Labor Market Impacts of School Interruptions:
A Dynamic, Structural Approach

Liang Guo

A Thesis

In the Department

of

Economics

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Economics) at

Concordia University

Montreal, Quebec, Canada

June 2018

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Liang Guo

Entitled: Labor Market Impacts of School Interruptions: A Dynamic, Structural Approach

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy (Economics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________ Chair
Dr. Amy Swiffen

______________________ External Examiner
Dr. Nicolai Kristensen

______________________ External to Program
Dr. Michel Magnan

______________________ Examiner
Dr. Tatyana Koreshkova

______________________ Examiner
Dr. Damba Lkhagvasuren

______________________ Thesis Supervisor (s)
Dr. Jorgen Hansen

____

Approved by ______________________
Dr. Damba Lkhagvasuren, Graduate Program Director

July 10th, 2018
Date of Defence

____
Dr. André Roy, Dean, Faculty of Arts and Science

ABSTRACT

Labor Market Impacts of School Interruptions: A Dynamic, Structural Approach

Liang Guo, Ph.D.

Concordia University, 2018

Schooling interruptions are common in North America and have recently attracted the attention of economists. In this paper, I provide structural estimates of a dynamic model of education and labor supply to investigate the impact of schooling interruptions on subsequent wage and employment outcomes. In this model, schooling interruptions evolve endogenously, and I approximate the expected value function using an approach suggested by Geweke and Keane (1996) that significantly reduces computational time. The results indicate that a temporary interruption to schooling attainment hurts post-graduation income relative to continuous investment in education, and consequently on the life-cycle income. However, even with an interruption, there are still significant wage gains from college and university relative to high school. Finally, the costs of interruptions and wage gains are heterogeneous and vary significantly across different racial groups and types of agents at both waves of NLSY. Using policy simulation, I find that under a post-secondary schooling subsidy policy, up to 4 percent more white males from cohort 1979 will make a choice to return to school, 9.6 percent more of them will obtain a university-level degree, and the hourly wage will also increase up to 4.1 percent. Under the same post-secondary schooling subsidy policy simulation, the white males from 1997 cohort been more affected by the subsidy than their earlier cohort. At each cohort, black and Hispanics are more affected by the policy.

Acknowledgment

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Jorgen Hansen for guiding and supporting me over the years, for your excellent patience, motivation, and immense knowledge. You guide me in all the time of research and writing of this thesis. You have set an example of excellence as a researcher, mentor, instructor, and role model.

Besides Prof. Jorgen Hansen, I would like to thank the rest of my thesis committee: Prof. Damba Lkhagvasuren, Prof. Tatyana Koreshkova, Prof. Michel Magnan (John Molson School of Business) and Prof. Nicolai Kristensen (VIVE - The Danish Centre for Applied Social Science) for all your guidance through this process and invaluable discussion, ideas, and feedback.

My sincere thanks also go to Prof. Christian Belzil, Prof. Xingfei Liu and Prof. Xingtong Han who provided me with their insightful comments and encouragement.

Finally, I would like to thank and dedicate this thesis to my parents. I undoubtedly could not have done this without your constant encouragement and support.

Table of Contents

List of Tables

List of Figures

1 Introduction

A fundamental result of the conventional human capital model is that agents have no incentive to delay or interrupt their educational investment, since the opportunity cost increases with age, and the time horizon when returns on the investment are realized is reduced. Nevertheless, the incidence of schooling interruptions is widespread, both in the U.S. and elsewhere. For example, in one of the first economic studies of this topic, Light (1995) used data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79) and reported that around 34 percent of white males experience at least one interruption to their educational attainment (regardless of the timing of the interruption). More recently, from North american data, Johnson (2013) shows that roughly 25 percent of young Americans take a break from schooling after completing high school and before enrolling in college or university. Similar patterns have been observed in other countries.[1]

There are many potential reasons for interrupting educational attainment. The conventional explanation is that students interested in pursuing a college education may face financial constraints and are unable to enroll in a post-secondary education directly after high school. Instead, they need to work a year or more to save up funds for their higher education. This issue has gained attention recently following the significant increases in direct costs of attending college and university in the U.S. and is explored in detail in Johnson (2013). Alternatively, students may not be immediately interested in continuing their education following high school graduation but realize after a period of non-enrollment that they would like to return, either to improve their labor market prospects or because of a direct benefit from the activity itself. In this paper, I seek to determine the impact of interrupting educational investments on sub-

[1] See for example Foley and Groes (2016), Dominic and Netz (2011) for Europen evidence and Finnie and Qiu (2008),Tomkowicz and Bushnik (2003)for Canadian experiences.

sequent wage outcomes. As mentioned above, a mechanism illustrated in the human capital literature is that any delay in investing in education will reduce the overall life-cycle benefit, as it reduces the time available to take advantage of the investment. That is, compared to continuous investment, an interruption is predicted to reduce the benefit of the obtained education. However, the effect does not necessarily have to be negative. For example, if an individual is gaining skills and experience during the interruption that complements the skills acquired during college, the effect may be positive. Students who return to school may also be more motivated, interested and engaged in their studies compared to those who continuously invest and thereby enjoy more significant returns. Further, as was also mentioned above, working during the interruption period may enable individuals to enroll in college; something they would otherwise not be able to do due to financial constraints. If we consider labor market outcomes for those with a high school degree only with those who return after an interruption and graduate from college, we obtain another measure of the benefit of returning to school. Thus, the potential benefit of an interruption depends on the comparison group, and there are two groups we need to consider when evaluating the returns to returning to school: those who achieve the same level of education but without interruption and those who choose not to return to school. Even if the life-cycle income for the returners does not entirely match the one for the first comparison group, it may well exceed that of the second group. Hence, the effect of interruption may be positive, both when holding total years of education constant and when considering the opportunity it gives to further invest in education. Regardless of the reasons for, and the potential effects of, an interruption, it is reasonable to believe that the decision is intrinsically linked to, and plays a significant role in, the overall human capital accumulation process. It is, therefore, necessary to include this option as a decision variable whose outcome is endogenously related to labor market outcomes, such as wages and labor supply. For example, a negative wage or employment shock for someone who finished high school but did not continue beyond that level may incentivize that person to return to school, an option he would not have exercised in the absence of the shock if his optimal choice then was to enter the labor market after high school. Thus, there may be a selectivity issue involved which needs to be considered when estimating the returns to returning.

Even though schooling interruptions are frequent, the economic literature on this topic is sparse. One of the first papers to investigate both the incidence of interruption as well as its effect on wages is Light (1995). Using data from the NLSY79 for the period 1979 to 1989, she estimated wage regressions allowing for different growth rates before and after students' re-enroll. Her results suggest that those who temporarily interrupted their educational attainment had slightly lower wages than their counterparts who acquired their education continuously. A more recent paper by Johnson (2013) also finds that a significant share of individuals delays their post-secondary education. He also used the NLSY but data from the younger 1997 cohort instead and reports that a one-year delay in college enrolment may cost agents more than $9,000 in lost lifetime income. He also finds that those who delayed their schooling were generally from low-income families and belonged to minority ethnic groups. However, unlike Light (1995), he does not consider the possibility that wage growth may differ between those who invest continuously and those who interrupt. In this thesis, I expand the analyses of both Light (1995) and Johnson (2013) in several critical dimensions. First, I develop an economic model for educational choices and labor supply with forward-looking behavior where the decision to interrupt is an integral part and where the wage function is defined in a very flexible fashion to allow for different growth rates before and after re-enrollment. Secondly, I introduce unobserved heterogeneity in a flexible way to address the potential selectivity and endogeneity issues discussed above. While Johnson (2013) also considers these issues, they are absent in Light (1995). Thirdly, I expand the time horizon in my model using the 1979 cohort to consider data collected up to, and including 2010. It is again an extension of Light (1995) who did not utilize data after 1989 and the adoption of a long panel (up to 30 periods) is also different from Johnson (2013). These expansions provide a general framework for analyzing the long-term impact on wages from schooling interruptions where the decisions to temporarily interrupt school is endogenous. Moreover, the use of a longer time horizon is critical to identify longer-term impacts of school interruptions. Fourth, I combine data from both the NLSY79 and NLSY97 to both describe if the incidence and duration of school interruptions have changed over time and to asses if the impact of interruption on wages and lifetime income has changed. Finally, using data from NLSY79 and NLSY97, I estimate the model separately for white, black and Hispanic males to analyze any racial differences in interruptions and their effects. Another unique aspect of the research conducted in this thesis is that it considers temporary disruptions to schooling at any level of study and is not limited to delayed college entries as much of the previous literature has focused on.

The NLSY data is well suited for the analysis in this thesis. In addition to providing data on labor market activities and outcomes over a long time horizon, it also contains detailed information on critical socio-economic characteristics, such as parental education and income, household composition and measures of cognitive ability. The data also allows me to follow each from the age of 16 when I start modeling choices (enroll in school, work or NEET (Not in Education, Employment or Training)) until age 33 or 46, depending on the cohort.[2]

[2]In the NLSY79, I observe individuals for up to 30 years while in the NLSY97, I can follow them for 17 years (due to the fact that they were initially interviewed in 1997). In the analysis that compares choices and outcomes of the two cohorts, I limited the NLSY79 data

Educational and labor supply decisions are generated from a model that has a finite horizon and is solved by backward recursions, where each chose the option in each period that maximizes the discounted present value of expected lifetime utility. I approximate the expected value function using an approach suggested by Geweke and Keane (1995) that significantly reduces computational time. The parameters of the model are estimated using Maximum Likelihood techniques and the model's predictive behavior is assessed by comparing actual and model-generated choices. I show that a temporary interruption to schooling has an adverse effect on an individual's future income compared to completing the same level of education without interruption. Precisely, the average difference in present value earnings for those without and with an interruption – assuming a one-year interruption between high school and college during which the individual is working – equals \$35,712. The impact of a gap year is quite heterogeneous, and the difference in present value earnings can be as low as \$5,000 over a 30-year period (for one group or type of individuals) and as high as \$48,000 over the same period (for another group). However, and as elaborated above, those who return to school do realize an economic benefit to that activity compared to those who do not return to school, even though it is smaller than the one that applies to those who acquire their education without any breaks. For example, based on NLSY79[3],the estimated average high school-college earnings difference is 26 percent without interruption and 12 percent with an interruption[4].

An important aspect and focus of this thesis is the inequality in educational outcomes and wages that exist across different ethnic groups in the U.S. This

---

to 17 periods.

[3]Data of NLSY79 White agents.

[4]In the NLSY97, the estimated average high school-college earning difference is 38 and 28 percent.

Again, this return varies across individuals, especially for those with an interruption where the college earnings premium is as low as 1.7 percent for some individuals.

is also a critical topic for economists with vast literature, primarily focused on black-white racial differences. However, attention to Hispanics, the third largest racial group in the United States, has been more limited. This research attempts to fill some of this gap by carefully analyzing racial differences in educational attainment, future economic outcomes, and labor supply as well as providing impact results from policy simulations designed to improve outcomes of the minority groups. The data from NLSY show that both the high school graduation rates and the university graduation rates differ enormously among these three major racial groups. For white males in the NLSY79, it is 85 percent, which is much higher than the corresponding rate for black males (72.5 percent) and Hispanic males (63 percent). The difference in the rate of university accomplishment based on the same data is even more significant. Approximately 26 percent of white males obtained a university degree compared to only 11 percent of black males and 12 percent of Hispanic males. Regarding school interruptions, data show that 36 percent of white males have at least one interruption which is less than the rate for both black males (close to 40 percent) and Hispanic males (39 percent). It indicates a possible relationship between the incidence of school interruptions and family income and access to financial support (which on average is much lower among blacks and Hispanics than among whites). When analyzing how the effect of a temporary interruption to schooling adversely impacts the present value of future earnings, I find that the differential is $21,586 for white males, $23,419 for black males and $15,329 for Hispanic males[5].

Base on NLSY79, the college premia for those without (with) and interruption are: 26.3 (12.8) percent for whites, 49.0 (32.7) percent for blacks and 50.1 (22.6) percent for Hispanics. Thus, the economic returns to a college degree for any of the two minority groups are substantial, regardless of the path towards that degree. Moreover, the racial gaps in hourly income at age 32 are substan-

[5]It based on NLSY79 (at age 49).

tial for those without any education beyond high school: close to 28 percent for the white-black difference and 25 percent for the white-Hispanic difference. Among college-educated workers, the corresponding numbers are 8-9 percent for the white-black earnings gap and 9-15 percent for the white-Hispanic wage gap.

The similar trend is found in NLSY97, black and Hispanics agents do benefit more than their white counterpart does from a college degree weather with or without interruption. The college premia for those without (with) and interruption are: 18.0 (12.0) percent for whites, 33.1 (23.2) percent for blacks and 34.1 (33.3) percent for Hispanics. From NLSY97, the racial gaps in wage at age 32 are substantial as in the NLSY79, for those without any education beyond high school: close to 29 percent for the white-black difference and 36 percent for the white-Hispanic difference. Among college-educated workers, the corresponding numbers are 14-17 percent for the white-black earnings gap and 14-19 percent for the white-Hispanic wage gap.

As is well known, an essential advantage of estimating a structural model of education and labor supply is that it provides an opportunity to conduct counter-factual policy simulations and analyze impacts of the policies on labor market outcomes. In this thesis, I consider two alternative policies, both with empirical relevance. The first is a policy where students are not allowed to leave school until they have graduated from high school and the second one is the introduction of an economic subsidy, conditional on college enrollment. Both policies may impact educational attainment and the decisions to interrupt and return to school. The forced high school graduation will eliminate the decision to drop out of high school and the possibility of returning to complete a high school diploma. The subsidy may provide sufficient funds for some to decide to continue directly from high school to college by lowering the cost thereby reducing the incidence of interruptions. It may also increase the fraction of individuals returning to college after having left school, again by increasing the affordability of college.

The results suggest that subsidies will increase the average level of education, life-cycle income, and lead to a significantly higher post-secondary graduation rate (a 17 percent increase in the graduation rate of four-year universities). The rate of school interruptions increases by about 4 percent. There are also critical racial dimensions to the estimated impacts. For instance, blacks and Hispanics from the 1979 cohort were twice as likely to be affected by the college subsidy. Average education among black males increases with three years to a total of 14 years (corresponding to an associated degree). Hispanic males gain a 2.5-year increase in average education and also reach 14. Also, importantly, the differences between wages and unemployment rates between whites, blacks and Hispanics are almost eliminated. Similar results were also found for the NLSY97 cohort.

The rest of this thesis is organized as follows. In the next section, I present a literature review that will describe in more detail some of the pertinent research in this area and define the context in which this thesis should be viewed. The structural, dynamic model is presented in Section 3 where I describe the specifications of the instantaneous utility functions, the evolution of the state variables and the individual's optimization problem. In Section 4, I present the data including sample selections, variable definitions and descriptive statistics and various features of the data. The estimation strategy is illustrated in the following section with a derivation of the likelihood function. Sections 6 contain presentations of the results. The model fit, counterfactual policy simulations are described and discussed in Section 7, and Section 8 concludes this thesis with a summary.

2 Previous Literature

2.1 School Interruptions and Delayed College Entry

Temporary interruptions to schooling are common in many parts of the world and have been a feature of human capital investment paths for many years. As such, it has also been the focus of research by economists who have attempted to understand the reasons behind these events. The primary explanation is due to the possibility of economic constraints preventing individuals from acquiring education in a continuous fashion, in particular, higher education. Consequently, much of the existing literature in this area has focused on interruptions following the completion of high school. For example, Heckman (1996) used the NSLY79 data to show that individuals are more likely to interrupt their college education in states with higher tuition fees.

However, data from NLSY79 also show that although the majority of interruptions occur after high school, many take place following the completion of other grades (from grade 9 up to grade 19). It suggests that other reasons may also contribute to these interruptions, as students are less likely to be financially constrained while attending grades 10 and 11. Further, the fact that these interruptions are also a standard feature in the Scandinavian countries, where higher education is provided at little cost, suggests that this is an issue that arises because of other (non-financial) reasons as well.

An early study on this topic that considered interruptions at different grade levels was conducted by Light (1995). Using data for the period from 1979 to 1989 from the NLSY79, she examined the wage effects of a series of non-traditional enrollment patterns. Specifically, she estimated a wage regression that allowed individuals to follow a different wage path before and after their re-enrollment. Using a Hausman and Taylor procedure to account for the potential endogeneity of education, she estimates coefficients that are associated both with schooling received before an interruption takes place (if at all) and with schooling that is obtained after an interruption (the second enrollment spell). Her results suggest that the decision to return to school is only related to time-invariant personal characteristics and not with transitory wage shocks (something that has been found to be relevant in more recent studies see discussion of Johnson (2013) below). Further, the return to a certain level of education is smaller for those who interrupt their schooling compared to those who acquired their education without any breaks and the more prolonged students wait before re-enrolling, the smaller is the return. Light (1995) also reports that although there exists a significant initial wage gap between interrupters (or returners) and non-interrupters, this wage difference gets smaller over time due to faster wage growth among the returners.

Also, using the NLSY79 but with a slightly more extended observation period than Light (1995), Monk (1997) studied the age at the time of college completion. The paper investigated the importance of the timing of schooling in determining earnings and reported that those who complete college at a later age receive a significantly smaller initial increase in earnings than those who acquire their education earlier in life. Similar to Light (1995) but unlike the analysis in this thesis, Monk (1997) did not consider or control for activities during the (schooling) gap years.

Stratton et al. (2008), based on the American 1990/94 Beginning Postsecondary Survey (BPS:90/94), analyzed factors related to three first-year enrollment choices: continuous schooling, temporary interruption (less than one year), and a long-term interruption to schooling (one year or more). They used a multinomial logit model to study which factors may explain the different choices made by the agents. They found significant differences between the factors associated with the short-term and long-term interruptions to schooling and argued that the probability of a long-term interruption is higher for those who receive loans and lower for those who receive work-study government aid as compared to those who receive no aid. Further, students with less educated parents (high school or less) are less likely to enroll continuously in college and more likely to interrupt (and dropout). A similar finding was reported for students with low GPAs.

In contrast to the results in Light (1995) and others as well as somewhat counter-intuitive, Ferrer and Menendez (2014) found that graduates (both college and university) who temporarily interrupted their education received a lifetime income premium compared to graduates who continuously invested in school, even after considering other factors such as work experience or labour market connections. Their results are based on data extracted from the 1995 cohort of the Canadian National Survey of Graduates which collects data on graduates in 1997 and 2000. Thus, the wages observed are those associated with labor market activities within the first few years of graduation. The authors speculate that higher wages for interrupters relative to the non-interrupters are due to the possibility for the former group to uncover some of the uncertainty about future labor market outcomes and consequently make a more informed educational decision than those who continue directly from high school.

Using data from the same survey (The Canadian National Graduate Survey) as Ferrer and Menendez (2014) but a more recent cohort (the 2007 cohort), Fortin and Ragued (2017) also argued that the delay of schooling is not necessarily associated with a penalty on future income. They investigate how temporary schooling interruptions affect future wages and how future outcomes depend on the activity during, and reasons for, a temporary interruption to schooling. They found a positive effect on wages from a temporary schooling interruption for males who worked full-time during the gap period but an adverse wage effect if the temporary schooling interruption was associated with health issues. Both Ferrer and Menendez (2014) and Fortin and Ragued (2017) analyze the impact of interruptions on wages soon (within two years) after graduation and are unable to explore any long-term effects. Further, the use of wages within the first few years after graduation may bias the wage impact of an interruption in favor of those who interrupt if the job-matching process takes time and those with previous work experience can obtain jobs that match their qualifications faster than those without such experiences.

In an exciting and relevant paper using Danish national population admission data for the years, 1981 to 2009, Foley and Gores (2012) also study the effects of a delay or a temporary interruption of education on income with a focus on differences across fields of study. They formulate a dynamic economic model whereby individuals form expectations about their future incomes before they decide whether or not to interrupt. They report overall adverse wage effects from interruptions, and since income profiles differ across fields, there is heterogeneity in the wage penalties across fields of study. It also impacts the incidence of interruptions. For example, students in humanities are three times more likely to delay university than engineering students, and the primary reason for this is that interruptions are costlier regarding lost earnings for engineering students than for humanities students. However, fields of study and financial status cannot entirely explain an agent's decision to interrupt schooling which is also determined by, in their data, unobserved variables (such as scholastic ability).

Some studies have investigated and focused on the relationship between the decision to delay schooling and the socio-economic background of students. For example, Wells and Lynch (2012) argued that the delay of schooling is related to family income, parental education, and parental occupation. They also reported that delaying entry into college significantly decreases the likelihood of completing a bachelor's degree. Similar findings are reported in Bozick and DeLuca (2005) and Kenyon (2007) who argued that students from low socioeconomic backgrounds are less likely to attend college immediately after high school.

Butler (2016), based on the Education Longitudinal Study of 2002 (ELS2002), suggested that only six percent of individuals who delayed college entry earned a bachelor's degree compared to 42 percent for those who began college immediately after high school graduation. Further, 21 percent of the sampled individuals delayed entry to college, and black and Hispanic students were more likely to delay than white students. Among those who delayed entry, the primary factors for this decision were parental education and college costs. On the labor market side, students who delayed enrollment earned 28 percent less than those who entered directly from high school. Those who delayed were also less likely to hold professional or managerial positions in the workplace.

Niu and Tienda (2013) investigated how college postponement is associated with four-year college enrollment using a representative longitudinal survey of Texas high school seniors who graduated in 2002. They argued that the decision to delay college was affected not only by family background characteristics but also by the student's academic ability. They also found that delays of one year or more are associated with significantly lower odds of re-enrollment.

Finally, based on the NLSY97, Johnson (2013) proposed a dynamic model of college enrollment and completion, labor force participation, and savings to study schooling decisions. He finds that 25 percent of those who leave school for one or more semesters re-enroll. He also estimates that the loss of the present value of lifetime income of a one year delay in college enrollment equals \$ 9,000.

Further, Johnson (2013) argues that students who delay their schooling do not necessarily take this action because they are financially constrained. Instead, the decisions may be driven by shocks to preferences for schooling and changes in the opportunity cost of schooling caused by unemployment or negative wage shocks.

2.2 Racial differences in Education

The over-arching inequality problem for black and Hispanics has been intensively discussed in the United States and the topic has been vastly researched, especially by sociologists, and many of the references in this section are drawn from that literature. Some papers argue that differences in the family background could explain why black people in general, have lower educational attainment and lower income than their white counterparts. For example, Gamoran (2001) points out that the gap in high school graduation rates for whites and blacks is falling, but the graduation rates of the university or higher education for the same groups remain unbalanced. He believes that socioeconomic differences are the primary factor for black-white inequality in educational attainment. He argues (and demonstrates) that white parents tend to have higher levels of educational attainment and income than blacks' parents.

From the analysis of six large data sets collected between 1965 and 1996 using national probability samples of adolescents, Hedges and Nowell (1999) found that about a third of the gap in test scores was accounted for by racial differences in social class. Jencks and Phillips (1998) also support the idea that racial differences in social class are the main reason for educational inequality. They believe that policies designed to increase the funding for schools and reduce class sizes in the early grades of schooling will benefit blacks more than whites and may, therefore, improve educational equality.

Grissmer et al. (1994) estimate the net effect of changing family characteristics and demographics on aptitude scores and find that changing family characteristics would boost scores by about seven percentile points. Higher parental education and smaller family size are the main factors that affect a student's educational scores. They also claim that more government investment in school resources will benefit educational attainment, especially for Hispanic and black students.

Heckman (2011) generally agrees with this proposal. He argues that long-term factors (including parental education) are important determinants of student achievements while short-term financial support at the post-secondary level will have a more limited effect. A couple of older studies provide an interesting historical perspective on this issue. Coleman and Blum (1970 and 1972) note that unlike non-black males, whose educational attainment is not affected by family background characteristics, the educational attainment of black males is significantly affected by their mothers' education. They also show that wage growth for blacks and non-blacks differ, with higher growth rates among non-blacks, despite a similar starting wage.

Finally, Keane and Wolpin (2000) provide an analysis of different policies aimed at reducing the disparity in schooling and earnings across racial groups. They estimate a dynamic, structural model using data on males from NLSY79 and test two policy proposals: first, the provision of a high school graduation bonus to males from low-income families; and second, offering wage subsidies to low-wage workers. The first policy proposal has a significant effect on both black and white males, whereby black males can benefit more from the policy with the rate of having a cited degree even higher than for white males.

3 Model

The model described in this section provides a framework for analyzing the long-term impact of schooling interruptions on wages where the decision to return to school is endogenous. It deviates from how the existing literature has approached this issue, where many studies have ignored the possible selectivity embedded in the interruption decision. It is inspired by the model presented in Belzil et al. (2017) and shares many similarities with that one. However, it has been significantly modified to accommodate the unique features that the focus on interruptions entails.

To accomplish this, I assume that individuals make a choice at the beginning of each period in order to maximize the expected value of their state-specific lifetime utility. I define a time period as a calendar year, with a total of up to 30 time periods included in the model.[6] The initial period, when an individual starts to make choices(t_1), is the year he turns 16. However, because of attrition and different ages at the initial interview, not all individuals are observed for the entire sample period (around 66 percent of respondents in the 1979 cohort were observed for all 30 periods).

3.1 Choice set

There are three mutually exclusive and exhaustive options available in each period: school enrollment, employment and a residual, absorbing state reflecting NEET time and . I refer to this state as Not Enrolled, Employed or in Training (NEET). For each individual i, the choice in period t is represented by the binary

[6] In the initial analysis, where I focus an a sample of white males from the 1979 cohort (similar to that used by Light (1995)), I am able to track a majority of the respondents for 30 periods. For the analysis that compares choices and outcomes over time, using data from both NLSY cohorts, I limit the number of time periods in the 1979 cohort to 17 (which is the maximum available for the 1997 cohort). This is done so that the time frame is similar for the two cohorts.

indicators $d_{i,t}^k$, where $d_{i,t}^k = 1$ when option k is chosen in period t, where $k \in K$, and $K = \{s, e, NEET\}$. Thus, $d_{i,t}^s = 1$ if individual i is enrolled in school in period t, $d_{i,t}^s = 0$ otherwise; $d_{i,t}^e = 1$ if he is employed in period t, $d_{i,t}^e = 0$ otherwise; and finally $d_{i,t}^{NEET} = 1$ if individual i is not enrolled in school or employed in period t, $d_{i,t}^{NEET} = 0$ otherwise.

3.2 State variables

The vector of state variables for each individual in period t is denoted as:

$$\begin{aligned}
\Omega_{i,t} =& \{X_i, cog, type_i, S_{1_{i,t-1}}, S_{2_{i,t-1}}, E_{1_{i,t-1}}, E_{2_{i,t-1}}, NEET_{i,t-1}, ds_{i,t-1}^{t-1}, \\
& d(s_{i,t-1} = 1), d(s_{i,t-1} = 2), d(s_{i,t-1} = 3), d(s_{i,t-1} = 4), \\
& d(s_{1_{i,t-1}} = 1), d(s_{1_{i,t-1}} = 2), d(s_{1_{i,t-1}} = 3), d(s_{1_{i,t-1}} = 4), \\
& d(s_{2_{i,t-1}} = 1), d(s_{2_{i,t-1}} = 2), d(s_{2_{i,t-1}} = 3), d(s_{2_{i,t-1}} = 4), \\
& dgap_{i,t-1}, dgapdur_{i,t-1}, S_{i,t-1}, E_{i,t-1}\},
\end{aligned}$$

where X_i includes observed socioeconomic information, such as number of siblings, presence of both biological parents, parental education, parent's income (in 1997 U.S. dollar), and geographical location:[7]

$$\begin{aligned}
X_i = \quad & \{NSIB_i, I\left(NUCLEAR_i = 1\right), I\left(SOUTH_i = 1\right), I(RURAL_i = 1), \\
& I(RURAL_i = 1), HGCF_i, HGCM_i, HINCOME_i\},
\end{aligned}$$

The agent's cognitive and non-cognitive scores are represented by $ability_i$

[7] $NSIB_i$ is the number of siblings, $NUCLEAR_i$ is an indicator that is equal to 1 if individual i was living in a nuclear family (living with both biological parents at age 14), 0 otherwise. $SOUTH_i$ is an indicator that is equal to 1 if individual i was living in a southern state, 0 otherwise. $RURAL_i$ is an indicator that is equal to 1 if individual i was living in a rural area, 0 otherwise. $HGFM_i$ is the highest education individual $i's$ father obtained. $HGCM_i$ is the highest education individual $i's$ mother obtained. $HINCOME_i$ is the total income of individual $i's$ parents. With the exception of $HINCOME$, all variables are measured in the first period.

while his permanent unobserved heterogeneity is given by $type_i$. $S_{1,i,t}$,$S_{2,i,t}$, $E_{1,i,t}$,$E_{2,i,t}$and $H_{i,t}$ describe the accumulated number of periods occupied in school, employment and NEET, respectively. $dgap_{i,t}$ is a binary indicator variable that equals one if the individual has experienced a schooling interruption (or a gap) in period t and $dgap\_du_{i,t}$ is the duration of such an interruption, also measured in period t.

