Very preliminary. Please do not cite or distribute.

Income, Identity, and Marriage

Sara McLanahan and Tara Watson*
March 19, 2004

This paper explores the income gradient in the likelihood of marriage among men. The relationship between income and marriage is strongly positive for men below the median income. We propose an identity model to explain non-marriage among low-income men, in which the utility of marriage is a function of how closely the couple approximates an idealized marriage norm. We predict poor men will marry less when the married men around them have high incomes. Furthermore, the ratio of an individual man's income to the "marriage bar" is expected to be a strong predictor of marriage for low-income men, but a weak predictor for high-income men. We find empirical support for the model using the 1970-2000 U.S. Censuses.

*Princeton University. Please contact Tara Watson at tw@princeton.edu.

It is well known that marriage rates are correlated with income. Men with lower earnings are less likely to be married, but the underlying behavior driving this relationship is unclear. A model by Willis (1999) suggests that when sex ratios favor men, low-income men will choose multiple childbearing partners rather than marriage. Gould and Paserman (2003) examine search in the marriage market, suggesting that women refuse to marry low-income men when income inequality is high. Others have emphasized the role of government transfer programs in reducing the desirability of marriage for low-income groups. Furthermore, an individual man's earnings tend to rise around the time he marries, leading researchers to consider the role of a "marriage premium" in the labor market.

Here, we propose a model suggesting an additional reason why low-income men are less likely to marry. The model borrows heavily from Akerlof and Kranton's (2000, 2002) model of identity. We hypothesize that people marry in part to gain utility from thinking of themselves in the category of "married people". This category entails certain prescriptions for behavior and characteristics, including a material standard of living associated with marriage. When couples are far from achieving this standard of living, they benefit less from marrying, and therefore are less likely to do so.

In this paper, we develop a simple model of marriage behavior. We then provide preliminary empirical evidence that is consistent with the model.

I. A Model of Identity and Marriage

Suppose a locality has an equal number of men and women in the marriage market. Each person is endowed with income Y_i drawn from the same distribution. Suppose further that the desirability of men and women is represented by their income Y_i. We abstract from the matching process and assume men and women are matched by the level of income such that within each couple the man and the woman have equal levels of income. The couples may decide to cohabit or marry. The value of marrying is

determined by background characteristics (such as age, education, race, income, characteristics of peers) that in turn affect the financial returns and personal returns to marriage. For example, married couples might receive financial benefits or incur costs because of tax and welfare policies that interact with their level of income. The personal returns include social rewards for marriage from family and friends as well as the effect of marrying on one's self-image.

Following the model of identity outlined in Akerlof and Kranton (2000, 2002), we describe "married people" as one group c in a set of social categories **C** with which men and women may choose to identify. Prescriptions **P** describe the ideal characteristics and behavior for each category. For our purposes, we assume these prescriptions for "married people" are formed by the typical characteristics and behavior of the group members. For example, married people might be expected to have a high level of income, to live in their own residence apart from extended family, to stay home instead of going to the bar, and to exhibit high levels of father's involvement in childrearing. We assume the category "cohabiting people" has no set of prescriptions. An individual's self-image I_i depends on the match between his or her behavior and characteristics with the ideals prescribed for his or her category.

In our simple model, we focus on the prescription that married people have a certain level of income. We also allow a random error term $\varepsilon_{ic} \sim N(0,\sigma)$ to affect an individual's self image associated with any given category. Thus, an individual's utility can be described by:

$$U_i = U_i(Y_i, I_i),$$
$$\text{where } I_i = I_i(Y_i, c_i, \mathbf{P}, \varepsilon_{ic}),$$
$$dU/dY_i > 0,$$
$$\text{and } dU/dI_i > 0.$$

That is, in general an individual's utility depends on his or her income and self-image. Self-image, in turn, is a function of an interaction between an individual's income, the

category with which he or she identifies, the prescriptions for that category, and a random error term.

Suppose that the ideal income for a married person is greater than or equal to Y_{mmed}, where Y_{mmed} is the median income of currently married people. The identity payoff for a married person is then:

$$I_i = I_{mar} - t(\max(0, 1 - (Y_i / Y_{mmed})) + \varepsilon_{imar},$$

where t is a positive scalar describing the identity loss associated with falling below the "marriage bar". The identity payoff for cohabiting is:

$$I_i = I_{cohab} + \varepsilon_{icohab},$$

and we assume $I_{mar} > I_{cohab}$. In other words, on average an ideal married person has a higher self-image than an average cohabiting person.

