

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Economics of the Family: Marriage, Children, and Human Capital

Volume Author/Editor: Theodore W. Schultz, ed.

Volume Publisher: UMI

Volume ISBN: 0-226-74085-4

Volume URL: <http://www.nber.org/books/schu74-1>

Publication Date: 1974

Chapter Title: A Theory of Marriage

Chapter Author: Gary S. Becker

Chapter URL: <http://www.nber.org/chapters/c2970>

Chapter pages in book: (p. 299 - 351)

A Theory of Marriage

Gary S. Becker

University of Chicago and National Bureau of Economic Research

I

1. Introduction

In recent years, economists have used economic theory more boldly to explain behavior outside the monetary market sector, and increasing numbers of noneconomists have been following their examples. As a result, racial discrimination, fertility, politics, crime, education, statistical decision making, adversary situations, labor-force participation, the uses of "leisure" time, and other behavior are much better understood. Indeed, economic theory may well be on its way to providing a unified framework for *all* behavior involving scarce resources, nonmarket as well as market, nonmonetary as well as monetary, small group as well as competitive.

Yet, one type of behavior has been almost completely ignored by economists,¹ although scarce resources are used and it has been followed in some form by practically all adults in every recorded society. I refer to marriage. Marital patterns have major implications for, among other things, the number of births and population growth, labor-force participation of women, inequality in income, ability, and other characteristics among families, genetical natural selection of different characteristics

I benefited from the discussion of several earlier drafts of this paper at the Workshop in Applications in Economics of the University of Chicago and in seminars at the National Bureau of Economic Research, Northwestern University, and the Population Council. Very helpful suggestions were received from William Brock, Isaac Ehrlich, Alan Freiden, H. Gregg Lewis, Robert T. Michael, Marc Nerlove, Richard Posner, George J. Stigler, T. W. Schultz, and two referees. Michael Keeley provided valuable research assistance. Research was supported by a grant from the Ford Foundation to the National Bureau of Economic Research for the study of the economics of population. This paper is not an official NBER publication since it has not been reviewed by the NBER Board of Directors.

¹ To the best of my knowledge, the only exception prior to my own work is an unpublished paper by Gronau (1970a). His paper helped stimulate my interest in the subject.

over time, and the allocation of leisure and other household resources. Therefore, the neglect of marriage by economists is either a major oversight or persuasive evidence of the limited scope of economic analysis.

In this essay, it is argued that marriage is no exception and can be successfully analyzed within the framework provided by modern economics. If correct, this is compelling additional evidence on the unifying power of economic analysis.

Two simple principles form the heart of the analysis. The first is that, since marriage is practically always voluntary, either by the persons marrying or their parents, the theory of preferences can be readily applied, and persons marrying (or their parents) can be assumed to expect to raise their utility level above what it would be were they to remain single. The second is that, since many men and women compete as they seek mates, a *market* in marriages can be presumed to exist. Each person tries to find the best mate, subject to the restrictions imposed by market conditions.

These two principles easily explain why most adults are married and why sorting of mates by wealth, education, and other characteristics is similar under apparently quite different conditions. Yet marital patterns differ among societies and change over time in a variety of ways that challenge any single theory. In some societies divorce is relatively common, in others, virtually impossible, and in Western countries it has grown rapidly during the last half-century. Some societies adjust to legal difficulties in receiving divorces by delaying marriage, whereas others adjust by developing more flexible "consensual," "common-law," or "trial" marriages. In many the bride brings a dowry, in others the groom pays a bride-price, and in still others couples marry for "love" and disdain any financial bargaining. In some the newly married usually set up their own household, in others they live with one set of parents.

I do not pretend to have developed the analysis sufficiently to explain all the similarities and differences in marital patterns across cultures or over time. But the "economic" approach does quite well, certainly far better than any available alternative.² It is hoped that the present essay will stimulate others to carry the analysis into these uncharted areas.

Section 2 of Part I considers the determinants of the gain from marriage compared to remaining single for one man and one woman. The gain is shown to be related to the "compatibility" or "complementarity" of their time, goods, and other inputs used in household production.

Section 3 of Part I considers how a group of men and women sort themselves by market and nonmarket characteristics. Positive assortive mating—a positive correlation between the values of the traits of husbands and wives—is generally optimal, one main exception being the sorting by the earn-

² Some of the best work has been done by Goode (1963), but there is no systematic theory in any of his fine work.

ing power of men and women, where a negative correlation is indicated. Empirically, positive assortive mating is the most common and applies to IQ, education, height, attractiveness, skin color, ethnic origin, and other characteristics.

Section 4 of Part I considers how the the total output of a household gets divided between the husband and wife. The division is not usually fixed, say at 50–50, or determined mechanically, but changes as the supply of and demand for different kinds of mates changes.

Part II develops various extensions and modifications of the relatively simple analysis in this part. “Caring” is defined, and some of its effects on optimal sorting and the gain from marriage are treated. The factors determining the incidence of polygamous marital arrangements are considered. The assumption that the characteristics of potential mates are known with certainty is dropped, and the resulting “search” for mates, delays in marriage, trial marriage, and divorce are analyzed. Divorce and the duration of marriage are also related to specific investments made during marriage in the form of children, attachments, and other ways. Also briefly explored are the implications of different marital patterns for fertility, genetical natural selection, and the inequality in family incomes and home environments.

2. *The Gain from Marriage*

This section considers two persons, M and F , who must decide whether to marry each other or remain single. For the present, “marriage” simply means that they share the same household. I assume that marriage occurs if, and only if, both of them are made better off—that is, increase their utility.³

Following recent developments in the theory of household behavior, I assume that utility depends directly not on the goods and services purchased in the market place, but on the commodities produced “by” each household.⁴ They are produced partly with market goods and services and partly with the own time of different household members. Most important for present purposes, commodities are not marketable or transferable among households, although they may be transferable among members of the same household.

Household-produced commodities are numerous and include the quality of meals, the quality and quantity of children, prestige, recreation, companionship, love, and health status. Consequently, they cannot be

³ More precisely, if they *expect* to increase their utility, since the latter is not known with certainty. Part II discusses some consequences of this uncertainty, especially for the time spent searching for an appropriate mate and the incidence of divorce and other marital separations.

⁴ An exposition of this approach is given in Michael and Becker (1973).

identified with consumption or output as usually measured: they cover a much broader range of human activities and aims. I assume, however, that all commodities can be combined into a single aggregate, denoted by Z . A sufficient condition to justify aggregation with fixed weights is that all commodities have constant returns to scale, use factors in the same proportion, and are affected in the same way by productivity-augmenting variables, such as education. Then different commodities could be converted into their equivalent in terms of any single commodity by using the fixed relative commodity prices as weights.⁵ These weights would be independent of the scale of commodity outputs, the prices of goods and the time of different members, and the level of productivity.

Maximizing utility thus becomes equivalent for each person to maximizing the amount of Z that he or she receives. Moreover, my concentration on the output and distribution of Z does not presuppose transferable utilities, the same preference function for different members of the same household, or other special assumptions about preferences.

Each household has a production function that relates its total output of Z to different inputs:

$$Z = f(x_1, \dots, x_m; t_1, \dots, t_k; E), \quad (1)$$

where the x_i are various market goods and services, the t_j are the time inputs of different household members, and E represents "environmental" variables. The budget constraint for the x_i can be written as:

$$\sum^m p_i x_i = \sum^k w_j l_j + v, \quad (2)$$

where w_j is the wage rate of the j th member, l_j the time he spends working in the market sector, and v property income. The l_j and t_j are related by the basic time constraint

$$l_j + t_j = T \quad \text{all } j, \quad (3)$$

where T is the total time of each member. By substituting equation (3) into (2), the goods and time constraints can be combined into a single "full" income constraint:

$$\sum^m p_i x_i + \sum^k w_j t_j = \sum^k w_j T + v = S, \quad (4)$$

where S stands for full income, the maximum money income achievable, if the w_j are constants.

I assume that a reduction in the household's total output of Z makes

⁵ One serious limitation of these assumptions is that they exclude the output of commodities from entering the production functions of other commodities. With such "joint production," the relative price of a commodity would depend partly on the outputs of other commodities (Grossman 1971). Joint production can result in complementarity in consumption, and thereby affect the gain from marriage and the sorting of mates. See the brief discussion which follows in section 3.

no member better off and some worse off.⁶ Consequently, each member would be willing to cooperate in the allocation of his time and goods to help maximize the total output of Z . Necessary conditions to maximize Z include

$$\frac{MP_{t_i} \equiv (\partial Z / \partial t_i)}{MP_{t_j} \equiv (\partial Z / \partial t_j)} = \frac{w_i}{w_j}, \quad \text{for all } 0 < t < T. \quad (5)$$

If the household time of the k th member = T , then

$$\frac{MP_{t_k}}{MP_{t_j}} = \frac{\mu_k}{w_j}, \quad (6)$$

where $\mu_k \geq w_k$ is the "shadow" price of the time of k . Also

$$\frac{MP_{x_i}}{MP_{t_j}} = \frac{p_i}{w_j} \quad \text{for all } x_i > 0 \text{ and } 0 < t_j < T. \quad (7)$$

Each member must cooperate and allocate his time between the market and nonmarket sectors in the appropriate proportions.

If M and F are married, their household is assumed to contain only the two time inputs t_m and t_f ; for simplicity, the time of children and others living in the same household is ignored. As long as they remain married, $T_m = T_f = 24$ hours per day, 168 hours per week, and so forth, and conditions (5) to (7) determine the allocation of the time of M and F between the market and nonmarket sectors. More time would be allocated to the market sector by M than by F (less to the nonmarket sector) if $w_m > w_f$ and if $MP_{t_f} \geq MP_{t_m}$ when $t_f = t_m$. Indeed, F would specialize in the nonmarket sector ($t_f = 0$) if either w_m/w_f or MP_{t_f}/MP_{t_m} were sufficiently large.

A singles household is taken to be exactly the same as a married one except that $T_f = 0$ when M is single and $T_m = 0$ when F is single. A singles household allocates only its own time between the market and nonmarket sectors to satisfy equation (7). Single persons generally allocate their time differently than married persons because the former do not have time and goods supplied by a mate. These differences depend partly on the elasticities of substitution among the x_i , t_f , and t_m , and partly on the differences between the market wage rates w_m and w_f . For example, single F are more likely to "work" more than married F and single M less than married M , the greater the percentage excess of w_m over w_f . Empirically, single women clearly "work" more than married women and single men less than married men.⁷

If Z_{m0} and Z_{0f} represent the maximum outputs of single M and F , and m_{mf} and f_{mf} their incomes when married, a necessary condition for

⁶ This assumption is modified in the following section and in Part II.

⁷ See, e.g., *Employment Status and Work Experience* (U.S., Bureau of the Census 1963c), tables 4 and 12.

M and F to marry is that

$$\begin{aligned} m_{mf} &\geq Z_{m0} \\ f_{mf} &\geq Z_{0f}. \end{aligned} \quad (8)$$

If $m_{mf} + f_{mf}$, the total income produced by the marriage, is identified with the output of the marriage,⁸ a necessary condition for marriage is then that

$$m_{mf} + f_{mf} \equiv Z_{mf} \geq Z_{m0} + Z_{0f}. \quad (9)$$

Since most men and women over age 20 are married in all societies, equation (9) must generally hold because of fundamental reasons that are not unique to time or place. I have a useful framework for discovering these reasons.

The obvious explanation for marriages between men and women lies in the desire to raise own children and the physical and emotional attraction between sexes. Nothing distinguishes married households more from singles households or from those with several members of the same sex than the presence, even indirectly, of children. Sexual gratification, cleaning, feeding, and other services can be purchased, but not *own* children:⁹ both the man and woman are required to produce their own children and perhaps to raise them. The physical and emotional involvement called "love" is also primarily between persons of the opposite sex. Moreover, persons in love can reduce the cost of frequent contact and of resource transfers¹⁰ between each other by sharing the same household.

Economies of scale may be secured by joining households, but two or more males or females could equally well take advantage of these economies and do so when they share an apartment and cooking. Consequently, the explanation of why men and women live together must go beyond economies of scale.

The importance of own children and love implies that, even with constant returns to scale, M (now standing for a man) and F (now standing for a woman) gain from marriage because t_m and t_f are not perfect substitutes for each other or for goods and services supplied by market firms or households. When substitution is imperfect, single persons cannot produce small-scale equivalents of the optimal combination of inputs achieved by married couples.

Consequently, the "shadow" price of an hour of t_f to a single M —the price he would be willing to pay for t_f —would exceed w_f , and the "shadow" price of t_m to a single F —the price she would be willing to pay

⁸ Income and output can differ, however, because some output may be jointly consumed. See the discussion in the following section and in Part II.

⁹ The market in adoptions is used primarily by couples experiencing difficulties in having their own children and by couples paid to raise other persons' children.

¹⁰ The relation between love and such transfers is discussed in Part II.

for t_m —would exceed w_m . Both gain from marriage because M then, in effect, can buy an hour of t_f at w_f and F can buy an hour of t_m at w_m , lower prices they then would be willing to pay. Of course, this is also why married households use positive amounts of t_f and t_m .

My explanation of the gain from marriage focuses on the complementarity between M and F . The gain from complementary can be illustrated in much-exaggerated measure by assuming that the production function relating Z to t_m , t_f , and x has the Cobb-Douglas form

$$Z = kx^a t_m^b t_f^c. \quad (10)$$

Clearly, $Z_{m0} = Z_{0f} = 0$ since both t_m and t_f are needed to produce Z ($Z = 0$ if t_m or $t_f = 0$), whereas Z_{mf} can take any value. Other functions have less extreme “complementarity” and permit positive production when some inputs are absent but less “efficiently” than when all are present.

Some sociological literature also suggests that complementarity between men and women is the major source of the gain from marriage (Winch 1958, 1967; Goode 1963), but the meaning of “complementarity” is left rather vague and ill defined. By building on the substantial economic literature that analyzes complementarity and substitution in production, I have shown how “complementarity” determines the gain from marriage.

Can this analysis also explain why one man is typically married to one woman, rather than one man to several women, several men to one woman, or several men to several women? The importance of own children is sufficient to explain why marriages of several men to one or several women are uncommon since it would be difficult to identify the father of a child if many men had access to the same woman, whereas the identity of the mother is always known. The marriage of several women to one man does not suffer from this defect, and, indeed, such marriages have been more common. However, if the sex ratio equalled about unity, each household having several women and one man would have to be balanced by households having only men. If I assume that all men and all women are identical, and if I make the rather plausible assumption of “diminishing returns” from adding persons to a household having one man and one woman, the total output from say two single male households and one household with three women and one man would be smaller than the total output from three households each having one man and one woman.¹¹ Consequently, monogamous unions—one man married to one woman—predominate because it is the most efficient marital form.

¹¹ For example, assume that singles households have an output of 5 units of Z , one man and one woman 13 units, one man and two women 20 units, and one man and three women 26 units. Three households each with one man and one woman would produce 39 units, whereas two single male households and one household having three women and one man would produce only 36 units.

Polygamy is encouraged when the sex ratio is significantly different from unity and when men or women differ greatly in wealth, ability, or other attributes.¹²

My definition of marriage in terms of whether a man and a woman share the same household differs from the legal definition because my definition includes persons in "consensual" and casual unions and excludes legally married persons who are separated. However, my analysis does have useful implications about the choice between legally recognized and other unions (Kogut 1972), as well as about the decisions to remain married, divorce, remarry legally, remarry "consensually," remain single, and so forth, that must be made in the course of a lifetime (see Part II).

The gain from marriage has to be balanced against the costs, including legal fees and the cost of searching for a mate, to determine whether marriage is worthwhile. The larger the gain is relative to costs, the larger the net gain from marriage; presumably, therefore, the larger too is the fraction of persons who marry. I now consider the more important determinants of this net gain.