Further, $d\left(s_{i,t}=k\right)$,$k=1,2,3$ and 4 are binary indicators for the highest grade (high school, some college, college and post-college) completed at the beginning of period t:

$$\begin{aligned}
d(s_{i,t}=1) &= I\left(S_{i,t}=12\right) \\
d(s_{i,t}=2) &= I\left(13\leq S_{i,t}\leq 14\right) \\
d(s_{i,t}=3) &= I\left(15\leq S_{i,t}\leq 16\right) \\
d(s_{i,t}=4) &= I\left(S_{i,t}>16\right),
\end{aligned}$$

and $d\left(s_{1_{i,t}}=k\right)$ are binary indicators for the highest grade (high school, some college, college and post-college) completed at the beginning of period t in the first schooling span:

$$\begin{aligned}
d(s_{1_{i,t}}=1) &= I\left(S_{1_{i,t}}=12\right) \\
d(s_{1_{i,t}}=2) &= I\left(13\leq S_{1_{i,t}}\leq 14\right) \\
d(s_{1_{i,t}}=3) &= I\left(15\leq S_{1_{i,t}}\leq 16\right) \\
d(s_{1_{i,t}}=4) &= I\left(S_{1_{i,t}}>16\right).
\end{aligned}$$

Finally, $d\left(s_{2_{i,t}}=k\right)$ are binary indicators for the highest grade (high school, some college, college and post-college) completed at the beginning of period t in the second schooling span as:

$$
\begin{aligned}
d(s_{2_{i,t}} = 1) &= I\left(S_{2_{i,t}} = 12\right) \\
d(s_{2_{i,t}} = 2) &= I\left(13 \leq S_{2_{i,t}} \leq 14\right) \\
d(s_{2_{i,t}} = 3) &= I\left(15 \leq S_{2_{i,t}} \leq 16\right) \\
d(s_{2_{i,t}} = 4) &= I\left(S_{2_{i,t}} > 16\right).
\end{aligned}
$$

3.3 Initial conditions

At the beginning of the first period, each agent is endowed with a set of both observed and unobserved characteristics. The individual's socio-economic status is permanent and given by his family environment at this time (X_i). The ability measures ($cognitive_i$) are also assumed to be time-invariant. Permanent unobserved heterogeneity is represented by the individual's type (detailed below).

The initial value of S_i, ($S_{i,0}$) is the grade completed at age 16, while $E_{i,0}$ and $NEET_{i,0}$ are both equal to zero.

3.4 Laws of motion

The laws of motion of accumulated education, work experience, and NEET are defined as follows:

$$
\begin{aligned}
S_{1,i,t} &= S_{1,i,t-1} + I\left(d^s_{i,t} = 1\right) \ if \ dgap_{i,t} = 0 \\
S_{2,i,t} &= S_{2,i,t-1} + I\left(d^s_{i,t} = 1\right) \ if \ dgap_{i,t} = 1 \\
E_{1,i,t} &= E_{1,i,t-1} + I\left(d^e_{i,t} = 1\right) \ if \ dgap_{i,t} = 0 \\
E_{2,i,t} &= E_{2,i,t-1} + I\left(d^e_{i,t} = 1\right) \ if \ dgap_{i,t} = 1 \\
NEET_{i,t} &= NEET_{i,t-1} + I\left(d^{NEET}_{i,t} = 1\right).
\end{aligned} \tag{1}
$$

3.5 Preferences

In this model, individuals are assumed to maximize the expected value of lifetime utility when they make their choices at the beginning of each period. The state-specific utilities are described below.

3.5.1 Utility of school

The instantaneous utility associated with attending school in period t is denoted U^s_t and is defined as:

$$
\begin{aligned}
U^s_{i,t} \quad = \quad & \alpha^s_m + X_i\beta^s_x + \beta^s_g ds^{t-1}_{i,t} + \beta^s_{gapdur} dgapdur_{i,t} \\
& + \beta^s_{ds1} d(s_{i,t} = 1) + \beta^s_{ds2} d(s_{i,t} = 2) + \beta^s_{ds3} d(s_{i,t} = 3) + \beta^s_{ds4} d(s_{i,t} = 4) \\
& + \varepsilon^s_{i,t},
\end{aligned} \tag{2}
$$

where α^s_m is the permanent observed and unobserved individual heterogeneity for type m in terms of the utility of attending school. β^s_g captures the potential cost associated with returning to school while β^s_{gapdur} captures the possibility that the cost a school interruption depends on its duration.β^s_{ds1}, β^s_{ds2}, β^s_{ds3} and

β^s_{ds4} capture the differences in the utility of attending school across the four different schooling levels. The four parameters describe how the cost of attending school evolves at the different levels of education.β^s_X, $x = 1, ..., 7$ describe how the utility of enrolling in school varies across different socioeconomic background characteristics. Finally, $\varepsilon^s_{i,t}$ is a stochastic utility shock.

3.5.2 Utility of working

The instantaneous utility of working, denoted as $U^e_{i,t}$, is defined as:

$$U^e_{i,t} = \alpha^e_m + lnw_{i,t} + \beta^e_{s1} S_{1_{i,t}} + \beta^e_{s2} S_{2_{i,t}} + \beta^e_e E_{i,t} + \varepsilon^e_{i,t}, \tag{3}$$

where α^e_m is the time-invariant, type-specific utility shifter and $lnw_{i,t}$ is the expected log-wage in period t. The parameters β^e_{s1} and β^e_{s2} captures the direct effects of education on the utility of working (in addition to the indirect effect of education through the expected log-wage), whereas β^e_e measures the state-dependence in working. $\varepsilon^e_{i,t}$ is a stochastic utility shock.

3.5.3 Utility of NEET

The utility of NEET is denoted as U^{NEET}_t and is normalized to zero and serves as a reference group.

3.6 Wages

Individual wages depend on schooling, labor market experience and any interruption in schooling. In addition to the occurrence of an interruption, the duration of the interruption or gap is included since longer gaps may reduce subsequent wages. In order to capture potential differences in wage growth for those with and without a gap, education and experience before and after a schooling interruption are included in the wage function. Specifically, the log-wage function is given by:

$$
\begin{aligned}
lnw_{i,t} \quad = \quad & \alpha_m^w + \beta_{e1}^w E_{1_{i,t}} + \beta_{e2}^w E_{2_{i,t}} + \beta_{e1^2}^w E_{1_{i,t}}^2 + \beta_{e2^2}^w E_{2_{i,t}}^2 \\
& +\beta_{s1_1}^w d(s_1 = 1)_{i,t} + \beta_{s1_2}^w d(s_1 = 2)_{i,t} \\
& +\beta_{s1_3}^w d(s_1 = 3) + \beta_{s1_4}^w d(s_1 = 4)_{i,t} + \beta_{s2_1}^w d(s_2 = 1)_{i,t} \\
& +\beta_{s2_2}^w d(s_2 = 2)_{i,t} + \beta_{s2_3}^w d(s_2 = 3)_{i,t} + \beta_{s2_4}^w d(s_2 = 4)_{i,t} \\
& +S_{1_{i,t}} E_{1_{i,t}} \lambda_{s1e1} + S_{1_{i,t}} E_{2_{i,t}} \lambda_{s1e2} + S_{2_{i,t}} E_{2_{i,t}} \lambda_{s2e1} + \varepsilon_{i,t}^w,
\end{aligned}
\tag{4}
$$

where $S_{1_{i,t}} E_{1_{i,t}}$, $S_{1_{i,t}} E_{2_{i,t}}$ and $S_{2_{i,t}} E_{1_{i,t}}$ are the arrays of interaction terms of pre- and post-interruption schooling (at different educational levels) with working experience at time period t:

1)The interaction term of pre-gap educational levels and pre-gap work experience:

$$
\begin{aligned}
S_{1_{i,t}} E_{1_{i,t}} \lambda_{s1e1} \quad = \quad & \beta_{e1s1_1}^w E_{1_{i,t}} d(s_1 = 1)_{i,t} + \beta_{e1s1_2}^w E_{1_{i,t}} d(s_1 = 2)_{i,t} \\
& +\beta_{e1s1_3}^w E_{1_{i,t}} d(s_1 = 3)_{i,t} + \beta_{e1s1_4}^w E_{1_{i,t}} d(s_1 = 4)_{i,t}
\end{aligned}
\tag{5}
$$

2)The interaction term of pre-gap educational levels and post-gap work experience:

$$
\begin{aligned}
S_{1_{i,t}} E_{2_{i,t}} \lambda_{s1e2} \quad = \quad & \beta_{e2s1_1}^w E_{2_{i,t}} d(s_1 = 1)_{i,t} + \beta_{e2s1_2}^w E_{2_{i,t}} d(s_1 = 2)_{i,t} \\
& +\beta_{e2s1_3}^w E_2 d(s_1 = 3)_{i,t} + \beta_{e2s1_4}^w E_{2_{i,t}} d(s_1 = 4)_{i,t}
\end{aligned}
\tag{6}
$$

3) The interaction term of post-gap educational levels and post-gap work experience:

$$
\begin{aligned}
S_{2_{i,t}}E_{2_{i,t}}\lambda_{s2e2} \quad = \quad & \beta^w_{e2s2_1}E_{2_{i,t}}d(s_2=1)_{i,t}+\beta^w_{e2s2_2}E_{2_{i,t}}d(s_2=2)_{i,t} \\
& +\beta^w_{e2s2_3}E_{2_{i,t}}d(s_2=3)_{i,t}+\beta^w_{e2s2_4}E_{2_{i,t}}d(s_2=4)_{i,t},
\end{aligned} \tag{7}
$$

Further, α^w_m is a type-specific, permanent unobserved heterogeneity. $\beta^w_{s1_p}$ and $\beta^w_{s2_p}$ where $p=$ 1,2,3 and 4 capture the return to pre- and post-interruption education according to different education levels. β^w_{e1} and β^w_{e2} represent the growth shifts of the log-wage for the pre- and post-interruption work experience.λ_{s1e1} and λ_{s1e2} show how pre- and post-interruption wage growth depend on pre-interruption levels of schooling. λ_{s2e2} show the extent to which working during an interruption to schooling impacts the return to post-interruption schooling. $\beta^w_{e1^2}$ and $\beta^w_{e2^2}$ allow for non-linear growth rates in E_1 and E_2, respectively.

Finally, $\varepsilon^w_{i,t}$ is an idiosyncratic, stochastic wage shock that follows a first order time series process. Specifically:

$$
\begin{aligned}
\varepsilon^w_{i,t} &= \rho\varepsilon^w_{i,t-1}+v^w_{i,t} \\
v^w_{i,t} &\sim N\left(0,\ \sigma^2_\nu\right) \\
\rho &\in (0,1).
\end{aligned} \tag{8}
$$

3.7 Permanent observed and unobserved heterogeneity

All type-specific heterogeneity terms described above $(\alpha^s_m, \alpha^e_m, \alpha^w_m)$ [8]depend on the agent's cognitive test score, cog_i, and is specified as follows:

[8] m is the type, $m=$ 1,2,3,4.

$$\alpha_m^s = \breve{a}_m^s + \breve{a}^s \cdot cog_i \tag{9}$$

$$\alpha_m^e = \breve{a}_m^e + \breve{a}^e \cdot cog_i \tag{10}$$

$$\alpha_m^w = \breve{a}_m^w + \breve{a}^w \cdot cog_i. \tag{11}$$

Following the standard approach in the structural, dynamic discrete choice literature, the distribution of unobserved heterogeneity distribution is approximated by a multivariate discrete distribution. Thus, each type individual i is endowed with the following set:

$$\begin{aligned} \Psi &= \{\breve{a}^s_{i,type=m},\ \breve{a}^e_{i,type=m},\ \breve{\alpha}^w_{i,type=m}\} \\ m &\in M \\ M &= \{1,\ 2,\ 3,\ 4\} \end{aligned}$$

where the probability of belonging to a specific type is defined by the following logistic transformation:

$$Pr(type = m|S_{i,0}) = p_m = \frac{\exp(\lambda_m + \lambda_{ms} S_{i,0})}{\sum_{j=1}^{M} \exp(\lambda_j + \lambda_{js} \cdot S_{i,0})}, \tag{12}$$

where $S_{i,0}$ is the initial level of schooling for individual i. λ_m and λ_{ms} are parameters to be estimated and represent the intercept term and the weight of the initial schooling for individual i, respectively, in terms of the probability of belonging to type m. Parameters λ_M and λ_{Ms} are normalized to zero.

3.8 Value functions

As mentioned above, the model assumes that agents are forward-looking and considers the impact of current choices on future outcomes. Therefore, the choice-specific, instantaneous utility functions described above need to be adjusted to account for expectations in terms of future utility and wage shocks. Specifically, the objective of individual i in any period t, $t = 1, ..., T$ is to choose option k in order to maximize the expected present value of his future utility stream:

$$\begin{array}{c} \max_k E\left\{\sum_{t=1}^{T} \delta^{t-1} U\left(\Omega_{i,t} | d_{i,t}^k = 1\right)\right\} \\ k = \{s,\, e,\, NEET\} \end{array}, \tag{13}$$

where δ is the discount factor and E is the expectations operator. Using a Bellman equation (Bellman, 1957), this complex dynamic optimization problem can be reduced to sequences of two period problems and solved recursively. Thus, the optimization problem above can be formulated as:

$$V_{i,t}\left(\Omega_{i,t}\right) = max_{k=\{s,e,NEET\}} \left\{U_{i,t}^k + \delta EV_{i,t+1}\left(\Omega_{i,t+1} | d_{i,t}^k = 1, \Omega_{i,t}\right)\right\} \tag{14}$$

$$EV_{i,t+1}\left(\Omega_{i,t+1} | d_{i,t}^k = 1, \Omega_{i,t}\right) = E \max\left\{V_{i,t+1}^s, V_{i,t+1}^e, V_{i,t+1}^{NEET} | d_{i,t}^k = 1, \Omega_{i,t}\right\} \tag{15}$$

for $t = 1, ..., T-1$ and

$$V_{i,T}\left(\Omega_{i,T}\right) = max_{k=\{s,e,NEET\}} U_{i,T}^k, \tag{16}$$

for period T and where $V_{i,t}\left(\Omega_{i,t}\right)$ denotes the value function. Starting from the last period, it is straightforward to use backwards recursions to solve the model given the conditional expectation $EV_{i,t+1}\left(\Omega_{i,t+1}|d^k_{i,t}=1,\Omega_{i,t}\right)$. While this expectation does not in general have a convenient solution, by assuming that the utility shocks $\left\{\varepsilon^s_{i,t},\, \varepsilon^e_{i,t},\, \varepsilon^{NEET}_{i,t}\right\}$ follow an i.i.d. extreme value distribution, $EV_{i,t+1}\left(\Omega_{i,t+1}|d^k_{i,t}=1,\Omega_{i,t}\right)$, conditional on the wage shocks $\left(\varepsilon^w_{i,t}\right)$, can be expressed as:

$$\mathit{\Phi}_{i,t+1}\left(\Omega_{i,t+1}\right) \quad = \quad \tau\gamma+\tau lnG, \tag{17}$$

where G is specified as:

$$G=\left(\sum_k exp\left\{\frac{\left(U^k_{i,t+1}+\delta EV_{i,t+2}\left[\Omega_{i,t+2}|d^k_{i,t+1}=1,\Omega_{i,t+1},\varepsilon^w_{i,t+1}\right]\right)}{\tau}\right\}\right), \tag{18}$$

where τ is a parameter of the extreme value distribution and γ is Euler's constant. What remains is to integrate out the wage shocks from the expression above, and this can be done by approximating the integral using simulated draws from the distribution of wage shocks and averaging the simulated conditional expectations over the draws. That is:

$$\widehat{\mathit{\Phi}}_{i,t+1}\left(\Omega_{i,t+1}\right)=\frac{1}{R}\sum_{r=1}^{R}\mathit{\Phi}^r_{i,t+1}\left(\Omega_{i,t+1}\right). \tag{19}$$

Finally, normalizing τ to one, the choice probabilities implied by the model assumptions, $Pr(d^k_{i,t}=1)$, are defined as:

$$\Pr(d^k_{i,t} = 1) = \frac{\exp\left(U^k_{i,t} + \delta\widehat{\varPhi}_{i,t+1}\left(\Omega_{i,t+1}\right)\right)}{\sum_j \exp\left(U^j_{i,t} + \delta\widehat{\varPhi}_{i,t+1}\left(\Omega_{i,t+1}\right)\right)} \tag{20}$$

$$k, j = s, e, NEET$$

$$k \neq j.$$

Evaluating $\varPhi_{i,t+1}\left(\Omega_{i,t+1}\right)$ at all possible combinations of state variables for all applicable time periods and for all draws of the wage error distribution is computationally very demanding, despite the limited number of options available in each time period. To overcome this issue, I approximate the $\varPhi_{i,t+1}\left(\Omega_{i,t+1}\right)$ function using a polynomial in state variables, following Geweke and Keane (1996) and Geweke et al. (1999). Specifically:

$$\begin{aligned}\widetilde{\varPhi}_{i,t+1}\left(\Omega_{i,t+1}\right) = \quad & \varrho_0 + \varrho_1 S_{1_{i,t+1}} + \varrho_2 S_{2_{i,t+1}} + \varrho_3 E_{1_{i,t+1}} + \varrho_4 E_{2_{i,t+1}} + \\ & \varrho_5 S^2_{1_{i,t+1}} + \varrho_6 E^2_{1_{i,t+1}} + \varrho_7 S^2_{2_{i,t+1}} + \varrho_8 E^2_{2_{i,t+1}} + \\ & \varrho_9 S_{1_{i,t+1}} S_{2_{i,t+1}} + \varrho_{10} E_{1_{i,t+1}} E_{2_{i,t+1}} + \\ & \varrho_{11} S_{1_{i,t+1}} E_{1_{i,t+1}} + \varrho_{12} S_{2_{i,t+1}} E_{2_{i,t+1}} + \\ & \varrho_{13} S_{1_{i,t+1}} E_{2_{i,t+1}} + S_{2_{i,t+1}} E_{1_{i,t+1}}.\end{aligned} \tag{21}$$

4 Data

The empirical part of this thesis is based on samples from the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth (NLSY79, NLSY97). In this section, I first provide a brief description of each survey. This is followed by details on variable definitions and descriptions of different aspects of the data.

4.1 NLSY79

NLSY79 is a nationally representative sample of 12,686 males and females who were between 14 to 22 years old when they were first surveyed in 1979. After 1979, follow-up interviews were conducted annually until 1994 after which it became a biannual survey. NLSY79 gathers information in an event history format, in which dates are collected for the beginning and end of significant life events. Information includes the start and end dates for each job or school period held since the last interview. Because an individual's work and schooling histories are collected in this manner, measures of actual labor market experience can be constructed. Prior to any selections, NLSY79 contained 7,510 white agents, 3,174 black agents , and 2,002 Hispanic agents.

Since I am focusing on educational attainment, including high school dropouts, I remove individuals who were older than 18 at the time of the first survey in 1979. Those who were not living with their parents at the time of the survey were also excluded from the analysis.I also exclude respondents with missing information on included observed characteristics such as family income, Armed Forces Qualification Test (AFQT) scores, parental education, family stability (whether the individual report having been raised within a nuclear family or not), number of siblings, the area of residence (urban vs. rural), and ethnic background. After imposing these restrictions, the sample consists of of 1,199 white males, 625 black males, and 425 Hispanic males.

4.2 NLSY97

Similar to NLSY79, NLSY97 is also a nationally representative sample and share many similarities with the earlier cohort. Precisely, it consists of a nationally representative sample of 8,984 youths who were 12-16 years old as of late December 1996. For both NLSY cohorts, there is detailed information on family background and income as well as on individual scholastic ability (measured by Armed Forces Qualification Test (AFQT) scores). Interviews are ongoing for both cohorts and are conducted on an annual or biannual basis. After imposing similar restrictions on age as in NLSY79 and after removing those with missing information on key variables, I ended up with a sample of 1,265 white males, 425 black males, and 373 Hispanic males.

4.3 Variable descriptions

4.3.1 Family income and AFQT scores

For both surveys, information on parental income was collected at multiple interviews, and I use as much information as possible and create an average (over time) of parental income.For instance, if income information is available for four years (1978-1981), I use the average of those income measures. If income is only available for one of the years, the average income is replaced by that income. I chose this approach as it provides a measure closer to permanent income and it also minimizes the number of individuals dropped because of missing income. For both cohorts, I express income in the year 1997 dollars using the CPI for all urban consumers.

Like the earlier literature on education using NLSY data, I use AFQT scores to control for cognitive ability. The scores, available for both samples, are adjusted to account for age differences at the time of the tests. Moreover, the test scores for the 1997 cohort were adjusted to improve comparability with the scores for the 1979 cohort, following Altonji et al. (2012).

4.3.2 Individual characteristic variables

Individual characteristic variables are reported in Table-1.

The variable, *Number of siblings*, is the number of brother and sister in the individual's family. Individuals of all racial groups from NLSY79 cohort have significantly less sibling compared to their count parts from NLSY97. The numbers of siblings of white individuals from NLSY79 is 2.76 that is by far less than their black and Hispanic counterparts from the same cohort, which are 4.56 and 4.37, respectively. In NLSY97, the number of siblings of individuals from all racial groups converged to around 2.5. The number of siblings of black and Hispanic males decreased 78 percent and 74 percent respectively, while white males' number sibling just reduced 22 percent. That may indicate the opportunity cost of having more kids was increased dramatically from cohort 1979 compared to cohort 1997, especially for black and Hispanics families.

The variable, *Intact family*, is the proportion of males that who live in an intact family. The proportion of males that who live in a complete family in NLSY97 is dropped a lot compared to the ratio in NLSY79 for all racial groups except Hispanics males. The proportion in NLSY79 are 83 percent and 58 percent for white and black males respectively and reduced to 69 percent and

42 percent in NLSY97 which is the ratio of intact family decreased 22 percent and 38 percent respectively. The proportion of an intact family only reduced 5 percent for Hispanic males.

The variable, *South*, is the proportion of males who that live in south region states[9]. That is only a quart of white and Hispanics males live in the south region states, and over half of the males living the south region states in NLSY79 and the ratios are similar in NLSY97, the ratios of live in south region states are 26 percent , 64 percent and 31 percent for white, black and Hispanics males respectively.

The variable, *Rural*, is the proportion of males who live in rural area. In NLSY79, the proportion is 26 percent, 22 percent and 14 percent for white, black and Hispanics males, respectively. In NLSY97, the ratio for white males rise to 30 percent but both black and Hispanics males reduced to 21 percent and 10 percent respectively. That shows white families moving from urban area to rural area and black and Hispanics males moving from rural area to urban area.

The variable, *Father's education*, is the average total schooling of individuals' father. In NLSY79, the average total schooling of white males' fathers is 12.55 years which is a bit above high school degree. The average total schooling for black and males' fathers is about 10.36 years and 8.61 years, which is 21 percent and 45 percent lesser than white males'. In NLSY97, the average total schooling of father are all increased. White, black and Hispanics males' verge total schooling of fathers increased to 13.78 years, 12.2 years and 10.76 respectively.

The variable, *Mother's education*, is the average total schooling of individuals' mother. In NLSY79, the average total schooling of white males' mothers is

[9] Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee Texas, Virginia and West Virginia.

Table 1: Descriptive Statistics

	NLSY79					
	White		Black		Hispanic	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Periods observed in panel	13.84	5.57	12.51	4.37	12.77	4.18
Number of siblings	2.76	1.76	4.56	2.96	4.37	2.79
Intact family	0.83	0.38	0.58	0.49	0.72	0.45
South	0.25	0.43	0.53	0.50	0.26	0.44
Rural	0.26	0.44	0.22	0.36	0.14	0.35
Father's education	12.55	3.19	10.36	3.27	8.61	4.68
Mother's education	12.14	2.29	11.05	2.49	8.26	4.32
Family income	66.92	33.82	28.04	17.51	42.15	23.49
Number of observations	1,199		625		418	

	NLSY97					
	White		Black		Hispanic	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Periods observed in panel	13.49	3.18	13.62	3.29	13.60	3.25
Number of siblings	2.26	1.04	2.55	1.35	2.51	1.14
Intact family	0.69	0.46	0.42	0.49	0.68	0.47
South	0.26	0.44	0.64	0.48	0.31	0.46
Rural	0.30	0.46	0.21	0.41	0.10	0.31
Father's education	13.78	2.83	12.2	2.32	10.76	3.73
Mother's education	13.69	2.44	12.60	2.09	11.08	3.41
Family income	99.63	70.94	36.39	38.96	58.78	45.59
Number of observations	1,265		425		373	

Note: Family income is measured in $10,000 and expressed in 1997 U.S. dollars.

12.14 years which is a bit above high school degree. The average total schooling for black and males' mothers are about 11.05 years and 8.26 years, which is less than high school degree. In NLSY97 , the average total schooling of mother are increased for all racial groups. White, black and Hispanics males' verge total schooling of mothers increased to 13.69 years, 12.60 years and 11.08 respectively. Compare to the average mothers' schooling in NLSY79, white, black and Hispanics increased 12 percent, 14 percent and 34 percent respectively.

Table 2: Endogenous Variables

Variable Name	NLSY79 White	NLSY79 Black	NLSY79 Hisp.
	Mean	Mean	Mean
pre-gap schooling (years)	12.8	11.6	11.3
post-gap schooling (years)	2.4	1.7	2.4
pre-gap working (years)	9.3	7.8	8.5
post-gap working (years)	3.4	2.4	3.2
NEET (years)	1.6	3.9	1.9
gap (%)	36.5	35	38
gap_du (years)	1.4	1.6	1.5

Variable Name	NLSY97 White	NLSY97 Black	NLSY97 Hisp.
	Mean	Mean	Mean
pre-gap schooling (years)	13.4	12.4	12.5
post-gap schooling (years)	3.0	2.1	2.7
pre-gap working (years)	7.3	6.6	6.9
post-gap working (years)	3.7	2.1	3.2
NEET (years)	1.7	3.6	1.7
gap (%)	40	43	41
gap_du (years)	1.3	1.7	1.3

	NLSY79 ALL	NLSY97 ALL
	Mean	Mean
pre-gap schooling (years)	11.9	12.8
post-gap schooling (years)	2.1	2.6
pre-gap working (years)	8.5	6.9
post-gap working (years)	3.0	3.0
NEET (years)	2.5	2.3
gap (%)	36	41
gap_du (years)	1.5	1.5

4.3.3 School enrollment and grade attainment

In both surveys, respondents are continuously asked about their current school enrollment status, the highest grade they have attended (and completed), as well as the dates they were last enrolled in school. The school enrollment state variable I use, $d^s_{i,t}$, is constructed from the school enrollment and grade attainment questions in time period t. Thus, $d^s_{i,t} = 1$ if the individual was enrolled in school in period t and had an increment in his grade from period $t-1$. If the individual was not enrolled or was enrolled but had no grade increment, $d^s_{i,t} = 0$.

4.3.4 Accumulated schooling before and after an interruption

In my analysis, I need to separate educational attainment during different spells. Specifically, I define $S_{1_{i,1}}$ as the highest grade completed by individual i in period t before any schooling interruption, while $S_{2_{i,t}}$ measures the additional grades completed in period t after a schooling interruption. Thus, $S_{2_{i,1}}$ equals $S_{i,t} - S_{1_{i,1}}$ if the individual has experienced a schooling interruption by the time period t, and equals 0 otherwise (in the case of no interruption).

To illustrate, assume the individual interrupted his schooling after completing grade 13 in period 3 and returns to school in period 5 and completes grade

16 before permanently leaving school. In this case:

$$S_{1_{i,1}} = 11$$

$$S_{1_{i,2}} = 12$$

$$S_{1_{i,3}} = 13$$

$$S_{1_{i,s}} = 13 \; for \; s = 4, ..., 30.$$

The values of $S_{2_{i,1}}$ will be as follows:

$$S_{2_{i,t}} = 0 \; for \; t = 1, ..., 4$$

$$S_{2_{i,5}} = 1$$

$$S_{2_{i,6}} = 2$$

$$S_{2_{i,t}} = 3 \; for \; t = 7, ..., 30.$$

The accumulated schooling in NLSY79, which are reported in Table-2, reported that the white agents obtain 12.8 years of schooling in the pre-gap schooling period and around 2.4 years in the post-gap schooling period. Black agents were obtained about 11.6 years in the pre-gap schooling and 1.7 years after schooling. Hispanics agents were obtained only 11.3 years pre-gap schooling but obtained 2.4 years post-gap schooling as the white agents.

From the table-2, The accumulated schooling in NLSY97 is report that all racial groups obtained more schooling on the pre-gap schooling, all Hispanics agents' pre-gap schooling is increased 1.2 years compared to its younger cohort. The post-gap schooling is also increased for all racial groups. In general, the post-gap schooling is increased 0.5 years in the cohort NLSY97 compare to NLSY79 which is about 10 percent increase. The pre-gap schooling is increased 0.9 years in the cohort NLSY97 compared to the number in NLSY79 which is about 20 percent increase.

