5-21-2002

Wage Growth Differentials Between White and Non-White Males by Industry

Stephen Ransom Barnes

Wage Growth Differentials Between White and Non-White Males by Industry

An Undergraduate Thesis:

by

Stephen R. Barnes

Under the direction of

Dr. Robert J. Newman

Submitted to the Honors College of
Louisiana State University and Agricultural and Mechanical College
in partial fulfillment of the requirements for
Upper Division Honors Distinction
Economics

May 21, 2002

Wage Growth Differentials Between White and Non-White Males by Industry

Stephen R. Barnes

Senior Thesis
Louisiana State University
2001-2002

This paper studies the differences in wage growth from accumulated work experience between white and non-white workers within different industries. The study uses CPS data from 1990 to 1997, which is restricted to males who are terminal high school graduates working in non-government jobs. A simple human capital earnings model is specified within each industry to measure the difference in returns to experience between white and non-white workers. The results indicate that there are differences in the return to experience in some industries while returns to experience have converged in others.

I. Introduction

Economists have long been interested in learning what causes wage growth over time. One particular area that has received a great deal of attention in recent years has been the study of differences in wage growth across races. Historically, blacks have suffered from lower wages in their careers with lower starting salaries than whites and lower wage growth over time than whites. This combination continues to drive a wedge of inequality between the wages of different races and has thus become a topic of much study. While differences in starting salaries between races may seem to be important as a basis for equality in compensation among races, how those wages grow over time is much more relevant to the overall wage equality and earning power of black workers. By examining the issue at a greater level of detail, it may be possible to determine with more precision what causes these differences to persist. This paper seeks to expand on the

current body of literature by focusing on the differences in returns to experience stratified by industry.

In order to effectively study differences in wage growth by race, a solid understanding of the nature of wage differentials is necessary. Rees and Shultz conducted a very comprehensive study of wages in the 1960's. Their book looks for an explanation of wage differentials and structural characteristics of markets for different occupations, among other topics. They begin by presenting the dominant theory of wage differentials, which they call the classical theory of compensating differentials (most notably expressed by Adam Smith in his chapter "On Wages and Profits in the Different Employments of Labour and Stock" – chapter 10 of book 1). This theory states that within a neighborhood there is a tendency toward equality in the advantages and disadvantages of different jobs with the wage taken as one of several potential advantages. This theory provides more flexibility than simply considering the absolute wage and allows for the assumption that man seeks to maximize utility rather than money. A problem with this theory, however, is that economists are then expected to define these advantages and disadvantages in order to be able to make useful predictions of behavior, and some consider this to be ultimately impossible. They cite several papers suggesting that these advantages and disadvantages do not tend toward equality even when they have been carefully defined.

Rees and Shultz go on to discuss other issues that are traditionally seen as possible contributors to wage differentials. Alfred Marshall's "efficiency earnings", which measures earnings per unit of output rather than earnings per hour, is brought up as a way to better account for differences in the quality of the worker. Under this concept, earnings per unit for workers with the same occupation tend to equalize. Also, perceived worker qualities, rather than actual worker qualities, are mentioned as a possible source

of wage differentials (especially among races, genders, or ethnic groups where prejudice is a potential problem). The potential contribution of union activity to differences in wages is brought up as an example of a modern anti-competitive force inhibiting the classical equilibrium. These examples highlight just a few of the different measures and factors that have been developed to try to isolate the driving forces of wage differentials, but they represent what the body of research had found to be the most significant at the time.

Rees and Shultz test if these theories hold in empirical data. In order to accomplish this, they interview seventy-five employers in Chicago and the surrounding areas. From these interviews and personnel records, they create their own sample of workers in fourteen selected occupations. In analyzing this data, they regress earnings per hour on independent variables for the worker such as age, race, education, distance the place of work is from home, and establishment variables such as size, industry, or location using ordinary least squares. They find that wage differentials can be accounted for in part by location (distance traveled), and also to some extent by differences in schooling and previous work experience, but seniority seems to be the most important factor. They also find clear evidence of wage differentials in favor of males over females and whites over nonwhites (this is after being hired so their results do not show discrimination to entry). These differences are found after allowing for the effects of differences among workers in age, seniority, schooling, and experience, and they are therefore able to conclude that significant differences do in fact exist. This evidence supports the classical theory of compensating differentials, especially with regard to those differentials associated with location. This evidence also supports Alfred Marhsall's concept of "efficiency wages" because of a clear relationship between wages and worker

quality measures. But, the largest determining factor for differentials in this study is seniority, which is the general focus for the more recent literature. I take this book as a starting point for my study of wage differentials. Though there had been numerous studies before Rees and Shultz, this book's econometric approach combined with its inclusion of both race and job specific variables provides a good example for a paper such as this to follow.

Since this study, the body of research surrounding wage differentials has continued to grow. The more recent literature has focused on trying to pinpoint exactly what effects cause the largest proportion of observed differentials. This is a challenging question to answer, because many of the proposed explanations are highly related to one another, making it difficult to determine which factors are actually driving wage growth. Early researchers on the topic, like Rees and Shultz, concluded that job seniority has a strong positive relationship with wage rates, but other researchers have proposed numerous explanatory factors including experience, race, gender, marital status, region, occupation, and union activity among others. Recent race-wage studies have focused most on the measurement of the return to experience.

Altonji and Shakotko examine the returns to experience in more detail, using a simple instrumental variables scheme to deal with the fact that tenure is likely to be related to unobserved individual and job characteristics that affect the wage. They assume that the wage of individual i in job j at period t is determined according to the following equation:

$$W_{ijt} = b_0 X_{ijt} + b_1 T_{ijt} + b_2 T^2_{ijt} + b_3 OLDJOB_{ijt} + \varepsilon_{ijt},$$

$$\varepsilon_{ijt} = \varepsilon_i + \varepsilon_{ij} + \eta_{ijt}$$

The variable W_{ijt} is the log of the real wage. X_{ijt} is a vector of characteristics of the person and the job and includes labour market experience. Tenure is measured by T_{ijt}, T^2_{ijt}, and $OLDJOB_{ijt}$. T_{ijt} is the level and T^2_{ijt} is the square of the number of years person I has held job j as of time t. $OLDJOB_{ijt}$ is 1 if $T_{ijt} > 1$ and O otherwise and is included so that the wage response to the first year of tenure is not restricted by the quadratic specification of the tenure profile. The variable ε_{ijt} is the error term and consists of a fixed individual effect ε_i, a fixed job match effect ε_{ij}, and a transitory component η_{ijt}.

Altonji and Shakotko state that even this instrumental variables scheme, however, is subject to bias in overestimating the partial effect of experience and subsequently underestimating the partial effect of tenure. They also recognize the potential problem with the instrumental variables, which concerns measurement error in the tenure variable for their data. This is likely to bias downward their estimates of the tenure profile. A third possible source of bias lies in the possibility that the tenure profile of wages varies across people and jobs. They do find, however, that these biases are most likely small. Based on this new analysis, they conclude that the partial effect of tenure on wages is small, with general work experience accounting for most of the wage growth in a career. Altonji and Shakotko further find that for blacks, the effect of tenure is very similar to what it is for whites, but the effect of experience is much smaller than for whites.

Another article, written by Light and Ureta, studies the issue of wage growth with a more carefully specified measure of experience. When conducting studies on the partial effects of experience and tenure on wages, problems within the data, or even small measurement errors, can contribute to drastically different results. They use a new measure of experience giving continuous experience by taking into consideration time spent not working throughout a career. Their paper focuses on wage differentials by

gender and is therefore quite concerned with interpersonal differences in the timing of past work experience and differences in the total amount of experience accumulated. Studying this measure of experience not only paints a clearer picture of the nature of gender differentials in wage growth, but also more clearly defines the nature of wage growth in general. Their work history model includes variables that measure the fraction of time worked in each of the previous years since the beginning of the career (called X_m describing the fraction of time worked m years ago). To account for years when the worker did not work, they also include dummy variables $O_1 - O_5$ for the last five years of work (not including more years of history because they failed to provide any explanatory value), which are set to one if X_m is zero and zero otherwise. The O_m measure the wage penalty incurred for not working for an extended period of time. Because of missing information, they also include dummy variables called $Q_1 - Q_m$. When X_m equals zero simply because of a lack of information on weeks worked, Q_m is set to one, and it equals zero otherwise.

This work history model is compared to potential experience and actual experience models to estimate the effect of timing of employment on wages. The actual experience model replaces X_m and O_m with the cumulative number of years worked (obtained by summing the X_m). The potential experience model defines experience as "age – schooling – 6" to provide a crude measure of experience with which to compare. The potential experience model is commonly used with Current Population Surveys, the decennial censuses, and other data sets that lack information on actual experience and school enrollment dates.

Light and Ureta find that the conventional measures of actual or potential experience may produce biased estimates of the effects of early-career work experience

on earnings. For both men and women, they find that using the work history array results in higher returns to experience and lower returns to tenure. The work history model both measures experience more precisely and imposes far less constraint (such as the use of quadratic form) on the shape of the wage-experience profile than the traditional specifications. This new specification also adds insight into the gender wage gap. For men and women who have the same cumulative experience, differences in timing account for as much as 12% of the raw wage gap and for as much as 30% of the portion of the gap explained by differences in the timing of and returns to experience. If there were any differences in the timing of work between blacks and whites, these considerations would certainly be relevant in creating a more accurate picture of the structure of race wage differentials.

There is an increasingly large body of literature that focuses on the effects of tenure and experience as the primary causes of wage differentials. In particular, the difference in wage growth between races has been a topic of much study in recent years. It is clear that ensuring equality in job opportunities and compensation has long been a goal of many public policies (most notably affirmative action). But evidence continues to show that wage growth for blacks is lower than wage growth for whites. There is no consensus about the causes for these differences. This is clearly an area that requires further study to discover why these differences persist so that more effective programs or controls can be developed to encourage equality.

Wolpin studied the determinants of black-white differences in early employment focusing on search, layoffs, quits, and endogenous wage growth. He focuses on the transition from school to full-time employment and subsequent labor mobility during the

first five post-schooling years to find what differences may exist between races. In order to do this, he developed a constrained optimization model of labor force dynamics. It is implemented empirically, integrating features of models previously described in the literature. He modeled employee decisions as maximizing wealth subject to the following constraints:

1. For the unemployed, wage offers arrive stochastically from a known distribution. The probability per period of receiving an offer depends on the employment history.
2. Once an offer is accepted, there is no further stochastic variation in the wage at that employer. Wage offers are still received from other firms as described above.
3. For the employed, there is a history-dependent positive probability per period of layoff and while on layoff a positive probability per period of a recall offer that depends on the employee's career to date in a known way.
4. There is a fixed and known unemployment insurance benefit available during periods of layoff (for a fixed duration).
5. Wage offers depend on the level of general human capital as measured by the number of prior periods of employment and on the level of specific human capital.

Using this specification, Wolpin finds that work history has an important effect on the labor market environment faced by an individual both through its effect on the likelihood of receiving job offers and on the level of the wage offer received. He does further conclude that workers can alter their current constraints by past behavior but the labor market environment and the effect of history on it evidently differ by race.

A recent article by Bratsberg and Terrell studies the differences in returns to tenure and experience across races. This paper uses three estimation techniques (ordinary least squares, Altonji and Shakotko, and Topel estimators) to minimize potential biases that might arise due to correlations of experience and tenure to unobservable determinants of wage. In light of recent warnings that estimates of returns may be sensitive to measurement errors, they study annual data for young black and white males from 1979-1991 drawn from nonmilitary sub samples of the National Longitudinal Survey of Youth and the accompanying workhistory data set, which provides them with weekly data on work status of all subjects, thereby allowing precise measures of on-the-job tenure and labor market experience. In addition, to avoid problems arising from the sensitivity of estimates to assumptions in constructing on-the-job tenure and labor market experience, they follow Wolpin by focusing on terminal high school graduates (excluding GEDs and anyone who enrolled in college during any period in the sample) and assume that the career begins at high school graduation. To ensure data quality and avoid problems in measuring tenure and experience, they count only weeks the subjects report working so there is no question about measurement. When comparing this definition of experience (actual experience) to the NLSY definition of experience as time elapsed since the start of the career (potential experience) they found that white and black workers had equal potential experience, but white workers had significantly higher actual experience than black workers. This large difference shows how important exact measures of experience are to obtaining useful results in measuring wage growth differences.

Some of the important theoretical considerations of wage growth are examined carefully in this paper. Job match quality has been studied as a driving force of wage

growth. As workers quit jobs that are poor matches and stay in jobs that are good matches, over time workers will improve the quality of their matches. One explanation of the relationship between tenure and experience focuses on job match quality. With this consideration, they specify the following wage equation:

$$w_{ijt} = Z_i\pi + T_{ijt}\alpha + X_{ijt}\beta + \varepsilon_j + \eta_{ij} + \nu_{ijt}$$

where i is an individual index; j indexes jobs; and t denotes time period. The equation uses w_{ijt} to denote jog wages; Z_i is a vector of observable individual characteristics; T_{ijt} denotes tenure on the job; and X_{ijt} represents total accumulated labor market experience. The constant ε_j is included to capture unobserved heterogeneity across individuals and η_{ij} to reflect the quality of job match between worker and firm.

For all three estimators, they found that black workers receive far lower returns to general experience than white workers. This conclusion, however, still leaves the question open as to what causes these differences in returns. Some of their proposed answers include Lazear's hypothesis that employers respond to affirmative action by providing on-the-job training in the place of higher wages for white workers leading to an accumulation of general rather than specific human capital for white workers. Becker's discrimination model is also discussed as well as a model presented by Farmer and Terrell in which employers incorrectly perceive differences in worker ability across races.

