Job Matching: The Effects of Job Search on Match Quality[1]

Kristin L. Terris
Department of Economics
Georgetown University
Washington, DC 20057

November 2004

In this paper, I analyze a large sample from the National Longitudinal Survey of Youth to characterize the duration of employment for an individual's first job after leaving school. I find that employment duration has a significant positive correlation with the length of the preceding non-employment spell. I take this as evidence that time spent in job search improves the quality of the match. I also find that there are positive returns to schooling through match quality. In addition, factors that negatively impact an individual's ability to spend time to search for a job increase the hazard out of employment and factors related to family stability decrease the hazard out of employment.

JEL Classification: J00; J21; J64
Keywords: Employment, Unemployment; Match Quality; Returns to Schooling, Job Search

[1] I would like to thank my advisors, Susan Vroman, James Albrecht, and Daniel Westbrook for their significant and valuable contributions to the development of this paper. All remaining errors are my own.

I. Introduction

This paper is an empirical study of factors that contribute to match quality. Match quality itself is difficult to measure empirically; however, there are literary precedents that equate match quality and employment duration, Jovanovic (1979) and Bowlus (1995). A job match is an experience good in that firms and workers must experience the match to evaluate its quality. Firms and workers negotiate an employment contract that leads to a balance between compensation and productivity. The length of employment duration, which arises from that contract, is an empirically observed measure of the success of the match. Holding other effects constant, high quality matches are characterized by long employment durations.

In order to investigate match quality, I divide the relevant covariates into factors that contribute to match quality and control variables. Employment duration is easily measurable and is a primary factor in the determination of match quality. I explore the relationship between employment duration and non-employment duration. I focus on the transition between non-employment and employment to evaluate the impact that the time an individual spends in non-employment has on the duration of the subsequent employment spell. I restrict the analysis to the duration of a single job where job-to-job changes represent the dissolution of one match and the formation of another match. I analyze the effects of workers' behavior during non-employment as another contribution to the determination of match quality and control for personal characteristics, including level of education.

One could easily conclude that the correlation between employment duration and unemployment duration should be negative. High ability workers are more proactive, get jobs faster and then keep jobs longer whereas low ability workers take a long time to find work and their matches dissolve more quickly as they reveal their low ability status to their employers. On the other hand, people who spend more time in job search make a better quality match. A positive relationship between non-employment duration and employment duration may be evidence that job search has a larger contribution to good match quality than ability alone.

Different types of search models yield competing predictions with respect to the relationship between the duration of employment and the duration of non-employment. For example, a Burdett-Mortensen Model [Burdett and Mortensen (1998)] predicts that the duration of employment is uncorrelated with the duration of the preceding non-employment spell.[2] In contrast, an Assignment Model [Shimer (2000)] predicts that the duration of employment is positively correlated with the duration of the preceding non-employment spell.[3] Between these two models, I find evidence to support the prediction of the Assignment Model.

The data source for this paper is the National Longitudinal Survey of Youth, (NLSY79). This is a comprehensive panel data set that provides detailed information on individual

[2] The Burdett-Mortensen search model predicts that the duration of employment is solely determined by the exogenous job arrival rate and the exogenous job destruction rate and will have no correlation to the duration of the previous non-employment spell.
[3] The assignment model predicts that the duration of employment is determined by the probability of competing for a job with a more qualified worker along with the rate of job arrival and job destruction. Therefore, this model predicts that a longer time in non-employment or more job search will improve the workers probability of making a good match.

employment durations for a group of individuals over an extended period of time. These data have been frequently used and are well suited for labor applications.

My study produces three main results. First, there is a significant positive correlation between the length of the non-employment spell and the length of the subsequent employment spell. Second, as levels of education increase, the hazard out of employment decreases. Finally, factors that negatively affect an individual's ability to conduct a job search during the non-employment spell (such as part time work, having children, and spending time out of the labor force) increase the hazard out of employment. In addition, among the control variables, factors related to family stability (such as marital status and children) decrease the hazard out of employment.

The remainder of this paper is organized as follows. In the next section, I provide a brief review of the related literature. In section III, I describe the data. Section IV contains the results of the hazard rate analysis and the proportional hazard model. The last section contains concluding remarks.

II. Literature Review

Most empirical studies that use duration modeling techniques focus on either unemployment or employment. Many authors have characterized the distributions of unemployment and wages, and the job stability literature focuses on the distribution of employment. Among those who study employment, none have attempted to study the correlation between unemployment duration and the duration of the subsequent employment spell.

Eckstein and Wolpin, in their 1990 *Econometrica* paper, were the first to use the NLSY to characterize the duration of non-employment and wages for young workers in their first jobs. Their work primarily aims to demonstrate the feasibility of estimating an equilibrium search model [Albrecht and Axell (1994)]. Their analysis of the data is limited to observed hazard rates for non-employment.[4] I have extended their initial work by using their data sample, addressing the link between employment and non-employment and by exploring specific behaviors of individuals during non-employment.

Previous work on match quality focuses on the effect of the business cycle on employment duration [Bowlus (1995)]. She uses the NLSY with the unemployment rate as the cyclical variable and discovers that there is variation in match quality over the business cycle. She also finds that controlling for wages reduces this effect significantly. I use Bowlus's definition of match quality to document evidence on the correlation between job search and match quality. [5]

The cornerstone of the applied empirical literature in this field is the Meyer (1990) *Econometrica* article that introduces a semiparametric duration analysis technique and applies it to test the effects of the level and length of unemployment insurance benefits (UI) on unemployment durations. Meyer uses duration modeling and hazard rate analysis

[4] Eckstein and Wolpin use the term unemployment and non-employment interchangeably.
[5] I have excluded wages from my analysis due to the high degree of correlation between wages, unemployment duration, and employment duration.

to conclude that increasing UI benefits has a negative effect on the probability of leaving unemployment. His conclusions imply that high UI benefits implicitly decrease the cost of job search and leisure, but he restricts his analysis to the duration of the unemployment spell. Meyer introduces a multiplicative form of unobserved heterogeneity to the proportional hazard model. He concludes that the coefficients with gamma distributed unobserved heterogeneity are similar to those obtained with the no heterogeneity specification. He also concludes that the non-parametric specification of the baseline hazard substantially reduces the inconsistency effects of misspecifying the baseline hazard. I apply some of Meyer's estimation techniques in this paper.

Previous work on the characterization of employment focuses on job stability. The job stability literature documents the evolution of job retention rates and wages in the United States [Diebold, Neumark, and Polsky (1997)], but does not provide any evidence on the relationship between unemployment duration and job stability. In addition, job stability is measured by the job retention rate, which is averaged over individuals and does not control for heterogeneity.

As is the case in this study, previous empirical studies on employment and unemployment have been "one-sided" due to the fact that individual characteristics are well defined in the data, but firm characteristics are not. In an analysis that includes both individual and firm characteristics, Abowd, Kramarz and Margolis (1999) find that person effects are more important than firm effects for explaining wage variation in a French data set. In a related study, Postel-Vinay and Robin (2002) find that person effects are important, accounting for up to 40% of wage variation, but that market imperfections caused by search frictions account for up to 50% of wage variation. These papers provide evidence that studies using “one-sided” data sources may still yield useful results.