4.3.5 Not Enrolled, Employed or in Training (NEET)

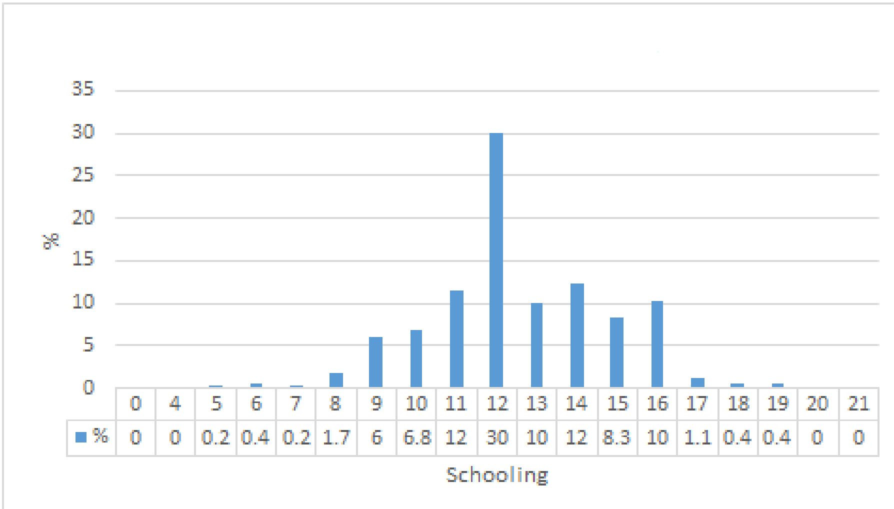
In my model, an individual may choose not to enroll in school or work in any given period. The indicator for this option is $d_{i,t}^{neet}$ and the accumulated number of periods in this state in period t is described by $NEET_{i,t}$. In NLSY79, White agents spend 1.6 years in NEET, Hispanics spend 1.9 years, and black agents spend 3.9 years on NEET[10]. In NLSY97, the trend of NEET is similar to the NLSY79 cohort.

4.3.6 Temporary interruptions to schooling

To identify individuals who have experienced schooling interruptions in period t, I use the notation $dgap_{i,t}$. The value of this variable changes from zero to one in the period when the individual re-enrolls in school. From this point forward it equals one, while it equals zero for all periods prior to re-enrollment. For individuals who never re-enroll, $dgap_{i,t} = 0 \forall t$. As showed in Table-2, among white males, 36.5 percent had at least one school interruption in the 1979 cohort while the corresponding figure for the 1997 cohort is 40 percent. Thus, the incidence of a school interruption for white males is virtually the same and the figure for the older cohort is also very similar to that reported by Light (1995). For black males, the proportion with an interruption was higher than for any other group in the 1979 cohort (35 percent) but decreased to 43 percent in the 1997 cohort, close to that for white males in the same cohort. Finally, for Hispanics, the interruption rate is constant at close to 38 percent in NLSY79 and 41 in the NLSY97[11].

[10] Table-2
[11] Table-2

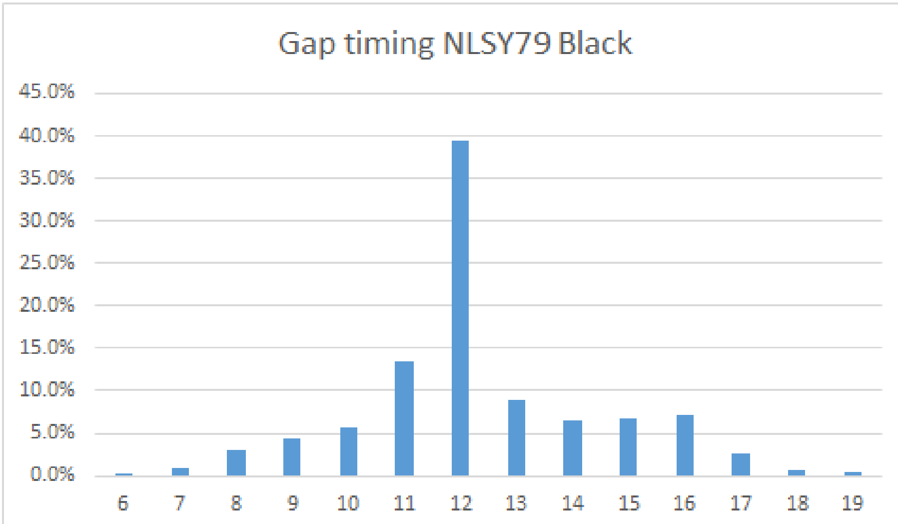
Figure 1: Highest grade completed at the time of the first school interruption (white males in NLSY79)



In Figure 1, I illustrate the distribution of highest grade completed at the time of a first schooling interruption[12]. Not surprisingly, most of the interruptions (30 percent of all interruptions) occur after high school graduation (grade 12). As discussed above, this school interruption between high school and college is also the one that has attracted the most attention in the previous literature. However, interruptions are also frequent during college years with over 30 percent occurring after the completion of grades 13-15. Figures 2 and 3 show the distribution of highest grade completed at the time of a first schooling interruption for black and Hispanic males belonging to the 1979 cohort. As can be seen, the timing of the interruptions is similar across the groups with a significant peak at grade 12. However, in all cases, the majority of interruptions occur after completion of other grades.

[12]This figure is based on data for white males in NLSY79

Figure 2: Highest grade completed at the time of first schooling interruption (black males in NLSY79)



In Figure 4, I present the distribution of the length or duration of school interruptions for white males from the 1979 cohort. As can be seen, most interruptions are short (11 percent of all individuals (or about a third of those with an interruption) had an interruption of one year only. Long school gaps are uncommon and, as expected, very few returns after gaps longer than 10 years. Again, similar profiles and patterns are observed for black and Hispanic males in the 1979 cohort.

4.3.7 Wages

Consistent with the previous literature on this topic, I use hourly wages as my earnings measure to avoid contamination due to differences in working hours. The natural logarithm of individual i's hourly wage in period t, $lnw_{i,t}$ is the

Figure 3: Highest grade completed at the time of first schooling interruption (Hisp. males in NLSY79)

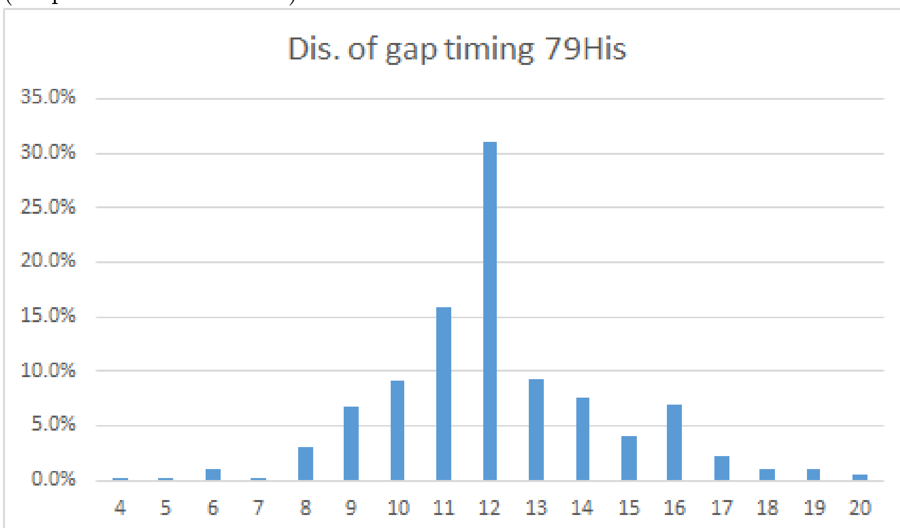
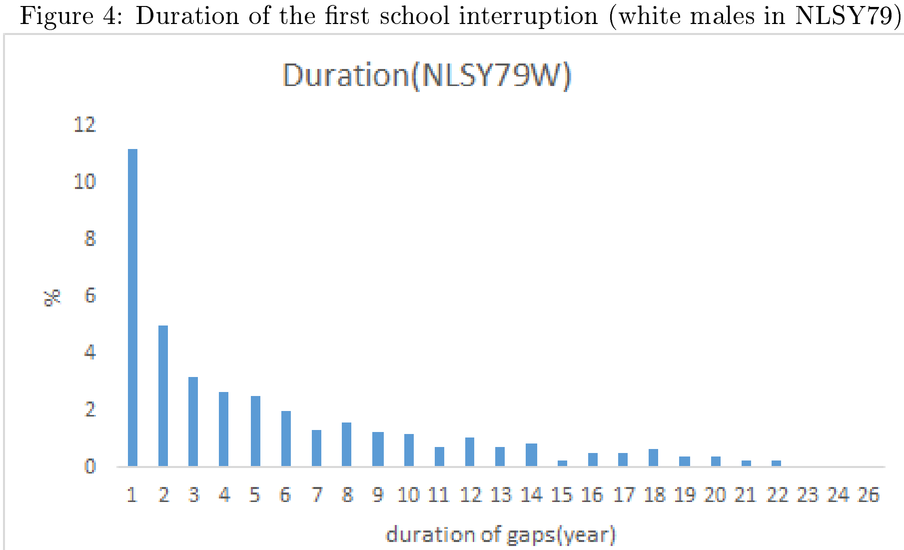


Figure 4: Duration of the first school interruption (white males in NLSY79)



outcome variable in the wage function. To construct this variable, information from questions about the individuals' hourly wages at their primary jobs are used. For those who did not provide an hourly wage rate, this figure is computed using their total annual earnings in period t divided by hours worked in that period. Nominal hourly wages are adjusted to real terms using the Consumer Price Index with 1997 as the reference year. Reported wages below the federal minimum wage rate in any given time period were discarded.

4.3.8 Employment and work experience

Both NLSY surveys provide detailed information on respondents' work history. Employment on a week-by-week basis is available from the work history files from the initial survey up to and including 2012. The employment variable $d^e_{i,t}$ is defined based on the individual's work status and working hours. Specifically, $d^e_{i,t} = 1$ if individual i worked during period t and was not enrolled in school. Work experience $E_{i,t}$ is the accumulation of employment states as of period t.

4.3.9 Work experience between first and second schooling spell

The wage function in my model is very general and I separate work experience before and after an interruption in order to allow for different wage growth for those with and without an interruption. Specifically, I define $E_{1_{i,t}}$ as the work experience acquired between the first and the second schooling spell. This begins to accumulate after the individual starts working. For those who never return to school, $E_{1_{i,t}}$ equals $E_{i,t}$. For those who do interrupt, $E_{1_{i,t}}$ equals $E_{i,t}$

during the first working spell and $E_{1_{i,t}} = E_{1_{i,t-1}}$ after the completion of the first working spell.

4.3.10 Work experience after second schooling spell

For those who return to school, I define $E_{2_{i,t}}$ as the work experience accumulated starting after the completion of the second schooling spell. Specifically, $E_{2_{i,t}} = E_{i,t} - E_{1_{i,t}}$ if the person has had an interruption in period t, and $E_{2_{i,t}} = 0$ otherwise. For example, assume that an individual interrupts his schooling after completing grade 13 in period 3, returns to school in period 5, completes grade 16 in period 7, and then works for the remaining time periods. In this case, $E_{1_{i,t}} = E_{1_{i,2}} = E_{1_{i,3}} = 0$ and from period 4 to 30, it will take on the value one, that is $E_{1_{i,4}} = 1, ..., E_{1_{i,30}} = 1$. The variable representing experience accumulated after an interruption will evolve as follows in this example: $E_{2_{i,1}} = ... = E_{2_{i,7}} = 0$ and from period 8 onwards, $E_{2_{i,t=t-7,...,30}}$.

5 Estimation

The parameters of the model are estimated using maximum likelihood techniques where the likelihood contribution for a given individual consists of the probability of choosing option k in period t, given by $Pr\left(d_{i,t}^{k}=1\right)$ above and the density function of log-wages $(f\left(w_{i,t}\right))$, derived using the distribution assumption of the wage shocks. The permanent unobserved heterogeneity terms are integrated out from the conditional probabilities and the overall, unconditional log-likelihood function for individual i is:

$$LogL_i = \sum_{m=1}^{M} \prod_{t=1}^{T_i} Pr\left(d_{i,t}^{k}=1|type=m\right) f\left(w_{i,t}\right)^{I_{i,t}(w>w_{min})} Pr\left(type=m\right), \tag{22}$$

where T_i denotes the number of time periods individual i is observed in the data and $I_{i,t}\left(w>w_{min}\right)$ is an indicator function that equals one if a wage is observed (above the minimum wage) in period t, and zero otherwise.

5.1 Identification

The model parameters are identified from the distributional assumptions made on the utility and wage shocks and from data on choices and wages in each period. The discount rate is set to 0.95, consistent with the norm in the literature. The parameters in the utility of attending school are also identified from data on the socioeconomic status of the individual, while information on initial education help to identify the type probabilities. In this paper, individuals are observed making different choices because they differ in types.

I use the highest level of education obtained by individual's mother and father and other personal characters to help explain the choice of schooling and the degree of education to help explain employment outcomes. The parameters of the wage equation are identified by agents' work experience and schooling choices, including the decisions to interrupt education. I also assume that individuals' s type is one of the facts that that affects their wages which allow individuals with the high unobserved ability to have persistently higher than average wages.

6 Empirical results

In this section, only selected estimates of interest will be discussed. All parameter estimates and their asymptotic standard errors for all racial groups and both cohorts appear in Tables 10 to 32 in the Appendix. Before discussing selected parameter estimates, I describe how I select the appropriate number of types used to approximate the distribution of unobserved heterogeneity. Following this, the results for white males from the 1979 cohort with 30 observation periods will be discussed. This discussion will be followed by presentations and comparisons of results for white, black and Hispanic males from both cohorts (using 17 periods of data in both NLSY79 and NLSY97). Lastly, I compare my results to those in Light (1995).

6.1 Selection of the number of types

In my model, to control for general serial correlation and make the dynamic model computational feasible, I assumed that individuals could be distinguished to M types where$M$ is determined by comparing two different, likelihood-based information criteria - the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). AIC and BIC are both common statistics in studies that need to discriminate between different model alternatives. In Table 2 and Figures 5 and 6, I show the AIC, [13] BIC[14] scores as well as the computational time as a result of different assumptions regarding the number of types.

As can be seen, both AIC and BIC are minimized for the 4-type model. Consequently, the results in this thesis are based on the 4-type model. The

[13] AIC=2 (number of estimated parameters in the model) + 2 (number of observations)L

[14] BIC= (number of estimated parameters in the model) (log of number of observation) + 2 (number of observations)L

Table 3: AIC and BIC Scores of different models

Models	no-type	2-type	3-type	4-type	5-type
AIC	20527.51	19673.64	19269.73	19094.47	19107.65
△AIC	-	-4.16	-2.05	-0.91	0.07
BIC	20781.85	19953.41	19574.93	19425.11	19463.72
△BIC	-	-3.99	-1.90	-0.77	0.20
Time	20hrs	36hrs	52hrs	64hrs	82hrs

Figure 5: AIC and BIC Scores

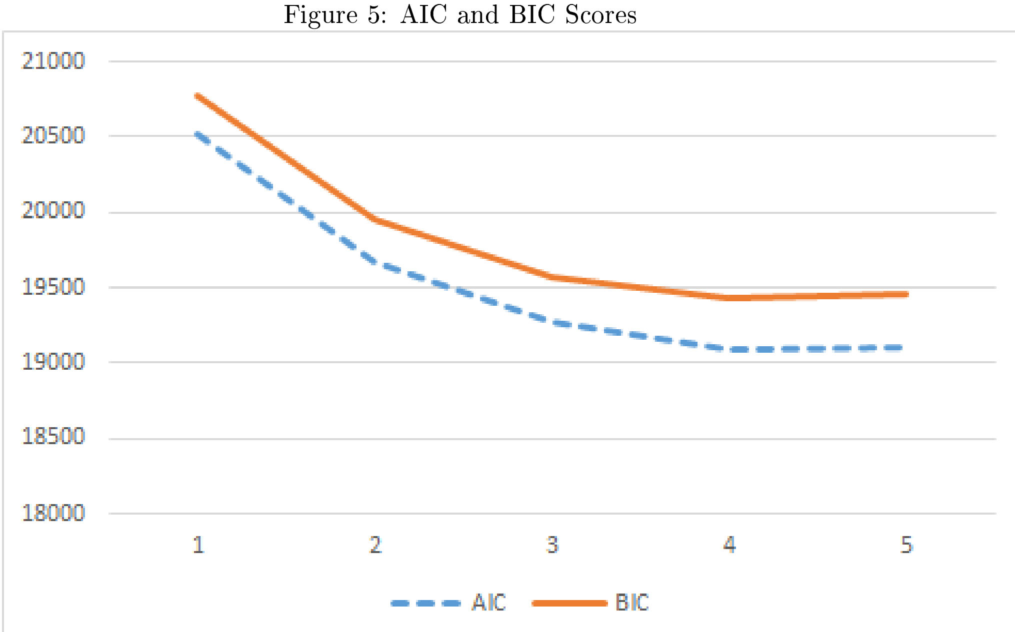
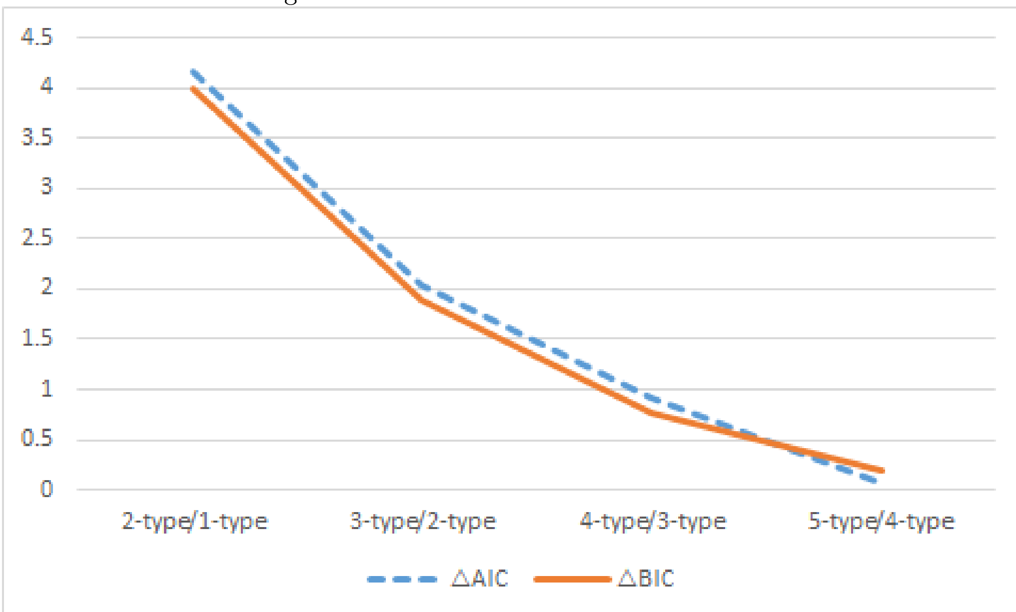


Figure 6: △AIC and △BIC Scores



computational times described in Table 1 refers to the model estimated on white males from the 1979 cohort where choices can be made for up to 30 periods. This long time horizon significantly impacts computational time, in addition to the incorporation of discretely distributed unobserved heterogeneity. Adding four types, compared to the specification that is based on a single, common type increases time till convergence with 44 hours.

6.2 Model fit

Model fit study is important in structural modeling. In this paper, I proved the model-fit study to the data of all racial groups in all cohorts. The result of all model fits of all data sets are provide in the appendix [15]. I conducted six

[15] figure 13 to figure 55

separate model-fit studies on six different data sets: white males (NLSY79), black males (NLSY79), Hispanics males (NLSY79), white males (NLSY97), black males (NLSY97) and Hispanic males (NLSY97). I compare and study endogenous variables from the generated data[16] to the observed data: 1) The incident of pre-gap schooling; 2) the incidence of the post-gap schooling; 3) the incidence of the pre-gap working status; 4) the incidence of the post-gap working status; 5) the incidence of being NEET; 6) the rate of schooling interruption; and 7) the wage path. The result shows the model fits most of the data sets properly.

6.3 Results for White Males from the 1979 Cohort

6.3.1 Unobserved heterogeneity

As mentioned above, I assume that the portion of the unobserved heterogeneity that is time-invariant is orthogonal to initial schooling, and as described above, discretely distributed with four support points.

Using the estimated parameters associated with the type probabilities and Bayes's rule, each has been allocated into one of the four groups. Selected details and features of the four types are shown in Table 4. From this table, we can see that type three is the most common type representing 62 percent of the sample, while type two is the least frequent one, with only five percent of the sample.

Further, the estimated utility and wage parameters that are type specific differ significantly across the four types with type two being the category with the lowest utility of both school and work, followed by type four. The incidence

[16]Generated by my structural model.

Table 4: Distribution of Type-Specific Heterogeneity (white males in NLSY79)

Type	1	2	3	4
α_m^s	6.7	1.7	7.0	2.8
α_m^e	4.1	-1.8	4.3	0.5
α_m^w	2.1	2.0	1.6	1.8
$gap\,\%$	29.8	33.3	31.3	46.4
$Cognitive(score)$	0.17	-0.2	-0.04	0.05
$Freq.\,\%$	13	5	62	20

of a schooling interruption or a gap also varies across the types and is highest for type four individuals (over 46 percent) and lowest among type one males (those with the highest cognitive skills). [17]

Since it is a non-linear model, the magnitude of coefficients of some parameters are difficult to interpret, and I will illustrate the impact of certain policies in counterfactual simulations below. Nevertheless, the signs of the estimates are interpretable and of interest, and a brief discussion will be devoted to this.

6.3.2 Parameters of the utility of schooling

After controlling for the unobservable terms, the coefficients associated with the grade level ,β_{ds1}^s, β_{ds2}^s, β_{ds3}^s and β_{ds4}^s, are -1.9122, -0.5961, -0.7583 and 0.8403, respectively. All the coefficients are significantly negative suggesting that white males' utility of schooling is the lowest when they first enroll in a postsecondary education institution. The utility of schooling is decreasing as the schooling increase[18]. The coefficient for enrolled in a graduate program is lower than the coefficients for enrolled in a four-year university program or a two-year university/college program.

[17] The cognitive skills measure is standardized with mean zero and variance equal to one.
[18] Expect if the person is enrolled in high school

The coefficient of parental income $\beta^s_{homeincome}$ is 0.0036 and it is significant. The coefficients for parental educations β^s_{hgcf} and β^s_{hgcm} are 0.0577 and 0.0575, respectively and significant. It indicates that both parents' education plays an equally important role in individual's utility of schooling. The coefficient of living in a nuclear family $\beta^s_{nuclear}$ is 0.2647, and the coefficient is significant. It confirms findings in other studies and suggests family stability play an important role in educational outcomes. The coefficient on siblings β^s_{nsib} is-0.0505 and is also significant. It also confirms previous patterns where a negative association between family size and education has been identified. From an economic perspective, this suggests that families with many children have fewer resources to devote to each child's education. The coefficient of the duration of the interruption to schooling β^s_{gapdur} is 0.0541, and it is significant. It suggests that students will be more likely to return to school the longer the duration of the schooling interruption. The the coefficient of return to school indicator $\beta^s_{I_{t-1}}$is -2.2025 and it is significant. It suggests that if a white male is not in school at the last period, then he is less likely to choose to go to school the current term. The coefficients of living in a southern state and living in a rural area are not significant factors related to the utility of schooling.

6.3.3 Parameters of the utility of working

The coefficients of education, β^e_{eudc} and work experience, β^e_{exp} are 0.3053 and 0.2037, respectively and they are both significant. These suggest, not very surprisingly that the probability of working improves with the level of education and past work experience.

6.3.4 Parameters of the wage function

As mentioned above, the wage function is specified in a very general way in order to allow for different returns to education depending on how the education was acquired and the highest level that was achieved. It also needs to allow for different growth rate post-graduation or after completion of education. As such, the specification used in this thesis differs from the more traditional, Mincer-type regressions that are common in labor economics. It also means that the coefficients are not as straightforward to interpret. Consequently, in this section, I will present the main estimates and comment on their significance and general effect but refer more refined analysis of the wage effect of school interruptions for example to the section on simulations below.

The estimates associated with pre-gap schooling levels, $\beta^w_{s1_1}$, $\beta^w_{s1_2}$, $\beta^w_{s1_3}$ and $\beta^w_{s1_4}$, are 0.1340, 0.1890, 0.5049 and 0.7357, respectively. All of them are significant, positive and increasing as the schooling levels increase suggesting that increases in pre-gap education levels significantly increase entry wages and there is evidence of so-called "sheep-skin effects" as the estimates suggest discrete increases when certain grade levels have been completed.

In terms of work experience, the variables measuring experience before and after an interruption enter the wage equation both in levels as well as squared (to allow for the wage-experience profiles to be concave). Each of them is also interacted with education before and after a school gap, differentiated by the highest level completed. Starting with the level variables (β^w_{e1} and β^w_{e2}), the estimates are 0.0492 and 0.0538, respectively and both are highly significant. The coefficients of the squared terms of these experience variables, $\beta^w_{e1^2}$ and $\beta^w_{e2^2}$ are-0.0011 and -0.0014, both negative and significant, which shows that both experience profiles are concave.

The estimated interaction terms are very interesting. The interaction between pre-gap schooling and work experience, $\beta^w_{e1s1_1}$, $\beta^w_{e1s1_2}$, $\beta^w_{e1s1_3}$and $\beta^w_{e1s1_4}$ are all positive and significant. This suggests that wages grow faster with work experience for those with more pre-gap schooling. However, the interaction terms of pre-gap work experience and post-gap educational levels,$\beta^w_{e1s2_1}$, $\beta^w_{e1s2_2}$, $\beta^w_{e1s2_3}$and $\beta^w_{e1s2_4}$ are all negative and significant. That indicates that there is no complementarity between work experience gained during the schooling interruption and subsequent education. The final interaction term between post-gap work experience and post-gap educational levels, $\beta^w_{e2s2_1}$, $\beta^w_{e2s2_2}$, $\beta^w_{e2s2_3}$and $\beta^w_{e2s2_4}$, are all positive but $\beta^w_{e2s2_1}$ are $\beta^w_{e2s2_2}$ are not significant. Thus, as expected wage growth are higher for those with more education but it is lower for the returners than for the non-returners.

6.4 Results for White Males from the 1979 and 1997 Cohorts

As mentioned above, I only observe respondents for up to 17 years in NLSY97 and to be consistent; I compare the NLSY97 results with results from using only the first 17 years of the NLSY79.

6.4.1 Parameters of the utility of schooling

As expected, all of the coefficients associated with the grade level are negative, with the exception of the coefficient of enrolled in high school in NLSY97, and they are all significant in NLSY97 but not in NLSY79, where only the first estimate that is associated with completing high school is significant. There are some interesting differences between the two cohorts. For instance, the coefficient of enrolled in a graduate program is about 2.25 times[19] lower than the coefficient of enrolled in an undergrad program in the 1997 cohort. For the 1979 cohort, the difference is only about 1.13 times[20]. It may indicate greater costs with higher education for those in the younger cohort, something that is consistent with the increases in direct costs observed over the last few decades. A similar difference exists for lower levels of post-secondary education.

For both cohorts, the coefficients of parental income, parental education and the effect of living in a nuclear family are strongly positively related to the utility of schooling. Also for both cohorts, the coefficient for the number of siblings, the effect of the return to school indicator, are strongly significant and negatively related to the schooling utility. Living in a southern state and living in a rural area are not significant factors related to the utility of schooling.

The estimated effects of cognitive ability on the utility of enrolling in school are, as expected, positive and significant for both cohorts. However, and consistent with recent and emerging work on comparative analyses between NLSY79 and NLSY97, the impact appears to be more substantial for the 1979 cohort (0.5860 compared to 0.3774 for the 1997 cohort).[21] Although it is not possible to directly associate and compare these two estimates, the difference in magnitudes suggests a decline in the importance of cognitive skills in educational decisions. Part of this reduced effect may be due to improved access to higher education in more recent years.

6.4.2 Parameters of the utility of working

[19]-3.12869 vs. -1.3869

[20]-0.4315 vs. -0.3755

[21]See for instance Castex and Dechter (2014) for an in-depth analysis of the changing roles of ability in explaining wage distributions.

The coefficients of educational level and work experience are significantly and positively related to the utility of working. The coefficient of working experience is about 2.8 times larger than the coefficient of schooling for the 1997 cohort, while the same ratio is about 1.2 for the 1979 cohort. It indicates, in the 1997 cohort, that each additional year of work experience will provide 2.8 times more utility units to working than each additional year of schooling. Hence the agent with more working experience will be more likely to choose to keep working instead of choosing study or consume NEET (i.e., chose NEET).

On the other hand, for the cohort of 1979, each additional year of working experience will provide about the same utility units to working as each additional year of schooling. Hence, for agents in the 1979 cohort, the decision to work is not as sensitive to prior working experience (or as persistent) as it is for agents in the younger cohort.