When making the decision whether to marry, couples compare the utility from cohabiting and marriage, and ignore the small effect their decision has on Y_{mmed}. The self-image gained through marriage (relative to cohabitation) is

$$(I_{mar} - I_{cohab}) + (\varepsilon_{imar} - \varepsilon_{icohab}) - t(\max(0, 1 - (Y_i / Y_{mmed}))).$$

Therefore, the gains to self-image through marriage tend to increase with the average gain in self-image from marriage, the idiosyncratic gain from marriage, and an individual's income, while the gains decrease with a higher "marriage bar" and a higher penalty t for deviating from the norm.

This framework generates some simple comparative statics. The gain to marriage is increasing in Y_i for $Y_i < Y_{mmed}$ ($dI/dY_i = t/Y_{mmed}$) and constant in Y_i for $Y_i > Y_{mmed}$ ($dI/dY_i = 0$). Similarly, an increase in the marriage bar Y_{mmed} is associated with a decrease in the gain to marriage for low values of Y_i ($dI/dY_{mmed} = -tY_i / Y_{mmed}^2$) but no change in the gain

4

for high values of Y_i. An increase in Y_{mmed} is also associated with an increase in the population falling below the marriage bar, thus reducing the overall marriage rate. A higher level of t strengthens the relationship between Y_i and marriage below the median, and reduces the overall marriage rate holding other factors constant.

In the next section we will test several predictions of the model. Holding other factors constant, we expect

(i) A high level of the "marriage bar" Y_{mmed} is associated with a lower marriage rate.

(ii) The decline in marriage is concentrated in the bottom half of the income distribution.

(iii) The ratio of an individual's income to the marriage bar is strongly predictive of marriage below the marriage bar, and unrelated to marriage above the marriage bar.

II. Empirical Evidence

We use the 1970-2000 Censuses to investigate the determinants of marriage. We focus on a 0.5% sample of men ages 30-49, and limit our analysis to native non-Hispanic white men.[1] For each state and year, we determine the median total income of native non-Hispanic white married men ages 30-49. After adjusting for inflation, we refer to this median income of married men as the "marriage bar" for the state and year. We use states as the relevant geographic units for consistency over time. In future work, we will also consider norms at the metropolitan area level.

Our model leaves us agnostic as to whether we should consider incomes of men, women, or both together. Emprically, however, female labor market participation and income are highly endogenous to the marriage decision. This is much less true for men. Therefore, we focus on male income and male probability of marriage. The final sample used in the individual-level analysis excludes the top and bottom tails of the income distribution in

---

[1] This sample was drawn randomly from the 1% IPUMS for 1970 and 2000 and the 5% IPUMS for 1980 and 1990.

each year.[2] We estimate the probability that a man is currently married using probit models. The reported coefficients represent marginal effects evaluated at the mean.

Table 1 describes the sample characteristics. Notably, the fraction of men in the sample who are married has declined over time. The "marriage bar", the median income of a typical married man in an individual's state-year, has been declining slightly over time. However, because the income of men in the bottom half of the income distribution has also been falling, the fraction of men below the marriage bar is growing over time.

Figure 1 illustrates the raw probability of marriage at different levels of income for the 1970-2000 Census years. Marriage has decreased at every point in the income distribution. Furthermore, within each year there is a strong positive relationship between income and marriage. This relationship appears to be stronger (i.e. the line is steeper) at the bottom of the income distribution.

In Table 2, we show the base specification describing the relationship between income and marriage. There is a strong positive relationship between the log of real income and marriage after controlling for a number of individual and state-year characteristics, state fixed effects, and year fixed effects. The positive relationship is stronger for individuals below the median of the income distribution in any given year. In other words, additional income is more strongly associated with marriage for low-income men than for high-income men. We also consider a series of twenty income categories and find that the coefficients on the control variables do not change very much compared to using a linear specification for income.[3]

In Table 3, we explore the relationship between the "marriage bar" in a state and the likelihood of marriage. According to our model, an exogenous increase in the median income of married men in the state (the marriage bar) should reduce the probability of

---

[2] We exclude the top and bottom 5 percent of all those with positive income, as well as those with non-positive income, from the analysis. The men with non-positive total income never exceed 3 percent of the sample in a given year.

[3] The pooled incomes of the sample for 1970-2000 are divided into twenty groups of equal size. These categories are used for the twenty income dummies.

6

marriage of low-income men. After controlling for the many factors included in Table 2, we do see the predicted negative relationship. This negative relationship holds even after controlling for log income linearly, including a square term, allowing the effect of income to be different below the median, or using a series of twenty income categories. Columns V and VI of Table 3 shows that the marriage bar is most negatively associated with marriage among those in the bottom half of the income distribution, as predicted by the model.