All of the above-mentioned work examines either the distribution of employment or the distribution of non-employment, but it does not address the question of whether the amount of time spent in non-employment has an effect on match quality. To investigate the link between employment and non-employment requires that both distributions be examined to determine their degree of correlation.

III. Data Description

The NLSY79[6] is a national survey of men and women born in the years 1957-64. Respondents were ages 14-22 when first interviewed in 1979. The survey is composed of 12,686 individuals who resided throughout the United States. The survey is ongoing and was conducted each year until 1994. After 1994 the participants were interviewed every two years. Interviewers collected detailed information to describe the individuals’ demographics, family backgrounds, labor market participation, education, etc. Individuals reported information on their work history but were asked to limit their report to a maximum of five of their most recent jobs in each survey period through 2002. In each survey year there were links in the data to previously held jobs in order to provide the ability to generate a continuous work history for each respondent.

[5] See http://www.bls.gov/nls/home.htm.

Firms are partners in the matches that produce employment spells, therefore I would like to be able to control for firm characteristics. There are firm identifiers such as Employer ID and Industry, but unfortunately, the NLSY79 contains few indicators of firm characteristics prior to 1998. The identity of the firm is strictly confidential. As a result of this limitation of the NLSY79 data, I cannot control for firm heterogeneity in this empirical model. However, studies that include both characteristics of firms and characteristics of workers have found that firm effects are not as important as person effects when estimating labor models [Abowd, Kramarz and Margolis (1998)], and [Postel-Vinay and Robin (2002)].

In the sample I use, an employment spell begins with an individual's job start date and ends with the job stop date for a given employer. I construct the variable EMP, which records the length of the employment spell measured in weeks. The occupation category assigned to each job is the 3-digit CPS code (Current Population Survey) defined by the 1970 and 1980 Census occupational classification system. A job change by an individual who continues working for the same employer can be identified only if the CPS code also changes. I also construct the variable NEMP, which measures the length of non-employment in weeks. Non-employment is the number of weeks between an individual's departure from school and the first reported date of employment. A non-employment spell for an individual may include time spent unemployed, time spent out of the labor force, or some combination of the two.

There are over 100,000 employment spells that are preceded by non-employment in this data. There can be multiple employment spells for each individual that are likely to be correlated with the spell durations. These correlations may contain useful information, however, exploiting these correlations is complicated. For now I limit the analysis to the first spell per individual.

The sample I used contains the duration of the first "real" job held by the individual after leaving school. The definition of a real job in this study is the same as the definition of a real job specified by Eckstein and Wolpin (1990). A real job is one characterized by employment hours of at least 30 hours per week and with employment duration of at least three months. An employment spell ends when the individual leaves the position either to enter into a non-employment spell or to accept a job with a different occupation code or at a different firm. This sample contains 8,590 observations, and includes data through 2002.[7] The potential ages of these individuals are between 14 and 44, and the mean age is 22. The sample includes spells for those who graduated as well as for those who left school prior to graduation at both the high school and college level. A small number of individuals, 14 men and 42 women, are excluded because they never reported holding a job. Table 1 contains a summary of the duration variables.

[7] At each interview, respondents were asked about the period of time between the current interview and the last interview, so a continuous work history has been maintained for all individuals who remained in the sample even if there were years in which the individual was not interviewed.

Summary of Duration Variables in (weeks)

	N	Mean	Median	SD	Min	Max
Some High School						
Employment	860	54	18	105	1	948
Non-employment	860	156	109	171	0	1241
High School						
Employment	5061	93	27	181	1	1383
Non-employment	5061	92	45	136	0	1164
Some College						
Employment	1269	103	34	179	1	1211
Non-employment	1269	93	42	148	0	1098
College						
Employment	1400	134	53	195	1	1100
Non-employment	1400	63	18	118	0	1029
Total Sample						
Employment	8590	97	30	178	1	1383
Non-employment	8590	94	43	141	0	1241

Table 1

The education cohorts are defined relative to the highest degree reported by the respondents. The category *Some High School* includes all those who did not report graduating from high school, nor did they report earning a GED certification. The *High School* category includes those whose maximum education training captures high school graduates and those with GED certifications. Those in the *Some College* category graduated from high school and then reported attending college, junior college, or trade school but did not earn a bachelor's degree. And finally the *College* category includes those who graduated from college with at least a bachelor's degree.

The employment duration data are censored in two ways. Employment spells are right-censored at the last valid survey date for those individuals who report that they are currently employed at the date of their last survey. Employment spells are also right-censored at the last reported survey date for individuals who leave the sample. Only three percent of the employment spells are right-censored.

The major advantages of the NLSY data source are the large number of covariates available for each individual and the ability to use a panel of these covariates to identify a pattern of job search behavior over a specific time period. To characterize a worker, I use the values of time-varying covariates defined at the beginning of the employment spell. Factors that influence a worker's ability to conduct a job search during non-employment include whether the individual actively searched for a job, the intensity of that search, whether the individual was out of the labor force, whether the individual had children, and whether the individual worked part time. I have assigned the value of each of the respective dummy variables to be one if the individual exhibited the behavior for at least one week during the interval and zero otherwise. I constructed the variable that characterizes search intensity by including how many weeks the individual searched for a job while not employed added to a measure of the number of different search sources they utilized. A search source includes tasks such as consulting a newspaper, networking, hiring an employment agency, etc. The dummy variable takes on the value one if the individual had a constructed search index higher than average and zero otherwise.

It is interesting to note that the mean employment duration of an individual's first real job is 97 weeks and the average time to acceptance of that job is 94 weeks, which indicates that respondents spent nearly as much time looking for work as they did working. This finding is somewhat different from the results of Eckstein and Wolpin (1990), who observed a mean employment spell of 34 weeks and a mean non-employment spell of 45 weeks. The difference in these two results is due to the extended sample time. This sample includes observations on each individual for 23 years. In their paper, Eckstein and Wolpin considered just three years of data and a large percentage of their sample observations were censored.

In some studies, females are excluded from the sample due to concerns that females have higher probabilities of terminating a match due to reasons other than match quality [Bowlus (1995)]. To evaluate the duration of employment, I include both female and male workers. The results do not differ significantly between genders. Table 2 contains descriptive statistics for the sample including an individual's personal characteristics and characteristics of an individual's behavior on the non-employment interval. I make a distinction between an individual who had children in the household, and an individual who had children during the non-employment interval. The dummy variable characterized by "Children in the Household" is one if there were children present in the household at the time of employment. The dummy variable characterized by "Children Born During the Interval" is one if the individual had a child born or adopted during the non-employment interval.

Male workers make up 49% of the spells of employment. Over 20% of the employment spells are attributed to married workers. High school graduates work 56% of employment spells. Those with some high school comprise 12% of the workforce. Respondents with some college education and college graduates make up 15% and 17% of the workforce respectively. Over 22% of individuals in the sample engaged in part-time work before being hired into a real job.