Within the study of race wage differentials, researchers have also attempted to explain whether those differentials are tending to equality, as we would expect in either the classical model or Alfred Marshall's "efficiency wages" as discussed in Rees and Shultz. Grogger studies wage convergence and the underlying causes of wage differentials in a paper on school quality. He notes that the trend toward racial wage

convergence stopped around 1980. To estimate if this can be explained by school quality, he analyzes data drawn from the NLSY using the following equation.

$$Y_{it} = \alpha_t BLACK_{it} + X_{it}\beta_t + F_{it}\theta_{Ft} + A_{it}\theta_{At} + Q_{it}\theta_{Qt} + \varepsilon_{it},$$

$$i = 1,...,n_t;\ t = 1979,\ 1986$$

In this equation, Q is a variable that represents an index of school quality. BLACK is a dummy variable for black workers. F represents an index of family background. A is an index of ability measures.

He finds that school quality fails to explain the recent break in the trend toward racial wage equality and that measurable school inputs have little effect on wages. Because there were only small differences in the schools attended by blacks and whites in the 1970's, it was difficult to find any effect on wages. Differences in schools since that time had become even less noticeable up until Grogger's study, but since then any effects of school quality have not been measured. There still has not been any conclusive evidence to explain why these trends exist so there clearly exists the need for further research.

These articles highlight the recent study of wage differentials by race. In general, there seems to be agreement that experience has a large explanatory value for wage growth over time and is therefore an important factor in studying differentials across races. Researchers have also continued to search for other factors relevant to differences among races that could explain why there has been little or no convergence in the wages of workers doing similar work in similar areas who are simply of different races.

To shed more light on the subject of race wage differentials, I plan to measure the effect on wages of experience (among other variables) and stratify the data by industry.

Though this may not provide any explanations for the persistence of these differences among races, it will create a clearer picture of where these differences persist (assuming that they will not be uniform across all industries). This in turn, will provide direction for future research on why they tend to be larger in some areas than in others and hopefully provide some intuition as to what, in general, causes the differences.

The paper is organized as follows. Section II contains a description of the data and discusses important features of the Current Population Survey from 1979-1997. Section III presents the model for wages. Section IV discusses the preliminary results. Section V presents conclusions and discusses the potential for future research.

II. Data

The data are taken from the Current Population Survey (CPS). The CPS is a monthly survey of about 60,000 households. An adult at each household is asked to answer questions about all persons in the household. Since 1979, houses interviewed in months 4 and 8 have been asked questions about earnings and hours worked. These observations are gathered by the BLS and merged into a single file that they call the "Annual Earnings Files". The data for this study are drawn from this file for the years 1990 through 1997. This provides more recent data than the studies previously discussed, but also presents problems that arise from the changes in the data over those years.

Initially, the full data set for each year was filtered to exclude persons under the age of 16 and over the age of 64. All persons who were not currently working were also excluded during the filtering to ensure more accurate results. See appendix A for the SAS program used in this process.

In addition to this first initial rough cut, I narrowed the sample size even further for the analysis to control for some of the other potential explanatory factors of wage growth discussed above. Following Wolpin and Bratsberg and Terrell, I focused only on terminal high school graduates to control for the effects of schooling. All females were then deleted from the data set to control for the differences in accumulated experience of men and women. I then removed all farming and related workers and those employed in government jobs.

Once I had narrowed down the sample for each year, I checked to ensure that each industry still had enough observations to provide reliable results. Three industries (Mining, Personal Services, and Entertainment and Recreational Services) were removed from the sample due to insufficient sample size (for example, Mining had 11 non-white workers for 1997). It was decided that any industry that had less than 50 observations for either race in 1997 would be deleted. The eight industries that were analyzed are Construction, Durable Goods, Nondurable Goods, Transportation and Communication, Wholesale and Retail Trade, Finance, Insurance, and Real Estate, Business and Related Services, and Professional and Related Services. Finance, Insurance and Real Estate was close to this number in 1997 and even dropped below this number in other years, but was still included and so the results from this industry must be qualified. The sample size of each group within each industry is provided with the results in Section IV.

After the sample was defined, I began constructing variables to use in the model. This was a challenging process, however, because during the time period under study, the questions that were asked changed from year to year. In addition, even for those questions that didn't change, the choices for answers changed for some years. It was therefore necessary to carefully test each definition for the given measure of interest to

ensure that there would be no large errors from choosing one definition over another. This was done through computing descriptive statistics and then by visually comparing the min, max, mean and standard deviations of the different definitions across years. By the end of the process, I concluded that the differences that might exist between possible measures were negligible for my purposes.

The limitations of the CPS data led to a few decisions about variable definition and exclusion of some potentially important factors in the wage equation. The previously discussed literature repeatedly shows the importance of a carefully defined measure of experience. However, the CPS does not contain a detailed measure of experience so a rough measure of potential experience is used following the standard estimate described in Light and Ureta of "age – schooling – 6." In addition, tenure is a factor that is consistently included in estimations of this sort and since the CPS has no measure of tenure, there was no way to incorporate this into the model. These are two important factors to consider when discussing the results of the regression, but what may be lost in measurement error using the CPS is at least somewhat offset by the large sample that allows for the separate analysis of each industry. I also defined dummy variables to control for occupation, marital status, and region.

III. Model Specification

The objective of this paper is to estimate if there exist any differences in the return to experience for white and non-white workers. The definition of variables was done with great care to control for as many potential. As described above, I considered multiple measures for a given variable where available and chose the one that was most consistent across years. The model is as follows:

$$W = \alpha_0 + \alpha_1 EXP + \alpha_2 RACE*EXP + \alpha_3 EXP^2 + \alpha_4 RACE*EXP^2 + \mathbf{X}\beta + \varepsilon$$

where W denotes log wages, the α_i measure experience, **X** is a vector of observable individual characteristics, and ε is the error term. RACE is a dummy variable that is 1 if the worker is white and 0 otherwise. This equation allows for the simultaneous estimation of the return to experience of white and non-white workers. The coefficients for EXP and EXP^2 measure the returns to experience of non-white workers while RACE*EXP and $RACE*EXP^2$ measure the additional return to experience that we observe for whites over non-whites. The vector **X** includes the following observable characteristics: marital status, occupation, and region.

IV. Preliminary Results

The results of the regression analysis are presented in tables 1 through 8 and representations of these results are shown graphically in charts 1 through 8. These tables and charts are presented beginning on the next page as pages 16-27 and correspond to each of the eight industries studied.

Table 1

Construction Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0087	0.0123	0.0171**	0.0269*	0.0325*	0.0337*	0.0367*	0.0302*
	(1.03)	(1.37)	(1.68)	(3.03)	(3.64)	(3.80)	(4.01)	(3.64)
Race*Exp	0.0223*	0.0187*	0.0246*	0.0097	0.0052	0.0002	-0.0017	0.0020
	(2.56)	(1.99)	(2.35)	(1.05)	(0.56)	(0.02)	(0.18)	(0.23)
Exp^2	-0.0001	-0.0001	-0.0001	-0.0003	-0.0006*	-0.0006*	-0.0006*	-0.0004*
	(0.38)	(0.41)	(0.49)	(1.50)	(2.52)	(2.65)	(2.68)	(2.16)
$Race*Exp^2$	-0.0004*	-0.0004	-0.0006*	-0.0002	-0.0001	0.0000	-0.0000	-0.0001
	(1.97)	(1.62)	(2.07)	(1.04)	(0.30)	(0.06)	(0.01)	(0.26)
Adj. R^2	0.2519	0.2850	0.2924	0.3000	0.2703	0.2435	0.2729	0.2901
No. of Whites	2704	2405	2230	2334	2271	2252	2039	2202
No. of Non-Whites	242	219	179	235	210	247	206	217

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 2
Wholesale and Retail Trade Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0217*	0.0245*	0.0168*	0.0151*	0.0185*	0.0346*	0.0242*	0.0266*
	(4.44)	(5.44)	(3.71)	(3.20)	(3.63)	(7.60)	(4.93)	(5.86)
Race*Exp	0.0082	0.0091**	0.0102*	0.0154*	0.0120*	-0.00411	0.0030	0.0020
	(1.58)	(1.90)	(2.13)	(3.07)	(2.20)	(0.83)	(0.57)	(0.41)
Exp^2	-0.0003*	-0.0004*	-0.0002**	-0.0002**	-0.0002	-0.0006*	-0.0004*	-0.0005*
	(2.72)	(3.54)	(1.79)	(1.85)	(1.47)	(5.11)	(2.92)	(4.31)
$Race*Exp^2$	-0.0002	-0.0002*	-0.0003**	-0.0003*	-0.0003*	0.0000	-0.0001	-0.0001
	(1.55)	(2.02)	(2.40)	(2.59)	(2.48)	(0.37)	(0.86)	(0.43)
Adj. R^2	0.3223	0.3341	0.3333	0.3445	0.2844	0.3059	0.3187	0.2920
No. of Whites	3294	3275	3429	3337	3091	3062	2859	2812
No. of Non-Whites	542	543	558	550	579	630	483	506

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 3

Professional and Related Services Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0220*	0.0195*	0.0243*	0.0214*	0.0049	0.0288*	0.0173	0.0209**
	(2.34)	(2.26)	(2.46)	(2.01)	(0.40)	(2.67)	(1.43)	(1.93)
Race*Exp	-0.0006	0.0027	0.0053	-0.0192	0.0256**	-0.0028	0.0138	0.0030
	(0.06)	(0.27)	(0.49)	(1.65)	(1.89)	(0.23)	(1.06)	(0.24)
Exp^2	-0.0004*	-0.0003*	0.0005*	-0.0004	0.0002	-0.0004**	-0.0001	-0.0003
	(2.03)	(1.32)	(2.35)	(1.47)	(0.78)	(1.67)	(0.52)	(1.33)
$Race*Exp^2$	0.0001*	-0.0001	0.000	-0.0003	-0.0007*	0.0000	-0.0004	-0.0000
	(0.37)	(0.51)	(0.07)	(1.03)	(2.37)	(0.04)	(1.43)	(0.12)
Adj. R^2	0.3497	0.3133	0.3610	0.2833	0.2632	0.2733	0.3023	0.2076
No. of Whites	535	518	572	538	537	516	457	475
No. of Non-Whites	167	170	130	158	169	165	117	160

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 4

Nondurable Goods Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0357*	0.0235*	0.0295*	0.0271*	0.0410*	0.0405*	0.0357*	0.0302*
	(5.97)	(4.15)	(4.32)	(4.67)	(6.53)	(6.54)	(4.82)	(4.17)
Race*Exp	-0.0032	0.0125*	0.0057	0.070	-0.0102	-0.0116**	-0.0030	0.0018
	(0.51)	(2.03)	(0.78)	(1.11)	(1.49)	(1.70)	(0.38)	(0.23)
Exp^2	-0.0007*	-0.0003*	-0.0005*	-0.0004*	-0.0007*	-0.0007*	-0.0006*	-0.0005*
	(4.84)	(2.48)	(3.00)	(2.82)	(5.00)	(5.13)	(3.34)	(2.84)
$Race*Exp^2$	0.0001	-0.0003**	-0.0001	-0.0001	0.0003**	0.0003**	0.0000	-0.0000
	(0.95)	(1.90)	(0.38)	(0.81)	(1.74)	(1.91)	(0.20)	(0.25)
Adj. R^2	0.2241	0.2459	0.2402	0.2741	0.2368	0.2200	0.2307	0.2009
No. of Whites	2136	2092	2107	1966	1847	1779	1600	1516
No. of Non-Whites	405	369	334	360	377	368	263	303

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 5
Transportation and Communication Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0276*	0.0400*	0.0327*	0.0329*	0.0426*	0.0078*	0.0229*	0.0256*
	(3.29)	(4.50)	(3.66)	(3.68)	(4.21)	(4.74)	(2.61)	(3.27)
Race*Exp	0.0093	0.0034	0.0042	0.0031	-0.0115	0.0056	0.0199*	0.0143**
	(1.04)	(0.37)	(0.44)	(0.32)	(1.07)	(0.65)	(2.08)	(1.66)
Exp^2	-0.0005*	-0.0008*	-0.0005*	-0.0005*	-0.0008*	-0.0006*	-0.0004**	-0.0004*
	(2.27)	(3.76)	(2.76)	(2.45)	(3.50)	(3.69)	(1.89)	(2.37)
$Race*Exp^2$	-0.0002	-0.0000	-0.0001	-0.0001	0.0003	-0.000	-0.0004	-0.0002
	(0.74)	(0.04)	(0.33)	(0.32)	(1.23)	(0.19)	(1.61)	(1.20)
Adj. R^2	0.2319	0.2279	0.2351	0.2509	0.2147	0.2236	0.2632	0.2894
No. of Whites	1682	1613	1617	1490	1498	1372	1242	1229
No. of Non-Whites	258	218	238	242	230	272	225	230

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 6

Durable Goods Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0362*	0.0287*	0.0251*	0.0311*	0.0268*	0.0240*	0.0185*	0.0172*
	(7.38)	(5.55)	(4.52)	(5.59)	(4.86)	(4.46)	(3.23)	(2.98)
Race*Exp	-0.0010	0.0053	0.0092	0.0003	0.0089	0.0113*	0.0115*	0.0140*
	(0.19)	(0.97)	(1.57)	(0.05)	(1.53)	(1.99)	(1.89)	(2.30)
Exp^2	-0.0006*	-0.0004*	-0.0003*	-0.0004*	-0.0004*	-0.0002**	-0.0002**	-0.0002
	(5.59)	(3.03)	(2.22)	(3.23)	(3.07)	(1.75)	(1.67)	(1.50)
$Race*Exp^2$	-0.0001	-0.0002	-0.0003*	-0.0001	-0.0002	-0.0003*	-0.0002**	-0.0003*
	(0.48)	(1.52)	(2.04)	(0.45)	(1.33)	(2.62)	(1.67)	(2.10)
Adj. R^2	0.2657	0.2698	0.2609	0.2671	0.2518	0.2672	0.2314	0.2491
No. of Whites	4128	3742	3748	3538	3372	3336	2959	2956
No. of Non-Whites	472	472	485	473	486	476	388	363

