DO HUSBANDS WANT TO BE SHORTER THAN THEIR WIVES?
THE HAZARDS OF INFERRING PREFERENCES FROM MARRIAGE MARKET OUTCOMES

Martha Bailey[12], Ariel Binder[1], and David Lam[12]
University of Michigan
Date of draft: March 24, 2017

CONTACT INFORMATION

[1] University of Michigan, Department of Economics and Institute of Social Research. [2] National Bureau of Economic Research and IZA. Bailey: baileymj@umich.edu and http://www-personal.umich.edu/~baileymj; Binder: ajbinder@umich.edu; and Lam: davidl@umich.edu.

ABSTRACT

Spousal characteristics such as age, height, and income are often used in social science research to infer social preferences. For example, a "male taller" norm has been inferred from the fact that fewer wives are taller than their husbands than would occur with random matching. The fact that more husbands out-earn their wives than vice versa has been used as evidence that husbands prefer that their wives earn less or wives prefer that their husbands earn more. This paper argues that it is difficult and potentially misleading to infer social preferences from marriage market outcomes. We first show how standard economic theory predicts that positive assortative matching on a characteristic in equilibrium is consistent with a wide variety of preferences. This theoretical result is applied to an empirical investigation of income differences between spouses, where a persistent gender gap also exists. Simulations which sort couples positively on permanent income can largely replicate the observed distribution of spousal income differences in US Census data—including the sharp drop-off in mass as the wife begins to out-earn her husband—without assuming a norm against the wife out-earning her husband.

ACKNOWLEDGEMENTS

This project was generously supported by the services and facilities of the Population Studies Center at the University of Michigan (R24 HD041028). Binder acknowledges research support from an NICHD training grant to the Population Studies Center at the University of Michigan (T32 HD007339).

I. INTRODUCTION

Do men prefer to be taller than their wives? Do women prefer to earn less than their husbands? Patterns in the characteristics of spouses are often used by social scientists to infer preferences and social norms. For example, a number of researchers have investigated the extent to which there is a "male-taller" norm in marriage in various populations (Gillis and Avis 1980, Stulp et al. 2013) and the extent to which preferences about height may affect other choices such as inter-ethnic marriage patterns (Belota and Fidrmuc 2010). A large literature looks at income differences between spouses (Winkler 1998, Brennan, Barnett, and Gareis 2001, Raley, Mattingly, and Bianchi 2002), inferring preferences about spousal income differences and the impact of those differences on time allocation decisions, consumption decisions, and marital stability (Schwartz and Gonalons-Pons 2016). Bertrand, Kamenica, and Pan (2015) (hereafter BKP) argue that the drop-off in the density of the wife's share of total spousal income at 50 percent provides evidence of a social norm that husbands strictly out-earn their wives.

This paper argues that it is very difficult and potentially misleading to infer preferences from observed pairings in the marriage market. The challenge comes from the fact that the underlying distributions of spousal characteristics will impose constraints on the possible set of matches. In the context of a Beckerian marriage model, we demonstrate that a wide variety of social preferences over a characteristic are consistent with the same marriage market equilibrium of strict positive sorting on that characteristic. Given this, gaps between the male and female distributions of given characteristics readily give rise to skewed distributions of spousal characteristics. These skewed distributions would lead to many conclusions reached in the literature *even if such social norms do not exist*. For example, a low share of husbands shorter than their wives is consistent with a male-taller norm as well as with men preferring to be *shorter* than their wives.

Our theoretical exploration considers simple models of marital sorting on height. We then generalize the model to apply to a broad range of distributions and preferences and consider its implications for analyses of spousal income differences. Next, we use data on incomes of U.S. husbands and wives in the 2000 Census to discipline several simulation and calibration exercises. These exercises investigate how closely we can replicate observed spousal income differences using simple models of marital sorting and labor supply which do not impose an explicit preference for husbands to out-earn their wives.

Our simulations produce density functions of the share of total family income earned by the wife which closely match the observed density, and which drop sharply across the 50 percent threshold—the point at which the wife begins to out-earn her husband. This pattern is similar to (though less dramatic than) the pattern documented in BKP, who interpret it as evidence of a social norm that the wife should not out-earn her husband.

This discrepancy between the model's predictions and the data could suggest that spouses with near-equal earning potential are manipulating their work efforts so as not to violate such a social norm. Indeed, one key difference between the simulations and the actual data is the presence of a point mass of couples earning exactly identical incomes. BKP report a discontinuous drop in the density function just to the right of this mass point at 50 percent, which is consistent with testing for the presence of a social norm that the wife should not *strictly* out-earn her husband. We obtain a very similar sample and replicate this result almost exactly. However, when we test for a discontinuity just to the *left* of 50 percent, we find a discontinuous *upward* jump, which is consistent with a social norm that a wife should earn *at least as much* as her husband.

Thus the mass point at 50 percent exerts a strong influence on the discontinuity test, even though the mass amounts to only about one quarter of one percent of all couples. Further investigation reveals that couples earning identical incomes are disproportionately joint owners of a business who chose to split revenues equally on their Schedule C tax forms, which means their presence in the data might confound statistical tests aiming to detect social norms relating to spousal income differences. We, therefore, remove the mass point and test for a discontinuity exactly at 50 percent. Omitting these couples, the results show *no evidence* of a discontinuity at 50 percent. In fact, the resulting estimates are similar in magnitude to those generated from our simulations.

Together, the results indicate that observed spousal income differences appear to be largely explained by the gender gap in earnings in the labor market, combined with a tendency for positive sorting on income in marriage. Although it is possible that social norms stigmatize wives who earn more than their husbands, definitive evidence for or against this claim cannot be found by analyzing the distribution of the wife's share in total spousal income. These results suggest that drawing inferences about preferences from observed marriage market outcomes may prove misleading.

II. BECKER'S THEORY OF MARRIAGE AND A SIMPLE MODEL OF SORTING ON HEIGHT

Our theoretical discussion requires that we make predictions about how men and women are sorted in a marriage market. We build on Becker's (1973) economic theory of marriage, which provides well-known predictions about assortative matching on traits. Consider a man *M* and a woman *F* who are considering marriage. We assume they marry if and if only if it makes both better off compared to alternatives. Denote the "output" of the marriage by Z_{mf}. For now assume output can be divided $Z_{mf} = m_{mf} + f_{mf}$, where m_{ij} indicates what man *i* consumes when married to woman *j*. Although this may not be a minor assumption, since "household public goods" like children – or the income difference between spouses – cannot literally be divided in this way, Lam (1988) shows that the model can be applied to the case of household public goods under the assumption of transferable utility. Because output (or utility) can be divided up between husbands and wives, it is possible for men to make offers to potential wives (and women to make offers to potential husbands) of some division of output. This means that a man can in principle use "side payments" to attract a particular wife, and a woman can use side payments to attract a particular husband, making that person better off than he or she would have been with some other partner.

Suppose we have a set of *N* women and *N* men, with marital output between woman *i* and man *j* denoted by Z_{ij}, and we consider all possible sortings of men and women. Drawing on results from other matching models in mathematics and economics, Becker showed that a competitive equilibrium in the marriage market will be the set of assignments that maximizes the sum of output across all marriages. The argument is a standard argument about the Pareto optimality of competitive markets. If an existing set of pairings does not maximize total output, then there must be at least two couples for which we could switch partners and increase total output. Given this, there must be an incentive for the individuals in those couples to capture that increase by a set of new matches and new division of output. This will be illustrated below for a simple example of two couples sorting on height.