6.4.3 Parameters of the wage function

After controlling for unobserved ability, the coefficients of pre-gap schooling levels for the 1979 cohort ($\beta^w_{s1_1}$, $\beta^w_{s1_2}$, $\beta^w_{s1_3}$ and $\beta^w_{s1_4}$) are 0.0982, 0.1858, 0.4209 and 0.6114, respectively. Compared to the longer version of the same cohort (shown in the first column of Table 12), these estimates are lower, especially at higher levels of education. This clearly illustrates the importance of incorporating wage data from much of the life-cycle when estimating the returns to higher education and that the focus or usage (intended or unintended) of short-term data may seriously underestimate the economic benefits of higher education (as the benefits of these educational outcomes take longer to materialize).

For the younger cohort, the corresponding coefficients of pre-gap schooling levels display a similar, increasing pattern (-0.0496, 0.0162, 0.1140 and 0.3372, respectively) but the magnitudes are significantly lower at each level. The negative estimate for high school level of pre-gap schooling must be interpreted in its proper context where the variable is also interacted with work experience. Thus, it is not necessarily suggesting that there is a negative return to graduating from high school.

The coefficients of post-gap schooling levels ($\beta^w_{s2_1}$, $\beta^w_{s2_2}$, $\beta^w_{s2_3}$ and $\beta^w_{s2_4}$) are -0.0246, 0.0519, 0.1164, and 0.1031 for the 1979 cohort and

0.1558, 0.0314, 0.0632, and 0.0696 for the 1997 cohort. Two out four ($\beta^w_{s2_3}$ and $\beta^w_{s2_4}$) are significant for the 1979 cohort while the last three are significant for the 1997 cohort. This pattern may, in part, be due to the timing of interruptions with relatively few students returning after having stopped schooling before high school graduation.

When compared to the estimates for pre-gap schooling discussed above, the magnitudes of the post-gap schooling are substantially lower, especially for higher levels of education. For instance, the estimates of $\beta^w_{s1_3}$ and $\beta^w_{s1_4}$ (0.4209 and 0.6114, respectively, are between six and nine time larger than the corresponding estimates for post-gap schooling. This indicates that the return to education acquired after an interruption is significantly lower than the corresponding return is education is acquired continuously. However, there are still some modest wage gains associated with improvements in education. It should also be noted that both pre- and post-gap education are interacted with work experience which needs to be considered. As mentioned above, a more complete and comprehensive assessment of the impact on life-cycle income from an educational interruption will be presented below.

The estimates for work experience (when not interacted with education) are 0.0771 and 0.0983 for pre-gap experience for the 1979 and 1997 cohorts, respectively. For post-gap experience, the estimates are 0.0808 and 0.0997. The coefficients of the squared terms of these experience variables are -0.0029 and -0.0043 for pre-gap experience and -0.0047 and -0.0028 for post-gap experience. The first numbers in each set refer to the 1979 cohort while the second numbers are for the 1997 cohort. These are all statistically significant and show that the experience profiles are concave in both cohorts.

For NLSY79, the interaction terms between pre-gap schooling and work experience, $\beta^w_{e1s1_1}$and $\beta^w_{e1s1_2}$are not significant while $\beta^w_{e1s1_3}$and $\beta^w_{e1s1_4}$ are positive and significant. This means that wages grow faster with work experience for those with more pre-gap schooling in 4-year university or graduate school while is not the case for lower levels of pre-gap schooling (high school or 2-year college). The interaction terms of pre-gap work experience and post-gap educational levels,$\beta^w_{e1s2_1}$, $\beta^w_{e1s2_2}$, $\beta^w_{e1s2_3}$and $\beta^w_{e1s2_4}$ are-0.0086,-0.0208, -0.0303 and-0.0355. The signs of these are the same as those discussed above and indicate that there is no complementarity between work experience gained during the schooling interruption and subsequent education.

The first two interaction term between post-gap work experience and post-gap educational levels, $\beta^w_{e2s2_1}$and $\beta^w_{e2s2_2}$ are positive and significant (0.0270 and 0.0225) but the last two, $\beta^w_{e2s2_1}$ and $\beta^w_{e2s2_2}$ are not significant. This indicates that wages grow faster with work experience for those who return to complete high-school or a two-year college degree but not necessarily for those who return to complete a higher degree. This finding, combined with the result above where those who return and complete a higher degree do realize a (statistically) significant return, show that there is an immediate reward to return to school but that reward does not grow over time and with experience.

For NLSY97, the corresponding interaction terms between pre-gap schooling and work experience show a similar pattern with one exception. First, like the 1979 results, the first interaction (for high school) $\beta^w_{e1s1_1}$is not significant but the rest $\beta^w_{e1s1_2}$, $\beta^w_{e1s1_3}$and $\beta^w_{e1s1_4}$ are both positive and significant. Like before, wages grow faster with work experience for those with more pre-gap schooling and more so for the 1997 cohort than for the 1979 cohort.

The interaction terms of pre-gap work experience and post-gap educational levels,$\beta^w_{e1s2_1}$, $\beta^w_{e1s2_2}$, $\beta^w_{e1s2_3}$and $\beta^w_{e1s2_4}$ are-0.0185,-0.0308, -0.0184 and-0.0196. Similar to the 1979 results, these indicate that there is no complementarity between work experience gained during the schooling interruption and subsequent education. Finally, and in contrast to the result from NLSY79, the final interaction terms between post-gap work experience and post-gap educational levels, $\beta^w_{e2s2_1}$, $\beta^w_{e2s2_2}$ $\beta^w_{e2s2_1}$and $\beta^w_{e2s2_2}$ are all significant and negative. Taken outside their context, this indicates that for the 1997 cohort, wages decrease with work experience for those with more post-gap schooling. This counter-intuitive result may however depend on the estimates of the other variables involving post-gap human capital and I will show below that wages grow with experience also for those who return to school.

6.5 Black males from NLSY79 and NLSY97 with 17 periods of observations

The results in this section were obtained utilizing information on black respondents from both the original sample and from the over-sample of black respondents in both cohorts. Despite including the over-sample, the number of respondents is somewhat low (625 for the 1979 cohort and 425 for the 1997 cohort). Nevertheless, I believe it is important to estimate the parameters separately for each racial group in order to allow them to differ across groups. Similar to the presentation above, I will first present the estimates for the utility of school, followed by the utility of employment, wage function and finally, the estimates associated with cognitive scores.

6.5.1 Parameters of the utility of schooling

For the 1979 cohort, the coefficients associated with grade levels β^s_{ds1} ,β^s_{ds2}, β^s_{ds3} and β^s_{ds4} are -1.4791,-0.9658,-1.0504 and -1.1734, respectively All these estimates are significant. For the 1997 cohort, the coefficients associated with the same grade levels are also all negative but the magnitudes are lower in each case: -1.1545,-0.2459,-0.1173 and -0.5885, respectively. However, for the younger cohort, the estimates are not significant, with the exception of the first one that relates to high school.

To the extent that these estimates are capturing costs of attending a certain grade level and progressing through the educational system, the difference in estimates across the two cohorts suggests that the cost was higher in the 1980s (when the respondents of the 1979 cohort attended senior high school and college) than in the first decade of the new millenia.

When comparing these results with the corresponding ones for white males, we note that the pattern is opposite for the two groups. While education appears to have become less costly (or more accessible) for blacks over time, the contrary is true for whites where these estimates where relatively small and not significant for the 1979 cohort but large, negative and significant for the 1997 cohort.

A further indication that there has been a shift or change in the cost of and access to higher education for black males is the coefficient of parental income $\beta^s_{homeincome}$ which is 0.2102 and significant in NLSY79 and 0.0908 and not significant in NLSY97. Thus, the relationship between family income and educational attainment is weaker for the younger cohort. This is also a result that contrasts that reported for whites above.

The coefficients associated with parental education β^s_{hgcf} and β^s_{hgcm} are 0.0090 and 0.0812 for NLSY79 and 0.0631 and 0.0845, respectively for the NLSY97. With the exception of the coefficient for father's schooling in NLSY79, these are all significant and comparable to those reported above for whites. The coefficient of living in a nuclear family $\beta^s_{nuclear}$ is 0.0452 for the NLSY79 and 0.2377 for the NLSY97. The estimate is significant for the cohort in 1997 but not for the cohort in 1979 and suggests that family stability has become much more important for educational success and attainment than before. The magnitude of the estimate for NLSY97 is similar to that for white males of the same cohort (0.28).

The coefficient for number of siblings β^s_{nsib} is -0.0064 in NLSY79 and -0.0652 in NLSY97. The coefficient is significant for the 1997 cohort but not for the 1979 cohort. The negative association between family size and educational attainment may be due to lack of financial resources as the cost increases with family size. However, as discussed above, I have included measures of family income in the utility of school and this is arguably a better measure of financial resources of the household and its economic ability to support their children through school, suggesting that other factors may be at play in driving the change in the estimate for number of siblings.

Finally, the coefficient of return to school indicator $\beta^s_{I_{t-1}}$ is -3.1655 in NLSY79 and -3.1057 in NLSY97. The coefficient is negative and significant in both cohorts and larger in absolute terms than for whites. These suggest that, in both cohorts, black males are less likely to choose to go to school if they were not in school at the previous period. The coefficient of living in a southern state or in a rural area are not significant factors related to the utility of schooling.

6.5.2 Parameters of the utility of working

The coefficients for education and work experience, β^e_{educ} and β^e_{exp} are 0.2958 and 0.3628 in the NLSY79 and 0.1117 and 0.2881 in the NLSY97. All the coefficient are significantly and positively related to the utility of working. Similar to the ratio for white males, the coefficient for work experience is about 2.5 times larger than the coefficient for schooling for the 1997 cohort while the ratio is about 1.2 for the 1979 cohort. This indicates that, similar to their white counterparts in the NLSY97, additional work experience will provide 2.5 times units of work utility than additional schooling. Also, for the results for NLSY79, the estimated experience-education ratio in utility gains is similar to the one estimated for their white counterparts in NLSY79.

6.5.3 Parameters of the wage function

The coefficients of pre-gap schooling levels, $\beta^w_{s1_1}$, $\beta^w_{s1_2}$, $\beta^w_{s1_3}$ and $\beta^w_{s1_4}$, are 0.0406, 0.6781, 0.1134 and 0.1235, respectively in NLSY79. All of the coefficients of pre-gap schooling levels are significant. The corresponding coefficients for NLSY97 are 0.0605, 0.5153, 0.1131 and 0.2341, respectively and they are also significant.

The pattern with a peak for the completion of a two-year college and then lower returns for higher education is differs from what was observed for white males where the return increased gradually with completed education, which is the case for black males except for the large estimate for a two-year college. Aside from that estimate, there a gradual but small increase and although the magnitudes are higher for the 1997 cohort, at least for the first and last coefficients (high school and graduate school, respectively), they are much below the corresponding ones for white males (again, with the exception of the estimate for a two-year college degree).

The coefficients of post-gap schooling levels, $\beta^w_{s2_1}$, $\beta^w_{s2_2}$, $\beta^w_{s2_3}$ and $\beta^w_{s2_4}$, are 0.0131, 0.0122, 0.0431, and 0.0145, respectively in NLSY79. For the NLSY97, the coefficients are 0.0567, 0.0256, 0.0561, and 0.0231, respectively. With the exception of the estimate for high school in NLSY97, they are all significant.

These estimates are somewhat lower the those for white males but still suggest a positive wage impact of returning to school and complete a higher degree. For example, ignoring for the moment the interaction terms of post-gap schooling and other work experience, returning to complete a four-year college degree in NLSY79 increases the wage with over 4 percent. For NLSY97, that number is even higher (around 5.5 percent) and similar to that of returning to complete high school. However, and as was documented for white males, the returns to post-gap education are lower than those to pre-gap education. This will be further illustrated using model simulations below.

The estimates for work experience (when not interacted with education) are 0.0163 and 0.0528 for pre-gap experience for the 1979 and 1997 cohorts, respectively. For post-gap experience, the estimates are 0.0141 and 0.0461. In all cases, these estimates are significantly below those obtained for white males above. The coefficients of the squared terms of these experience variables are -0.0003 and -0.0027 for pre-gap experience and -0.0004 and -0.002 for post-gap experience. The first numbers in each set refer to the 1979 cohort while the second numbers are for the 1997 cohort. The estimates associated with pre-gap experience are both statistically significant and show that the experience profiles are concave in both cohorts. For post-gap experience, the estimates are not significantly different from zero.

For NLSY79, the interaction terms between pre-gap schooling and work experience, $\beta^w_{e1s1_1}$, $\beta^w_{e1s1_2}$, $\beta^w_{e1s1_3}$and $\beta^w_{e1s1_4}$ are all positive and significant. This means that wages grow faster with work experience for those with more pre-gap schooling. The interaction terms of pre-gap work experience and post-gap educational levels,$\beta^w_{e1s2_1}$, $\beta^w_{e1s2_2}$, $\beta^w_{e1s2_3}$and $\beta^w_{e1s2_4}$ are 0.023,-0.001, 0.114 and-0.0145 for NLSY79 and -0.0185,-0.0031, -0.1123 and-0.0112 for NLSY97. The estimates for NLSY79 are generally not significant while they are for NLSY97 which, like the results for white males, indicate that there is no complementarity between work experience gained during the schooling interruption and subsequent education.

For NLSY79, the interaction terms between post-gap work experience and post-gap educational levels, $\beta^w_{e2s2_1}$, $\beta^w_{e2s2_2}$, $\beta^w_{e2s2_3}$ and $\beta^w_{e2s2_4}$ are all significant (-0.0003, 0.0012, 0.1124 and 0.0781). These estimates suggest a slightly smaller wage growth for those who return to complete high school and a substantially greater wage growth for those returning to complete a 4-year college degree as well as for those returning to school in order to get a graduate degree. This pattern is also observed for NLSY97, although the magnitudes of the estimates are lower, closer to zero and with only one exception, not significantly different from zero. Those who return to school at grade levels above 16 see a significant wage gain, like the 1979 cohort, but almost half the size of that group. Further, this finding, combined with the result above where those who return and complete a higher degree do realize a (statistically) significant return, show that there is an immediate reward to return to school for the 1997 cohort but that reward does not grow over time and with experience

6.6 Hispanic males from NLSY79 and NLSY97 with 17 periods of observations

The results in this section were obtained utilizing information on Hispanic respondents from both the original sample and from the over-sample of Hispanic respondents in both cohorts. Despite including the over-sample, the number of respondents is somewhat low (425 for the 1979 cohort and 373 for the 1997 cohort). Nevertheless, as I mentioned above, it is important to estimate the parameters separately for each racial group in order to allow them to differ across groups. Similar to the presentation above, I will first present the estimates for the utility of school, followed by the utility of employment, wage function and finally, the estimates associated with cognitive scores.

6.6.1 Parameters of the utility of schooling

For the 1979 cohort, the coefficients associated with grade levels β^s_{ds1} ,β^s_{ds2}, β^s_{ds3} and β^s_{ds4} are -0.2290,-0.9255,-1.0648 and -1.4581, respectively. Except for the first estimate, these estimates are significant. For the 1997 cohort, the coefficients associated with the same grade levels are also all negative but the magnitudes are lower for higher levels of education: -1.0831,-0.4857,-0.4991 and -0.7015, respectively. For the younger cohort, the estimates are significant, with the exception of the third one that relates to 4-year college.

As for the results for black males just discussed, the difference in estimates across the two cohorts suggests that the cost was higher in the 1980s (when the respondents of the 1979 cohort attended senior high school and college) than in the first decade of the new millenia.

Also consistent with the results for black males, the coefficient of parental income $\beta^s_{homeincome}$ is larger in NLSY79 than in NLSY97, providing further evidence that there has been a shift or change in the cost of, and access to, higher education for non-white males. The estimate is 0.1059 in NLSY79 and

0.0476 in NLSY97. Both are significant. Thus, the relationship between family income and educational attainment is weaker for the younger cohort in contrast to the results for whites above.

The coefficients associated with parental education β^s_{hgcf} and β^s_{hgcm} are 0.0052 and -0.017 for NLSY79 and 0.0316 and 0.0152, respectively for the NLSY97. With the exception of the coefficient for father's schooling in NLSY97, none of these are significant. The coefficient of living in a nuclear family $\beta^s_{nuclear}$ is 0.2647 for the NLSY79 and 0.2617 for the NLSY97. The estimates are significant for both cohorts and comparable to those for white males, suggesting that family stability is important for educational attainment. The coefficient for number of siblings β^s_{nsib} is-0.0325 in NLSY79 and -0.0261 in NLSY97 but none of them are significant.

Finally, the coefficient of return to school indicator $\beta^s_{I_{t-1}}$is -2.888 in NLSY79 and -2.698 in NLSY97. The coefficients are negative and significant in both cohorts and similar to those for whites and blacks presented above.

6.6.2 Parameters of the utility of working

The coefficients for education and work experience, β^e_{educ} and β^e_{exp} are 0.3029 and 0.275 in the NLSY79 and 0.0634 and 0.2791 in the NLSY97. All the coefficient are significantly and positively related to the utility of working. Similar to the ratio for black and white males, the coefficient for work experience is substantially larger than the coefficient for schooling for the 1997 cohort while the ratio is close to one for the 1979 cohort. This indicates that, similar to their black and white counterparts in the NLSY97, additional work experience will provide more units of work utility than additional schooling. Also, for the results for NLSY79, the estimated experience-education ratio in utility gains is similar to the one estimated for their white counterparts in NLSY79.

6.6.3 Parameters of the wage function

The coefficients of pre-gap schooling levels, $\beta^w_{s1_1}$, $\beta^w_{s1_2}$, $\beta^w_{s1_3}$ and $\beta^w_{s1_4}$, are -0.0428, 0.2341, 0.3211 and 0.6435, respectively in NLSY79. All of the coefficients of pre-gap schooling levels are significant. The corresponding coefficients for NLSY97 are -0.0526, 0.2333, 0.2414 and 0.7675, respectively and they are also significant. The pattern with gradually increasing returns is similar to the one observed for white males but the magnitude of the estimates are a bit higher for Hispanics.

The coefficients of post-gap schooling levels, $\beta^w_{s2_1}$, $\beta^w_{s2_2}$, $\beta^w_{s2_3}$ and $\beta^w_{s2_4}$, are 0.0121, 0.0134, 0.1213, and 0.1723, respectively in NLSY79. For the NLSY97, the coefficients are -0.1231, 0.0144, 0.234, and 0.2321, respectively. With the exception of the estimate for high school in NLSY79, they are all significant.

These estimates are higher than the those for white males and suggest a positive wage impact of returning to school and complete a higher degree, especially at higher levels (4-year college and above). For example, ignoring for the moment the interaction terms of post-gap schooling and other work experience, returning to complete a four-year college degree in NLSY79 increases the wage with over 12 percent. For NLSY97, that number is even higher (over 23 percent). However, and as was documented for the other groups, the returns to post-gap education are lower than those to pre-gap education. This will be further illustrated using model simulations below.

The estimates for work experience (when not interacted with education) are 0.0236 and 0.0684 for pre-gap experience for the 1979 and 1997 cohorts, respectively. For post-gap experience, the estimates are 0.0658 and 0.0943. In all cases, these estimates are significantly above those obtained for black males above. The coefficients of the squared terms of these experience variables are -0.0023 and -0.0042 for pre-gap experience and -0.0028 and -0.003 for post-gap experience. The first numbers in each set refer to the 1979 cohort while the second numbers are for the 1997 cohort. The estimates associated with pre-gap experience are both statistically significant and show that the experience profiles are concave in both cohorts. For post-gap experience, the estimates are not significantly different from zero.

For NLSY79, the interaction terms between pre-gap schooling and work experience, $\beta^w_{e1s1_1}$, $\beta^w_{e1s1_2}$, $\beta^w_{e1s1_3}$and $\beta^w_{e1s1_4}$ are all significant and increasing with education. This means that wages grow faster with work experience for those with more pre-gap schooling. Similar results, but with slightly larger effects at each level, are observed for NLSY97. The interaction terms of pre-gap work experience and post-gap educational levels,$\beta^w_{e1s2_1}$, $\beta^w_{e1s2_2}$, $\beta^w_{e1s2_3}$and $\beta^w_{e1s2_4}$ are -0.0241, 0.0453, 0.0121 and 0.0247 for NLSY79 and -0.0191, 0.0245, 0.0133 and 0.0144 for NLSY97. Two of the four estimates for NLSY79 are significant while all the NLSY97 estimates are significant. Unlike the results for black and white males, there is complementarity between work experience gained during the schooling interruption and subsequent education for some levels of education, primarily for higher degrees.

For NLSY79, the interaction terms between post-gap work experience and post-gap educational levels, $\beta^w_{e2s2_1}$, $\beta^w_{e2s2_2}$, $\beta^w_{e2s2_3}$ and $\beta^w_{e2s2_4}$ are not significant (0.0036, 0.0038, 0.1163 and 0.1356), except for the one corresponding to 4-year college. That estimate, 0.1163, suggests a larger wage growth for those who return to complete a 4-year college degree. This pattern is also observed for NLSY97, although the magnitudes of the estimates are a bit lower.

7 Counterfactual Policy Simulations

As mentioned already, the estimates of the model are not always directly interpretable and to complement the discussion in the Results section above, I will use the model and the estimated parameters to simulate outcomes under different assumptions in order to illustrate more clearly the effects of interrupting schooling. Specifically, I consider three different scenarios to illustrate changes in wages and lifetime income across educational levels, with and without a schooling interruption.

In the first scenario, agents are forced to finish high school but are not allowed to continue investing beyond this level. Following high school graduation, they are allowed to optimize over the work and home options until they reach age 45 (when I use data on white males from the 30 periods of NLSY79) or age 32 (when I use data on all racial groups and for both NLSY79 and NLSY97). In the former case, I consider discounted, lifetime (until age 45) income while in the latter case, I use hourly wages instead since I only observe them until age 32.

In the second scenario, agents are forced to finish a university degree before they have to leave school and, again, are not allowed to re-enrolled in school. As in the first scenario, they are allowed to optimally choose work or home time from graduation until ages 32 or 45. Finally, in the third scenario, agents are also forced to have a university degree before they can permanently leave school, but are also forced to take a gap year between high school and university, a year during which they are assumed to be working. The rationale for this simulation exercise is to illustrate both the economic benefits with a university degree relative to a high school diploma as well as showing the impact a gap year may have on the estimated university return.

I will also use the model to show expected impacts of a hypothetical policy that provides financial support to families if the respondent enrolls in college. In this analysis, I increase the family income of agents taking college-level education by different amounts from \$3,500 per year up to \$30,000 per year. Such a subsidy may both increase and decrease the incidence of school interruptions. For example, the subsidy may make college more attractive to students who decided to leave school after graduating from high school, that is improving the re-enrollment rates and the proportion of students who interrupt school. It may also provide the means for students who take a gap year to save money for college, thereby decreasing the incidence of interruptions.

The rest of this section is organized as follows. I the next subsection, I will discuss results for white males from the long (30 periods) NLSY79 data. This is followed by a presentation of differences in returns and impacts across racial groups and over time. The following subsection will illustrate heterogeneity in returns and the effect of ignoring unobserved heterogeneity when estimating the model. Finally, this section ends with an analysis of the policy simulation described above.

7.1 White Males from the long NLSY79

7.1.1 Simulated discounted lifetime income

The reason why I devote an entire subsection to the analysis of white males from the 30 period (long) NLSY79 is that some, perhaps most, of the benefits to re-enroll in education will only materialize later on their careers and access to panel data over a long horizon is required. For the shorter (17 periods) panels, students who go to a 4-year college do not have enough time to recover the lost

Table 5: Simulated discounted, lifetime income (in 1997 U.S. dollars)
NLSY White Agents

	Income			
	Proportion	High School	College - no gap	College - gap
Type 1	16%	\$404,465	\$458,989	\$411,260
Type 2	5%	\$234,420	\$265,738	\$260,083
Type 3	60%	\$239,585	\$284,197	\$255,013
Type 4	19%	\$224,973	\$254,294	\$232,886
All		\$274,085	\$329,522	\$293,810

Note: The entries are based on a model that use up to 30 periods of data on white males from the NLSY79.

earnings while in college, and I consequently focus on the hourly wages for these groups instead.

In Table 4, I present generated discounted lifetime income associated with the three educational options described above: high school, college without any interruption or gap, and finally college following a one year gap between high school and college.[22] The results are illustrated separately for each type category as well as aggregated across all types and are expressed in 1997 \$US.

Overall, and averaged across types, the discounted present value of the earnings from age 16 to 45 for agents with a high school diploma but no further education is \$274,085, while it is \$329,522 if they have completed college without a gap year. This corresponds to a return of around 20 percent. If, however, the individual takes a gap year between high school and college, the discount earnings reduce to \$293,810, and this is only about 7 percent higher than the scenario of which the individual holds only a high school degree.

However, and as clearly illustrated in Table 3, the cost of a schooling interruption (as well as the return to college) varies significantly across individual types. For type 2 individuals, the return to college is similar, regardless of an interruption or not (13.4 percent versus 10.5 percent), while the difference is

[22]Income at any given age is given by the product of simulated hourly wages and 2,000 if they are predicted to work at that age. I chose 2,000 hours per year as this corresponds to full-time, full-year work and is the most common hours value reported in the data.

significantly higher for type 1 individuals, 13.5 percent as opposed to 1.7 percent. The relative return to a college degree for those who take a gap is highest among type 2 individuals while for those who enroll in college directly after high school, it is highest among type 3 individuals.

7.1.2 Simulated hourly wage rates at different ages

As mentioned above, I used simulated hourly wages when I generated simulated lifetime income. However, it is also interesting to illustrate the average wages at different ages (I choose 30, 40 and 45) across the three educational options considered in this section.

The results are presented in Table 5 below where the information on age and average hourly wages are supplemented by the percentage changes in wages for a certain age relative to average wages for those who go to college without a gap (third row) and those who take a gap year (fourth row). In all three options, average wages increase with age. However, the wage growth rates are almost twice as high for those with continuous college than for those with high school only. For those with a gap before college, the difference is much smaller but wages still grow faster for them than for high school graduates.

7.1.3 Simulation of wage growth and type effects

In this thesis, I have presented results suggesting that an agent's expected hourly wage rate when schooling is interrupted is much lower than he had completed his education without any interruption. However, expected wages and wage growth is still higher for those with a college education following a gap

Table 6: Simulated hourly wages, by education and age (in 1997 U.S. Dollars)

	High School			College - no gap			College - gap		
Age	30	40	45	30	40	45	30	40	45
Hourly wage	14.6	16.5	18	20.2	26.5	29.2	18.4	22.6	24.4
$\triangle COL\%$*	-27.7	-37.7	-38.4	0	0	0	-8.9	-14.7	-16.4
$\triangle Colgap\%$**	-20.7	-27.0	-26.2	9.8	17.3	19.7	0	0	0

Note: The entries are based on a model that use up to 30 periods of data on white males from the NLSY79.
*Percentage difference in average wages relative to those who acquired a 4-year college degree without a gap.
**difference in average wages relative to those who acquired a 4-year college degree without a gap.

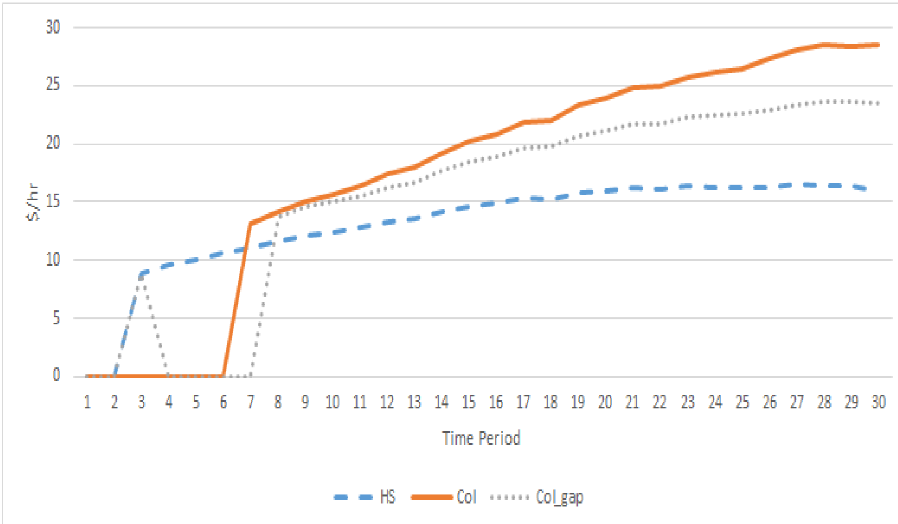
year than what they are for those who only completed high school. I also showed that there exists a substantial type variation in wages. Models ignoring types tend to underestimate wages in the options with college education, regardless of a gap or not.

To further illustrate the importance of controlling for unobserved heterogeneity, I show simulated wages across different ages under two cases: i) using my preferred model specification with four types and ii) based on a naive model specification that ignores unobserved heterogeneity in Figures 10 and 11. For this exercise, all agents completed their high school level education at age 18 and had no earnings prior to that. Hence, the income for periods one and two (ages 16 and 17) is 0.