The gain is greater the more complementary are the inputs: the time of spouses and market goods. Since I have argued that these inputs are complementary in good part because of the desire to raise own children, the gain would be positively related to the importance of children. Hence, persons desiring relatively few or low-"quality" children either marry later, end their marriages earlier, or do both.¹³

The gain from marriage also depends on market opportunities. The effect of a change in opportunities can be analyzed most easily by equating the maximum output of any household to its full income deflated by the average cost of producing a unit of output. For example, with constant returns to scale, the output of a married household with both members participating in the labor force can be written as

$$Z_{mf} = \frac{\text{full income}}{\text{average cost of production}} \equiv \frac{S_{mf}}{C_{mf}(w_m, w_f, p)} \equiv \frac{S_m + S_f}{C_{mf}}, \quad (11)$$

where C_{mf} depends on the wage rates of t_m and t_f and the price of x .¹⁴ The output of a singles household can be written in the same form except that only one price of time enters the average cost functions C_m and C_f .¹⁵

What is the effect of an increase in income on the incentive to marry? If only the property incomes of M and F , v_m and v_f , rose exogenously

¹² See the more extensive discussion of polygamy in Part II.

¹³ A further discussion can be found in Keeley (1974).

¹⁴ Duality theory shows that C is the dual of the production function.

¹⁵ Or, alternatively, the shadow price of F to M enters C_m , and the shadow price of M to F enters C_f .

by the same percentage, and if $v_m/S_m = v_f/S_f$, then S_m , S_f , and S_{mf} would all rise by the same percentage. With constant returns to scale, Z_{m0} , Z_{0f} , and Z_{mf} , and thus the absolute gain from marriage, would also rise by the same percentage as full income since neither C_{mf} , C_m , nor C_f would be affected by the rise in property incomes, as long as both M and F continue to participate in the labor force,¹⁶ and assuming that property income is unaffected by the allocation of time.¹⁷ Since a rise in property income should not greatly affect the cost of getting married, the incentive to marry would also rise.

The effect of a rise in wage rates alone¹⁸ on the incentive to marry is less clear-cut. A rise in the wage rates of M and F by the same percentage would increase outputs by smaller percentages than full incomes, even with constant returns to scale, because costs of production also rise.¹⁹ Moreover, the cost of getting married rises to the extent that the own time of M and F enters into search and other marital costs. Consequently, the effect on the net gain from marriage is not clear a priori and depends on the relative importance of own time in marriage costs and in the production of output in single and married households.

Consequently, my analysis predicts that a rise in property income, necessarily, and a rise in wage rates, possibly, increase the incentive to marry. This implication runs counter to the popular opinion that poor persons marry earlier and divorce less than rich persons but is consistent with the empirical evidence. In the United States, at least, the probability of separation and divorce is negatively related to income (U.S., Bureau of the Census 1971). Keeley (1974) finds too that when years of schooling and a few other variables are held constant, higher-wage persons appear to marry earlier than others.

My analysis implies that a rise in w_f relative to w_m , F 's wage rate relative to M 's, with the productivity of time in the nonmarket sector held constant, would decrease the gain from marriage if w_f were less than w_m : the gain from substituting M 's time in the market for F 's time (and F 's time in the household for M 's time) is greater the lower w_f is relative to w_m . As a proof, consider an increase in w_f "compensated" by a sufficient decrease in w_m to maintain constant the combined output of the two singles households. The increase in w_f would not increase married output as

¹⁶ Even if married F did not participate in the labor force, the percentage rise in Z_{mf} would still equal the share of property income in full income (see section 2, Part I of the Appendix).

¹⁷ The gain from marriage would increase even more if the income from nonhuman capital, i.e., property income, was positively related to the time allocated to "portfolio management" (see the discussion in the following section).

¹⁸ By alone is meant in particular that the productivity of time in household production or marital search is unchanged.

¹⁹ The percentage rise in output equals the percentage rise in wage rates multiplied by the ratio of total earnings to full income. Although this relation holds whether or not married F is in the labor force (see section 2, Part I of the Appendix), the ratio of total earnings to full income can depend—positively or negatively—on her participation.

much as the decrease in w_m would decrease it if married F worked sufficiently fewer hours in the market sector than single F , and married M worked at least as much as single M . Since married women do work much less than single women and married men work more than single men, an increase in the wage rate of women relative to men would decrease the incentive to marry.²⁰ As supporting evidence, note that American states that have higher wage rates of women relative to men also have smaller fractions of men and women who are married (Santos 1970; Freiden 1972).

The gain from marriage also depends on traits, such as beauty, intelligence, and education, that affect nonmarket productivity as well, perhaps, as market opportunities. The analysis of sorting in section 3b implies that an increase in the value of traits that have a positive effect on nonmarket productivity, market productivity held constant, would generally increase the gain from marriage. Presumably this helps explain why, for example, less attractive or less intelligent persons are less likely to marry than are more attractive or more intelligent persons.²¹

3. The Marriage Market and Sorting of Mates

a) Optimal Sorting

I now consider not one M and F who must decide whether to marry or remain single, but many M 's and F 's who must decide whom to marry among numerous potential candidates, as well as whether to marry. If there are n M 's and n F 's (unequal numbers of M and F are discussed in section 4), each is assumed to know all the relevant²² entries in an $(n + 1) \times (n + 1)$ payoff matrix showing the maximum household commodity output that can be produced by any combination of M and F :

	F_1	\dots	F_n	
M_1	Z_{11}	\dots	Z_{1n}	Z_{10}
		Z_{ij}		
M_n	Z_{n1}	\dots	Z_{nn}	Z_{n0}
	Z_{01}	\dots	Z_{0n}	

(12)

The last row and column give the output of single M and F . Each person has $n + 1$ possibilities and the $2n$ persons together have $n^2 + 2n$ pos-

²⁰ A fortiori, if married women were not in the labor force, a compensated increase in their wage rate would decrease the incentive to marry since an increase in their wage rate would not affect married output, whereas a decrease in the male wage rate would decrease output. This footnote as well as the text assumes that compensated changes in w_f and w_m do not much affect the cost of getting married.

²¹ Evidence on marriage rates by intelligence can be found in Higgins, Reed, and Reed (1962) and Bajema (1963). The statement on marriage rates by attractiveness is not based on any statistical evidence.

²² That is, all the entries relevant to their decisions. This strong assumption of sufficient information is relaxed in Part II, where "search" for a mate is analyzed.

sibilities. I assume that each person gains from marriage, so that the singles row and column of the payoff matrix can be ignored.

There are $n!$ different combinations that permit each M to marry one F and vice versa; that is, there are $n!$ ways to select one entry in each married row and column. The total output over all marriages produced by any one sorting can be written as

$$Z^k = \sum_{i \in M, j \in F} Z_{ij}, \quad k = 1, \dots, n! \tag{13}$$

Number one of the sortings that maximizes total output so that its entries lie along the diagonal and write

$$Z^* = \sum_{i=1}^n Z_{ii} = \max_k Z^k \geq Z^k \quad \text{all } k. \tag{14}$$

If the total output of any marriage is divided between the mates,

$$m_{ij} + f_{ij} = Z_{ij}, \tag{15}$$

where m_{ij} is the income of the i th M from marriage to the j th F , and similarly for f_{ij} . If each chooses the mate who maximizes his or her "income," the optimal sorting must have the property that persons not married to each other could not marry and make one better off without making the other worse off. In game theoretic language, the optimal sorting is in the "core" since no "coalition" outside the core could make any of its members better off without making some worse off.

Persons entering noncore marriages could not produce more together than the sum of their incomes in the core. For, if they could, and if any division of output between mates were feasible, they could find a division of their output that would make each better off, a contradiction of the definition of the core. If the sorting along the diagonal were in the core, this condition states that

$$m_{ii} + f_{jj} \geq Z_{ij} \quad \text{all } i \text{ and } j. \tag{16}$$

Conditions (15) and (16) immediately rule out any sorting that does not maximize the total output of commodities over all marriages, for at least one M and one F would then be better off with each other than with their mates.²³ Moreover, the theory of optimal assignments, which has

²³ If M_i married F_j and F_i married M_p in an optimal sorting that did not maximize total output, condition (16) requires that $m_{ij} + f_{pi} \geq Z_{ii}$, all ij, pi , or, by summation,

$$Z_p = \sum_{\text{all } ij, pi} m_{ij} + f_{pi} \geq \sum_i Z_{ii} = Z^*.$$

Since Z^* is the maximum total output, it must exceed Z_p , by assumption less than the maximum. Hence, a contradiction, and a proof that the optimal sorting cannot produce less than the maximum total output.

the same mathematical structure as the sorting of persons by marriage, implies the existence of a set of incomes that satisfy conditions (15) and (16) for sortings that maximize total output.²⁴

The solution can be illustrated with the following 2×2 matrix of payoffs:

$$\begin{array}{cc} & F_1 & F_2 \\ M_1 & \left[\begin{array}{cc} 8 & 4 \end{array} \right] \\ M_2 & \left[\begin{array}{cc} 9 & 7 \end{array} \right] \end{array} \quad (17)$$

Although the maximum output in any marriage is between M_2 and F_1 , the optimal sorting is M_1 to F_1 and M_2 to F_2 . For, if $m_{11} = 3$, $f_{11} = 5$, $m_{22} = 5$, and $f_{22} = 2$, M_2 and F_1 have no incentive to marry since $m_{22} + f_{11} = 10 > 9$, and neither do M_1 and F_2 since $m_{11} + f_{22} = 5 > 4$. In other words, the marriage market chooses not the maximum household commodity output of any single marriage but the maximum sum of the outputs over all marriages, just as competitive product markets maximize the sum of the outputs over all firms. Let me stress again that the commodity output maximized by all households is not to be identified with national output as usually measured, but includes conversation, the quantity and quality of children, and other outputs that never enter or enter only imperfectly into the usual measures. Put still differently, the marriage market acts as if it maximizes not the gain from marriage compared to remaining single for any particular marriage, but the average gain over all marriages.²⁵

Each marriage can be considered a two-person firm with either member being the "entrepreneur" who "hires" the other at the "salary" m_{ij} or f_{ij} and receives residual "profits" of $Z_{ij} - m_{ij}$ or $Z_{ij} - f_{ij}$. Another interpretation of the optimal sorting is that only it enables each "entrepreneur" to maximize "profits" for given "salaries" of mates because only the optimal sorting satisfies condition (16). With all other sortings, some "entrepreneurs" could do better by "hiring" different mates than those assigned to them.

²⁴ For a proof, see Koopmans and Beckman (1957).

²⁵ Clearly,

$$\left[\sum_i^n Z_{ii} - \sum_{j=1}^n (Z_{0j} + Z_{j0}) \right] / n = \left\{ \sum_i [Z_{ii} - (Z_{0j} + Z_{j0})] \right\} / n$$

is maximized if

$$\sum Z_{ii}$$

is, since Z_{0j} and Z_{j0} are given and independent of the marital sorting.

b) Assortive Mating

I now consider the optimal sorting when M and F differ in a trait, or set of traits, such as intelligence, race, religion, education, wage rate, height, aggressiveness, tendency to nurture, or age. Psychologists and sociologists have frequently discussed whether likes or unlikes mate, and geneticists have occasionally assumed positive or negative assortive mating instead of random mating. But no systematic analysis has developed that predicts for different kinds of traits when likes or unlikes are motivated to mate.²⁶ My analysis implies that likes or unlikes mate when that maximizes total household commodity output²⁷ over all marriages, regardless of whether the trait is financial (like wage rates and property income), or genetical (like height and intelligence), or psychological (like aggressiveness and passiveness).

Assume that M differs only in the quantitative trait A_m , and F only in A_f , that each trait has a monotonic effect on the output of any marriage, and that higher values have the larger effect:

$$\frac{\partial Z_{ij}(A_m, A_f)}{\partial A_m} > 0, \quad \frac{\partial Z_{ij}(A_m, A_f)}{\partial A_f} > 0. \quad (18)$$

If increasing both A_m and A_f adds the same amount to output as the sum of the additions when each is increased separately, all sortings of M and F would give the same total output. On the other hand, if increasing both adds more to output than the sum of the separate additions, a sorting of large A_m with large A_f and small A_m with small A_f would give the greatest total output since an increase in A_m reinforces the effect of an increase in A_f . The converse holds if increasing both adds less to output than the sum of the separate additions. Mathematically, this states that positive or negative assortive mating—mating of likes or unlikes—is optimal as

$$\frac{\partial^2 Z(A_m, A_f)}{\partial A_m \partial A_f} \geq 0 \quad (19)$$

(proofs in Appendix, Part I, section 1).

Consider, as an example, a matrix of outputs when $n = 2$:

$$\begin{matrix} & A_1 & A_2 \\ A_1 & \left[\begin{matrix} Z_{11} & Z_{12} \end{matrix} \right] \\ A_2 & \left[\begin{matrix} Z_{21} & Z_{22} \end{matrix} \right] \end{matrix}, \quad \text{with } A_2 > A_1. \quad (20)$$

²⁶ Winch (1958) essentially assumes that each person tries to maximize utility ("In mate selection each individual seeks within his or her field of eligibles for that person who gives the greatest promise of providing him or her with maximum need gratification" [pp. 88-89]) and stresses complementary needs as a prerequisite for mating (especially in chap. 4), but he only considers psychological traits, brings in "eligibles" as a *deus ex machina*, and nowhere shows how mating by complementary needs brings equilibrium into the marriage market.

²⁷ Let me emphasize again that commodity output is not the same as national product as usually measured, but includes children, companionship, health, and a variety of other commodities.

If $Z_{22} - Z_{12} > Z_{21} - Z_{11}$, if equality (19) is positive, then obviously $Z_{11} + Z_{22} > Z_{12} + Z_{21}$, and a positive correlation between A_m and A_f maximizes total output, as predicted from (19).

One tradition in production theory distinguishes substitution from complementarity by the sign of the cross-derivative of output with respect to different inputs into a production function. Although condition (19) is not defined in terms of household production functions, duality theory implies that the same condition holds when A_m and A_f are treated as inputs into these production functions.²⁸ Condition (19) says, therefore, that the association of likes is optimal when traits are complements and the association of unlikes is optimal when they are substitutes, a plausible conclusion since high values of different traits reinforce each other when they are complements, and offset each other when they are substitutes.

Economists have generally considered the sorting of different *quantities* of different traits, such as labor and capital, not different *qualities* of the same trait. Although sorting by quantity and quality are related analytically, many applications of sorting by quality are also directly available in economics, such as the optimal sorting of more able workers and more able firms,²⁹ more "modern" farms and more able farmers, or more informed customers and more honest shopkeepers. As already mentioned (n. 26 above), some sociologists have considered "complementarity" to be an important determinant of sorting, but have not given a rigorous analysis of the effects of "complementarity" or embedded their discussions in the context of a functioning marriage market.

Mating of likes—positive assortive mating—is extremely common, whether measured by intelligence, height, skin color, age, education, family background, or religion, although unlikes sometimes also mate, as measured, say, by an inclination to nurture or succor, to dominate or be deferential. This suggests that traits are typically but not always complements.

The determinants of complementarity and substitutability are best discovered by going explicitly to the household production function and the maximization process. All households are assumed to have the same production *function*; that is, if the inputs of time, goods, and *all* traits were exactly the same, the output of commodities would be exactly the same. Different families can, of course, produce different outputs from the same input of goods and time if their education, ability, and so forth, differ.

I consider a number of determinants in turn. First, if M and F differ *only* in their market wage rates—each M and each F are identical in all

²⁸ Wage rates or other monetary variables, however, cannot be treated as productive inputs.