Becker applied this very general result to the case of sorting on some trait *A*, where we will consider woman *f* to have a trait value A_f and man *m* to have trait value A_m, where *A* might be height, age, education, income, etc. We will characterize marital output (which might be some measure of joint marital happiness) as a function of the values of *A* for each partner,

$Z_{mf} = Z(A_m, A_f)$. Becker showed that the marriage market equilibrium will be characterized by positive assortative matching on *A* if

$$\frac{\partial Z(A_m, A_f)}{\partial A_m \partial A_f} > 0. \tag{1}$$

There will be positive assortative matching if the cross-partial in (1) is positive, and negative assortative matching if the cross-partial is negative. A positive cross-partial derivative can be interpreted as implying that the value of *A* for the husband and wife are complements, while a negative cross-partial implies they are substitutes. If, for example, having a higher educated husband raises the impact of the wife's education on marital output, then we will tend to see positive assortative matching on education. We draw on the result in (1) extensively below.

Illustrative Model of Sorting on Height

Some of the key theoretical points can be demonstrated with a very simple model of sorting on height in the marriage market. Denote female height by H_f and male height by H_m. Suppose there are two women: F_1 is 60" tall and F_2 is 66" tall. There are two men: M_1 is 66" tall and M_2 is 72" tall. There are two possible pairings, 1) F_1M_1, F_2M_2, which is positive assortative matching on height, and 2) F_1M_2, F_2M_1, which is negative assortative matching on height.

In order to find the marriage market equilibrium, we describe how the heights of couples affect marital utility. Assume that people get utility from their individual consumption and some bonus that comes from being married. The gains from marriage take the very simple form of some bonus *K* (representing, say, economies of scale in consumption or benefits of household public goods) that is offset by some penalty that depends on the height difference between spouses. *K* can be thought of in monetary or consumption units, representing in the simplest example the amount of money the couple saves by being married. The penalty associated with the height difference between couples can also be given a monetary interpretation, representing the amount of additional consumption that would be required to compensate for the disutility from a sub-optimal height difference between spouses.

Now, consider various alternative cases for the loss function associated with the height difference between spouses. For the first case, suppose that all men and women agree that the ideal marriage is one in which the husband is 6" taller than his wife. Couples in which this is not the case experience some loss of utility that increases at an increasing rate as the height difference between spouses increases. A simple example is a quadratic loss function:

$$Z(H_m, H_f) = K - (H_m - H_f - 6)^2 . \tag{2}$$

If the husband is 6” taller than the wife then there is no loss of utility from marriage. If the husband is the same height as the wife then the loss is is $(0\text{-}6)^2 = 36$. As a concrete and very literal example, this could mean that the couple would need an additional $36 worth of consumption to make them as happy as a couple with the ideal height difference. If the husband is 12” taller than the wife then the penalty is $(12\text{-}6)^2 = 36$. With these payoff functions, we can consider the two possible sortings of couples. If the taller man marries the taller woman and the shorter man marries the shorter woman, then each husband is 6” taller than his wife, generating a total marital utility of *2K* (zero penalty in either marriage). If we switch partners, then one couple (same height) has a penalty of 36 and the other couple (taller man and shorter woman) also has a penalty of 36, for a total penalty of 72. Total marital utility is obviously highest with perfect rank-order sorting, and this is the competitive equilibrium we would expect to observe. If we started with the alternative sorting, everyone could be made better off by switching partners. If we observe the perfect rank-order sorting equilibrium and conclude that everyone prefers that husbands are taller than their wives, our inference would be correct.

Now consider a different payoff function in which the ideal couple is one in which the husband and wife have equal heights, with a penalty for height differences that is increasing in the difference:

$$Z(H_m, H_f) = K - (H_m - H_f)^2 \tag{3}$$

With perfect rank-order sorting the total penalty is now 36 + 36 = 72, since each couple is 6” from the ideal height difference. In the alternate sorting we can create one ideal couple of equal heights, generating a penalty of zero. But the other couple (the tall man and the short woman) has a height difference of 12”, creating a penalty of 144 (which we can think of as 72 per spouse). Perfect rank-order sorting produces higher total marital utility (lower total penalties). This follows from the convex penalty function, which penalizes very large differences in height more than small differences.

The logic in terms of a competitive marriage market is as follows: Suppose we began with the sorting in which one couple has equal heights while the other couple has a 12” height difference. The individuals in the mismatched couple, F_1 and M_2 see that they would each be much happier if they could switch partners and have a 6” height difference instead of a 12” height difference. The question is whether F_1 would be able to induce M_1 to switch from F_2 to her. Her penalty would decline from 72 (half of 144) to 18 (half of 36) if she changed partners. The penalty for M_1 would increase from 0 to 18 (half of 36) if he switched partners. Clearly F_1 can more than compensate M_1 for changing, making him a side payment of at least 18, leaving herself better off after the switch. The exact same story can be told for M_2 inducing F_2 to switch to him. Every person will be better off after the re-sorting, so the positive assortative matching equilibrium is the one we should observe.

The resulting sorting of spouses with the preferences in (3) is exactly the same as the sorting with the preferences in (2)–the sorting with positive assortative matching on height. In this second case we would be drawing an incorrect inference if we interpreted the equilibrium as resulting from a preference for men to be taller than their wives. In fact the preference is for equal heights, and the distribution of heights allows for such a case. The reason we do not see it is because creating that match leads to another match of extremely unequal heights.

Taking this case even further, consider a payoff function in which the ideal couple is one in which the wife is 6” taller than her husband, with, once again, a penalty for deviations from the ideal that is increasing in the difference:

$$Z(H_m, H_f) = K - (H_f - H_m - 6)^2 \tag{4}$$

With perfect rank-order sorting the total penalty is 144 + 144 = 288, since each couple is 12” from the ideal height difference. In the alternate sorting the total penalty is 36 + 324 = 360. Once again it is positive assortative matching that produces the maximum total payoff across all marriages. If we started with negative assortative matching, a process of renegotiation analogous to the one just described should lead to a re-sorting. We will therefore expect that positive sorting will be observed as the equilibrium outcome. This then, is the interesting case in which the underlying preferences are that men prefer to be shorter than their wives. We never observe this in the actual marital outcomes, however. The reason is that the convex payoff function pushes the equilibrium toward a sorting that has small average differences between spouses. It is better to have everyone slightly off from the ideal rather than have some couples that are close to the ideal and other couples that are very far from the ideal.

A General Model of Marriage Matching on Characteristics

The conclusions reached in the special case discussed above of two men and two women sorting on heights is very simple to apply much more generally. It is straightforward to show that the model generalizes to cases with a large number of women and men covering a large range of heights.

As long as the distribution of heights for men is shifted to the right from the distribution of heights for women, and as long as the penalty function to height differences between spouses is convex in the height difference (that is, the penalty to a 2" height difference is more than twice as large as the penalty to a 1" height difference), we should observe strong positive assortative matching on height, with the same set of matches implied by a wide range of preferences. Notably, assortative matching on height would occur *even* in cases when men prefer to be strictly shorter than their wives.