The long dashed line shows the growth path for agents who only received high school level education. The short dashed line represents the scenarios where agents completed their college degree without a gap, and the solid line represents the wage growth path for the scenario where the agents complete the same college degree but had a gap year before they started their college studies (after graduating from high school).

From studying Figure 7and 8, we can identify three significant results. First, agents with college degrees will earn around 40 percent more than those with

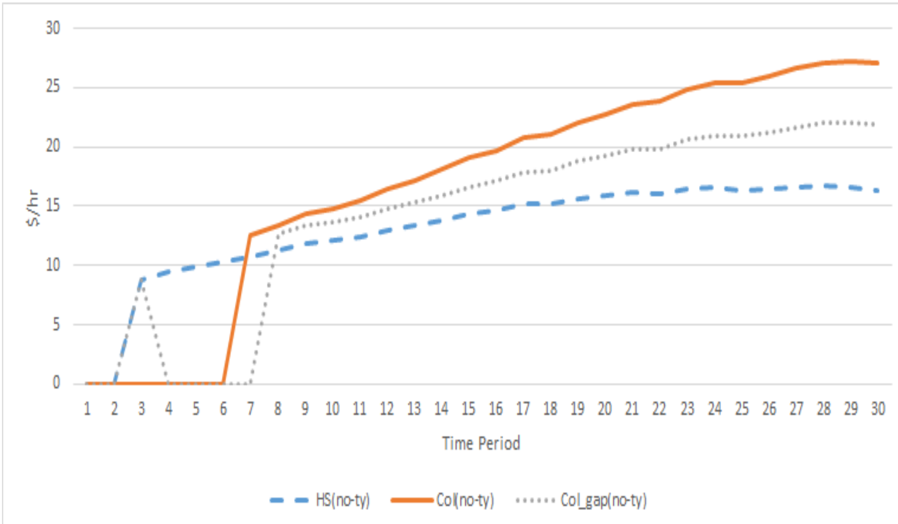
Figure 7: Wage paths for those with a high school degree, college degree without a gap and college degree with a gap year (4-type model)



only a high school degree. Second, agents who return to school after a gap year will earn around 25 percent more than those with only a high school degree. Third, those with a schooling interruption will initially have a higher income in the first years after college, but wage growth rate is much lower for them than for those without a gap so that, at age 46, wages among the 'non-interrupters' has surpassed wages among those who took a gap.

By comparing Figures 7 and 8, we can also see that the results from the naive model provide reasonable values for those with high school only but underestimate wage growth among those with a college degree, both with and without a gap year. This bias arises even when I include a wide range of observable characteristics, including AFQT scores and illustrates the importance of controlling for unobserved heterogeneity in wages that is allowed to be correlated with unobserved heterogeneity in the choices of education and working.

Figure 8: Wage paths for those with a high school degree, college degree without a gap and college degree with a gap year (single-type model)



7.2 Racial Differences - Males from NLSY79 and NLSY97

In this section, I will analyze both differences in the costs and benefits of taking a gap year across racial groups but also how these have changed over time and if the changes are equal across race. As is well documented, the cost of attending higher education such as obtaining a Bachelor's degree has increased significantly since the 1970s and this may have impacted the incidence of school interruptions (students may have to take a break from studies to work and save money for tuition and other fees).[23] At the same time, the returns to higher education have increased which may also impact school interruptions and the decision to re-enroll.

Regarding the racial differences, this analysis is, to my knowledge, the first to detail the incidence of and costs-benefits from school interruptions across the three major racial groups in the U.S. and how these have changed over time as

[23]See for instance Abel and Deitz (2014).

access to higher education has changed.

In Table 6, I present the average hourly wages for the three educational categories defined above and separately for white, black and Hispanic males. They are measured at age 32 which is the latest age I observe them in each cohort and, consequently, as far from entry wages following school leaving as I can get. The top panel shows the entries for respondents to the NLSY79 while the bottom panel displays the same information for the respondents of NLSY97. All wages are expressed in constant 1997 dollars.

Starting with the results for the 1979 cohort, the wage return to a 4-year college degree without any gap in acquiring it varies from 13.6 percent for Hispanics (\$16 versus \$12.6) to 45 percent for blacks (\$15 versus \$10.4). For white males, the return is close to 26 percent (\$16.8 versus \$13.3).

The corresponding wage returns for a college degree with a one year gap show a similar racial pattern with the lowest return for Hispanics and highest for blacks. However, as we have seen before, the magnitudes of the returns are lower compared to college without a gap in each case. Specifically, the return for Hispanics is 6.4 percent for Hispanics (\$15 versus \$12.6), 23.4 percent for blacks (\$12.8 versus \$10.4) and 12.6 percent (\$15 versus \$13.3) for whites. This is, the relative wage differences are only about half the size of those when enrolling in college directly from high school. Nevertheless, there are positive and significant returns to college, even if it follows a gap year.

For the younger, 1997 cohort, the wages are presented in the lower panel of Table 6. Generally, as expected, there is an increase in average wages across all groups and educational levels. However, for Hispanics the wage differences are modest compared to those for whites. High school wages grew over this 20 year period (the 1979 wages are observed in the first half of the 1990s while the 1997 wages are observed during the years 2011-2014) by a modest 2.7 percent for Hispanics and by 22 percent for whites. For blacks, the growth was even higher at 26.9 percent but despite the higher growth, their average wage was still below that of Hispanics.

Looking at average college wages (without interruptions), the grew by 34.5 percent for whites but much less for blacks (5.9 percent) and Hispanics (7 percent). The average wages for college with a gap year, the corresponding increases between the two cohorts are 38.9 percent for whites, 17 percent for blacks and 3.5 percent for Hispanics. One implication of these changes is that the difference in mean wages for those with a college degree has increased and while white males saw significant growth in real wages, this was not the case for the other groups.

The racial differences in returns to a college degree for the younger cohort more pronounced than they are for the older cohort. The wage return to a 4-year college degree without any gap varies from 18.4 percent for Hispanics ($17.2 versus $14.5) to 39 percent for whites ($22.6 versus $16.3). Although there is an increase for both groups, it is much larger for whites. For black males, the return is 21 percent ($15.9 versus $13.2) which, in relative terms, is below the corresponding return for the older cohort.

The corresponding wage returns for a college degree with a one year gap show a similar racial pattern with the lowest return for Hispanics and highest for whites. The magnitudes of the returns are again lower than those when college is acquired without a gap. Specifically, the return for Hispanics is 7.2 percent ($15.6 versus $14.5), 13.8 percent for blacks ($15 versus $13.2) and 28.2 percent ($20.9 versus $16.3) for whites. Compared to the 1979 cohort, the college return with and without a gap are higher for blacks and whites but lower for Hispanics. In the older cohort, the college return with a gap year was around 50 percent of the return without a gap. For whites from the younger cohort, the

Table 7: Average Hourly Wages at Age 32 (in 1997 U.S. dollars)

	High School	College - no gap	College - gap
NLSY79			
White	\$13.3	\$16.8	\$15.0
Black	\$10.4	\$15.5	\$13.8
Hispanic	\$10.6	\$16.0	\$13.0
NLSY97			
	High School	College - no gap	College - gap
White	\$18.3	\$21.6	\$20.5
Black	\$14.2	\$18.9	\$17.5
Hispanic	\$13.5	\$18.1	\$18.0

Note: The entries are based on a model that use up to 17 periods of data on from the NLSY79 and NLSY97, respectively.

difference between the two college returns is smaller, around 25 percent while for blacks, it is around 33 percent.

To summarize, there is a wage gain associated with completing a 4-year college degree and it is higher if students do not interrupt their studies. However, even if they do, there is a significant wage return. Moreover, the returns have become closer over time for whites and blacks making the return to college relatively more beneficial now than in the past. Finally, the college returns are lowest for Hispanics, regardless of the presence of a gap or not and this group also experience the lowest growth in average wages over time.

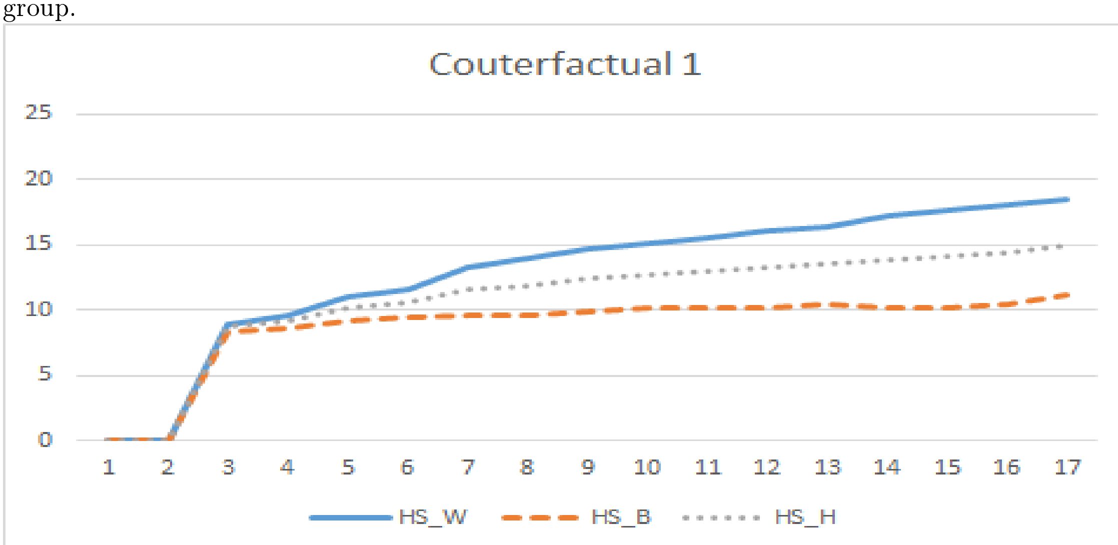
Figures 5, 6 and 7 show expected wage growth paths for the three different racial groups under the three different counterfactual scenarios discussed above using data from NLSY97. The vertical axis shows the hourly wage (in U.S. 1997 dollars), and the horizontal axis shows the decision horizon which is starts at age 16 and ends at age 33. In the first scenario, all agents complete their high school level education on time at age 18 and have no wages before that age. Hence, wages for periods 1 and 2 (ages 16 and 17) are 0. The solid line is the growth path for white males, the dashed line is the wage growth path for black males and the dotted line is the wage path for Hispanic males. Over the 15 years following high school, wages grow the most for whites (from just below \$10 the first year after high school to almost \$20 at age 33, corresponding to an annual growth rate of around 6.5 percent) and the least (in fact, almost not at all) for blacks.

In scenario 2, agents enroll in college directly after high school and complete a 4-year degree. Again, it is assumed that they are not working while in school. The pattern across the racial groups from above remains, with white showing the highest growth and black the lowest. Over the first 11 years after college, expected wages grow by around 6 percent per year for white males and by about half of that for black males. Compared to the paths for the high school scenario, the variation in wage growth rates across the racial groups is much smaller in the college educated case. Further, like the high school case, initial wages are compressed and very similar for whites, blacks and Hispanics.

Finally, in scenario 3, agents complete a 4-year college but have a gap year before they enroll in college. The wage received between high school and college is similar for the three groups as this corresponding to the initial wage following high school described above, which display very little variation across race. Following college graduation, wage grow somewhat less than those for college graduates without a gap year. For whites (the group with the highest growth rates), wages grow by around 4.5 percent per year, and again by about half of this rate for black males. Even though the growth rates are lower in this case, the initial wages after college graduation are somewhat higher that for those without a gap.

To summarize, there are two main conclusions from this exercise. First, individuals start with similar wages in all counterfactual scenarios, but the growth rate of white males is much higher than that of black and Hispanic males. Sec-

Figure 9: Simulated wage paths for the case with high school only, by racial group.



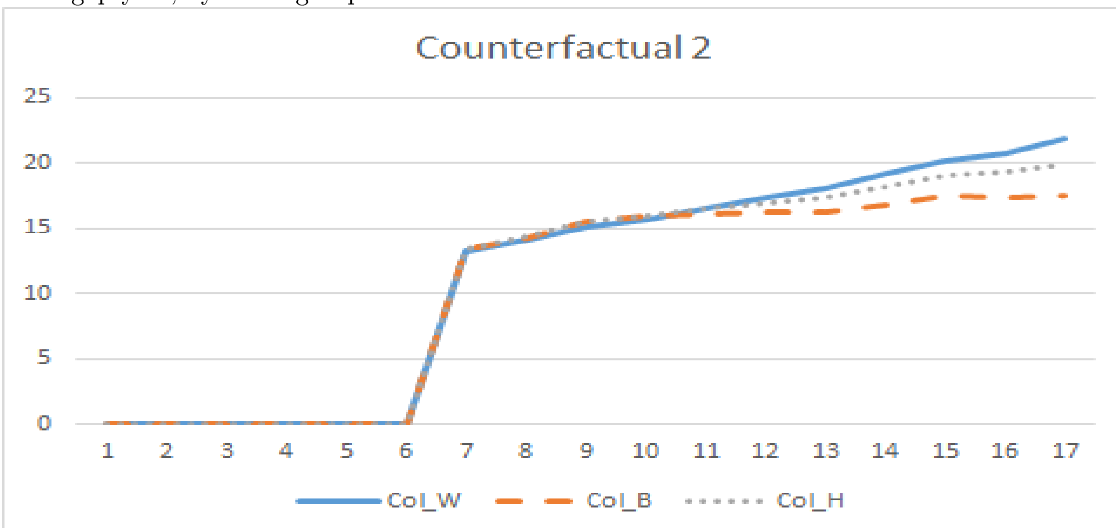
Note: The entries are based on a model that use up to 17 periods of data on from NLSY97.

ond, the difference in growth rates is more significant in the high school only scenario than in those involving a college education, which suggests that as educational levels increase, the wage gap between racial groups is reduced.

7.3 Policy Simulation

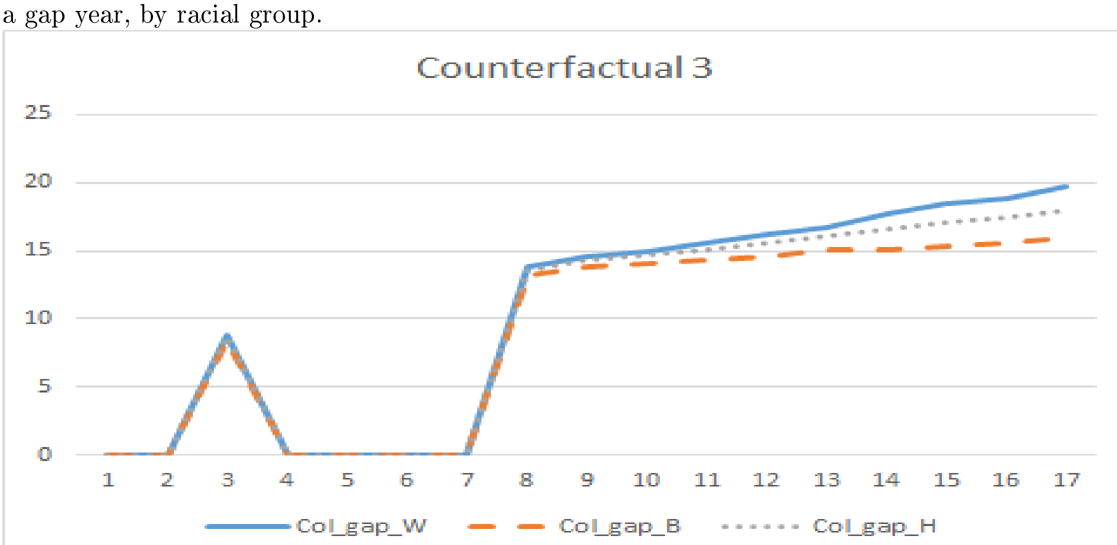
The first set of results presented in this section for white males from the NLSY79 showed that individuals gained a 7 percent increase in expected, discounted lifetime income if they return to school and increase their educational attainment, which also means increased output for the entire economy. Hence, it is natural for policymakers to consider alternative ways to encourage indi-

Figure 10: Simulated wage paths for the case with a 4-year college degree without a gap year, by racial group.



Note: The entries are based on a model that use up to 17 periods of data on from NLSY97.

Figure 11: Simulated wage paths for the case with a 4-year college degree with a gap year, by racial group.



Note: The entries are based on a model that use up to 17 periods of data on from NLSY97.

viduals without higher education to return to school and gain a college-level education. In my model, individuals choose to enroll in school if the present value of schooling at the beginning of that period (t),$V^s_{i,t}$, is larger than the present values of the other options, working $(V^e_{i,t})$ and NEET$(V^h_{i,t})$. From the literature, the most cost-effective treatment to increase educational attainment is school performance-based subsidies which aim to reduce the educational cost.

In this subsection, I present results from a policy simulation whereby I increase family income of individuals enrolling in college-level education. I simulate the effects of annual government subsidies in the amounts of: \$20,000, \$25,000 and \$30,000. Specifically, I modify the utility of enrolling in school by increasing family income by one of the amounts above if individuals enroll in a college-level education and for a maximum of four years. In the case of a \$20,000 subsidy, the instant utility function of schooling in time period t is:

$$\begin{aligned}\hat{U}^s_{i,t} \quad = \quad & \widehat{\alpha}^s_m + \acute{X}_i\widehat{\beta}^s_x + ability_i\widehat{\beta}_{ab} + \widehat{\beta}^s_g ds^{t-1}_{i,t} + \widehat{\beta}^s_{gd} dgapdu_{i,t} \\ & + \hat{\beta}^s_{ds1} ds_{i,t}\,(1) + \beta^s_{ds2} ds_{i,t}\,(2) + \hat{\beta}^s_{ds3} ds_{i,t}\,(3) + \hat{\beta}^s_{ds4} ds_{i,t}\,(4) \\ & + \widehat{\beta}^s_{inc}(Homeincome + \$20,000) + \varepsilon^s_{i,t},\end{aligned} \tag{23}$$

where $\acute{X}$ is the array of all the background characteristic variables except home (or family) income.

The results from this exercise are presented in Table 7 below. The first row shows how the highest grade completed change as the subsidy increases and the impacts are quite modest which is not surprising as this subsidy have little impact on the majority of individuals who at most complete high school.[24] With a \$30,000 annual subsidy, the highest grade completed is predicted to increase

[24]It is possible that a college subsidy may increase high school graduation rates among forward looking individuals. In my simulations, the high school graduation rates were however unchanged.

Table 8: Simulated impacts of alternative college subsidies - white males from NLSY79

	Baseline	Annual subsidy amounts		
		\$20,000	\$25,000	\$30,000
Highest Grade Completed	13.29	13.45	13.47	13.51
%△ in HGC	-	1.2	1.4	1.7
Hourly Wage	24.6	25.6	25.7	26
%△ Hourly Wage	-	4.1	4.5	5.7
Two-year College	37.95	40.95	41.28	41.76
%△ Two-year College	-	7.9	8.8	10.0
Four-year College	19.02	20.85	21.52	22.19
%△ Four-year College	-	9.6	13.1	16.7
Proportion with a gap year	40.63	41.92	42.26	42.16
%△ Proportion with a gap	-	3.20	4.0	3.8

Note: The entries are based on a model that use up to 33 periods of data on from NLSY79. In 1997 U.S. dollars.

by 1.7 percent.

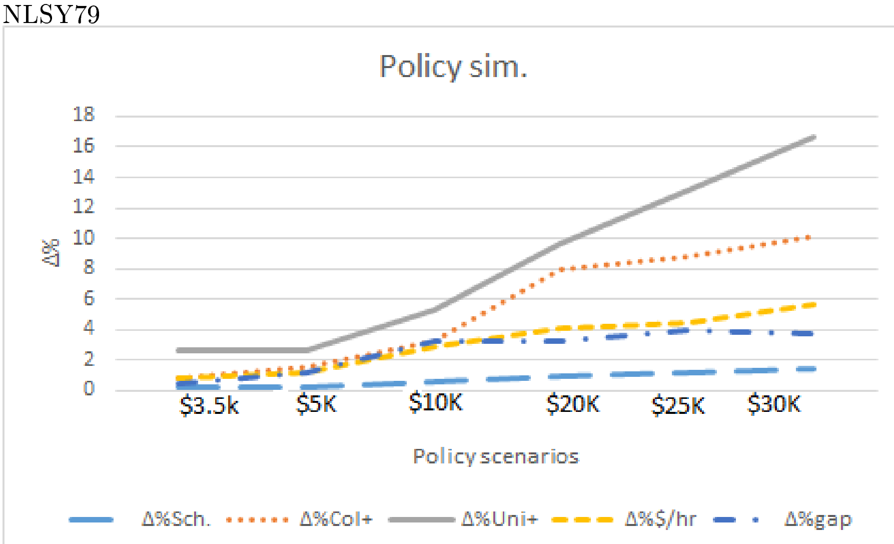
In terms of wage impacts, the subsidies are expected to increase hourly wages as education increases, measured as late as possible which for these data means at age 46, by 4.1 percent (for the \$20,000 subsidy) and by 5.7 percent for the \$30,000 subsidy). These wage increases follow from improved enrollment and graduation rates from college and this, not surprisingly, where the largest impacts are observed. For the \$30,000 subsidy, the proportion of individuals with at least two years of college increase with 10 percent. For the same subsidy level, the proportion of individuals with at least four years of college increase with close to 17 percent.

Lastly, I use the simulations to assess the impact on school interruptions between high school and college. As hypothesized before, a college subsidy may both increase and decrease the incidence of an interruption. An increase may occur if individuals have strong preferences against higher education after completing high school and choose to leave school. However, the subsidy may bring them back to school by increasing the value of the school option. Similarly, a college subsidy may decrease the interruption rates by affording potential college students to enroll directly after high school instead of taking a year of to work and save money to pay for college. The entries in the table show that interruptions increase marginally, by 3-4 percent depending on the subsidy amount. Thus, making college more affordable appears to bring students back to school, increase the proportion of workers with a college degree and improve their average wages.

In Figure 12 I plot the percentage changes in the outcomes considered in Table 7 for six different subsidy amounts. In addition to those presented in table 7, I also simulated the effects of subsidies equal to \$3,500, \$5,000 and \$10,000. The figure clearly shows that the impact, regardless of subsidy level, is largest for the four-year college option (the solid line) and that its effect grows with the magnitude of the subsidy. The proportions with two-year and four-year college degrees increase as well and, in magnitudes, in the same way (most for Hispanics and least for whites).

The results above were obtained for the long panel of white males from the NLSY79 cohort. Below, I present the corresponding information for the three racial groups for both cohorts (NLSY79 and NLSY97) for the subsidy amount of \$20,000. Starting with NLSY79, the entries in Table 8 below show an increase in the average highest grade completed for all three groups, with the largest increase for Hispanics (2.3 percent). The wage effects are also positive for all groups and again largest for Hispanics (an increase with 7.3 percent, followed by blacks at 5.3 percent and smallest for white males at 4.1 percent). The proportion of individuals with at least two years of college increase with 7.9 percent for whites, 10.4 percent for blacks and 13.3 percent for Hispanics. The corresponding proportion of individuals with at least four years of college increase with close to 9.6 percent for whites, 13.7 percent for blacks and 15.6 percent for Hispanics. Finally, the table entries show that interruptions increase

Figure 12: Simulated impacts of alternative college subsidies - white males from NLSY79



Note: The graph is based on a model that use up to 33 periods of data on from NLSY79. The different lines show predicted changes in different outcomes (highest grade completed (Sch.), two-year college (Col+), four-year college (Uni+), hourly wage ($/hr) and school interruption (gap)) relative to the baseline with no subsidy as the value of the subsidy increases. In 1997 U.S. dollars.

Table 9: Simulated impacts of alternative college subsidies - males from NLSY79

	White		Black		Hispanic	
Outcome	Baseline	$20,000	Baseline	$20,000	Baseline	$20,000
Highest Grade Completed	13.3	13.5	12.3	12.3	11.9	12.2
%△ in HGC	-	1.0	-	0.7	-	2.3
Hourly Wage	24.6	25.6	15.9	16.8	17.0	18.2
%△ Hourly Wage	-	4.1	-	5.3	-	7.3
Two-year College	38.0	41.0	16.0	17.6	13.6	15.4
%△ Two-year College	-	7.9	-	10.4	-	13.3
Four-year College	19.0	20.8	11.9	13.5	9.9	11.5
%△ Four-year College	-	9.6	-	13.7	-	15.6
Proportion with a gap year	40.6	41.9	38.5	39.1	36.7	38.7
%△ Proportion with a gap	-	3.2	-	1.5	-	5.3

Note: The entries are based on a model that use up to 17 periods of data on from NLSY79. In 1997 U.S. dollars.

marginally, by 1.5 percent for blacks, 3.2 percent for whites and 5.3 percent for Hispanics. Thus, much of the educational improvement following the subsidy arises because students are incentivized to re-enroll in school and attend (and complete) college and this effect is the largest for Hispanic students, who have the lowest college graduation rates in the data.

In Table 9 below, I present the corresponding information for the three racial groups for the younger cohort (NLSY97) for the subsidy amount of $20,000. The entries in the first row show a modest increase in the average highest grade completed for all three groups, around one percent for all groups. The wage effects are also positive for all groups and largest for Hispanics (an increase with 10.1 percent, followed by blacks at 9.8 percent and smallest for white males at 7.1 percent). These wage gains are larger than those obtained for the older cohort above. The proportion of individuals with at least two years of college increase with 11.3 percent for whites, 14.3 percent for blacks and 18.3 percent for Hispanics. The corresponding proportion of individuals with at least four years of college increase with close to 12.1 percent for whites, 9.1 percent for blacks and 10.2 percent for Hispanics.

Table 10: Simulated impacts of alternative college subsidies - males from NLSY97

	White		Black		Hispanic	
Outcome	Baseline	\$20,000	Baseline	\$20,000	Baseline	\$20,000
Highest Grade Completed	14.1	14.3	12.5	12.6	12.7	12.8
%△ in HGC	-	1.0	-	1.1	-	1.1
Hourly Wage	26.7	28.6	20.8	22.8	22.8	25.1
%△ Hourly Wage	-	7.1	-	9.8	-	10.1
Two-year College	45.7	50.9	29.4	33.6	27.1	32.1
%△ Two-year College	-	11.3	-	14.3	-	18.3
Four-year College	28.6	32.1	16.1	17.6	16.5	18.2
%△ Four-year College	-	12.1	-	9.1	-	10.2
Proportion with a gap year	41.5	43.6	38.4	40.8	44.8	48.1
%△ Proportion with a gap	-	5.1	-	6.3	-	7.4

Note: The entries are based on a model that use up to 17 periods of data on from NLSY97. In 1997 U.S. dollars.

Finally, the table entries show that interruptions increase more for this cohort than for the 1979 cohort. As before, the increase is smallest for whites (5.1 percent), larger for blacks (6.3 percent) and largest for Hispanics (7.4 percent). Thus, close to half the Hispanic males are predicted to return to school following an interruption with the introduction of a \$20,000 college subsidy. Generally, the estimated reactions are larger for the 1997 cohort than for the older one, possibly due to the increased value associated with education more recently.

8 Conclusion

In this thesis, I have provided new evidence on the cost and benefits associated with an interruption to schooling investments. This topic has recently attracted much attention, especially given the increase in the cost of higher education over the last couple of decades. I specify and estimate a dynamic, discrete choice model that includes a very flexible wage function. It allows the estimation of different wage growth rates before and after a schooling interruption (or between those who interrupt and those who do not).

In one set of analysis, I use data from the 1979 cohort of the NLSY which allows me to analyze the long-term effects of a schooling interruption since I can follow workers from age 16 to their late 40s. It contrasts many of the previous studies which have utilized data with information on labor market outcomes immediately after graduation. I also utilize data from the 1997 cohort of the NLSY to analyze if the incidence, as well as the cost and benefits of an interruption, has changed over time. I also use data from both cohorts to compare racial differences in school interruptions and their effects.

The predictions generated by my model, along with other dimensions including the decision to interrupt school and wages, are very similar to those observed in the data. I am therefore confident that the model can be used to analyze different counterfactual policies. In this paper, I argue that the decision to interrupt schooling is endogenous, as is education and work experience, and ignoring this may contaminate the estimation results. I show that a temporary interruption to schooling has a negative effect on an agent's future income compared to completing the same level of education without interruption. The difference in present value earnings, assuming a one-year interruption during which the individual is working, varies across individuals, and can be as low as \$5,000 over a 30 year period and as high as \$48,000, over the same period. However, it is considerably higher for agents with only high school degrees.

From policy simulation, I find that under a post-secondary study subsidy, more agents will make a choice to return to school and obtain a post-secondary degree, and the hourly wage will also increase considerably. This thesis provides quantitative evidence on the effect of a college-level subsidy on the racial differences in wages and schooling between white, black and Hispanic males. The results demonstrated that a conditional (upon college enrollment) subsidy could significantly increase schooling attainment for all three racial groups but more so for the underprivileged groups and consequently help reduce racial wage and employment gaps. Although most of the results I present focus on interruptions between high school and college, my model is more general than that, and it allows for the possibility of interruption at any school level.

