LONG-TERM CONSEQUENCES OF YOUTH UNEMPLOYMENT

Jack Moore

Stanford University
May 2003

ABSTRACT

LONG-TERM CONSEQUENCES OF YOUTH UNEMPLOYMENT

Jack Moore

Men in the United States who spend time out of work during their youth tend to be employed less and earn lower wages in their adult years. Why? Is this because the state of being out of work as a youth causes workers to earn less and work fewer weeks later in life? Or is it because individuals who work fewer weeks during youth have personal traits that lower their probability of employment and their wages during both youth and adulthood? Researchers refer to the first possible explanation as state dependence and call the second explanation heterogeneity. Using successive observations of the same people over time from the National Longitudinal Survey of Youth, this paper identifies a causal relationship (state dependence) between early employment experience and employment and wages ten years later among high school graduates and individuals with higher measured ability but not among high school dropouts or those with low measured ability. Previous research on this topic has not examined whether the impact of state dependence varies with schooling and measured ability. Among those who are affected by state dependence, those who as youths experience especially long periods of unemployment are particularly prone to long-term effects on future wages and employment.

LONG-TERM CONSEQUENCES OF YOUTH UNEMPLOYMENT

Jack Moore [*]

I. Introduction

As the national unemployment rate in the United States reaches 6 percent in the month of December 2002, and one million people drop out of the labor force during the past six months, some declare that "the economy has fallen into the worst hiring slump in almost twenty years."[1] As in most economic downturns, the group hit hardest by this hiring slump is young people with low levels of education. This group's job prospects typically have the highest sensitivity to the business cycle. A recent study by economist Andrew Sum, director of the Center for Labor Market Studies at Northeastern University in Boston identified a 12 percent increase in "joblessness among out-of-school youths between 16 and 24 since the year 2000,"[2] which brings the total number of non-working, out-of-school youths in that age range to 5.5 million as of January 2003.

These figures, according to Sum, the study's lead author, need "clear recognition by economic policy makers" so that youth joblessness can be addressed and its remedy given high policy priority. The special attention to youth joblessness for which the study calls arises from the fact that there exists some special problem created by such high

[*] This paper has benefited from numerous conversations with John Pencavel. His helpful comments on previous drafts and his patience in discussing the issues studied here have contributed enormously both to this work and to my own research skills.

[1] Leonhardt, New York Times, Feb 6,2003.

[2] Herbert, New York Times, Feb. 6, 2003.

current non-employment, and that simply waiting for the business cycle to pick up again will not fully solve the problem. As the study's title, "Left Behind in the Labor Market," suggests, Sum believes that youths' being "left out of the ranks of the employed" creates a "substantial risk of [their] being permanently left behind."[3] It is this risk of permanence or at least persistence of joblessness and low wages that causes Sum to think youth joblessness warrants attention above and beyond policymakers' efforts to improve the economy in general.

While the Northeastern University study is rich with data on current unemployment and other labor market trends, the link between the current situation and future outcomes is conjectural alone. It presumes that many key decisions individuals make during the early years of adulthood affect future outcomes.[4] It is often hypothesized that people who experience more unemployment as teenagers tend to experience more unemployment when they are older. Or, similarly, non-employment (the union of unemployment and time out of the labor force) during an individual's teen years will result in lower wages and higher levels of unemployment further into that person's career. The issue here is: does this persistence in labor market experiences indicate a *causal* relationship from early experiences to outcomes later in life?

The theory of state dependence, also known as structural dependence, argues that this lasting effect of youth unemployment is a result of the experience of the individual while unemployed. That is, something inherent in the state of being unemployed lowers an

---

[3] Phelan (2003).
[4] Sum, et al. (2002), pg. 2.

individual's future potential earnings and raises the amount of time he will be unemployed in the future. Various mechanisms are identified as creating this effect. Unemployment could create a psychological discouragement of an individual while he is unemployed that affects his future attitudes towards work. For instance, an employee who puts forth great effort into doing a job that he eventually loses may feel less secure in his relationship with his next employer. Worried about the stability of this job, he may be hesitant to work hard or to invest in firm-specific human capital. Secondly, a deterioration (or lack of accumulation) of job skills during the time of unemployment could also lead to lower wages and higher unemployment in the future. If work experience makes a worker more productive, then time spent not working lowers the individual's productivity compared to what it would be had he been employed during that time. This lost opportunity to accumulate skills lowers future wages.

Finally, regardless of a worker's actual condition, time not employed can decrease potential employers' opinions of the worker's skills and attitude and make the employers less interested in hiring that worker. As David Ellwood describes this mechanism, employers, who have a limited amount of time and information with which to select employees, may rely on a potential employee's job history as an indicator of that worker's job skills and attitude.[5] Such employers often choose not to hire workers who have spent a long period of time not working after leaving school. As a result, it is possible in weak job markets that, if employers "are slow to adjust their expectations for

[5] Ellwood (1982), pg. 351.

work experience from young applicants, cohorts entering a weak labor market will suffer." [6]

An alternative hypothesis accounting for this persistence is that there are permanent differences among people that make some more likely to experience a particular event at different stages of their life cycle. This competing hypothesis, labeled "heterogeneity" by many researchers, postulates that workers have characteristics, many of which are unobserved by the economist, that make the individual less likely to get a job at a young age and have the same effect later in life. Even among those with the same level of education, age, etc., different individuals will be more or less attractive employees to hire both in their teen years and later in life. A worker's punctuality, for instance, could affect employers' desires to hire or to retain that worker. If a person has a habit of arriving late throughout his life, then both early and later in his career, this person will be more likely to be unemployed than a different worker who is identical to the first except that he is meticulous about arrival times and deadlines. Observation of these two workers over a long period of time would reveal persistence in the employment level of each of these two workers. The cause of this persistence, however, would be their heterogeneous level of punctuality rather than the fact that one worker had more employment early in his career. According to this line of analysis, the pattern of labor market experiences we observe over an individual's lifetime are not causally related to previous experience but rather, experiences at all ages result from the permanent individual differences that may

[6] Ibid., pg. 351.

be unobserved to the researcher. This unobserved heterogeneity simply makes some individuals "unusually prone" to being not employed.[7]

Empirically distinguishing between these two hypotheses is not simple. David Ellwood notes that "the fundamental problem is separating differences in employment and wages which are causally related to early unemployment, from the differences due to unobserved personal characteristics correlated with early unemployment."[8] The largest problem is the inability to observe these personal characteristics. If the researcher could accurately observe all the characteristics that affect individuals' likelihood of being unemployed, then those variables could be controlled for in a regression, and the degree of state dependence could be accurately computed. However, since many of these factors influencing each individual's work experiences are unobservable, other statistical tools must be used to disentangle heterogeneity from state dependence.

This study will make use of over 12 years of panel data collected by the National Longitudinal Survey of Youth 1979. This wave of the survey began tracking men and women of eight birth cohorts (1957-1964) beginning in 1979 (when the '57 cohort was 21-22 years old and the '64 cohort was 14-15. The survey first interviewed individuals once per year until 1994, after which individuals still participating have been interviewed every other year through the present. The survey began with over 12,000 youths who were selected to represent an ethnic and geographic cross section of the United States population. Blacks and Hispanics were oversampled in the survey but the NLSY

[7] Gregg (2001), pg. 627.
[8] Ellwood (1982), pg. 349.

provides sample weights that compensate for this oversampling. It is interesting that, during the survey, many of these youth were looking for jobs in the economic recession of the early eighties. As the U.S. currently has the "worst job market in nearly two decades," seeing what became of the youth leaving school during the worst job market since the 1930s may have extra significance at this time. However, it also should be noted that that recession was followed by one of the longest periods of growth in the nation's history, and it would be courageous to assume similar future job markets to be on the horizon for current youth.

II. Conceptual Framework

Distinguishing state dependence from heterogeneity is of tremendous importance for policy. Identifying the relative importance of state dependence and heterogeneity has tremendous value because of the disparate implications for policy that are suggested by each theory. State dependence proposes that the early labor market experiences of an individual causally affects his/her welfare and income in more than just the current period, and individuals who experience low employment in their teen years really do risk being left behind their peers who have jobs at the same age. In this case, if there are policy programs that effectively increase the employment of youth, their benefits extend for that individual beyond the current period. That is, if persistence in low employment and wages resulted from the state of being unemployed, then lowering teen unemployment today would raise the future wages and employment opportunities for those teens in the future.

Alternatively, if the employment patterns we see over the life cycle are largely the result of permanent individual differences among workers that exist at both young and older ages, then the current hiring slump only affects a worker's employment prospects currently, and attempts by policy makers to raise youth employment are not needed to avert enduring damage or "scarring" of currently unemployed individuals. In this case, individuals who are unemployed now may go on to have less employment later in life but that fact is not the consequence of current low employment.

This paper seeks to identify the relative importance of heterogeneity and state dependence in explaining persistence in joblessness and low wages using panel data from the National Longitudinal Survey of Youth 1979. This analysis will begin with a simple description of the labor market experiences of the teenagers and young adults in the survey who never attend college. We will attempt to measure the persistence in joblessness and low wages among this group of individuals by examining probability trees that track the autocorrelation of employment for these individuals during their first few years of work. We document the fraction of these young men and women who are employed in every year after leaving school and the fraction who never get a job for many years. We will also analyze the changes in real wages of subsections of this group to measure the correlation between wages at a young age and wages later in the life cycle. Finally, we will use econometric techniques of individual fixed effects and first differencing to determine how much of this correlation is due to some causal relationship

between early and later labor market experiences as opposed to permanent individual differences among workers.

Heterogeneity and State Dependence

Heckman and Borjas, in an article titled "Does Unemployment Cause Future Unemployment? Definitions, Questions, and Answers from a Continuous Time Model of Heterogeneity and State Dependence," argue that state dependence is "rooted in economic theory" while individual differences is an explanation "based solely on statistical considerations." In defining state dependence they write that, regardless of the mechanism used to explain it, "prior unemployment experience has a genuine behavioural effect in the sense that an otherwise identical individual who did not experience unemployment would" be more likely to earn lower wages and be unemployed again than would an individual with all the same personal characteristics but a more favorable work history.[9]

The authors go on to identify four types of state dependence of employment, each associated with different mechanisms for producing this effect. The first type is Markovian state dependence, a form of short-term autocorrelation. In a short time interval, an unemployed worker is more likely to remain unemployed than an employed worker is to lose his or her job. For example, it is obvious that a person who is unemployed on one particular day is more likely to be unemployed two days later than a person who

[9] Heckman and Borjas (1980), pg 247-48.

currently has a job is to get fired in that same two day period, and hence the probability of a person being unemployed two days from the current date is dependent on that individual's current employment status. This effect attenuates over time, but still could influence one's likelihood of employment for a year or more into the future. Many self-evident reasons explain this type of dependence, including the fact that firms incur some cost to replacing an employee, and will attempt to avoid incurring that cost if the worker's productivity is adequate. Due to its shorter-term nature, this is the least interesting type of state dependence from a policy point of view.

Second, "occurrence dependence," identifies the *number* of prior spells of unemployment as a key factor in determining future unemployment. Occurrence dependence may be a result of employers' methods of selecting workers based on their work history, as workers who have been separated from their jobs many times may have suspect work attitudes and behavior.

Third, "duration dependence" links the probability of finding employment for an individual without a job to the length of his/her current unemployment spell. This type of dependence may be the result of deterioration of job skills while unemployed, or of psychological changes in a worker's attitude when he is not working.

Finally, "lagged duration dependence" sets the probability of becoming or remaining unemployed as dependent on the *lengths* of *many* previous unemployment *spells*. This type of dependence takes into account both occurrence dependence and the duration of

those past occurrences of unemployment, and could more generally be stated that more unemployment will probably lead to more unemployment in the future. A worker with high unemployment in his past has had fewer opportunities over that time to acquire job skills than a person who has spent more time working during that time. The missed opportunities to develop these job skills could make the worker less valuable to a firm. Regardless of the actual level of job skills a worker has acquired, employers may use amount of past employment as a proxy for these skills when making hiring decisions. In this case, a worker with more unemployment is less likely to be hired for future jobs. We are focused on exploring whether experiencing less unemployment as a youth makes a person less likely to be unemployed later in life. Therefore, in this paper, we are most concerned with the effects of lagged duration dependence.

Alternatively, Heckman and Borjas characterize heterogeneity (when used as a way to explain persistence of unemployment) as differences across individuals "in certain unmeasured variables that influence the probability of experiencing unemployment but that are not influenced by the experience of unemployment." Furthermore, if these variables are not controlled for, correlation between the variables over time may cause previous unemployment to "appear to be a determinant of future unemployment solely because it is a proxy variable for temporally correlated unobservables," producing a "conditional relationship between future and past unemployment due solely to uncontrolled heterogeneity."

III. Empirical Methodology

Probability Tree

Our data analysis consists first of probability trees of the employment experiences of a group of youths. That is, we follow the youths in our sample over a number of years distinguishing those who are employed during each year for at least one week, and for 40 or more weeks. This way of describing the data allows us to identify these individuals who end up employed in all years and those who remain not employed in all years. Both the relative size and the characteristics of the individuals in these groups are important. If these two groups represent a large portion of the cohort, then persistence in experience is a relevant characterization. Also, if many individuals that compose either group have characteristics that make them easily identifiable, then we might be able to identify those characteristics as possible causes of heterogeneity.

Multiple Regression

Secondly, we examine the wage and employment effects of these early labor market experiences on the individuals later in life. The greatest challenge to performing this analysis is that many characteristics inherent in an individual may be correlated with both the early labor market experience and the wage and employment probability of that individual later in life. Hence, a simple regression of, for instance, wages at age 30 on weeks employed at age 20, does not necessarily indicate a ***causal*** relationship.

Two primary methods exist for solving this problem. The first is multiple regression: ***if*** we had observations of all ***enduring*** characteristics (in addition to employment history) of an individual that cause him to be employed or unemployed early in working life and also affect his wages and employment later in life (call this set of enduring characteristics *Xi* for individual *i*), then we could estimate a multiple regression similar to the one mentioned above, but also including all the variables in *Xi* as control variables. Because not all the variables in *Xi* are measured by the NLSY (or even observable), this procedure is not fully effective. Despite this shortcoming, we examine the regressions using control variables for family background, education, race, and an indicator of ability to gain some understanding of relationships within the data.

In the following equations, *i* denotes an individual, *t* a calendar year, and *a* is age. A birth cohort is, therefore, defined as *t-a*, so if *t* and *a* are both specified, then a particular cohort is also identified. These variables are defined as follows:

$a \in A = \{19, 20, 29, 30\}$.

$t \in T = \{1979, 1980, ..., 1994\}$.

E(i,a,t) = weeks employed of individual *i* in year t when the individual is aged *a*.

ln w(i,a,t) = natural logarithm of real hourly earnings of individual *i* in year *t*.

N(i,a,t) = dummy variable for zero weeks of work, =1 if *E(i,a,t)* =0, =0 if *E(i,a,t)* > 0.

Xi= a vector of permanent attributes of individual *i* observed to the researcher. Dummy variables for different birth cohorts of workers in our sample are included in *Xi*

Zi= permanent attributes unobserved by the researcher.

These definitions allow us to examine key relationships between employment probability and real wages at early and later life stages for the individuals in our sample.

Employment and Unemployment

We seek to measure how much more an individual would be expected to work later in life than an otherwise identical individual who worked fewer weeks as a youth. This relationship can be measured by regressing the weeks worked during a later year in a worker's career on the weeks worked at an age during the worker's youth and on the permanent attributes of the worker. This equation is written as:

(1) $E(i,t,a) = a_0 + a_1 E(i, t-k, a-k) + a_2 X(i) + e_1(i,t,a)$

where $e_1(i,t,a)$ is the error term of this regression. α_1 measures the partial association between the weeks individual *i* was employed *k* years ago and the weeks that this individual worked this year.

Equation (1) treats the effect on $E(i,t,a)$ of a change in $E(i, t-k, a-k)$ as independent of whether the individual was completely out of work during year $t-k$. That is, the effect of a change in $E(i, t-k, a-k)$ is restricted in equation (1) to being the same whether the individual is at zero weeks worked in year $t-k$ (i.e. was out of work for the entire year) or at 40 weeks worked. In fact, the consequences of getting a job and raising $E(i,t-k,a-k)$ from zero to a few weeks may be substantially greater than raising $E(i, t-k, a-k)$ from forty to forty-five weeks. The following equation allows for the additional impact of not working at all during year $t-k$:

(2) $E(i,t,a) = a'_0 + a'_1 E(i, t-k, a-k) + ?'_1 N(i, t-k, a-k) + a'_2 X(i) + e_2(i,t,a)$

where $e_2(i,t,a)$ is the error term of this regression. α'_1 measures the partial association between the weeks individual *i* was employed *k* years ago and the weeks that this

individual worked this year. $?'_1$ measures the partial association between whether individual *i* was not employed at all *k* years ago and the weeks that this individual worked this year.

