching is stronger than what appears to be the case if we look only at data from existing marriages. Although this conclusion is tempered somewhat by the

brief span of years in our sampling frame, it suggests the potential value of additional research.

VI. Summary

Married couples tend to possess similar demographic and economic characteristics. Evidence in this study, which addresses a fundamental weakness in previous research, indicates that they tend to be economically similar before marriage, at least in the dimension of earnings. Their measured earnings are correlated before marriage, which suggests positive marital matching: there is a tendency for high earners to marry one another. Their earnings residuals are also correlated, to an extent that is nontrivial relative to the correlation in earnings. In general, residual correlations prior to marriage are larger in magnitude than the post-marriage correlations. The residual correlation is informative if we interpret residuals, as other researchers have done, to convey statistical content about individuals' unobserved earnings traits. Some individual characteristics cannot be observed in conventional data sets, yet prospective spouses seem to incorporate them, to a limited but nonetheless significant extent, into the matching process.

References

Akerlof, George. 1998. "Men without Children." *Economic Journal* 108(March):287–309.

Behrman, Jere, Nancy Birdsall, and Anil Deolalikar. 1995. "Marriage Markets, Labor Markets, and Unobserved Human Capital: An Empirical Exploration of South-Central India." *Economic Development and Cultural Change* 43 (April):585–601.

Becker, Gary. 1974. "A Theory of Marriage." In *Economics of the Family*, ed. T.W. Schultz, , 299–344. Chicago: University of Chicago Press.

———. 1981. *A Treatise on the Family*. Cambridge, Mass.: Harvard University Press.

Chadwick, Laura, and Gary Solon. 2002. "Intergenerational Income Mobility Among Daughters." *American Economic Review* 92 (March): 335–44.

Cornwell, Christopher, and Peter Rupert. 1997. "Unobservable Individual Effects, Marriage and the Earnings of Young Men." *Economic Inquiry* 35 (April): 285–94.

Freiden, Alan. 1974. "The United States Marriage Market." *Journal of Political Economy* 82 (March/April): S34–S56.

Hamermesh, Daniel, and Jeff Biddle. 1994. "Beauty and the Labor Market." *American Economic Review* 84:1174–94.

Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (January):153–61.

Kendall, Maurice, and Alan Stuart. 1961: *The Advanced Theory of Statistics, Vol. 2*. New York: Hafner Publishing Company.

Korenmann, Sanders, and David Neumark. 1991. "Does Marriage Really Make Men More Productive?" *Journal of Human Resources* 26(2):282–307.

Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality?" *Quarterly Journal of Economics* 112 (February):115–39.

Lam, David. 1988. "Marriage Markets and Assortative Mating with Household Public Goods: Theoretical Results and Empirical Implications." *Journal of Human Resources* 23(4):426–87.

Lam, David, and Robert Schoeni. 1994. "Family Ties and Labor Markets in the United States and Brazil." *Journal of Human Resources* 29(4):1235–58.

———. 1993. "Effects of Family Background and Returns to Schooling: Evidence from Brazil." *Journal of Political Economy* 101 (August):710–40.

Nakosteen, Robert, and Michael Zimmer. (2001) " Spouse Selection and Economic Traits: Evidence of Marital Sorting and Earnings Capacity." *Economic Inquiry* 39 (April):201–13.

Smith, James. 1979. "The Distribution of Family Earnings." *Journal of Political Economy 87* (October): S163–192.

Van der Klauww, Wilbert. 1996 "Female Labour Supply and Marital Status Decisions: A Life Cycle Model." *Review of Economic Studies* 63 (April):199–235.

Zimmer, Michael. 1996 "Assortative Mating and Ethnicity in the Low Wage Population: An Examination of Spouses' Earnings." *Applied Economics Letters* 3(May):311–15.