References

[1] Abel, R. Jaison and Richard Deitz. "Do the Benefits of College Still Outweigh the Costs?" *Current Issues in Economics and Finance* 20, (2014): 1-11.

[2] Altonji, G. Joseph. "The Demand for and Return to Education When Education Outcomes are Uncertain." *Journal of Labor Economics* 11, issue 1, (1991): 48-51.

[3] Altonji, G. Joseph and Thomas A. Dunn. "The Effects of School and Family Characteristics on the Return to Education." *The Review of Economics and Statistics* 8, no. 4, (1995): 665-662.

[4] Belley, Philippe and Lance Lochner. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *Journal of Human Capital.* (2007): 37-39.

[5] Bellman, Richard. "Dynamic Programming" Princeton, NewJersey, Princeton University Press.(1957)

[6] Belzil, Christian. "Testing the Specification of the Mincer Wage Equation." *SSRN Electronic Journal*, (2006): 427-451.

[7] Belzil, Christian, Jorgen Hansen and Xingfei Liu, "Dynamic skill accumulation, education policies, and the return to schooling" *Quantitative Economics.*Volume8, Issue3 (2017):895-927

[8] Blake, Judith. "Number of siblings and educational attainment." *Science* 245, no. 4913 (1989): 32-36.

[9] Bozick, Robert and Stefanie DeLuca. "Better Late Than Never? Delayed Enrollment in the High School to College Transition." *Social Forces* 84, no. 1 (2005): 531-54.

[10] Butler, Rebecca A. "An Examination of Academic, Financial, and Societal Factors Impacting the Decision to Delay Entry to College and Subsequent Workforce Implications"(2016)

[11] Cameron, Stephen and James Heckman. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts." *Journal of Political Economy* 106, issue 2 (1998): 262-333.

[12] Coleman, S. James, Berry C. Charles and Zahava D. Blum, "White and Black careers during the first decade of labor force experience. Part III: Occupational status and income together." *Social Science Research,*Volume 1, Issue 3, (1972): 293-304

[13] Dominic, Christoph Gwosc Orr and Nicolai Netz, "Social and Economic Conditions of Student Life in Europe." Eurostudent IV 2008-2011.

[14] Ferrer, M. Ana and Alicia Menendez. "The Puzzling Effects of Delaying Schooling on Canadian Wages." *Canadian Public Policy* 40, no. 3 (2014): 197-208.

[15] Finnie, Ross and Hanqing Qiu. "The Patterns of Persistence in Post-Secondary Education in Canada: Evidence from the YITS-B Data set" Toronto, Ontario: Canadian Education Project (2008): 89.

[16] Finnie, Ross and Richard E. Mueller . "The Effects of Family Income, Parental Education and Other Background Factors on Access to Post-

Secondary Education in Canada: Evidence from the YITS." *SSRN Electronic Journal*, (2008): 89-97.

[14] Foley, Kelly and Fane Gores. “Field of Study and the Decision to Delay University”, Working paper (2016)

[17] Fortin, Bernard, and Safa Ragued. "Does temporary interruption in post-secondary education induce a wage penalty? Evidence from Canada." *Economics of Education Review* 58 (2017): 108-22.

[18] Geweke, F. John , Michael P. Keane, Roberto Mariano, Til Schuermann, and Melvyn J. Weeks. "Bayesian inference for dynamic discrete choice models without the need for dynamic programming." *Simulation-based Inference in Econometrics*(2005): 100-31.

[19] Gamoran, Adam. "American Schooling and Educational Inequality: A Forecast for the 21st Century." *Sociology of Education* 74 (2001): 135-141.

[20] Grissmer, W. David, Sheila Nataraj Kirby, Mark Berends, Stephanie Williamson, “Student Achievement and the Changing American Family” *Rand* (1994)

[21] Heckman, James. “The American Family in Black and White: A Post-Racial Strategy for Improving Skills to Promote Equality." *Sociology of Education* 89, no.6 (2011): 45-47.

[22] Heckman, James. “Human Capital Pricing Equations with an Application to Estimating the Effect of Schooling Quality on Earnings.” *The Review of Economics and Statistics* 78, no. 4 (1996): 562-610

[23] Hedges, V. Larry and Amy Nowell. "Changes in the Black-White Gap in Achievement Test Scores." *Sociology of Education* 72, no. 2 (1999): 111.

[24] Jencks, Christopher and Meredith Phillips. The black-white test score gap. Washington (D.C.): *Brookings institution* 64 (1998): 127.

[25] Johnson, T. Matthew "Borrowing Constraints, College Enrollment, and Delayed Entry." *Journal of Labor Economics* 31, no. 4 (2013): 669-725.

[26] Keane, P. Michael , and Kenneth I. Wolpin. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *SSRN Electronic Journal* 42, no. 4, (2001): 1051-1103.

[27] Keane, P. Michael and Kenneth I. Wolpin. "Eliminating Race Differences in School Attainment and Labor Market Success." *Journal of Labor Economics* 18, no. 4 (2000) 34-45.

[28] Kenyon A. Daphne, " The Property Tax-School Funding Dilemma" *Policy Focus Report* • Lincoln Institute of Land Policy(2007)

[29] Lareau, Annette. "Social Class Differences in Family-School Relationships: The Importance of Cultural Capital." *Sociology of Education* 60, no. 2 (1987): 73-85.

[30] Light, Audrey. "The Effects of Interrupted Schooling on Wages." *The Journal of Human Resources* 30, no. 3 (1995): 472-478.

[31] Lynn, Richard. "Student achievement and the changing American family." *Personality and Individual Differences* 20, no. 1 (1996): 128-29.

[32] Mincer, Jacob. "Investment in Human Capital and Personal Income Distribution." *Journal of Political Economy* 66, no. 4 (1958): 281-302.

[33] Mincer, Jacob. "Schooling, Experience, and Earnings". *Human Behavior & Social Institutions* No. 2.(1974): 67.

[34] Monks, James. "The impact of college timing on earnings." *Economics of Education Review* 16, no. 4 (1997): 419-23.

[35] Nathan, Mitchell J. "Rethinking Formalisms in Formal Education." *Educational Psychologist* 47, no. 2 (2012): 125-48.

[36] Niu, Sunny and Marta Tienda. "Delayed Enrollment and College Plans: Is There a Postponement Penalty?" *The Journal of Higher Education* 84, no. 1 (2013): 1-26.

[37] Niu, Sunny and Marta Tienda. "Delayed Enrollment and College Plans: Is There a Postponement Penalty?" *The Journal of Higher Education* 84, no. 1 (2013): 1-26.

[38] O'Toole, M. Dennis, Stratton, S. Leslie and James N. Wetzel, "A longitudinal analysis of the frequency of part-time enrollment and the persistence of students who enroll part-time", *Research in Higher Education* 44, no.5 (2003): 519-537.

[39] O'Toole, M. Dennis, Stratton, S. Leslie and James N. Wetzel, "A multinomial logit model of college stopout and dropout behavior." *Economics of Education Review* 27, no. 3 (2008): 319-31.

[40] Rowan-Kenyon, T. Heather."Predictors of Delayed College Enrollment and the Impact of Socioeconomic Status." *The Journal of Higher Education* 78, no. 2 (2007): 188-214.

[41] Rust, John. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55, no. 5 (1987): 999-1013.

[42] Tomkowicz, Joanna and Tracey Bushnik, "Who goes to post-secondary education and when: pathways chosen by 20 year-olds," Catalogue no. 81-595-MIE, Statistics Canada (2003).

[43] Wells, S. Ryan and Cassie M. Lynch. "Delayed College Entry and the Socioeconomic Gap: Examining the Roles of Student Plans, Family Income, Parental Education, and Parental Occupation." *The Journal of Higher Education* 83, no. 5 (2012): 671-97.

APPENDIX

Table 11: MLE estimates from the dynamic model

	NLSY79 (30priods)
Parameters	Estimates (Std.err)
$\breve{a}^s_{m=1}$	6.1185*
	(0.3014)
$\breve{a}^s_{m=2}$	1.0375 *
	(0.2759)
$\breve{a}^s_{m=3}$	6.2248 *
	(0.2969)
$\breve{a}^s_{m=4}$	2.1085 *
	(0.2840)
β^s_{ds1}	-1.9122*
	(0.0854)
β^s_{ds2}	-0.5961*
	(0.1150)
β^s_{ds3}	-0.7583*
	(0.1606)
β^s_{ds4}	-0.8403*
	(0.2242)
$\breve{a}^s_{cog}$	0.6444 *
	(0.0515)
β^s_{Snisb}	-0.0505*
	(0.0152)
$\beta^s_{Snuclear}$	0.2647*
	(0.0726)
β^s_{Ssouth}	-0.0060
	(0.0447)
β^s_{Srural}	0.0231
	(0.0579)
β^s_{Shgcf}	0.0577*
	(0.0108)
β^s_{Shgcm}	0.0575*
	(0.0142)
$\beta^s_{Shomeincome}$	0.0036*
	(0.0009)

* Significant at 1% level.

Table 12: MLE estimates from the dynamic model Con't 1
NLSY79 (30periods)

Parameters	Estimates (Std.err)
$\beta^s I_{t-1}$	-2.2025*
	(0.0647)
$\beta^s gapdur$	0.0541*
	(0.0250)
$\breve{a}^e_{m=1}$	0.2389
	(0.3841)
$\breve{a}^e_{m=2}$	-5.7032*
	(0.3896)
$\breve{a}^e_{m=3}$	0.2682
	(0.4037)
$\breve{a}^e_{m=4}$	-3.3910*
	(0.3805)
β^e_{educ}	0.3053*
	(0.0258)
β^e_{exp}	0.2037*
	(0.0104)
β^e_{cog}	0.0524
	(0.0451)
$\breve{a}^w_{m=1}$	2.1116
	(0.0237)
$\breve{a}^w_{m=2}$	1.9143
	(0.0340)
$\breve{a}^w_{m=3}$	1.6126*
	(0.0281)
$\breve{a}^w_{m=4}$	1.8476*
	(0.0225)
$\beta^w_{S1=1}$	0.1340*
	(0.0222)
$\beta^w_{S1=2}$	0.1890*
	(0.0298)
$\beta^w_{S1=3}$	0.5049*
	(0.0305)

* Significant at 1% level.

Table 13: MLE estimates from the dynamic model Con't 2
NLSY79 (30periods)

Parameters	Estimates (Std.err)
$\beta^w_{S1=4}$	0.7357*
	(0.0495)
$\beta^w_{S2=1}$	0.0767*
	(0.0313)
$\beta^w_{S2=2}$	0.0834*
	(0.0213)
$\beta^w_{S2=3}$	0.1205*
	(0.0136)
$\beta^w_{S2=4}$	0.1275*
	(0.0095)
β^w_{cog}	0.0420*
	(0.0083)
β^w_{e1}	0.0492*
	(0.0030)
β^w_{e2}	0.0538*
	(0.0041)
$\beta^w_{c1^2}$	-0.0011*
	(0.0001)
$\beta^w_{c2^2}$	-0.0014*
	(0.0002)
$\beta^w_{e1s1=12}$	0.0040
	(0.0020)
$\beta^w_{e1s1=13/14}$	0.0061*
	(0.0029)
$\beta^w_{e1s1=15/16}$	0.0201*
	(0.0026)
$\beta^w_{e1s1=16+}$	0.0114*
	(0.0048)
$\beta^w_{e1s2=12}$	-0.0327*
	(0.0102)
$\beta^w_{e1s2=13/14}$	-0.0058*
	(0.0030)

* Significant at 1% level.

Table 14: MLE estimates from the dynamic model Con't 3
NLSY79 (30periods)

Parameters	Estimates (Std.err)
$\beta^w_{e1s2=15/16}$	-0.0254*
	(0.0031)
$\beta^w_{e1s2=16+}$	-0.0247*
	(0.0027)
$\beta^w_{e2s2=12}$	0.0049
	(0.0039)
$\beta^w_{e2s2=13/14}$	0.0036
	(0.0033)
$\beta^w_{e2s2=15/16}$	0.0156*
	(0.0033)
$\beta^w_{e2s2=16+}$	0.0074*
	(0.0034)

* Significant at 1% level.

Table 15: MLE estimates from the dynamic model

	NLSY79 (30priods)	NLSY79 (17periods)	NLSY97 (17periods)
Parameters	Estimates (Std.err)	Estimates (Std.err)	Estimates (Std.err)
$\breve{a}^s_{m=1}$	6.1185*	5.1276*	4.8108*
	(0.3014)	(0.5564)	(0.4220)
$\breve{a}^s_{m=2}$	1.0375 *	1.7910*	1.5277*
	(0.2759)	(0.5899)	(0.4307)
$\breve{a}^s_{m=3}$	6.2248 *	3.9839*	8.2042 *
	(0.2969)	(0.5536)	(1.4001)
$\breve{a}^s_{m=4}$	2.1085 *	5.2616*	2.4760 *
	(0.2840)	(0.6996)	(0.5452)
β^s_{ds1}	-1.9122*	-1.7751*	0.2937*
	(0.0854)	(0.1653)	(0.1038)
β^s_{ds2}	-0.5961*	-0.2501	-1.6519*
	(0.1150)	(0.2698)	(0.1144)
β^s_{ds3}	-0.7583*	-0.3755	-1.3869*
	(0.1606)	(0.3950)	(0.1474)
β^s_{ds4}	-0.8403*	-0.4315	-3.1281*
	(0.2242)	(0.6040)	(0.1778)
$\breve{a}^s_{cog}$	0.6444 *	0.5860*	0.3774*
	(0.0515)	(0.0676)	(0.0582)
β^s_{Snisb}	-0.0505*	-0.0595*	-0.0166
	(0.0152)	(0.0164)	(0.0243)
$\beta^s_{Snuclear}$	0.2647*	0.2604*	0.2795*
	(0.0726)	(0.0750)	(0.0584)
β^s_{Ssouth}	-0.0060	-0.0105	-0.1088
	(0.0447)	(0.0612)	(0.0589)
β^s_{Srural}	0.0231	0.0174	-0.0034
	(0.0579)	(0.0627)	(0.0563)
β^s_{Shgcf}	0.0577*	0.0616*	0.0796*
	(0.0108)	(0.0122)	(0.0116)
β^s_{Shgcm}	0.0575*	0.0688*	0.0768*
	(0.0142)	(0.0157)	(0.0127)
$\beta^s_{Shomeincome}$	0.0036*	0.0034*	0.0056*
	(0.0009)	(0.0009)	(0.0015)

* Significant at 1% level.

Table 16: MLE estimates from the dynamic model Con't 1

	NLSY79 (30periods)	NLSY79 (17periods)	NLSY97 (17periods)
Parameters	Estimates (Std.err)	Estimates (Std.err)	Estimates (Std.err)
$\beta^s I_{t-1}$	-2.2025*	-1.9866*	-2.8391*
	(0.0647)	(0.0772)	(0.0773)
$\beta^s gapdur$	0.0541	0.0317	0.0920
	(0.1350)	(0.1060)	(0.1531)
$\breve{a}^e_{m=1}$	0.2389	-1.7467	3.5424*
	(0.3841)	(1.0630)	(0.4669)
$\breve{a}^e_{m=2}$	-5.7032*	-5.2139*	0.0838
	(0.3896)	(1.1360)	(0.5429)
$\breve{a}^e_{m=3}$	0.2682	-2.4784*	7.7656*
	(0.4037)	(1.0827)	(1.3864)
$\breve{a}^e_{m=4}$	-3.3910*	-1.0887	2.5239*
	(0.3805)	(1.2739)	(0.6436)
β^e_{educ}	0.3053*	0.3580*	0.1070*
	(0.0258)	(0.0668)	(0.0179)
β^e_{exp}	0.2037*	0.4380*	0.2869*
	(0.0104)	(0.0249)	(0.0244)
β^e_{cog}	0.0524	-0.0475	- 0.0053
	(0.0451)	(0.0647)	(0.0562)
$\breve{a}^w_{m=1}$	2.1116	2.2864*	2.7148*
	(0.0237)	(0.0334)	(0.2585)
$\breve{a}^w_{m=2}$	1.9143	1.9043*	1.9166*
	(0.0340)	(0.0339)	(0.0309)
$\breve{a}^w_{m=3}$	1.6126*	1.5134*	1.9494*
	(0.0281)	(0.0281)	(0.0374)
$\breve{a}^w_{m=4}$	1.8476*	1.8951*	1.8219*
	(0.0225)	(0.0306)	(0.0251)
$\beta^w_{S1=1}$	0.1340*	0.0982*	-0.0496*
	(0.0222)	(0.0270)	(0.0173)
$\beta^w_{S1=2}$	0.1890*	0.1858*	0.0162
	(0.0298)	(0.0382)	(0.0171)
$\beta^w_{S1=3}$	0.5049*	0.4209*	0.1140*
	(0.0305)	(0.0375)	(0.0214)

* Significant at 1% level.

Table 17: MLE estimates from the dynamic model Con't 2

	NLSY79 (30periods)	NLSY79 (WhiteMale17periods)	NLSY97 (17periods)
Parameters	Estimates (Std.err)	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{S1=4}$	0.7357*	0.6114*	0.3372*
	(0.0495)	(0.0640)	(0.0236)
$\beta^w_{S2=1}$	0.0767*	-0.0246	0.1558
	(0.0313)	(0.0379)	(0.0976)
$\beta^w_{S2=2}$	0.0834*	0.0519	0.0314*
	(0.0213)	(0.0304)	(0.0474)
$\beta^w_{S2=3}$	0.1205*	0.1164*	0.0632*
	(0.0136)	(0.0163)	(0.0177)
$\beta^w_{S2=4}$	0.1275*	0.1031*	0.0696*
	(0.0095)	(0.0210)	(0.0107)
β^w_{cog}	0.0420*	0.0394*	0.0345*
	(0.0083)	(0.0101)	(0.0119)
β^w_{e1}	0.0492*	0.0771*	0.0983*
	(0.0030)	(0.0057)	(0.0067)
β^w_{e2}	0.0538*	0.0808*	0.0997*
	(0.0041)	(0.0088)	(0.0136)
$\beta^w_{c1^2}$	-0.0011*	-0.0029*	-0.0043*
	(0.0001)	(0.0004)	(0.0005)
$\beta^w_{c2^2}$	-0.0014*	-0.0047*	-0.0028*
	(0.0002)	(0.0007)	(0.0013)
$\beta^w_{e1s1=12}$	0.0040	0.0011	0.0114
	(0.0020)	(0.0034)	(0.0068)
$\beta^w_{e1s1=13/14}$	0.0061*	0.0049	0.0065*
	(0.0029)	(0.0054)	(0.0068)
$\beta^w_{e1s1=15/16}$	0.0201*	0.0112*	0.0205*
	(0.0026)	(0.0054)	(0.0026)
$\beta^w_{e1s1=16+}$	0.0114*	0.0059*	0.0100*
	(0.0048)	(0.0121)	(0.0076)
$\beta^w_{e1s2=12}$	-0.0327*	-0.0086	-0.0185
	(0.0102)	(0.0128)	(0.0227)
$\beta^w_{e1s2=13/14}$	-0.0058*	-0.0208*	-0.0308*
	(0.0030)	(0.0063)	(0.0159)

* Significant at 1% level.

Table 18: MLE estimates from the dynamic model Con't 3

	NLSY79 (30periods)	NLSY79 (17periods)	NLSY97 (17periods)
Parameters	Estimates (Std.err)	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{e1s2=15/16}$	-0.0254*	-0.0303*	-0.0184*
	(0.0031)	(0.0063)	(0.0055)
$\beta^w_{e1s2=16+}$	-0.0247*	-0.0355*	-0.0196*
	(0.0027)	(0.0095)	(0.0042)
$\beta^w_{e2s2=12}$	0.0049	0.0270*	-0.0531*
	(0.0039)	(0.0078)	(0.0178)
$\beta^w_{e2s2=13/14}$	0.0036	0.0225*	-0.0103
	(0.0033)	(0.0071)	(0.0123)
$\beta^w_{e2s2=15/16}$	0.0156*	0.0134	-0.0019 *
	(0.0033)	(0.0075)	(0.0110)
$\beta^w_{e2s2=16+}$	0.0074*	0.0198	-0.0265*
	(0.0034)	(0.0107)	(0.0104)

* Significant at 1% level.

Table 19: MLE estimates from the dynamic model

	NLSY79 (Black)	NLSY97 (Black)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\breve{a}^s_{m=1}$	5.5326*	6.8605*
	(0.5081)	(0.9328)
$\breve{a}^s_{m=2}$	3.8340*	3.0524*
	(0.7699)	(0.7685)
$\breve{a}^s_{m=3}$	5.4848*	5.5006*
	(0.4673)	(0.6112)
$\breve{a}^s_{m=4}$	3.8892*	2.9922*
	(0.4682)	(0.6065)
β^s_{ds1}	-1.4791*	-1.1545*
	(0.0388)	(0.1477)
β^s_{ds2}	-0.9658*	-0.2459
	(0.0756)	(0.1851)
β^s_{ds3}	-1.0504*	-0.1173
	(0.2524)	(0.2583)
β^s_{ds4}	-1.1734*	-0.5885
	(0.1532)	(0.3553)
$\breve{a}^s_{cog}$	0.4509*	0.4581*
	(0.0714)	(0.0848)
β^s_{nisb}	-0.0064	-0.0652*
	(0.0152)	(0.0335)
$\beta^s_{nuclear}$	0.0452	0.2377*
	(0.0858)	(0.0937)
β^s_{south}	-0.1000	-0.0307
	(0.0828)	(0.1041)
β^s_{rural}	-0.1648	0.0313
	(0.1066)	(0.0563)
β^s_{hgcf}	0.0090	0.0631*
	(0.0154)	(0.0225)
β^s_{hgcm}	0.0812*	0.0845*
	(0.0201)	(0.0250)
$\beta^s_{homeincome}$	0.2102*	0.0908
	(0.0623)	(0.0590)

* Significant at 1% level.

Table 20: MLE estimates from the dynamic model Con't 1

	NLSY79 (Black)	NLSY97 (Black)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^s I_{t-1}$	-3.1655*	-3.1057*
	(0.1154)	(0.1227)
$\beta^s gapdur$	0.1192	0.1664
	(0.1643)	(0.1511)
$\breve{a}^e_{m=1}$	-5.3625*	2.7764*
	(0.5500)	(0.9207)
$\breve{a}^e_{m=2}$	-5.7032*	-1.5249*
	(0.8272)	(0.7905)
$\breve{a}^e_{m=3}$	-0.6244	2.1002*
	(0.5768)	(1.3864)
$\breve{a}^e_{m=4}$	-2.8546*	-0.8481
	(0.5777)	(0.7344)
β^e_{educ}	0.2958*	0.1117*
	(0.0391)	(0.0416)
β^e_{exp}	0.3628*	0.2881*
	(0.0343)	(0.0361)
β^e_{cog}	0.1472*	0.2736*
	(0.0668)	(0.0841)
$\breve{a}^w_{m=1}$	1.6792*	1.7844*
	(0.0237)	(0.1047)
$\breve{a}^w_{m=2}$	1.9143	3.0823*
	(0.0340)	(0.1458)
$\breve{a}^w_{m=3}$	1.2679*	1.2069*
	(0.0685)	(0.0822)
$\breve{a}^w_{m=4}$	0.0574*	1.2633*
	(0.0054)	(0.0794)
$\beta^w_{S1=1}$	0.0406*	0.0605*
	(0.0007)	(0.0066)
$\beta^w_{S1=2}$	0.6781*	0.5153*
	(0.0083)	(0.0162)
$\beta^w_{S1=3}$	0.1134*	0.1131*
	(0.0457)	(0.0314)

* Significant at 1% level.

Table 21: MLE estimates from the dynamic model Con't 2

	NLSY79 (Black)	NLSY97 (Black)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{S1=4}$	0.1235*	0.2341*
	(0.0125)	(0.0112)
$\beta^w_{S2=1}$	0.0131*	0.0567
	(0.0012)	(0.0231)
$\beta^w_{S2=2}$	0.0122*	0.0256*
	(0.0071)	(0.0051)
$\beta^w_{S2=3}$	0.0431*	0.0561*
	(0.0046)	(0.0111)
$\beta^w_{S2=4}$	0.0145*	0.0231*
	(0.0013)	(0.0081)
β^w_{e1}	0.0163*	0.0528*
	(0.0011)	(0.0223)
β^w_{e2}	0.0141*	0.0461*
	(0.0023)	(0.0010)
$\beta^w_{c1^2}$	-0.0003*	-0.0027*
	(0.003)	(0.0010)
$\beta^w_{c2^2}$	-0.0004	-0.0020
	(0.0001)	(0.003)
$\beta^w_{e1s1=12}$	0.0012*	0.0013*
	(0.0011)	(0.0016)
$\beta^w_{e1s1=13/14}$	0.0013*	0.0025*
	(0.0007)	(0.0012)
$\beta^w_{e1s1=15/16}$	0.0016*	0.0032*
	(0.012)	(0.0032)
$\beta^w_{e1s1=16+}$	0.0017*	0.0021*
	(0.0530)	(0.0352)
$\beta^w_{e1s2=12}$	0.0230	-0.0185*
	(0.008)	(0.0012)
$\beta^w_{e1s2=13/14}$	-0.001*	-0.0031*
	(0.0003)	(0.0011)

* Significant at 1% level.

Table 22: MLE estimates from the dynamic model Con't 3

	NLSY79 (Black)	NLSY97 (Black)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{e1s2=15/16}$	0.1140*	-0.1123
	(0.0135)	(0.0043)
$\beta^w_{e1s2=16+}$	-0.0145	-0.0112*
	(0.0034)	(0.0032)
$\beta^w_{e2s2=12}$	-0.0003*	-0.0012*
	(0.0001)	(0.0013)
$\beta^w_{e2s2=13/14}$	0.0012*	0.0103
	(0.0003)	(0.0131)
$\beta^w_{e2s2=15/16}$	0.1124*	0.0019*
	(0.0121)	(0.0312)
$\beta^w_{e2s2=16+}$	0.0781	0.0431*
	(0.0082)	(0.0012)

* Significant at 1% level.

Table 23: MLE estimates from the dynamic model

	NLSY79 (Hisp.)	NLSY97 (Hisp.)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\breve{a}^s_{m=1}$	8.0503*	6.1598*
	(0.7847)	(0.6283)
$\breve{a}^s_{m=2}$	6.3014*	4.4336*
	(0.7668)	(0.7259)
$\breve{a}^s_{m=3}$	7.1096*	6.4416*
	(0.7562)	(0.5975)
$\breve{a}^s_{m=4}$	5.2720*	3.6599*
	(0.6896)	(0.5178)
β^s_{ds1}	-0.2290	-1.0831*
	(0.1576)	(0.1479)
β^s_{ds2}	-0.9255*	-0.4857*
	(0.2221)	(0.2011)
β^s_{ds3}	-1.0648*	-0.4991
	(0.2980)	(0.2828)
β^s_{ds4}	-1.4581*	-0.7015*
	(0.4612)	(0.3805)
$\breve{a}^s_{cog}$	0.8404*	0.5144*
	(0.0185)	(0.1192)
β^s_{Snisb}	-0.0325	-0.0261
	(0.0185)	(0.0404)
$\beta^s_{Snuclear}$	0.2647*	0.2617*
	(0.0726)	(0.1049)
β^s_{Ssouth}	-0.1304	-0.0206
	(0.1097)	(0.0969)
β^s_{Srural}	-0.2501	-0.0267
	(0.1394)	(0.1748)
β^s_{Shgcf}	0.0052	0.0316*
	(0.0140)	(0.0152)
β^s_{Shgcm}	-0.0170	0.0152
	(0.0149)	(0.0168)
$\beta^s_{Shomeincome}$	0.1059	0.0476
	(1.5018)	(0.0566)

* Significant at 1% level.

Table 24: MLE estimates from the dynamic model Con't 1

	NLSY79 (Hisp.)	NLSY97 (Hisp.)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^s I_{t-1}$	-2.8880*	-2.6980*
	(0.1409)	(0.1325)
$\beta^s gapdur$	0.2563	0.1741
	(0.2123)	(0.1900)
$\breve{a}^e_{m=1}$	1.0377	1.6636*
	(0.9435)	(0.6224)
$\breve{a}^e_{m=2}$	-1.3450	-0.6069
	(0.9426)	(0.6618)
$\breve{a}^e_{m=3}$	0.4167	2.2855*
	(0.9753)	(0.5641)
$\breve{a}^e_{m=4}$	-2.2097*	-0.8903*
	(0.9401)	(0.5641)
β^e_{educ}	0.3029*	0.0634*
	(0.0490)	(0.0490)
β^e_{exp}	0.2750*	0.2791*
	(0.0389)	(0.0378)
$\breve{a}^w_{m=1}$	1.7478	1.7158*
	(0.0823)	(0.1071)
$\breve{a}^w_{m=2}$	2.2566	2.4769*
	(0.1000)	(0.1411)
$\breve{a}^w_{m=3}$	1.3448*	1.2893*
	(0.0803)	(0.0955)
$\breve{a}^w_{m=4}$	1.5033*	1.3450*
	(0.0749)	(0.0849)
$\beta^w_{S1=12}$	-0.0428*	-0.0526*
	(0.0067)	(0.0073)
$\beta^w_{S1=13/14}$	0.2341*	0.2333*
	(0.0143)	(0.0111)
$\beta^w_{S1=15/16}$	0.3211*	0.2414*
	(0.0221)	(0.0124)

* Significant at 1% level.