²⁹ This sorting is discussed for Japanese firms by Kuratani (1972). Hicks (1948, chap. 2, sec. 3) asserts that more able workers work for more able firms without offering any proof. Black (1926) discusses the sorting of workers and firms with a few numerical examples.

other market and in nonmarket traits—according to equation (11), the optimal output between M and F who are both participating in the labor force can be written as

$$Z = \frac{S}{C(w_m, w_f, p)}, \quad (21)$$

where the subscripts on Z , S , and C have been omitted and constant returns to scale assumed. Then, by differentiation and by using equation (4),

$$\left. \begin{aligned} Z^m &= \frac{T}{C} - \frac{S}{C^2} C^m, \\ Z^m &= \frac{\partial Z}{\partial w_m} \quad \text{and} \quad C^m \equiv \frac{\partial C}{\partial w_m} \end{aligned} \right\} \quad (22)$$

where

Since

$$C^m = t_m Z^{-1}, \quad (23)$$

where t_m is the time spent by M in the household,

$$Z^m = l_m C^{-1} > 0 \quad (24)$$

if l_m , the time spent at work, is greater than zero. Similarly,

$$Z^f = \frac{T}{C} - \frac{S}{C^2} C^f = l_f C^{-1} > 0. \quad (25)$$

Positive or negative assortive mating by wage rates is optimal as

$$\frac{\partial^2 Z}{\partial w_m \partial w_f} \equiv Z^{mf} \equiv Z^{fm} \geq 0. \quad (26)$$

Differentiate Z^f with respect to w_m to get

$$Z^{fm} = -C^{-2} C^m l_f + C^{-1} \frac{\partial l_f}{\partial w_m}. \quad (27)$$

The first term on the right is clearly negative, so Z^{fm} will be negative if the second term, $\partial l_f / \partial w_m \leq 0$, is nonpositive, that is, if t_m and t_f are not gross complements, as these terms are usually defined.³⁰ Consequently, a perfectly negative rank correlation between w_m and w_f would maximize total commodity output if the time of M and F were not such gross

³⁰ This definition is different from the one given earlier in terms of the sign of the cross-derivative of profit or production functions. The definition in equation (28) is preferable, at least as a predictor of responses to changes in input prices. By "gross" rather than "net" complements is meant in the usual way that the income effect is included along with the substitution effect. Even if t_m and t_f were net complements they could still be gross substitutes since the income effect of an increase in w_m would tend to increase t_f .

complements as to swamp the first term in (27). Considerable empirical evidence supports the conclusion that t_m and t_f are not gross complements (Ofek 1972; Smith 1972a).

A negative correlation between w_m and w_f maximizes total output because the gain from the division of labor is maximized. Low-wage F should spend more time in household production than high-wage F because the foregone value of the time of low-wage F is lower; similarly, low-wage M should spend more time in household production than high-wage M . By mating low-wage F with high-wage M and low-wage M with high-wage F , the cheaper time of both M and F is used more extensively in household production, and the expensive time of both is used more extensively in market production.

All persons have been assumed to participate in the labor force. During any year, however, most married women in the United States do not participate, and a significant number never really participate throughout their married life. My analysis does predict that many women would have only a weak attachment to the labor force since low-wage women would be discouraged from participation both by their low wage and by the high wage of their husbands.³¹

If some women are not in the labor force, however, the wage rates of men and women need not be perfectly negatively correlated to maximize total output. For assume that all women with wage rates below a certain level would not participate in the labor force with a perfectly negative correlation between the wage rates of men and women. These women have $\partial Z/\partial w_f = 0$,³² and, thus, $Z^{f^m} = 0$; therefore, up to a point, they could switch mates without lowering total output. Consequently, other sortings having weaker negative, and conceivably even positive, correlations would also maximize total output; that is, many sortings would be equally good, and wage rates would not be a decisive determinant of the optimal sorting.

If M and F differ only in their stock of nonhuman capital, K_m and K_f , and if everyone participates in the labor force, $\partial C/\partial K_m = \partial C/\partial K_f = 0$ since the value of time is measured by the market wage rates. If the rate of return on K , denoted by r , depended positively on the amount of time allocated to "portfolio management," r would be positively related to K .³³ It then follows that

³¹ Low-wage men also would be encouraged to work less both because of their low wage and the relatively high wage of their wives. They would not leave the labor force in large numbers, however, partly because average wage rates of men are so much higher than those of women and partly because the nonmarket productivity of women is higher than that of men.

³² As long as they are not indifferent at the margin to working in the market sector.

³³ For this result and a more complete analysis of the allocation of time to portfolio management, see Ben-Zion and Ehrlich (1972).

and
$$\left. \begin{aligned} \frac{\partial Z}{\partial K_m} = \frac{\partial Z}{\partial K_f} = rC^{-1} > 0 \\ \frac{\partial^2 Z}{\partial K_m \partial K_f} = \frac{dr}{dK} C^{-1} > 0 \end{aligned} \right\} .^{34} \tag{28}$$

A perfectly positive correlation between the nonhuman capital of *M* and *F* would be optimal, an implication that is consistent with evidence on sorting by, say, parental wealth.

If some *F* did not participate in the labor force, the value of their time would be measured by a “shadow” price that exceeded their wage rate and was not constant but positively related to the sum of their nonhuman capital.³⁵ Moreover, a perfectly positive correlation of this capital is no longer necessarily optimal because of diminishing returns to an increase in the time of *M* and goods for a given amount of the time of *F* (for proof, see Appendix, Part I, section 2).

All differences in the output of commodities, by assumption the only determinant of behavior, not related to differences in wage rates or nonhuman capital are, by definition, related to differences in nonmarket productivity.³⁶ The widespread differences between men and women in nonmarket productivity are caused by differences in intelligence, education, health, strength, height, personality, religion, and other traits. I now consider the optimal sorting of traits that affect nonmarket productivity, while assuming that wage rates and nonhuman capital are the same for all *M* and for all *F*.

To demonstrate the tendency toward complementarity of nonmarket traits in the context of household commodity outputs, rewrite the optimal output equation given by (21) as

$$Z = \frac{S}{C(w_m, w_f, p, A_m, A_f)}, \tag{29}$$

where *A_m* and *A_f* are the traits of *M* and *F*. Then using the assumption that *w_m*, *w_f*, and the rate of return on nonhuman capital are independent of *A_m* and *A_f*,

³⁴ If time is allocated to portfolio management, $S = wT + Kr(\ell_p) - w\ell_p$, where ℓ_p is the time so allocated. Then $\partial S/\partial K = r + (K dr/d\ell_p)(d\ell_p/dK) - w(d\ell_p/dK) = r + d\ell_p/dK[(K dr/d\ell_p) - w]$. Since, however, $K dr/d\ell_p = w$ is one of the first-order maximization conditions, then $\partial S/\partial K = r$.

³⁵ See the discussion in section 2, Part I of the Appendix.

³⁶ Differences in the earning power of children are assumed to be derived from differences in either the nonmarket productivity or incomes of their parents, and are not considered separately.

$$\left. \begin{aligned} \frac{\partial C}{\partial A_m} &\equiv C_{a_m} \\ \frac{\partial C}{\partial A_f} &\equiv C_{a_f} \end{aligned} \right\} < 0 \quad \text{and} \quad \frac{\partial S}{\partial A_f} = \frac{\partial S}{\partial A_m} = 0. \quad (30)$$

Then,

$$\left. \begin{aligned} \frac{\partial Z}{\partial A_m} &= -SC^{-2}C_{a_m} \\ \frac{\partial Z}{\partial A_f} &= -SC^{-2}C_{a_f} \end{aligned} \right\} > 0, \quad (31)$$

and

$$\frac{\partial^2 Z}{\partial A_m \partial A_f} > 0 \quad \text{if } 2C^{-1}C_{a_m}C_{a_f} > C_{a_m a_f}. \quad (32)$$

Since the term on the left is positive, equation (32) necessarily holds if A_m and A_f have either independent or reinforcing effects on productivity, for then $C_{a_m a_f} \leq 0$; moreover, (32) might hold even if they had offsetting effects. Therefore, perfectly positive assortive mating is definitely optimal if the traits have reinforcing effects; less obvious and more impressive, however, is the conclusion that positive assortive mating is also optimal if they have independent effects because C enters inversely in the equation for Z , or even if they have offsetting effects if these are weaker than a multiple of the direct ones.³⁷

The reasons for the prevalence of a complementary relation between traits that raise nonmarket productivity can be seen more transparently by considering a couple of special cases. If the percentage effect on output of a trait were independent of the quantities of goods and time, the optimal output equation could be written as

$$Z = \frac{S}{b(A_m, A_f)K(w_m, w_f, \rho)}, \quad (33)$$

where $\partial b/\partial A_m \equiv b_{a_m} < 0$, and $\partial b/\partial A_f \equiv b_{a_f} < 0$. Hence,

$$\frac{\partial^2 Z}{\partial A_m \partial A_f} > 0 \quad \text{as } 2b^{-1}b_{a_m}b_{a_f} > b_{a_m a_f}, \quad (34)$$

³⁷ Equation (32) can be written as

$$2|\varepsilon_{cam}| > \varepsilon_{ca_f a_m},$$

where $\varepsilon_{cam} = (C_{a_m} \cdot A_m)/C < 0$, and $\varepsilon_{ca_f a_m} = C_{a_f a_m} \cdot A_m/C_{a_f} > 0$ if the effects are offsetting. The cross-elasticity must be smaller than twice the absolute value of the direct elasticity.

which must hold if $b_{a_m, a_f} \leq 0$ and can easily hold even if $b_{a_m, a_f} > 0$. Positive assortive mating is optimal even when these productivity effects are independent because productivity is raised multiplicatively: higher A_m (or A_f) have bigger *absolute* effects when combined with higher A_f (or A_m). A fortiori, this multiplicative relation encourages the mating of likes when the effects are reinforcing and can do so even when they are offsetting.³⁸

The effect of most traits on nonmarket output is not independent of goods and time, but generally operates through the time supplied to the household; for example, if the time supplied became zero, so would the effect. A simple way to incorporate this interaction is to assume that each trait affects outputs only by augmenting the effective amount of own household time. It is shown in section 3, Part I of the Appendix that positive assortive mating would still be optimal as long as the elasticity of substitution between the household time of M and F was not very high.³⁹ Negative assortive mating can be expected for own-time-augmenting traits only if they augment dimensions that are easily substitutable between M and F . Dominant and deferential persons tend to marry each other (Winch 1958), perhaps, therefore, because the dominant person's time can be used when households encounter situations calling for dominance and the deferential person's time can be used when they call for deference.

Note that it is shown in section 2 that the gain from marriage is also greater when substitution between the time of M and F is more difficult. Therefore, the mating of likes should be more common when marriage is more attractive, an important and subtle implication of the analysis.

How do the nonmarket traits of one sex combine with the market traits of the other? In particular, does my analysis justify the popular belief that more beautiful, charming, and talented women tend to marry wealthier and more successful men? Section 4 in Part I of the Appendix shows that a positive sorting of nonmarket traits with nonhuman wealth always, and with earning power usually,⁴⁰ maximizes commodity output over all marriages. The economic interpretation is basically that non-market productivity and money income tend to combine multiplicatively, so that higher values of a trait have larger absolute effects when combined with higher income.

Scattered references have been made to the empirical evidence on sorting, and this evidence is now considered a little more systematically. The simple correlations between the intelligence, education, age, race,

³⁸ Section 3, Part I of the Appendix shows that positive assortive mating of A_m and A_f is still optimal even when F do not participate in the labor force.

³⁹ The elasticity estimates of Ofek (1972) and Smith (1972a) are only of modest size.

⁴⁰ By "usually" is meant that a positive sorting with earnings always maximizes total output when an increase in a trait does not decrease the spouses' hours worked in the market sector and *could* maximize output even when they do decrease.

nonhuman wealth, religion, ethnic origin, height, and geographical propinquity of spouses are positive and strong.⁴¹ A small amount of evidence suggests that the correlations between certain psychological traits, such as a propensity to dominate, nurture, or be hostile, are negative.⁴² The correlation by intelligence is especially interesting since, although intelligence is highly inheritable, the correlation between mates is about as high as that between siblings (Alstrom 1961). Apparently, the marriage market, aided by coeducational schools, admissions tests, and the like, is more efficient than is commonly believed.

This evidence of positive simple correlations for a variety of traits, and of negative correlations for some, is certainly consistent with my theory of sorting. A more powerful test of the theory, however, requires evidence on partial correlations, when various other traits are held constant. For example, how strong is the correlation by intelligence, when years of schooling and family background are held constant? I do not yet have results on partial correlations by intelligence, but do have some on years of schooling, wage rates, and age, for samples of white and black families.⁴³ Even when age and wage rates are held constant, the correlation between years of schooling is high, $+ .53$ for whites and virtually the same ($+ .56$) for blacks. Although the partial correlations between wage rates are much lower, they are also positive, $+ .32$ for whites and a bit lower ($+ .24$) for blacks.

The strong positive partial correlation between years of schooling is predicted by the theory, but the positive correlation between wage rates is troublesome since the theory predicts a negative correlation when nonmarket productivity is held constant. Note, however, that the sample is biased because it is restricted to women in the labor force in a particular year. Since the higher the husband's wage rate the higher must be his wife's wage rate to induce her to enter the labor force, a negative correlation across all mates is consistent with a positive one for those having wives in the labor force.⁴⁴ Indeed, Gregg Lewis has shown⁴⁵ that a correlation of about $+ .3$ for mates who are participating almost certainly implies a negative one (about $- .25$) for all mates, given the relatively small

⁴¹ Many of the relevant studies are listed in Winch (1958, chap. 1).

⁴² See Winch (1958, chap. 5). Deference is treated as negative values of dominance, succorance as negative values of nurturance, and abasement as negative values of hostility.

⁴³ A 20 percent random sample of the approximately 18,000 married persons in the 1967 Survey of Economic Opportunity was taken. Families were included only if the husband and wife both were less than age 65 and were employed, the wife for at least 20 hours in the survey week.

⁴⁴ Also, nonmarket productivity varies even when years of schooling and age are held constant. If investments that raise nonmarket productivity also raise, somewhat, market earning power (Heckman [1974] finds that the education of women raises their nonmarket productivity almost as much as their market earning power), the positive correlation between wage rates may really be picking up the predicted positive correlation between husband's wage rate and wife's nonmarket productivity.

⁴⁵ Via an unpublished memorandum extending some work of Gronau (1972).

fraction of married women who participate. If his calculations hold up, this would be striking confirmation of my theory since it is counter to common impressions and is one of the few examples (and a predicted one!) of negative associative mating.

Other evidence, probably less affected by unobserved differences in nonmarket productivity, does suggest that the gain from marriage is greater when differentials between male and female wage rates are greater. For example, a larger percentage of persons are married in American states that have higher wages of males and lower wages of females, even when age, years of schooling, the sex ratio, the fraction Catholic, and other variables are held constant (Santos 1970; Frieden 1972). Or a larger fraction of black households are headed by women in metropolitan areas with higher earnings of black women relative to black men (Reischauer 1970).

Quantitative evidence on the association of traits that affect nonmarket productivity with earnings and other income is scarce. The evidence I put together and referred to earlier indicates that husband's wage rate and wife's education are significantly positively correlated, even when husband's education and wife's wage rate are held constant.⁴⁶ One interpretation, stressed by Benham in his paper which follows, is that a wife's education contributes to her husband's earnings, just as a mother's education is said to contribute to her children's earnings (Leibowitz 1972). An alternative suggested by our theory of sorting is that a wife's education is a proxy for traits affecting her nonmarket productivity, especially when her wage rate is held constant⁴⁷ and that women with higher nonmarket productivity marry men with higher earning power. Although the relative importance of these alternative interpretations has not been determined, Benham does find that hours worked by husbands are positively related to wife's education, a sufficient condition for positive sorting (see n. 40 above).