Real Wages

Labor market experience in one's youth may also significantly affect a worker's real wages later in life. This impact is estimated by the following equation:

(3) $ln\ w(i,t,a) = b_0 + b_1 E(i, t-k, a-k) + b_2 X(i) + v_1(i,t,a)$

where $v_1(i,t,a)$ is the error term of the regression. b_1 measures the partial association between the weeks individual *i* was employed *k* years ago and the real hourly wages he earned this year.

As defined in equation (3), the effect on $ln\ w(i,t,a)$ of given change in $E(i,t-k,a-k)$ is the same regardless of whether the individual is without work for the entire year or worked for 40 weeks at age $a-k$. However, the impact on $ln\ w(i,t,a)$ of obtaining a job and raising $E(i,t-k, a-k)$ from zero weeks to one week may be significantly greater than the effect of increasing $E(i, t-k, a-k)$ from forty to forty-five weeks. The following equation allows for an additional impact on wages in year *t* of not working at all in year $t-k$:

(4) $ln\ w(i,t,a) = b'_0 + b'_1 E(i, t-k, a-k) + ?'_2 N(i, t-k, a-k) + b'_2 X(i) + v_2(i,t,a)$

where $v_2(i,t,a)$ is the error term of the regression. b'_1 measures the partial association between the weeks individual *i* was employed *k* years ago and the real hourly wages he earned this year. $?'_2$ measures the partial association between whether individual *i* was not employed at all *k* years ago and the real wages that this individual worked this year.

Lastly, even if a youth finds employment, the type of job and its wages may influence his wage rates in future jobs later in life. The following equation attempts to capture this relationship between wages received early and later in life:

(5) $ln\ w(i,t,a) = ?_0 + ?_1 ln\ w(i, t-k, a-k) + ?_2 X(i) + t_1\ (i,t,a)$

where $t_1\ (i,t,a)$ is the error term. $?_1$ measures the partial association between the individual *i*'s real hourly wages *k* years ago and the real hourly wages he earned this year.

Interaction Effects

It is also possible that, in some cases, the relationship between current employment experience and future employment outcomes have stronger effects on some types of individuals than on individuals with different personal characteristics. Individuals with different educational levels, for instance, may experience different amounts of employment persistence. The following equation allows for the interaction between employment persistence and individual characteristics.

(6) $E(i,t,a) = ?_0 + ?_1 E(i, t-k, a-k) + ?_2 [E(i, t-k, a-k)*X(i)] + ?_3 X(i) + h_1(i,t,a)$

For this equation, $h_1(i,t,a)$ is the error term. $?_1$ measures the partial association between the weeks individual *i* was employed *k* years ago and the weeks that this individual worked this year.

Similarly, the amount of impact that current employment has on future wages could depend on individual characteristics. The following equation allows for this interaction:

(7) $ln\ w(i,t,a) = ?_0 + ?_1 E(i, t-k, a-k) + ?_2 [E(i, t-k, a-k)*X(i)] + ?_3 X(i) + g_1(i,t,a)$

In this regression, g_1 is the error term, and $?_1$ measures the partial association between the weeks individual *i* was employed *k* years ago and the real hourly wages he earned this year.

Unobserved Heterogeneity

As noted earlier, unobserved heterogeneity among workers that is not fully captured by *X(i)* is also likely to affect employment levels and wages. Because they are unobserved, in the previous equations, these variables *(Z(i))* are assumed to be part of the error term. To account for the impact of this unobserved heterogeneity, we rewrite the equations as follows.

Employment Equations:

(8) $E(i,t,a) = g_0 + g_1E(i, t-k, a-k) + g_2 X(i)+ Z(i) + u_1(i,t,a)$

where $Z(i) + u_1(i,t,a) = e_1(i,t,a)$.

(9) $E(i,t,a) = g'_0 + g'_1E(i, t-k, a-k) + f'_1N(i, t-k, a-k) + g'_2 X(i)+ Z(i) + u_2(i,t,a)$

where $Z(i) + u_2(i,t,a) = e_2(i,t,a)$.

Wage Equations:

(10) $\ln w(i,t,a) = d_0 + d_1E(i, t-k, a-k) + d_2 X(i)+ Z(i) + h_1(i,t,a)$

where $Z(i) + h_1(i,t,a) = v_1(i,t,a)$.

(11) $\ln w(i,t,a) = d'_0 + d'_1E(i, t-k, a-k) + f'_2N(i, t-k, a-k) + d'_2 X(i) + Z(i) + h_2(i,t,a)$

where $Z(i) + h_2(i,t,a) = v_2(i,t,a)$.

(12) $\ln w(i,t,a) = ?_0 + ?_1 \ln w(i, t-k, a-k) + ?_2X(i)+ Z(i) + h_3(i,t,a)$

where $Z(i) + h_3(i,t,a) = t_1(i,t,a)$.

Interaction Equations:

(13) $E(i,t,a) = g''_0 + g''_1 E(i, t-k, a-k) + g''_2[E(i, t-k, a-k)*X(i)] + g''_3 X(i) + Z(i) + u_3(i,t,a)$

where $Z(i) + u_3(i,t,a) = h_1(i,t,a)$.

(14) $\ln w(i,t,a) = d''_0 + d''_1 E(i, t-k, a-k) + d''_2[E(i, t-k, a-k)*X(i)] + d''_3 X(i) + Z(i) + h_4(i,t,a)$

where $Z(i) + h_4(i,t,a) = g_1(i,t,a)$.

Now, $u_1(i,t,a)$, $u_2(i,t,a)$, $u_3(i,t,a)$, $h_2(i,t,a)$, $h_3(i,t,a)$, and $h_4(i,t,a)$ are the error terms and are properly distinguished from unobserved heterogeneity.

First Differences

Because the $Z(i)$ variables are not observed, the researcher cannot include them in empirical analysis. However, because they are enduring characteristics of the individual, they do not affect each individual's changes in wages and weeks employed. The following variables are used to represent the year-to-year difference in weeks worked and in real wages in adjacent years:

$DE(i,t,a) = E(i,t,a) - E(i,t-1,a-1)$

$DN(i,t,a) = N(i,t,a) - N(i,t-1,a-1)$

and

$D\ln w(i,t,a) = \ln w(i,t,a) - \ln w(i,t-1,a-1)$

Taking the first difference of of equation **(7)** results in the following:

(15) $DE(i,t,a) = g_1 DE(i, t-k, a-k) + u'_1(i,t,a)$

where $u'_1(i,t,a) = u_1(i,t,a) - u_1(i,t-1,a-1)$

The first difference of eqation **(8)** is written as:

(16) $DE(i,t,a) = g'_1 DE(i, t-k, a-k) + f'_1 DN(i, t-k, a-k) + u'_2(i,t,a)$

where $u'_2(i,t,a) = u_2(i,t,a) - u_2(i,t-1,a-1)$

Similarly, the first differences of the three wage equations are:

(17) $Dln\ w(i,t,a) = d_1 DE(i, t-k, a-k) + m_1(i,t,a)$

where $m_1(i,t,a) = h_1(i,t,a) - h_1(i,t-1,a-1)$

(18) $Dln\ w(i,t,a) = d'_1 DE(i, t-k, a-k) + f'_2 DN(i, t-k, a-k) + m_2(i,t,a)$

where $m_2(i,t,a) = h_2(i,t,a) - h_2(i,t-1,a-1)$

(19) $Dln\ w(i,t,a) = ?_1 Dln\ w(i, t-k, a-k) + m_3(i,t,a)$

where $m_3(i,t,a) = h_3(i,t,a) - h_3(i,t-1,a-1)$

Finally the first differences of equations **(13)** and **(14)** are written, respectively, as:

(20) $DE(i,t,a) = + g''_1 DE(i, t-k, a-k) + g''_2 [DE(i, t-k, a-k)*X(i)] + u'_3(i,t,a)$

where $u'_3(i,t,a) = u_3(i,t,a) - u_3(i,t-1,a-1)$

(21) $Dln\ w(i,t,a) = d''_1 DE(i, t-k, a-k) + d''_2 [DE(i, t-k, a-k)*X(i)] + m_4(i,t,a)$

where $m_4(i,t,a) = h_4(i,t,a) - h_4(i,t-1,a-1)$.

These first difference equations contain neither the observable nor the unobservable permanent attributes of the individuals (*X(i)* and *Z(i)*) because these attributes are time-invariant. The coefficients $g_1, g'_1, f'_1, d_1, d'_1, f'_2, ?_1, g''_1, g''_2, d''_1$, *and* d''_2 on the regressor in each of these equations measure the impact of labor market experiences as a youth on labor market outcomes about a decade later. If the attributes characterized by *Z(i)* do not affect our dependent variables, then the impact of unobserved heterogeneity on our level equations is negligible. In that case, our estimates of the following parameters should be equal:

g_1 and a_1, d_1 and b_1, and d_1' and b'_1, etc.

A large disparity in the estimate from the first difference equations and that of the level equations first estimated would imply that unobserved heterogeneity represented by *Z(i)* was an important cause of persistence in unemployment and wages. Alternatively, if, after first differencing, we see a strong influence of the lagged employment or wage term (*DE(i, t-k, a-k)* or *Dln w(i, t-k, a-k)*) then state dependence may be said to exist in a person's work experience.

IV. Data and Selection

The National Longitudinal Study of Youth 1979 (NLSY79) provides a highly useful set of panel data for examining the long-term consequences of not being employed.[10] In 1979, the NLSY79, one of six waves of longitudinal surveys conducted by the Bureau of Labor Statistics, first interviewed a national sample of young men and women who were born between January 1, 1957 and December 31, 1964.[11] Of the youths first interviewed, the BLS selected 6,111 individuals to represent a cross-section of civilian youths living in the United States in 1979. An addition group of 5,295 youths, who were black, Hispanic, or economically disadvantaged non-black/non-Hispanic was also interviewed so that these parts of the national population would be over-sampled in the survey. Lastly, 1,280 youths serving in the military and aged between 17 and 21 at the beginning of 1979 were included in the survey.

---

[10] Much of the following information is drawn from the NLSY Handbook (2002), published by the BLS. http://www.bls.gov/nls/handbook/nlshndbk.htm.

[11] They were aged between 14 and 21 on December 31, 1979.

The BLS interviewed this same set of youths (with the exception of those who chose to quit the survey) annually through 1994. Since then, those who remain with the survey continue to be interviewed biennially through the present. After the 1990 interview, the economically disadvantaged non-black/non-Hispanic members of the supplementary sample were dropped. Similarly, in 1985 the NLSY79 stopped interviewing all but 200 members of the military sample.

As with most longitudinal surveys, the NLSY79 has experienced attrition in observations. Through 1994, 5,457 of the 6,111 original members of the cross-sectional sample were still being interviewed. Of the 2,172 over-sampled blacks and 1,480 over-sampled Hispanics in the supplementary sample, 1,960 and 1,296, respectively, remained in the survey through 1994. 178 of the 201 members of the military sample also remained at that time. Though these retention rates, over 85% in all cases, appear high given the length of time spanned by the interviews, this figure may be misleading. Not all persons who are interviewed in later years have been interviewed in every preceding year of the survey, and even in years that an individual is interviewed, questions are sometimes skipped. Therefore, the 85% retention statistic should be seen as the maximum possible number of observations available for analysis requiring information through 1994. Moreover, even with an 85% retention rate, it would be presumptuous to assume that a random 15% of the individuals in the survey dropped out. To the extent that certain characteristics made a given type of person more likely to drop out of the survey, fewer

people of that type will be observed in later years of the survey. In other words, sample attrition is non-random.

To correct for the oversampling of blacks and Hispanics, the NLSY79 contains a set of sample weights. The variables for an individual can be multiplied by his sample weight to attempt to "provide the researcher with an estimate of how many individuals in the U.S. each respondent's answers represents."[12] The National Opinion Research Center (NORC) at the University of Chicago constructs sample weights in each survey year so that the weighted NLSY sample will be representative of the U.S. population as a whole.[13] These sample weights are also adjusted in each year to account for non-response interviewees. That is, the NORC constructs its sample weights to correct for misrepresentation that would occur in a given year as a result of certain groups of people not being interviewed.

Often when a researcher uses the longitudinal data of the NLSY to analyze changes in variables over time, he selects individuals only if they completed surveys in multiple years. Given that many respondents leave for some time and then re-enter the sample in later years, the ability of any single year's sample weights to make the selected group represent the national population is questionable. Such difficulties in applying the sample weights have been noticed by other economists, most notably, MaCurdy, Mroz, and Gritz, who emphasize that "the weights currently supplied with the NLSY are inapplicable for use with most sample compositions that are analyzed in the literature."[14] If all

[12] NLSY79 User's Guide (1998), pg. 36.
[13] NLS Handbook (2002), pg. 32.
[14] MaCurdy, Mroz, and Gritz. (1998), pg. 434-35.

respondents who left the survey left permanently, then an analysis that needed observations on individuals for four different years could select those individuals who stayed in the survey through the latest year that was needed for analysis. In this case, the researcher could use the sample weight from the latest year to weight correctly his sample. However, since a number of respondents drop out and then return to the survey, use of the survey weights of the latest year would not be taking into account that many of the respondents in the group used to create the most recent weights may have been missing from previous years that the researcher needs to observe.

Because we are analyzing individuals' employment experiences over time, we must select only individuals we observe in multiple years. Constructing our own sample weights to correct for the effects of attrition on our selected sample is beyond the scope of this project. Thus, we instead use the sample weights for the original 1979 sample so that we can compensate for the oversampling issue and use members of both the cross-sectional and supplementary black and Hispanic samples in our analysis.

A large amount of information about each individual's family and personal background is observed in the first years of the survey. For instance, the data include the individual's region of birth, mother and father's years of schooling, religion as a youth, and detailed race and ethnic background. Additionally, the data include retrospective descriptions of the individual's labor market experience. Each survey year has variables for the number of weeks in the previous calendar year in which the individual was employed,

unemployed and out of the labor force. Annual interviews ask the individual to calculate the wages per hour that he earned in his current or most recently held job.

The NORC also administered an *Armed Services Vocational Aptitude Battery (ASVAB)* test to most of the NLSY79 participants during the summer and fall of 1980. Approximately 94% of the NLSY79 participants completed the test. From the results of this test, the Center for Human Resources Research (CHRR) at Ohio State University, which maintains the NLSY data set and creates many important variables from the raw interview data, has constructed an approximate and unofficial *Armed Forces Qualifying Test (AFQT)* score for each individual who was tested. The AFQT at school leaving age, is designed to measure one's ability to be trained, and thus may be useful in determining one's ability to become proficient at a job, regardless of whether the person is a civilian or a military serviceman.[15]

Sample Selection for Analysis in this Paper

During the late teens and the twenties of the individuals in this sample, the work experiences of men and women are likely to be significantly different even when many of their family background characteristics and early work experiences are very similar. Due to the fact that so many more women than men choose during this time to work in the home and, therefore, do not participate in the labor force, many factors which influence

[15] Note, however, that the NLSY79 User's Guide urges caution in using the AFQT score as members of the 1964 and 1963 birth cohorts took the test when they were under age 17, which is the accepted year that one can be given the test. See NLSY79 User's Guide (1998), pg. 95.

work decisions, such as non-wage income and educational levels, tend to affect men and women differently. As a result, analysis of long-term labor experience that treats men and women separately is preferable, and is commonly employed by economists addressing similar employment topics, such as David Ellwood[16] and Paul Gregg.[17] Consequently, we will focus only on men.

In seeking to measure and analyze the long-term impact of youth unemployment, we need to be able to observe the labor market experience of individuals both near the time that they enter the labor market and at a time later in their career. Throughout most of their teen years, many young men in the NLSY79 attended school full-time, and many of them were not looking for a job or reporting the number of weeks they worked during the year. Though we wish to observe one's labor market experience early in the individual's working life, we also seek to have a large number of individuals in the group to analyze. Thus, we select individuals who reported the number of weeks they worked at ages 19 and 20. Only individuals born after June of 1960 are first interviewed at or before age 19 by the NLSY79. Thus, the sample's earlier birth cohorts must be excluded from our analysis. This restriction leaves four full annual birth cohorts (individuals born between July 1960 and June 1961, and the same months of 1961-62, 1962-63, and 1963-64) and one birth cohort with half as many individuals (born from July 1-December 31,1964). Of the 6,043 men in the cross-sectional sample and the Black and Hispanic supplementary sample of the NLSY79, 3,461 are in one of these five birth cohorts. This selection should

[16] Ellwood (1982).
[17] Gregg (2001).

not diminish the quality of the sample because the NLSY was not designed to have a higher percentage of any particular type of individual in any certain age group.