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 7
Business and Repair Services Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0267*	0.0137**	0.0221*	0.0102	0.0328*	0.0180*	0.0322*	0.0077
	(3.17)	(1.77)	(2.70)	(1.15)	(3.57)	(2.32)	(4.32)	(1.01)
Race*Exp	0.0027	0.0143**	0.0088	0.0146	0.0014	0.0137	0.0054	0.0184*
	(0.29)	(1.67)	(0.99)	(1.52)	(0.14)	(1.60)	(0.64)	(2.19)
Exp^2	-0.0006*	-0.0002	-0.0005*	-0.0001	-0.0007*	-0.0003**	-0.0005*	-0.0000
	(2.89)	(1.13)	(2.53)	(0.49)	(3.32)	(1.72)	(2.99)	(0.19)
$Race*Exp^2$	-0.0001	-0.0003	-0.0001	-0.0003**	0.0001	-0.0002	-0.0002	0.0004*
	(0.40)	(1.52)	(0.70)	(1.36)	(0.43)	(1.22)	(0.79)	(2.23)
Adj. R^2	0.2769	0.2927	0.2624	0.2498	0.2421	0.2531	0.3004	0.2494
No. of Whites	988	942	997	920	940	945	864	851
No. of Non-Whites	186	207	184	191	201	220	209	205

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Table 8
Finance, Insurance, and Real Estate Parameter Estimates

	1990	1991	1992	1993	1994	1995	1996	1997
Exp	0.0389*	0.0128	0.0306**	0.0352**	0.0444*	0.0336*	0.0160	0.0408*
	(2.95)	(0.92)	(1.71)	(1.89)	(2.39)	(1.65)	(0.74)	(2.70)
Race*Exp	0.0048	0.0183	0.0279	-0.0064	-0.0047	-0.0009	0.0251	-0.0155
	(0.33)	(1.18)	(1.44)	(0.32)	(0.23)	(0.04)	(1.05)	(0.93)
Exp^2	-0.0006**	-0.0001	-0.0006*	-0.0007**	-0.0008**	-0.0008*	-0.0001	-0.0009*
	(1.91)	(0.30)	(1.63)	(1.75)	(1.74)	(1.59)	(0.12)	(2.75)
$Race*Exp^2$	-0.0003	-0.0005	-0.0006	0.0000	0.0002	0.0002	0.0006	0.0005
	(0.88)	(1.36)	(1.49)	(0.11)	(0.30)	(0.30)	(1.11)	(1.39)
Adj. R^2	0.2868	0.2259	0.2970	0.2238	0.2676	0.0632	0.1347	0.2121
No. of Whites	211	217	203	224	200	199	195	214
No. of Non-Whites	57	64	45	43	42	72	46	55

Note: Other variables in model include marital status, occupation and region dummies.
* Significant at .05
** Significant at .10
(|t| – value)

Chart 1.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Construction

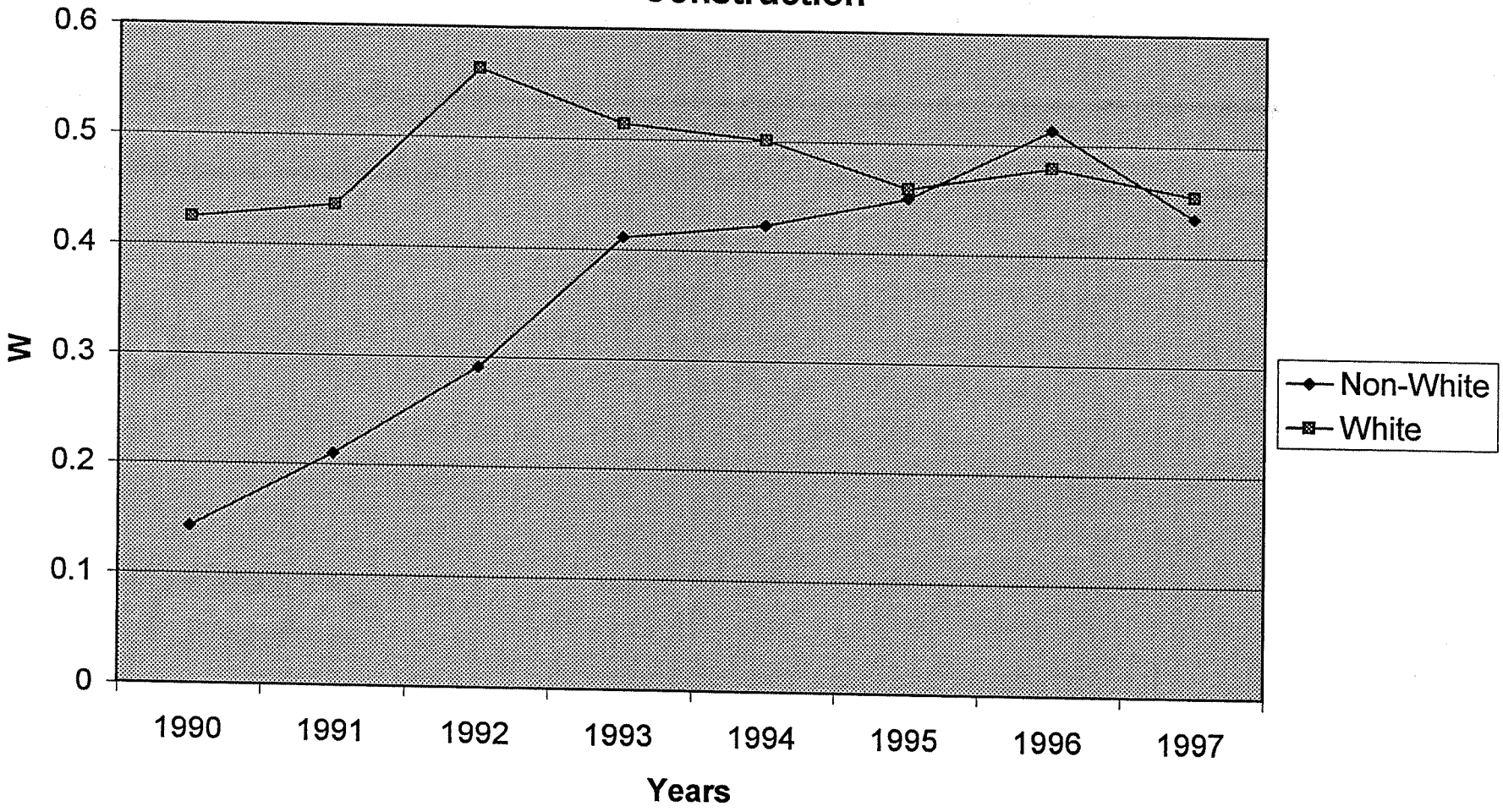


Chart 2.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Wholesale and Retail Trade

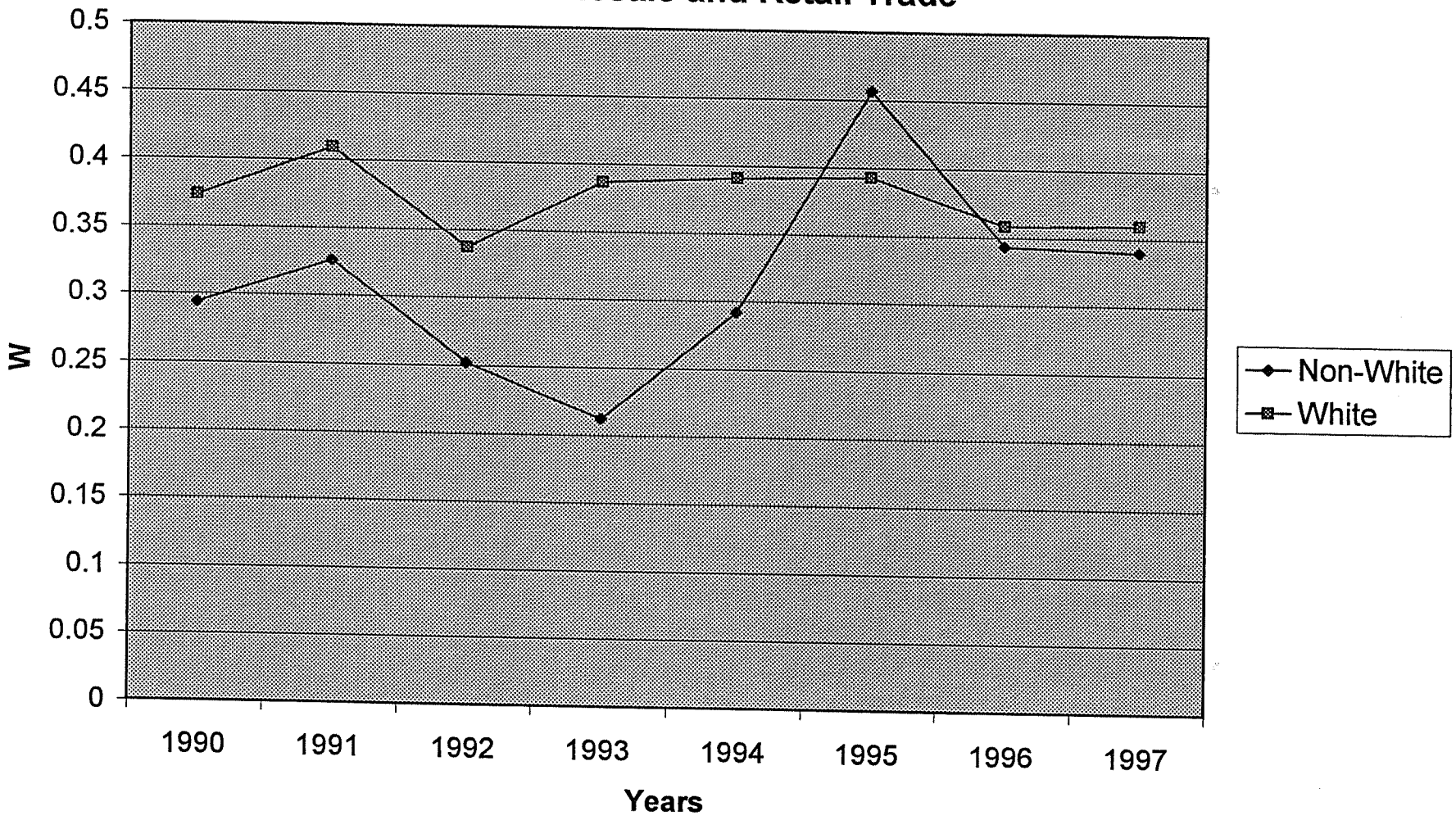


Chart 3.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Professional and Related Services