My analysis of mating and sorting has assumed perfect certainty in the production of household commodities. Uncertainty surrounds the production of many commodities, but my concern here is only with uncertainty about the "quality" of own children since children are a major source of the gain from marriage. An important result in population genetics is that positive assortive mating of inheritable traits, like race, intelligence, or height, increases the correlation of these traits among siblings; the increase would be greater the more inheritable the trait is and the greater the degree of assortive mating (Cavalli-Sforza and Bodmer 1971, chap. 9, sec. 7). Therefore, inheritable traits of M and F

⁴⁶ In his more detailed analysis in this book, Benham finds similar results, after several additional variables are also held constant. Note, however, that the husband's wage rate is much more strongly related to his own than to his wife's education.

⁴⁷ I argued earlier that her wage rate also is a proxy for such traits, when her education is held constant.

can be said to be complements in reducing the uncertainty about one's children. Positive assortive mating of inheritable traits would increase the utility of total output if more certainty about the "quality" of children is desirable—perhaps because friction between siblings or the cost of raising them is increased by uncertainty.

My analysis of sorting is based on several other simplifying assumptions that ought to be modified in a fuller treatment. For example, the conclusion in section 2, that the gain from marriage is independent of preferences, assumes, among other things, no joint production and constant returns to scale in households. With beneficial joint production⁴⁸ or increasing returns, mating of persons with similar preferences would be optimal and conversely with detrimental production or decreasing returns. Similarly, the conclusion in section 2, that a monogamous union is always optimal, which is taken for granted in the discussion of sorting, should be modified to consider polygamy (I do this in Part II) and remaining single (see the discussion of search in Part II). Further, I have considered only one trait at a time, holding all other traits constant. But since people differ in many interdependent traits, optimal sortings should be determined for a set of traits, perhaps using the canonical correlation coefficient or related statistics as the measure of association.

Probably the assumption that would be most questioned is that any division of output between mates is feasible. Some of the output may not be divisible at all and may constitute a "public," or better still, a "family" commodity. Children might be said to be largely a family commodity, and, as shown in Part II, "caring" can convert the whole output into family commodities. Or some divisions may not be feasible because they are not enforceable. For example, even though the marriage market might dictate a 2/5 share for a particular husband, he may receive a 3/5 share because his wife cannot "police" the division perfectly.

Although the rigidities resulting from family commodities and enforcement problems can often be overcome (through dowries and other capital transfers), it is instructive to consider a model of sorting that incorporates these rigidities in an extreme fashion. How robust are the conclusions about optimal sorting when complete rigidity in the division of output replaces the assumption of complete negotiability?

Rigidity is introduced by assuming that M_i would receive a constant fraction e_i of commodity output in *all* marriages, and F_j receive d_j . Note that e_i and e_k ($k \neq i$) or d_j and d_k ($k \neq j$) need not be equal, and that

$$e_i + d_j \geq 1, \quad (35)$$

as family commodities or enforcement costs were dominant. The matrix showing the incomes for all combinations of M and F would then be

⁴⁸ Grossman (1971) distinguishes beneficial from detrimental production by the effect of an increase in output of one commodity on the cost of producing others.

	F_1	F_j	F_n
M_1	$e_1 Z_{11}, d_1 Z_{11}$	$e_1 Z_{1n}, d_n Z_{nn}$
M_i		$e_i Z_{ij}, d_j Z_{ij}$	
M_n	$e_n Z_{nj}, d_1 Z_{n1}$	$e_n Z_{nn}, d_n Z_{nn}$

(36)

If

$$\hat{Z}_1 \equiv Z_{st} > Z_{ij}, \quad \text{all } i \neq s, \text{ all } j \neq t, \tag{37}$$

were the maximum output in any possible marriage and if each person tried to maximize his commodity income, M_s would marry F_t since they could not do as well in any other marriage.⁴⁹ Now exclude M_s and F_t from consideration, and if

$$\hat{Z}_2 = Z_{uv} > Z_{ij}, \quad \text{all } i \neq u \text{ or } s, \text{ all } j \neq v \text{ or } t, \tag{38}$$

were the maximum output in all other marriages, M_u would marry F_v . This process can be continued through the $\hat{Z}_3, \dots, \hat{Z}_n$ until all the M and F are sorted.

How does this sorting, which combines the various maxima, compare with that obtained earlier, which maximizes total output? As the example in (17) indicates, they are not necessarily the same: combining the maxima in that example sorts M_2 with F_1 and M_1 with F_2 , whereas maximizing total output sorts M_1 with F_1 and M_2 with F_2 . Yet, in perhaps the most realistic cases, they are the same, which means that the sum of the maxima would equal the maximum of the sums.

Assume that an increase in trait A_m or A_f always increases output and that M and F are numbered from lower to higher values of these traits. Then, \hat{Z}_1 is the output of M_n with F_n , \hat{Z}_2 is that of M_{n-1} with F_{n-1} , and \hat{Z}_n that of M_1 with F_1 . Consequently, when traits have monotonic effects on output, the most common situation, combining the various maxima implies perfectly positive assortive mating.

I showed earlier that, in a wide variety of situations, namely, where traits are "complementary," maximizing total output also implies perfectly positive assortive mating. In these situations, permitting the market to determine the division of output and imposing the division a priori gives exactly the same sorting. Therefore, the implication of the theory about the importance of positive assortive mating is not weakened, but rather strengthened, by a radical change in assumptions about the determinants of the division of output.

When maximizing total output implies negative assortive mating, as it does between wage rates (with nonmarket productivity held constant), and between own-time augmenting traits that are close substitutes, these assumptions about the division of output have different implications. The

⁴⁹ Clearly, $e_s Z_{st} > e_s Z_{sj}$, all $j \neq t$, and $d_t Z_{st} > d_t Z_{it}$, all $i \neq s$ by condition (37).

empirical evidence on sortings cannot yet clearly choose between these assumptions, however, because positive sortings are so common: perhaps the positive correlation between observed wage rates is evidence of rigidities in the division, but several alternative interpretations of this correlation have been suggested that are consistent with a negative "true" correlation, and some psychological traits are apparently negatively correlated. Moreover, dowries and other capital transfers provide more effective fluidity in the division than may appear to the casual observer.

4. *The Division of Output between Mates*

With complete negotiability the division of output is given by condition (15) and (16). The m_{ii} and f_{ii} are determined by their marginal productivity in the sense that if $Z_{ki} > Z_{kk}$, necessarily $f_{ii} > f_{kk}$,⁵⁰ and similarly for the m_{ii} . Also, if $f_{ii} > f_{kk}$, necessarily $Z_{ii} > Z_{ik}$.⁵¹ The following limits are easily derived:

$$\left. \begin{aligned} Z_{ii} - \text{Max}_k (Z_{ki} - Z_{kk}) &\geq m_{ii} \geq \text{Max}_k (Z_{ik} - Z_{kk}) \\ Z_{ii} - \text{Max}_k (Z_{ik} - Z_{kk}) &\geq f_{ii} \geq \text{Max}_k (Z_{ki} - Z_{kk}) \end{aligned} \right\} .52 \quad (39)$$

The division of output resulting from conditions (15) and (16) is not unique, however. For if a set of m_{ii} and f_{ii} satisfies these conditions with all $0 < m_{ii} < Z_{ii}$, a positive quantity λ exists, such that $m_{ii} + \lambda$ and $f_{ii} - \lambda$ also satisfy these conditions. The range of indeterminacy in the division would narrow as the sum of $\text{Max}_k (Z_{ik} - Z_{kk})$ and $\text{Max}_k (Z_{ki} - Z_{kk})$ approached closer to Z_{ii} .

Clearly, the indeterminacy would vanish if the distribution of Z_{ik} became continuous. It could also vanish in a second case to which I turn. Assume v_i identical M_i and u_i identical F_i ; by identical is meant that they would produce the same output with any mate or while single, so that they would receive the same income in market equilibrium. If the number of v_i were sufficiently large for a competitive equilibrium, there would be a supply curve of M_i to the marriage market: it would be horizontal at the singles income Z_{i0} until all v_i^0 were married, and then would rise vertically (see S_0 in fig. 1). Similarly, if the number of u_i were sufficiently large, there would be a market supply curve of F_i : it would be horizontal at Z_{0i} until all u_i^0 were married, and then would rise vertically. If initially I assume, for simplicity, that the M_i and F_i either marry each other or remain single, the supply curve of F_i would

⁵⁰ Since $f_{kk} + m_{kk} = z_{kk}$, all k , and $f_{ii} + m_{kk} \geq z_{ki}$, all i and k , then $f_{ii} - f_{kk} \geq z_{ki} - z_{kk} > 0$ by assumption.

⁵¹ That is, if $f_{ii} > f_{kk}$, then $Z_{ii} = m_{ii} + f_{ii} > m_{ii} + f_{kk} \geq Z_{ik}$.

⁵² Given conditions (15) and (16), $m_{ii} - m_{kk} \geq Z_{ik} - Z_{kk}$, all k , or, since $m_{kk} \geq 0$, $m_{ii} \geq Z_{ik} - Z_{kk}$, all k . The other conditions in (39) can be proved in a similar way.

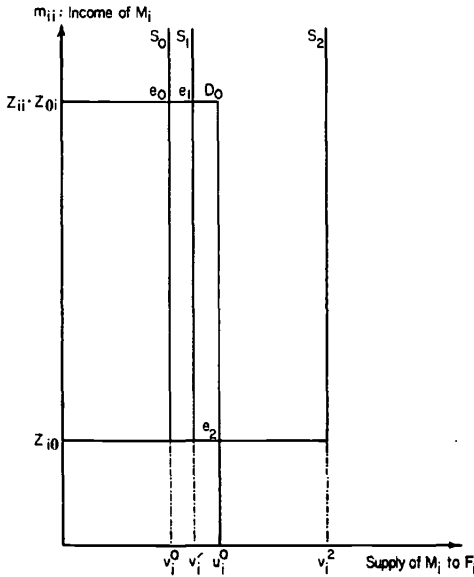


FIG. 1

also be a derived demand curve for M_i that would be horizontal at $Z_{ii} - Z_{oi}$ until all u_i^o were married, and then would fall vertically (D_o in fig. 1); moreover, the supply curve of M_i to the market would be its supply curve to F_i .

The equilibrium income to each M_i is given by point e_o , the intersection of S_o and D_o . If the sex ratio (v_i^o/u_i^o) were less than unity, the equilibrium position is necessarily on the horizontal section of the derived demand curve, as is e_o . All the M_i would marry and receive the whole difference between their married output and the singles output of F_i . All the F_i would receive their singles output and, therefore, would be indifferent between marrying and remaining single, although market forces would encourage v_i^o of them to marry.

An increase in the sex ratio due to an increase in the number of M_i would lengthen the horizontal section of the supply curve and shift the equilibrium position to the right, say, to e_1 . All the M_i would continue to marry and a larger fraction of the F_i also would. If the sex ratio rose above unity, equilibrium would be on the horizontal section of the supply rather than the derived demand curve (see e_2). Now all the F_i would marry and receive the whole difference between their married output and the singles output of M_i ; market forces would induce u_i^o of the M_i to marry, and $v_i^2 - u_i^o$ to remain single.

The importance of sex ratios in determining the fraction of men and women who marry has been verified by numerous episodes and in several studies. An aftermath of a destructive war is many unmarried young

women pursuing the relatively few men available, and men usually either marry late or not at all in rural areas that have lost many young women to cities. Statistical studies indicate that the fraction of women currently married at different ages is positively related to the appropriate sex ratio.⁵³

I know of only highly impressionistic evidence on the effects of the sex ratio, or for that matter any other variable, on the division of output between mates. This division usually has not been assumed to be responsive to market forces, so that no effort has been put into collecting relevant evidence. Admittedly, it is difficult to separate expenditures of goods and time into those that benefit the husband, the wife, or both, but with enough will something useful could be done. For example, the information giving the separate expenditures on husband's and wife's clothing in some consumer surveys, or on the "leisure" time of husbands and wives in some time budget studies could be related to sex ratios, wage rates, education levels, and other relevant determinants of the division of output.

If I drop the assumption that all the M_i and F_i must either marry each other or remain single, M_i 's supply curve to F_i would differ from its market supply curve because marriage to other persons would be substituted for marriage to F_i ; similarly, F_i 's supply curve to M_i would differ from its market supply curve. To demonstrate this, suppose that, at point e_0 in figure 1, M_i does better by marrying F_i than by marrying anyone else; that is, condition (16) is a strict inequality for M_i . If M_i 's income from marrying F_i were less than at e_0 , the difference between the sum of M_i 's income and that of other $F_j \neq F_i$, and what they could produce together would be reduced. At some income, this difference might be eliminated for an F , say, F_k : then all the M_i would be indifferent between marrying F_i and F_k .

At lower values of M_i 's income from marrying F_i , some of the M_i would try to marry F_k . The increase in the supply of mates to F_k would raise M_i 's income and reduce that of M_i 's mates. In equilibrium, just enough M_i would marry F_k to maintain equality between the income M_i receives with F_i and F_k . The important point is that if some M_i marry F_k , the number marrying F_i would be less than the number supplied to the marriage market (v_i). Moreover, the number marrying F_i might fall still further as M_i 's income with F_i fell further because some might marry, say, F_p , if they could then do as well with F_p as with F_i or F_k .

The net effect of these substitutions toward other F is a rising supply curve of M_i to F_i , shown by S_0 in figure 2, with an elasticity determined both by the distribution of substitute F and by the effect on the income of

⁵³ See the studies essentially of whites by Santos (1970) and Freiden (1972), of blacks by Reischauer (1970), of Puerto Rico by Nerlove and Schultz (1970), and of Ireland by Walsh (1972). By "appropriate" is meant that a group of women must be matched with the men they are most likely to marry, e.g., college-educated women with college-educated men, or women aged 20-24 with men aged 25-29.

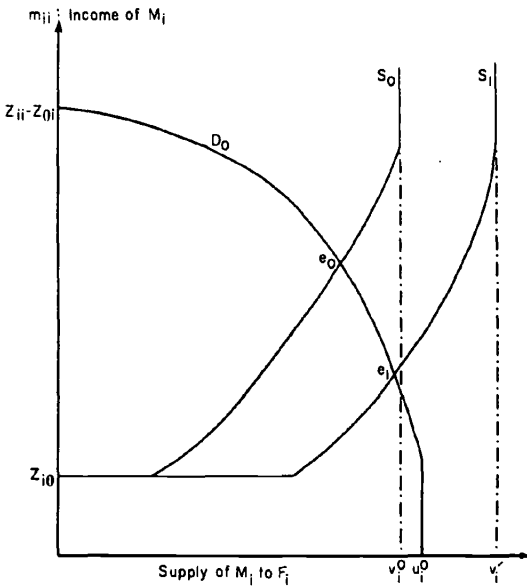


FIG. 2

these F of a given increase in the number of M_i available to marry them. Since F_i would also substitute toward other M , its derived demand curve for M_i would also fall, as D_0 does in figure 2. The equilibrium position e_0 determines both the division of output between M_i and F_i and the number marrying each other. The difference between the total number of M_i , v_i^0 , and the number marrying F_i no longer measures the number of M_i remaining single, since at e_0 all M_i marry, but rather it measures the number marrying other F and receiving the same income as the M_i marrying F_i ; similarly, for the F_i .

An increase in the number of M_i from v_i^0 to v_i' would shift their supply curve to F_i to the right and lower the equilibrium position to e_1 in figure 2. The reduction in M_i 's income (equal to the increase in F_i 's income) is negatively related to the elasticities of the demand and supply curves, which are determined by the availability of substitute M and F . The additional M_i all marry, some to F_i and some to other F ; a larger fraction of the F_i are induced to marry M_i by the increase in F_i 's income.