There is a clear potential for reverse causality in the regressions in Table 3. In particular, any unobserved factor that discourages marriage among low-income men would tend to increase the "marriage bar." In future versions of this paper we plan to investigate the determinants of the marriage bar and address this possible source of bias.

Table 4 considers the ratio of an individual's income to the marriage bar in his state-year. Even after controlling for the log of real income, the ratio of a man to the marriage bar is a strong predictor of marriage for men below the bar. This relationship is much weaker for men above the marriage bar. Indeed, if one controls non-linearly for income as in column IV, the income-to-marriage bar ratio is highly associated with marriage below the bar, but not significantly related above the bar. The regressions in Table 4 are consistent with the model.

III. Conclusions

In this paper, we develop a model of marriage that is based on identity. Couples marry in part for the identity utility of marriage, which is larger when they approach a norm or an ideal set of characteristics. We hypothesize that low-income couples choose to avoid marriage because they are far from the "marriage bar." The empirical work suggests that the level of the marriage bar in an individual's state and year is negatively associated with marriage probabilities of low-income men. The ratio of an individual's income to the "marriage bar" is positively associated with marriage for men below the bar.

7

We plan to develop this work in several ways. First, we would like to explore the determinants of the marriage bar and try to establish a causal relationship and marriage rates. Using the Fragile Families data set, an in-depth study of more than 3000 couples with newborn children, we may be able to garner some insight into the exact mechanism at work. For example, it would be interesting to understand the role of housing costs and income inequality. We also plan to consider the role of the "marriage bar" in the African-American community.

IV. References

Akerlof, George A., and Rachel E. Kranton, 2000 "Economics and Identity." *Quarterly Journal of Economics,* 115(3), 715-53.

Akerlof, George A., and Rachel E. Kranton, 2002. "Identity and Schooling: Some Lessons for the Economics of Education," Journal of Economic Literature, 40(4), 1167-1201.

Gould, Eric D. and Daniele M. Paserman, 2003. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates," *Journal of Urban Economics,* 53(2), 257-81.

Willis, Robert J., 1999. "A Theory of Out of Wedlock Childbearing," *Journal of Political Economy,* 107(6), S33-64.

DEMO VERSION

Table 1. Sample Characteristics

	1970	1980	1990	2000
Individual Characteristics (excludes tails of income distribution)				
Number of Observations	78,869	92,564	122,504	127,703
Married	0.91	0.84	0.76	0.71
Median Total Income ($2000)	43,402	44,841	41,662	40,518
25th Percentile Total Income ($2000)	33,079	30,965	27,774	26,874
Log Total Income ($2000)	10.65	10.64	10.56	10.56
State-Year Marriage Bar	44,588	45,999	44,305	43,783
Under Marriage Bar	0.51	0.53	0.54	0.55
Age	39.53	38.28	38.50	39.59
Employed	0.96	0.94	0.94	0.92
At Least High School Grad	0.67	0.83	0.92	0.94
At Least Some College	0.30	0.46	0.60	0.61
At Least College Grad	0.17	0.27	0.29	0.29
Control Variables				
State-Year Marriage Rate	0.88	0.83	0.74	0.70
State-Year Sex Ratio	1.03	1.01	1.00	1.00
State-Year Fraction Black	0.09	0.11	0.11	0.11
State-Year Fraction Foreign Born	0.10	0.07	0.09	0.11
State-Year Fraction Hispanic	0.05	0.06	0.08	0.11
State-Year Fraction Under 20	0.38	0.32	0.29	0.28
State-Year Fraction Under 65	0.90	0.89	0.87	0.87
State-Year Fraction At Least Some College	0.21	0.32	0.45	0.52
State-Year Fraction At Least College Grad	0.11	0.16	0.20	0.24
State-Year Male Emp.-to-Pop. Ratio (ages 18-64)	0.85	0.81	0.80	0.77

Notes: Sex ratio is ratio of native non-Hispanic white men ages 30-49 to women of the same description.
All statistics based on estimates after excluding tails of distribution in each year.