Summary of Individual Characteristics

	Mean	Median	SD	Min	Max
Personal Characteristics					
Age	21.69	20	4.39	14	44
Gender	0.49		0.50		
White	0.68		0.47		
Black	0.27		0.44		
Other	0.05		0.22		
Married	0.20		0.40		
Children in the Household	0.24		0.43		
Job Characteristics					
Wage	8.63	4.4	6.39	0	98
Hours	42.80	40	8.49	31	96
Union	0.13		0.34		
Interval Behavior					
Active Search	0.82		0.38		
Search Intensity	0.29		0.46		
Part Time Work	0.22		0.41		
Children Born During the Interval	0.33		0.81		
Time Spent Out of the Labor Force	0.33		0.47		

Table 2

Since the employment spells are for an individual's first job, the sample deals primarily with young people. It is interesting to note that 47% of the employment spells terminate in the first 26 weeks and 63% of employment spells are completed by the end of the first year. The empirical frequency of employment duration is given in Figure 1. The empirical frequency of non-employment duration is in Figure 2. Non-employment spells are very uniformly distributed in this sample. Nearly 36% percent of individuals obtained real jobs within six months of leaving school. However, only nine percent were employed in their first real jobs within a month of the date they left school. By the end of the first year, 54% of the individuals in the sample were employed, and by the end of the second year, 73% of individuals had real jobs. Only 17% of the respondents did not have real jobs three years after leaving school. Fifty-two percent of the individuals who had non-employment spells greater than 104 weeks were female, and 80% of those women, had children in their household.

Distribution of Employment Duration

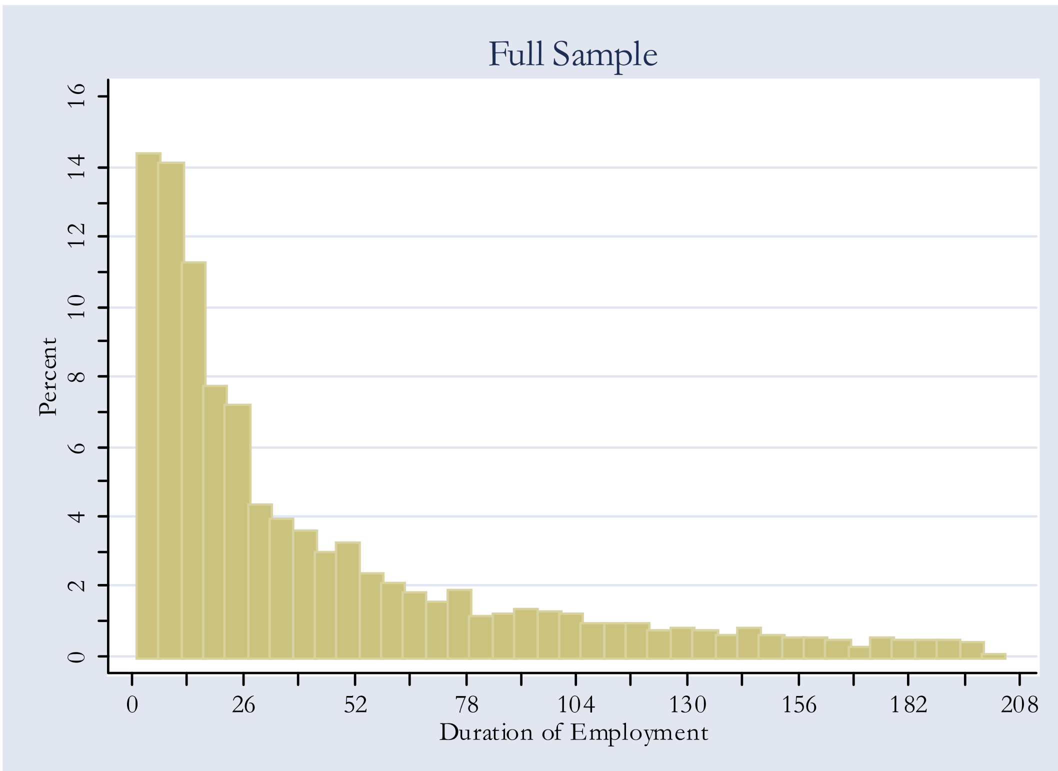


Figure 1

Distribution of Non-employment Duration

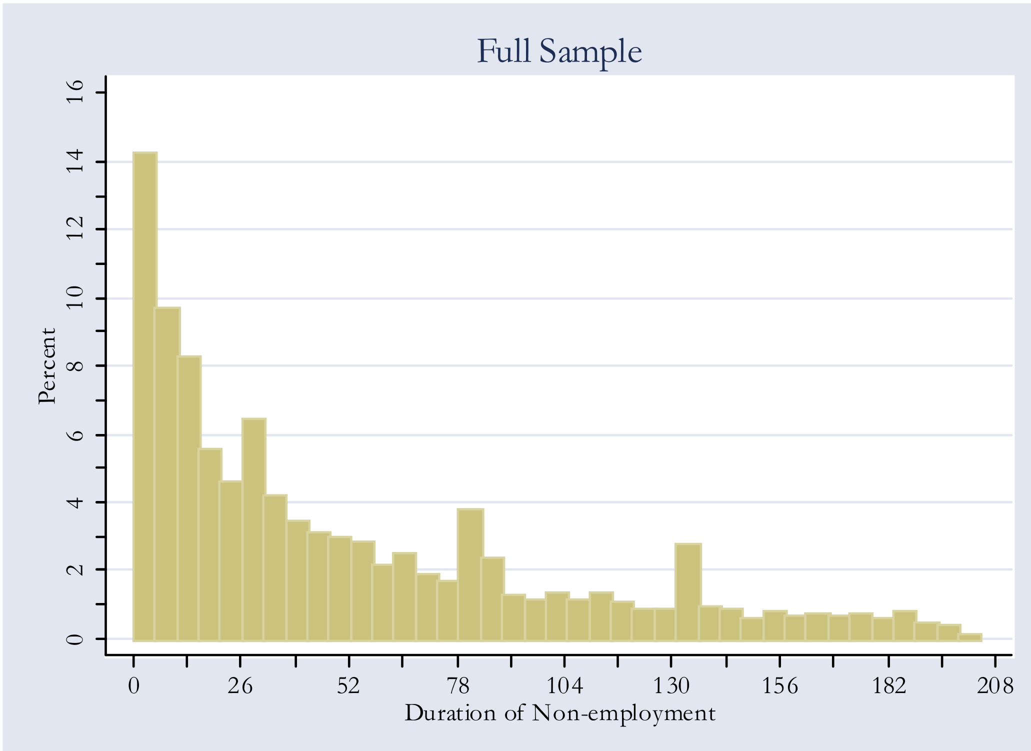


Figure 2

IV. Estimation Methodology and Results

Figure 3 shows the distribution of employment duration using Kaplan-Meier estimates of the survivor function by education cohort. The survivor function is a measure of the probability that an individual will be employed in period *t+1* given he is employed at time *t*. The survivor function is calculated as:

$$\hat{S}(T_j) = \prod_{i=1}^{j} \frac{n_i - d_i}{n_i}$$

The index *i* is the number of discrete spells in rank order of duration, n_j represents the number of employment spells at risk at time T_j, and d_j is the number of completed employment spells at T_j. Clearly the hazard out of employment decreases as education increases. This represents initial statistical evidence of returns to schooling through employment duration.