An increase in the sex ratio between M_i and F_i would not necessarily increase the fraction of F_i or decrease the fraction of M_i who marry since all can marry if some marry other F or M . However, if all F_i and M_i married, an increase in their sex ratio would tend to decrease the number of other M or increase the number of other F who marry, if the quantity of other M and F were fixed. For an increase in the ratio of M_i to F_i not only lowers M_i 's and raises F_i 's income, but also lowers the incomes of substitute M and raises those of substitute F . Some of these M

would thereby be induced not to marry because their gain from marriage would be eliminated, and some F would be induced to marry because a gain from marriage would be created. Consequently, an increase in the ratio of M_i to F_i would still decrease the fraction of M and increase the fraction of F marrying, if substitute M and F as well as M_i and F_i were considered.

To illustrate these effects, assume an autonomous increase (perhaps due to selective immigration) in the size of a group of identical men, aged 24, who initially were indifferent between marrying women aged 22 and those slightly older or younger, although most married 22-year-olds. The increase in their numbers would decrease their income and the proportion marrying women aged 22. For if the percentage increase in the number marrying women aged 22 were as large as the increase in the number marrying other women, the income of those marrying 22-year-olds would fall by more than others, since men aged 24 are a larger fraction of all men marrying women aged 22 than of all men marrying women of other ages. Moreover, the income of women aged 22 would increase and more of them would marry men aged 24; the income of older or younger men marrying women aged 22 would fall and they would be encouraged to marry women of other ages; the income of women somewhat older or younger than 22 would increase too, and so on.⁵⁴

5. Summary

In Part I above I have offered a simplified model of marriage that relies on two basic assumptions: (1) each person tries to find a mate who maximizes his or her well-being, with well-being measured by the consumption of household-produced commodities; and (2) the "marriage market" is assumed to be in equilibrium, in the sense that no person could change mates and become better off. I have argued that the gain from marriage compared to remaining single for any two persons is positively related to their incomes, the relative difference in their wage rates, and the level of nonmarket-productivity-augmenting variables, such as education or beauty. For example, the gain to a man and woman from marrying compared to remaining single is shown to depend positively on their incomes, human capital, and relative difference in wage rate.

The theory also implies that men differing in physical capital, education or intelligence (aside from their effects on wage rates), height, race, or many other traits will tend to marry women with like values of these traits, whereas the correlation between mates for wage rates or for traits of men and women that are close substitutes in household production will tend to be negative.

⁵⁴ The permanence of these effects depends on whether the immigration continues or is once and for all.

My theory does not take the division of output between mates as given, but rather derives it from the nature of the marriage market equilibrium. The division is determined here, as in other markets, by marginal productivities, and these are affected by the human and physical capital of different persons, by sex ratios, that is, the relative numbers of men and women, and by some other variables.

II

1. *Introduction*

In the discussion which follows I extend the simplified analyses in Part I in several directions. My purpose is both to enrich the analysis in Part I and to show the power of this approach in handling different kinds of marital behavior.

The effect of "love" and caring between mates on the nature of equilibrium in the marriage market is considered. Polygamy is discussed, and especially the relation between its incidence and the degree of inequality among men and the inequality in the number of men and women. The implications of different sorting patterns for inequality in family resources and genetic natural selection are explored. The assumption of complete information about all potential mates is dropped and I consider the search for information through dating, coeducational schools, "trial" marriages, and other ways. This search is put in a life-cycle context that includes marriage, having children, sometimes separation and divorce, remarriage, and so forth.

2. *Love, Caring, and Marriage*

In Part I, I ignored "love," that cause of marriage glorified in the American culture. At an abstract level, love and other emotional attachments, such as sexual activity or frequent close contact with a particular person, can be considered particular nonmarketable household commodities, and nothing much need be added to the analysis, in Part I, of the demand for commodities. That is, if an important set of commodities produced by households results from "love," the sorting of mates that maximizes total commodity output over all marriages is partly determined by the sorting that maximizes the output of these commodities. The whole discussion in Part I would continue to be relevant.

There is a considerable literature on the effect of different variables such as personality, physical appearance, education, or intelligence, on the likelihood of different persons loving each other. Since I do not have anything to add to the explanation of whether or why one person would love another, my discussion concentrates on some effects of love on marriage. In particular, since loving someone usually involves caring

about what happens to him or her,⁵⁵ I concentrate on working out several implications, for marriage, of “caring.”

An inclusive measure of “what happens” is given by the level of commodity consumption, and the natural way for an economist to measure “caring” is through the utility function.⁵⁶ That is, if M cares about F, M’s utility would depend on the commodity consumption of F as well as on his own; graphically, M’s indifference curves in figure 3 are negatively inclined with respect to Z_m and Z_f , the commodities consumed by M and F respectively.⁵⁷ If M cared as much about F as about himself (I call this “full” caring), the slopes of all the indifference curves would equal unity (in absolute value) along the 45° line;⁵⁸ if he cared more about himself, the slopes would exceed unity, and conversely if he cared more about F.

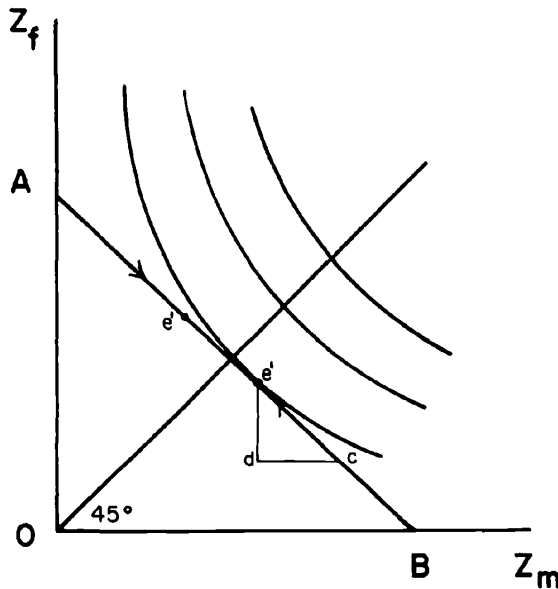


FIG. 3

⁵⁵ The *Random House Dictionary of the English Language* includes in its definitions of love, “affectionate concern for the well-being of others,” and “the profoundly tender or passionate affection for a person of the opposite sex.”

⁵⁶ This formulation is taken from my paper, “A Theory of Social Interactions” (1969).

⁵⁷ Since there is only a single aggregate commodity, saying that M’s utility depends on F’s consumption is equivalent to saying that M’s utility depends on F’s utility (assuming that F does not care about M). If many commodities Z_1, \dots, Z_q , were consumed, M’s utility would depend on F’s utility if $U^m = U^m[Z_{1m}, \dots, Z_{qm}, g(Z_{1f}, \dots, Z_{qf})]$ where g describes the indifference surface of F. Hence $(\partial U^m / \partial Z_{1f}) / (\partial U^m / \partial Z_{1m}) = (\partial g / \partial Z_{1f}) / (\partial g / \partial Z_{1m})$; this ratio is F’s marginal rate of substitution between Z_1 and Z_f .

⁵⁸ “Full” caring might also imply that the indifference curves were straight lines with a slope of unity, that Z_f was a perfect substitute for Z_m .

Point c in figure 3 represents the allocation of commodities to M and F that is determined by equilibrium in the marriage market. Only if M were married to F could he transfer commodities to F, since household commodities are transferable within but not between households. If the terms of transfer are measured by the line AB , he moves along AB to point e : he transfers cd and F receives de . Presumably commodities can be transferred within a household without loss, so that AB would have a slope of unity. Then the equilibrium position after the transfer would be on the 45° line with full caring, and to the right of this line if M preferred his own consumption to F's.

Most people no doubt find the concept of a market allocation of commodities to beloved mates strange and unrealistic. And, as I have shown, caring can strikingly modify the market allocation between married persons. For example, the final allocation (point e) after the transfer from M to F has more equal shares than does the market allocation (point c).⁵⁹ Moreover, if F also cared about M, she would modify the market allocation by transferring resources to M from anywhere in the interval Ae' until she reached a point e' ,⁶⁰ generally to the left of e . The market completely determines the division of output only in the interval $e'e$: positions in Be are modified to e , and those in Ae' are modified to e' . Furthermore, if each fully cares for the other, points e and e' are identical and on the 45° line. Then the total amount produced by M and F would be shared equally, regardless of the market-determined division. This concept of caring between married persons, therefore, does imply sharing—equal sharing when the caring is full and mutual—and is thus consistent with the popular belief that persons in love “share.”

Sharing implies that changes in the sex ratio or other variables considered in section 4 of Part I would not modify the actual distribution of output between married M and F (unless the market-mandated distribution were in the interval ee'). This is another empirical implication of caring that can be used to determine its importance.

I indicated earlier that total income would be less than total output in a marriage if resources were spent “policing” the market-mandated division of output, whereas total income would exceed total output if some output were a “family” commodity, that is, were consumed by both mates. Caring raises total income relative to total output both by reducing policing costs and by increasing the importance of family commodities.

Consider first the effect of caring on policing costs. “Policing” reduces the probability that a mate shirks duties or appropriates more output than is mandated by the equilibrium in the marriage market.⁶¹ Caring reduces

⁵⁹ Provided it were in the interval Ae , M would not modify the market allocation.

⁶⁰ I assume that AB also gives the terms of transfer for F, and that e' is the point of tangency between AB and her indifference curves.

⁶¹ Policing is necessary in any partnership or corporation, or, more generally, in any cooperative activity (see Becker 1971*b*, pp. 122–23; Alchian and Demsetz 1972).

the need for policing: M's incentive to "steal" from his mate F is weaker if M cares about F because a reduction in F's consumption also lowers M's utility. Indeed, caring often completely eliminates the incentive to "steal" and thus the need to police. Thus, at point e in the figure, M has no incentive to "steal" from F because a movement to the right along AB would lower M's utility.⁶² Therefore, if M cares about F sufficiently to transfer commodities to her, F would not need to "police" M's consumption.⁶³ Consequently, marriages with caring would have fewer resources spent on "policing" (via allowances or separate checking accounts?) than other marriages would.

M's income at e exceeds his own consumption because of the utility he gets from F's consumption. Indeed, his income is the sum of his and F's consumption, and equals OB (or OA), the output produced by M and F. Similarly, F's income exceeds her own consumption if she benefits from M's consumption.⁶⁴ Caring makes family income greater than family output because some output is jointly consumed. At point e , all of F's and part of M's consumption would be jointly consumed. Since both e and e' are on the 45° line with mutual and full caring, the combined incomes of M and F would then be double their combined output: all of M's and all of F's consumption would be jointly consumed.

Love and caring between two persons increase their chances of being married to each other in the optimal sorting. That love and caring cannot reduce these chances can be seen by assuming that they would be married to each other in the optimal sorting even if they did not love and care for each other. Then they must also be married to each other in the optimal sorting if they do love and care for each other because love raises commodity output and caring raises their total income by making part of their output a "family" commodity. Hence, their incomes when there is love and caring exceed their incomes when there is not. Consider the following matrix of outputs:

$$\begin{array}{cc} & \begin{array}{cc} F_1 & F_2 \end{array} \\ \begin{array}{c} M_1 \\ M_2 \end{array} & \left[\begin{array}{cc} 8 & 4 \\ (3, 5) & \\ 9 & 7 \\ & (5, 2) \end{array} \right]. \end{array} \quad (1)$$

⁶² A fortiori, a movement along any steeper line—the difference between AB and this line measuring the resources used up in "stealing"—would also lower M's utility.

⁶³ With mutual and full caring, neither mate would have to "police." On the other hand, if each cared more about the other than about himself (or herself), at least one of them, say M, would want to transfer resources that would not be accepted. Then F would "police" to prevent undesired transfers from M. This illustrates a rather general principle; namely, that when the degree of caring becomes sufficiently great, behavior becomes similar to that when there is no caring.

⁶⁴ F's income equals the sum of her consumption and a fraction of M's consumption that is determined by the slope of F's indifference curve at point e . See the formulation in section I of the Mathematical Appendix.

With no caring, this is also the matrix of total incomes,⁶⁵ and M_1F_1 and M_2F_2 would be the optimal sorting if incomes were sufficiently divisible to obtain, say, the division given in parenthesis. With mutual and full caring between M_1 and F_1 , m'_{11} , the income of M_1 , would equal $8 > 3$, and f'_{11} , the income of F_1 , would equal $8 > 5$;⁶⁶ clearly, M_1 would still be married to F_1 in the optimal sorting.

That love and caring can bring a couple into the optimal sorting is shown by the following matrix of outputs:

$$\begin{matrix}
 & F_1 & F_2 & F_3 \\
 M_1 & \left[\begin{array}{c} 10 \\ (4, 6) \end{array} \right. & & \\
 M_2 & \left[\begin{array}{c} 9 \\ \\ (6, 4) \end{array} \right. & 10 & 4 \\
 M_3 & \left[\begin{array}{c} 2 \\ \\ 3 \end{array} \right. & & \left. \begin{array}{c} 10 \\ (5, 5) \end{array} \right]
 \end{matrix} \quad (2)$$

Without love and caring the optimal sorting is M_1F_1 , M_2F_2 , and M_3F_3 , with a set of optimal incomes given in parenthesis. If, however, M_1 and F_2 were in love and had mutual and full caring, the optimal sorting would become M_1F_2 , M_2F_1 , and M_3F_3 because the incomes resulting from this sorting, $m_{12} = f_{21} = k > 6$,⁶⁷ and, say, $m_{21} = f_{21} = 4\frac{1}{2}$, and $m_{33} = f_{33} = 5$, can block the sorting along the diagonal.

Does caring per se—that is, as distinguished from love—encourage marriage: for example, couldn't M_1 marry F_1 even though he receives utility from F_2 's consumption, and even if he wants to transfer resources to F_2 ? One incentive to combine marriage and caring is that resources are more cheaply transferred within households: by assumption, commodities cannot be transferred between households, and goods and time presumably also are more readily transferred within households. Moreover, caring partly results from living together,⁶⁸ and some couples marry partly because they anticipate the effect of living together on their caring.

Since, therefore, caring does encourage (and is encouraged by) marriage, there is a justification for the economist's usual assumption that even a multiperson household has a single well-ordered preference function. For, if one member of a household—the "head"—cares enough about all other members to transfer resources to them, this household would act *as if* it maximized the "head's" preference function, even if the preferences of other members are quite different.⁶⁹

⁶⁵ I abstract from other kinds of "family" commodities because they can be analyzed in exactly the same way that caring is.

⁶⁶ The output of love raises these incomes even further.

⁶⁷ The difference between k and 6 measures the output of love produced by M_1 and F_2 .

⁶⁸ So does negative caring or "hatred." A significant fraction of all murders and assaults involve members of the same household (see Ehrlich 1970).

⁶⁹ For a proof, see section 1, Part II of the Appendix; further discussions can be found in Becker (1969).

Output is generally less divisible between mates in marriages with caring than in other marriages⁷⁰ because caring makes some output a family commodity, which cannot be divided between mates. One implication of this is that marriages with caring are less likely to be part of the optimal sorting than marriages without caring that have the same total income (and thus have a greater total output).⁷¹

Another implication is that the optimal sorting of different traits can be significantly affected by caring, even if the degree of caring and the value of a trait are unrelated. Part I shows that when the division of output is so restricted that each mate receives a given fraction of the output of his or her marriage, beneficial traits are always strongly positively correlated in the optimal sorting. A negative correlation, on the other hand, is sometimes optimal when output is fully divisible. Caring could convert what would be an optimal negative correlation into an optimal positive one because of the restrictions it imposes on the division of output.

For example, assume that a group of men and women differ only in wage rates, and that *each* potential marriage has mutual and full caring, so that the degree of caring is in this case uncorrelated with the level of wage rates; then the optimal correlation between wage rates would be positive, although I showed in Part I that it is negative when there is no caring.⁷² The (small amount of) evidence presented there indicating that wage rates are negatively correlated suggests, therefore, that caring does not completely determine the choice of marriage mates.

3. Polygamy

Although monogamous unions predominate in the world today, some societies still practice polygamy, and it was common at one time. What determines the incidence of polygamous unions in societies that permit them, and why have they declined in importance over time?

