A Contribution to the Marriage Premium Analysis

Campbell, J., Kim, J., Low, XK. & Rybak, C.

Section I: Introduction and Literature Review

This paper has the modest goal of showing the existence of the marriage premium in the United States, and then explaining possible causes for said premium. We hypothesize that the premium can be explained by both the specialization, and the selection hypothesis. Our estimates provide evidence for the existence of unobservable characteristics that account for a percentage of the premium. Further research will go into accounting for the remaining, unaccounted for, discrepancy.

The idea of the marriage premium isn't new to the 21st century. Findings from the late 1970's show that ever-married white men earned 25-30 percent higher wages than did single white men (Hill 1979). This marriage premium has been explained variously by three main hypothesis. The first is the idea that marriage makes men more productive, which has come to be known as the specialization or productivity hypothesis. The second is that more productive men tend to get married, perhaps due to unobservable characteristics valued by both the workforce and potential spouses, now known as the selectivity hypothesis. The third is the idea of discrimination in the work force, where bosses tend to prefer married men to non-married men.

In his 1981 work, "A Treatise on the Family," Gary Becker helped lay the groundwork for the specialization hypothesis. It was his idea that if members of an efficient household had different comparative advantages, no more than one member would allocate time to both the market and the household sectors (Becker 1981). If single men had to dedicate time to both the market and household sector, that same man, when married, could spend all of his time in the market, potentially accruing more human capital in the process. In the Coleman-Rossi Retrospective Life Histories Study, 1598 males were asked questions regarding their employment past and marital

status. Analysis of the survey suggested that human capital is accumulated more rapidly when a male is married than when that same male is single, other things equal (Kenny 1983), a step towards the productivity hypothesis.

Using data from the National Longitudinal Survey (NLS) of Young Men, which followed men aged 14-24 in 1996 for fifteen years, economists used fixed effects analysis to test the selectivity hypothesis. Taking unobservable characteristics into consideration only helped to explain less than 20% of the marriage premium (Korenman and Neumark 1991). They found some evidence of the productivity hypothesis, but they recommended further study into the matter.

Delving more into the productivity hypothesis, Joni Hersch and Leslie S. Stratton (2000) used fixed effects estimation, specifically first differences, to analyze whether or not the marriage premium could be explained by specialization within the household. Per Becker's 1981 work, an efficient marriage wouldn't have multiple people specializing in both the market and household sector. One spouse would focus on the household, where another, presumably the male, would specialize in the market. Taking data from the National Survey of Families and Households, which surveyed 13,008 households in 1978-1988 and again in 1992-1994, they concluded that specialization does a poor job in explaining the premium in question.

The results from Hersch and Stratton (2000) have been questioned recently, with papers by Hyunbae Chun and Injae Lee (2001), and Elena Bardasi and Mark Taylor (2006), both finding evidence that the specialization hypothesis does a good job at explaining a significant part of the marriage premium. Chun and Lee (2001) showed that wage gains from marriage decreased by 0.6% for each additional hour per week that a wife works in the labor market. Bardasi and Taylor (2006) found evidence that perhaps both the selectivity hypothesis and the specialization hypothesis are correct when it comes to explaining the premium. They showed that the premium decreased by one third when taking into account intra-household specialization. The marriage

premium also decreased by half when controlling for unobservable characteristics. This still leaves approximately 15% of the premium unexplained.

Section II: Theoretical Model

This paper is built on a basic linear regression model with the dependent variable being the log of wages. For each individual, there are observable ($\boldsymbol{X}$ being the observable time-variant unobservables) and unobservable characteristics (z_i serving as vector of time-invariant individual constants) determining the logged wages. For the purpose of analyzing the 'marriage premium', we include M_{it}, the vector of marital status at time t for individual i.

The Basic Model

$$\ln(w_{it}) = X_{it}\beta + M_{it}\gamma + z_i\alpha + \varepsilon_{it}$$

In this basic model, in order to simplify analysis, we have assumed that there is no significant time-varying unobservable characteristic that might be correlated with any of the explanatory variables and thereby would be appropriately subsumed in the error term of the model. 'z' represents the vector of unobservable time-invariant individual characteristics and α is the coefficient for these unobserved variables. Therefore, β gives the vector of coefficients for the other determinants of wages; γ provides the estimate of the marriage premium.