We also must be able to observe the men's experiences at a point after the general time of transition from school and adolescence into working adulthood. Again, there is a tradeoff between using observations of individuals at a much later stage in life, and selecting a sample that minimizes the number of individuals lost due to attrition over time in NLSY79. Therefore, we have selected individuals for whom we have employment information when they are aged 29 and 30.

Of the 3,461 men in the last five birth cohorts of the NLSY's cross-sectional sample and the Black and Hispanic supplementary sample of the NLSY79, 2,627 men give information on the number of weeks worked at ages 19, 20, 29 and 30. These individuals have a variety of educational levels. Because the individuals who did not take the AFQT appear to be randomly selected with respect to the characteristics that we do observe for these individuals, we omit the 61 who do not have an AFQT score. This leaves a sample of 2,389 men. Roughly 15 percent are high school dropouts, 45 percent graduated from high school but did not complete a single year of college, and 40 percent have one or more years of college education.[18]

The heterogeneous educational levels of these individuals significantly affect the importance of their early labor market experiences in obtaining future jobs. High school

[18] This means there are 362 drop-outs, 1,082 high school graduates, and 945 men who completed one year of college or more.

dropouts, graduates, and those with some college education may be looking for different kinds of jobs when they begin their working career. Additionally, any employer making hiring decisions will look at a job candidate's past performance in both school and previous jobs for a prediction of how well that candidate will be able to perform the job in question. If candidates have high school diplomas, many employers will infer that the individuals have the intellectual ability to learn various subjects and the emotional discipline to study and complete a goal of passing the necessary high school classes. For job candidates without a diploma, employers wishing for empirical evidence of their ability to perform the job successfully must look principally to a candidate's previous employment record and work references. Thus, on average, previous work experience may well have a larger impact on the likelihood of obtaining a future job for high school dropouts than for high school graduates.[19] To test if the returns to early experience are different for high school dropouts and graduates, F-tests were used with our previously specified regression equations. While some of the equations showed no significant difference in returns to early experience for dropouts and graduates, graduating from high school does affect these returns in many of these regressions. See Appendix A for the results. Because there are differences in some cases, we decided consistently to distinguish between high school graduates and high school dropouts so that their early labor market experiences could have a different impact on their employment later in their careers.

[19] For instance, a job advertisement that lists as requirements that an individual "must be a high school graduate or have 3 years of work experience" reflects this type of screening method by employers that treats a high school diploma and work experience as somewhat substitutable.

A similar argument applies when comparing the importance of work experience for job candidates possessing a high school diploma but no college education with those who also go on to college. Employers are more willing to accept an applicant without much job experience if he has attended college, particularly if he has also graduated from college. However, among individuals who attended college, there is a very wide variety of types of schools, degrees, and areas of concentration. If we were to treat these individuals as one single group when analyzing their labor market experiences, there would be no effective way of controlling for the different effects of types of college education. On the other hand, separating individuals with different levels and types of college experiences would soon create many separate groups that are too small to draw any significant conclusions from analyzing them. Moreover, the fact that someone is attending college at age 19 or 20 is likely to affect his decision of whether to look for a job at that time in a variety of different ways. Some individuals must work in order to pay their tuition, while others often view studying in college as their only current job if they can afford to do so. In the interest of clarity, we have decided to focus our analysis only on men who dropped out of high school and those who graduated from high school but never completed a year of college. While the challenges faced when analyzing how the differences among college students influence the effects of unemployment scarring on them is beyond the scope of this paper, it would be an interesting subject for future research and could extend much of the analysis used here.

Because a greater fraction of the youth population attends college for at least some time than did youths in the early 1980s, we would expect the current group of non-college

attenders to be a smaller percentage of their age group, and, on average, they may be a less skilled group of individuals than non-college attenders in the '80s. Hence using the experiences of that group of men never attending college to predict the experiences of current young men may create an overly optimistic outlook. Thus, it is important to be careful in generalizing results to other labor markets in other time periods.

Individuals who reported that they had worked in the past calendar year are also asked to report their hourly wages. If someone is not paid by the hour, CHRR used the information that he did provide about his wages or salary to calculate his wages on a hourly basis.[20] For certain individuals in some years, this is impossible using the information they provided, but the majority of those who do work during the year also have a hourly wage variable. Among the sample of 362 male high school dropouts with AFQT scores who reported the number of weeks they worked at age 19, 20, 29 and 30, a subset of 134 reported their wages in those years. Similarly, of our sample of 1,082 men with AFQT scores who graduated high school but did not attend college and who reported their number of weeks worked at ages 19, 20, 29, and 30, a total of 400 worked and reported their wages in those years. We have used the annual Personal Consumption Expenditure (PCE) statistic to deflate these nominal wages reported in the NLSY79 to real hourly wages in 1996 dollars.[21] Much of the analysis that will follow uses ordinary least-squares regression procedures. To minimize the sensitivity of our estimates to outliers, we have chosen to use only those individuals whose reported real hourly wages are between $1 and $100 in any of the years observed (at ages 19, 20, 29, and 30). Of

[20] NLSY79 User's Guide (1998), pg. 279.
[21] PCE from The Economic Report of the President (Feb. 2003), Table B-6, pg. 284.

141 dropouts who reported their wages at all four ages, 7 men were omitted for not having wages in the required range. Of 412 graduates reporting their wages at all four ages, 12 were omitted as outliers.

Equation (4) measures the impact on future wages of not working during youth for an entire year. This equation can be estimated for all individuals who worked for pay at ages 29 and 30. However, this estimation does not require those individuals to have worked for pay at ages 19 and 20. Of the sample of 362 dropouts who reported their AFQT scores, 268 also reported earning postive wages at ages 29 and 30. Of the sample of 1,082 graduates, 873 also reported their wages at ages 29 and 30. Again, individuals who reported wages that were above $100 and below $1 were omitted as outliers.

In short, the analysis in this paper is based on six samples of men from the NLSY79. The individuals in all four samples must have reported the number of weeks they worked in the calendar years in which they turned ages 19, 20, 29, and 30. The first sample (Group I) includes the men who never graduated from high school. The second sample (Group II) includes the men who graduated from high school but did not complete a year of college by age 30. The third and fourth samples (Groups III and IV) are subsets of the first and second samples respectively. Groups III and IV omit all individual from the first two samples who did not work or did not report a reasonable real wage (between $1/hr. and $100/hr.) in the calendar years when they were aged 19, 20, 29, and 30.[22] Finally,

[22] Individuals who were part of the military sample (dropped from the NLSY79 in 1985 except for 300 individuals) or part of the economically disadvantaged non-black, non-

Groups V and VI are also subsets of the first two samples. These last two groups omit the individuals from Group I and II who did not work for pay (or whose reported real wages were greater than $100/hr. or less than $1/hr.) at ages 29 and 30. Table 1.1 provides a list of restrictions that were used to select these six groups and the number of observations that remained after each restriction. The six numbers in bold correspond to the six data groups described above. Table 1.2 includes descriptive statistics for Groups I through IV, and Table 1.3 provides descriptive statistics for the final two groups.

V. Probability Trees

Trees: Employment Experience From Ages 19-22

The greatest benefit of using the NLSY79 for analyzing employment and wage persistence is that the data set allows us to track the experiences of the same individuals over many years. By partitioning the individuals who work more than a certain number of weeks during each year from those who do not, we are able to analyze work patterns as individuals age. Probability trees provide a graphical representation of this pattern, or persistence, in employment. The probability tree in Figure 1.0 follows high school dropouts for four years beginning when they are aged 19. Of our sample of 362 high school dropouts, we observe the weeks worked from ages 19 through 22 for 352 young men. In each year, the tree partitions these men into 2 groups: those who work at least one week out of the year (represented by a "1" in the tree) and those who do not work at all during the year (represented by a "0").

Hispanic group of the supplementary sample of the NLSY79 (who were dropped from the NLSY79 after 1990) were also omitted from all 6 groups.

The tree has four stages of partitions corresponding to the four different years. In each year, the figure above the horizontal line represents the number of individuals in a particular group. In parentheses underneath the line is the probability of an individual being a member of that group conditional upon being at the node that precedes it. For instance, the probability of a worker working at least one week during the second year conditional having worked one or more weeks during the first year is .92.[23] The number below the conditional probability is the average number of weeks worked for each group of individuals. Figure 1.1 depicts a similar partitioning of the 1,046 high school graduates for whom we have observations of the weeks worked at ages 19, 20, 21 and 22.

As both the graduates and dropouts age, patterns emerge in their employment experiences. First of all, individuals who work during one year are much more likely to work during the following year than those who do not work initially. This process appears to have a collective effect over many years as emphasized by the fact that the graduates who have worked at ages 19, 20, and 21 are 53 percentage points more likely to work when aged 22 than are those who did not work at all during the previous 3 years.[24] Additionally, even among those who obtain a job during a particular year, the individuals who have a history of working in every year on average work a greater number of weeks than those who have experienced entire years of non-employment. For instance, among dropouts, the individuals in the group who worked all four years worked an average of 14 additional

[23] According to Figure 1.0, 263 dropouts worked at least one week at ages 19 and 20, and 285 dropouts worked at least one week when aged 19. (263/285 = .923)
[24] From Figure 1.1, of 691 graduates who worked at ages 19, 20, and 21, 652 also worked at age 22 (652/691 =.94). 22 out of the 54 graduates who never worked at ages 19, 20, and 21 worked at least one week at age 22 (22/54 =.41).

weeks at age 22 compared with those who worked at ages 21 and 22 but were not employed at ages 19 and 20.[25]

The other most noticeable feature of the probability trees is that, for both graduates and dropouts, over half of the individuals in each sample work for at least one week in each of the four years. The remaining individuals are relatively evenly dispersed among the other ending groups in the tree, meaning these individuals have experienced between one and four complete years of non-employment. The high concentration in the "worked in each year" classification highlights the strong likelihood for individuals to be employed at least one week of every year. This high likelihood makes sense when we consider that if a person acquires and keeps a job even for a short period of time at the end of the year, he would likely be employed for one or more weeks in both the current year and the next, allowing him to be counted in the employed group for years. Thus, whether someone works one week in a given year may not be telling us much about his overall labor market experience. It may be more useful to analyze a tree that is partitioned in a way that creates a less skewed distribution of individuals into the groups.

Figures 2.0 and 2.1 partition high school dropout and graduates, respectively, based on whether the individuals had worked forty or more weeks during each year. Forty weeks is a useful threshold for classifying individuals because some statisticians and policy makers view it as an indicator that an individual was employed for the full year. Any

[25] The 230 dropouts who worked in all four years worked an average of 39.3 weeks at age 22. The 8 dropouts who worked at ages 21 and 22 but had no employment at ages 19 and 20 worked an average of 25.3 weeks at age 22 (39.3-25.3 =14 weeks).

worker who is employed less than forty weeks spends over one-quarter of the year unemployed or out of the labor force. When we partition individuals according to forty work weeks, the trees of both the high school graduates and the high school dropouts have large numbers of individuals in the two "end" groups. That is, the two largest groups of individuals are those who never worked forty or more weeks in a year between ages 19 and 22 and those who were employed forty or more weeks in every year during that time. For graduates, 15 percent of the sample worked forty or more weeks during every year, [26] and among dropouts, over 10 percent of our sample worked for forty or more weeks in all four years.[27] Similarly approximately 30 percent of each schooling group never worked forty weeks in any year when they were aged 19 through 22.[28]

As mentioned above, a high relative number of workers in the edges of these probability trees is a strong indication of persistence in employment. However, these probability trees do not distinguish whether a causal mechanism is inducing the individuals with jobs in earlier years to work more weeks in later years or whether the persistence is merely a sorting that occurs due to permanent individual differences that make some people more likely to be employed for many weeks during each of the years observed.

Individual Characteristics: Probability Tree Grouping

Tables 2.1 and 2.2 report personal characteristics and employment experiences of three distinct groups of individuals in the probability trees. These groups are: (1) individuals

[26] 162 of 1046 graduates worked 40+ weeks in every year (162/1046=.154).
[27] 41 out of 352 dropouts worked for 40+ weeks in all four years (41/352=.116).
[28] 305 of 1046 graduates never worked (305/1046=.292). Out of 352 dropouts, 111 individuals never worked (111/305=.315).

who worked for the selected number of weeks in all four years they were observed, (2) individuals who worked the selected number of weeks in some but not all of the relevant years, and (3) individuals who never worked the selected number of weeks in any year between ages 19 and 22. These groups are further divided between high school dropouts and graduates, and two different required number of weeks of work are used to partition the men: forty or more weeks and one or more week.

From these tables, it is apparent that the individuals who work for either of the specified number of weeks in every year have significantly higher AFQT scores than those who do not work that number of weeks in any year. For instance, among graduates, the AFQT scores of workers who worked the one or more weeks in all four years are on aveage 0.07 higher on average than the scores for men who worked one or more weeks in only some years, and 0.15 higher than the scores of men who never worked in all four years.[29] Additionally, a worker's mother's schooling (as measured by the variable "momhsgrd,") also is correlated with employment probabilities. High school graduates who worked forty or more weeks in every year are 23 percentage points more likely to have a mother who completed high school than are workers who never worked the specified number of weeks.[30]

[29] From Table 2.2, the average AFQT scores for graduates who worked one or more weeks in all four years, some years, and no year are .386, .313 and .231, respectively.
[30] From Table 2.2, 76.2% of graduates who worked 40 or more weeks in all four years had mothers who also graduated. Comparatively, 52.9% of graduates who never worked 40 weeks have mothers who graduated high school.

Ethnicity is also correlated with employment probabilities, though less strongly than AFQT score. For instance, blacks represent 26 percent of the dropouts who never work 40 weeks, but only 6 percent of dropouts who work 40 or more weeks in every year are black. On the other hand, roughly an equal percentage of dropouts who worked at least one week in each of the four years and dropouts who did not work at all in all four years are black (14.3 percent and 14 percent, respectively). Approximately 15 percent of the overall sample of dropouts and 17 percent of the sample of graduates are black. Graduates who did not work in any of the four years are also over four times more likely to be Hispanic than are dropouts who worked at least one week in all four years.[31] However, the percentage of dropouts who worked 40 or more weeks that are Hispanic (16.9 %) is slightly greater than the percentage of dropouts who never worked 40 weeks that are Hispanic(13.8%).

The strong correlations between AFQT scores, mother's schooling, race, and consistently high weeks worked in each year suggest that heterogeneity could be significantly influencing the persistence in employment that we observe in these probability trees. From these tables, we can easily recognize the possibility that mother's schooling and ability as measured by AFQT scores could be greatly affecting the number of weeks individuals work at all four ages. Assuming that a causal relationship between individual characteristics and weeks worked at different ages exits, ignoring mother's schooling and an individual's AFQT score would lead us to mistake a mere correlation in weeks worked at different ages for a causal relationship between labor market experience

[31] 6.3% of graduates who worked 1+ weeks each year are Hispanic, while 26.3% of those who never worked in any of the four years are Hispanic.

at different ages. On the other hand, both individual differences and past experiences could be affecting later employment experience simultaneously. To attempt to parse the effects of heterogeneity and state dependence, we need to use the statistical tools of regression analysis.

Another weakness of probability tree analysis is that it becomes tedious when the number of years analyzed increases because the number of branches will grow exponentially. As we are interested in the long-term effects of youth unemployment, we wish to examine the persistence in employment and wages ten years into an individual's career. By looking at the variables 'wkwkat29' and 'wkwkat30', our groupings do provide one view of long-term effects of employment. On average, dropouts who work 40 or more weeks at ages 19, 20, 21 and 22 work over 18 additional weeks per year at ages 29 and 30 than dropouts who never worked 40 weeks when aged 19 to 22.[32] Because all people do not work for pay, we do not have observations of wages of all individuals in our data set. Among the individuals whose wages we do observe, we have calculated the average hourly real wages.[33] There exits some positive correlation between the number of weeks a worker is employed when aged 19 to 22 and his wages 8 years later. For instance, among graduates, who worked 40 weeks in all four years earned an average wage of $13.20 at age 30, while graduates who never worked 40 weeks in a year earned $10.01 on average at age 30.