Table 25: MLE estimates from the dynamic model Con't 2

	NLSY79 (Hisp.)	NLSY97 (Hisp.)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{S1=16+}$	0.6435*	0.7675*
	(0.0231)	(0.0245)
$\beta^w_{S2=12}$	0.0121	-0.2341*
	(0.1220)	(0.0135)
$\beta^w_{S2=13/14}$	0.0134*	0.0144*
	(0.0011)	(0.0021)
$\beta^w_{S2=15/16}$	0.1213*	0.2341*
	(0.0141)	(0.0011)
$\beta^w_{S2=16+}$	0.1723*	0.2321*
	(0.0012)	(0.0011)
β^w_{e1}	0.0236	0.0684*
	(0.0145)	(0.0197)
β^w_{e2}	0.0658*	0.0943*
	(0.0197)	(0.0193)
$\beta^w_{c1^2}$	-0.0023*	-0.0042*
	(0.0006)	(0.0008)
$\beta^w_{c2^2}$	-0.0028	--0.0030
	(0.0021)	(0.0020)
$\beta^w_{e1s1=12}$	-0.0041*	-0.0031*
	(0.0011)	(0.0001)
$\beta^w_{e1s1=13/14}$	0.0066*	0.0061*
	(0.0021)	(0.0012)
$\beta^w_{e1s1=15/16}$	0.0412*	0.0556*
	(0.0131)	(0.0012)
$\beta^w_{e1s1=16+}$	0.1321*	0.1756*
	(0.0124)	(0.0323)
$\beta^w_{e1s2=12}$	-0.0241*	-0.0191*
	(0.0012)	(0.0011)
$\beta^w_{e1s2=13/14}$	0.0453	0.0245*
	(0.0215)	(0.0011)

* Significant at 1% level.

Table 26: MLE estimates from the dynamic model Con't 3

	NLSY79 (Hisp.)	NLSY97 (Hisp.)
Parameters	Estimates (Std.err)	Estimates (Std.err)
$\beta^w_{e1s2=15/16}$	0.0121*	0.0133*
	(0.0011)	(0.0014)
$\beta^w_{e1s2=16+}$	0.0247	0.0144*
	(0.0131)	(0.0011)
$\beta^w_{e2s2=12}$	0.0036	-0.0043*
	(0.0011)	(0.0003)
$\beta^w_{e2s2=13/14}$	0.0038	0.0241*
	(0.0027)	(0.0034)
$\beta^w_{e2s2=15/16}$	0.1163*	0.0978*
	(0.0012)	(0.0135)
$\beta^w_{e2s2=16+}$	0.1356	0.1244
	(0.5146)	(0.0843)

* Significant at 1% level.

Table 27: Wage Simulation (4-type model)

Age	HS (4-type)	Col (4-type)	Col_gap (4-type)
16	0	0	0
17	0	0	0
18	8.856	0	8.856
19	9.562	0	0
20	10.037	0	0
21	10.565	0	0
22	11.021	13.219	0
23	11.604	14.191	13.774
24	12.161	15.11	14.56
25	12.455	15.727	15
26	12.822	16.463	15.552
27	13.259	17.38	16.252
28	13.56	18.009	16.729
29	14.201	19.168	17.722
30	14.63	20.208	18.422
31	14.884	20.796	18.886
32	15.34	21.878	19.671
33	15.246	22.073	19.77
34	15.818	23.367	20.714
35	15.938	23.908	21.08
36	16.215	24.815	21.688
37	16.067	25.017	21.695
38	16.344	25.756	22.296
39	16.483	26.149	22.545
40	16.516	26.508	22.668
41	16.778	27.295	22.895
42	16.881	28.054	23.352
43	17.1	28.566	23.698
44	17.33	28.786	23.749
45	17.959	29.171	24.446

Table 28: Wage Simulation (no-type model)

Age	High School	University	University (gap)
16	0	0	0
17	0	0	0
18	8.814	0	8.814
19	9.474	0	0
20	9.875	0	0
21	10.354	0	0
22	10.772	12.568	0
23	11.343	13.429	12.717
24	11.919	14.369	13.421
25	12.167	14.857	13.69
26	12.476	15.484	14.078
27	13.029	16.448	14.754
28	13.37	17.185	15.282
29	13.85	18.088	15.961
30	14.397	19.075	16.658
31	14.661	19.721	17.103
32	15.171	20.749	17.852
33	15.17	21.123	18.057
34	15.666	22.091	18.797
35	15.893	22.799	19.239
36	16.148	23.6	19.776
37	16.082	23.822	19.871
38	16.498	24.879	20.66
39	16.574	25.344	20.907
40	16.689	25.435	20.931
41	16.478	26.006	21.271
42	16.634	26.61	21.659
43	16.803	27.117	22.08
44	17.551	27.25	22.087
45	18.274	27.592	22.903

Table 29: Present Values (4-type)

Age	HS (4-type)	Col (4-type)	Col_gap (4-type)
Total	$789,204	$1,061,028	$929,840
npv30years	$197,434.92	$173,300.16	$160,787.29
npv40years	$314,603.50	$349,864.65	$316,258.71
npv45years	$356,310.91	$421,787.66	$376,077.36
ln(npv30years)	5.30	5.24	5.21
ln(npv40years)	5.50	5.54	5.50
ln(npv45years)	5.55	5.63	5.58

Table 30: Present values (no-type)

Age	HS (4-type)	Col (4-type)	Col_gap (4-type)
Total	$785,664	$1,010,282	$855,136
npv30years	$193,834.98	$164,082.18	$147,804.47
npv40years	$310,779.46	$332,647.99	$289,999.90
npv45years	$356,310.91	$401,143.05	$345,699.66
ln(npv30years)	5.29	5.242	5.17
ln(npv40years)	5.49	5.52	5.46
ln(npv45years)	5.55	5.60	5.54

Table 31: Policy implication on hourly wages (in 1997 U.S. dollar)

Age	Obs.Wage	Wage (Policy)
16	7.54	7.35
17	7.39	7.83
18	8.32	7.75
19	8.68	8.58
20	9.14	8.77
21	9.21	9.32
22	9.88	10.23
23	10.86	10.65
24	12.07	11.23
25	13.31	11.76
26	14.98	12.18
27	14.34	13.14
28	14.87	13.91
29	15.38	14.66
30	15.84	15.94
31	15.45	16.47
32	16.98	17.45
33	17.79	17.59
34	18.46	18.76
35	19.22	19.52
36	19.96	20.21
37	20.54	21
38	21.29	21.74
39	21.72	22.25
40	22.31	22.63
41	23.13	23.67
42	22.73	24.28
43	23.15	24.92
44	23.86	25.59
45	23.86	25.42

Table 32: Wage Simulation (White_17periods)

Age	79Col	79Col_gap	97Col	97Col_gap
19	0	7.958	0	8.856
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	10.699	0	12.745	0
24	11.442	11.285	13.499	13.647
25	11.837	11.578	14.702	14.912
26	12.37	12.012	15.633	16.003
27	13.1	12.638	16.686	17.022
28	13.779	13.204	17.458	17.870
29	14.569	13.888	18.603	19.253
30	15.357	14.548	18.391	19.548
31	15.908	15.004	19.868	21.629
32	16.767	15.741	22.557	14.185

Table 33: Wage Simulation (Black_17periods)

Age	79Col	79Col_gap	97Col	97Col_gap
19	0	8.252	0	8.381
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	10.639	0	9.864	0
24	11.771	11.879	11.309	11.717
25	12.126	11.831	11.344	11.339
26	12.246	12.144	12.397	12.326
27	12.615	12.573	12.660	12.346
28	13.279	13.250	13.652	13.632
29	13.577	13.577	16.262	15.632
30	13.968	13.968	14.918	14.801
31	14.588	14.588	15.570	15.471
32	14.990	14.990	15.912	15.995

Table 34: Wage Simulation (Hisp._17periods)

Age	79Col	79Col_gap	97Col	97Col_gap
19	0	8.277	0	8.429
20	0	0	0	0
21	0	0	0	0
22	0	0	0	0
23	10.475	0	10.639	0
24	10.800	10.844	11.740	11.785
25	11.649	11.531	12.276	12.330
26	12.736	12.590	13.457	13.593
27	13.300	13.193	14.024	14.216
28	14.338	14.195	15.125	15.576
29	15.447	15.328	16.054	16.272
30	15.822	15.729	17.684	17.787
31	17.183	17.057	17.471	17.606
32	16.941	16.736	17.171	17.556

Figure 13: NLSY79 (White) Proportion of Pre-gap School

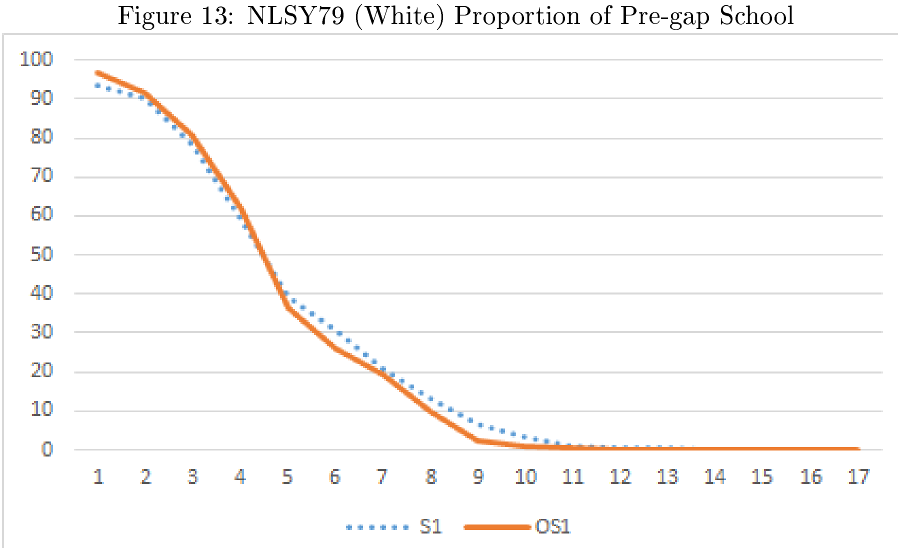


Figure 14: NLSY79 (White) Proportion of Post-gap School

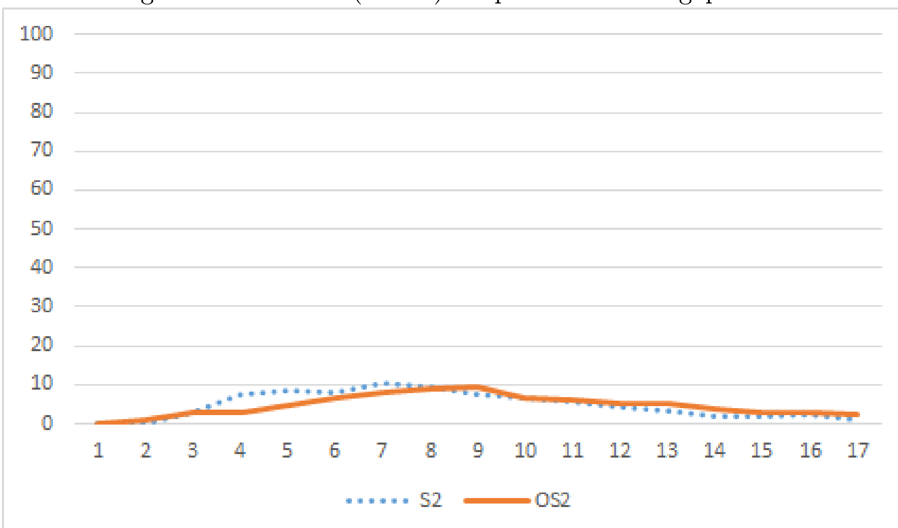


Figure 15: NLSY79 (White) Proportion of Pre-gap Work

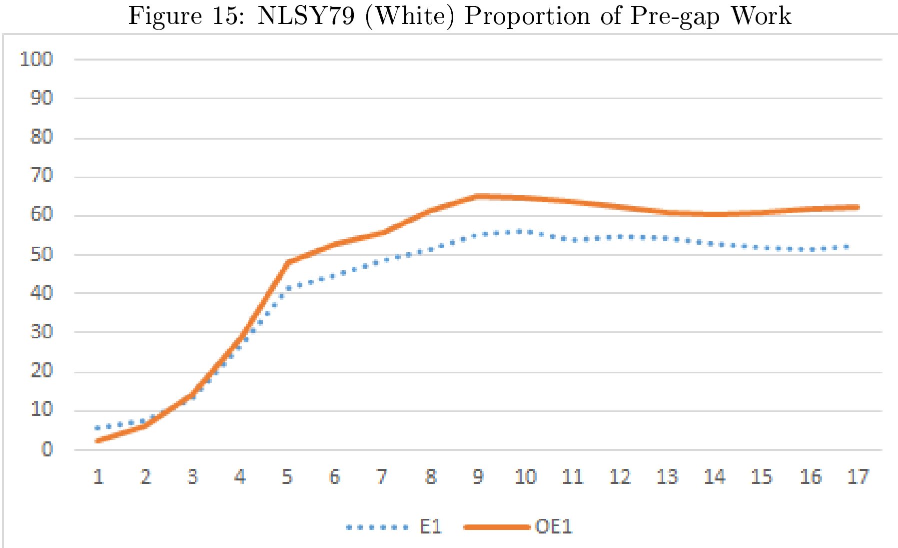


Figure 16: NLSY79 (White) Proportion of Post-gap Work

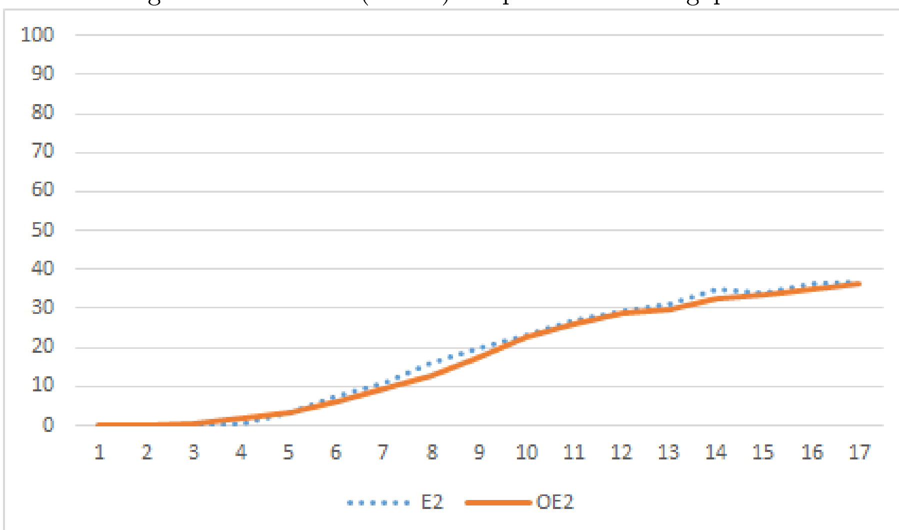


Figure 17: NLSY79 (White) Proportion of NEET

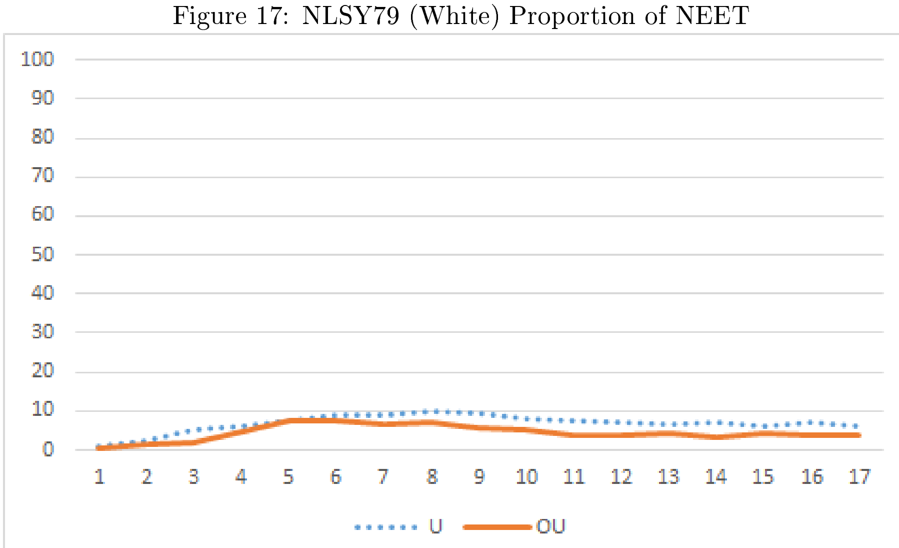


Figure 18: NLSY79 (White) Proportion of Schooling Interruption

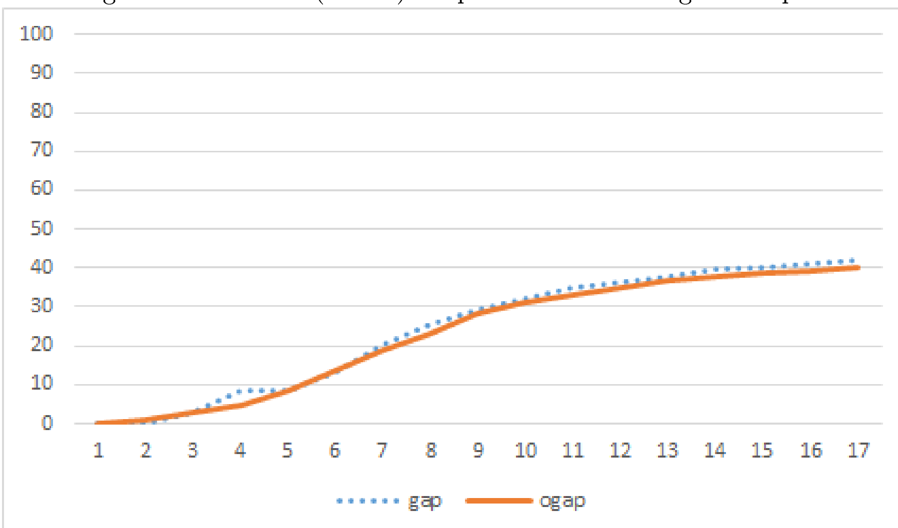


Figure 19: NLSY79 (White) Hourly Wage (1997 US dollars)

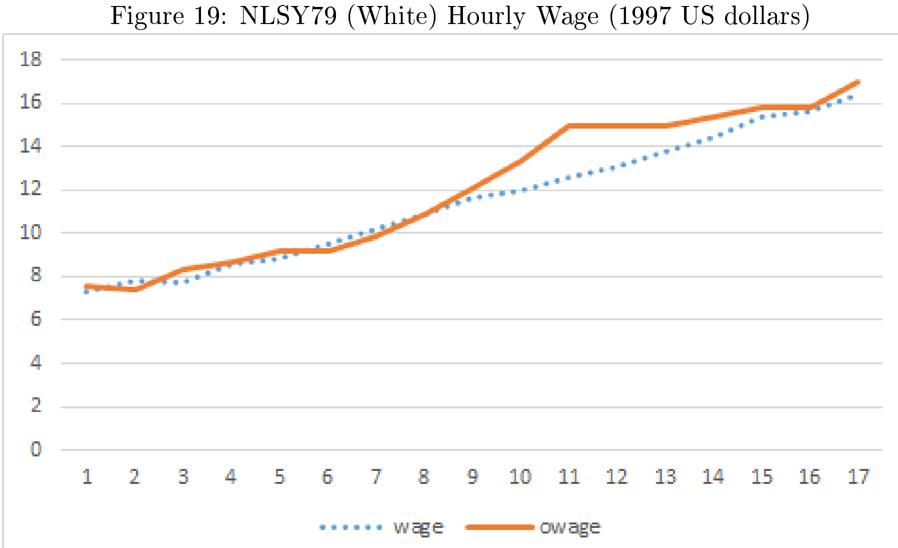


Figure 20: NLSY97 (White) Proportion of Pre-gap School

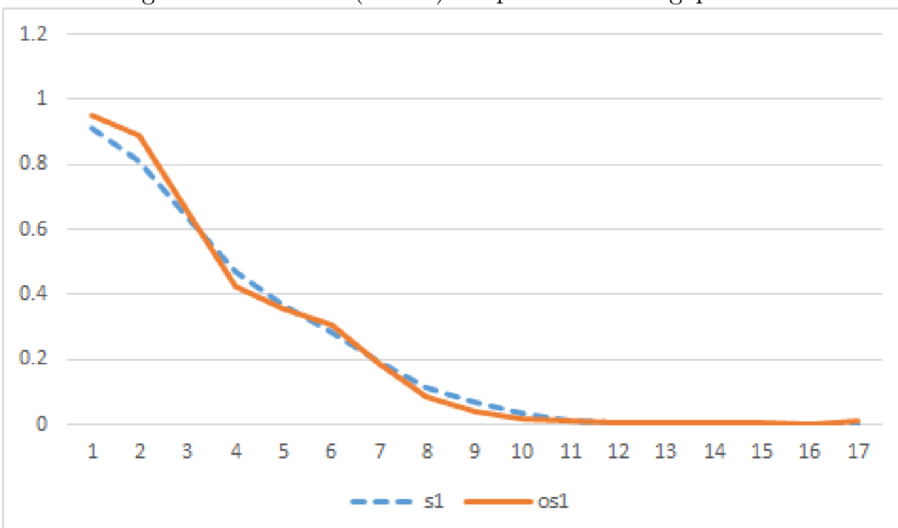


Figure 21: NLSY97 (White) Proportion of Post-gap School

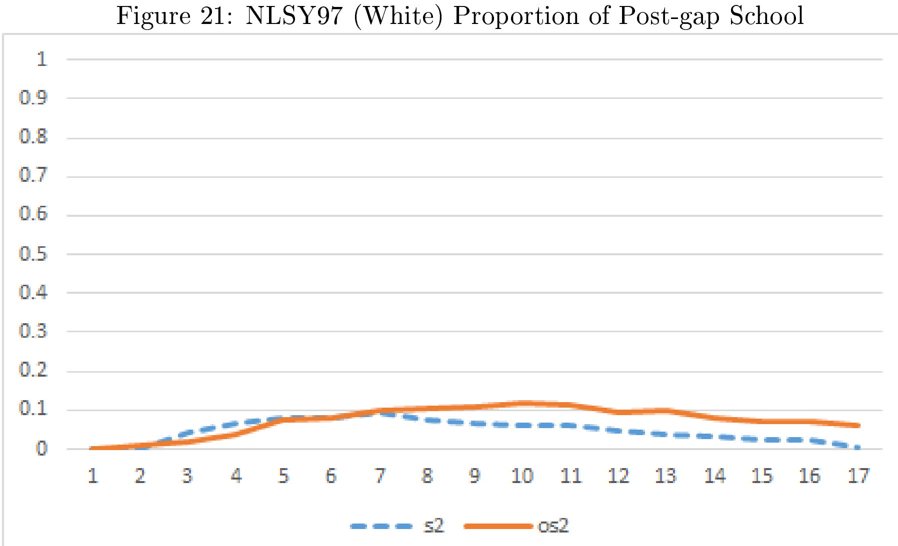


Figure 22: NLSY97 (White) Proportion of Pre-gap Work

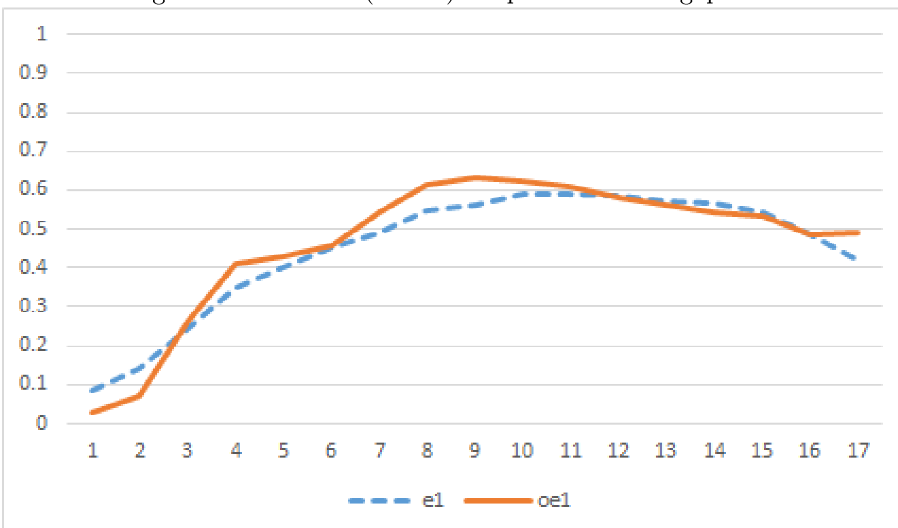


Figure 23: NLSY97 (White) Proportion of Post-gap Work

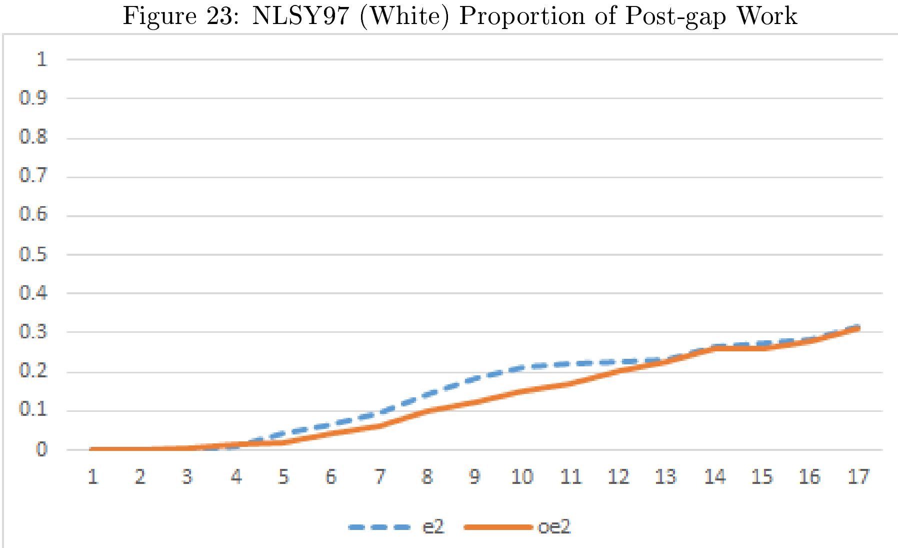


Figure 24: NLSY97 (White) Proportion of NEET

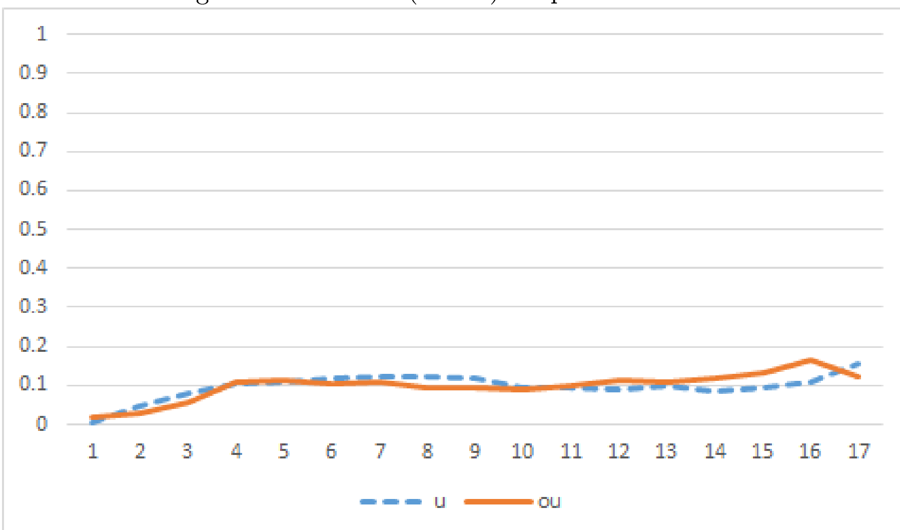


Figure 25: NLSY97 (White) Proportion of Schooling Interruption

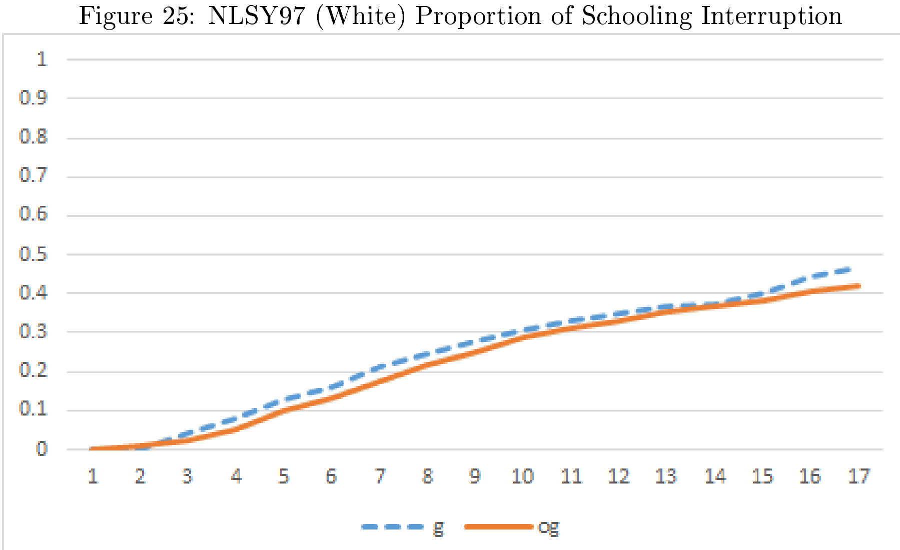


Figure 26: NLSY97 (White) Hourly Wage (1997 US dollars)