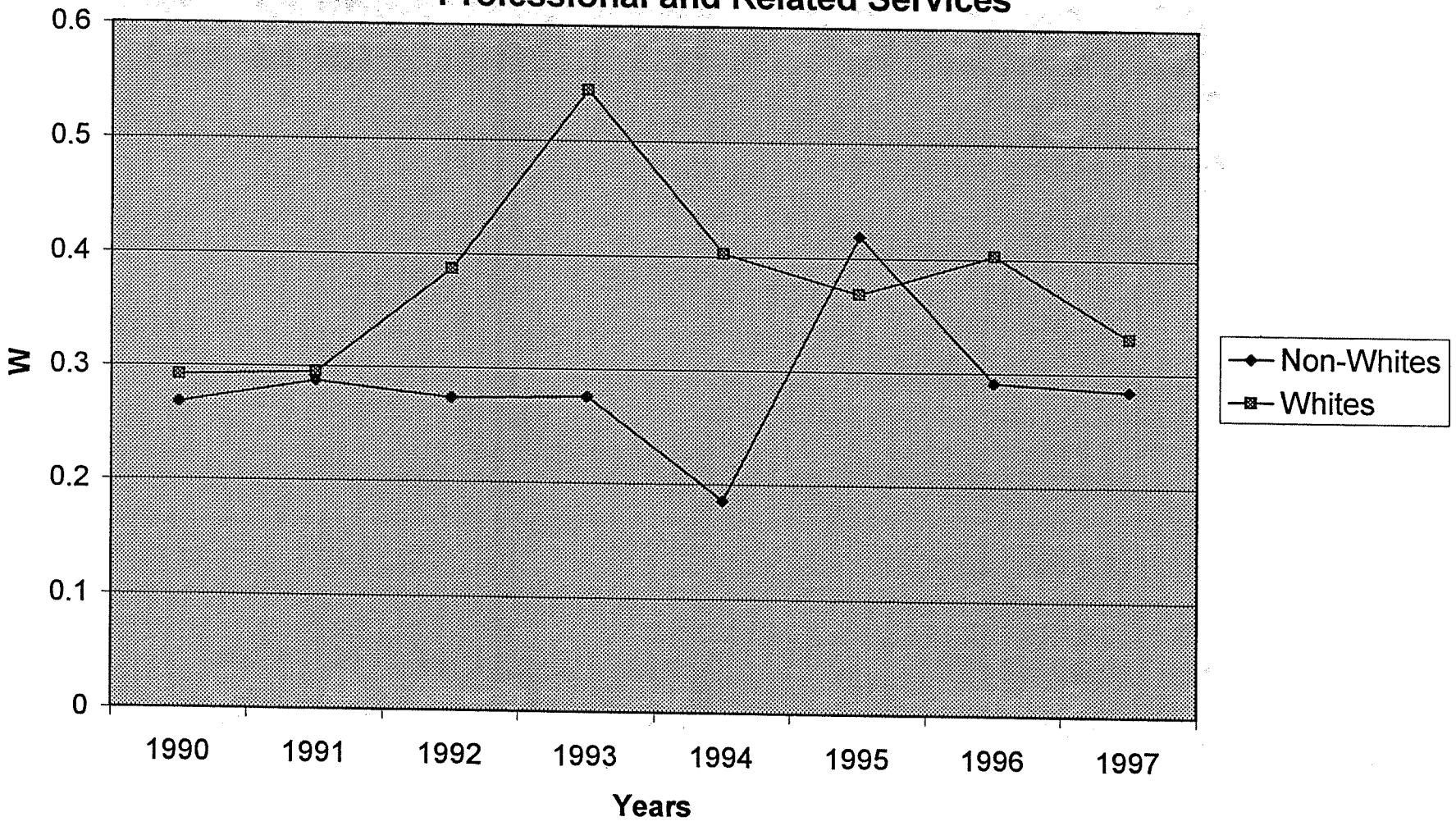


Chart 4.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Nondurable Goods

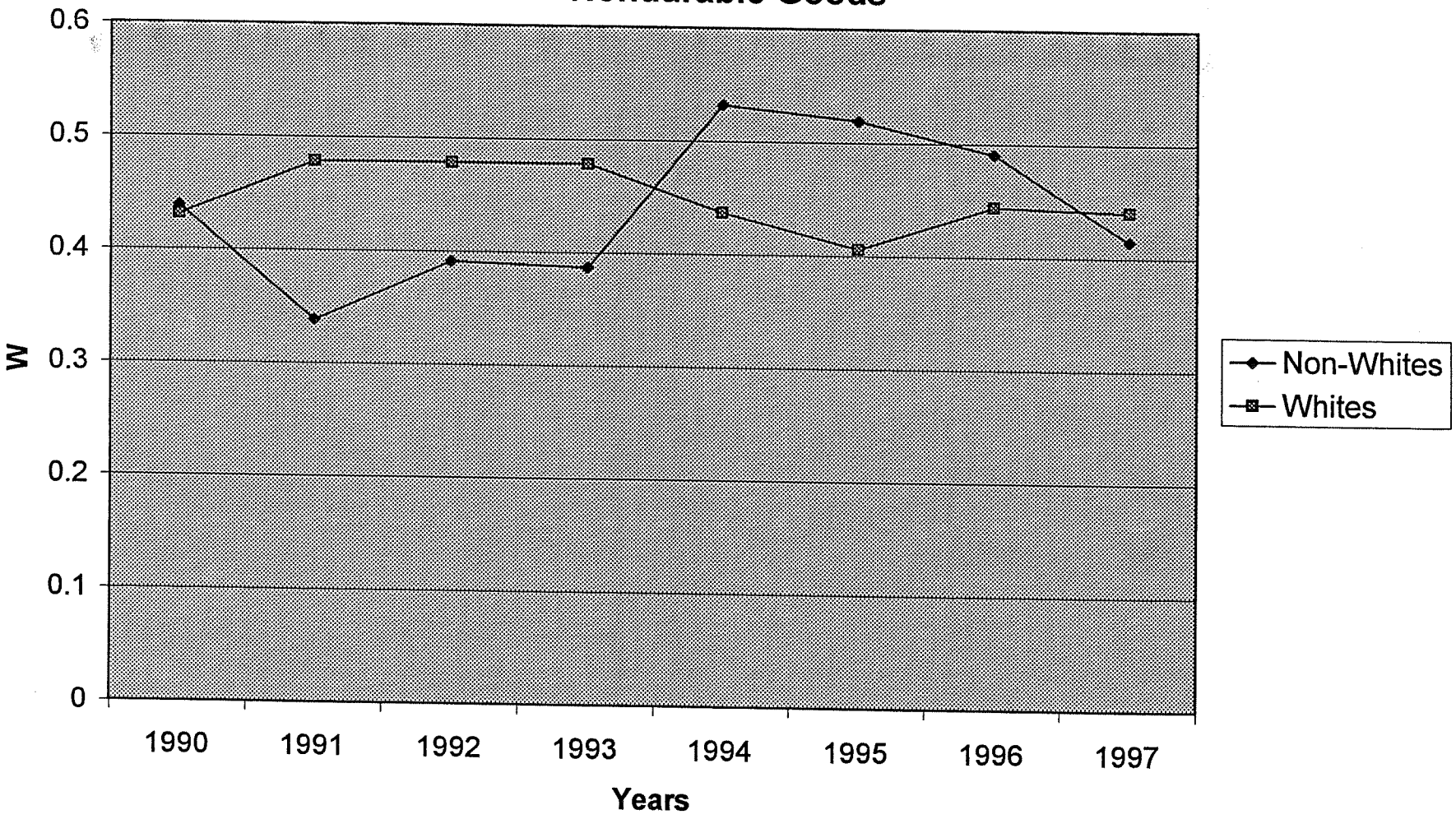


Chart 5.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Transportation and Communication

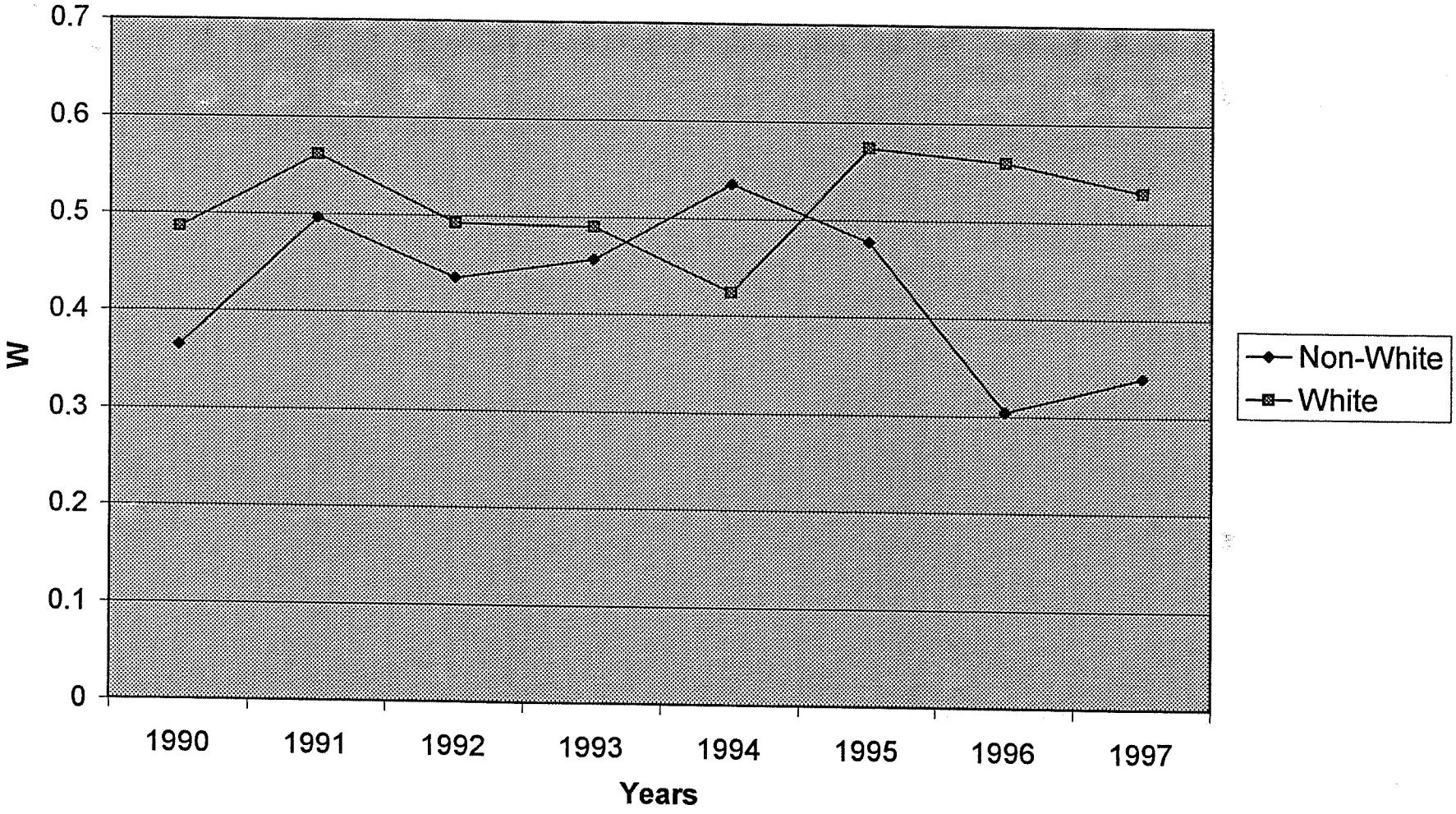


Chart 6.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Durable Goods

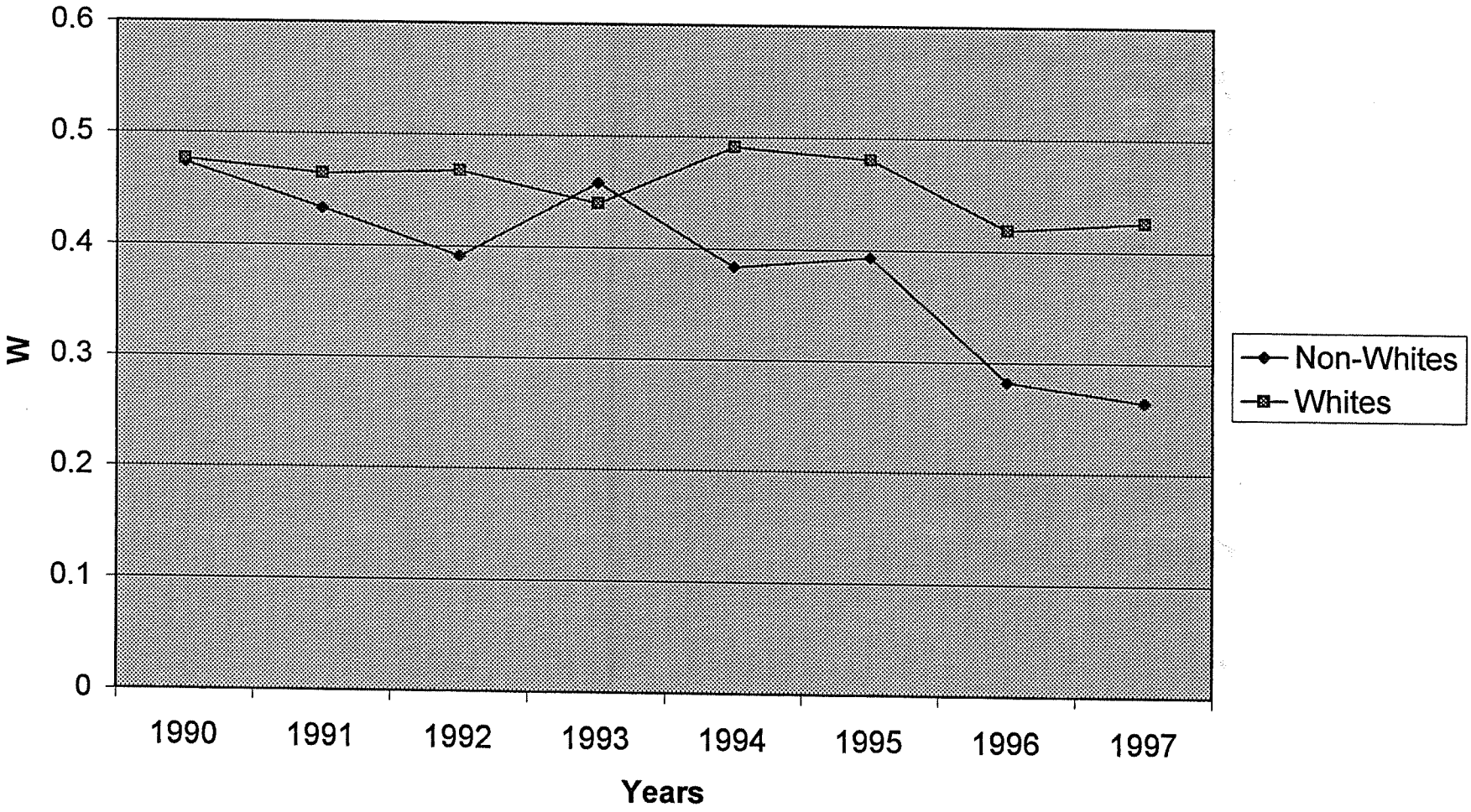


Chart 7.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Business and Repair Services