[32] Dropouts who worked 40+ weeks in all four years worked 47.4 and 46.8 weeks at ages 29 and 30, respectively. Dropouts who never worked 40 weeks at ages 19-22 worked 29.1 and 28.2 weeks on average at ages 29 and 30, respectively.

[33] Note: The number of individuals observed for who we calculated average real wages is in parentheses following the mean of the wages for the group.

VI. Regression Analysis

The probability trees and correlations discussed above indicate a strong presence of unemployment persistence when workers are aged in their twenties, but these measures do not show whether this persistence is a result of individual differences or is causally related to the unemployment experienced during youth. To make this distinction, the regression techniques described above were used.

Employment

We first estimated equation (1) to analyze persistence in the amount of employment experienced by the men in our sample. Because the sum of weeks worked and weeks not worked is 52, the same regression result can be applied to weeks not worked by simply changing the intercept term and reversing the signs on the regression coefficients. For both high school graduates and high school dropouts, the number of weeks worked at age 29 was regressed on the weeks worked at age 19, and the number of weeks worked at age 30 was regressed on weeks worked at age 20. The estimates for the high school dropouts are contained in Tables 3.0 and 3.1 and those for high school graduates are in Table 3.2 and 3.3. The 1960 birth cohort is the cohort that is omitted as a regressor from these estimations. The coefficients on the dummy variables representing the later cohorts in these equations are negative probably because the 1960 cohort, whose members were aged 29 in 1989 and aged 30 in 1990, were observed at this time during a higher employment stage of the business cycle than the later cohorts, who were observed during the recession of the early 1990s.

In the first column of Tables 3.0 through 3.3, we have controlled only for the birth cohort of the worker. The estimates in column (II) of each table add other control variables. These equations imply that, for high school dropouts, an individual working ten extra weeks at age 19 or 20 would be expected to have worked more than two additional weeks ten years later. High school graduates working ten extra weeks at ages 19 or 20 would be expected to have between one and two extra weeks of work ten years later in their career. Though this impact of lagged employment might seem small, the fact that the events of 10 years had taken place in the intervening years makes these results more striking that they might otherwise be. If all of the relationship between employment at different ages could be attributed to state dependence, then these results imply that unemployment as a youth affects an individual's experiences in the labor market 10 years later.

Controlling for ability (as measured by AFQT scores), mother's educational level, and race in these equations reduces the regression coefficient on the earlier weeks worked variable by 10 to 20 percent in both the sample of high school graduates and the sample of high school dropouts. This decrease implies that at least some of the association between employment at different ages is attributed to observed differences among these men. Other things equal, black men work two to five fewer weeks than white men at ages 29 and 30. Also, among high school dropouts, a worker with an AFQT score that is 0.10 higher than a worker with otherwise identical characteristics is likely to work two to two and a half more weeks per year. Other things equal, high school dropouts whose mothers graduated from high school work approximately 3 more weeks at ages 29 and 30 than

dropouts with mothers who were also dropouts. A graduate with a mother who also graduated high school works one more week per year than an otherwise identical graduate whose mother did not graduate. This impact of mother's schooling suggests that a worker's family background matters to his long-term employment experiences.

Equation (2) is estimated in the third column of Tables 3.0 through 3.3. In these estimations, the coefficients on "nowkat19" and "nowkat20" represent the additional negative impact on future employment of not working at all as a youth *beyond* what is captured by the linear slope of the relationship between weeks employed at age 30 and weeks employed ten years earlier. According to these results, dropouts who do not work for the entire year at ages 19 or 20 are employed an average of five to six and a half fewer weeks ten years later than otherwise identical workers who work even one week out of the year when aged 19 or 20. Among graduates, men who are employed for at least one week at age 19 work between one and three and a half more weeks a decade later than those who never work when aged 19.

The key question is whether this correlation between early and later weeks of employment is due to the actual labor market experiences during youth, or whether it merely indicates that important heterogeneity within these samples remains that we do not observe and for which we have not controlled. As specified in equations (15) and (16), first differencing the employment equations should help us answer this question. For high school dropouts and graduates, Tables 3.4 and 3.5 report the results of regressing the change in weeks worked at age 30 on the change in weeks worked at age 20. With this specification, all invariant characteristics of an individual that determine

weeks worked at 30 drop out of the equation. Columns (I) and (II) report estimations of equation (15) with and without the constant term. As the first two columns of these tables demonstrate, omitting the constant term, which actually is omitted from the specification of equation (15), creates no significant difference for our results. With or without the constant term in the equation, the estimated relationship between the change in weeks worked at age 30 and ten years earlier is far from statistically significant. Moreover, the coefficient on the lagged change in weeks worked regressor is one-tenth of the size of the coefficient on the lagged weeks worked term in the level equations.

According to equations (8) and (15), if all heterogeneity were controlled for in the level equations by including permanent individual characteristics as regressors, then the coefficient on lagged weeks worked in the level equation should be the same as the coefficient on the lagged change in weeks worked in the first difference equation. The large disparity in the values of these coefficients provides strong support for the presences of heterogeneity in our data. That is, the disparity implies that much of the correlation between weeks worked during one's youth and later in one's career is probably due to unobserved heterogeneity that was not fully controlled for by the other variables in the level equations. Though including AFQT score, race, and mother's schooling as regressors does somewhat diminish the strength of the relationship between the dependent variable and lagged weeks worked, a large amount of heterogeneity must still be driving the results in Tables 3.0 through 3.3.

The third column of Table 3.4 and 3.5 reports estimates of equation (16) for high school dropouts and graduates. In addition to using the change in weeks worked as a regressor, we also include the year-to-year change in the state of being out of work for the entire year. For dropouts, the large negative association observed in the level equations (Tables 3.0 and 3.1) between not working at all at age 20 and weeks worked ten years later becomes statistically insignificant and positive in first differences (Table 3.4). That is, no statistically or economically significant relationship exists between the year-to-year change in working zero weeks out of the year at age 20 and the change in number of weeks employed a decade later.

Equations (9) and (16) imply that, if the other regressors in the level equations adequately control for the effects of heterogeneity, the coefficient estimating the relationship between the change in weeks worked at age 30 and the year-to-year change at age 20 of being out of work for the entire year is the same coefficient that measures the partial association between weeks worked at age 30 and not working any weeks at age 20. Because these coefficients are not the same in the results, the estimates of these parameters suggest that unobserved heterogeneity significantly influences the relationship between weeks worked at age 30 and the lagged state of working zero weeks for a full year for high school dropouts. Thus, we infer that, among dropouts, state dependence is not present in the relationship between not working at all at age 20 and the weeks one works at age 30.

This is not the case for high school graduates. For these individuals, the relationship between the annual change in weeks worked at age 30 and the change in being out of work for an entire year at age 20 is negative and statistically significant at the 0.08 level. The significance of this result implies that not working for an entire year at age 19 *causes* high school graduates to work an average of one and half fewer weeks at age 29 than they would have worked had they been employed for at least one week ten years earlier. As mentioned above, an estimation of the first difference equation (16) is not susceptible to the effects of unobserved heterogeneity because permanent individual differences do not change from one year to the next. Therefore, we can infer that state dependence is present in the relationship estimated here in first differences. Because the coefficient on the change in being out of work for a full year at age 20 in Column (III) of Table 3.5 is very close to the estimates of the relationship between the weeks worked at ages 29 and 30 and being out of work for a full year a decade earlier (presented in Tables 3.2 and 3.3), we infer that the estimates from the level equations are also displaying a relationship characterized by state dependence. That is, a causal association exists between the weeks a graduate works at age 30 and whether or not he worked any weeks out of the year one decade earlier.

Our results give mixed indications of the state dependence in the relationship between the weeks one works at age 30 and one's employment experience a decade earlier. It appears that whether an individual works at all during a year of youth causes a noticeable impact on his future employment probability, but that the number of weeks that he works does not. Also, high school graduates are affected by state dependence in this relationship

while dropouts are not. We must compare these results to the effects of work experience during youth on one's wages later in life. [34]

Wages

The wage equations are estimated using data groups III and IV, which include the high school dropouts and graduates who worked for pay at ages 19, 20, 29, and 30. Table 3.6 and 3.7 report the estimates of equation (3) and (5) for the high school dropouts. The first and third columns of each table estimate the equations controlling only for cohort effects while the second and fourth column also control for AFQT score, mother's schooling, and ethnicity. According to these results, a worker with an AFQT score that is 0.1 higher should have 7.5 to 9 percent higher hourly real wages at ages 29 and 30 than an otherwise identical worker. Other things equal, among high school dropouts, estimates of black men's wages are between 2 percent higher and 4.5 percent lower than the wages of white men at ages 29 and 30.

In these results, neither wages nor weeks worked during youth have a consistent and statistically significant effect on the future wages of high school dropouts. Though the low statistical significance could be a result of the small sample size (134 men), the coefficients on lagged weeks worked and lagged logarithm of real wages are also extremely small in magnitude and negative in two of the four regressions. For instance,

[34] Equation (6) was also estimated for both graduates and dropouts to analyze if the relationship between employment at age 30 and lagged weeks worked was influenced by race, mother's schooling, or AFQT score. Such interactions, however, did not reveal any relationships in the data significantly different than those already noted and, thus, were not reported.

these regressions estimate that, during a high school dropout's youth, a 10 percent higher hourly wage would, at best, raise the person's expected wage rate at age 29 by 0.16 percent, other things being equal. Similarly, 10 extra weeks per year of employment during youth raises a worker's probable wages at age 29 and 30 by anywhere from 0 to 7 percent.

High school graduates show much stronger signs of wage persistence than do the dropouts. As reported in columns (II) and (IV) of Table 3.9, the persistence in wages, which is specified by equation (5), is statistically significant in both cases. These regressions estimate that high school graduates who have 10 percent higher wages during their youth than otherwise identical men are likely to earn 1.3 to 2.5 percent more per hour than the other men by age 30. For graduates, the second column of Tables 3.11 and 3.12 display estimates of equation (3) for graduates. These results imply that graduates who work 10 extra weeks per year during youth earn an estimated 2 to 3.5 percent higher wages ten years later in life than otherwise identical men with fewer weeks of early employment.

Among men who graduate from high school, black men earn between 8 and 14 percent less per hour than white men at ages 29 and 30, This disparity is nearly twice as large as the difference between the wages of black and white high school dropouts. Also, the AFQT score of high school graduates, while statistically significant in these regressions, has on average only one-half to one-third of the impact on later wages that AFQT scores have on wages at age 30 of high school dropouts. Why do the AFQT scores of dropouts

have a stronger relationship with later wages than do the AFQT scores of graduates? It is conceivable that high school dropouts must rely more on their abilities (which AFQT scores attempt to capture) than graduates because dropouts lack a diploma to signify their quality as a worker. It this is true, then AFQT and ability would be more important for dropouts.

Similarly, we proposed earlier in this paper that past employment experience may be more important for dropouts than graduates because dropouts lack a diploma that show that they had the discipline and ability to complete high school. Comparing the results in Column (II) and (IV) of Tables 3.6 with those in the second column of Tables 3.11 and 3.12 shows that this is not the case for the relationship between wages in one's youth and later in life. Graduates actually experience higher correlation in early wages and wages ten years later than do dropouts.

Similar to our regressions of equation (1), the estimated wage equations (3) and (5) could be affected by unobserved heterogeneity. As represented in equations (17) and (19), first differencing equations (3) and (5) allows us to distinguish the effects of state dependence from unobserved heterogeneity in wage persistence. For dropouts, the estimates of equations (19) with and without the intercept term are in columns (I) and (II), respectively, of Table 3.8 for dropouts. Similarly, equation (19) is estimated with and without the intercept term in the third and fourth columns of Table 3.8.

Among high school dropouts, the first difference equation (19) reveals no statistically significant relationship between the change in a worker's wages at age 20 and the change in his wage rate at age 30. Likewise, the change in weeks worked at 20 does not significantly affect the change in wages at 30 for dropouts. These results should come as no surprise because even the level equations for the group of dropouts (Table 3.6 and Table 3.7) showed no consistent relationship between either wages during youth and ten years later or between weeks worked as a youth and wages later in life. High school graduates, on the other hand, do exhibit strong wage persistence in the level equations, and a significant relationship between weeks worked as a youth and wages at ages 29 and 30. Thus, we focus on the results of the first differences equations for graduates' wages at age 30 in Table 3.10 and 3.13.

Equation (19) is estimated with an intercept term in column (I) of Table 3.10 and without an intercept term in Column (II) of that table. Similarly, the estimates of equation (17) with and without an intercept term are reported in the first and second columns of Table 3.14. Estimating equations (19) and (17) with or without the intercept creates little difference in the results for high school graduates. Thus, we will discuss the result of the estimations with the intercept term, but the same analysis applies to the regressions that omitted the intercept term as a regressor.

The results reported in Table 3.10 are probably one of the most clear indications of the effects of unobserved heterogeneity. In the level equations for wage persistence in high school graduates, which are estimated in columns (II) and (IV) of Table 3.9, the

relationship between wages in one's youth and wages later in life had both statistical and economic significance. However, when we estimate the first difference of these equations (Table 3.10), the relationship between the change in wages in one's youth and ten years later in life becomes statistically insignificant and slightly negative.

When estimating the first difference equation (17), the disappearance of a positive association between wages during youth and ten years later leads us to conclude that the wage persistence revealed in level equation (5) for high school graduates is *not* a causal relationship. That is, nothing about the *state* of earning high (or low) wages during ages 19 and 20 is affecting these high school graduates in a way to make them more likely to earn high (or low) wages later in life, but rather members of the group have certain unobserved permanent characteristics, such as punctuality and discipline, that make them prone to earning higher wages as youths and also raise their expected earnings during adulthood.

The results in columns (I) and (II) of Table 3.13 show a very different result for the effect of weeks worked during youth on later wages. Estimated either with or without an intercept term, the impact on the change in the logarithm of real wages at age 30 of the change in weeks worked at age 20 (equation 17) is statistically significant with a magnitude of 0.003. This means that a worker who worked 10 more weeks at age 20 than he worked at age 19 would be likely to see his wage increase three percent more than a worker whose was employed for the same number of weeks at ages 19 and 20.

Furthermore, if equations (17) and (10) are correctly specified, this result also implies that a graduate who works 10 more weeks at age 19 or 20 than an otherwise identical worker should earn a three percent higher hourly wage ten years later.

What is important about this result is that the size of this regression coefficient is very close to the value of the coefficients associated with the relationship between weeks worked during youth and wages later in life from the level equations (equation 8), which are reported in the second column Tables 3.11 and 3.12. According to the equations in section III, if the first difference equation (17) gives the same result as the estimated level equation (3), then the regressors in the level equation are accurately controlling for individual heterogeneity. Thus, these results provide evidence of a small but significant presence of state dependence in the relationship between weeks worked as a youth and real wages later in life. Working 10 extra weeks at age 19 affects a high school graduate in a way that *causes* him to earn 3 percent higher wages on average ten years later than he otherwise would have earned at that time. Similarly, six months of extra work at age 20 is predicted actually to *cause* the worker to earn 8.1 percent higher wages at age 30 than if he had missed the additional employment opportunity.

The state dependence we do see is rather small, but the fact that effects on wages are still present *ten years* after the cause, which is the amount worked as a youth, is still striking. Moreover, it is important from a policy standpoint because, even a 3 percent increase in wages, if the increase lasts for ten or more years, could have a large dollar value when a worker's income is summed over many years. For instance, if a worker earns $20,000

annually, a three percent increase represents an additional $600. An extra $600 per year over 40 years constitutes $24,000, not a trivial sum.