I argued in Part I that polyandrists—women with several husbands—have been much less common than polygynists—men with several wives—because the father's identity is doubtful under polyandry. Todas of India did practice polyandry, but their ratio of men to women was much above

⁷⁰ See the proof in section 2, Part II of the Appendix.

⁷¹ See the example discussed in section 2, Part II of the Appendix.

⁷² As an example, let the matrix of outputs from different combinations of wage rates be

$$\begin{matrix} & F_{w_1} & F_{w_2} \\ M_{w_1} & \begin{bmatrix} 5 \\ (5, 5) \end{bmatrix} & \begin{bmatrix} 10 \\ (10, 10) \end{bmatrix} \\ M_{w_2} & \begin{bmatrix} 12 \\ (12, 12) \end{bmatrix} & \begin{bmatrix} 15 \\ (15, 15) \end{bmatrix} \end{matrix}.$$

If outputs were fully divisible, the optimal sorting would be $M_{w_1}F_{w_2}$ and $M_{w_2}F_{w_1}$, since that maximizes the combined output over all marriages. With mutual and full caring in all marriages, the income of each mate equals the output in his or her marriage; these incomes are given in parenthesis. Clearly, the optimal sorting would now be $M_{w_2}F_{w_2}$ and $M_{w_1}F_{w_1}$.

one, largely due to female infanticide.⁷³ They mitigated the effects of uncertainty about the father by usually having brothers (or other close relatives) marry the same woman.

I showed in Part I that if all men and all women were identical, if the number of men equaled the number of women, and if there were diminishing returns from adding an additional spouse to a household, then a monogamous sorting would be optimal, and therefore would maximize the total output of commodities over all marriages.⁷⁴ If the plausible assumption of diminishing returns is maintained, inequality in various traits among men or in the number of men and women would be needed to explain polygyny.

An excess of women over men has often encouraged the spread of polygyny, with the most obvious examples resulting from wartime deaths of men. Thus, almost all the male population in Paraguay were killed during a war with Argentina, Brazil, and Uruguay in the nineteenth century,⁷⁵ and apparently polygyny spread afterward.

Yet, polygyny has occurred even without an excess of women; indeed, the Mormons practiced polygyny on a sizable scale with a slight excess of men.⁷⁶ Then inequality among men is crucial.

If the "productivity" of men differs, a polygynous sorting could be optimal, even with constant returns to scale and an equal number of men and women. Total output over all marriages could be greater if a second wife to an able man added more to output than she would add as a first wife to a less able one. Diminishing marginal products of men or women within each household do not rule out that a woman could have a higher marginal product as a second wife in a more productive household than as the sole wife in a less productive household.

Consider, for example, two identical women who would produce 5 units of output if single, and two different men who would each produce 8 and 15 units, respectively, if single. Let the married outputs be 14 and 27 when each man has one wife, and 18 and 35 when each has two.⁷⁷ Clearly, total output is greater if the abler man takes two wives and the other remains single than if they both take one wife: $35 + 8 = 43 > 14 +$

⁷³ See Rivers (1906). Whether the infanticide caused polyandry, or the reverse, is not clear.

⁷⁴ An optimal sorting has the property that persons not married to each other could not, by marrying, make some better off without making others worse off. I show in Part I (1973) that an optimal sorting maximizes total output of commodities.

⁷⁵ After the war, males were only 13 percent of the total population of Paraguay (see *Encyclopaedia Britannica*, 1973 ed., s.v. "Paraguay"). I owe this reference to T. W. Schultz.

⁷⁶ See Young (1954, p. 124). The *effective* number of women can exceed the number of men, even with an equal number at each age, if women marry earlier than men and if widowed women remarry. The number of women married at any time would exceed the number of men married because women would be married longer (to different men—they would be sequentially polyandrous!). This apparently was important in Sub-Saharan Africa, where polygyny was common (see Dorjahn 1959).

⁷⁷ These numbers imply diminishing marginal products, since $18 - 14 = 4 < 6$, and $35 - 27 = 8 < 12$.

27 = 41. If the abler man received, say, 21 units and each wife received, say, 7 units, no one would have any incentive to change mates.

My analysis implies generally that polygyny would be more frequent among more productive men—such as those with large farms, high positions, and great strength—an implication strongly supported by the evidence on polygyny. For example, only about 10–20 percent of the Mormons had more than one wife,⁷⁸ and they were the more successful and prominent ones. Although 40 percent of the married men in a sample of the Xavante Indians of Brazil were polygynous, “it was the chief and the heads of clans who enjoyed the highest degree of polygyny” (Salzano, Neel, and Maybury-Lewis 1967, p. 473). About 35 percent of the married men in Sub-Saharan Africa were polygynous (Dorjahn 1959, pp. 98–105), and they were generally the wealthier men. Fewer than 10 percent of the married men in Arab countries were polygynous, and they were the more successful, especially in agriculture (Goode 1963, pp. 101–4).

I do not have a satisfactory explanation of why polygyny has declined over time in those parts of the world where it was once more common.⁷⁹ The declines in income inequality and the importance of agriculture presumably have been partly responsible. Perhaps the sex ratio has become less favorable, but that seems unlikely, wartime destruction aside. Perhaps monogamous societies have superior genetic and even cultural natural selection (see the next section). But since more successful men are more likely to be polygynous, they are more likely to have relatively many children.⁸⁰ If the factors responsible for success are “inherited,” selection over time toward the “abler” might be stronger in polygynous than in monogamous societies. I have even heard the argument that Mormons are unusually successful in the United States because of their polygynous past! However, if the wives of polygynous males were not as able, on the average, as the wives of equally able monogamous males, selection could be less favorable in polygynous societies.

The decline in polygyny is usually “explained” by religious and legislative strictures against polygyny that are supposedly motivated by a desire to prevent the exploitation of women. But the laws that prevent men from taking more than one wife no more benefit women than the

⁷⁸ Young (1954, p. 441) says that “in some communities it ran as high as 20–25 percent of the male heads of families,” but Arrington (1958, p. 238) says about 10 percent of all Mormon families were polygynous.

⁷⁹ Polygyny was more common in Islamic and African societies than in Western and Asian ones, although in China and Japan concubines had some of the rights and obligations of wives (see Goode 1963, chap. 5).

⁸⁰ Salzano, Neel, and Maybury-Lewis (1967, p. 486) found evidence among the Xavante Indians of “similar means but significantly greater variance for number of surviving offspring for males whose reproduction is completed than for similar females.” This indicates that polygynous males (the more successful ones) have more children than other males.

laws in South Africa that restrict the ratio of black to white workers (see Wilson 1972, p. 8) benefit blacks. Surely, laws against polygyny reduce the "demand" for women, and thereby reduce their share of total household output and increase the share of men.⁸¹

4. Assortive Mating, Inequality, and Natural Selection

I pointed out in Part I that positive assortive mating of different traits reduces the variation in these traits between children in the same family (and this is one benefit of such mating). Positive assortive mating also, however, increases the inequality in traits, and thus in commodity income, between families. Note that the effects on inequality in commodity and money incomes may be very different; indeed, if wage rates, unlike most other traits, are negatively sorted (as argued in Part I), assortive mating would reduce the inequality in money earnings and increase that in commodity income.

Positive sorting of inherited traits, like intelligence, race, or height, also increases the inequality in these traits among children in different families, and increases the correlation between the traits of parents and children (see proofs in Cavalli-Sforza and Bodmer [1971, chap. 9]). Moreover, positive sorting, even of noninherited traits such as education, often has the same effect because, for example, educated parents are effective producers of "education-readiness" in their children (see Leibowitz [1972] and the papers by her and Benham in this volume). The result is an increase in the correlation between the commodity incomes of parents and children, and thereby an increase in the inequality in commodity income among families spanning several generations. That is, positive assortive mating has primary responsibility for noncompeting groups and the general importance of the family in determining economic and social position that is so relevant for discussions of investment in human capital and occupational position.

Since positive assortive mating increases aggregate commodity income over all families, the level of and inequality in commodity income are affected in different ways. Probably outlawing polygyny has reduced the

⁸¹ An alternative interpretation of the religious and legislative strictures against polygyny is that they are an early and major example of discrimination *against* women, of a similar mold to the restrictions on their employment in certain occupations, such as the priesthood, or on their ownership of property. This hypothesis has been well stated by (of all people!) George Bernard Shaw: "Polygamy when tried under modern democratic conditions as by the Mormons, is wrecked by the revolt of the mass of inferior men who are condemned to celibacy by it; for the maternal instinct leads a woman to prefer a tenth share in a first rate man to the exclusive possession of a third rate." See his "Maxims for Revolutionists" appended to *Man and Superman* (Shaw 1930, p. 220). Shaw was preoccupied with celibacy; he has three other maxims on celibacy, one being "any marriage system which condemns a majority of the population to celibacy will be violently wrecked on the pretext that it outrages morality" (1930, p. 220).

inequality in commodity income among men at the price of reducing aggregate commodity income. Perhaps other restrictions on mating patterns that reduce inequality would be tolerated, but that does not seem likely at present.

Since positive assortive mating increases the between-family variance, it increases the potential for genetic natural selection, by a well-known theorem in population genetics.⁸² The actual amount of selection depends also on the inheritability of traits, and the relation between the levels of the traits of mates and the number of their surviving children (called "fitness" by geneticists). For example, given the degree of inheritability of intelligence, and a positive (or negative) relation between number of children and average intelligence of parents, the rate of increase (or decrease) per generation in the average intelligence of a population would be directly related to the degree of positive assortive mating by intelligence.

Moreover, the degree of assortive mating is not independent of inheritability or of the relation between number of children and parental traits. For example, the "cost" of higher-"quality" children may be lower to more-intelligent parents, and this affects the number (as well as quality) of children desired.⁸³ In a subsequent paper I expect to treat more systematically the interaction between the degree of assortive mating and other determinants of the direction and rate of genetic selection.

5. *Life-Cycle Marital Patterns*

To life-cycle dimensions of marital decisions—for instance, when to marry, how long to stay married, when to remarry if divorced or widowed, or how long to stay remarried—I have paid little attention so far. These are intriguing but difficult questions, and only the broad strokes of an analysis can be sketched at this time. A separate paper in the not-too-distant future will develop a more detailed empirical as well as theoretical analysis.

A convenient, if artificial, way to categorize the decision to marry is to say that a person first decides when to enter the marriage market and then searches for an appropriate mate.⁸⁴ The age of entry would be earlier

⁸² This theorem was proved by Fisher (1958, pp. 37–38) and called "the fundamental theorem of natural selection." For a more recent and extensive discussion, see Cavalli-Sforza and Bodmer (1971, sec. 6.7).

⁸³ For a discussion of the interaction between the quantity and quality of children, see Becker and Lewis in this book.

⁸⁴ This categorization is made in an important paper by Coale and McNeil, "The Distribution by Age of the Frequency of First Marriage in a Female Cohort" (1972). They show that the frequency distribution of the age at first marriage can be closely fitted in a variety of environments by the convolution of a normal distribution and two or three exponential distributions. The normal distribution is said to represent the distribution of age at entry into the marriage market, and the exponential distributions, the time it takes to find a mate.

the larger the number of children desired, the higher the expected lifetime income, and the lower the level of education.⁸⁵

Once in the marriage market, a person searches for a mate along the lines specified in the now rather extensive search literature.⁸⁶ That is, he searches until the value to him of any expected improvement in the mate he can find is no greater than the cost of his time and other inputs into additional search. Some determinants of benefits and costs are of special interest in the context of the marriage market.

Search will be longer the greater the benefits expected from additional search. Since benefits will be greater the longer the expected duration of marriage, people will search more carefully and marry later when they expect to be married longer, for example, when divorce is more difficult or adult death rates are lower. Search may take the form of trial living together, consensual unions, or simply prolonged dating. Consequently, when divorce becomes easier, the fraction of persons legally married may actually *increase* because of the effect on the age at marriage. Indeed, in Latin America, where divorce is usually impossible, a relatively small fraction of the adult population is legally married because consensual unions are so important (see Kogut 1972); and, in the United States, a smaller fraction of women have been married in those states having more-difficult divorce laws (see Freiden [1972] and his paper in this volume).⁸⁷

Search would also be longer the more variable potential mates were because then the expected gain from additional "sampling" would be greater. Hence, other determinants being the same, marriage should generally be later in dynamic, mobile, and diversified societies than in static, homogeneous ones.

People marry relatively early when they are lucky in their search. They also marry early, however, when they are unduly pessimistic about their prospects of attracting someone better (or unduly optimistic about persons they have already met). Therefore, early marriages contain both lucky and pessimistic persons, while later marriages contain unlucky and optimistic ones.

The cost of search differs greatly for different traits: the education, income, intelligence, family background, perhaps even the health of persons can be ascertained relatively easily, but their ambition, resiliency under pressure, or potential for growth are ascertained with much greater difficulty.⁸⁸ The optimal allocation of search expenditures implies that marital decisions would be based on fuller information about more-easily searched traits than about more-difficult-to-search traits. Presumably,

⁸⁵ For a theoretical and empirical study of these and other variables, see Keeley (1974).

⁸⁶ The pioneering paper is by Stigler (1961). For more recent developments, see McCall (1970) and Mortensen (1970).

⁸⁷ These results are net of differences in income, relative wages, and the sex ratio.

⁸⁸ In the terminology of Nelson (1970), education, income, and intelligence are "search" traits, whereas resiliency and growth potential are "experience" traits.

therefore, an analysis of sorting that assumes perfect information (as in Part I) would predict the sorting by more-easily searched traits, such as education, better than the sorting by more-difficult-to-search traits, such as resiliency.⁸⁹

Married persons also must make decisions about marriage: should they separate or divorce, and if they do, or if widowed, when, if ever, should they remarry? The incentive to separate is smaller the more important are investments that are "specific" to a particular marriage.⁹⁰ The most obvious and dominant example of marriage-specific investment is children, although knowledge of the habits and attitudes of one's mate is also significant. Since specific investments would grow, at least for quite a while, with the duration of marriage, the incentive to separate would tend to decline with duration.

The incentive to separate is greater, on the other hand, the more convinced a person becomes that the marriage was a "mistake." This conviction could result from additional information about one's mate or other potential mates. (Some "search" goes on, perhaps subconsciously, even while one is married!) If the "mistake" is considered large enough to outweigh the loss in marriage-specific capital, separation and perhaps divorce will follow.

The analysis in Part I predicts sorting patterns in a world with perfect information. Presumably, couples who deviate from these patterns because they were unlucky in their search are more likely than others to decide that they made a "mistake" and to separate as additional information is accumulated during marriage. If they remarry, they should deviate less from these patterns than in their first marriage. For example, couples with relatively large differences in education, intelligence, race, or religion, because they were unlucky searchers, should be more likely to separate,⁹¹ and should have smaller differences when they remarry. Subsequently, I plan to develop more systematically the implications of this analysis concerning separation, divorce, and remarriage, and to test them with several bodies of data.

6. Summary

The findings of Part II include:

a) An explanation of why persons who care for each other are more likely to marry each other than are otherwise similar persons who do not.

⁸⁹ See the discussion in section 3, Part II of the Appendix.

⁹⁰ The distinction between general and specific investment is well known, and can be found in Becker (1964, chap. 11). Children, for example, would be a specific investment if the pleasure received by a parent were smaller when the parent was (permanently) separated from the children.

⁹¹ If they have relatively large differences because they were less efficient searchers, they may be less likely to separate.

This in turn provides a justification for assuming that each family acts as if it maximizes a single utility function.

b) An explanation of why polygyny, when permitted, has been more common among successful men and, more generally, why inequality among men and differences in the number of men and women have been important in determining the incidence of polygyny.

c) An analysis of the relation between natural selection over time and assortive mating, which is relevant, among other things, for understanding the persistence over several generations of differences in incomes between different families.

d) An analysis of which marriages are more likely to terminate in separation and divorce, and of how the assortive mating of those re-marrying differs from the assortive mating in their first marriages.