For a more general version of the problem, consider a population with *N* men and *N* women, with all men and women getting married. Assume that there is first order stochastic dominance in the distribution of heights, implying that the *i*th man is taller than the *i*th woman for all *i*, where *i* is the rank order, with *i*=1 indicating the tallest man and woman.[1] Suppose that the payoff from man *i*, with height H_{mi}, marrying woman *j*, with height H_{fj}, can be characterized by some general payoff function that includes a convex penalty function in terms of their height difference, which we will define as $g_{ij} = H_{mi} - H_{fj}$:

$$Z(H_{mi}, H_{fj}) = Z(H_{mi} - H_{fj}) = K - f(g_{ij}), \tag{5}$$

with $f''(g) > 0$. The *f(g)* penalty function could include all of the examples given above, implying a preference for taller husbands, a preference for taller wives, a preference for equality of heights, or many other preferences related to the height difference. Note that the first derivatives $\partial Z / \partial H_M$ and $\partial Z / \partial H_F$ can be either positive or negative. For example, if the penalty function is minimized with equal heights, then increasing the man's height will reduce the penalty (and thus increase the payoff from marriage) if the man is shorter than the woman, but will increase the penalty if the man is taller than the woman. Height is not necessarily a "good"

[1] Although this may sound like a strong assumption, it is quite realistic. For example, the income distributions of husbands and wives in the 2000 US census, which we use below, are characterized by first-order stochastic dominance, with the CDF for women lying everywhere above the CDF for men, implying that at any rank in the distribution, male income is higher than female income.

or "bad" trait in the sense of having an unambiguously positive or negative impact on the payoff from marriage. Nonetheless, it is still true that Becker's result about assortative matching in Equation (1) holds based on the cross-partial derivative:

$$\frac{\partial Z(H_m, H_f)}{\partial H_m \partial H_f} > 0. \tag{6}$$

To see this, consider a case in which the husband is shorter than the wife and there is a preference for equality. Increasing his height reduces the height gap and thus increases the total payoff from marriage. The impact of reducing the gap is larger when the initial gap is larger (from the convexity of the penalty function), so the positive impact of increasing his height is increasing in the height of the wife. Conversely, the impact of the husband's height is negative when he is taller than his wife, but this effect will be smaller when the initial gap is smaller. So the negative impact of husband's height becomes less negative as the wife's height increases, once again implying a positive cross-partial. This implies that there will be positive assortative matching on height.

In fact, we can show that there will be strict positive assortative matching in the sense that *i*th tallest man will be matched to the *i*th tallest woman, for all *i*. To prove this, compare that sorting with some alternative set of partners for the *i*th tallest man and woman. Consider the *i*th ranked man and woman and the *j*th ranked man and woman, $i<j$, with heights H_{mi},$>H_{fi}$, $H_{mj}>H_{fj}$, and $H_{mi}>H_{mj}$ We do not make any assumption about the ranking of H_{fi} and H_{mj}, the tallest woman and the shortest man. There are two possible pairings of these two men and two women, with the following total payoffs from the two marriages:

Payoff from pairing *A*: $Z(H_{Mi} - H_{Fi}) + Z(H_{Mj} - H_{Fj})$.

Payoff from pairing *B*: $Z(H_{Mi} - H_{Fj}) + Z(H_{Mj} - H_{Fi})$.

Pairing *A* is positive assortative matching, with the tallest man married to the tallest woman. Pairing *B* is negative assortative matching, with the tallest man married to the shortest woman. Given the height rankings, it must be the case that $g_{ji} < g_{ii} < g_{ij}$ and $g_{ji} < g_{jj} < g_{ij}$. That is, the height gap between man *i* (the tallest man) and woman *i* (the tallest woman) must be larger than the height gap between woman *i* and man *j*, which must in turn be smaller than the gap between man *i* and woman *j* (the shortest woman). And the gap between woman *i* and man *j* must be smaller than the gap between woman *j* and man *j*, which must be smaller than the gap between between woman j and man i. In other words, g_{ji} is the smallest gap, g_{ij} is the largest gap, with g_{ii} and g_{jj} intermediate.

The sum of the gaps for Pairing A can be written and rearranged as:

$$(H_{Mi} - H_{Fi}) + (H_{Mj} - H_{Fj}) = (H_{Mi} - H_{Fj}) + (H_{Mj} - H_{Fi})$$
$$g_{ii} + g_{jj} = g_{ij} + g_{ji}$$

This simply states that the sum of the height gaps must be the same under either sorting. Since g_{ji} is the smallest gap and g_{ij} is the largest gap, it follows from concavity of the payoff function (convexity of the penalty function) that

$$Z(g_{ii}) + Z(g_{jj}) > Z(g_{ij}) + Z(g_{ji})\,. \tag{7}$$

Equation (7) gives the key result. It states that for men and women at any two ranks in the distribution, we will always get higher total payoff with perfect positive sorting, matching men and women with the same rank. More generally, if we take any two men and women who are not matched with positive sorting, we will always get higher total payoff if we rearrange them with positive sorting. Given this, it follows that the only sorting for the full distributions that will maximize total payoffs is perfect rank order sorting, given stochastic dominance of the distributions and given a concave payoff function.

These results obviously extend to differences in other characteristics such as income. Income has the additional complication that, unlike height, it is not an exogenous trait. The incomes of husbands and wives will be affected by decisions about labor supply and investments in human capital. But assortative matching also plays a fundamental role in determining the income difference between spouses. Our results imply that if the male income distribution stochastically dominates the female income distribution, and if there is a tendency for strong positive assortative matching on income, it will tend to be relatively rare for women to earn more than their husbands. This tendency will exist even if the underlying norm is to have equal incomes between husbands and wives, or even if the norm is for wives to earn more than their husbands. Of course many other factors may lead to strong positive assortative matching on income or income-related characteristics. Lam (1988), for example, demonstrates that there will tend to be positive assortative matching on income whenever the gains from marriage results from household public goods, such as children. Individuals may have no preferences regarding the difference between spousal incomes at all, but the equilibrium set of pairings in the marriage market may look as if there is a norm that a husband should earn more than his wife.

III. IMPLICATIONS FOR SPOUSAL INCOME DIFFERENCES

The theoretical results of Section II indicate that the presence of arbitrary social preferences over spousal income differences creates a strong tendency for positive sorting on income in marriage market equilibrium. Given a sizable gender gap in earnings, this positive sorting result indicates that it should be relatively rare for wives to out-earn their husbands, regardless of underlying preferences. We now demonstrate that if we calibrate the male and female income distributions according to Census data, and assume positive sorting on potential income, we can very closely replicate the empirical distribution of the wife's share of earned household income. We use BKP's results to guide our investigation.

Simulated Distributions

BKP assemble a strong case that "women are bringing personal glass ceilings from home to the workplace" (p. 574). Both in Census and administrative data, they find that the density function of the share of total household income earned by the wife drops sharply and discontinuously (based on McCrary's (2008) statistical test) at 50 percent, the point at which the wife starts to out-earn the husband. This discontinuity, they argue, is evidence of a social norm that women not out-earn their husbands. They supplement these discontinuities with other findings: in marriage markets in which women are likelier to out-earn men, marriage rates are lower; when the wife's full earning potential exceeds her husband's she is less likely to work full-time; and when the wife does out-earn the husband the marriage is less stable and likelier to end in divorce.

We begin by replicating BKP's key finding using a sample of couples drawn from the 2000 U.S. Census 5-percent sample (Ruggles et al. 2015). Following BKP, we restrict the sample to couples aged 18-65, process earned income variables following the procedure outlined in the paper's main text and appendix,[2] and keep only couples in which both spouses report positive income. Figure 1 displays two 20-bin histograms of the distribution of the share of total household income earned by the wife: the one presented in BKP and our replication. Following

[2] Among other things, the procedure removes the unnatural spike of couples reporting exactly identical incomes.

BKP, we apply a local linear smoother to the histogram bins, allowing for a break in the smoothed distribution at 50 percent. The two distributions (smoothed and unsmoothed) look very similar, and both display a sharp reduction in probability mass to the right of 50 percent.