Fixed Effects Marriage-Premium Identification

$$\ln(w_{it}) - \ln(\overline{w_i}) = (X_{it} - \overline{X_i})\beta + (M_{it} - \overline{M_i})\gamma + \nu_{it}$$

Since the unobservable traits of the individual driving the logged wages are likely to be correlated with other wage-determinants, we would have to identify the marriage premium using a fixed effects regression and test out further hypothesis using this fixed effects model instead.

One could think of how a certain personality characteristic that is unobserved that influences wages could also affect educational attainment (which is in the $\boldsymbol{X}$ matrix). As we are mainly interested in the marriage dummy, which varies across time and individual, it will not drop out of the model and its within-individual estimator would provide a consistent estimate of the premium we are studying.

Testing Specialization Hypothesis

By assuming $M_{it} = 1$ in our basic model we arrive at the following statement:

$$\ln(w_{it}) = X_{it}\mathrm{B} + S_{it}\delta + \varepsilon_{it}$$

In this modification of the basic model, we include S_{it} which is a measure of the number of hours a spouse works in a week. A higher number of hours worked per week by a spouse indicates a lower degree of specialization within the marriage in term of wage-earning and household-stewardship. By including this variable as a proxy for the degree of specialization within the marriage relation, we measure the extent of specialization's impact on an individual's wages. In this model, we assume that there are no unobservable individual characteristics that are correlated with the number of hours the spouse works which affects the log of wages. The average hours worked by the spouse within the set of observations where individuals are married would give us an estimate of the average premium contributed by the average degree of specialization present.

Fixed-effect identification in this model would help to eliminate unobservable heterogeneity and thus eliminate any potential bias resulting from omitted variables that are correlated with both the explanatory and dependent variables. However, that would also remove cross-individual variation in spouse work hours that could potentially explain differences in the premium. We have thus decided to stick to a model that exploits the heterogeneity across individuals' spouse working-hours impact on individuals.

Section III: Data and Estimation

Our analysis uses data from the US National Longitudinal Survey of Youth (NLS). The survey performed annual surveys of 12686 individuals annually from 1979 to 1994 and then biannually from 1994 to 2010. All participants were between the ages of 14 and 22 when the survey began in 1979. After some subgroups were eliminated the final sample contains 24 panels of approximately 10,000 individuals. The surveys contained hundreds of variables about all areas of the individual's life. The questions on the survey were not consistent throughout the life of the NLS. Many questions were added or removed from the survey creating difficulties in taking full advantage of the data set. However, the survey consistently focused on labor market questions and household demographics, making it possible to generate a useable, albeit unbalanced, panel for our paper.

Data treatment

Our final sample contained 6213 individuals with approximately 6 panels per individual. In order to create a usable panel a number of adjustments needed to be made to the overall data set. Some obvious steps were taken, such as removing women from the sample and individuals who didn't clearly report their marital status on surveys. However, we also eliminated individuals who are self-employed. Theoretically, the marriage premium should apply to self-employed individuals, however, literature suggests self-employed individuals under report earnings. Additionally, individuals were not included during periods of sudden unemployment. This allowed for a continuous data set to regress on.

Dependent Variable

Our dependent variable was the log of earnings (*LEARN*). Fortunately, earnings data was collected extensively throughout the survey. Specifically, we looked at hourly wage. Annual salary information was available however it does not control for things like hours worked per week or time spent on vacation. Beyond consistency hourly wages are an economically sound choice because we are interested in efficiency gains rather strict earnings gains. Individuals can increase their income by taking a second job or working overtime. The marriage premium states that individuals actually become more efficient following marriage, and thus command a higher wage for the same amount of work. We've also taken the log of wages rather than just wages. Wages have a positive skew and taking logs allows it to fit better into the model.