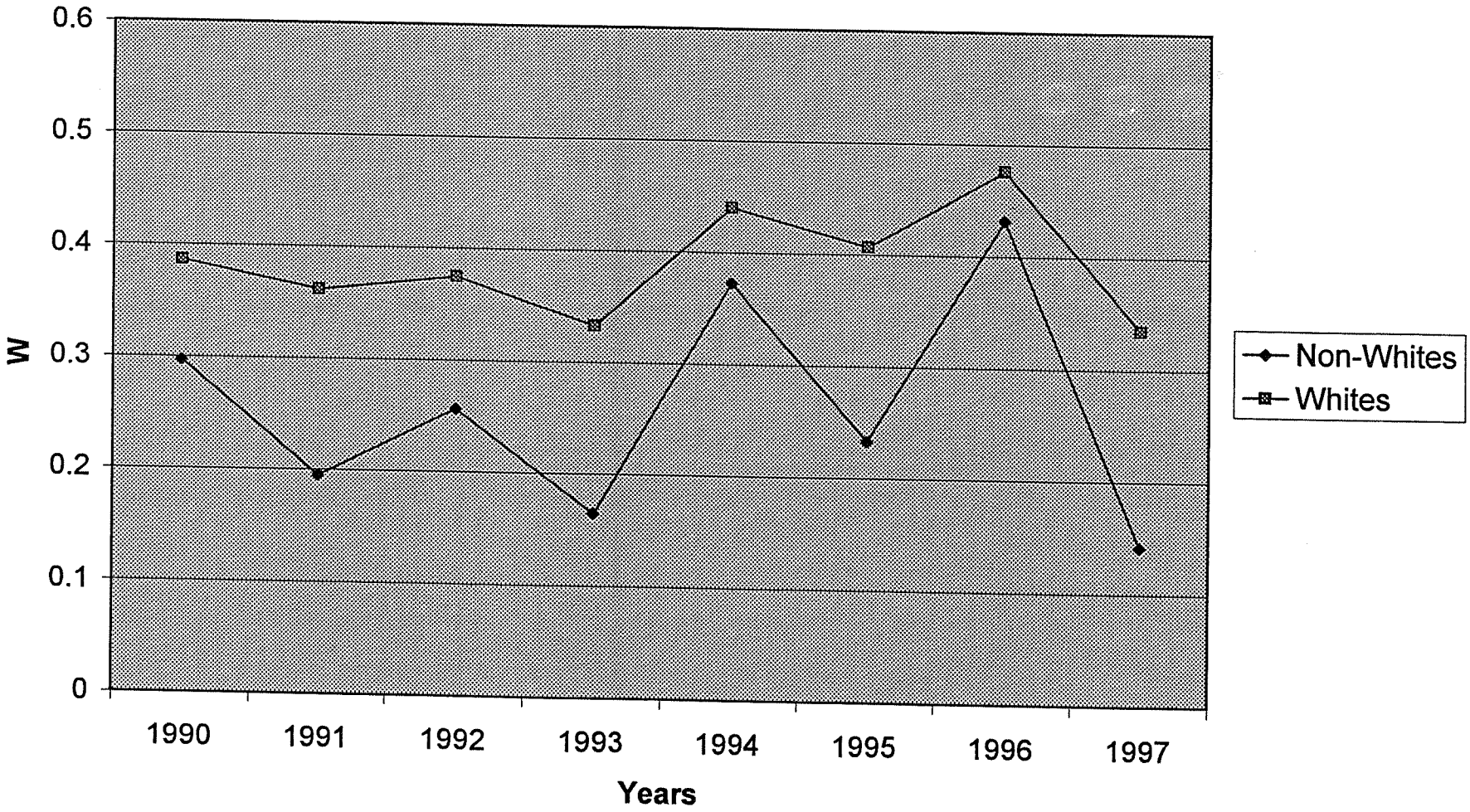
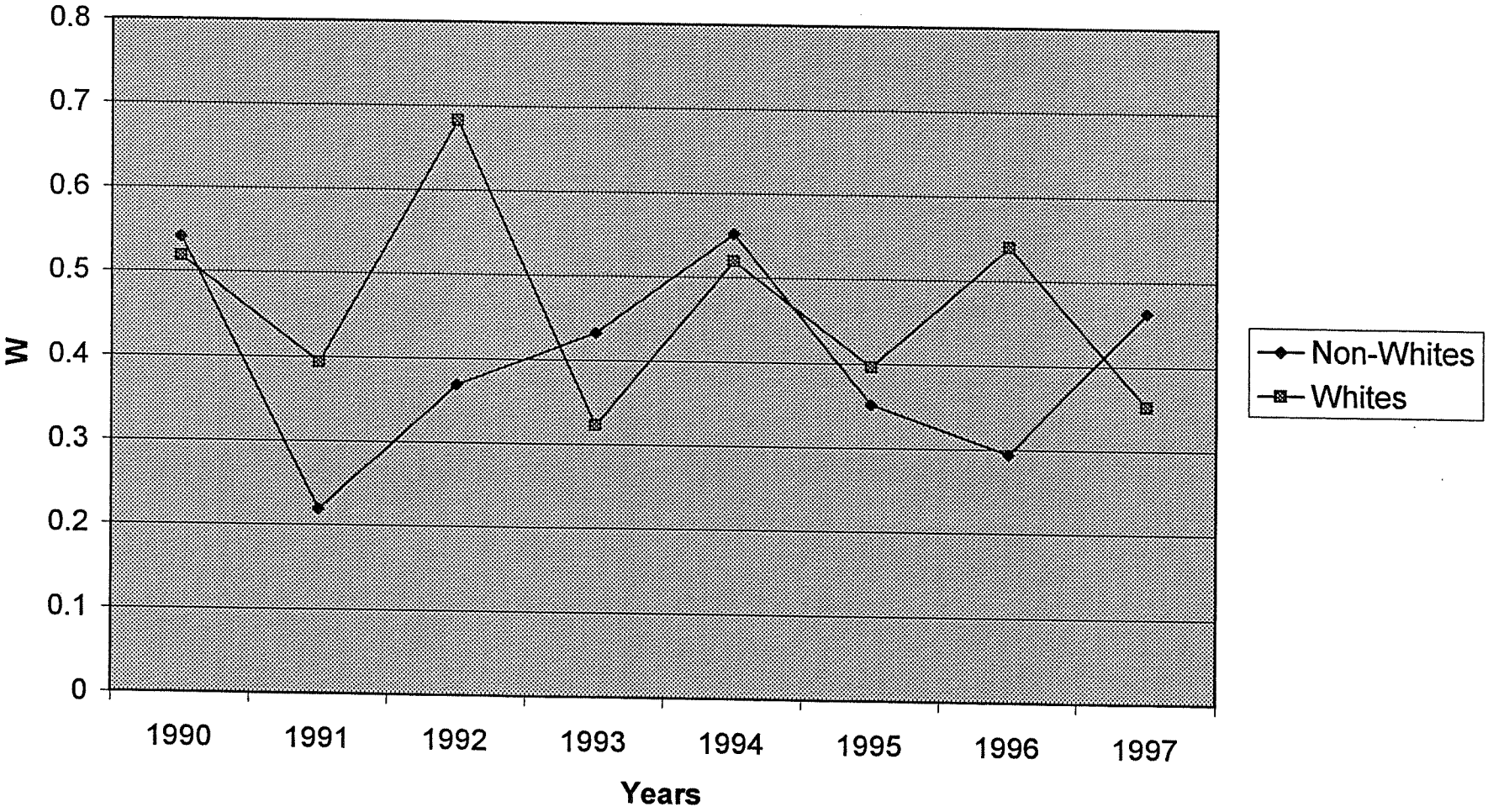


Chart 8.
Cumulative Effect of Experience on Log Wages of 20 Years of Experience
Finance, Insurance, and Real Estate



There are two rough categories into which each industry falls. The first consists of those industries that have shown a pattern of convergence of the return to experience for white and non-white workers. Construction, Wholesale and Retail Trade, Professional and Related Services, and Nondurable Goods. Of these, Construction showed the clearest trend toward convergence with the statistical difference between white and non-white workers becoming insignificant toward the end of the sample. Wholesale and Retail Trade also showed significant differences toward the beginning of the sample, but those differences have become insignificant in the late 1990's. The return to experience for workers in Professional and Related Services and Nondurable Goods does not necessarily show a clear trend toward convergence, but they do show insignificant differences for whites and non-whites. Professional and Related Services shows a small, insignificant advantage for white and Nondurable Goods shows an advantage for whites early in the sample but then non-whites seem to have a small advantage.

The second category of industries is made up of those industries that show no trend toward convergence. Transportation and Communication, Durable Goods, Business and Repair Services, and Finance, Insurance, and Real Estate fall into this category. Most notably, Transportation and Communication and Durable Goods have shown a divergence in the return to experience for white and non-white workers in recent years with the difference becoming significant for the last few years of the sample. Business and Repair Services does not necessarily show a divergent trend, but it does show a consistent advantage to white workers over black with significant differences in 1997. Finance, Insurance, and Real Estate shows no real trend and these erratic results are most likely due to the small sample size.

V. Conclusions

There clearly exist differences among the various industries studied in the return to experience for white and non-white workers. Explanations for why this is true would most likely be similar to those presented by Bratsberg and Terrell. With more information about the differences among these industries, it might become more obvious which of their potential explanations seems to hold the most weight. However, because of the limitations of the CPS data and the econometric approach used in this study, I can provide only limited results. While there does seem to be strong evidence to support the claim that there are significant differences among the various industries studied, there clearly exists a need for further research on the topic to more accurately estimate the size of these differentials within each industry and the differences between industries.

References

Altonji, Joseph G. and Robert A. Shakotko, "Do Wages Rise with Job Seniority?," NBER Working Paper No. 1616 (May 1985).

Altonji, Joseph G. and Robert A. Shakotko, "Do Wages Rise with Job Seniority?," Review of Economic Studies 54 (July 1987): 437-459.

Becker, Gary S., The Economics of Discrimination, Chicago: University of Chicago Press (1971).

Bratsberg, Bernt and Dek Terrell. "Experience, Tenure, and Wage Growth of Young Black and White Men," The Journal of Human Resources 33(3) (Summer 1998): 658-82.

Bratsberg, Bernt and Dek Terrell. "School Quality and Returns to Education of U.S. Immigrants." Economic Inquiry, 40(2) (April 2002): 177-198.

Farber, Stephen C. and Robert J. Newman. " Accounting for South/Non-South Real Wage Differentials and for Changes in Those Differentials Over Time," The Review of Economics and Statistics, 1987: 215-223.

Grogger, Jeff, "Does School Quality Explain the Recent Black/White Wage Trend?," Journal of Labor Economics 14 (April 1996): 231-253.

Jaeger, D. A., "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers," Journal of Business and Economic Statistics, 15(3), 1997: 300-309.

Jaeger, D. A., "Reconciling Educational Attainment Questions in the CPS and the Census," Monthly Labor Review (August 1997): 36-40.

Lazear, Edward, "The Narrowing of Black-White Wage Differentials is Illusory," American Economic Review 69 (September 1979): 553-564.

Light, Audrey and Manuelita Ureta, "Early Career Work Experience and Gender Wage Differentials," Journal of Labor Economics (January 1995): 121-154.

Wolpin, Kenneth, "The Determinants of Black-White Differences in Early Employment Careers," Search, Layoffs, Quits, and Endogenous Wage Growth," Journal of Political Economy 100 (June 1992): 535-560.

Appendix A: SAS Program—Data Filter

```
dm 'log;clear;output;clear';

data d79;
set "c:\STEPHEN\cps79.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS79f.sd2";
set d79;
run;


data d80;
set "c:\STEPHEN\cps80.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS80f.sd2";
set d80;
run;


data d81;
set "c:\STEPHEN\cps81.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS81f.sd2";
set d81;
run;


data d82;
set "c:\STEPHEN\cps82.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS82f.sd2";
set d82;
run;


data d83;
set "c:\STEPHEN\cps83.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
```

```
run;

data "C:\STEPHEN\CPS83f.sd2";
set d83;
run;


data d84;
set "c:\STEPHEN\cps84.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS84f.sd2";
set d84;
run;


data d85;
set "c:\STEPHEN\cps85.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS85f.sd2";
set d85;
run;


data d86;
set "c:\STEPHEN\cps86.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS86f.sd2";
set d86;
run;


data d87;
set "c:\STEPHEN\cps87.sd2";
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS87f.sd2";
set d87;
run;


data d88;
set "c:\STEPHEN\cps88.sd2";
```

```
if age < 16 then delete;
if age > 64 then delete;
if esr > 1 then delete;
run;

data "C:\STEPHEN\CPS88f.sd2";
set d88;
run;


data d89;
set "c:\STEPHEN\cps89.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr89 > 1 then delete;
run;

data "C:\STEPHEN\CPS89f.sd2";
set d89;
run;


data d90;
set "c:\STEPHEN\cps90.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr89 > 1 then delete;
run;

data "C:\STEPHEN\CPS90f.sd2";
set d90;
run;


data d91;
set "c:\STEPHEN\cps91.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr89 > 1 then delete;
run;

data "C:\STEPHEN\CPS91f.sd2";
set d91;
run;


data d92;
set "c:\STEPHEN\cps92.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr89 > 1 then delete;
run;

data "C:\STEPHEN\CPS92f.sd2";
set d92;
run;
```

```
data d93;
set "c:\STEPHEN\cps93.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr89 > 1 then delete;
run;

data "C:\STEPHEN\CPS93f.sd2";
set d93;
run;


data d94;
set "c:\STEPHEN\cps94.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr94 > 1 then delete;
run;

data "C:\STEPHEN\CPS94f.sd2";
set d94;
run;


data d95;
set "c:\STEPHEN\cps95.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr94 > 1 then delete;
run;

data "C:\STEPHEN\CPS95f.sd2";
set d95;
run;


data d96;
set "c:\STEPHEN\cps96.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr94 > 1 then delete;
run;

data "C:\STEPHEN\CPS96f.sd2";
set d96;
run;


data d97;
set "c:\STEPHEN\cps97.sd2";
if age < 16 then delete;
if age > 64 then delete;
if lfsr94 > 1 then delete;
run;

data "C:\STEPHEN\CPS97f.sd2";
```

```
set d97;
run;
```