Figure 1.
Fraction Married By Log Income,
Native Non-Hispanic White Men 30-49

Fraction Currently Married (y-axis)

Log Real Total Income ($2000)
(excludes tails in each year)
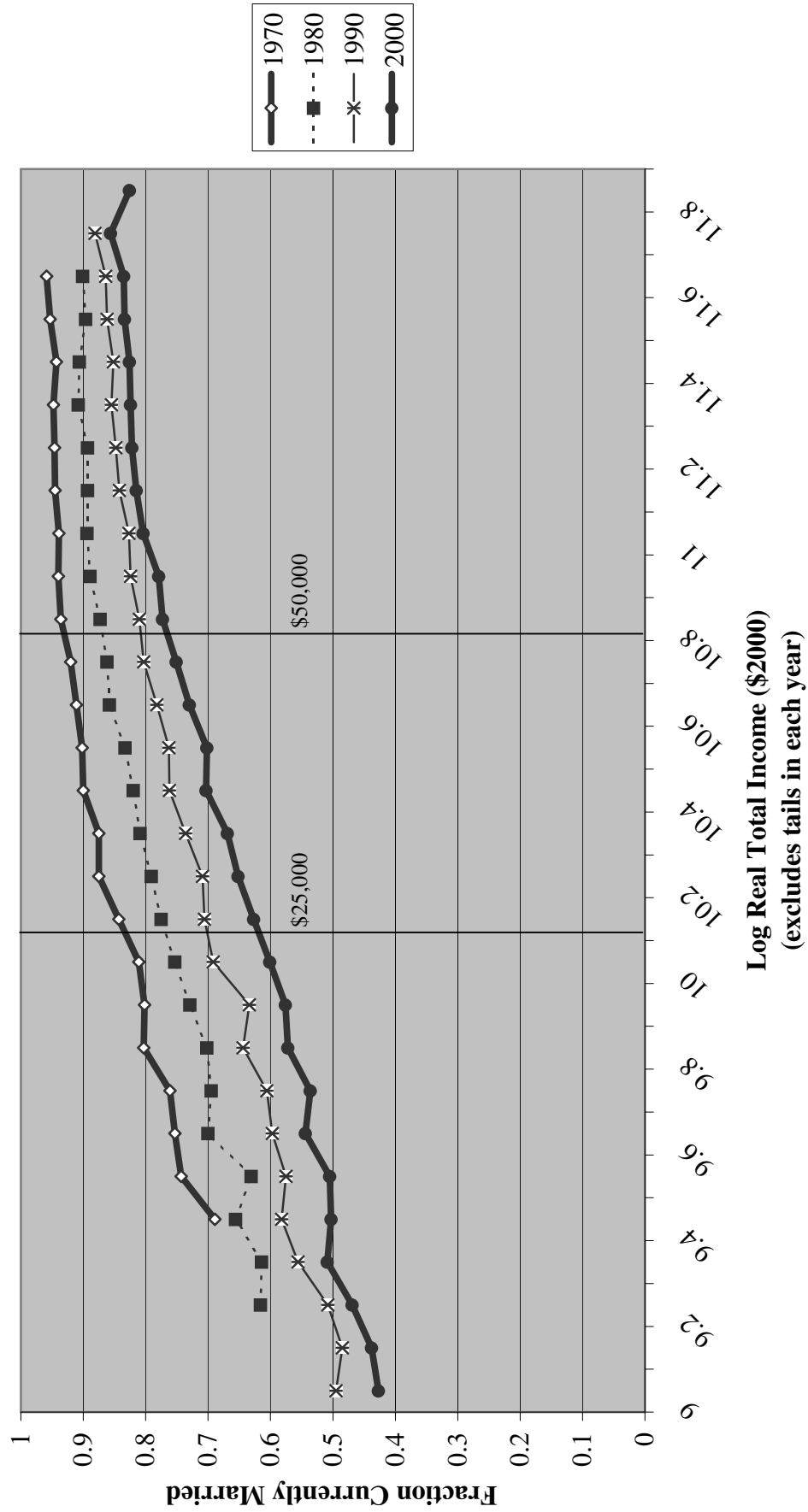
Legend: 1970, 1980, 1990, 2000

Table 2. Determinants of Marriage for Native Non-Hispanic White Men Aged 30-49,
Census Years 1970-2000, Probit Models