Interaction Effects and Wages

It is notable that state dependence is present in the relationship between wages and lagged weeks worked for high school graduates but not for dropouts. To more closely examine this selective presence of state dependence, we estimate equation (7) in the third column of table 3.11 and 3.12 while allowing for an interaction between the weeks one worked at age 19 (or age 20 in Table 3.12) and the worker's AFQT score.[35]

Column (III) of Table 3.13 reports the estimates of equation (21), which is the first difference of specification (7). The coefficient on the interaction of AFQT score and year-to-year change in lagged weeks worked is statistically significant and is slightly greater in magnitude than the estimate of the same interaction effect in the level equations. Equations (14) and (21) imply that this parameter is the correct estimate of the interaction effect of AFQT score on the relationship between wages and lagged weeks worked when properly controlling for heterogeneity. Thus, it is clear that the magnitude of the effect of weeks worked on later wages is highly influenced by one's AFQT score, and that state dependence is present in this effect. Among high school graduates in the workforce, the mean AFQT score is .394. According to the results in Column (III) of Table 3.13, a high school graduate who has the mean AFQT score would be expected to have 3.4 percent

[35] We also estimated equation (7) allowing for interaction effects between mother's schooling and lagged weeks worked, and between ethnicity and lagged weeks worked. However, such interactions do not appear to significantly affect the data. Therefore, only the AFQT-lagged weeks worked interaction is reported and discussed.

higher wages at age 30 for every additional 10 weeks he works ten years earlier.[36] Moreover, according to these results, for every additional increase of 0.10 in a worker's AFQT score, the effect on his wages at age 30 of his lagged weeks worked is expected to increase by 1.1 percentage points. Therefore, a worker otherwise identical to the one described above but with an AFQT score of 0.60 should see his wages at age 30 increase by approximately 5.6 percent for every additional 10 weeks he works a decade earlier.[37] On the other hand, an otherwise similar worker who has an AFQT score of 0.20 is expected to only earn only a 1.2 percent higher wage at age 30 if he works an additional 10 weeks at age 20.[38]

The strong interaction between AFQT scores and lagged employment's effect on wages implies that workers with lower ability (as measured by the AFQT) experience weaker long-term effects of unemployment than do workers with higher ability. As mentioned above, the AFQT was administered to the respondents during the year 1980, when the many of the responders were still in high school. Therefore, the AFQT measures the workers' ability at or before the beginning their careers. Perhaps this type ability deteriorates if it is not actively used. If this is the case, the workers with higher AFQT scores and more ability have more to lose by not working than do workers who start out with lower ability.

---

[36] From Table 3.13: -0.00098 + (afqt * .011) = -0.00098 + (.394 * 0.011) = 0.00335.
[37] From Table 3.13: -0.00098 + (afqt * .011) = -0.00098 + (.60 * 0.011) = 0.00562.
[38] From Table 3.13: -0.00098 + (afqt * .011) = -0.00098 + (.20 * 0.011) = 0.00122.

It could be hypothesized that the interaction discussed above explains why high school dropouts do not exhibit a significant relationship between wages and lagged weeks worked. According to this conjecture, dropouts, whose average AFQT score is .20 lower than that of the graduates, may experience less state dependence in the relationship between their wages and lagged weeks worked as a result of having lower ability (as measured by their AFQT scores). However, when we estimated equations (7) and (21) for high school dropouts (not presented in tables), the dropouts' AFQT scores showed no significant correlation with the presence of state dependence in the relationship between their weeks worked during youth and their later wages.. In those regressions, the coefficient measuring the interaction term between AFQT and weeks worked during youth was not significant and very small in magnitude. Therefore, something other than their lower ability must be causing dropouts not to experience state dependence.

Wages and Non-Employment

Last of all, we estimate equation (9) to measure the impact of not working for an entire year during youth on wages ten years later. This estimation uses data groups V and VI, which include of high school dropouts and graduates who worked for pay at ages 29 and 30 but were not restricted to having to work for pay at ages 19 and 20. The dropouts (group V), who showed no significant state dependence in the relationship between their wages at 30 and lagged weeks worked, produced similarly insignificant estimates for the relationship between wages at age 30 and not working at all during youth. Thus, they were not reported in the tables of this paper.

Column (I) of these tables reports estimates of equation (8) controlling only for cohort effects, and column (II) of the tables gives estimates of equation (8) while controlling for race, mother's schooling, and AFQT score. Equation (9) is estimated for the graduates in column III of Tables 3.14 and 3.15. According these results, graduates who work for zero weeks at age 20 are expected to earn between 3.7 and 4.6 percent lower wages at age 30 than graduates who do not spend age 20 without any employment. This effect is in addition to the association measured by the slope of relationship between weeks worked and later wages estimated in equation (8).

While these results do show that workers who do not work at age 20 have lower wages a decade later than otherwise identical workers, the interesting issue is whether the state of not being employed actually *causes* workers to have lower wages in the long term. Table 3.16 answers this question by providing first difference equations unaffected by unobserved heterogeneity. Equation (17) is estimated with and without an intercept term in columns (I) and (II), while equation (18) is reported with an intercept in column (III). In first differences, there is a positive and statistically insignificant relationship between the year-to-year change in wages at age 30 and the year-to-year change in the state of working zero weeks for a full year a decade earlier. This insignificance in the first difference relationship indicates that heterogeneity, not state dependence is the source of the negative association between wages at age 30 and whether someone worked for any week during the entire year a decade earlier.

This result is particularly interesting because it implies that whether someone is out of work for an entire year of their youth *causes* no additional negative effect on their later wages but it does *cause* a negative impact on the number of weeks that they work later in life. This may indicate that skill acquisition, which is often linked with earnings, is not significantly affected by whether someone works zero or one week out of a year during their youth. However, working a larger number of weeks can create a positive impact on future wages (and possibly on the acquisition of skills).

VII. Conclusions

Andrew Sum's argument[39] that currently unemployed young people are at risk of becoming permanently "left behind" assumes that there exists long-term persistence in youths' employment and wages and that a significant amount of this persistence is due to state dependence. As represented in the employment probability trees, the NLSY79 data clearly show high persistence in an individual's weeks worked and wages during youth and weeks worked and wages more than a decade into that individual's career. This part of Sum's argument is consistent with the data analyzed in this paper. However, Sum's assumption that state dependence is the cause of this persistence is not consistently supported by our analysis of the NLSY79 men.

In particular, we can infer the following from the data:

- The correlation between weeks worked in youth and later in life (equation 1) is due to unobserved heterogeneity. That is, differences among individuals in their

[39] Sum, et al. (2002), pg 85-87.

permanent characteristics give rise to variations in labor market outcomes that produce this correlation.

- The relationship between wage rates in youth and later in life (equation 5) is similarly the result of heterogeneity among individuals.
- For high school graduates, the number of weeks that a individual works during his youth (equation 3) causally affects his wage rate later in life. That is, there is state dependence in this aspect of individuals' labor market experiences.
- Among those graduates, weeks worked during youth have a greater impact on later wages for individuals with higher ability as measured by their AFQT scores.
- Among high school graduates, state dependence may also be present in the relationship between not working at all during a year of youth and weeks worked ten years later in life (equation 2). However, the relationship between not working at all during a year of youth and wages later in life (equation 4) shows no evidence of state dependence.

These results imply that, while state dependence affects neither the persistence in number of weeks worked nor the persistence in wages, state dependence is present in the relationship between wages at age 29 and 30 and weeks worked a decade earlier. Working with the Parnes data set, an earlier wave of the NLSY, Ellwood similarly found that, after attempting to control for heterogeneity, there was a greater amount of state dependence in the relationship between youth employment and later wages than in the relationship between youth employment and later employment.[40]

[40] Ellwood (1982), pg. 383.

Weeks of employment during youth appear to affect wages ten years later but weeks of employment during youth do not appear to have an effect on weeks of employment ten years later. Why does this difference exist? This contrast may suggest that different mechanisms are at work in the state dependence we observe. Because state dependence is not present in employment persistence (equation 1), it is less likely that we are seeing youth unemployment create a permanent cycle of repeated job losses for those who work fewer weeks as youths, which would be suggested by strong occurrence dependence. Instead, state dependence created by youth employment has an evident role in affecting the long-term wage rate of the worker. This may imply that, by working more during youth, workers acquire more "on the job" skills that can make them more productive and valuable to employers. Workers who spend less time working and developing these skills during youth can find jobs with equal likelihood in their later twenties, but the *quality* of these jobs and pay rates might be lower because they have missed the early opportunities to develop important work skills.

If the state dependence we observe were the result of workers becoming psychologically disaffected from the labor force and not working as hard in the in their jobs, then we would expect also to see state dependence in the effect of weeks worked on later weeks worked. Because the data do not suggest this type of state dependence, we should examine other possible causes of the state dependence in the relationship between wages and earlier weeks worked. Future researchers may wish to distinguish between a worker's time out of the labor force and time unemployed to obtain a clearer perspective

of workers' experiences and attitudes. The risk of this approach, however, is that workers may not accurately recall during which weeks they were looking for a job. As a result, the result that the number of weeks unemployed is likely to be measured with more error than weeks worked.

A closer examination of *who* experiences state dependence in the effect of employment on later wages provides further information about the type of mechanism that causes this state dependence. As noted above, state dependence is evident only in the relationship between weeks worked and later wages for high school graduates. In Section IV of this paper, we chose to separate high school dropouts from graduates because we had assumed that workers without a high school diploma would be more affected by their employment during youth than graduates.[41] We conjectured that potential employers place a greater emphasis on the work history of dropouts than that of graduates when making hiring decisions because the dropouts do not also have a diploma to indicate their level of skill and discipline. That is, we had assumed that skills and recognition acquired through schooling and those acquired through work experience were substitutes. Our results reveal, however, that it is graduates, not dropouts, whose future outcomes are more strongly affected by employment during youth. Thus, education and work experience appear to be complements rather than substitutes, that is, the value of work experience is not as large to those with less schooling.

[41] See above, pg.26.

Similarly, ability, as measured by AFQT score, also appears to be complementary with work experience. We noted above that the impact of weeks worked on later wages is stronger among those with higher AFQT scores. This may imply that workers of high ability and with more schooling have a greater capacity for learning from their work experiences and for developing skills valuable to their firm. If this is the case, the state dependence we observed in the data is probably the consequence of workers employed during youth acquiring skills that will raise their wage trajectory compared to otherwise identical workers who work less during youth. Further research could attempt to further identify skill acquisition as the mechanism causing of state dependence by examining whether education, ability, and weeks worked during youth also yield complementary benefits for those workers who attend college. The equations in this paper, however, would have to be modified to accommodate the greater variation in age of labor force entry for workers who attend college.

Also, to examine the implications of skill acquisition being the cause of the state dependence found in the data, it would be useful to examine workers who remain for many years in industries that require a large amount of training (for instance, those who operate heavy machinery). If skill acquisition is the mechanism creating state dependence in these data, then individuals who work in industries that require a large amount of training should experience a greater impact of youth employment on later earnings than workers who require less training to perform their job.

It is difficult to determine the cause of the state dependence evident in the relationship between being out of work for an entire year during youth and weeks worked ten years later. It is interesting that the number of weeks worked during youth affect one's future wages but not their future employment probability, while being out of work for an entire year does just the opposite, causing a negative impact on the future number of weeks one works but not significantly affecting that person's wages.

One distinction that could explain this phenomenon is that wages may be more closely related to skill acquisition while the amount of employment one experiences is affected more directly by employers' judgments about the worker's quality. Acquiring on-the-job skills requires more time than just one week of work. Hence we should not expect to find a significant difference between the acquired level of on-the-job skills of individuals who are out of work for an entire year as a youth and the level of those skills of individuals who worked for only a very short period of time. Therefore, if wages are linked to the level of skill that a worker has acquired, a worker who worked no weeks as a youth and one with only a few weeks of work should have similar wages.

On the other hand, the number of weeks that a person works at age thirty is affected by the probability of a person remaining in his job if he is working and by the probability of obtaining a job if he is out of work. For potential employers who are interviewing job candidates, the number of weeks that a person worked during their youth may be of less importance than whether or not the person was employed at all during that time. If a worker has worked at all, at least he may have some work references to provide a new

employer, and, if he does, his chances of getting that job will be greater than those of an otherwise identical worker who did not work at all during a given time period. It is difficult, however, to know exactly how a job candidate's work experience ten years earlier influences the opinions of potential employers. More research into the selection and hiring process for major industries and typical jobs in the United States could greatly improve our understanding of this form of state dependence.

Regardless of the mechanism that causes state dependence, to what extent do the data suggest that workers who experience high unemployment during youth become "permanently left behind" their peers who do find employment? When discussing "left behind" workers, it is common to think of the group of young workers with low education who consistently are unemployed or employed in low-wage jobs. However, our results imply that these workers are not currently at risk of experiencing long-term impact from their current unemployment. These workers with low education are more likely simply to be "unusually prone" to unemployment and especially exhibit that trait at a time of high unemployment such as the present. On the other hand, it is young workers with higher education and higher ability whose future outcomes in wages could be significantly affected by the current economic downturn. Thus, though some workers may be affected in the long-term by their current unemployment, these workers are not the ones thought of as most in need of employment assistance.

Given these results, should policy makers pay particular attention to youth unemployment and target it in order to change long-term outcomes? Ten years in the future, what long-term effect on employment and wages would we expect from a policy aimed at reducing

current youth unemployment? Ten years from now, such a policy may well have the most significant impact on the wages of high school graduates with high ability. In other words, among those who do not attend college, this policy would be helping those who are already more advantaged and successful than dropouts with low ability. While on average, graduates with high measured ability are not necessarily rich individuals, neither are they the most disadvantaged group. As a result, there would be a greater earnings disparity between more highly skilled graduates and lower skilled dropouts than if the policy were not implemented. Though the most disadvantaged workers with less ability would not actually be harmed, their employment situation would appear increasingly distinct from graduates who do benefit from this type of policy.

High school dropouts earn lower wages on average and are unemployed more often in the long-term than high school graduates. Policy intended to help dropouts should recognize that the skills acquired on-the-job appear to complement the general skills learned in school, not to substitute for them. Thus, a policy designed to raise youth employment among dropouts *only* is unlikely to be very effective in achieving long-term goals. For assisting the dropouts who experience the greatest amount of unemployment throughout their careers, trying to boost employment during youth is a tool for intervention that appears to be too late to be effective. If the intended effect is increased long-term employment and wages among those dropouts, a policy that focuses on increasing employment of those individuals needs to be coupled with policy to increase high school completion rates. In other works, to improve long-term outcomes, more attention ought to be focused on preventing such people from dropping out of high school.

Additionally, policy that encourages individuals to obtain GEDs may also be an effective method for helping high school dropouts benefit from greater employment as youths. It should be noted that Cameron and Heckman have found strong evidence suggesting that dropouts with GEDs on average have labor market experiences that are similar to those of dropouts without a GED and very different from the experiences of "traditional" high school graduates who completed twelve years of schooling.[42] Despite these results, it would be useful to conduct research examining whether, among dropouts with GEDs, *state dependence* exists in employment persistence and the effect of employment on wages ten years later. If, in this respect, individuals with GEDs are able to benefit more than other dropouts from work experience during youth, then perhaps acquiring a GED has an indirect value not recognized by Cameron and Heckman. Having a GED sends a "mixed signal" to employers because workers with GEDs are often more intelligent than other dropouts, but they also tend to "have lower levels of noncognitive skills"[43], such as discipline and persistence. If a high level of work experience is an indicator of a job candidate's non-cognitive skills, then perhaps a GED and work experience could complement each other, together sending a signal that the worker has both discipline and ability. In this case, policy intended to help those who have already dropped out of high school would be more effective to combine assistance for finding a job with encouragement to obtain a GED.

This is not to say that there is no role that policy can take to improve the long-term employment of high school dropouts and individuals with lower ability. Neither do we

---

[42] Cameron and Heckman (1993), pg. 2.
[43] Heckman and Rubinstein (2001), pg. 146.

recommend not attempting to assist young workers who experience large amounts of unemployment. On the contrary, our analysis shows that these workers have the greatest probability of earning low wages and experience the most unemployment in the long term, and, therefore, they need of assistance. However, what is important to recognize is that not being employed during youth is *not* the cause of the future unemployment and low wages predicted for these workers, and, thus, attempting to raise the level of employment among youths is not an effective tool for this intervention. Other means must be investigated for changing the long term situation of these individuals.

Finally, we can also see from the data that there is little state dependence in the persistence of wages for graduates or dropouts. Moreover, the effect on future outcomes of increasing weeks of employment has its largest impact when going from zero weeks to one week of work. Therefore, the type of job and the wages paid, in addition to the length of time the job is held, appear less important than whether someone had a job at all, even for one week out of the year, during youth. Policy motivated to increase youth employment should attempt to create the greatest impact by reducing long term unemployment for as many different individuals as possible and should be less concerned with the type of job or whether the individual is employed for the full year. This type of "employment occurrence dependence," as opposed to unemployment occurrence dependence, does play a significant role, and may imply that an efficient allocation of resources for increasing long-term employment should focus on getting every young individual a job of any duration.

References

Bureau of Labor Statistics, U.S. Department of Labor, NLS Handbook, 2002.