Distribution of Employment Duration by Level of Education

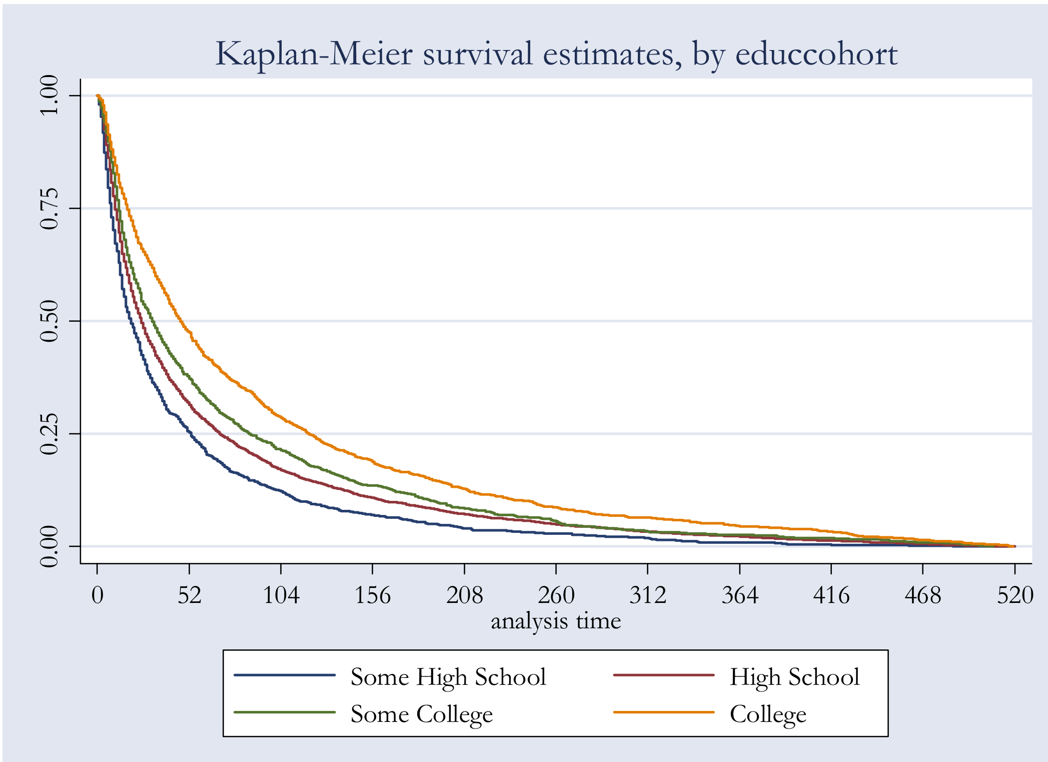
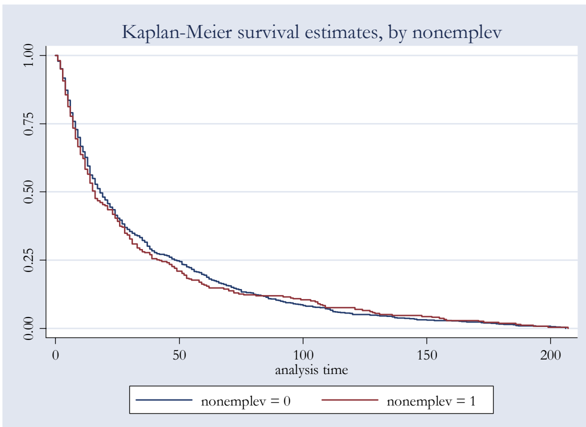


Figure 3

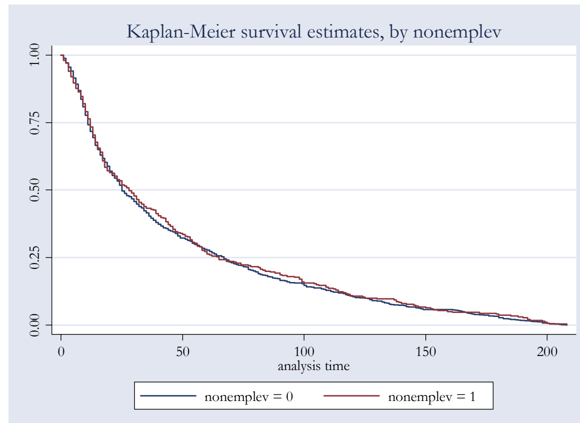
To illustrate the relationship between the length of non-employment spells and the length of subsequent employment spells, Figure 4 shows two estimated survivor functions: one for individuals who experienced shorter-than-average non-employment durations and one for individuals who experienced longer-than-average non-employment durations for each education cohort. The important feature of this figure is that the survivor function for employment is consistently higher for longer non-employment durations. A simple log-rank test indicates that the differences in these two distributions are statistically significant at or below the 10% level.

Survivor Estimates by Level of Non-Employment

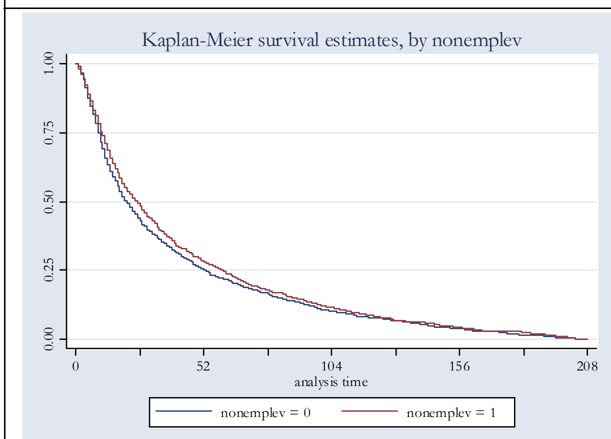
Some High School



Some College



High School Graduates



College Graduates

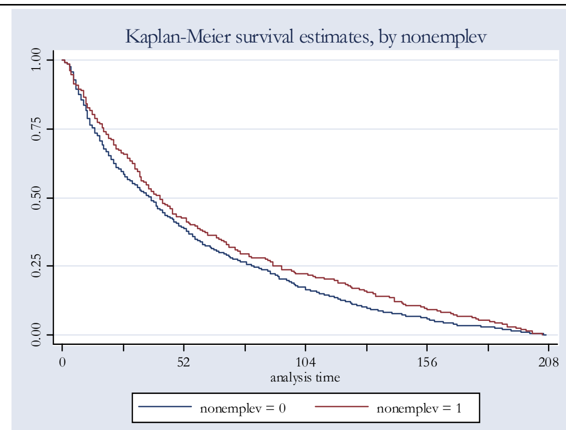


Figure 4

The same basic pattern can be seen in Table 3, which summarizes the hazard rate for the entire sample. The hazard rate $\hat{\lambda}(t)$ is the probability that an individual exits employment at time *t+1* given that the individual was employed at time *t*. The integrated hazard $\hat{H}(t)$ is simply the sum of the hazard rates to date. The calculations are as follows:

$$\hat{\lambda}(t_j) = \frac{d_j}{n_j} \qquad \hat{H}(t_j) = \sum_{i=1}^{j} \lambda(t_i)$$