Key Explanatory Variables

It is worth discussing our primary variable, marital status (*MARRIED*), in more detail. For the purposes of this paper we are interested solely in the variation between married and unmarried individuals. However, unmarried individuals fall into several categories that may be significant with regard to the specialization hypothesis. Unmarried individuals fall into three categories: those who have never been married, those who have been married but divorced and those who have been widowed. Our hypothesis suggests that marriage allows the male to devote more time to work and thus earn a higher hourly wage. A quick survey of the data suggests that this may be the case, as married men earn substantially more than an unmarried man in any group. This anecdotal look at our data is confirmed by our regression analysis in table 2 below. There is also guidance with regard to whether or not the time spent at work accrues as a form of human capital. Again, our anecdotal look at the data suggests that this is not the case. While further study would be necessary to truly examine such a correlation a quick summary of our data supports our choice of solely looking at married vs unmarried individuals.

Variable	Mean	Std. Dev.	Min	Max
Log of Earnings	6.733849	0.7879757	0	15.60727
Marriage Dummy	0.4269685	0.4946404	0	1
Spouse Hours	12.00656	18.22261	0	144
Children Dummy	0.2548787	0.4357954	0	1
Number of children	0.46987	0.9276564	0	8
Age	29.69901	8.596399	14	53
Family Size	3.188381	1.859485	1	15
Disability Dummy	0.0101516	0.100243	0	1
Tenure in Job	191.2531	251.1592	1	1763
Black Dummy	0.2599278	0.4385972	0	1
Hispanic Dummy	0.1780961	0.3825957	0	1
Years of Schooling	12.5768	2.418367	0	20
Degree Dummy	0.1528942	0.3598875	0	1
Urban Local Dummy	0.7554947	0.4297959	0	1
Northeastern Dummy	0.1740589	0.3791622	0	1
Southern Dummy	0.3777088	0.4848169	0	1
Western Dummy	0.1999868	0.3999923	0	1
Lives in Parents Household Dummy	0.1036235	0.3047732	0	1
Divorce Dummy	0.0772776	0.2670329	0	1
Seperated Dummy	0.0329791	0.1785829	0	1
North Central Dummy	0.2482454	0.4319973	0	1
Single Dummy	0.4606407	0.4984512	0	1
Widowed Dummy	0.0021341	0.046147	0	1

<Table 1. Summary Statistics>

Spouse Hours

Beyond confirming the existence of the marriage hypothesis our model also requires we test the specialization hypothesis. Specifically, does marriage allow men to invest more time into work, and therefore develop more human capital? Some longitudinal studies have provided very ways of testing this, such as listing the number of chores for which each member was responsible. Our data set did not have this information, however, a reliable proxy is spouse hours worked

(*SPHRS*). While imperfect, ample research (Daniel 1992, Gray 1997) has demonstrated that household responsibilities are dependent on spousal hours. it is also a very natural assumption that as spouses hours increase their time spent on household chores decreases. So, a primary regressor of interest was how spousal working hours impacted the wage premium.

Key Controls

The number of observable and unobservable variables that affect wage growth is myriad. A large portion of our chosen variables was meant to account for wage growth in order to isolate the marriage effect. Many of these controls are intuitive. Obviously things such as age, age squared (*AGE, AGE2*), years of schoolings, (*YSCH*) degree (*DEGREE*) and job tenure (*TENURE*) are expected to have direct positive impacts on wages. However, recent literature has demonstrated that other things may have an effect on hourly wages. Time constraints due to presence of children (*CHILD*) may induce subjects to work more efficiently leading to an increased wage per hour. We also included things that may impact wages negatively or in less certain ways such as disability status (*DISB*), location (*LOCNE, LOCSU, LOCWT, URBAN, PRTHH*) and ethnicity (*BLACK, HISP*). We erred on the side of caution when considering explanatory variables. It is central to this paper that we not overestimate the explanatory power of marriage.

An additional factor that must be taken into account is time. Some time effects are obvious, such as inflation, while other important time effects may be unknown. Things such as business cycles and external shocks on the economy will directly impact wages. In order to control for this we also included time dummies for each year of the sample. Thus each wage entry was associated with a particular year. While more complex data techniques could also be used, time dummies allowed us to control for these effects while not limiting the various estimation techniques that were open to us when we began empirical testing.