| | *Dependent Variable: Married Now* | | | |
	I	II	III	IV
Log Total Real Income ($2000)	0.149**	0.118**	0.128**	
	(0.003)	(0.005)	(0.005)	
Income Dummies (20 Categories)				yes
Log Total Real Income ($2000)*Bottom 50% of Year		0.022**		
		(0.006)		
In Bottom 50% of Year		-0.250**		
		(0.067)		
Log Total Real Income ($2000)*Bottom 50% of State-Year			0.006	
			(0.005)	
In Bottom 50% of State-Year			-0.092	
			(0.061)	
Employed	0.076**	0.076**	0.077**	0.080**
	(0.005)	(0.005)	(0.005)	(0.005)
At Least High School Grad	-0.021**	-0.022**	-0.022**	-0.022**
	(0.002)	(0.002)	(0.002)	(0.002)
At Least Some College	-0.005**	-0.006**	-0.006**	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
At Least College Grad	-0.032**	-0.032**	-0.032**	-0.032**
	(0.005)	(0.005)	(0.005)	(0.005)
State-Year Marriage Rate	0.294**	0.287**	0.319**	0.268**
	(0.101)	(0.102)	(0.100)	(0.102)
State-Year Sex Ratio	0.022	0.021	0.024	0.022
	(0.061)	(0.061)	(0.060)	(0.062)
State-Year Fraction Black	0.273	0.283	0.257	0.306*
	(0.174)	(0.175)	(0.173)	(0.181)
State-Year Fraction Foreign Born	-0.233*	-0.233*	-0.209	-0.257*
	(0.135)	(0.136)	(0.132)	(0.139)
State-Year Fraction Hispanic	0.159	0.163	0.161	0.184
	(0.114)	(0.115)	(0.112)	(0.115)
State-Year Fraction Under 20	0.601**	0.605**	0.570**	0.622**
	(0.142)	(0.143)	(0.139)	(0.147)
State-Year Fraction Under 65	-0.206	-0.200	-0.202	-0.206
	(0.202)	(0.202)	(0.199)	(0.206)
State-Year Fraction At Least Some College	-0.276**	-0.273**	-0.281**	-0.261**
	(0.105)	(0.107)	(0.103)	(0.110)
State-Year Fraction At Least College Grad	0.317**	0.312**	0.346**	0.308**
	(0.105)	(0.105)	(0.104)	(0.109)
State-Year Male Emp.-to-Pop. Ratio	-0.287**	-0.286**	-0.277**	-0.293**
	(0.099)	(0.099)	(0.098)	(0.099)
Age Dummies	yes	yes	yes	yes
State Dummies	yes	yes	yes	yes
Year Dummies	yes	yes	yes	yes
Number of Observations	421640	421640	421640	421640

Notes: * and ** refer to statistical significance at the 10 and 5 percent levels, respectively. Standard erro
in parentheses clustered on state.

Table 3. Effect of Marriage Bar on Marriage of Native Non-Hispanic White Men Aged 30-49,
Census Years 1970-2000, Probit Models

Dependent Variable: Married Now

	I	II	III	IV	V	VI
Log Marriage Bar in State-Year	-0.100**	-0.100**	-0.102**	-0.107**	-0.073**	-0.076**
	(0.034)	(0.034)	(0.034)	(0.035)	(0.033)	(0.034)
Log Marriage Bar in State-Year*In Bottom 50% of Year					-0.050*	
					(0.026)	
Log Marriage Bar in State-Year*In Bottom 50% of State-Year						-0.059**
						(0.027)
Log Total Real Income ($2000)	0.149**	0.087*	0.118**			
	(0.003)	(0.053)	(0.005)			
Log Income Squared		0.003				
		(0.003)				
In Bottom 50% of Year			-0.250**		0.511**	
			(0.067)		(0.244)	
In Bottom 50% of State-Year						0.507**
						(0.191)
Log Total Real Income ($2000)*In Bottom 50% of Year			0.022**			
			(0.006)			
Income Dummies (20 Categories)				yes	yes	yes
All controls	yes	yes	yes	yes	yes	yes
Number of Observations	421640	421640	421640	421640	421640	421640

Notes: * and ** refer to statistical significance at the 10 and 5 percent levels, respectively. Standard errors
in parentheses clustered on state.

Table 4. Determinants of Marriage for Native Non-Hispanic White Men Aged 30-49,
Census Years 1970-2000, Probit Models

| | *Dependent Variable: Married Now* | | | | |
	I	II	III	IV	V
Log Total Real Income ($2000)	0.138**	0.094**	0.050**		
	(0.003)	(0.007)	(0.013)		
Income Dummies (20 Categories)				yes	
Under Marriage Bar	-0.017**	-0.118**	-0.157**	-0.149**	-0.209**
	(0.003)	(0.014)	(0.017)	(0.022)	(0.007)
Ratio Income to Marriage Bar if Under Marriage Bar		0.104**	0.183**	0.174**	0.273**
		(0.013)	(0.025)	(0.023)	(0.006)
Ratio Income to Marriage Bar if Over Marriage Bar			0.033**	0.018	0.064**
			(0.009)	(0.015)	(0.004)
All Controls	yes	yes	yes	yes	yes
Number of Observations	421640	421640	421640	421640	421640

Notes: * and ** refer to statistical significance at the 10 and 5 percent levels, respectively. Standard errors
in parentheses clustered on state.