In Table 3, it is interesting to note that the integrated hazard rate for the first stratum exceeds the hazard rate for the second stratum for employment durations up to three and one half years. In practical terms, the integrated hazard of the first stratum lies above the integrated hazard distribution of the second stratum for 98% of the observations in the sample.[8]

Kaplan Meier Estimate
Observed Hazard Rates

Employment Number of Weeks	Hazard	Integrated Hazard	Standard Error	First Stratum Integrated Hazard	Second Stratum Integrated Hazard
13 - 26	0.461	0.604	0.010	0.715	0.423
27 - 52	0.650	1.633	0.023	2.006	1.177
53 - 78	0.568	4.682	0.024	2.608	2.215
79 - 104	0.343	2.878	0.047	2.951	2.741
105 - 130	0.231	3.149	0.055	3.163	3.109
131 - 156	0.105	3.270	0.059	3.284	3.233
157 - 182	0.121	3.409	0.066	3.434	3.334

Table 3

Kaplan-Meier estimators do not control for individual heterogeneity. Duration modeling provides methods to introduce covariates into the model. The estimation approach is the Cox partial likelihood technique [Cox (1972) and (1975)]. With this method, no parametric assumptions about the baseline hazard are required. The basic proportional hazards model assumes the relationship:

$$\lambda(t_i \mid \mathbf{X}) = \lambda_0(t_i) \exp\{\beta_1 x_{i1} + ... + \beta_k x_{ik}\}$$

[8] See Figure 1 for the distribution of the duration of employment.

The function λ_0 is the baseline hazard function for each individual. The vector $\mathbf{x_{i1}}$ is the duration of non-employment for individual i in the sample. The remaining $k-1$ vectors in the matrix **X** represent the personal characteristics of all the individuals in the sample and factors that represent an individual's ability to conduct a job search.[9] The parameter β is a vector of unknown coefficients that scales the hazard rate.

The likelihood function using duration data can be expressed as the product of the likelihood contribution $L(j)$ over all individuals where:

$$L(j) = \frac{1}{d_j} \prod_{k \in H_j} \frac{\lambda(t_k, x_k, \beta)}{\sum_{i \in Rj} \lambda(t_k, x_i, \beta)}$$

The likelihood contribution $L(j)$ of the jth observation is the conditional probability that observation j concludes a spell at duration t_j, given that any of the J observations could have been concluded at duration t_j. In other words, the contribution to likelihood is the individual hazard at t_j divided by the sum of hazards in progress at t_j and can be expressed as:

$$\ln L = \sum_{j=1}^{J} \left[\sum_{l \in H_j} x_l \beta - d_j \ln \left[\sum_{i \in R_j} \exp\{x_i \beta\} \right] \right]$$

where j is the index of completed spells in ascending order: $j = \{1, \ldots, k\}$, d_j is the number of completed spells at t_j, H_j is the set of observations that complete at t_j, and R_j is the set of observations that are at risk at t_j. If an observation is a, tie $d_j > 1$, then the contribution to likelihood is the same for each of the tied observations.

[9] A complete list of the variables contained in **X** and the full regression results appear in Appendix A Table 1.

The estimated ($\hat{\boldsymbol{\beta}}$) is found using numerical maximization techniques. A one-unit increase in the covariate of interest will affect the hazard out of employment by $\exp\{\hat{\boldsymbol{\beta}}\}$ which represents the hazard rate for each covariate. Estimation results are presented in Table 4 below.

Maximum Likelihood Estimates - Cox Proportional Hazards Model

Variable Description	Exp β - Coeff	Standard Error	
Personal Characteristics			
Non-employment Duration	0.9992	0.0001	***
Some High School	1.5040	0.0620	***
Some College	0.8664	0.0312	***
College	0.7409	0.0255	***
Family Stability			
Married	0.9113	0.0337	**
Children in the Home	0.9013	0.0526	*
Behavior While Not Employed			
Active Job Search	0.8974	0.0313	***
Intense Job Search	0.7848	0.0286	***
Out of the Labor Force	1.1897	0.0325	***
Part Time Work	1.2709	0.0629	***
Children Born While Not Employed	1.0664	0.0326	**
Number of Observations		8590	
Log Likelihood		-36329	
LR Chi2(51)		479.04	
Prob>Chi2		0.0000	

*** Significance of less than 1%
** Significance of less than 5%
* Significance of less than 10%

Table 4

The results indicate a highly significant positive effect for non-employment duration on the length of the subsequent employment duration. The exponentiated β-coefficient on the non-employment duration variable is less than one, and is significant at the one percent level.[10] An increase in the duration of the non-employment spell results in a reduction of the empirical employment hazard. In other words, an increase in the length of time spent in non-employment reduces the probability that an individual leaves employment. This positive correlation may be interpreted as an indication that workers who have longer search times stay employed longer as well. This finding is robust to changes in the model specification and to exclusion of the censored observations.

The magnitude of the coefficient on the duration of non-employment is not large, but the direction is surprising. Intuitively a negative relationship seems more likely, but the positive result is highly significant. In this model, I have made the assumption that the unobserved components are random and uncorrelated across individuals. In addition, I have also assumed that observed and unobserved personal characteristics are independent. To test the validity of these assumptions, it is necessary to make a provision to include unobserved heterogeneity in the model. But if these assumptions are not valid the resulting coefficient estimate will be biased to zero. Therefore, if correlation exists, or if independence doesn't hold then correcting for these effects will yield coefficients with higher magnitudes. Therefore the uncertainty in this estimate actually increases the positive correlation evident in this data.

In testing the relationship between the duration of non-employment and the duration of employment it is also necessary to consider the possibility that the factors that contribute to non-employment duration may be correlated with the factors that contribute to the

[10] Imposing the assumption that employment is distributed exponentially, for the mean individual, a one percent increase in non-employment duration leads to a corresponding 0.2% increase in employment duration.

duration of employment. In a linear estimation it would be necessary to instrument for non-employment duration, however, this estimate is non-linear. If correlation exists the effect is again, to bias the coefficient estimate to zero.