For the purposes of our simulation, we further restrict the sample to relatively young couples (aged 18-40) without children. This additional restriction is motivated by the fact that women disproportionately reduce their working hours or exit the labor force to raise young children, and re-enter the workforce with lower earnings potential (Mincer and Ofek 1982, Hotchkiss and Pitts 2007, Attanasio, Low, and Sanchez-Marcos 2008, Bertrand, Goldin, and Katz 2010). Our simple treatment of the income process and marital sorting will not address the dynamics of household fertility and how they interact with labor supply decisions.[3]

Our final sample consists of 109,569 dual-earning couples, and for each couple we calculate the share of family income earned by the wife. Figure 2 plots the observed distribution of the wife's share of family income. The main difference between this distribution and the full sample distribution is the reduced presence of couples where the wife earns below 25 percent of total family income, which likely reflects the impact of children on the wife's labor supply and earned income. Otherwise, the distribution based on our sub-sample broadly matches the full sample distribution and, importantly, also exhibits a sharp drop at 50 percent. According to the McCrary test the estimate of the sharp drop is 12.4 percent, with a standard error of 1.8 percent. This plot and discontinuity estimate serve as benchmarks for our subsequent empirical exercises.

In our first simulation, we randomly match men and women in our sample into couples. Figure 3 displays a smoothed distribution of the wife's share of family income based on random matching, again allowing for a break at 0.5, overlaid on the observed distribution. The distribution generated by random matching is not too dissimilar from the observed distribution. Even this overly simple and surely unrealistic sorting process produces a mode of wife's share of income around 0.42 and a drop-off in mass above that point. Notably, significantly fewer wives slightly out-earn their husbands than vice versa: 0.5 corresponds to the 70th percentile of the

[3] The effect of the presence of children and marital tenure on the observed distribution of the wife's share of total spousal income can be readily inferred from BKP's Appendix Figures A.1 and A.2. These figures show that the leftward spike near 0 is much more pronounced in couples with children versus couples without, and as marriage tenure increases. Accurately reproducing this spike would require a dynamic labor supply model with a careful specification of the earnings process and how it is affected by a period of reduced or non-participation. This exercise is beyond the scope of this paper.

distribution of wife's share of earned income. This exercise demonstrates that the prevailing male and female income distributions exert a strong influence on spousal income differences.

Our next exercise investigates whether improvements over random sorting can be made by sorting couples positively on income. Reflecting the discussion of Section II and Lam (1988), positive sorting could arise from a variety of explicit preferences over spousal income differences, the presence of household public goods, or positive sorting on other characteristics related to income (such as education).

In our Census sample, observed male log income y_i^m is distributed nearly normally with mean 10.35 and standard deviation 0.75. Female log income y_i^f also follows a roughly normal distribution with mean 10.00 and standard deviation 0.87. We use these log-normal parameters to simulate 100,000 male incomes Y_i^m and 100,000 female incomes Y_i^f. Next, we create couples by matching individuals not according to observed income rank, but rather the rank of observed income perturbed with noise. That is, for each individual of gender g we assign $W_i^g = Y_i^g + u_i$, where u is normally distributed white noise, and pair up males and females according to their rank of W. This is consistent with at least two interpretations. One interpretation is that couples are perfectly sorted on the basis of permanent incomes and the white noise represents transitory income shocks. A second is that men and women care about other characteristics as well as income, or that marital matching is imperfect, for example due to the presence of search frictions.[4] Under the latter interpretation, equilibrium sorting on observed income plus noise is the reduced form of a more complicated and unspecified matching process.

Figure 4 displays a simulated distribution of the wife's share of family income based on this very simple model, with the standard deviation of u set to 16,000, overlaid on the actual distribution. Other than failing to replicate the extreme left tail of the distribution, the simulated distribution is remarkably similar to the actual. Importantly, the simulated distribution also exhibits a sharp drop in mass across the 50 percent threshold.

We next test whether the drop is actually discontinuous in nature via a Monte Carlo version of the McCrary (2008) test. We simulate 500 distributions of independently from the data-generating process and test for a discontinuity at 50 percent in each distribution. The

---

[4] Common specifications of earnings processes (see, e.g., Moffit and Gottschalk, 2002) assume transitory white-noise shocks enter log-linearly, rather than linearly as we have assumed. This might lead one to prefer the second interpretation over the first, although the first interpretation is simpler.

average point estimate is a 2.6 percent drop in mass, and the average *t* statistic is -1.1. Thus we cannot reject the null hypothesis that our simple data-generating process produces a distribution that is smooth at 50 percent. In a sense this should be expected, as our model does not specify any discrete choices or impose any discontinuous functional forms. Nonetheless, it is important to stress the point made by Figure 4: given the underlying male and female income distributions, the observed distribution of the share of family income earned by the wife is largely consistent with positive sorting on observed income plus noise. As Section II indicates, the observed matching could be consistent with a wide variety of underlying preferences. It could be based on a desire for equality in spousal incomes, a preference for wives to be richer than their husbands, or gains from marriage related to household public goods that lead to positive sorting on permanent income (i.e. with no explicit preference at all for equal or unequal spousal incomes).

In the next sub-section we closely examine the discontinuity in the observed data and the failure of our model to replicate it. Before doing so, however, we address one key shortcoming of the previous exercise: the implicit assumption that observed female income is an exogenous trait. Even despite focusing on a sample of childless couples of prime working age, this is likely an unrealistic assumption. For a variety of reasons, including specialization incentives, or the very social norm BKP argues exists, the maximization of household objectives may lead the wife not to work full-time. To address this, we endogenize the wife's income via a simple labor supply model and explore the model's predictions about the distribution of *wifeshare*.

We assume that, for a given male *m* and female *f*, the match output function is given by

$$Z_{mf} = Z\left(Y_m, Y_f, P\right) = \frac{C^{1-\gamma}}{1-\gamma} - \psi P, \tag{8}$$

with $C = 0.61(Y_m + Y_f P)$, where Y_m and Y_f denote each spouse's permanent income, P is the wife's labor supply decision (constrained to be in the unit interval), γ is the CRRA parameter, and ψ is the disutility incurred by the household if the wife works. This specification of household utility has been used in recent work investigating determinants of wives' labor supply (e.g. Attanasio et al., 2008). It assumes household consumption of earned income is a public good with congestion; the 0.61 is a McClements scale calibration capturing consumption economies of scale in marriage.[5] We depart from the framework of Section II and now assume

[5] To illustrate, suppose P=1 and Y_m=Y_f. Then the couple enjoys a higher level of joint consumption in marriage than either member would as single.

fully *non*-transferable utility. With this assumption positive sorting on permanent income occurs in marriage market equilibrium so long as each member's permanent income positively affects match output.[6] It is trivial to show that this holds here (regardless of the wife's eventual labor supply decision). Assuming that each individual's income from full-time work (full income) in a given period is the sum of his or her permanent income and a transitory shock, positive sorting on full income plus noise will arise in equilibrium.

After marriage, the wife takes her husband's and her own full income as given and chooses $P \in [0,1]$ to maximize the above utility function. Assuming an interior solution, she optimally chooses

$$P^* = \frac{\frac{1}{0.61}\left(\frac{\psi}{0.61Y^f}\right)^{-\frac{1}{\gamma}} - Y^m}{Y^f}; \qquad (9)$$

if P^* lies outside of the unit interval, the appropriate corner solution applies.