Cameron, Stephen and James J. Heckman, "The Nonequivalence of High School Equivalents," Journal of Labor Economics, 11, January 1993, 1-47.

Center for Human Resource Research (CHRR), NLSY79 User's Guide: A Guide to the 1979-1998 National Longitudinal Survey of Youth Data, Ohio State University, July 1999.

Ellwood, David T., "Teenage Unemployment: Permanent Scars or Temporary Blemishes?", The Youth Labor Market Problem, ed. by Richard B. Freeman and David A. Wise, University of Chicago Press, 1982, 349-390.

Executive Office of the United States of America, The Economic Report of the President. United States Government Printing Office, February 2003.

Gregg, Paul, "The Impact of Youth Unemployment on Adult Unemployment in the NCDS", Economic Journal,111, November 2001, F626-F653.

Heckman, James J. and George J Borjas, "Does Unemployment Cause Future Unemployment? Definitions, Questions, and Answers from a Continuous Time Model of Heterogeneity and State Dependence", Economica, 47, August 1980, 247-283.

Heckman, James J. and Yona Rubinstein, "The Importance of Noncognitive Skills: Lessons from the GED Testing Program, American Economic Review, 91, May 2001, 145-149.

Herbert, Bob, "Young, Jobless Hopeless", The New York Times, February 6, 2003.

Leonhardt, David, "Hiring In Nation Hits Worst Slump in Nearly 20 Years", The New York Times, February 6, 2003.

MaCurdy, Thomas, Thomas Mroz, and R. Mark Gritz, "An Evaluation of the National Longitudinal Study of Youth", Journal of Human Resources, 33, Spring 1998, 345-436.

Phelan, Christine, "Youth Shut Out by Labor Market", News from Northeastern, Northeastern University, February 3, 2003.

Sum, Andrew, et al., Left Behind in the Labor Market: Labor Market Problems of the Nation's Out-of-School Young Adult Populations, Center for Labor Market Studies, Northeastern University, November 2002.

Appendix A. Testing for Differences between High School Graduates and Dropouts

The purpose of this section is to determine whether it is necessary to consider high school graduates and dropouts as two separate groups when estimating our equations or whether we could pool these sets of individuals and estimate our equations for all of them together. Pooling would allow us the benefit of a large sample size. However, if the effects of the regressors are different for high school dropouts and high school graduates, then it is inappropriate to constrain the coefficients to be the same. The following F-tests measure the interaction of wage and employment persistence with one's educational attainment.

The regressions in the restricted models (Res) in the even columns of this table are the same specification as equations (1), (2), and (3) used elsewhere in this paper, except that the estimates here were based on the pooled group of high school graduates and dropouts. In the unrestricted model (UR) in the even columns of this table, variables such as 'wkwkat19 & hsgrad' were added. These variables are equal to the preceding variables in the column if the respondent is a high school graduate. The variables are set to zero otherwise.

As Table A.1 shows, the influence of being a high school graduate on the relationship between youth employment and wages and later employment and wages is significant in three of our six estimations. In some of the regressions, such as specification (3), the impact is significant, and pooling graduates and dropouts would be unwise. Alternatively,

in other regressions, particularly for specification (2), being a high school graduate seems to make little difference on the relationship between later wages and early employment.

In the interest of presenting a consistent explanation of the data, we presented all results for separate educational groups rather than pooling graduates and dropouts in some specifications and keeping them separate in others.

Table A.1
High School Graduates and Dropouts: F-test Results

Model Specification	(1) UR	(1) Res	(1) UR	(1) Res	(2) UR	(2) Res	(2) UR	(2) Res	(3) UR	(3) Res	(3) UR	(3) Res
Dep. Variable	wkwkat19		wkwkat20		lrwage19		lrwage20		lrwage19		lrwage20	
Regressors Used:												
constant	x	x	x	x	x	x	x	x	x	x	x	x
wkwkat19	x	x			x	x						
wkwkat19 x hsgrad	x				x							
wkwkat20			x				x	x				
wkwkat19 x hsgrad			x				x					
lrwage19									x	x		
lrwage19 x hsgrad									x			
lrwage20											x	x
lrwage20 x hsgrad											x	
afqt	x	x	x	x	x	x	x	x	x	x	x	x
cohort61	x	x	x	x	x	x	x	x	x	x	x	x
cohort62	x	x	x	x	x	x	x	x	x	x	x	x
cohort63	x	x	x	x	x	x	x	x	x	x	x	x
cohort64	x	x	x	x	x	x	x	x	x	x	x	x
black	x	x	x	x	x	x	x	x	x	x	x	x
hispanic	x	x	x	x	x	x	x	x	x	x	x	x
momhsgrd	x	x	x	x	x	x	x	x	x	x	x	x
Num Regressors	12	11	12	11	12	11	12	11	12	11	12	11
NOOBS	1444	1444	1444	1444	534	534	534	534	534	534	534	534
DF	1432	1433	1432	1433	522	523	522	523	522	523	522	523
R-squared	0.0815	0.078	0.1122	0.1109	0.1273	0.1241	0.1093	0.1062	0.1418	0.1351	0.1079	0.1014
F-Statistic		5.456		2.097		1.914		1.817		4.075		3.803
Pr (F<0)		<.05		<.25		<.25		<.25		<.05		0.05

Table 1.0
Variable Definitions

afqt	Decimal score (in range of [0,1]) on Armed Forces Qualifying Test taken in 1980 (unofficial).
wkwkat*XX*	Number of weeks worked in calendar year when person was *XX* years of age.
cohort*XX*	Dummy for birth cohort (1 if person was born in year 19*XX*).
black	1 if black, =0 otherwise.
hispanic	1 if Hispanic, =0 otherwise.
momhsgrd	Dummy for education of mother of respondent (1 if mother graduated from high school, =0 otherwise).
hgcby30	Highest grade level respondent had completed by age 30.
rhwage*XX*	Real hourly wage of respondent received at age *XX*.
lrwage*XX*	Natural log of rhwage*XX*.
cgww*XXYY*	Year-to-year change in weeks worked (Weeks worked at age *XX* - Weeks worked at age *YY*).
cglw*XXYY*	Year-to-year change in log of real hourly wages (lrwage*XX* - lrwage*YY*).
wkxafq*YY*	Interaction Effect: (Weeks worked at age YY * AFQT score).
wkxmhs*YY*	Interaction Effect: (Weeks worked at age YY * momhsgrd).
wkxblk*YY*	Interaction Effect: (Weeks worked at age YY * black).
dwxafq*XX*	Year-to-year change in interaction effect (wkxafq*XX* - wkxafq*YY*).
dwxmhs*XX*	Year-to-year change in interaction effect (wkxmhs*XX* - wkxmhs*YY*).
dwxblk*XX*	Year-to-year change in interaction effect (wkxblk*XX* - wkxblk*YY*).
nowkat*XX*	1 if individual worked no weeks at age *XX*, =0 otherwise.
cgnww*XXYY*	Year-to-year change in not working any weeks (nowkat*XX* - nowkat*YY*).

Table 1.1
Selection of Data Groups and Number of Observations

Group Restriction	All	Graduates	Dropouts
1. All in NLSY	12686		
2. All Males	6403		
3. Males in Birth Cohorts 1960-1964	3461		
4. (3) who reported highest grade compl eted by age 30	2805	1279	484
5. (4) who reported AFQT score	2684	1221	448
6. (4) who reported highest grade completed (HGC) by mother	2613	1182	424
7. (4) reporting AFQT and Mother's HGC	2506	1131	393
8. (7) reporting wkwkat19, wkwkat20	2453	1106	377
9. (7) reporting wkwkat29, wkwkat 30	2439	1106	377
10. (7) reporting weeks worked at 19,20, 29, and 30	2389	**1082**	**362**
11. (10) reporting positive weeks worked at 29, 30	2127	957	305
12. (11) reporting hourly wages at 29, 30	1971	891	275
13. (12) with real wages at 29, 30 >= $1.00	1950	880	271
14. (13) with real wages at 29, 30 <= $100.00	1930	**873**	**268**
15. (14) with positve weeks worked at 19, 20	1517	660	211
16. (15) reporting hourly wages at 19, 20	916	404	136
17. (16) with real wages at 19, 20 >=$1.00	904	400	134
18. (17) with real wages at 19, 20 <=$100.00	904	**400**	**134**

Note: Numbers of observations in bold correspond to our six data groups.

Table 1.2
Descriptive Statistics For Data Groups
(Note: Data are weighted based on NLSY 1979 sample weight)

	Group I: All High School Dropouts (N=362)				Group II: All High School Graduates (N=1,082)			
Variable	**Mean**	**Std Dev**	**Min**	**Max**	**Mean**	**Std Dev**	**Min**	**Max**
afqt	.15	.138	.01	.64	.359	.236	.01	.99
wkwkat19	23	18.5	0	52	27.8	19.2	0	52
wkwkat20	28.8	18.8	0	52	31.1	18.8	0	52
wkwkat29	39.8	18.1	0	52	43.2	15.8	0	52
wkwkat30	38.8	19	0	52	43	16.1	0	52
black	0.199	0.399	0	1	0.175	0.379	0	1
hispanic	0.131	0.337	0	1	0.065	0.247	0	1
momhsgrd	0.405	0.491	0	1	0.634	0.481	0	1
hgcby30	9.75	1.22	5	11	12	0	12	12
cohort60	0.195	0.396	0	1	0.232	0.422	0	1
cohort61	0.22	0.414	0	1	0.214	0.41	0	1
cohort62	0.258	0.437	0	1	0.209	0.407	0	1
cohort63	0.233	0.423	0	1	0.257	0.437	0	1
cohort64	0.094	0.292	0	1	0.088	0.283	0	1

	Group III: High School Dropouts In Workforce (N=134)				Groups IV: High School Graduates In Workforce (N=400)			
Variable	**Mean**	**Std Dev**	**Min**	**Max**	**Mean**	**Std Dev**	**Min**	**Max**
afqt	.169	.143	.01	.59	.394	.238	.01	.97
wkwkat19	31.5	17.4	2	52	38	15.1	3	52
wkwkat20	39.9	13.8	4	52	43.2	12.4	2	52
wkwkat29	47.9	9.09	3	52	48.1	8.7	3	52
wkwkat30	47.5	9.79	1	52	48.6	7.97	2	52
cohort60	0.147	0.354	0	1	0.208	0.406	0	1
cohort61	0.194	0.395	0	1	0.166	0.372	0	1
cohort62	0.221	0.415	0	1	0.217	0.412	0	1
cohort63	0.307	0.461	0	1	0.305	0.46	0	1
cohort64	0.132	0.338	0	1	0.105	0.307	0	1
black	0.146	0.354	0	1	0.09	0.286	0	1
hispanic	0.144	0.351	0	1	0.06	0.237	0	1
momhsgrd	0.395	0.489	0	1	0.69	0.464	0	1
hgcby30	9.7	1.17	6	11	12	0	12	12
rhwage19	7.02	4.61	1.2	35.1	6.13	2.02	1.12	22.6
rhwage20	8.11	7.01	1.45	55.2	6.85	2.89	1.26	30.9
rhwage29	10.1	4.11	1.55	40.8	12	5.05	2.41	50
rhwage30	10.82	7.01	2.13	53.6	12.2	5.67	1.3	51.1
lrwage19	1.84	0.42	0.184	3.56	1.76	0.334	0.11	3.12
lrwage20	1.95	0.466	0.373	4.01	1.86	0.362	0.233	3.43
lrwage29	2.24	0.393	0.436	3.71	2.4	0.419	0.879	3.91
lrwage30	2.25	0.482	0.754	3.98	2.41	0.429	0.26	3.93

Table 1.3
Descriptive Statistics For Data Groups
(Note: Data are weighted based on NLSY 1979 sample weight)

	Group V: High School Dropouts In Workforce at Ages 29 and 30 (N=268)				Groups VI: High School Graduates In Workforce at Ages 29 and 30 (N=873)			
Variable	**Mean**	**Std Dev**	**Min**	**Max**	**Mean**	**Std Dev**	**Min**	**Max**
afqt	.163	.144	.01	.61	.367	.238	.01	.99
wkwkat19	23.5	18.34	0	52	29	18.99	0	52
wkwkat20	31	18.06	0	52	32.6	18.56	0	52
wkwkat29	46.1	11.2	1	52	47.4	9.75	1	52
wkwkat30	46	11.07	1	52	47.7	9.21	1	52
cohort60	0.196	0.397	0	1	0.238	0.426	0	1
cohort61	0.214	0.41	0	1	0.211	0.408	0	1
cohort62	0.249	0.433	0	1	0.213	0.409	0	1
cohort63	0.233	0.423	0	1	0.251	0.433	0	1
cohort64	0.108	0.31	0	1	0.088	0.284	0	1
black	0.16	0.367	0	1	0.156	0.363	0	1
hispanic	0.137	0.344	0	1	0.063	0.243	0	1
momhsgrd	0.416	0.493	0	1	0.064	0.479	0	1
hgcby30	9.8	1.2	5	11	12	0	12	12
rhwage29	9.62	3.94	1.55	40.8	11.12	4.93	1.09	50
rhwage30	10.74	7.06	1.16	53.6	11.34	5.18	1.3	53.5
lrwage29	2.18	0.399	0.436	3.71	2.32	0.44	0.085	3.91
lrwage30	2.24	0.495	0.145	3.98	2.34	0.426	0.26	3.98

Figure 1.0. Probability Tree of Working At Least One Week at Ages 19-22 (High School Dropouts)

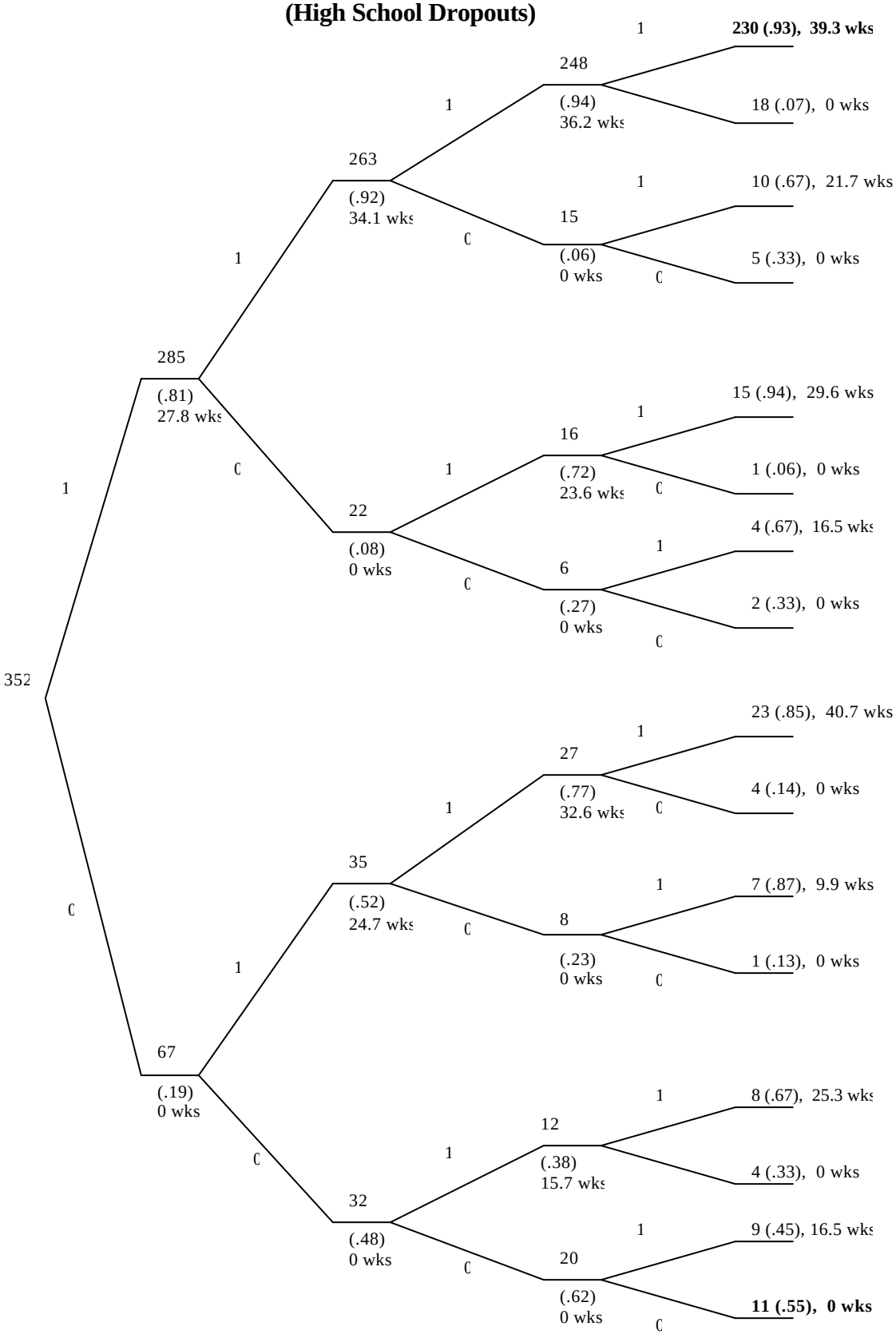

Figure 1.1. Probability Tree of Working At Least One Week at Ages 19-22 (High School Graduates)