Empirical Model

Taking all of our variables into consideration the empirical specification becomes:

$$LEARN = \beta cons + \beta_1 MARRIED + \beta_2 AGE + \beta_3 AGE2 + \beta_4 DISB + \beta_5 TENURE + \beta_6 BLACK + \beta_7 HISP + YSCH + \beta_9 DEGREE + \beta_{10} URBAN + \beta_{11} LOCNE + \beta_{12} LOCSU + \beta_{13} LOCWT + \beta_{14} PRTHH + \beta_{15} SPHRS + \beta_{16} CHILD + \beta_{17} Y1983 + \cdots + \beta_{35} Y2010$$

We will be considering this model under OLS and under fixed effects. Obviously, the effects of the data will differ slightly in each case. The time invariant variables will not be considered under the fixed effects model, so we will lose information on dummy variables and the demographic variables such as *BLACK* and *HISP*. The coefficients of primary interest will be on the variables married and spouse hours. A statistically significant positive coefficient on *MARRIED* will be necessary to establish the existence of the marriage premium. While fixed effects should lower it slightly, if the specialization hypothesis is valid we should also see a significant positive coefficient on *SPHRS*. While the control variables are less central to the hypothesis of this paper we also hope to see intuitive coefficients in order to verify that the data set conforms to modern research in the field of earnings research.

Section IV: Emprical Results

Table 1 summarizes the results of our OLS and fixed effects regressions for the marriage premium hypothesis with natural log of wage as the dependent variable and marital status as the primary explanatory variable. We ran a fixed effects regression as opposed to a random effects regression because (1) intuitively, there are time-invariant unobservable characteristics correlated with both wages and marital states and; (2) after running the Hausman test to see if the individual

time-invariant traits are correlated with the other explanatory variables, we rejected the null hypothesis that they are not correlated (chi2(33)=518.93, Prob>chi2=0.000).

VARIABLES	(OLS) LEARN	(FE) LEARN
MARRIED	0.170***	0.108***
	(0.0044500)	(0.0064800)
AGE	0.0821***	0.0793***
	(0.0032200)	(0.0084900)
AGE2	-0.00104***	-0.00118***
	(0.0000549)	(0.0000873)
DISB	-0.167***	-0.0869***
	(0.0191000)	(0.0184000)
TENURE	0.000462***	0.000315***
	(0.0000091)	(0.0000151)
BLACK	-0.140***	
	(0.0047100)	
HISP	-0.0847***	
	(0.0058500)	
YSCH	0.0443***	0.0401***
	(0.0012800)	(0.0047500)
DEGREE	0.167***	0.322***
	(0.0083000)	(0.0203000)
URBAN	0.113***	0.0542***
	(0.0046700)	(0.0078700)
LOCNE	0.133***	0.120***
	(0.0060900)	(0.0262000)
LOCSU	0.0272***	0.0596***
	-0.00507	-0.0197
LOCWT	0.135***	0.140***
	(0.0062200)	(0.0235000)
PRTHH	-0.0807***	-0.0256***
	(0.0067400)	(0.0073100)
Constant	4.054***	4.158***
	(0.0434000)	(0.1380000)
Observations	89,439	89,439
R-squared	0.503	0.525
Number of CASEID		6,211

*** p<0.01, ** p<0.05, * p<0.1, Robust S.E. in parenthesis

<Table 1. Results of OLS and FE for marriage premium. Dummy variables for the year were included in ALL regressions to control for inflation but have been omitted from ALL tables.>

In the OLS regression, the marital status coefficient, *MARRIED*, had a significant positive coefficient of 0.170 suggesting that being married increased earnings by 100(e^ 0.170 - 1) = 18.53%. This, however, does not give insights as to whether the earnings premium is coming from the marital status or from other unobservable characteristics. The fixed effects regression had a coefficient for *MARRIED* of 0.108; even after allowing for time-invariant individual characteristics, there was still a 10.8% premium on *LEARN*. Therefore, a little less than two thirds of the premium observed in the OLS estimate is due to other unobservable characteristics, such as employer discrimination and or specialization, that is correlated with both marital status and wage premium.