To use the above model to draw valid conclusions about the distribution of the share of spousal income earned by the wife in marriage market equilibrium, we must reasonably calibrate it. Outside of the calibration we impose $\gamma = 1.5$, a standard value estimated in the macro literature. We assume log-normally distributed full incomes and allow the work disutility parameter, ψ, to be heterogeneous in the population and negatively correlated with Y_f.[7] The model in total contains 8 parameters, which we calibrate by targeting 8 moments in our observed data: the means and standard deviations of male and female log observed income, the observed mean gender earnings ratio conditional on earning positive income (P^*>0), the observed mean gender earnings ratio conditional on full-time work (defined in the data as at least 1600 hours worked in the last calendar year; defined in the model as P^*>0.95), the female employment rate (defined in the data as the share of wives working positive hours in the last calendar year), and the female full-time employment rate.

Table 1 summarizes the calibration—overall the model does a very good job of replicating the targets in the data. Notice also that we did not explicitly target any moment

---

[6] Starting from perfectly positive sorting, it is easy to show that no two individuals can become better off by dissolving their current matches and matching with each other. The inability of individuals to make transfer payments means we no longer need the cross-partial assumption on the match output function to generate positive sorting on the given trait in marriage market equilibrium.

[7] Imposing a negative correlation, as has been estimated in the literature (Eckstein and Lifshitz 2011), ensures that positive sorting on potential income is not disturbed.

related to spousal income differences in the calibration, so we are not knowingly biasing the model in favor of our purpose.

With the calibrated model we can now simulate the distribution of the wife's share of spousal income, as displayed in Figure 5.[8] The simulated distribution again matches the actual distribution very closely. Observe that because of the endogenous labor supply decision, some wives choose to work very few hours, and so the region of the distribution below 25 percent is much better replicated here than in Figure 4. On the other hand, this simulation produces a distribution in which slightly too many wives outearn their husbands relative to the actual data. Nonetheless, the overall shape of the distribution, including the sharp drop in mass across the 50 percent threshold, is accurately replicated.[9]

Overall these empirical simulations indicate that gender differences in the wage structure combined with any match utility function that delivers positive sorting on permanent income can sufficiently explain the surprisingly low incidence of wives out-earning their husbands. However, our simple models are not able to account for the discontinuous nature of the distribution of wives' relative income across the equality threshold.

A Closer Look at the Drop-Off in Mass at 50 Percent

BKP estimate the sharp drop in mass at 50 percent to be significantly discontinuous in a variety of Census samples as well as in administrative data. Their analysis is complicated by the fact that unlike in our simulated data, the actual data contain a point mass of couples with identical earned incomes. This point mass is quite small in the administrative data (between 0.2 and 0.3 percent of all couples), and rather large in the Census samples (around 3 percent of all couples, even after removing individuals with imputed income and couples where both spouses' incomes are top-coded).[10] Without the point mass, the straightforward way to proceed would be to test for the discontinuity of the distribution exactly at 50 percent, and interpret a negative and

[8] The simulation uses a sample size of 120,000 men and 120,000 women. Since around 90 percent of wives choose to work, an initial sample of 120,000 returns around 108,000 dual-earning couples, which closely matches the sample size observed in the 2000 Census.

[9] Performing the same Monte Carlo version of the McCrary test as in the previous exercise we also estimate a small and statistically insignificant drop-off in mass at 50 percent. In the next sub-section we take a closer look at BKP's discontinuity results and illustrate that they are driven by the small spike of couples earning identical incomes, which our simple simulations do not take into account.

[10] This point mass has been removed in Figures 2-4 to maximize comparability to the simulations as well as to what BKP present in their Figure III.

significant finding as evidence that couples are manipulating their earnings so as not to have the wife out-earn the husband. The presence of the point mass presents a challenge. Consistent with the wife *strictly* out-earning the husband to be a violation of a social norm, BKP test for a discontinuity just to the right of 0.5. The negative and significant result, combined with the presence of the spike, might suggest that a non-trivial portion of couples manipulate their earnings so that the wife earns *the same as or less than* her husband.

While this treatment of the data is sensible and in line with the hypothesis test BKP wished to conduct, it is unclear whether asymptotic inference based on the McCrary test is robust when there is a point mass close to the supposed breakpoint.[11] Like a non-parametric regression discontinuity design, the test involves local linear smoothing of a finely-binned histogram on either side of the supposed breakpoint, and asymptotic inference is based on the size of the bins shrinking to zero at the correct rate as the number of observations increases to infinity. In BKP's application of the test, for a small bin size, the bin immediately before the breakpoint will (by virtue of containing the point mass) be much taller than the bin immediately after the breakpoint. This could exert undue influence on the discontinuity estimate, especially if a small bin size and bandwidth is used to perform the test.

To investigate the sensitivity of the discontinuity test to the presence of the point mass, we replicate BKP's administrative sample and analysis. The data used are Survey of Income and Program Participation (SIPP) data linked to the Social Security Administration's (SSA's) detailed earnings records.[12] BKP construct a sample of administrative earnings data for all dual-earning couples aged 18 to 65 observed in the first year they were in the SIPP panel. They consider SIPP panels 1990 through 2004. We construct a sample according to the same conditions, but include the 1984 and 2008 SIPP panels as well, which are available in the most recent version of the SIPP/SSA data product. We obtain a sample of around 83,000 couples—

[11] In fact, one of the assumptions required to perform the test is that the distribution is continuous everywhere except possibly at the supposed breakpoint (McCrary 2008).

[12] The particular data we and BKP use come from a pre-linked and cleaned Census Bureau data product called the Gold Standard File (GSF). Users work with synthetic versions of the data remotely and then have Census run final programs internally on the actual GSF, subject the output to a disclosure review, and then release the output. More information can be found in Benedetto, Stinson, and Abowd (2013) and here: http://www.census.gov/programs-surveys/sipp/guidance/sipp-synthetic-beta-data-product.html.

about 9,500 more than in BKP's sample.[13] Despite using a slightly different sample, our distribution of the wife's share of total spousal income is virtually identical to BKP's, as illustrated in Figure 6.

In our replicated sample, 0.21 percent of all dual-earning couples earn identical incomes, compared to 0.26 percent in BKP's sample. Using our sample we now perform 3 different versions of the McCrary test for a discontinuity in the distribution at 50 percent, based on three different treatments of the point mass: keeping the point mass and testing for a discontinuity at .500001, keeping the point mass and testing for a discontinuity at .499999, and deleting the point mass and testing for a discontinuity exactly at .50. For each version we use 4 different sets of tuning parameters to gauge the sensitivity of the test results. McCrary's test procedure involves an algorithm that automatically chooses a bin size for the histogram and a bandwidth within which to apply the local linear smoother to the histogram. McCrary (2008) recommends using a smaller bandwidth than the automatic one (around half the size) to conduct robust asymptotic inference. We consider the automatically selected bandwidth, which in this case is around .084; and then bandwidths of .045, .023, and .011. Especially the last bandwidth may be too narrow for optimal statistical inference, but using successively smaller bandwidths allows us to gauge the sensitivity of the test to the presence of the point mass (which becomes increasingly dominant as the bandwidth shrinks).

Table 2 reports the discontinuity estimates, which equal the estimated log increase in the height of the density function as one travels from just to the left of the supposed breakpoint to just to the right. A negative number thus indicates a sharp drop and a positive number indicates a sharp gain. Bolded estimates are statistically significant at the 5 percent level; italicized estimates are significant at the 1 percent level. Standard errors appear below estimates in parentheses.