Appendix B: SAS Program—Proc Means, Variable Definition, and Model Specification

Note: Some lines of this program have been marked as comments so that all steps in the data analysis process can be seen without having the program run all lines when simply testing the model.

```
/* ********************************************************  */
/* * THIS PROGRAM WILL BE  USED TO DO SIMPLE REGRESSION     *  */
/* * ANALYSIS,WAGE GROWTH DIFFERENTIALS AMONG RACES         *  */
/* * BY INDUSTRY                                    *  */
/* * STEPHEN BARNES LSU SENIOR THESIS 2002                 *  */
/* * ---------------------------------------------------- *  */
/* * THIS PROGRAM IS BASED ON SAMPLE PROVIDED BY BOB NEWMAN*  */
/* * REGRESSIONS ON APRIL 20, 2002                         *  */
/* * PROGRAMMER : STEPHEN BARNES                           *  */
/* * PURPOSE : RUN REGRESSION FOR MODEL                 *  */
/* *           1. LOG OF WAGE AGAINST DUMMIES BY INDUSTRY  *  */
/* *                INCLUDING DUMMIES FOR OCCUPATION,        *  */
/* *                REGION, AND MARITAL STATUS           *  */
/* ********************************************************  */
/*                                                           */
/* USING PREVIOUSLY FILTERED DATA FROM CPS LABOR EXTRACTS     */
/* FOR THE YEARS 1979 - 1997 INCLUDING ONLY THOSE BETWEEN     */
/* THE AGES OF 16 AND 64 WHO ARE CURRENTLY WORKING          */

*************************************************************;

* VARIABLE DEFINITION *;

* ------------------------------------------------------------ *;
* -------------VARIABLES USED FOR YEARS 1979-1982------------- *;

*Data CPS88; *set "C:\STEPHEN\CPS82f.sd2";

* YEARS OF EXPERIENCE *;
*EXP = (AGE-GRADEAT-6); *Assumes ind is out of the LF for the full
year*;
*IF EXP <= 0 THEN DELETE;
*EXP_SQ = EXP*EXP;

* HOURLY WAGE *;
*WAGE = EARNHRE/100; * HOURLY WAGE CONVERTED TO DOLLARS-ALWAYS 50 CENTS
OR MORE*;
*LNWAGE = LOG(WAGE);

* RACE *;
*DRACE = 0;
*IF RACE = 1 THEN DRACE = 1; * WHITE = 1 *;

* SEX *;
*DSEX = 0;
*IF SEX =1 THEN DSEX = 1; * MALE = 1 *;

* REGIONAL DUMMIES *;
*NE = 0; * NEW ENGLAND DIVISION *;
```

```
*MA = 0; * MIDDLE ATLANTIC DIVISION *;
*EN = 0; * EAST NORTH CENTRAL DIVISION *;
*WN = 0; * WEST NORTH CENTRAL DIVISION *;
*SA = 0; * SOUTH ATLANTIC *;
*ES = 0; * EAST SOUTH CENTRAL *;
*WS = 0; * WEST SOUTH CENTRAL *;
*MT = 0; * MOUNTAIN *;
*PC = 0; * PACIFIC *;

*IF STATE < 20 THEN NE = 1;
*IF STATE < 30 AND STATE > 20 THEN MA = 1;
*IF STATE < 40 AND STATE > 30 THEN EN = 1;
*IF STATE < 50 AND STATE > 40 THEN WN = 1;
*IF STATE < 60 AND STATE > 50 THEN SA = 1;
*IF STATE < 70 AND STATE > 60 THEN ES = 1;
*IF STATE < 80 AND STATE > 70 THEN WS = 1;
*IF STATE < 90 AND STATE > 80 THEN MT = 1;
*IF STATE < 100 AND STATE > 90 THEN PC = 1;

* INDUSTRY DUMMIES *;
*AGR = 0; * AGRICULTURE, FORESTRY, AND FISHERIES *;
*MIN = 0; * MINING *;
*CON = 0; * CONSTRUCTION *;
*DUR = 0; * DURABLE GOODS *;
*NDR = 0; * NONDURABLE GOODS *;
*TRN = 0; * TRANSPORTATION AND COMMUNICATION *; * CREATED TO MATCH
1983- *;
*TRD = 0; * WHOLESALE AND RETAIL TRADE *;
*FIR = 0; * FINANCE, INSURANCE, AND REAL ESTATE *;
*BRS = 0; * BUSINESS AND REPAIR SERVICES *;
*PER = 0; * PERSONAL SERVICES *;
*ENT = 0; * ENTERTAINMENT AND RECREATIONAL SERVICES *;
*PRO = 0; * PROFESSIONAL AND RELATED SERVICES *;
*PAD = 0; * PUBLIC ADMINISTRATION *;
*NON = 0; * WORKERS NOT CLASSIFIABLE BY INDUSTRY *;

*IF IND70 < 40 THEN AGR = 1;
*IF IND70 < 60 AND IND70 > 40 THEN MIN = 1;
*IF IND70 < 80 AND IND70 > 60 THEN CON = 1;
*IF IND70 < 268 AND IND70 > 100 THEN DUR = 1;
*IF IND70 < 399 AND IND70 > 267 THEN NDR = 1;
*IF IND70 < 500 AND IND70 > 400 THEN TRN = 1;
*IF IND70 < 700 AND IND70 > 500 THEN TRD = 1;
*IF IND70 < 719 AND IND70 > 700 THEN FIR = 1;
*IF IND70 < 768 AND IND70 > 718 THEN BRS = 1;
*IF IND70 < 800 AND IND70 > 768 THEN PER = 1;
*IF IND70 < 820 AND IND70 > 800 THEN ENT = 1;
*IF IND70 < 900 AND IND70 > 820 THEN PRO = 1;
*IF IND70 < 948 AND IND70 > 900 THEN PAD = 1;
*IF IND70 > 947 THEN NON = 1;

*IF AGR = 1 THEN DELETE; *DELETES AGRICULTURE, FORESTRY, AND
FISHERIES*;
*IF NON = 1 THEN DELETE; *DELETES WORKERS WITH NO INDUSTRY
CLASSIFICATION*;

* CREATING INDUSTRY VARIABLE FOR SORT *;
```

```
*IF AGR = 1 THEN IND = 1;
*IF MIN = 1 THEN IND = 2;
*IF CON = 1 THEN IND = 3;
*IF DUR = 1 THEN IND = 4;
*IF NDR = 1 THEN IND = 5;
*IF TRN = 1 THEN IND = 6;
*IF TRD = 1 THEN IND = 7;
*IF FIR = 1 THEN IND = 8;
*IF BRS = 1 THEN IND = 9;
*IF PER = 1 THEN IND = 10;
*IF ENT = 1 THEN IND = 11;
*IF PRO = 1 THEN IND = 12;
*IF PAD = 1 THEN IND = 13;
*IF NON = 1 THEN IND = 14;

* OCCUPATIONAL DUMMIES *;
*PTMS = 0; * PROFESSIONAL, TECHNICAL, AND KINDRED WORKERS *;
          * AND MANAGERS AND ADMINISTRATORS, EXCEPT FARM *;
          * AND SALES WORKERS *;
*SERV = 0; * SERVICE WORKERS EXCEPT PRIVATE HOUSEHOLD *;
          * AND PRIVATE HOUSEHOLD WORKERS *;
*CRFT = 0; * CRAFTSMEN AND KINDRED WORKERS *;
*OPLA = 0; * OPERATIVES, EXCEPT TRANSPORT *;
          * AND TRANSPORT OPERATIVES *;
          * AND LABORERS, EXCEPT FARM *;
*FARM = 0; * FARMERS AND FARM MANAGERS *;
          * AND FARM LABORERS AND FARM FOREMEN *;

*NOOC = 0; * WORKERS NOT CLASSIFIABLE BY OCCUPATION *;

*IF OCC70 < 200 AND OCC70 > 0 THEN PTMS = 1;
*IF OCC70 < 260 AND OCC70 > 199 THEN PTMS = 1;
*IF OCC70 < 400 AND OCC70 > 259 THEN PTMS = 1;
*IF OCC70 < 980 AND OCC70 > 900 THEN SERV = 1;
*IF OCC70 < 990 AND OCC70 > 979 THEN SERV = 1;
*IF OCC70 < 600 AND OCC70 > 399 THEN CRFT = 1;
*IF OCC70 < 700 AND OCC70 > 599 THEN OPLA = 1;
*IF OCC70 < 740 AND OCC70 > 699 THEN OPLA = 1;
*IF OCC70 < 800 AND OCC70 > 739 THEN OPLA = 1;
*IF OCC70 < 810 AND OCC70 > 799 THEN FARM = 1;
*IF OCC70 < 850 AND OCC70 > 820 THEN FARM = 1;
*IF OCC70 > 990 THEN NOOC = 1;

*IF FARM = 1 THEN DELETE; * DELETES ALL FARMERS AND FARM MANAGERS *;
                    * AND FARM LABORERS AND FARM FOREMEN *;
*IF NOOC = 1 THEN DELETE; *DELETES WORKERS NOT CLASSIFIABLE BY
OCCUPATION*;

* GOVERNMENT EMPLOYEE DUMMIES *;
*GOVT = 0;
*IF CLASS = 2 THEN GOVT = 1;
*IF CLASS = 3 THEN GOVT = 1;
*IF CLASS = 4 THEN GOVT = 1;

* MARRIED SPOUSE PRESENT *;
*MARRIED = 0;
```

```
*IF MARITAL = 1 THEN MARRIED = 1; *NOT INCLUDING ARMED FORCES IN
DEFINITION*;

*RACE_EXP = EXP*DRACE;
*RACE_ESQ = EXP_SQ*DRACE;

*PROC MEANS; *VAR EXP WAGE GRADEAT;


* -------------VARIABLES USED FOR YEARS 1983-1988------------- *;

*Data CPS88; *set "C:\STEPHEN\CPS88f.sd2";

* YEARS OF EXPERIENCE *;
*EXP = (AGE-GRADEAT-6); *Assumes ind is out of the LF for the full
year*;
*IF EXP <= 0 THEN DELETE;
*EXP_SQ = EXP*EXP;

* HOURLY WAGE *;
*WAGE = EARNHRE/100; * HOURLY WAGE CONVERTED TO DOLLARS-ALWAYS 50 CENTS
OR MORE*;
*LNWAGE = LOG(WAGE);

* RACE *;
*DRACE = 0;
*IF RACE = 1 THEN DRACE = 1;

* REGIONAL DUMMIES *;
*NE = 0; * NEW ENGLAND DIVISION *;
*MA = 0; * MIDDLE ATLANTIC DIVISION *;
*EN = 0; * EAST NORTH CENTRAL DIVISION *;
*WN = 0; * WEST NORTH CENTRAL DIVISION *;
*SA = 0; * SOUTH ATLANTIC *;
*ES = 0; * EAST SOUTH CENTRAL *;
*WS = 0; * WEST SOUTH CENTRAL *;
*MT = 0; * MOUNTAIN *;
*PC = 0; * PACIFIC *;

*IF STATE < 20 THEN NE = 1;
*IF STATE < 30 AND STATE > 20 THEN MA = 1;
*IF STATE < 40 AND STATE > 30 THEN EN = 1;
*IF STATE < 50 AND STATE > 40 THEN WN = 1;
*IF STATE < 60 AND STATE > 50 THEN SA = 1;
*IF STATE < 70 AND STATE > 60 THEN ES = 1;
*IF STATE < 80 AND STATE > 70 THEN WS = 1;
*IF STATE < 90 AND STATE > 80 THEN MT = 1;
*IF STATE < 100 AND STATE > 90 THEN PC = 1;

* INDUSTRY DUMMIES *;
*AGR = 0; * AGRICULTURE, FORESTRY, AND FISHERIES *;
*MIN = 0; * MINING *;
*CON = 0; * CONSTRUCTION *;
*DUR = 0; * DURABLE GOODS *;
*NDR = 0; * NONDURABLE GOODS *;
*TRN = 0; * TRANSPORTATION AND COMMUNICATION*;
*TRD = 0; * WHOLESALE AND RETAIL TRADE *;
```

```
*FIR = 0; * FINANCE, INSURANCE, AND REAL ESTATE *;
*BRS = 0; * BUSINESS AND REPAIR SERVICES *;
*PER = 0; * PERSONAL SERVICES *;
*ENT = 0; * ENTERTAINMENT AND RECREATIONAL SERVICES *;
*PRO = 0; * PROFESSIONAL AND RELATED SERVICES *;
*PAD = 0; * PUBLIC ADMINISTRATION *;
*NON = 0; * WORKERS NOT CLASSIFIABLE BY INDUSTRY *;

*IF IND80 < 40 THEN AGR = 1;
*IF IND80 < 60 AND IND80 > 39 THEN MIN = 1;
*IF IND80 < 61 AND IND80 > 59 THEN CON = 1;
*IF IND80 < 400 AND IND80 > 229 THEN DUR = 1;
*IF IND80 < 230 AND IND80 > 99 THEN NDR = 1;
*IF IND80 < 500 AND IND80 > 399 THEN TRN = 1;
*IF IND80 < 700 AND IND80 > 499 THEN TRD = 1;
*IF IND80 < 721 AND IND80 > 699 THEN FIR = 1;
*IF IND80 < 761 AND IND80 > 720 THEN BRS = 1;
*IF IND80 < 800 AND IND80 > 760 THEN PER = 1;
*IF IND80 < 812 AND IND80 > 799 THEN ENT = 1;
*IF IND80 < 900 AND IND80 > 811 THEN PRO = 1;
*IF IND80 < 990 AND IND80 > 899 THEN PAD = 1;
*IF IND80 > 990 THEN NON = 1;

*IF AGR = 1 THEN DELETE; * DELETES AGRICULTURE, FORESTRY, AND FISHERIES
*;
*IF NON = 1 THEN DELETE; * DELETES WORKERS WITH NO INDUSTRY
CLASSIFICATION *;

* CREATING INDUSTRY VARIABLE FOR SORT*;
*IF AGR = 1 THEN IND = 1;
*IF MIN = 1 THEN IND = 2;
*IF CON = 1 THEN IND = 3;
*IF DUR = 1 THEN IND = 4;
*IF NDR = 1 THEN IND = 5;
*IF TRN = 1 THEN IND = 6;
*IF TRD = 1 THEN IND = 7;
*IF FIR = 1 THEN IND = 8;
*IF BRS = 1 THEN IND = 9;
*IF PER = 1 THEN IND = 10;