As expected, *AGE* and *DEGREE* had positive effects on log earnings. Schooling had a small positive coefficient, which may arise because degrees are more of a signal to the job market than the actual number of years of schooling. Those with a degree had a 38.99% higher hourly wage than those who did not in the fixed effects regression. This was substantially higher than in the OLS (18.18%) regression suggesting that there may be a individual personality trait which would be viewed as detrimental to work progression (wages) if the individual did not have a degree but viewed favorable in the presence of a degree. In addition, those residing in urban areas earned more; in terms of location within the USA, those working in the Northeast and the West saw a greater effect on the wage premium than those working elsewhere. Those who were disabled earned less as expected and our last explanatory variable, *PRTHH*, showed a significant coefficient suggesting that those who live with their parents earn less.

Existence of the marriage premium

The following table is a summary of the results of our fixed effects regression which tested if individuals married due unobservable characteristics of the spouse (specialization hypothesis). The null hypothesis of this OLS regression was that spouse hours had no effect on the log of earnings of the male and for our characteristic-induced selection regression, the null hypothesis was that even after controlling for unobservable characteristics of the spouse, marriage would have no effect on the wage premium. The first OLS regression tests the effect of *SPHRS* on *LEARN* with *MARRIED* acting as a control for marriage and the second OLS regression adds a control for the presence of children (*CHILD*) in the household. The resulting coefficient estimates on SPHRS appears robust to controlling for the presence of children.

VARIABLES	(OLS) LEARN	(OLS) LEARN
MARRIED	0.187***	0.179***
	(0.0059800)	(0.0069500)
SPHRS	-0.000667***	-0.000617***
	(0.0001540)	(0.0001560)
CHILD		0.0152***
		(0.0058700)
AGE	0.0821***	0.0815***
	-0.00322	-0.00323
AGE2	-0.00104***	-0.00103***
	(0.0000549)	(0.0000549)
DISB	-0.167***	-0.167***
	(0.0191000)	(0.0191000)
TENURE	0.000462***	0.000462***
	(0.0000091)	(0.0000091)
BLACK	-0.140***	-0.141***
	(0.0047100)	(0.0047100)
HISP	-0.0855***	-0.0862***
	(0.0058500)	(0.0058500)
YSCH	0.0445***	0.0448***
	(0.0012800)	(0.0012800)
DEGREE	0.166***	0.167***
	(0.0083000)	(0.0083100)
URBAN	0.113***	0.113***
	(0.0046700)	(0.0046700)

LOCNE	0.133***	0.133***
	(0.0060900)	(0.0060900)
LOCSU	0.0278***	0.0280***
	(0.0050700)	(0.0050700)
LOCWT	0.135***	0.135***
	(0.0062100)	(0.0062100)
PRTHH	-0.0814***	-0.0795***
	(0.0067400)	(0.0068000)
Constant	4.052***	4.057***
	(0.0434000)	(0.0434000)
Observations	89,439	89,439
R-squared	0.503	0.503
Number of CASEID	6,211	6,211

*** p<0.01, ** p<0.05, * p<0.1, Robust S.E. in parenthesis

<Table 2. Results of OLS to test effect of SPHRS on LEARN>

After running the OLS regression, we found a significant effect of the workings hours of the spouse on the wages on the male, just like previous literature on the topic. Our primary independent variable, *SPHRS*, had a statistically significant, but very small, negative effect on the log of earnings of the individual; the statistically significant negative effect of -0.000617 (controlling for presence of children) implies that for a one hour increase in the working hour of the spouse, the hourly earnings of the male can be expected to drop 6.18%. We ran a standard OLS regression because the fixed effects regression would have meant testing with only the variation in *SPHRS* within the single individual and the resulting small variation would have underestimated the effect of *SPHRS* on *LEARN*. In addition, the results of a fixed effects regression did not vary much from the OLS results.

Although the coefficient is very small, our results suggest an increase in the hours worked of the spouse leads to a decline in the earnings of the individual; however, the coefficient is too small to support the specialization hypothesis. The number of hours worked by the spouse has

been used in past studies, as a proxy for household specialization but it could be that when both the male and female are employed they hire companies or individuals to take care of the majority of household chores. The *SPHRS* variable had a negative coefficient significant at the 1% level even after controlling for the presence of children. The presence of a slightly smaller magnitude effect after controlling with *CHILD* may suggest that men become more productive due to additional financial pressures from the need to support the child, which reduces the effect of an increase in *SPHRS*. However, the difference in the coefficient is too small to make any justifiable conclusions.