The first version of the test replicates BKP's choice of retaining the point mass of couples and testing for a discontinuity just to the right of 50 percent. With the standard bandwidth and bin size, we estimate that the density function drops by a statistically significant 12.4 percent across the threshold. This is remarkably similar to BKP's reported estimate of a 12.3 percent drop in their very similar sample (reported on p. 576). Observe that as the bandwidth shrinks, the

[13] BKP report a sample size of 73,654, although it is unclear whether this number refers to all couples in their sample or all dual-earning couples. We have obtained BKP's do-files and in future work will exactly replicate their sample.

estimate of the sharp drop rises in magnitude, such that with the smallest bandwidth we estimate a 57.5 percent drop—over 4 times as large as the first estimate. This suggests that the point estimates are sensitive to the existence of the point mass.

When we retain the point mass and test for a discontinuity just to the *left* of 50 percent, we find the exact opposite result: the density function jumps discontinuously *upward*. Once again, the estimate starts out reasonably small (6.4 percent), and becomes very large (45.1 percent) as the bandwidth shrinks. The finding of a sharp gain in the distribution just to the left of 50 percent would suggest that couples manipulate earnings so as not to have the husband strictly out-earn his wife (i.e. there is missing mass just to the left of 50 percent). This is nearly opposite to the social norm that the wife should not out-earn her husband, which is suggested by the first version of the results.

The third column of results derive from deleting the point mass and testing for a discontinuity exactly at 0.50. Two features stand out. First, while the estimates are negative they are no longer statistically significant—moreover, the estimate based on the standard bandwidth matches closely the estimates generated by performing the test with the standard bandwidth on our simulated data. Secondly, the estimates do not rise appreciably in magnitude or statistical significance as the bandwidth shrinks, because the point mass is no longer present. Overall these results illustrate the undue influence of the point mass of couples earning identical incomes.

Therefore, if we ignore the one quarter of one percent of couples earning identical incomes, the conclusions that i) gender income gaps dictate the distribution of spousal income differences in marriage, and that ii) the actual distribution of spousal income differences could be consistent with a wide variety of underlying social preferences, are strongly supported by the data and empirical simulations. A related conclusion, stemming directly from the estimates in Table 2, is that in light of the point mass, one should not use a discontinuity test at the point that the wife begins to out-earn the husband to robustly infer social preferences.

In light of these conclusions, it is worth exploring why the point mass exists in the first place, and what it means to remove it from the sample. For example, the existence of the point mass could indicate a social preference, in the population or a certain sub-population, for strict equality of spousal incomes.

Further exploration of the 2000 Census data reveals three interesting facts about the couples who report identical incomes in comparison to the full sample.[14,15] First, compared to a couple reporting different incomes for husband and wife, when a couple reports identical earned incomes, it is far likelier that both the husband and wife report being self-employed. Second, over half of the couples who report identical incomes report identical *business* incomes, while less than 20 percent of all dual-earning couples report positive business income for either spouse. Third, 21 percent of couples in which each spouse earns positive business income report identical business incomes. Overall these facts suggest that couples with identical administrative earnings records are disproportionately couples whose sole source of earned income is a jointly owned business. For these couples, income declarations on tax forms are easily influenced by tax or Social Security incentives, or may reflect the fact that both owners have actually contributed equally to the business operation. The earnings of such couples are clearly not reflective of standard income processes. Thus there is a case to be made for excluding couples with identical administrative incomes to draw valid inferences about spousal income differences for the general population.

IV. CONCLUSION

Our theoretical and empirical results demonstrate that it is very difficult, and potentially quite misleading, to infer preferences about spousal characteristics from the observed distribution of differences in spousal traits. Actual marriage market outcomes are affected not just by preferences, but also by the underlying distribution of traits on both sides of the market. If men are taller or richer than women on average, any preferences that lead to positive assortative matching will produce equilibrium sortings in which it is relatively rare for women to be taller or richer than their husbands. As we have shown, even a preference for men to be shorter than their wives will plausibly lead to an equilibrium in which most men are taller than their wives.

[14] The Gold Standard File provides very little occupational information about the couples, which is why we use the Census for this exploration. It is important to keep in mind that the point mass of couples with identical incomes is over 10 times as large in the Census data, due to rounding of reported income as well as possible reporting biases. That is, many couples who report identical incomes in the Census data do not have identical administrative earnings records. However, it is reasonable to assume that couples who report identical incomes are (much) likelier than those who do not to have identical administrative records.

[15] This analysis is incomplete and will be expanded in the final version of the paper.

In the final version of this paper we will provide a more complete analysis of the theoretical results describing how preferences interact with the underlying distributions of traits to produce equilibrium marriage market outcomes. These include results using very general continuous distributions of traits and a wide range of preferences. These results are entirely consistent with the simple models presented above. We will also use the theoretical results to discuss the ways in which preferences could be correctly inferred, including the use of data on actual resource distribution within couples or the use of information on individuals who end up unmarried in the marriage market. We will also provide additional simulations based on actual distributions of income in the U.S. All our analyses support the results above, which show that we can generate distributions of the spousal distribution of income that are almost identical to the observed distributions without imposing any preferences regarding the relative incomes of husbands and wives. Using these results, we will provide a more detailed discussion of the implications of the results for research on gender gaps within marriage and the role of social norms in driving the allocation of resources within households.

V. REFERENCES

Attanasio, Orazio, Hamish Low, and Virginia Sanchez-Marcos. 2008. "Explaining Changes in Female Labor Supply in a Life-Cycle Model." *American Economic Review* 98 (4):1517-1522.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81 (4):813-846.

Belota, Michele, and Jan Fidrmuc. 2010. "Anthropometry of love: Height and gender asymmetries in interethnic marriages." *Economics and Human Biology* 8 (3):361-372.

Benedetto, Gary, Martha Stinson, and John M. Abowd. 2013. "The Creation and Use of the SIPP Synthetic Beta." Washington, DC: United States Census Bureau. Retrieved from: https://www.census.gov/content/dam/Census/programs-surveys/sipp/methodology/SSBdescribe_nontechnical.pdf.

Bertrand, Marianne , Claudia Goldin, and Lawrence F. Katz. 2010. "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2 (3):228-55.

Bertrand, Marianne, Emir Kamenica, and Jessica Pan. 2015. "Gender Identity and Relative Income within Households." *Quarterly Journal of Economics* 130 (3):571-614.

Brennan, Robert T., Rosalind Chair Barnett, and Karen C. Gareis. 2001. "When She Earns More than He Does: A Longitudinal Study of Dual-Earner Couples." *Journal of Marriage and Family* 63 (1):168-182.

Eckstein, Zvi, and Osnat Lifshitz. 2011. "Dynamic Female Labor Supply." *Econometrica* 79 (6):1675-1726.

Gillis, John Stuart, and Walter E. Avis. 1980. "The Male-Taller Norm in Mate Selection." *Personality and Social Psychology Bulletin* 6 (3):396-401.

Hotchkiss, Julie L., and M. Melinda Pitts. 2007. "The Role of Labor Market Intermittency in Explaining Gender Wage Differentials." *American Economic Review* 97 (2):417-421.

Lam, David. 1988. "Marriage Markets and Assortative Mating with Household Public Goods: Theoretical Results and Empirical Implications." *Journal of Human Resources* 23 (4):462-487.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142 (2):698-714.

Mincer, Jacob, and Haim Ofek. 1982. "Interrupted Work Careers: Depreciation and Restoration of Human Capital." *Journal of Human Resources* 17 (1):3-24.