*IF ENT = 1 THEN IND = 11;
*IF PRO = 1 THEN IND = 12;
*IF PAD = 1 THEN IND = 13;
*IF NON = 1 THEN IND = 14;

* OCCUPATIONAL DUMMIES *;
*PTMS = 0; * MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS *;
         * AND TECHNICAL, SALES, AND ADMINISTRATIVE SUPPORT OCCUPATIONS
*;
*SERV = 0; * SERVICE OCCUPATIONS *;
*CRFT = 0; * PRECISION PRODUCTION, CRAFT, AND REPAIR OCCUPATIONS *;
*OPLA = 0; * OPERATORS, FABRICATORS, AND LABORERS *;
*FARM = 0; * FARMING, FORESTRY, AND FISHING OCCUPATIONS *;

*NOOC = 0; * WORKERS NOT CLASSIFIABLE BY OCCUPATION *;

*IF OCC80 < 200 AND OCC80 > 0 THEN PTMS = 1;
*IF OCC80 < 390 AND OCC80 > 199 THEN PTMS = 1;
```

```
*IF OCC80 < 470 AND OCC80 > 400 THEN SERV = 1;
*IF OCC80 < 700 AND OCC80 > 500 THEN CRFT = 1;
*IF OCC80 < 890 AND OCC80 > 700 THEN OPLA = 1;
*IF OCC80 < 500 AND OCC80 > 470 THEN FARM = 1;
*IF OCC80 > 900 THEN NOOC = 1;

*IF FARM = 1 THEN DELETE; * DELETES ALL FARMERS AND FARM MANAGERS *;
                    * AND FARM LABORERS AND FARM FOREMEN *;
*IF NOOC = 1 THEN DELETE; * DELETES WORKERS NOT CLASSIFIABLE BY
OCCUPATION *;

* GOVERNMENT EMPLOYEE DUMMIES *;
*GOVT = 0;
*IF CLASS = 2 THEN GOVT = 1;
*IF CLASS = 3 THEN GOVT = 1;
*IF CLASS = 4 THEN GOVT = 1;

* MARRIED SPOUSE PRESENT *;
*MARRIED = 0;
*IF MARITAL = 1 THEN MARRIED = 1; * DOES NOT INCLUDE ARMED FORCES IN
DEFINITION *;

*RACE_EXP = EXP*DRACE;
*RACE_ESQ = EXP_SQ*DRACE;


*PROC MEANS; *VAR EXP WAGE GRADEAT;


* -------------VARIABLES USED FOR YEARS 1989-1991------------- *;

Data CPS91; set "C:\STEPHEN\CPS90f.sd2";

* YEARS OF EXPERIENCE *;
EXP = (AGE - GRADEAT + 1 - GRADECP - 6); *ADJUSTS FOR COMPLETION OF
FINAL YEAR OF SCHOOL*;
IF EXP <= 0 THEN DELETE;
EXP_SQ = EXP*EXP;

* HOURLY WAGE *;
WAGE = EARNHRE/100; * HOURLY WAGE CONVERTED TO DOLLARS *;
LNWAGE = LOG(WAGE);

* RACE *;
DRACE = 0;
IF RACE = 1 THEN DRACE = 1; * WHITE = 1 *;

* SEX *;
DSEX = 0;
IF SEX = 1 THEN DSEX = 1; * MALE = 1 *;

* REGIONAL DUMMIES *;
NE = 0; * NEW ENGLAND DIVISION *;
MA = 0; * MIDDLE ATLANTIC DIVISION *;
EN = 0; * EAST NORTH CENTRAL DIVISION *;
WN = 0; * WEST NORTH CENTRAL DIVISION *;
SA = 0; * SOUTH ATLANTIC *;
```

```
ES = 0; * EAST SOUTH CENTRAL *;
WS = 0; * WEST SOUTH CENTRAL *;
MT = 0; * MOUNTAIN *;
PC = 0; * PACIFIC *;

IF STATE < 20 THEN NE = 1;
IF STATE < 30 AND STATE > 20 THEN MA = 1;
IF STATE < 40 AND STATE > 30 THEN EN = 1;
IF STATE < 50 AND STATE > 40 THEN WN = 1;
IF STATE < 60 AND STATE > 50 THEN SA = 1;
IF STATE < 70 AND STATE > 60 THEN ES = 1;
IF STATE < 80 AND STATE > 70 THEN WS = 1;
IF STATE < 90 AND STATE > 80 THEN MT = 1;
IF STATE < 100 AND STATE > 90 THEN PC = 1;

* INDUSTRY DUMMIES *;
AGR = 0; * AGRICULTURE, FORESTRY, AND FISHERIES *;
MIN = 0; * MINING *;
CON = 0; * CONSTRUCTION *;
DUR = 0; * DURABLE GOODS *;
NDR = 0; * NONDURABLE GOODS *;
TRN = 0; * TRANSPORTATION AND COMMUNICATION*;
TRD = 0; * WHOLESALE AND RETAIL TRADE *;
FIR = 0; * FINANCE, INSURANCE, AND REAL ESTATE *;
BRS = 0; * BUSINESS AND REPAIR SERVICES *;
PER = 0; * PERSONAL SERVICES *;
ENT = 0; * ENTERTAINMENT AND RECREATIONAL SERVICES *;
PRO = 0; * PROFESSIONAL AND RELATED SERVICES *;
PAD = 0; * PUBLIC ADMINISTRATION *;
NON = 0; * WORKERS NOT CLASSIFIABLE BY INDUSTRY *;

IF IND80 < 40 THEN AGR = 1;
IF IND80 < 60 AND IND80 > 39 THEN MIN = 1;
IF IND80 < 61 AND IND80 > 59 THEN CON = 1;
IF IND80 < 400 AND IND80 > 229 THEN DUR = 1;
IF IND80 < 230 AND IND80 > 99 THEN NDR = 1;
IF IND80 < 500 AND IND80 > 399 THEN TRN = 1;
IF IND80 < 700 AND IND80 > 499 THEN TRD = 1;
IF IND80 < 721 AND IND80 > 699 THEN FIR = 1;
IF IND80 < 761 AND IND80 > 720 THEN BRS = 1;
IF IND80 < 800 AND IND80 > 760 THEN PER = 1;
IF IND80 < 812 AND IND80 > 799 THEN ENT = 1;
IF IND80 < 900 AND IND80 > 811 THEN PRO = 1;
IF IND80 < 990 AND IND80 > 899 THEN PAD = 1;
IF IND80 > 990 THEN NON = 1;

IF AGR = 1 THEN DELETE; * DELETES ALL AGRICULTURE, FORESTRY, AND
FISHERIES *;
IF NON = 1 THEN DELETE; * DELETES WORKERS WITH NO INDUSTRY
CLASSIFICATION *;

* CREATING INDUSTRY VARIABLE FOR SORT*;
IF AGR = 1 THEN IND = 1;
IF MIN = 1 THEN IND = 2;
IF CON = 1 THEN IND = 3;
IF DUR = 1 THEN IND = 4;
IF NDR = 1 THEN IND = 5;
```

```
IF TRN = 1 THEN IND = 6;
IF TRD = 1 THEN IND = 7;
IF FIR = 1 THEN IND = 8;
IF BRS = 1 THEN IND = 9;
IF PER = 1 THEN IND = 10;
IF ENT = 1 THEN IND = 11;
IF PRO = 1 THEN IND = 12;
IF PAD = 1 THEN IND = 13;
IF NON = 1 THEN IND = 14;

* OCCUPATIONAL DUMMIES *;
PTMS = 0; * MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS *;
        * AND TECHNICAL, SALES, AND ADMINISTRATIVE SUPPORT OCCUPATIONS
*;
SERV = 0; * SERVICE OCCUPATIONS *;
CRFT = 0; * PRECISION PRODUCTION, CRAFT, AND REPAIR OCCUPATIONS *;
OPLA = 0; * OPERATORS, FABRICATORS, AND LABORERS *;
FARM = 0; * FARMING, FORESTRY, AND FISHING OCCUPATIONS *;

NOOC = 0; * WORKERS NOT CLASSIFIABLE BY OCCUPATION *;

IF OCC80 < 200 AND OCC80 > 0 THEN PTMS = 1;
IF OCC80 < 390 AND OCC80 > 199 THEN PTMS = 1;
IF OCC80 < 470 AND OCC80 > 400 THEN SERV = 1;
IF OCC80 < 700 AND OCC80 > 500 THEN CRFT = 1;
IF OCC80 < 890 AND OCC80 > 700 THEN OPLA = 1;
IF OCC80 < 500 AND OCC80 > 470 THEN FARM = 1;
IF OCC80 > 900 THEN NOOC = 1;

IF FARM = 1 THEN DELETE; * DELETES ALL FARMERS AND FARM MANAGERS *;
                    * AND FARM LABORERS AND FARM FOREMEN *;
IF NOOC = 1 THEN DELETE; * DELETES WORKERS NOT CLASSIFIABLE BY
OCCUPATION *;

* GOVERNMENT EMPLOYEE DUMMIES *;
GOVT = 0;
IF CLASS > 1 AND CLASS < 5 THEN GOVT = 1;

* MARRIED SPOUSE PRESENT *;
MARRIED = 0;
IF MARITAL = 1 THEN MARRIED = 1; * DOES NOT INCLUDE ARMED FORCES IN
DEFINITION *;

RACE_EXP = EXP*DRACE;
RACE_ESQ = EXP_SQ*DRACE;

*PROC MEANS; *VAR EXP WAGE GRADEAT;

* ALL BELOW ARE NOT PRESENT IN PREVIOUS VARIABLE DEFINITIONS *;
IF GRADEAT ^= 12 THEN DELETE;
IF GRADECP ^= 1 THEN DELETE; * DELETES ALL NON HIGH SCHOOL GRADS *;
IF GOVT = 1 THEN DELETE; * DELETES ALL GOVERNMENT EMPLOYEES *;
IF PAD = 1 THEN DELETE; * DELETES THE REST OF THE GOVERNMENT EMPLOYEES
*;
IF DSEX = 0 THEN DELETE; * DELETES ALL FEMALES *;
```

```
PROC SORT; BY IND DRACE;
*PROC MEANS; *VAR SCHOOL MARRIED;
PROC REG S; BY IND DRACE;
      *MODEL LNWAGE = EXP RACE_EXP EXP_SQ RACE_ESQ DRACE MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      MODEL LNWAGE = EXP EXP_SQ MARRIED
                   PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      TEST EXP, EXP_SQ;

*PROC SORT; *BY IND;
*PROC MEANS; *VAR SCHOOL MARRIED;
*PROC REG S; *BY IND;
      *MODEL LNWAGE = EXP RACE_EXP EXP_SQ RACE_ESQ DRACE MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *MODEL LNWAGE = EXP EXP_SQ SCHOOL PTMS SERV CRFT GOVT DRACE DSEX
                     MIN DUR NDR TRN TRD FIR BRS PER ENT PRO PAD
                  NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*PROC REG S; *BY IND;
      *MODEL LNWAGE = EXP EXP_SQ SCHOOL PTMS SERV CRFT GOVT DRACE DSEX
                  NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

run;


* -------------VARIABLES USED FOR YEARS 1992-1993------------- *;

*data CPS92; *set "C:\STEPHEN\CPS93F.sd2";

* YEARS OF EXPERIENCE *;
*IF GRADE92 = 31 THEN EXP = (AGE-0-6);
*IF GRADE92 = 32 THEN EXP = (AGE-2.5-6);
*IF GRADE92 = 33 THEN EXP = (AGE-5.5-6);
*IF GRADE92 = 34 THEN EXP = (AGE-7.5-6);
*IF GRADE92 = 35 THEN EXP = (AGE-9-6);
*IF GRADE92 = 36 THEN EXP = (AGE-10-6);
*IF GRADE92 = 37 THEN EXP = (AGE-11-6);
*IF GRADE92 = 38 THEN EXP = (AGE-12-6);
*IF GRADE92 = 39 THEN EXP = (AGE-12-6);
*IF GRADE92 = 40 THEN EXP = (AGE-13-6);
*IF GRADE92 = 41 THEN EXP = (AGE-14-6);
*IF GRADE92 = 42 THEN EXP = (AGE-14-6);
*IF GRADE92 = 43 THEN EXP = (AGE-16-6);
*IF GRADE92 = 44 THEN EXP = (AGE-18-6);
*IF GRADE92 = 45 THEN EXP = (AGE-18-6);
*IF GRADE92 = 46 THEN EXP = (AGE-18-6);
*IF EXP <= 0 THEN DELETE;
*EXP_SQ = EXP*EXP;

* HOURLY WAGE *;
*WAGE = EARNHRE/100; * HOURLY WAGE CONVERTED TO DOLLARS *;
*LNWAGE = LOG(WAGE);

* RACE *;
*DRACE = 0;
```

```
*IF RACE = 1 THEN DRACE = 1; * WHITE = 1 *;

* SEX *;
*DSEX = 0;
*IF SEX = 1 THEN DSEX = 1; * MALE = 1 *;

* REGIONAL DUMMIES *;
*NE = 0; * NEW ENGLAND DIVISION *;
*MA = 0; * MIDDLE ATLANTIC DIVISION *;
*EN = 0; * EAST NORTH CENTRAL DIVISION *;
*WN = 0; * WEST NORTH CENTRAL DIVISION *;
*SA = 0; * SOUTH ATLANTIC *;
*ES = 0; * EAST SOUTH CENTRAL *;
*WS = 0; * WEST SOUTH CENTRAL *;
*MT = 0; * MOUNTAIN *;
*PC = 0; * PACIFIC *;

*IF STATE < 20 THEN NE = 1;
*IF STATE < 30 AND STATE > 20 THEN MA = 1;
*IF STATE < 40 AND STATE > 30 THEN EN = 1;
*IF STATE < 50 AND STATE > 40 THEN WN = 1;
*IF STATE < 60 AND STATE > 50 THEN SA = 1;
*IF STATE < 70 AND STATE > 60 THEN ES = 1;
*IF STATE < 80 AND STATE > 70 THEN WS = 1;
*IF STATE < 90 AND STATE > 80 THEN MT = 1;
*IF STATE < 100 AND STATE > 90 THEN PC = 1;

* INDUSTRY DUMMIES *;
*AGR = 0; * AGRICULTURE, FORESTRY, AND FISHERIES *;
*MIN = 0; * MINING *;
*CON = 0; * CONSTRUCTION *;
*DUR = 0; * DURABLE GOODS *;
*NDR = 0; * NONDURABLE GOODS *;
*TRN = 0; * TRANSPORTATION AND COMMUNICATION*;
*TRD = 0; * WHOLESALE AND RETAIL TRADE *;
*FIR = 0; * FINANCE, INSURANCE, AND REAL ESTATE *;
*BRS = 0; * BUSINESS AND REPAIR SERVICES *;
*PER = 0; * PERSONAL SERVICES *;
*ENT = 0; * ENTERTAINMENT AND RECREATIONAL SERVICES *;
*PRO = 0; * PROFESSIONAL AND RELATED SERVICES *;
*PAD = 0; * PUBLIC ADMINISTRATION *;