The results obtained with our NLS 1979-2010 dataset suggest neither the selectivity hypothesis nor the specialization hypothesis is complete in explaining the wage premium in married individuals. However, our data is limited due to the large number of participants who dropped out of the survey, therefore leading to a highly unbalanced data set. There are many other time-variant unobservable factors that may play a significant part in determining the wage premium such as networks, luck, economic & corporate events, employer wage discrimination, and etc. A more balanced data set will be required to further test for sources of wage premiums. In addition, a probit or logit model with instrumental variables to minimize the effect of endogeneity to test for the likely of wage premiums may also be a viable way to find other sources of the wage premium.

Conclusion

In this paper, we provide evidence of the existence of the marriage premium in the United States and some of the underlying reasons for the premium. Our estimates suggest that there is a 17% marriage premium, consistent with previous literature. However, this falls to 10% in the fixed effects regression suggesting that approximately 40% of the observed premium is explained by

time-invariant individual characteristics. Further research with a more balanced dataset will be needed to determine the source of the remaining, unaccounted for, wage premium. Upon testing the specialization hypothesis using the hours worked by the spouse, we found that there is a negative effect on the earnings of the individual; a one hour increase in spouse work hours led to an approximately 6% decline on earnings. The magnitude of the effect stayed similar even after controlling for the presence of children.

Our results demonstrate the existence of the premium but do not show if higher wages induced marriage or marriage led to a higher wage; a Granger causality test may be performed to predict which caused which. Regression testing with more variables affecting one's wage (i.e. hours spent at home, hours spent on chores, costs associated with children, etc.) will be needed to draw more determinant conclusions.

References:

BARDASI, E & TAYLOR, M. (2006). Marriage and wages: a test of the specialization hypothesis. Economica, 75 (299), 569-591

BECKER, G. (1981). A Treatise on the Family. Cambridge, Mass.: Harvard University Press

CHUN, H. and LEE, I. (2001). Why do married men earn more: productivity or marriage selection? Economic Inquiry, 39, 307–19

DANIEL, K. (1992). Does marriage make men more productive? NORC Discussion Paper no. 92–2, University of Chicago.

GRAY, J. (1997). The fall in men's return to marriage: declining productivity effects or changing selection? Journal of Human Resources, 32, 481–504

HERSCH, J. and STRATTON, L. (1997). Housework, fixed effects and wages of married workers. Journal of Human Resources, 32, 285–307.

HILL, M. (1979). The wage effect of marital status and children. Journal of Human Resources, 14, 579–94

KENNY, L. (1983). The accumulation of human capital during marriage by males.Economic Inquiry, 22, 223–31

Appendix 1: Summary of Variables

Variable Name	Information/Question	Units of Measurement	Index
MARRIED	Constructed Marital Status	Dummy	1 = Currently Married 0 = Otherwise
DEGREE	Degree-holding Status of individual	Dummy	1 = With college degree 0 = Otherwise
EARN1	Earnings on main job (Job #1)	Actual Dollars & cents	
NUMCH	Number of children	Number (Count)	
CHILD	Presence of Children	Dummy	1 = 1 or more children 0 = no children
SPHRS	Hours per week worked by spouse	Hours per week	
TENURE	Total Tenure in weeks on current job	Number of Weeks	
YSCH	Number of years in school	Actual Years	
DISB	Work limited due to health dummy	Dummy	1=Yes 0=No
URBAN	Live in an urban area	Dummy	= 1 if Urban, 0 Rural
FAMSIZE	Size of Family	Number of members in family	
LOCNE	Region Dummy	Dummy	= 1 if region is Northeast, 0 otherwise
LOCNC	Region Dummy	Dummy	=1 if region is NorthCentral, 0 otherwise
LOCWT	Region Dummy	Dummy	=1 if region is West, 0 otherwise
LOCSU	Region Dummy	Dummy	= 1 if Region is South, 0 otherwise
PRTHH	Residence Dummy	Dummy	= 1 if Lives in Parent's Household
SINGLE	Single Dummy	Dummy	= 1 if single, 0 otherwise
WIDOW	Widow Dummy	Dummy	= 1 if widowed, 0 otherwise
DIVOR	Divorce Dummy	Dummy	= 1 if divorced, 0 otherwise
SEPA	Seperated Dummy	Dummy	= 1 if seperated, 0 otherwise
Y1980-Y2010	Time Dummies	Dummy	= 1 if observation is from YXXX, 0 otherwise
LOCNE	Region Dummy	Dummy	= 1 if individual lives in Northeast, 0 otherwise