Raley, Sara B., Marybeth J. Mattingly, and Suzanne M. Bianchi. 2002. "How Dual are Dual-Income Couples? Documenting Change from 1970 to 2001." *Journal of Marriage and Family* 68 (1):11-28.

Ruggles, S., and et al. 2015. *Integrated Public Use Microdata Series (Version 6.0) [Machine-Readable database]*. Retrieved from:

Schwartz, Christine R., and Pilar Gonalons-Pons. 2016. "Trends in Relative Earnings and Marital Dissolution: Are Wives Who Outearn Their Husbands Still More Likely to Divorce." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2 (4):218–236.

Stulp, Gert, Abraham P. Buunk, Thomas V. Pollet, Daniel Nettle, and Simon Verhulst. 2013. "Are Human Mating Preferences with Respect to Height Reflected in Actual Pairings?" *PLoS ONE* 8 (1).

Winkler, Anne W. 1998. "Earnings of Husbands and Wives in Dual-Earner Families." *Monthly Labor Review* 121 (4):42-48.

TABLE 1. MODEL CALIBRATION

Parameter	Symbol	Calibrated Value
Mean male log income	μ^m	10.35
Standard deviation of male log income	σ^m	0.75
Mean female log full income	μ^f	10.16
Standard deviation female log full income	σ^f	0.70
Mean disutility of work	ψ	.0019
Standard deviation of disutility of work	σ^ψ	$\psi/2$
Correlation, disutility of work and female log full inc	ρ	-0.4
Standard deviation of transitory income shock	σ^u	13,000
Targets in the data	**Data**	**Model**
Mean male log observed income	10.35	10.35
Standard devation male log observed income	0.75	0.75
Mean female log observed income	10.00	9.98
Standard deviation female log observed income	0.87	0.87
Mean gender earnings ratio, all	0.74	0.71
Mean gender earnings ratio, full-timers only	0.80	0.79
Female labor-force participation rate	0.88	0.91
Female full-time labor-force participation rate	0.67	0.67

Notes: Calibration of marital sorting and female labor supply model discussed in Section III.

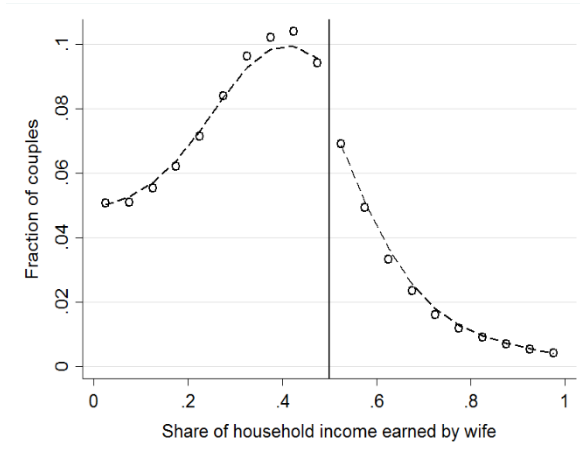
TABLE 2. DISCONTINUITY ESTIMATES IN THE GOLD STANDARD FILE

Bandwidth	Bin size	Treatment of point mass of couples at 0.5		
		Right of 0.5	Left of 0.5	Kick out 0.5 spike, test for break right at 0.5
.084	.0016	***-.124***	**.064**	-.034
		(.031)	(.031)	(.032)
.045	.0016	***-.184***	***.129***	-.031
		(.040)	(.040)	(.043)
.023	.0016	***-.310***	***.240***	-.040
		(.055)	(.055)	(.061)
.011	.0005	***-.575***	***.451***	-.078
		(.078)	(.081)	(.091)

Notes: The first reported bandwidth and bin size correspond to those automatically selected by the McCrary (2008) test algorithm. McCrary (2008) recommends using a smaller bandwidth than the automatically selected one, as is done in the second through fourth rows. Point estimates report the log difference in the height of the density function as one crosses from just left of the supposed breakpoint to just right of it. Bold estimates are statistically significant at the 5 percent level; italicized estimates achieve significance at the 1 percent level. Standard errors appear below point estimates in parentheses.

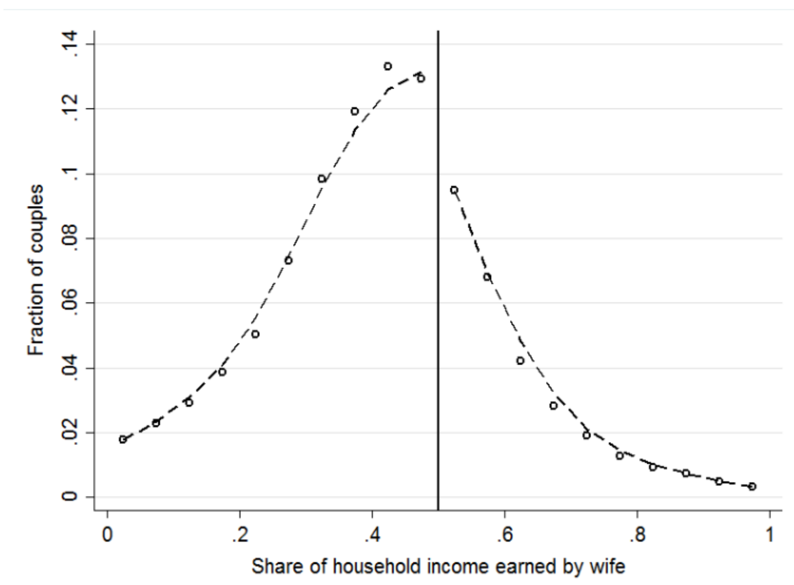
FIGURE 1. DISTRIBUTIONS OF RELATIVE INCOME, 2000 CENSUS

A. BKP Figure III *B. Replication of BKP Figure III*



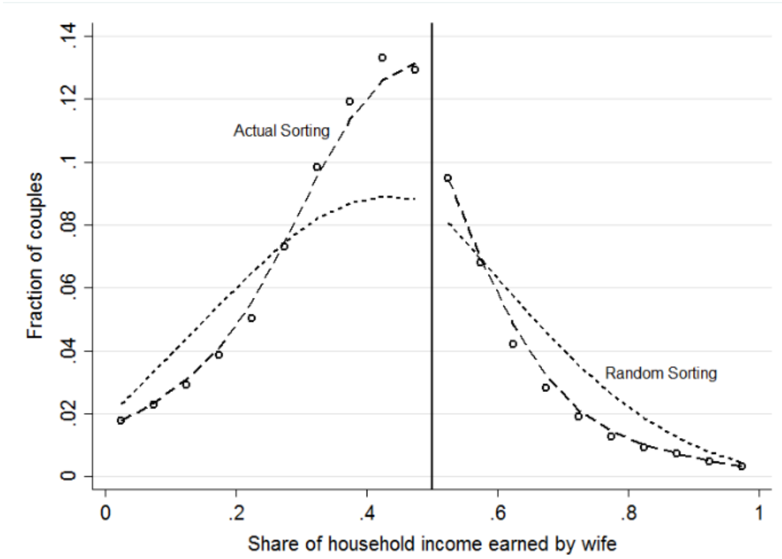
Notes: Graph A is a screenshot of part of Figure III of BKP. Graph B is our replication. Each graph is based on a sample drawn from the 2000 Census consisting of dual-earning couples, in which both the husband and the wife are between 18 and 65 years old. Each graph plots a 20-bin histogram of the distribution of the share of total household income earned by the wife. The dashed lines represent the lowess smoother applied to each histogram on either side of 0.5.