Job search behavior during non-employment adds supporting evidence to the effects of job search on employment duration. The magnitude of the coefficient on whether or not an individual conducted an official job search represents a 10% decrease in the hazard while the effect of search intensity reduces the hazard by nearly 23%. The effects of both of these variables are highly significant at the one percent level. Conducting a job search, decreases the hazard out of employment and the factors that detract from an individual's job ability to conduct a job search have a corresponding negative effect that is large in magnitude on the individual's resulting job tenure. Taken together, these factors substantiate the claim that job search is an important determinant of the length of the subsequent employment spell.

The effects of factors that influence an individual's ability to conduct a job search are also important. These include whether the individual spent time actively conducting a job search during the non-employment interval, whether the individual spent time out of the labor force, whether the individual spent time in part time work, and whether the individual had children. The main result for this group of covariates is that time spent on activities that distract an individual from an active job search negatively influences the length of the subsequent employment spell.

Working part time jobs during non-employment had a negative effect on the subsequent duration of employment. The estimated effect on the hazard rate is significant at one percent and represents a 27% increase in the probability of exiting employment. One interpretation of this result is that time spent working at part time jobs detracts from the quality of the job search for that individual and therefore increases the probability that the individual will leave the subsequent permanent job sooner.

Individuals who spent time out of the labor force during non-employment had a large and statistically significant increase of the hazard out of employment. Respondents who spent time out of the labor force increased their probability of exiting employment by 19%.

Individuals who had children during non-employment also had less job stability. The coefficient is significant at the five percent level and represents a 6% increase in the probability that an individual will leave employment. The impact of this factor is very strong for male workers. Men who become responsible for children during non-employment have a 20% reduction in job stability.

The results of this study give evidence that the length of a worker's employment spell is directly affected by his job search behavior. While the magnitude of the coefficient on non-employment duration is small, representing a 0.1% decrease in the hazard, it is highly significant regardless of the model specification and always has a positive effect on job tenure. The effects are consistent for both male and female cohorts.

The results on the education variables are also interesting. Compared to high school graduates, those with some high school education have a 50% higher probability of leaving employment. Individuals with some college are 13% less likely to leave their jobs and those with college degrees decrease the hazard out of employment by nearly 22%. All of these results are significant at the one percent level. Clearly as the level of education increases, the probability of leaving employment decreases.

It is interesting to note that personal characteristics related to family stability are highly significant and serve to increase job stability. Married respondents had more stable employment. In fact, marriage reduces the hazard out of employment by over nine percent. Similarly, respondents with children also showed more stable employment. Children present in the household led to improved job stability for both male and female workers. For men, having children present in the household reduced the hazard out of employment by nearly 14%, while the corresponding reduction in the hazard out of employment for women was over 6%.

V. Conclusion

In this study I explore employment duration for young workers in their first real jobs after leaving school. I find there is a significant positive correlation between the duration of non-employment and the duration of the subsequent employment spell. Workers who spent more time in non-employment before accepting jobs tend to have longer employment durations. These hazard rates decline as the level of education rises therefore, workers with more education also tend to have longer employment durations.

Since employment duration is one of the primary factors that determine match quality, I interpret these results as evidence that workers who spend longer time in job search and those with higher levels of education make better quality matches.

I use duration modeling techniques to evaluate the hazard rate out of employment for these workers. Among the variables that contribute to match quality, job search activities decrease the hazard, and factors that negatively impact an individual's ability to conduct a job search during the non-employment spell, such as part time work, having children, and spending time out of the labor force, increase the hazard. Workers who conduct effective job searches make higher quality matches and workers distracted from their job searches make lower quality matches. Among the control variables, a worker with a stable family life has a lower hazard out of employment than a single worker does. Workers with children also have a lower probability of leaving employment that those who do not.

One limitation of this study is the lack of firm specific data in the sample. It would be interesting to isolate the firm effects on match quality. It has also been suggested that there is a significant difference in the NLSY data between unemployment and non-employment. This study pertains specifically to the non-employment interval before the first real job for young workers. Since these workers are unlikely to be eligible for unemployment benefits due to their youth, it is difficult to make inferences about the effect of the structure of UI benefits on match quality. It may be interesting to evaluate an older cohort to include individuals who received unemployment benefits to evaluate if there is an effect on match quality.

References

[1] Abowd, John M., Kramarz, Francis, and Margolis, David N. “High Wage Workers and High Wage Firms”, Econometrica (March 1999): Volume 67, No. 2, 251-333.

[2] Albrecht, J.W., Axell, B. “An Equilibrium Model of Search Employment”, Journal of Political Economy, (1984): Volume 92, 824-840

[3] Bowlus, Audra J. “Matching Workers and Jobs: Cyclical Fluctuations in Match Quality”, Journal of Labor Economics (1995): Volume 3, No. 2, 335-350

[4] Burdett, Kenneth and Mortensen, Dale T. “Wage Differentials, Employer Size, and Unemployment”, International Economic Review (May 1998): Volume 39, No. 2, 257-273.

[5] Cox, D.R. “Regression Models and Life-Tables,” Journal of the Royal Statistical Society (1972): Volume B, No. 34, 187-202

[6] Cox, D.R. “Partial Likelihood,” Biometrika (1975): Volume 62, 269-276

[7] Diebold, Francis X. and Neumark, David and Polsky, Daniel “Job Stability in the United States”, Journal of Labor Economics, (April 1997): Volume 15, Issue 2, 206-233.

[8] Eckstein, Zvi and Wolpin, Kenneth I. “Estimating a Market Equilibrium Search Model from Panel Data an Individuals”, Econometrica (July 1990): Volume 58, No. 4, 783-808.

[9] Jovanovic, Boyan “Job Matching and the Theory of Turnover”, Journal of Political Economy (1979): Volume 87, No. 5, pt.1.

[10] Meyer, Bruce D. “Unemployment Insurance and Unemployment Spells”, Econometrica (July 1990): Volume 58, No. 4, 757-782.

[11] Mortensen, Dale T. “Unemployment Insurance and Job Search Decisions”, Industrial and Labor Relations Review (1977): Volume 30, 505-517.

[12] Postel-Vinay, Fabien and Robin, Jean-Marc “Equilibrium Wage Dispersion with Worker and Firm Heterogeneity”, Econometrica (November 2002): Volume 70, No. 6, 2295-2350.

[13] Shimer, Robert “The Assignment of Workers to Jobs In an Economy with Coordination Frictions”, NBER (September 2001).