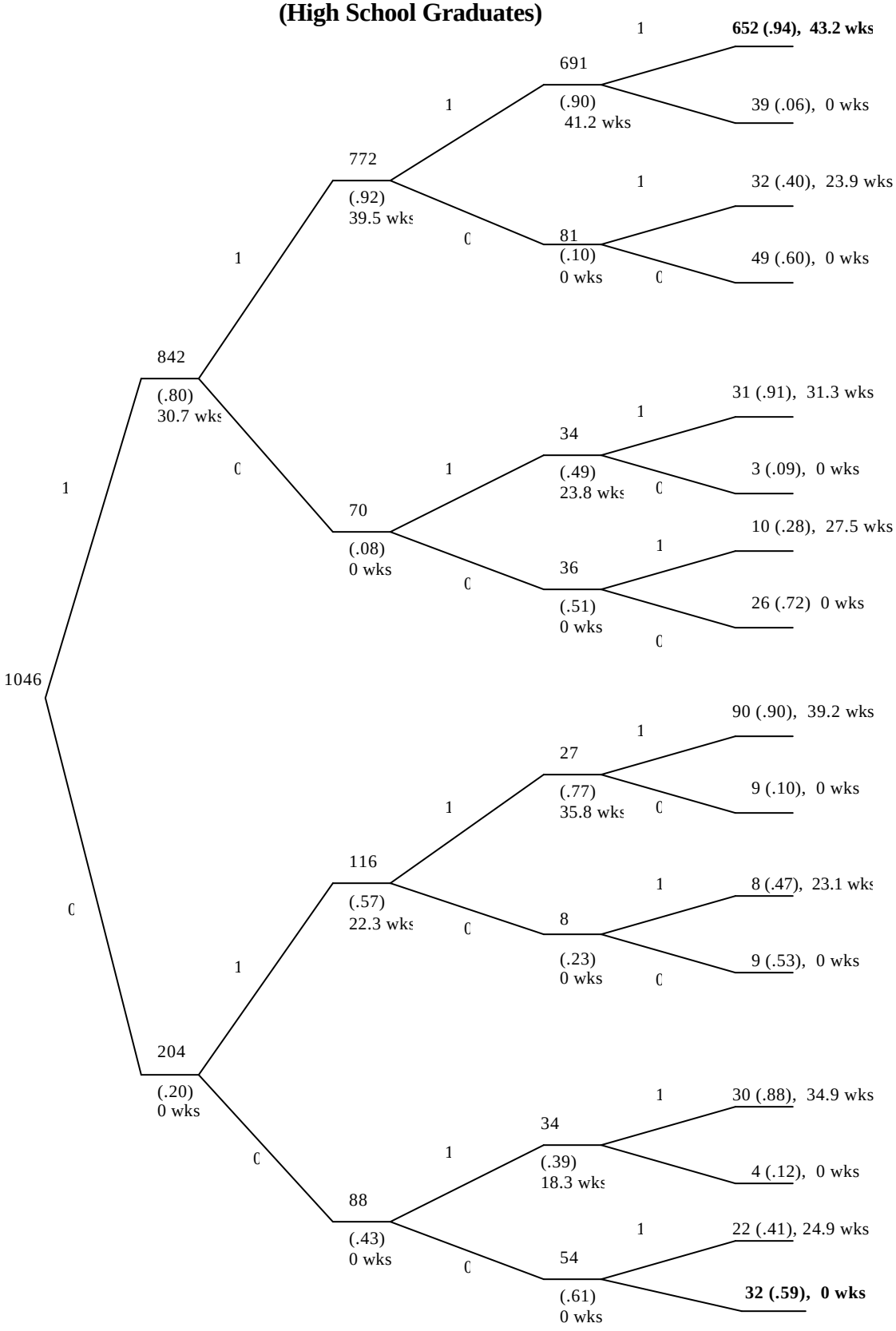

Figure 2.0. Probability Tree of Working At Least 40 Weeks at Ages 19-22 (High School Dropouts)

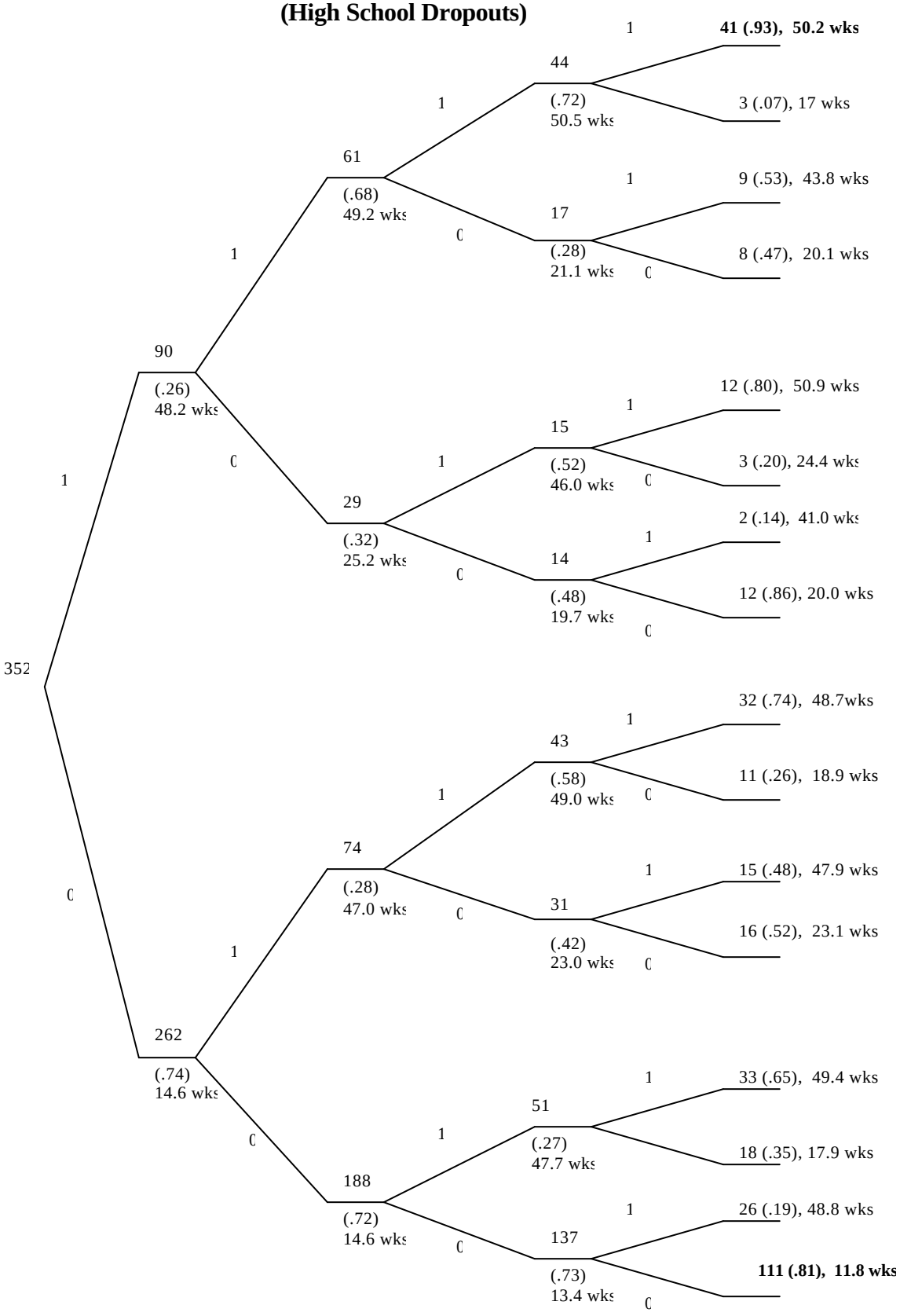

Figure 2.1. Probability Tree of Working At Least 40 Weeks at Ages 19-22 (High School Graduates)

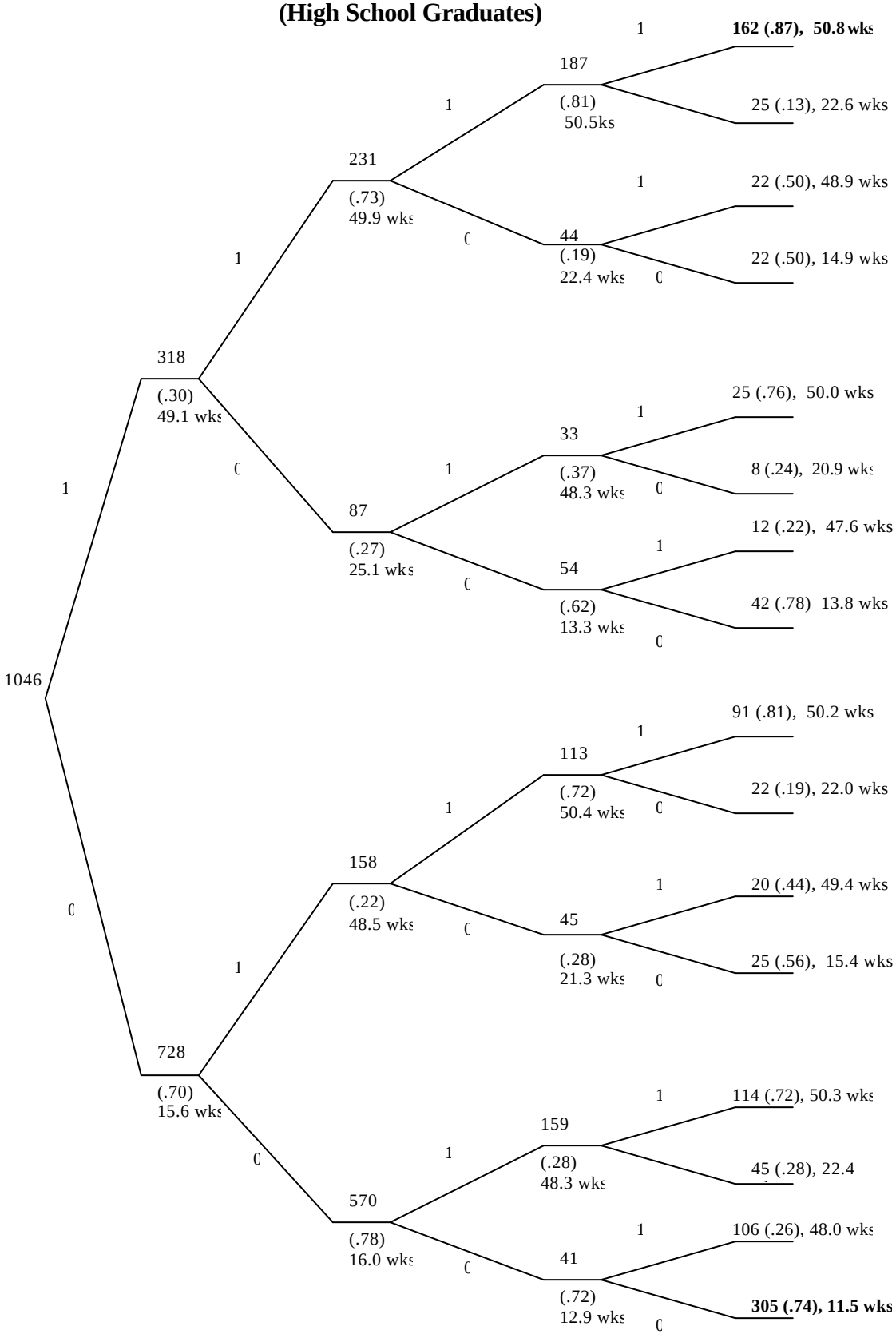

Table 2.1
Probability Tree Groupings for Dropouts Aged 19-22

	Was Employed– 1+wks/yr During:				40+wks/yr during:		
	Total Grp	**All 4 yrs**	**Some yrs**	**No Year**	**All 4 yrs**	**Some yrs**	**No Year**
NOOBS	352	230	111	11	41	200	111
afqt	0.151	0.173	0.101	0.113	0.195	0.159	0.119
black	0.195	0.143	0.334	0.14	0.057	0.19	0.26
hispanic	0.132	0.128	0.133	0.192	0.169	0.12	0.138
momhsgrd	0.408	0.455	0.311	0.255	0.149	0.459	0.336
cohort60	0.192	0.179	0.238	0.01	0.223	0.196	0.173
cohort61	0.221	0.245	0.184	0.045	0.242	0.216	0.221
cohort62	0.261	0.248	0.324	0.038	0.339	0.248	0.253
cohort63	0.233	0.229	0.177	0.727	0.182	0.204	0.308
cohort64	0.093	0.099	0.077	0.095	0.015	0.135	0.044
wkwkat19	22.8	28.8	10.8	0	49.3	24.1	9.54
wkwkat20	28.6	35.5	15.1	0	50	33.46	10.7
wkwkat21	30.1	37.4	15.7	0	50.5	35.9	10.8
wkwkat22	32	39.3	17.7	0	50.2	38.7	11.8
wkwkat29	39.7	44.2	32.2	11.6	47.4	43.7	29.1
wkwkat30	38.8	43.2	31.1	12.8	46.8	42.6	28.2
rhwage29	9.81 (296)	10.43 (208)	8.04 (84)	6.97 (4)	12.57 (38)	10.02 (183)	7.83 (75)
rhwage30	11.15 (294)	11.14 (210)	11.69 (78)	6.61 (6)	15.13 (41)	10.28 (173)	10.96 (80)

(Wages are averaged over all individuals who worked during year. Number of workers observed in parentheses.)

Table 2.2
Probability Tree Groupings for Graduates Aged 19-22

	Was Employed– 1+wks/yr During:				40+wks/yr during:		
	Total Grp	**All 4 yrs**	**Some yrs**	**No Year**	**All 4 yrs**	**Some yrs**	**No Year**
NOOBS	1046	652	362	32	162	579	305
afqt	0.360	0.386	0.313	0.231	0.421	0.364	0.301
black	0.174	0.121	0.275	0.454	0.071	0.156	0.309
hispanic	0.064	0.063	0.056	0.263	0.051	0.06	0.086
momhsgrd	0.634	0.675	0.552	0.542	0.762	0.634	0.529
cohort60	0.233	0.257	0.189	0.069	0.255	0.242	0.191
cohort61	0.213	0.203	0.24	0.133	0.2	0.215	0.221
cohort62	0.21	0.218	0.187	0.349	0.275	0.183	0.226
cohort63	0.256	0.243	0.281	0.345	0.207	0.271	0.262
cohort64	0.087	0.078	0.103	0.104	0.064	0.089	0.1
wkwkat19	27.8	34.6	14.6	0	49.8	26.4	13
wkwkat20	31	38.8	16	0	50.2	32.5	11.5
wkwkat21	33.2	42.2	15.6	0	50.8	36.8	9.56
wkwkat22	35.4	43.2	20.4	0	50.8	39.8	11.5
wkwkat29	43.2	46.1	38.3	21.8	48.8	44.7	34.8
wkwkat30	43.1	45.4	39.1	21	49	44.6	34.1
rhwage29	14.92 (922)	17.29 (601)	9.25 (304)	10.87 (17)	13.78 (157)	17.42 (521)	8.83 (244)
rhwage30	11.55 (932)	12.15 (605)	10.21 (308)	8.30 (19)	13.20 (157)	11.50 (528)	10.01 (247)

(Wages are averaged over all individuals who worked during year. Number of workers observed in parentheses.)