The discussion in this paper is mainly a series of preliminary reports on more extensive studies in progress. The fuller studies will permit readers to gain a more accurate assessment of the value of our economic approach in understanding marital patterns.

Mathematical Appendix

I

1. Optimal Sorting⁹²

Given a function $f(x,y)$, I first show that if $\partial^2 f/\partial x\partial y < 0$,

$$\frac{\partial[f(x_2, y) - f(x_1, y)]}{\partial y} \equiv \frac{\partial Q(x_2, x_1, y)}{\partial y} < 0 \quad \text{for } x_1 < x_2. \quad (A1)$$

Since $\partial Q/\partial y = (\partial f/\partial y)(x_2, y) - \partial f/\partial y(x_1, y)$, $\partial Q/\partial y = 0$ for $x_2 = x_1$. By assumption, $(\partial/\partial x_2)(\partial Q/\partial y) = (\partial^2 f/\partial x\partial y)(x_2, y) < 0$. Since $\partial Q/\partial y = 0$ for $x_2 = x_1$ and $\partial Q/\partial y$ decreases in x_2 , $\partial Q/\partial y < 0$ for $x_2 > x_1$; hence (A1) is proved. It follows immediately from (A1) that if $y_2 > y_1$,

$$f(x_2, y_1) - f(x_1, y_1) > f(x_2, y_2) - f(x_1, y_2). \quad (A2)$$

A similar proof shows that if $\partial^2 f/\partial x\partial y > 0$,

$$f(x_2, y_1) - f(x_1, y_1) < f(x_2, y_2) - f(x_1, y_2). \quad (A3)$$

I now am prepared to prove the following theorem: Let $f(x,y)$ satisfy $\partial^2 f/\partial x\partial y > 0$. Suppose $x_1 < x_2 < \dots < x_n$ and $y_1 < y_2 < \dots < y_n$. Then,

$$\left. \begin{aligned} \sum_{j=1}^n f(x_j, y_{i_j}) &< \sum_{i=1}^n f(x_i, y_i) \\ \text{for all permutations} & \\ (i_1, i_2, \dots, i_n) &\neq (1, 2, \dots, n) \end{aligned} \right\}. \quad (A4)$$

⁹² I owe the proofs in this section to William Brock.

Assume the contrary; namely, that the maximizing sum is for a permutation $i_1 \dots i_n$, not satisfying $i_1 < i_2 < \dots < i_n$. Then there is (at least) one j_o with the property $i_{j_o} > i_{j_o+1}$. Therefore,

$$f(x_{j_o}, y_{i_{j_o}}) + f(x_{y_{j_o+1}}, y_{i_{j_o+1}}) < f(x_{j_o}, y_{i_{j_o+1}}) + f(x_{j_o+1}, y_{i_{j_o}}), \quad (A5)$$

by (A3) since $y_{i_{j_o+1}} < y_{i_{j_o}}$. But this contradicts the optimality of i_1, \dots, i_n . QED.

A similar proof shows that if $\partial^2 f / \partial x \partial y < 0$, then

$$\left. \begin{aligned} \sum_{j=1}^n f(x_j, y_{i_j}) < \sum_{i=1}^n f(x_i, y_{n+1-i}) \\ \text{for all permutations} \\ (i_1, i_2, \dots, i_n) \neq (n, n-1, \dots, 1) \end{aligned} \right\} \quad (A6)$$

2. Women Not in the Labor Force

If F did not participate in the labor force,

$$S = Tw_m + T\hat{w}_f + r(l_{pm}, l_{pf})(K_m + K_f) - l_{pm}w_m - l_{pf}\hat{w}_f, \quad (A7)$$

where \hat{w}_f , the "shadow" price of F , is greater than w_f , her market wage rate, unless F is at the margin of entering the labor force,⁹³ and l_{pm} and l_{pf} are the time allocated to portfolio management by M and F , respectively. If the production function for Z were homogeneous of the first degree in time and goods, $Z = S/C(p, w_m, \hat{w}_f, A_f, A_m)$.

Then,

$$\begin{aligned} \frac{\partial Z}{\partial K_i} = C^{-1} \left[r + K \left(\frac{\partial r}{\partial l_{pm}} \frac{\partial l_{pm}}{\partial K_i} + \frac{\partial r}{\partial l_{pf}} \frac{\partial l_{pf}}{\partial K_i} \right) - \frac{\partial l_{pm}}{\partial K_i} w_m - \frac{\partial l_{pf}}{\partial K_i} \hat{w}_f \right] \\ + TC^{-1} \frac{d\hat{w}_f}{dK_i} - SC^{-2}C^f \frac{d\hat{w}_f}{dK_i} - C^{-1}l_{pf} \frac{\partial \hat{w}_f}{\partial K_i} \end{aligned} \quad (A8)$$

$$= rC^{-1} > 0, \quad i = m \text{ or } f \quad (A9)$$

since $C^f = l_f Z^{-1} = (T - l_{pf})Z^{-1}$, $K_m + K_f = K$, and $\hat{w}_f = (\partial r / \partial l_{pf})K$ and $w_m = (\partial r / \partial l_{pm})K$ with an optimal allocation of time. Similarly,

$$\begin{aligned} \frac{\partial Z}{\partial w_m} = TC^{-1} + \frac{TC^{-1}d\hat{w}_f}{dw_m} \\ + C^{-1} \left(\frac{\partial r}{\partial l_{pm}} \frac{\partial l_{pm}}{\partial w_m} K + \frac{\partial r}{\partial l_{pf}} \frac{\partial l_{pf}}{\partial w_m} K - l_{pm} - \frac{\partial l_{pm}w_m}{\partial w_m} - l_{pf} \frac{d\hat{w}_f}{dw_m} - \frac{\partial l_{pf}}{\partial w_m} \hat{w}_f \right) \\ - SC^{-2}C^m - SC^{-2}C^f \frac{\partial \hat{w}_f}{\partial w_m} = l_m C^{-1} > 0, \end{aligned} \quad (A10)$$

and

$$\begin{aligned} \frac{\partial Z}{\partial A_i} = -SC^{-2}C_{a_i} + TC^{-1} \frac{\partial \hat{w}_f}{\partial A_i} - SC^{-2}C^f \frac{\partial w_f}{\partial A_i} - l_{pf} \frac{\partial \hat{w}_f}{\partial A_i} \\ + \text{terms whose sum is zero} \\ = -SC^{-2}C_{a_i} > 0 \quad i = m \text{ or } f, \end{aligned} \quad (A11)$$

⁹³ An earlier draft of this section developed the analysis using the shadow price of F , but contained some errors. I owe the present formulation to H. Gregg Lewis.

if A_l does not directly affect r . Note that equations (A9)–(A11) are exactly the same as those when F does participate—equations (24), (28), and (31).

Then,

$$\frac{\partial^2 Z}{\partial K_f \partial K_m} = C^{-1} \left| \frac{\partial r}{\partial t_{pm}} \frac{\partial t_{pm}}{\partial K_m} + \frac{\partial r}{\partial t_{pf}} \frac{\partial t_{pf}}{\partial K_m} \right| - rC^{-2} C^f \frac{\partial \hat{w}_f}{\partial K_m}. \quad (A12)$$

The first term is positive, but the second one is negative since

$$\frac{\partial w_f}{\partial K_m} > 0, \quad \frac{\partial \hat{w}_f}{\partial K_f} > 0, \quad \left(\text{and } \frac{\partial \hat{w}_f}{\partial w_m} > 0 \right). \quad (A13)$$

A proof of (A13) follows from the derived demand equation for t_f . Of course,

$$\frac{\partial^2 Z}{\partial w_m \partial w_f} = 0. \quad (A14)$$

Moreover,

$$\frac{\partial^2 Z}{\partial K_m \partial A_f} = -rC^{-2} C_{af} - C^{-2} C^f \frac{\partial \hat{w}_f}{\partial A_f}. \quad (A15)$$

The first term is necessarily positive and the second would be nonnegative if $\partial \hat{w}_f / \partial A_f \leq 0$. It can easily be shown that $\partial \hat{w}_f / \partial A_f = 0$ if A_f has a factor-neutral effect on output and $\partial \hat{w}_f / \partial A_f < 0$ if A_f is own-time augmenting. Consequently, there is some presumption that

$$\frac{\partial^2 Z}{\partial K_m \partial A_f} > 0. \quad (A16)$$

The general expression for the cross-derivative of Z with respect to A_m and A_f can be found by differentiating equation (A11). I consider here only the case where the effects are factor-neutral, so that

$$Z = g(A_m, A_f) f(x, t_m, t_f), \quad (A17)$$

or the optimal Z is $Z = gS[K(p, w_m, \hat{w}_f)]$, with

$$g_i = \frac{\partial g}{\partial A_i} > 0, \quad \text{and} \quad g_{mf} = \frac{\partial^2 g}{\partial A_m \partial A_f} > 0. \quad i = m, f. \quad (A18)$$

By substituting into (A11),

$$\frac{\partial Z}{\partial A_i} = Z \frac{g_i}{g} > 0. \quad (A19)$$

Therefore,

$$\frac{\partial^2 Z}{\partial A_m \partial A_f} = \frac{g_m}{g^2} g_f Z + \frac{g_{mf} Z}{g} - \frac{g_m g_f Z}{g^2} = \frac{g_{mf} Z}{g} > 0. \quad (A20)$$

3. Own-Time-Augmenting Effects

By own-time augmenting is meant that the household production function can be written as $Z = f(x, t'_f, t'_m)$, where $t'_f = g_f(A_f)t_f$, and $t'_m = g_m(A_m)t_m$ are the time inputs of F and M in "efficiency" units, and

$$\frac{dg_f}{dA_f} = g'_f > 0, \quad \text{and} \quad \frac{dg_m}{dA_m} = g'_m > 0, \quad (A21)$$

indicates that an increase in the trait raises the number of efficiency units. The optimal Z can be written as $Z = S/C(p, w'_m, w'_f)$, where $w'_m = w_m/g_m$ and $w'_f = w_f/g_f$ are wage rates in efficiency units. Therefore,

$$\frac{\partial Z}{\partial A_m} = -t'_m C^{-1} \frac{\partial w'_m}{\partial A_m} > 0, \tag{A22}$$

since $\partial w'_m/\partial A_m < 0$. Hence,

$$\frac{\partial^2 Z}{\partial A_m \partial A_f} = - \frac{\partial w'_m}{\partial A_m} C^{-1} \left(\frac{\partial t'_m}{\partial A_f} - \frac{\partial w'_f}{\partial A_f} t'_m t'_f S^{-1} \right). \tag{A23}$$

The term outside the parenthesis and the second term in it are positive. The first term in the parenthesis might well be negative,⁹⁴ but Gregg Lewis has shown in an unpublished memorandum that $\partial^2 Z/\partial A_m \partial A_f$ is necessarily positive if the elasticity of substitution between the time of M and F is less than 2.

4. Sorting by Income and Nonmarket Productivity

If M differed only in K_m and F only in A_f , and if all M and F participated in the labor force, $\partial Z/\partial K_m = rC^{-1} > 0$, and

$$\frac{\partial^2 Z}{\partial K_m \partial A_f} = -rC^{-2} C_{a_f} > 0 \quad \text{since } C_{a_f} < 0. \tag{A24}$$

If M differed only in w_m , $\partial Z/\partial w_m = C^{-1} l_m > 0$, and

$$\frac{\partial^2 Z}{\partial w_m \partial A_f} = -C^{-2} C_{a_f} l'_m + C^{-1} \frac{\partial l_m}{\partial A_f}. \tag{A25}$$

The first term on the right is positive, and the second would also be if $\partial l_m/\partial A_f \geq 0$, that is, if an increase in A_f does not reduce the time M spends in the market sector. Even if it does, the cross-derivative is still positive if the first term dominates. In particular, equation (A25) is necessarily positive if the effect of A_f is independent of the input of goods and time. For, if A_f were independent, $C = b(A_f)K(p, w_m, w_f)$. Since $l_m = (\partial C/\partial w_m) Z = (\partial K/\partial w_m) SK^{-1}$, then,

$$\frac{\partial l_m}{\partial A_f} = 0. \tag{A26}$$

II

1. Formally, M (or F) maximizes his utility function

$$U_m = U_m(Z_m, Z_f) \tag{A1}$$

subject to the constraints

$$\left. \begin{aligned} Z_m^0 - C_m &= Z_m \\ Z_f^0 + C_m &= Z_f \\ C_m &\geq 0 \end{aligned} \right\}, \tag{A2}$$

where Z_m^0 and Z_f^0 are the market allocations of output to M and F , and C_m is the amount transferred by M to F . If $C_m > 0$, these constraints can be reduced to a single income constraint by substitution from the Z_f into the Z_m equation:

⁹⁴ There is some evidence suggesting, e.g., that men with more educated wives generally work more hours (see Benham's paper in this book).

$$m_{mf} = Z_{mf} = Z_f^0 + Z_m^0 = Z_m + Z_f, \tag{A3}$$

where Z_{mf} is the output produced by M and F, and m_{mf} is M's income. Maximization of U_m subject to this single income constraint gives

$$\frac{\partial U_m}{\partial Z_m} = \frac{\partial U_m}{\partial Z_f}. \tag{A4}$$

If $C_m = 0$, U_m is maximized subject to the two constraints $Z_m^0 = Z_m$ and $Z_f^0 = Z_f$. The equilibrium conditions are $\partial U_m / \partial Z_m = \lambda_m$, $\partial U_m / \partial Z_f = \mu_m$, where λ_m and μ_m are the marginal utilities of additional Z_m^0 and Z_f^0 , respectively. The income of M would then be

$$m_{mf} = Z_m^0 + (\mu_m / \lambda_m) Z_f^0, \tag{A5}$$

where μ_m / λ_m is the "shadow" price of Z_f to M in terms of Z_m .

Since $\mu_m / \lambda_m < 1$ (otherwise $C_m > 0$),

$$Z_m^0 + \frac{\mu_m}{\lambda_m} Z_f^0 < Z_{mf} = Z_m^0 + Z_f^0. \tag{A6}$$

If $C_m > 0$, the "family" consisting of M and F would act as if it maximized the single "family" utility function U_m subject to the single family budget constraint given by (A3), even if F's utility function were quite different from U_m . In effect, transfers between members eliminate the conflict between different members' utility functions.

2. Total income in a marriage between M and F is

$$m_{mf} + f_{mf} = I_{mf} = Z_{mf} + p_m Z_{mf}^f + p_f Z_{mf}^m,$$

where I_{mf} is the total income in the marriage, Z_{mf}^m and Z_{mf}^f are the outputs allocated to M and F, Z_{mf} ($= Z_{mf}^f + Z_{mf}^m$) is total output, p_m is the shadow price to M of a unit of Z_{mf}^f , and p_f is a shadow price to F of a unit of Z_{mf}^m . Their incomes must be in the intervals

$$\begin{aligned} Z_{mf}^m + p_m Z_{mf}^f &= m_{mf} \leq Z_{mf}, \\ Z_{mf}^f + p_f Z_{mf}^m &= f_{mf} \leq Z_{mf}. \end{aligned} \tag{A7}$$

If $p_m = p_f = 0$ —no caring— m_{mf} and f_{mf} can be anywhere between 0 and Z_{mf} . But if $p_m = p_f = 1$ —mutual and full caring—then $m_{mf} = f_{mf} = Z_{mf}$. And, more generally, if p_m and $p_f > 0$, then

$$\begin{aligned} Z_{mf}^m < m_{mf} &\leq Z_{mf} < I_{mf}, \\ Z_{mf}^m < f_{mf} &\leq Z_{mf} < I_{mf}. \end{aligned} \tag{A8}$$

Consider the following matrix of total incomes:

$$M_1 \begin{bmatrix} F_1 & F_2 \\ 8 & 8 \\ & (4, 4) \end{bmatrix} \\ M_2 \begin{bmatrix} 7 & 7 \\ (3, 4) \end{bmatrix}. \tag{A9}$$

On the surface, both sortings are equally optimal, but this is not so if only M_1 and F_2 have a marriage with caring, say full and mutual, so that $m_{12} = f_{12} =$

4.⁹⁵ The sorting M_1F_2 and M_2F_1 is not as viable as the sorting M_1F_1 and M_2F_2 because income is more divisible between M_1 and F_1 than between M_1 and F_2 .⁹⁶ For if, say, $m_{11} = 4\frac{1}{2}$, $f_{11} = 3\frac{1}{2}$, $m_{22} = 4\frac{1}{2}$, and $f_{22} = 2\frac{1}{2}$, no two persons have an incentive to change mates and marry each other.⁹⁷ On the other hand, since $m_{12} = f_{12} = 4$, unless $m_{21} = 3$ and $f_{21} = 4$, either M_1 and F_1 , or M_2 and F_2 would be better off by marrying each other. If $m_{21} = 3$ and $f_{21} = 4$, M_1 and F_1 , and M_2 and F_2 could be just as well off by marrying each other. Therefore, this sorting is not as viable as the sorting that does not have any marriages with caring.