Appendix A – Estimation Results

Table A1: Complete Regression Results

```
         failure _d:  1 (meaning all fail)
   analysis time _t:  firstten

Cox regression -- Breslow method for ties

No. of subjects =         8590                     Number of obs   =      8590
No. of failures =         8332
Time at risk    =       602123
                                                   LR chi2(50)     =    479.04
Log likelihood  =   -36328.725                     Prob > chi2     =    0.0000
```

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
firstnonemp	.9992871	.0001083	-6.58	0.000	.9990749	.9994994
firstwage	.9999022	.0000191	-5.11	0.000	.9998647	.9999397
ageatfirst	1.005564	.0052613	1.06	0.289	.9953052	1.015929
firsthrs	.9954775	.0015683	-2.88	0.004	.9924085	.998556
firstun2	.8468879	.0367739	-3.83	0.000	.7777943	.9221194
firsturban	.9674161	.0300047	-1.07	0.285	.9103598	1.028048
gender	1.070517	.0773338	0.94	0.346	.9291866	1.233344
black	1.073976	.0328816	2.33	0.020	1.011425	1.140396
other	1.000764	.0592558	0.01	0.990	.8911107	1.123911
e1	1.503524	.0620362	9.88	0.000	1.386721	1.630164
e3	.8664068	.0312289	-3.98	0.000	.8073111	.9298282
e4	.7409415	.0254539	-8.73	0.000	.6926952	.7925482
regNC	1.037017	.0383719	0.98	0.326	.9644724	1.115019
regS	.9728378	.0346305	-0.77	0.439	.9072769	1.043136
regW	1.133244	.0455571	3.11	0.002	1.047381	1.226147
married	.9112823	.0336824	-2.51	0.012	.8476006	.9797486
firstchinh2	.9012612	.0525902	-1.78	0.075	.8038621	1.010462
firstsearch	.8973857	.0312955	-3.10	0.002	.8380969	.9608687
firstsi2	.7848252	.0286362	-6.64	0.000	.7306592	.8430068
firstchbf	1.066364	.0326112	2.04	0.042	.9110899	1.248101
firstpt	1.270874	.0629243	4.84	0.000	1.15334	1.400385
firstolfd	1.189668	.0324645	6.36	0.000	1.12771	1.255029
ulev2	.8533636	.0876064	-1.54	0.122	.69783	1.043563
ulev3	.8276557	.0856234	-1.83	0.067	.6757568	1.013699
ulev4	.8063038	.0878264	-1.98	0.048	.6513012	.9981954
ulev5	.69957	.0834765	-2.99	0.003	.553683	.883896
ulev6	.6330963	.0846086	-3.42	0.001	.4872065	.8226716
ind1	4.122607	.9946754	5.87	0.000	2.569211	6.615219
ind2	4.600902	1.311158	5.36	0.000	2.631897	8.042982
ind3	4.143563	.8473975	6.95	0.000	2.775206	6.186608
ind4	3.162524	.6218112	5.86	0.000	2.151155	4.649389
ind5	3.049878	.6480001	5.25	0.000	2.011082	4.625249
ind6	3.649099	.7109657	6.64	0.000	2.490825	5.345988
ind7	2.793776	.5732335	5.01	0.000	1.868702	4.176795
ind8	4.044722	.8172487	6.92	0.000	2.72208	6.01003
ind9	2.986899	.6274606	5.21	0.000	1.978827	4.508512
ind10	4.852836	1.11973	6.85	0.000	3.087392	7.627804
ind11	3.534515	.6925731	6.44	0.000	2.407357	5.189424
ind12	4.09996	.8453405	6.84	0.000	2.737011	6.141615
occ1	2.286229	.6030011	3.14	0.002	1.363365	3.83378
occ2	2.10294	.5653611	2.76	0.006	1.241615	3.561776
occ3	3.666742	.9864654	4.83	0.000	2.164119	6.212688
occ4	3.790391	.9910146	5.10	0.000	2.270563	6.327534

_t	Haz. Ratio	Std. Err.	z	P>|z|	[95% Conf. Interval]	
occ5	2.968202	.7779868	4.15	0.000	1.775773	4.961345
occ6	2.448719	1.903322	1.15	0.249	.5337323	11.23452
occ7	4.017994	1.046024	5.34	0.000	2.412192	6.692784
occ8	4.269434	1.127897	5.49	0.000	2.543902	7.165397
occ9	.6248528	.6573105	-0.45	0.655	.0794989	4.911278
occ10	3.233686	1.061917	3.57	0.000	1.698908	6.154968
occ11	4.555695	1.196558	5.77	0.000	2.722611	7.622961
occ12	6.18864	1.933746	5.83	0.000	3.354462	11.41741

Appendix B – Construction of the Sample

In order to construct the sample of first jobs, it was necessary to first identify the date the individual left school. The process for this identification is defined in the following text. Once these assignments were made, the remaining adjustments account for obviously erroneous responses and missing information. The total sample is described in Table B.1 and Figure B.1 below.

Summary of Sample

	Included	Excluded	Total
Education Status			
some high school	860	834	1694
high school	5061	1651	6712
College Educated			
some college	1269		
college	1400		
Total College Educated	2669	1611	4280
Total Sample	8590	4096	12686

Table B.1

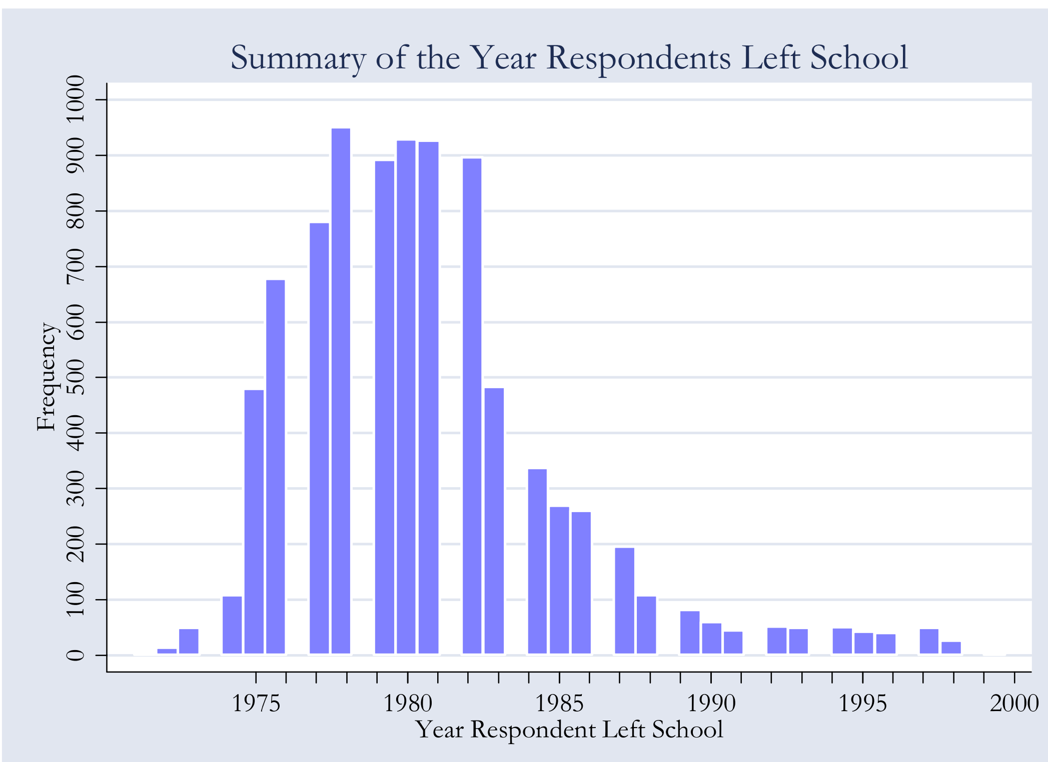


Figure B.1

Sample: Some High School

Sample Size: 860

Leave School Date

There are 55 individuals who answered yes when asked the question: "Has the respondent received any type of college degree or certificate?" This question was asked in 1979 and was followed up until 1984 with the question: "Has the respondent received any type of college degree or certificate since the last interview?" Of these, 5 reported that they never graduated from high school yet they have more than 12 years of schooling. I assume that these have some additional academic training past high school and have classified them as "some college". The remaining 50 respondents have less than 12 years of schooling and so are included with those with some high school. I am assuming that the source of their "yes" to this question is as a result of a certificate and not additional schooling.