*NON = 0; * WORKERS NOT CLASSIFIABLE BY INDUSTRY *;

*IF IND80 < 40 THEN AGR = 1;
*IF IND80 < 60 AND IND80 > 39 THEN MIN = 1;
*IF IND80 < 61 AND IND80 > 59 THEN CON = 1;
*IF IND80 < 400 AND IND80 > 229 THEN DUR = 1;
*IF IND80 < 230 AND IND80 > 99 THEN NDR = 1;
*IF IND80 < 500 AND IND80 > 399 THEN TRN = 1;
*IF IND80 < 700 AND IND80 > 499 THEN TRD = 1;
*IF IND80 < 721 AND IND80 > 699 THEN FIR = 1;
*IF IND80 < 761 AND IND80 > 720 THEN BRS = 1;
*IF IND80 < 800 AND IND80 > 760 THEN PER = 1;
*IF IND80 < 812 AND IND80 > 799 THEN ENT = 1;
*IF IND80 < 900 AND IND80 > 811 THEN PRO = 1;
*IF IND80 < 990 AND IND80 > 899 THEN PAD = 1;
*IF IND80 > 990 THEN NON = 1;
```

```
*IF AGR = 1 THEN DELETE; * DELETES AGRICULTURE, FORESTRY, AND FISHERIES
*;
*IF NON = 1 THEN DELETE; * DELETES WORKERS WITH NO INDUSTRY
CLASSIFICATION *;

* CREATING INDUSTRY VARIABLE FOR SORT *;
*IF AGR = 1 THEN IND = 1;
*IF MIN = 1 THEN IND = 2;
*IF CON = 1 THEN IND = 3;
*IF DUR = 1 THEN IND = 4;
*IF NDR = 1 THEN IND = 5;
*IF TRN = 1 THEN IND = 6;
*IF TRD = 1 THEN IND = 7;
*IF FIR = 1 THEN IND = 8;
*IF BRS = 1 THEN IND = 9;
*IF PER = 1 THEN IND = 10;
*IF ENT = 1 THEN IND = 11;
*IF PRO = 1 THEN IND = 12;
*IF PAD = 1 THEN IND = 13;
*IF NON = 1 THEN IND = 14;

* OCCUPATIONAL DUMMIES *;
*PTMS = 0; * MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS *;
         * AND TECHNICAL, SALES, AND ADMINISTRATIVE SUPPORT OCCUPATIONS
*;
*SERV = 0; * SERVICE OCCUPATIONS *;
*CRFT = 0; * PRECISION PRODUCTION, CRAFT, AND REPAIR OCCUPATIONS *;
*OPLA = 0; * OPERATORS, FABRICATORS, AND LABORERS *;
*FARM = 0; * FARMING, FORESTRY, AND FISHING OCCUPATIONS *;

*NOOC = 0; * WORKERS NOT CLASSIFIABLE BY OCCUPATION *;

*IF OCC80 < 200 AND OCC80 > 0 THEN PTMS = 1;
*IF OCC80 < 390 AND OCC80 > 199 THEN PTMS = 1;
*IF OCC80 < 470 AND OCC80 > 400 THEN SERV = 1;
*IF OCC80 < 700 AND OCC80 > 500 THEN CRFT = 1;
*IF OCC80 < 890 AND OCC80 > 700 THEN OPLA = 1;
*IF OCC80 < 500 AND OCC80 > 470 THEN FARM = 1;
*IF OCC80 > 900 THEN NOOC = 1;

*IF FARM = 1 THEN DELETE; * DELETES ALL FARMERS AND FARM MANAGERS *;
                    * AND FARM LABORERS AND FARM FOREMEN *;
*IF NOOC = 1 THEN DELETE; * DELETES WORKERS NOT CLASSIFIABLE BY
OCCUPATION *;

* GOVERNMENT EMPLOYEE DUMMIES *;
*GOVT = 0;
*IF CLASS > 1 AND CLASS < 5 THEN GOVT = 1;

* MARRIED SPOUSE PRESENT *;
*MARRIED = 0;
*IF MARITAL = 1 THEN MARRIED = 1; *NOT INCLUDING ARMED FORCES IN
DEFINITION*;

*RACE_EXP = EXP*DRACE;
*RACE_ESQ = EXP_SQ*DRACE;
```

```
*PROC MEANS; *VAR EXP WAGE;

*IF GRADE92 ^= 39 THEN DELETE; * DELETES NON HIGH SCHOOL GRADS-OR EQUIV
*;
*IF GOVT = 1 THEN DELETE; * DELETES GOVERNMENT EMPLOYEES *;
*IF PAD = 1 THEN DELETE; * DELETES THE REST OF THE GOVERNMENT EMPLOYEES
*;
*IF DSEX = 0 THEN DELETE; * DELETES FEMALES *;


*PROC SORT; *BY IND DRACE;
*PROC MEANS; *VAR SCHOOL MARRIED;
*PROC REG S; *BY IND DRACE;
      *MODEL LNWAGE = EXP RACE_EXP EXP_SQ RACE_ESQ DRACE MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *MODEL LNWAGE = EXP EXP_SQ MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*PROC SORT; *BY IND;
*PROC MEANS; *VAR SCHOOL MARRIED;
*PROC REG S; *BY IND;
      *MODEL LNWAGE = EXP RACE_EXP EXP_SQ RACE_ESQ DRACE MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *MODEL LNWAGE = EXP EXP_SQ SCHOOL PTMS SERV CRFT GOVT DRACE DSEX
                      MIN DUR NDR TRN TRD FIR BRS PER ENT PRO PAD
                  NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*PROC REG S; *BY IND;
      *MODEL LNWAGE = EXP EXP_SQ SCHOOL PTMS SERV CRFT GOVT DRACE DSEX
                  NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*run;


* -------------VARIABLES USED FOR YEARS 1994-1997------------- *;

*data CPS97; *set "C:\STEPHEN\CPS94F.sd2";

* YEARS OF EXPERIENCE *;
*IF GRADE92 = 31 THEN EXP = (AGE-0-6);
*IF GRADE92 = 32 THEN EXP = (AGE-2.5-6);
*IF GRADE92 = 33 THEN EXP = (AGE-5.5-6);
*IF GRADE92 = 34 THEN EXP = (AGE-7.5-6);
*IF GRADE92 = 35 THEN EXP = (AGE-9-6);
*IF GRADE92 = 36 THEN EXP = (AGE-10-6);
*IF GRADE92 = 37 THEN EXP = (AGE-11-6);
*IF GRADE92 = 38 THEN EXP = (AGE-12-6);
*IF GRADE92 = 39 THEN EXP = (AGE-12-6);
*IF GRADE92 = 40 THEN EXP = (AGE-13-6);
*IF GRADE92 = 41 THEN EXP = (AGE-14-6);
*IF GRADE92 = 42 THEN EXP = (AGE-14-6);
*IF GRADE92 = 43 THEN EXP = (AGE-16-6);
*IF GRADE92 = 44 THEN EXP = (AGE-18-6);
```

```
*IF GRADE92 = 45 THEN EXP = (AGE-18-6);
*IF GRADE92 = 46 THEN EXP = (AGE-18-6);
*IF EXP <= 0 THEN DELETE;
*EXP_SQ = EXP*EXP;

* HOURLY WAGE *;
*WAGE = EARNHRE/100; * HOURLY WAGE CONVERTED TO DOLLARS *;
*LNWAGE = LOG(WAGE);

* RACE *;
*DRACE = 0;
*IF RACE = 1 THEN DRACE = 1; * WHITE = 1 *;

* SEX *;
*DSEX = 0;
*IF SEX = 1 THEN DSEX = 1; * MALE = 1 *;

* REGIONAL DUMMIES *;
*NE = 0; * NEW ENGLAND DIVISION *;
*MA = 0; * MIDDLE ATLANTIC DIVISION *;
*EN = 0; * EAST NORTH CENTRAL DIVISION *;
*WN = 0; * WEST NORTH CENTRAL DIVISION *;
*SA = 0; * SOUTH ATLANTIC *;
*ES = 0; * EAST SOUTH CENTRAL *;
*WS = 0; * WEST SOUTH CENTRAL *;
*MT = 0; * MOUNTAIN *;
*PC = 0; * PACIFIC *;

*IF STATE < 20 THEN NE = 1;
*IF STATE < 30 AND STATE > 20 THEN MA = 1;
*IF STATE < 40 AND STATE > 30 THEN EN = 1;
*IF STATE < 50 AND STATE > 40 THEN WN = 1;
*IF STATE < 60 AND STATE > 50 THEN SA = 1;
*IF STATE < 70 AND STATE > 60 THEN ES = 1;
*IF STATE < 80 AND STATE > 70 THEN WS = 1;
*IF STATE < 90 AND STATE > 80 THEN MT = 1;
*IF STATE < 100 AND STATE > 90 THEN PC = 1;

* INDUSTRY DUMMIES *;
*AGR = 0; * AGRICULTURE, FORESTRY, AND FISHERIES *;
*MIN = 0; * MINING *;
*CON = 0; * CONSTRUCTION *;
*DUR = 0; * DURABLE GOODS *;
*NDR = 0; * NONDURABLE GOODS *;
*TRN = 0; * TRANSPORTATION AND COMMUNICATION*;
*TRD = 0; * WHOLESALE AND RETAIL TRADE *;
*FIR = 0; * FINANCE, INSURANCE, AND REAL ESTATE *;
*BRS = 0; * BUSINESS AND REPAIR SERVICES *;
*PER = 0; * PERSONAL SERVICES *;
*ENT = 0; * ENTERTAINMENT AND RECREATIONAL SERVICES *;
*PRO = 0; * PROFESSIONAL AND RELATED SERVICES *;
*PAD = 0; * PUBLIC ADMINISTRATION *;
*NON = 0; * WORKERS NOT CLASSIFIABLE BY INDUSTRY *;

*IF IND80 < 40 THEN AGR = 1;
*IF IND80 < 60 AND IND80 > 39 THEN MIN = 1;
*IF IND80 < 61 AND IND80 > 59 THEN CON = 1;
```

```
*IF IND80 < 400 AND IND80 > 229 THEN DUR = 1;
*IF IND80 < 230 AND IND80 > 99 THEN NDR = 1;
*IF IND80 < 500 AND IND80 > 399 THEN TRN = 1;
*IF IND80 < 700 AND IND80 > 499 THEN TRD = 1;
*IF IND80 < 721 AND IND80 > 699 THEN FIR = 1;
*IF IND80 < 761 AND IND80 > 720 THEN BRS = 1;
*IF IND80 < 800 AND IND80 > 760 THEN PER = 1;
*IF IND80 < 812 AND IND80 > 799 THEN ENT = 1;
*IF IND80 < 900 AND IND80 > 811 THEN PRO = 1;
*IF IND80 < 990 AND IND80 > 899 THEN PAD = 1;
*IF IND80 > 990 THEN NON = 1;

*IF AGR = 1 THEN DELETE; * DELETES AGRICULTURE, FORESTRY, AND FISHERIES
*;
*IF NON = 1 THEN DELETE; * DELETES WORKERS WITH NO INDUSTRY
CLASSIFICATION *;

* CREATING INDUSTRY VARIABLE FOR SORT *;
*IF AGR = 1 THEN IND = 1;
*IF MIN = 1 THEN IND = 2;
*IF CON = 1 THEN IND = 3;
*IF DUR = 1 THEN IND = 4;
*IF NDR = 1 THEN IND = 5;
*IF TRN = 1 THEN IND = 6;
*IF TRD = 1 THEN IND = 7;
*IF FIR = 1 THEN IND = 8;
*IF BRS = 1 THEN IND = 9;
*IF PER = 1 THEN IND = 10;
*IF ENT = 1 THEN IND = 11;
*IF PRO = 1 THEN IND = 12;
*IF PAD = 1 THEN IND = 13;
*IF NON = 1 THEN IND = 14;

* OCCUPATIONAL DUMMIES *;
*PTMS = 0; * MANAGERIAL AND PROFESSIONAL SPECIALTY OCCUPATIONS *;
         * AND TECHNICAL, SALES, AND ADMINISTRATIVE SUPPORT OCCUPATIONS
*;
*SERV = 0; * SERVICE OCCUPATIONS *;
*CRFT = 0; * PRECISION PRODUCTION, CRAFT, AND REPAIR OCCUPATIONS *;
*OPLA = 0; * OPERATORS, FABRICATORS, AND LABORERS *;
*FARM = 0; * FARMING, FORESTRY, AND FISHING OCCUPATIONS *;

*NOOC = 0; * WORKERS NOT CLASSIFIABLE BY OCCUPATION *;

*IF OCC80 < 200 AND OCC80 > 0 THEN PTMS = 1;
*IF OCC80 < 390 AND OCC80 > 199 THEN PTMS = 1;
*IF OCC80 < 470 AND OCC80 > 400 THEN SERV = 1;
*IF OCC80 < 700 AND OCC80 > 500 THEN CRFT = 1;
*IF OCC80 < 890 AND OCC80 > 700 THEN OPLA = 1;
*IF OCC80 < 500 AND OCC80 > 470 THEN FARM = 1;
*IF OCC80 > 900 THEN NOOC = 1;

*IF FARM = 1 THEN DELETE; * DELETES FARMERS AND FARM MANAGERS *;
                    * AND FARM LABORERS AND FARM FOREMEN *;
*IF NOOC = 1 THEN DELETE; * DELETES WORKERS NOT CLASSIFIABLE BY
OCCUPATION *;
```

```
* GOVERNMENT EMPLOYEE DUMMIES *;
*GOVT = 0;
*IF CLASS94 < 4 THEN GOVT = 1;

* MARRIED SPOUSE PRESENT *;
*MARRIED = 0;
*IF MARITAL = 1 THEN MARRIED = 1; *NOT INCLUDING ARMED FORCES IN
DEFINITION*;

*RACE_EXP = EXP*DRACE;
*RACE_ESQ = EXP_SQ*DRACE;

*IF GRADE92 ^= 39 THEN DELETE; * DELETES NON HIGH SCHOOL GRADS-OR EQUIV
*;
*IF GOVT = 1 THEN DELETE; * DELETES GOVERNMENT EMPLOYEES *;
*IF PAD = 1 THEN DELETE; * DELETES THE REST OF THE GOVERNMENT EMPLOYEES
*;
*IF DSEX = 0 THEN DELETE; * DELETES FEMALES *;

*PROC SORT; *BY IND DRACE;
*PROC MEANS; *VAR SCHOOL MARRIED;
*PROC REG S; *BY IND DRACE;
      *MODEL LNWAGE = EXP RACE_EXP EXP_SQ RACE_ESQ DRACE MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *MODEL LNWAGE = EXP EXP_SQ MARRIED
                  PTMS SERV CRFT     NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*PROC REG S; *BY IND;
      *MODEL LNWAGE = EXP EXP_SQ SCHOOL PTMS SERV CRFT GOVT DRACE DSEX
                  NE MA EN WN SA ES WS MT;
      *TEST EXP, EXP_SQ;

*run;
```