3. Assume that the gain from marriage of a particular person M is positively related to the expected values of two traits of his mate, as in $m = g(A_1, A_2)$, with $\partial g/\partial A_i = g_i > 0$, $i = 1, 2$. If the marginal costs of search were c_1 and c_2 for A_1 and A_2 , respectively, equilibrium requires that

$$\frac{g_1}{g_2} = \frac{c_1}{c_2}. \quad (\text{A10})$$

The lower c_1 is relative to c_2 , the higher generally would be the equilibrium value of A_1 relative to A_2 , since convexity of the isogain curves is a necessary condition for an internal maximum.

If g_1 and g_2 were invariant when search costs changed to all participants in the marriage market, not an innocuous assumption, then A_1^{\max} and A_2^{\max} would be the equilibrium values of A_1 and A_2 to M when everyone had perfect information about all traits. A reduction in the cost of searching A_1 , therefore, would move the equilibrium value of A_1 to M closer to A_1^{\max} , its value with perfect information.

⁹⁵ The output between M_1 and F_2 also equals four, half that between M_1 and F_1 .

⁹⁶ Or, put differently, the output between M_1 and F_1 exceeds that between M_1 and F_2 .

⁹⁷ F_2 would prefer to marry M_1 , but could not induce M_1 to do so because m_{12} cannot exceed four, the output produced by M_1 and F_2 (see eq. [A7]), which is less than $m_{11} = 4\frac{1}{2}$.

Comment: The Economics of Nonmonetary Variables

William J. Goode

Columbia University

In the history of science researchers have often borrowed theories, analogies, or metaphors from other fields, usually the better-developed ones; in economic terms, they invested their human capital by acquiring new and presumably more-advanced intellectual tools. The most conspicuous borrowing in nineteenth-century social science was the unfortunately imaginary set of the developmental sequences of societies, worked out by anthropology and sociology on the basis of findings from biological evolution. It is less often that scientists in a relatively developed field become restive with its constraints and invade another with the aspiration that their more powerful technical and theoretical tools will solve problems with which the less-developed field has not adequately coped. (This is a challenge I have sometimes hurled at physicists who believe that the findings of sociology are simple-minded.) Over the past two decades, physicists tried this successfully in their contributions to molecular biology, and now economists have been expanding their world by attempting to analyze problems usually brooded over by sociologists, social psychologists, anthropologists, and political scientists.

In this case, interestingly enough, they will encounter colleagues in these fields who have themselves been moving toward economic or quasi-economic analysis of the same phenomena, though to be sure they have not attempted many social analyses of purely market processes. Especially in social psychology and sociology, a small group of theorists have for nearly two decades been working out exchange and allocational problems with the aid of economic ideas.

They have been hampered by their lack of mastery over the tools of economics, their failure to use economics explicitly, and their unwillingness—thus shouldering a burden economists of the past did not wish to carry—to do much theorizing without the facts. In my own case, I have

been trying to understand social-control systems, of which monetary controls form one set, by considering how prestige, force and force threat, and love or affection are accumulated, allocated, or lost. Even in my monograph on divorce, written two decades ago, I was the first both to demonstrate rather fully the inverse relationship between class position and divorce rates (a prediction now made by Becker) and to give an essentially economic explanation for it (see Goode 1956, 1962).

The differing styles of economics and sociology prevent me from making an adequate, brief analysis of Becker's paper. I perceive many tautologies in the paper, but I know that disturbs an economist less than a sociologist. Where economists may be content with certain types of summary indexes, I want to see a large cross-table of percentages. Often I must respond by saying, in answer to an elegant set of mathematical formulas: "It is a beautiful flight but it is not reliable for transportation"; that is, it simply is not true—though that may seem to an economist only a crude answer. An example of this would be Becker's formulation that late marriage and the difficulty of divorce will be correlated. India, for example, has an early age of marriage but very difficult divorce; in the West the main line of division is between Catholic and non-Catholic countries, and in general Catholic countries have almost no divorce but a slightly earlier age at marriage (except for Ireland) than other Western countries; and so on. Of course, to an economist, that may seem at best an unimaginative answer.

In any event, though I am nervous about this invasion, I welcome it. In these few pages I shall look at only a few minor points in Becker's paper to illustrate a somewhat general problem that is often encountered in some explorations by economists.

A major source of weakness in this bold foray into intellectual fields such as sociology and political science that have been trying for generations to create an autonomous noneconomic body of theory is Becker's failure to be daring enough in a critical question, that is, whether any non-monetary variables actually enter the calculations. In fact, as we see in Part I, section 2, both the market and nonmarket variables are aggregated to explain S , which is the "full income, the maximum money attainable if the w_j [i.e., wage rates, of the j th member] are constants."

Although I believe this monetary emphasis may ultimately be a weakness, since the noncontractual structure of exchanges in those sectors may create different behaviors than in the monetary sector,¹ we are left with no assurance that the broadened economic formulation will be adequate, since in fact it has not been fully built into the equations or tested. That

¹ This matter is analyzed at length in my *The Celebration of Heroes: Prestige as a Control System* (forthcoming). Meanwhile, see my *Explorations in Social Theory* (1973).

is, the nonmonetary variables have not been given due weight, but have been monetized.

That monetary equivalent, so tempting to the economist because of his past training, can create both obstacles and factual errors in the analysis. Perhaps I should add a theoretical point, that since the income figure is purely hypothetical or imaginary, it has no greater degree of reality than an apparently subjective factor such as prestige.

To take a minor example, education *is* worth something on both the market and nonmarket exchange systems, and for some purposes a high education can be given an equivalent monetary value, but for a man with *little* education, a wife with very *much* education is *not* worth as much as one would suppose from her money value on the larger market. She is a less-fit wife for him, by nonmonetary calculations—which would still be economic—and he would be supported in this low evaluation by his social circle.

Let us consider at length a more elaborate instance of the problems this failure creates, specifically the lines of reasoning as to whether and why polygyny occurred and why it has usually been the upper-strata men who could or did take advantage of this opportunity. The case is instructive, for the focus on total monetary income leads to factual errors and obscurities, but a general focus on nonmarket economic analysis would clarify some of the facts we do know and leave one question unsolved but at least not obscure.

Becker reasons—correctly, I think—that the total monetary output of monogamy would be higher than that of either polyandry or polygyny. Here the facts are more powerful than he supposes, for whether or not multiple marriage was permitted or encouraged, it has not been common in any society, except under very special conditions. For most men, in all societies over time, polygyny was a statistically unlikely delight.

Obviously, a very low ratio of men to women—caused by successful conquests or by high death rates in war—might permit polygyny for a while, but such conditions are not likely to continue for long. Late marriage for men and early marriage for women raise the chances somewhat, but high mortality among women keeps the total number of woman-years available to each man rather limited. As a consequence, polygyny has not been general even where it is approved.

A purely market explanation may yield this aggregate result even if, as I believe, the reasoning in its favor (see Part II, section 3) is tautological. That is, for most men and women the net market payoff is higher if they pair together rather than make other types of arrangements. However, this reasoning fails at the next step, the explanation for the unequal distribution of women among men even where polygyny is permitted.

At this next step it is necessary to give due weight to strictly non-monetary factors. First, Becker argues, in the same section, that the

“total output over all marriages could be greater if a second wife to an able man added more to output than she would add as a first wife to a less able one.” Again leaving aside the tautological reasoning, this hypothetical result, an increase in total output over all marriages, is essentially empty as a motivating personal or social force in marital decisions.

As in many family decisions, especially those concerning fertility, the individual who decides does not ordinarily concern himself very much with his or her effects on the total output, or society as a whole. Thus, whether total output over all marriages is greater does not motivate people to enter polygyny. If Becker disclaims this as a motive, it is nevertheless obvious from the context that it is viewed as a partial explanation for the distribution of wives in a society.

Nevertheless, if we ask only about higher or lower monetary output of individual marriages, under Becker's assumptions we could expect that able men would try to get a second wife, and their chances of success would be greater than those of less able men.

If nothing else were involved, we could be satisfied that the unequal distribution of wives among the total number of husbands is now explained, at least from the man's side (from the woman's side, the matter is less clear, since under polygyny the husband takes a much larger share than the wife does). Indeed we might then take a step further and predict which economic strata of men would be able to reap this advantage, and even how much they would pay in order to make this investment in a wife, as compared with alternative investments in other types of capital goods. Where, for example, as in much of Central Africa, women may earn money by engaging in small-scale trading, or by tilling their own gardens for the family, we might even be able to put real monetary values on this kind of investment. Moreover, we might even give a better explanation for polyandry than Becker essays, since for the most part such systems occur where the productivity of a given man is low and he has little to invest in a wife, so that a set of men (often brothers) pool their assets in order to enjoy the benefits of one wife.²

However, reasoning from market or monetary total output, whether of the marriage or of all marriages, would miss the importance of women as goods, of women as producers of children, of women as links in purchasing political power, of women as prestige commodities, for monetary reasoning would *not* predict that women would be added where their monetary market output drops or almost stops upon entering marriage. For the stubborn fact remains that multiple marriage is most common among the

² In polyandry these are often brothers, perhaps because this causes less strife; and this alleviates somewhat the problem of producing one's *own* children, which Becker speaks of as a goal in marriage. In any event, whether or not the husbands are brothers, there are usually rules to decide who is sociologically defined as father of a given child.

stratum that gets the least monetary economic output from wives, that is, the most powerful and rich, who are most likely to wall them up and use them as pets, display them as objects that prove the man's high position, use them in power linkages, or even allow them to become parasites who destroy the family wealth by their extravagance. Indeed, it seems reasonable to assert that the higher the class position, the greater the likelihood of some polygynous arrangements, but equally the greater the likelihood that wives represent net losses, except in purely nonmarket goods. Like some other monetary goods, they are in part "bought in special markets"; they are also like such goods in another respect, for instance, furs, polo ponies, Rolls Royces, in that they can be put to other practical uses, and in revolutions sometimes are.

Thus, I am asserting that the reasoning of Becker's paper sometimes goes awry not because it uses economic reasoning in nonmarket areas, but because it fails to take note of the powerful nonmarket variables, whose effects may run counter to those of income variables, however these latter are aggregated. Nor am I as yet sure that market prices or decisions "prove" the total output is greater, whether of general nonmonetary or monetary goods, under these arrangements. As Frederick Knightsaid more than half a century ago, all suppliers are partially monopolies; and we know that these markets are rigged.

In a parallel fashion, the reasoning fails in its attempt to explain the decline of polygyny, that is, the system, in our time and to some extent (although the facts are less clear) over a longer period of time—less clear because we suppose that never in the history of the recent world have most nations been polygynous except under rare conditions.

After flirting with a biological explanation, Becker involves himself in a complex refutation of the claim that women would be better off economically if polygyny were abolished (see Part II, section 3). But his argument, and those he tries to refute, are essentially irrelevant, because all of them still seem to refer to the total market output of the household. It is much more likely that nonmonetary variables shape this decline in polygyny.

Here, as so often in such ventures into why social structures assume their sometimes peculiar shape, we have to ask, not merely about the non-market variables, but *who decides*: who organizes the market. In spite of Becker's graceful reference (n. 81) to Shaw's clever comment, the fact is that we have no evidence from any time or place that *women* ever created a polygynous system. It was *men* who established them, men whose command of force and force threat, prestige, and wealth was paramount in all the societies we know. Doubtless, some women did better economically under such systems, as some women in the United States would if they joined with other women to marry one rich and powerful man, but the system as a whole was not created by women; and wherever new laws,

social movements, revolutionary ideologies, industrialization, urbanization, or religion permit them a free market choice they move away from polygyny.³

I shall not presume, especially in the midst of our contemporary turmoil about who has the right to report what women feel, to know precisely why women object to this system, although I think Becker's earlier, excellent analysis of caring and sharing is relevant. From the known fact that the birth rate under polygyny is less than under monogamy, and from the rules about equal sexual rights in polygynous systems, we must infer that women do not get as much tenderness, affection, and sexual enjoyment in those systems—again, we can use economic theory to reason about these variables, but they are not monetary. Men are in a much more influential position in such systems, and do not have to share as much of the physical output of the family, as indeed they also monopolize much of the prestige-esteem output of the family. They have the power to impose their own prices, in such systems. Thus, when external constraints, including money-market constraints, permit, women decline to enter such unions, even when they are still permitted.

I suggested earlier that after this critical commentary an important puzzle would still remain, although it is one that Becker has really not perceived clearly. It is why some social *systems* are polygamous, and others are not. An explanation of a *system* may not be adequate as an explanation of how some individuals act within it. The possibility that some men will benefit from taking more than one wife does not explain why polygyny occurs here rather than there—and it especially fails if we remember that even in most such systems the typical marriage is a pairing, one man to one woman. Why are such systems approved and viewed as an *ideal*, though few can attain it?

I do not think that any current explanations work well, though doubtless when one does it will be some sort of nonmarket economic analysis. That explanation would have to confront the fact that Western societies, including Latin America, of course, have not been polygynous in any historical period—and of course I do not mean only in the statistical or distributional sense, but in the normative sense, that polygyny was never a socially approved pattern of marriage. Both Japan and China have been mainly monogamous, though in both cases wealthy or powerful men could purchase additional concubines, and in China some of these came close to being recognized as "secondary wives" with specific rights. India, similarly, has been mainly monogamous, though again with a much more open acceptance of concubines. Perhaps most societies outside those great civilizations—but remember they encompass most of the

³ See in this connection the relevant sections of my *World Revolution and Family Patterns* (1963), esp. chap. 4 on Sub-Saharan Africa.

world's population—have been polygamous in the specific but narrow sense that some form of polygamy was normative, something to which men and families aspired.

Whatever those explanatory variables will turn out to be when we locate them, I rather suppose they will not be monetary, and—more important—their impact will be seen primarily among the people who dominated those sociopolitical systems and those marriage markets, that is, the upper-strata men. Upper-strata men created a set of constraints that defined the marriage and family market so as to yield great monetary and nonmonetary advantages to them.

Few good theories in the history of science “fitted the facts” very well, and the “facts” at any given time were partly wrong, so that each explorer must follow his or her temperament rather than a rule which would ground new thinking only on the facts or only on free speculation. I do suspect, however, that a deeper sense of the crude regularities would not hinder these elegant flights, but might pose still more interesting puzzles for the imaginative economist. More fundamentally, I am at least partly persuaded by the work before us that in several fields we are moving toward a single structure of hypotheses about social behavior, hypotheses supported by research in several different fields under different labels; and that two decades hence we may discover we have created a rather impressive body of social science—if not “unified,” at least mutually confirmatory. Becker's explorations are an important step in that progress.