Variable	Obs	Mean	Std. Dev.	Min	Max
maxschl	1064	9.067669	1.869981	0	16

The leave school date for these individuals is calculated from their answer to the question: "When was the month and year you were last enrolled in regular school?" Of the 1064 respondents in the sample, 849 gave a consistent answer to this question between 1979 and 1998. 1050 of the initial 1064 responded to this question, however; there are 201 who gave multiple answers to this question over time, and there are 14 respondents who did not provide any information on the date they left school other than their years of schooling record. These individuals report completing the same level of schooling over the entire sample period with one exception (ID 6747).

Since children must legally attend school until they are 16 in this country I have chosen to assume these individuals left school on their 16th birthday. Individual 6747 changed the amount of schooling reported in 1981. This individual was 21 at the time and never reported graduating; therefore I will not include this respondent in the sample.

The following discussion refers to the 201 respondents who gave multiple answers to the question: "When was the month and year you were last enrolled in regular school?" These respondents fall into the following categories:

104 report no change in years of schooling: I use the minimum reported leave school date.
59 report a difference in max schooling of one year. The new date the individual reports leaving school rarely corresponds with the year the individual reports completing an extra year of education. There are only 7 of these who report a transition between 11 and 12 years of schooling. All of them report this transition later in their careers with no corresponding claim of graduation. I will use the minimum reported leave school date as it more accurately corresponds to the years of schooling reported. There are 14 individuals who are still in school in 1979 and report every year that they have left school at the completion of that year. I have used their maximum reported leave school date as it corresponds to their maximum attained years of school. The remaining 24 individuals are summarized below:

Variable	Obs	Mean	Std. Dev.	Min	Max
ageschlmin	24	16.70833	3.057125	10	26
ageschlmax	24	31.20833	4.180484	25	39
maxschl	24	10.95833	1.680558	6	14

I have used their minimum reported leave school date as these individuals have long intervals between reporting additional levels of schooling and never report graduating. These are not full time students. I am assuming that they are receiving certifications and not additional school credits.

I have eliminated an additional 203 observations for clearly erroneous responses and for failure to respond to key questions.

Sample: High School Certification Only

Sample Size: 5061

Leave School Date

The leave school date for these individuals is the date recorded when the respondents were asked "When did you receive your high school diploma or GED?" There are 5545 records where the individual reported receiving a high school diploma or GED and also reported the date it was received. These individuals reported later that high school was the maximum level of schooling they received through 2002. There are an additional 72 respondents who reported receiving only a high school diploma or GED but failed to report the date it was received. 66 of this last category reported the date they received their certification when asked the question: "In what month and year did you complete your highest degree?" For these 66 individuals I have included the date reported in the response to the latter question.

There are discrepancies in the report date of these two questions:

Variable	Obs	Mean	Std. Dev.	Min	Max
degdatedif	5432	-24.20214	429.9634	-5024	5661

The above difference is measured in days: 69% are within 1 month of each other, 79% within 2 months and 81% within 3 months.

Note: 5545-5432=113 which is the number of the 5545 respondents who did not answer the second question. In other words, there was no response for the variable *highdegdate*.

I believe there is less measurement error in the responses to the first question: "When did you receive your high school diploma or GED?" This question was asked consistently in all years and the bulk of respondents (99%) received their certification before 1988, which was the first year that the second question: "In what month and year did you complete your highest degree?" was asked. Therefore, answers to the first question were given closer in time to the actual event and are less subject to memory lapses.

Variable	Obs	Mean	Std. Dev.	Min	Max
lsyear	5611	1979.484	2.760795	1972	1999

There are 135 respondents who reported a high school degree or GED when asked the second question, but either answered no degree, or failed to answer the first question. Of these, 129 reported the date that they received their degree.

The total sample size is then 5545 + 66 + 129 = 5740 and includes all respondents that reported earning a high school diploma or GED in either of the two questions and also reported when the degree was awarded.

This category is reduced by the 568 high school graduates who reported having more than 12 years of schooling and were also older than 18 when they left school. These 568 did not report having obtaining a degree, but they clearly attended school past high school graduation. I have moved these respondents to the "some college" category. I have not adjusted the 1087 respondents who reported more than 12 years of school but were 18 or younger. I have assumed that respondents who are under the age of 18 and do not report attending college have either counted additional non-academic certifications or have made a mistake when answering the question: "What is the highest grade completed?"

The final adjustment to this category comes from those respondents who made obviously erroneous responses or who failed to respond to all of the relevant questions. I have eliminated an additional 109 observations for these reasons.

Sample: College Degree (BA or BS)

Sample Size: 1400

Leave School Date

There are 1557 Respondents who reported receiving a college degree; either a BA or a BS when answering the question: "What is the highest degree obtained?" Of those, 1549 reported their graduation date in answer to the question: "What is the month and year you received your highest degree?"

Variable	Obs	Mean	Std. Dev.	Min	Max
lsyear	1549	1984.749	4.421093	1975	1998

I have eliminated an additional 109 observations for clearly erroneous responses and for failure to respond to key questions.

Sample: Some College

Sample Size: 1269

Leave School Date

There are 5 respondents who reported that they never graduated from high school yet they have more than 12 years of schooling and answered "yes" to the question: "Has the respondent received any type of college degree or certificate?" I assume that these have some additional academic training past high school and have classified them as "some college".

These individuals are summarized in the table at the right. I estimated the leave school date by calculating the year that they attained their maximum years of schooling. Unfortunately, this doesn't seem to be a reasonable estimation technique as there are many years of work experience between their completions of additional years of schooling. Therefore, I have excluded them from the sample.

ID	Leave School	Birth Year	Age
1145	1983	1961	22
1847	1988	1960	28
7762	1982	1958	24
8001	1992	1957	35
8057	1998	1959	39

There are 804 individuals who reported receiving an associate degree; 793 of whom also reported the date that they received that degree. Therefore, the sample size is 793.

There are 568 high school graduates who reported having more than 12 years of schooling and were also older than 18 when they left school. These 568 did not report having obtaining a degree, but they clearly attended school past high school graduation. I have moved these respondents to this category.

```
    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+-------------------------------------------------------
      lsyear |       793    1985.685    5.617566       1974       1998
```

I have eliminated an additional 92 observations for clearly erroneous responses and for failure to respond to key questions.