FIGURE 2. DISTRIBUTION OF RELATIVE INCOME, 2000 CENSUS
COUPLES AGED 18-40 WITHOUT CHILDREN



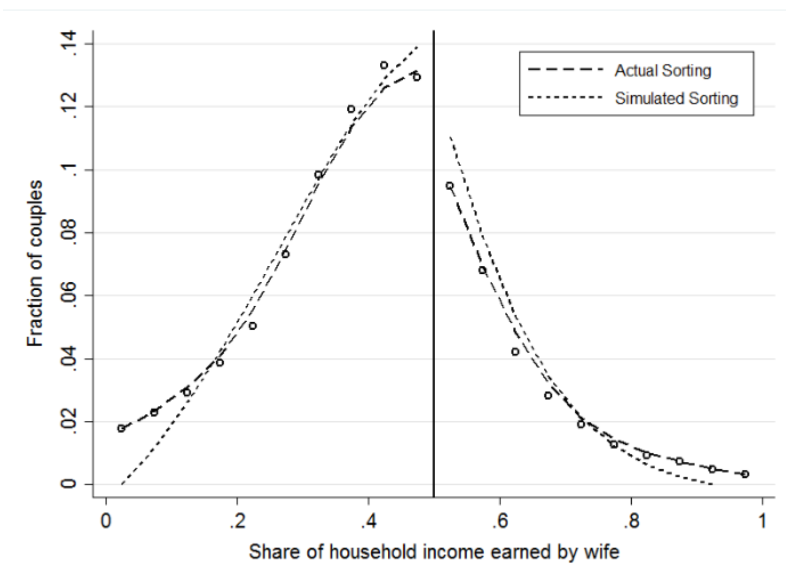
Notes: The sample includes dual-earning married couples who do not have children and where both the husband and wife are between 18 and 40 years of age. The figure plots a 20-bin histogram of the observed distribution of the share of total household income earned by the wife. The dashed lines represent the lowess smoother applied to the histogram on either side of 0.5.

FIGURE 3. RELATIVE INCOME DISTRIBUTIONS, 2000 CENSUS: ACTUAL AND RANDOM SORTING



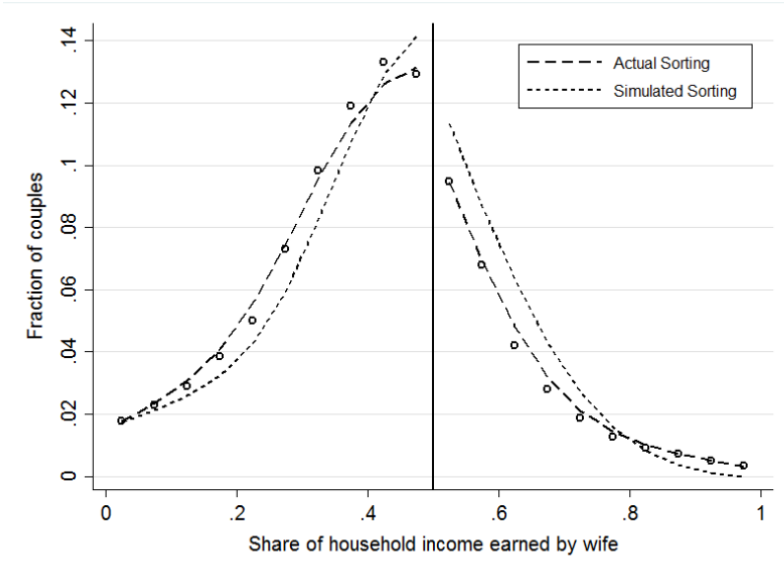
Notes: The sample is the same as in Figure 2. The figure plots 20-bin histograms of the observed distribution of the share of total household income earned by the wife ("Actual Sorting") and of a simulated distribution based on random sorting of couples in the sample ("Random Sorting"). The dashed lines represent the lowess smoother applied to the histogram on either side of 0.5.

FIGURE 4. RELATIVE INCOME DISTRIBUTIONS, 2000 CENSUS: ACTUAL AND SIMULATED SORTING



Notes: The sample is the same as in Figure 2. The figure plots 20-bin histograms of the observed distribution of the share of total household income earned by the wife ("Actual Sorting") and of a simulated distribution based on positive sorting of couples on income plus some noise ("Simulated Sorting"). See Section III for further detail on the simulation. The dashed lines represent the lowess smoother applied to the histogram on either side of 0.5.

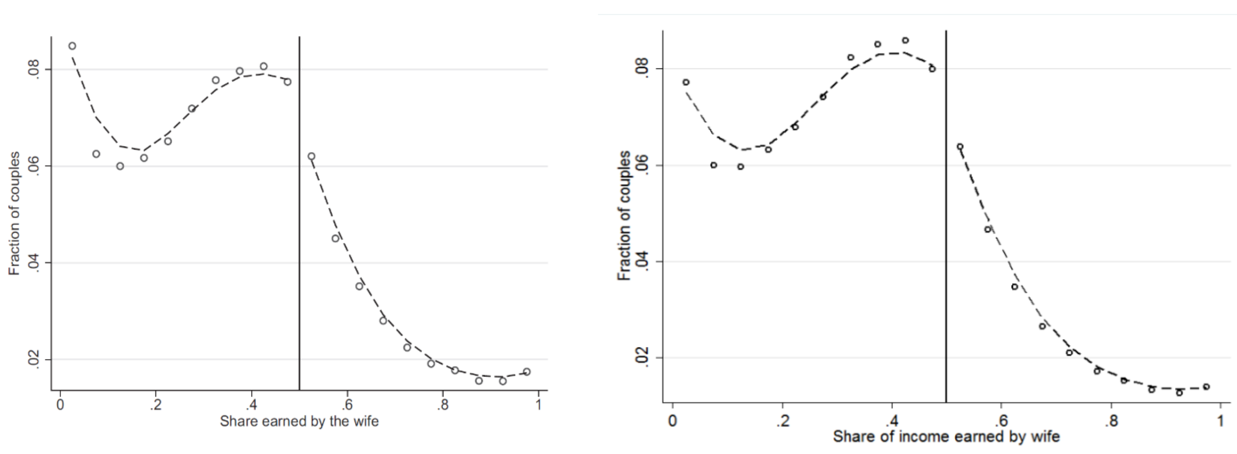
FIGURE 5. RELATIVE INCOME DISTRIBUTIONS, 2000 CENSUS: ACTUAL AND SIMULATED SORTING



Notes: The sample is the same as in Figure 2. The figure plots 20-bin histograms of the observed distribution of the share of total household income earned by the wife ("Actual Sorting") and of a simulated distribution based on positive sorting of couples on income plus some noise ("Simulated Sorting"), and in which the wife's income is endogenized via a labor supply decision. See Section III for further detail on the simulation. The dashed lines represent the lowess smoother applied to the histogram on either side of 0.5.

FIGURE 6. RELATIVE INCOME DISTRIBUTIONS IN ADMINISTRATIVE DATA

A. BKP Figure I | *B. Replication of BKP Figure I*



Notes: Graph A is a screenshot of Figure I of BKP. The data underlying this graph are administrative income data from the 1990 to 2004 SIPP/SSA Gold Standard File. Graph B is our replication of Figure I of BKP. We use the latest version of the Gold Standard File, which includes the 1984 and 2008 SIPP panels as well. For both graphs the sample includes all dual-earning couples aged 18 to 65, with income information taken from the first year the couple was observed in the SIPP panel. Both graphs plot 20-bin histograms of the observed distribution of the share of total household income earned by the wife. The dashed lines represent the lowess smoother applied to each histogram on either side of 0.5.