LOCSU	Region Dummy	Dummy	= 1 if individual lives in south, 0 otherwise
LOCNC	Region Dummy	Dummy	= 1 if individual lives in northcentral 0 otherwise
LOCWT	Region Dummy	Dummy	= 1 if individual lives in west, 0 otherwise
BLACK	Race Dummy	Dummy	= 1 if individual is black
HISP	Race Dummy	Dummy	= 1 if individual is hispanic

Appendix 2: Correlation Tables

Correlation Matrix

	LEARN	MARRIED	SPHRS	CHILD	NUMCH	AGE	FAMSIZE
LEARN	1.0000						
MARRIED	0.3581	1.0000					
SPHRS	0.2810	0.7012	1.0000				
CHILD	0.0863	0.5201	0.2812	1.0000			
NUMCH	0.0878	0.4654	0.2132	0.8660	1.0000		
AGE	0.6154	0.3324	0.2844	-0.0137	0.0268	1.0000	
FAMSIZE	-0.1328	0.1695	0.0278	0.2267	0.3158	-0.1562	1.0000
AGE2	0.5887	0.3036	0.2633	-0.0657	-0.0227	0.9899	-0.1326
DISB	-0.0053	-0.0187	-0.0140	-0.0162	-0.0052	0.0637	-0.0126
TENURE	0.4437	0.2906	0.2279	0.0390	0.0559	0.5212	-0.0149
BLACK	-0.0828	-0.1522	-0.0880	-0.0605	-0.0450	0.0559	0.0337
HISP	0.0025	0.0345	-0.0063	0.0585	0.0744	0.0281	0.1080
YSCH	0.3658	0.1300	0.1433	-0.0561	-0.0613	0.2379	-0.1808
DEGREE	0.3111	0.1332	0.1239	-0.0286	-0.0297	0.1734	-0.1151
URBAN	0.0434	-0.0973	-0.0687	-0.0413	-0.0290	-0.0381	-0.0206
LOCNE	0.0322	-0.0575	-0.0512	-0.0376	-0.0429	-0.0344	-0.0013
LOCSU	-0.0527	0.0185	0.0378	0.0073	-0.0025	0.0316	0.0019
LOCWT	0.0529	0.0060	-0.0098	0.0110	0.0223	0.0024	-0.0115
PRTHH	-0.0816	-0.2444	-0.1811	-0.1450	-0.1320	-0.0503	0.1410

	AGE2	DISB	TENURE	BLACK	HISP	YSCH	DEGREE
AGE2	1.0000						
DISB	0.0644	1.0000					
TENURE	0.5224	-0.0151	1.0000				
BLACK	0.0533	0.0189	-0.0512	1.0000			
HISP	0.0274	0.0079	0.0092	-0.2757	1.0000		
YSCH	0.2141	-0.0390	0.1310	-0.0506	-0.1280	1.0000	
DEGREE	0.1563	-0.0319	0.0935	-0.0879	-0.0861	0.7355	1.0000
URBAN	-0.0429	-0.0002	-0.0512	0.0742	0.1275	0.0712	0.0461
LOCNE	-0.0335	0.0035	0.0005	-0.0200	-0.0328	0.0262	0.0267
LOCSU	0.0303	0.0005	-0.0260	0.2493	-0.0825	-0.0520	-0.0297
LOCWT	0.0020	0.0033	-0.0172	-0.1773	0.3211	-0.0105	-0.0270
PRTHH	-0.0662	0.0213	-0.0852	0.0975	0.0212	-0.0427	-0.0571

	URBAN	LOCNE	LOCSU	LOCWT	PRTHH
URBAN	1.0000				
LOCNE	0.0973	1.0000			
LOCSU	-0.1494	-0.3586	1.0000		
LOCWT	0.1326	-0.2297	-0.3893	1.0000	
PRTHH	0.0355	0.0491	0.0067	-0.0267	1.0000