Table 3.0
Employment persistence, Level equations for High School Dropouts
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**
Dependent Variable:	**wkwkat29**	**wkwkat29**	**wkwkat29**
Regressors:			
Intercept	36.000 (2.326)	34.625 (2.726)	37.785 (3.053)
wkwkat19	0.269 (0.051)	0.246 (0.050)	0.171 (0.059)
nowkat19			-6.448 (2.867)
afqt		20.664 (7.183)	19.587 (7.157)
cohort61	-3.741 (2.890)	-5.086 (2.805)	-5.219 (2.790)
cohort62	-1.830 (2.773)	-2.716 (2.695)	-2.926 (2.682)
cohort63	-4.200 (2.823)	-5.175 (2.760)	-5.109 (2.744)
cohort64	-1.211 (3.664)	-2.107 (3.581)	-2.258 (3.561)
black		-5.182 (2.401)	-4.977 (2.389)
hispanic		-4.119 (2.736)	-3.846 (2.723)
momhsgrd		3.003 (1.962)	2.712 (1.956)
NOOBS	362	362	362
Dependent mean	39.795	39.795	39.795
R-squared	0.080	0.148	0.160

(Estimated by Ordinary Least Squares.)

Table 3.1
Employment persistence, Level equations for High School Dropouts
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**
Dependent Variable:	**wkwkat30**	**wkwkat30**	**wkwkat30**
Regressors:			
Intercept	31.389 (2.692)	29.644 (2.996)	31.763 (3.294)
wkwkat20	0.301 (0.051)	0.235 (0.052)	0.176 (0.065)
nowkat20			-5.317 (3.467)
afqt		27.705 (7.627)	28.268 (7.621)
cohort61	-1.830 (3.000)	-3.497 (2.915)	-3.553 (2.909)
cohort62	-2.002 (2.885)	-2.934 (2.812)	-2.574 (2.816)
cohort63	-2.156 (2.958)	-3.210 (2.902)	-2.958 (2.901)
cohort64	2.149 (3.825)	1.419 (3.743)	1.476 (3.736)
black		-3.087 (2.507)	-3.141 (3.503)
hispanic		-1.598 (2.862)	-1.155 (2.871)
momhsgrd		3.017 (2.052)	3.075 (2.049)
NOOBS	362	362	362
Dependent mean	39.795	38.824	38.824
R-squared	0.093	0.158	0.164

(Estimated by Ordinary Least Squares.)

Table 3.2
Employment persistence, Level equations for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**wkwkat29**	**wkwkat29**	**wkwkat29**
Regressors:			
Intercept	40.017 (1.215)	40.073 (1.625)	41.85 (1.83)
wkwkat19	0.150 (0.025)	0.118 (0.025)	0.077 (0.032)
nowkat19			-3.483 (1.652)
afqt		2.738 (2.240)	2.426 (2.241)
cohort61	-0.916 (1.414)	-1.046 (1.401)	-0.952 (1.399)
cohort62	-1.292 (1.422)	-1.177 (1.410)	-1.217 (1.408)
cohort63	-0.012 (1.352)	-0.043 (1.347)	-0.034 (1.345)
cohort64	-5.003 (1.877)	-4.605 (1.866)	-4.567 (1.863)
black		-5.090 (1.384)	-4.964 (1.383)
hispanic		-0.340 (1.983)	-0.408 (1.980)
momhsgrd		1.117 (1.031)	1.129 (1.029)
NOOBS	1082	1082	1082
Dependent mean	43.250	43.250	43.250
R-squared	0.042	0.056	0.068

(Estimated by Ordinary Least Squares.)

Table 3.3
Employment persistence, Level equations for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**wkwkat30**	**wkwkat30**	**wkwkat30**
Regressors:			
Intercept	36.228 (1.319)	34.572 (1.721)	35.248 (1.941)
wkwkat20	0.229 (0.025)	0.201 (0.026)	0.185 (0.033)
nowkat20			-1.362 (1.811)
afqt		6.523 (2.241)	6.509 (2.242)
cohort61	-0.070 (1.422)	-0.349 (1.411)	-0.341 (1.411)
cohort62	-1.833 (1.429)	-1.720 (1.411)	-1.733 (1.417)
cohort63	1.120 (1.361)	1.240 (1.357)	1.257 (1.358)
cohort64	-2.307 (1.882)	-1.720 (1.873)	-1.733 (1.873)
black		-2.662 (1.394)	-2.691 (1.395)
hispanic		-1.004 (1.984)	-1.012 (1.985)
momhsgrd		1.058 (1.033)	1.053 (1.033)
NOOBS	1082	1082	1082
Dependent mean	43.043	43.043	43.043
R-squared	0.077	0.097	0.097

(Estimated by Ordinary Least Squares.)

Table 3.4
Employment persistence, First Differences for High School Dropouts
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**
Dependent Variable:	**cgww3029**	**Cgww3029**	**cgww3029**
Regressors:			
Intercept	-1.170 (0.706)		-1.170 (0.704)
cgww2019	0.0340 (0.037)	0.016(0.035)	0.038 (0.040)
cgnw2019			0.540 (2.039)
NOOBS	362	362	362
Dependent mean	-0.971	-0.971	-0.971
R-squared	0.002	0.001	0.003

(Estimated by Ordinary Least Squares.)

Table 3.5
Employment persistence, First Differences for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**
Dependent Variable:	**cgww3029**	**cgww3029**	**cgww3029**
Regressors:			
Intercept	-0.168 (0.344)		-0.156 (0.344)
cgww2019	-0.012 (0.020)	-0.013 (0.020)	-0.031 (0.023)
cgnw2019			-1.759 (0.982)
NOOBS	1082	1082	1082
Dependent mean	-0.206	-0.206	-0.206
R-squared	0.001	0.001	0.003

(Estimated by Ordinary Least Squares.)

Table 3.6
Long-Term Effects of Weeks Worked on Wages,
Level Equations for High School Dropouts
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**	**IV**
Dependent Variable:	**lrwage29**	**lrwage29**	**lrwage30**	**lrwage30**
Regressors:				
Intercept	2.225 (0.108)	2.117 (0.118)	1.807 (0.170)	1.770 (0.184)
wkwkat19	-0.001 (0.002)	-0.001 (0.002)		
wkwkat20			0.008 (0.003)	0.007 (0.003)
afqt		0.752 (0.265)		0.844 (0.316)
cohort61	0.188 (0.118)	-0.083 (0.120)	0.298 (0.139)	0.185 (0.141)
cohort62	-0.061 (0.114)	-0.005 (0.114)	0.169 (0.135)	0.111 (0.136)
cohort63	-0.049 (0.108)	-0.107 (0.109)	0.051 (0.130)	-0.023 (0.136)
cohort64	0.002 (0.130)	-0.068 (0.130)	0.082 (0.154)	-0.006 (0.158)
black		-0.023 (0.100)		0.029 (0.123)
hispanic		0.059 (0.097)		0.003 (0.115)
momhsgrd		0.005 (0.078)		0.052 (0.093)
NOOBS	134	134	134	134
Dependent mean	2.237	2.237	2.254	2.254
R-squared	0.046	0.129	0.115	0.180

(Estimated by Ordinary Least Squares.)

Table 3.7
Wage Persistence, Level Equations for High School Dropouts
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**	**IV**
Dependent Variable:	**lrwage29**	**Lrwage29**	**lrwage30**	**lrwage30**
Regressors:				
Intercept	2.237 (0.173)	2.079 (0.182)	2.430 (0.211)	2.323 (0.217)
lrwage19	-0.019 (0.082)	0.0164(0.080)		
lrwage20			-0.136 (0.092)	-0.129 (0.090)
afqt		0.762 (0.268)		0.879 (0.318)
cohort61	0.185(0.118)	0.078 (0.120)	0.303 (0.141)	0.177 (0.142)
cohort62	0.061 (0.114)	0.005 (0.114)	0.199 (0.138)	0.121 (0.137)
cohort63	-0.051 (0.107)	-0.107 (0.109)	-0.040 (0.130)	-0.117 (0.131)
cohort64	0.007 (0.129)	-0.067 (0.130)	-0.019 (0.158)	-0.111 (0.158)
black		-0.020 (0.100)		-0.047 (0.120)
hispanic		-0.059 (0.097)		-0.011 (0.116)
momhsgrd		0.047 (0.077)		0.069 (0.933)
NOOBS	134	134	134	134
Dependent mean	2.237	2.237	2.254	2.254
R-squared	0.045	0.130	0.083	0.166

(Estimated by Ordinary Least Squares.)

Table 3.8
Wage Persistence and Long-Term Effect of Employment on Wages,
First Differences for High School Dropouts

(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	**I**	**II**	**III**	**IV**
Dependent Variable:	**cglw3029**	**cglw3029**	**cglw3029**	**cglw3029**
Regressors:				
Intercept	0.021 (0.034)		0.008 (0.037)	
cglw2019	-0.040 (0.058)	-0.033 (0.057)		
cgww2019			0.0011 (0.0020)	0.0013 (0.0018)
NOOBS	134	134	134	134
Dependent mean	0.017	0.017	0.017	0.017
R-squared	0.004	0.003	0.002	0.004

(Estimated by Ordinary Least Squares.)

Table 3.9
Wage Persistence, Level Equations for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III	IV
Dependent Variable:	**lrwage29**	**lrwage29**	**lrwage30**	**lrwage30**
Regressors:				
Intercept	2.021 (0.120)	1.875 (0.125)	2.183 (0.131)	2.083 (0.135)
lrwage19	0.251 (0.062)	0.235 (0.060)		
lrwage20			0.151 (0.061)	0.131 (0.060)
afqt		0.278 (0.094)		0.162 (0.098)
cohort61	0.043 (0.067)	0.044 (0.066)	-0.010 (0.070)	0.012 (0.069)
cohort62	-0.102 (0.063)	-0.093 (0.061)	-0.037 (0.066)	-0.036 (0.065)
cohort63	-0.125 (0.058)	-0.110 (0.057)	-0.091 (0.062)	-0.086 (0.061)
cohort64	-0.086 (0.078)	-0.068 (0.076)	-0.151 (0.081)	-0.145 (0.080)
black		-0.094 (0.074)		-0.141 (0.077)
hispanic		-0.054 (0.086)		-0.032 (0.090)
momhsgrd		0.098 (0.045)		0.125 (0.047)
NOOBS	400	400	400	400
Dependent mean	2.401	2.401	2.414	2.414
R-squared	0.071	0.127	0.039	0.068

(Estimated by Ordinary Least Squares.)

Table 3.10
Wage Persistence, First Differences for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II
Dependent Variable:	**cglw3029**	**cglw3029**
Regressors:		
Intercept	0.017 (0.020)	
cglw2019	-0.038 (0.044)	-0.029 (0.043)
NOOBS	400	400
Dependent mean	0.014	0.014
R-squared	0.002	0.001

(Estimated by Ordinary Least Squares.)

Table 3.11
Long-Term Effects of Weeks Worked on Wages,
Level Equations for High School Graduates

(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**lrwage29**	**lrwage29**	**lrwage29**
Regressors:			
Intercept	2.307 (0.073)	2.142 (0.087)	2.241 (0.120)
wkwkat19	0.0041 (0.0014)	0.0036 (0.0014)	0.0011 (0.0025)
afqt		0.296 (0.095)	0.258 (0.245)
wkxafqat19			0.0070 (0.0059)
cohort61	0.036 (0.068)	0.037 (0.066)	0.033 (0.066)
cohort62	-0.108 (0.063)	-0.098 (0.062)	-0.097 (0.062)
cohort63	-0.117 (0.059)	-0.104 (0.058)	-0.105 (0.058)
cohort64	-0.098 (0.078)	-0.080 (0.077)	-0.092 (0.078)
black		-0.083 (0.075)	-0.089 (0.075)
hispanic		-0.026 (0.087)	-0.029 (0.087)
momhsgrd		0.098 (0.045)	0.101 (0.045)
NOOBS	400	400	400
Dependent mean	2.401	2.401	2.401
R-squared	0.052	0.109	0.113

(Estimated by Ordinary Least Squares.)

Table 3.12
Long-Term Effects of Weeks Worked on Wages,
Level Equations for High School Graduates

(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**lrwage30**	**lrwage30**	**lrwage30**
Regressors:			
Intercept	2.347 (0.101)	2.229 (0.109)	2.350 (0.154)
wkwkat20	0.0029 (0.0018)	0.0023 (0.0018)	-0.00061 (0.0032)
afqt		0.168 (0.098)	-0.201 (0.348)
wkxafqat20			0.0084 (0.0076)
cohort61	0.012 (0.070)	0.014 (0.069)	0.017 (0.069)
cohort62	-0.056 (0.066)	-0.053 (0.064)	-0.048 (0.064)
cohort63	-0.102 (0.063)	-0.096 (0.062)	-0.090 (0.062)
cohort64	-0.143 (0.084)	-0.138 (0.082)	-0.136 (0.082)
black		-0.142 (0.078)	-0.146 (0.078)
hispanic		-0.039 (0.090)	-0.040 (0.090)
momhsgrd		0.126 (0.047)	0.130 (0.047)
NOOBS	400	400	400
Dependent mean	2.414	2.414	2.414
R-squared	0.030	0.060	0.084

(Estimated by Ordinary Least Squares.)

Table 3.13
Long-Term Effects of Weeks Worked on Wages,
First Differences for High School Graduates

(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**cglw3029**	**cglw3029**	**cglw3029**
Regressors:			
Intercept	-0.0039 (0.0203)		-0.0051 (0.020)
cgww2019	0.0034 (0.0014)	0.0033 (0.0013)	-0.00098 (0.0025)
dwxafq20			0.011 (0.0052)
NOOBS	400	400	400
Dependent mean	0.014	0.014	0.014
R-squared	0.015	0.017	0.026

(Estimated by Ordinary Least Squares.)

Table 3.14
Long-Term Effects of Weeks Worked (and No Weeks Worked) on Wages,
Level Equations for High School Graduates

(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**lrwage29**	**lrwage29**	**lrwage29**
Regressors:			
Intercept	2.191 (0.038)	2.059 (0.049)	2.08 (0.054)
wkwkat19	0.0060 (0.0007)	0.0049 (0.0007)	0.004 (0.001)
nowkat19			-0.046 (0.051)
afqt		0.380 (0.066)	0.378 (0.066)
cohort61	-0.025 (0.043)	-0.029 (0.042)	-0.280 (0.419)
cohort62	-0.058 (0.043)	-0.048 (0.042)	-0.485 (0.042)
cohort63	-0.084 (0.043)	-0.066 (0.040)	-0.066 (0.040)
cohort64	-0.116 (0.057)	-0.086 (0.056)	-0.085 (0.056)
black		-0.033 (0.043)	-0.030 (0.043)
hispanic		-0.006 (0.060)	0.005 (0.060)
momhsgrd		0.030 (0.031)	0.030 (0.031)
NOOBS	873	873	873
Dependent mean	2.316	2.316	2.316
R-squared	0.080	0.128	0.129

(Estimated by Ordinary Least Squares.)

Table 3.15
Long-Term Effects of Weeks Worked (and No Weeks Worked) on Wages, Level Equations for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**lrwage30**	**lrwage30**	**lrwage30**
Regressors:			
Intercept	2.184 (0.040)	2.076 (0.506)	2.094 (0.057)
wkwkat20	0.0050 (0.0077)	0.004 (0.001)	0.004 (0.001)
nowkat20			-0.037 (0.055)
afqt		0.341 (0.064)	0.340 (0.064)
cohort61	0.010 (0.042)	0.002 (0.041)	0.003 (0.041)
cohort62	0.005 (0.042)	0.011 (0.041)	0.011 (0.041)
cohort63	-0.031 (0.041)	-0.019 (0.040)	-0.018 (0.040)
cohort64		-0.050 (0.054)	-0.051 (0.054)
black		-0.099 (0.042)	-0.100 (0.042)
hispanic		0.017 (0.058)	0.016 (0.058)
momhsgrd		0.044 (0.030)	0.043 (0.030)
NOOBS	873	873	873
Dependent mean	2.339	2.339	2.339
R-squared	0.030	0.116	0.116

(Estimated by Ordinary Least Squares.)

Table 3.16
Long-Term Effects of Weeks Worked (and No Weeks Worked) on Wages, First Differences for High School Graduates
(Estimated standard errors in parentheses. NOOBS = Number of observations.)

Model	I	II	III
Dependent Variable:	**cglw3029**	**cglw3029**	**cglw3029**
Regressors:			
Intercept	0.019 (0.013)		0.019 (0.013)
cgww2019	0.0009 (0.0008)	0.0011 (0.0008)	0.0011 (0.0009)
cgnw2019			0.017 (0.038)
NOOBS	873	873	873
Dependent mean	0.023	0.023	0.023
R-squared	0.002	0.003	0.002

(Estimated by Ordinary Least Squares.)