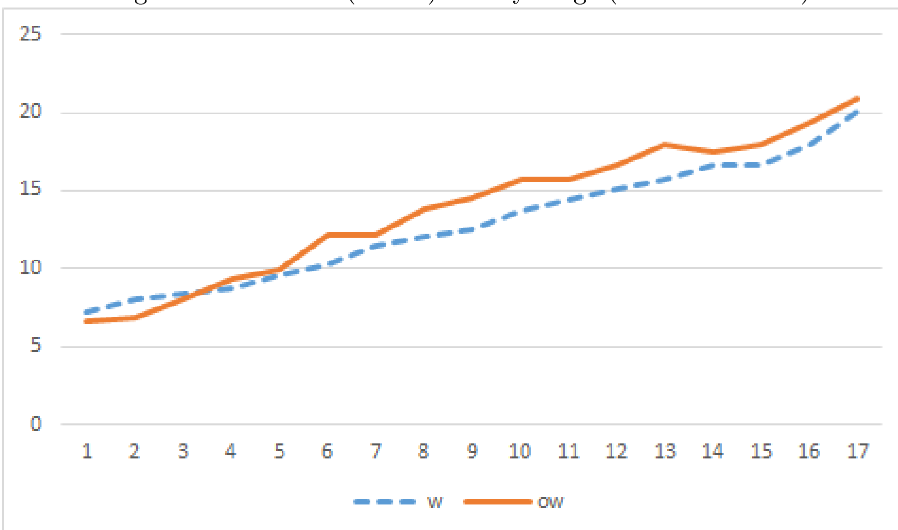


Figure 27: NLSY79 (Black) Proportion of Pre-gap School

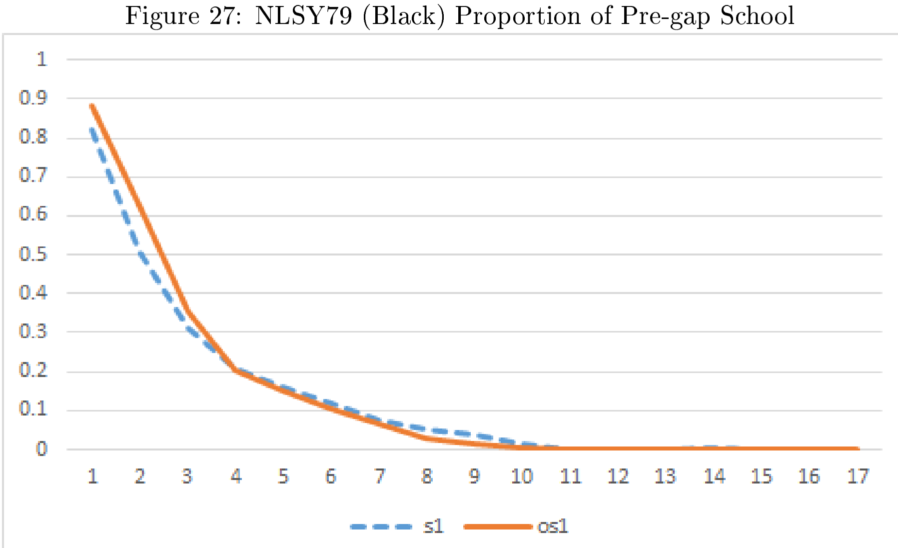


Figure 28: NLSY79 (Black) Proportion of Post-gap School

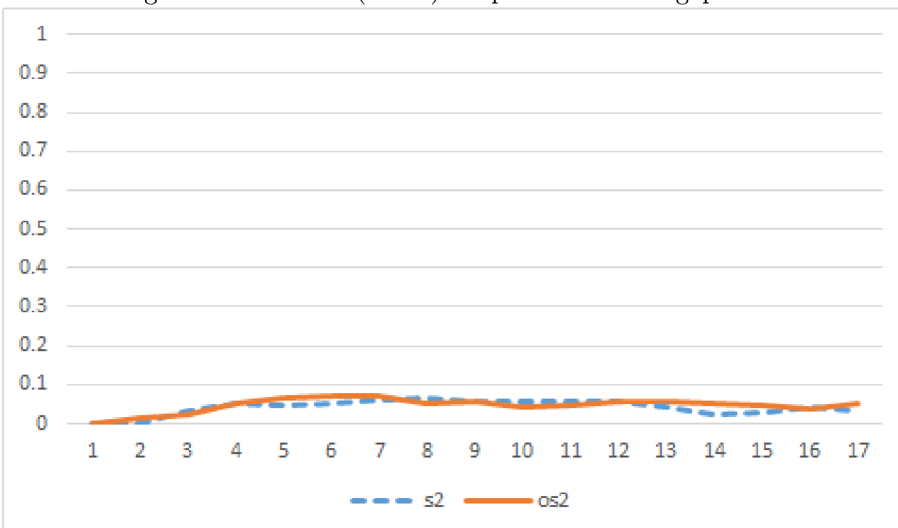


Figure 29: NLSY79 (Black) Proportion of Pre-gap Work

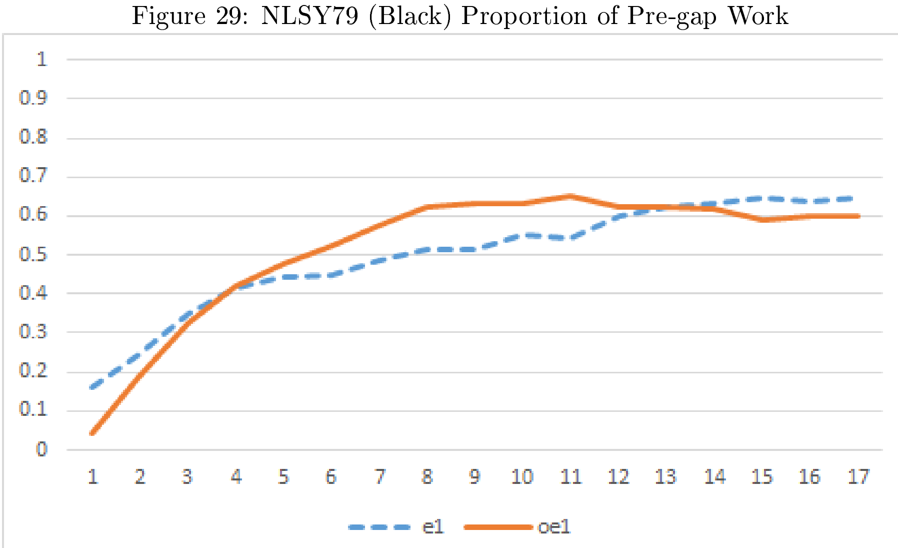


Figure 30: NLSY79 (Black) Proportion of Post-gap Work

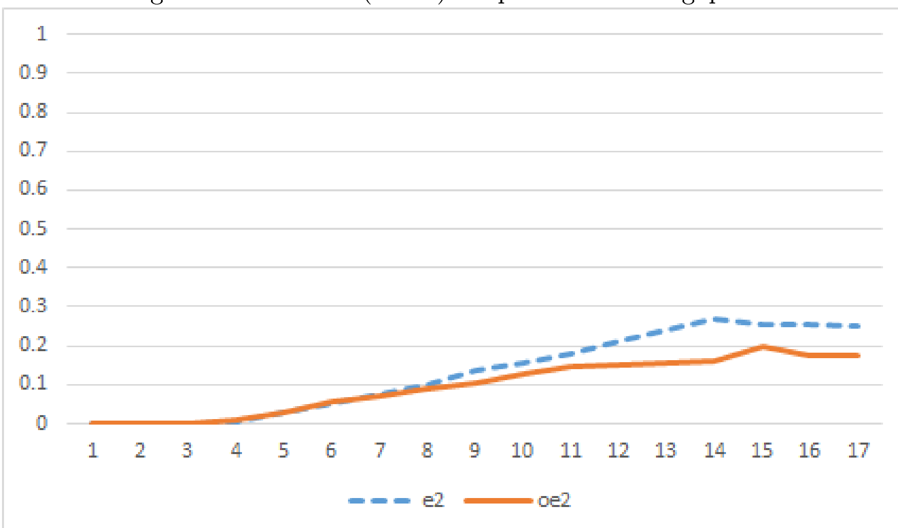


Figure 31: NLSY79 (Black) Proportion of NEET

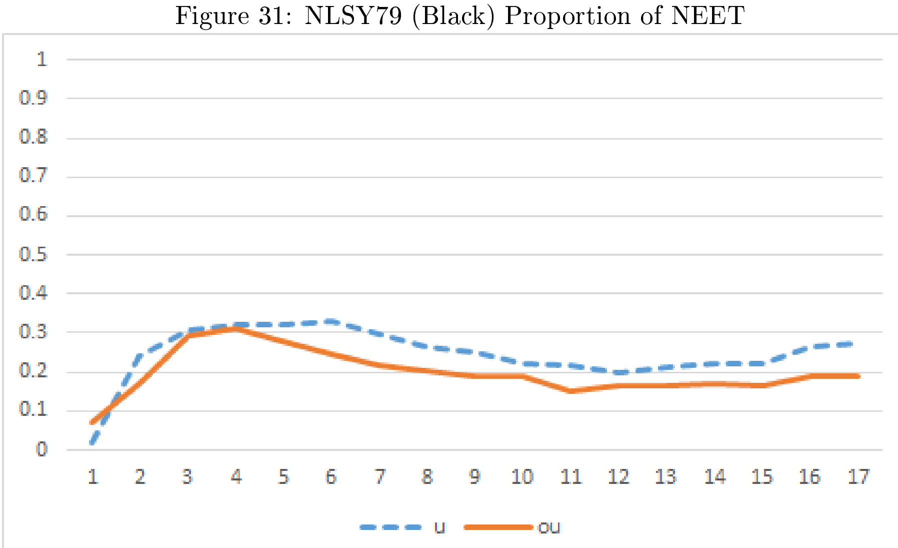


Figure 32: NLSY79 (Black) Proportion of Schooling Interruption

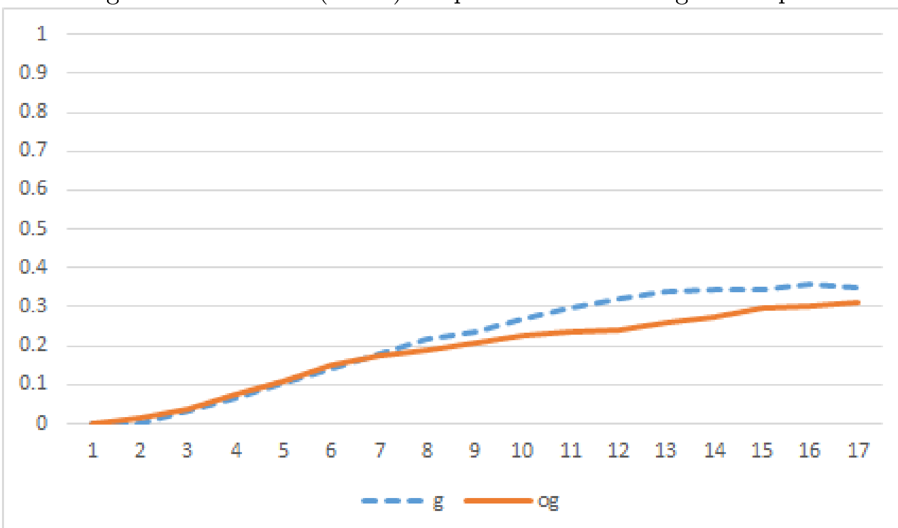


Figure 33: NLSY79 (Black) Hourly Wage (1997 US Dollar)

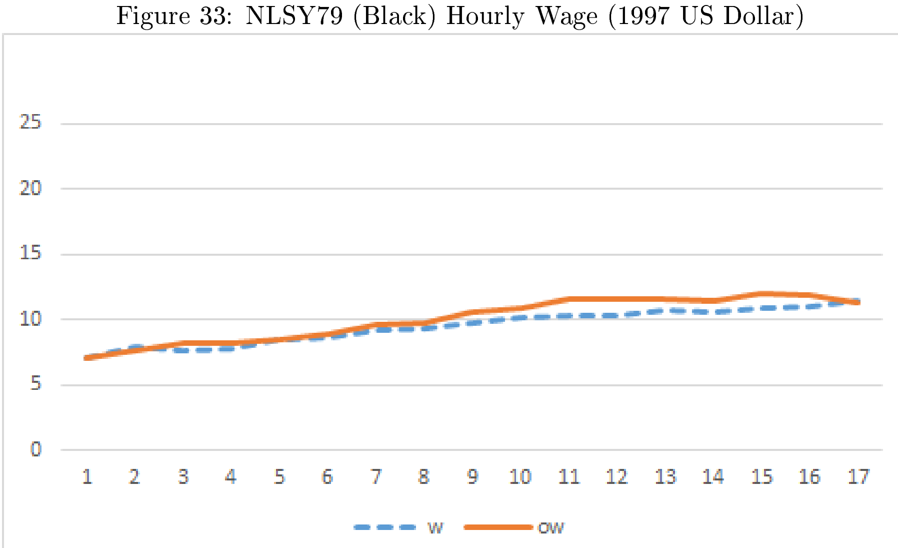


Figure 34: NLSY97 (Black) Proportion of Pre-gap School

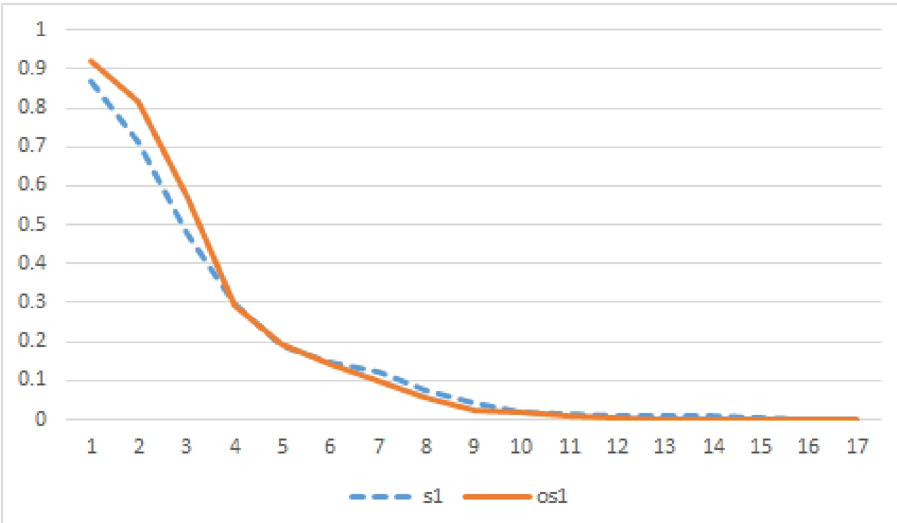


Figure 35: NLSY97 (Black) Proportion of Post-gap School

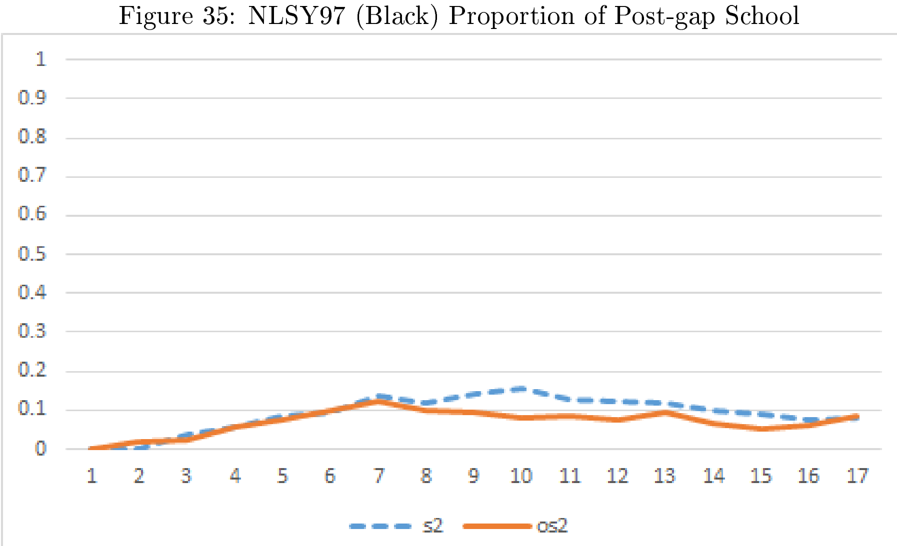


Figure 36: NLSY97 (Black) Proportion of Pre-gap Working

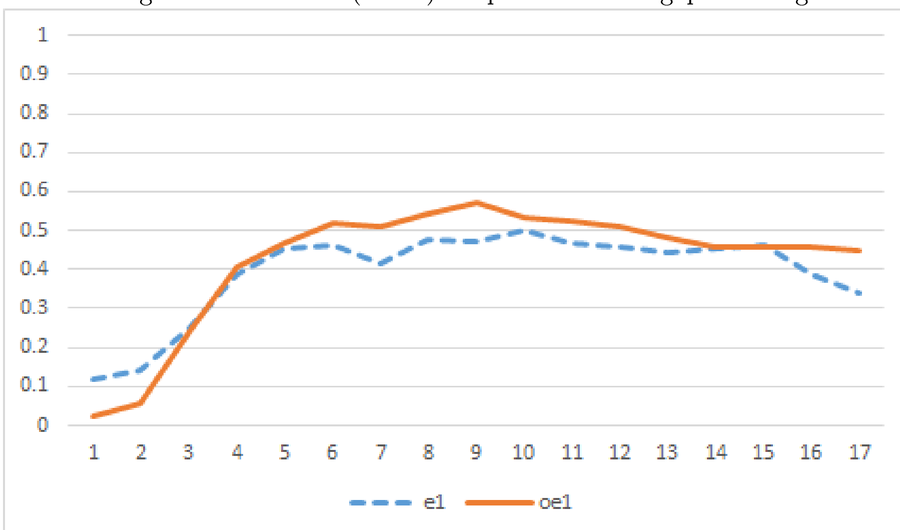


Figure 37: NLSY97 (Black) Proportion of Post-gap Working

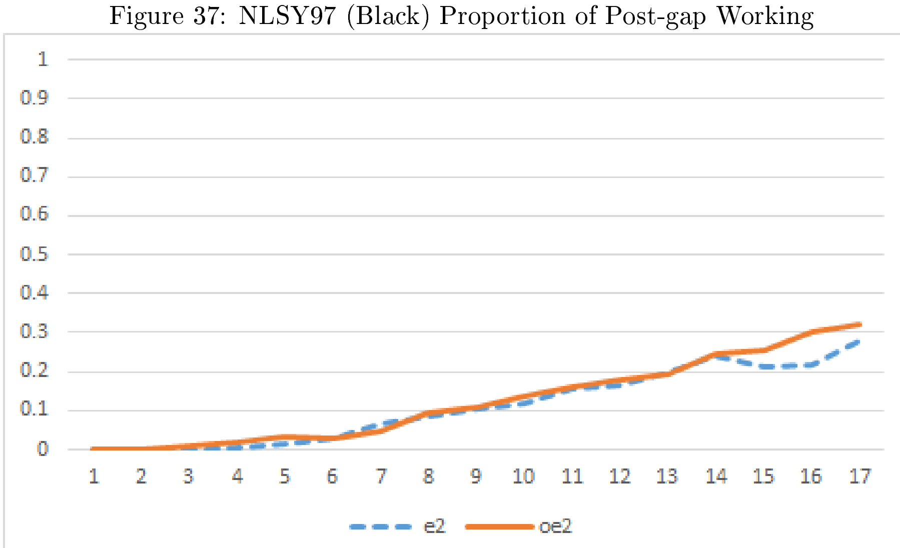


Figure 38: NLSY97 (Black) Proportion of NEET

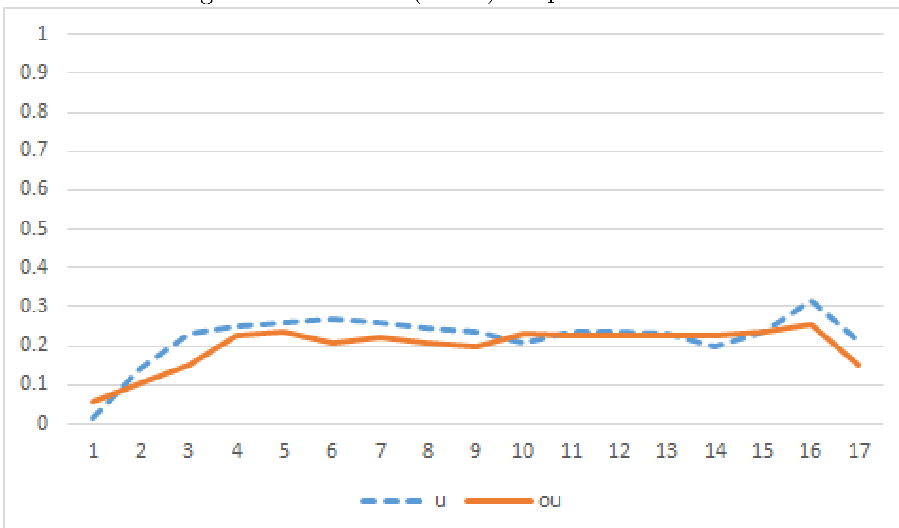


Figure 39: NLSY97 (Black) Proportion of Schooling Interruption

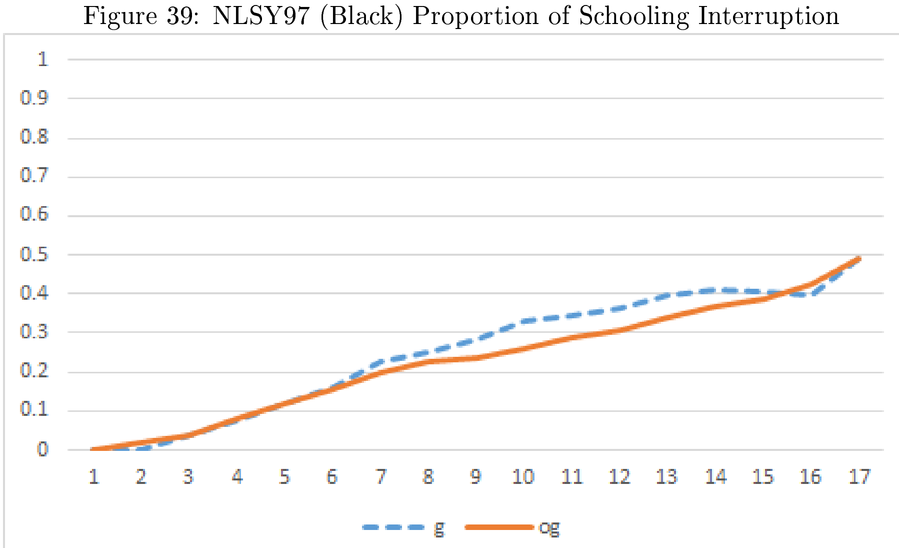


Figure 40: NLSY97 (Black) Hourly Wage (1997 US Dollar)

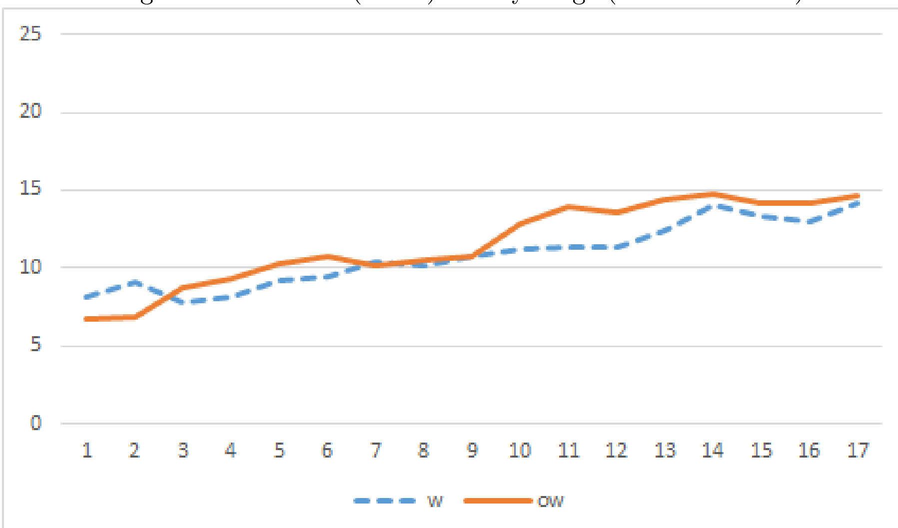


Figure 41: NLSY79 (Hisp.) Proportion of Pre-gap School

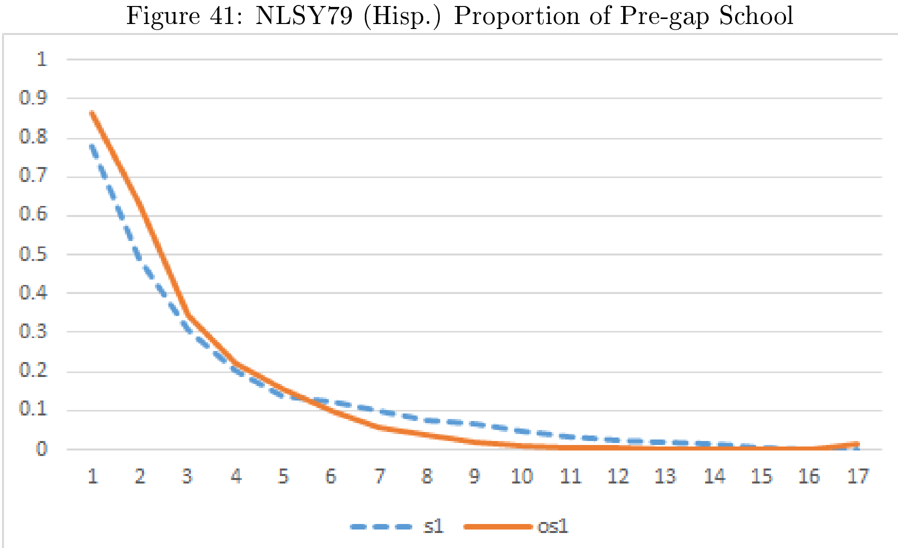


Figure 42: NLSY79 (Hisp.) Proportion of Post-gap School

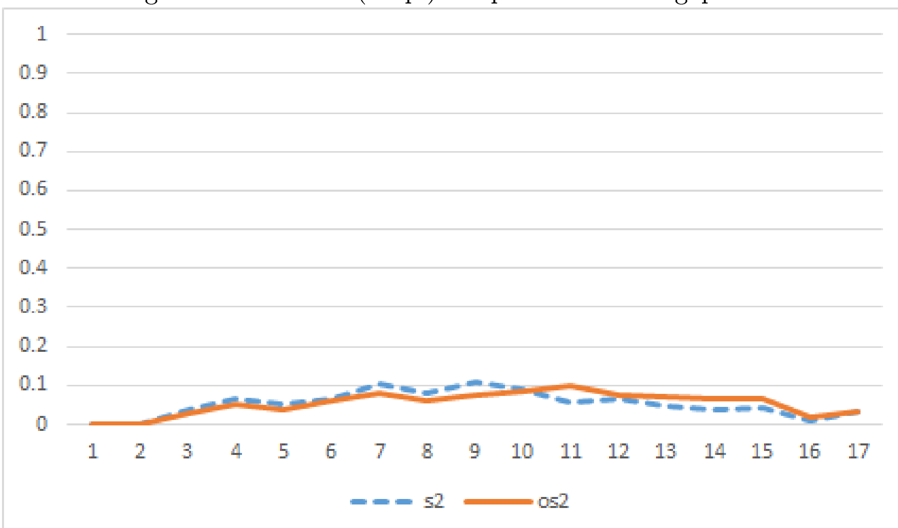


Figure 43: NLSY79 (Hisp.) Proportion of Pre-gap Working

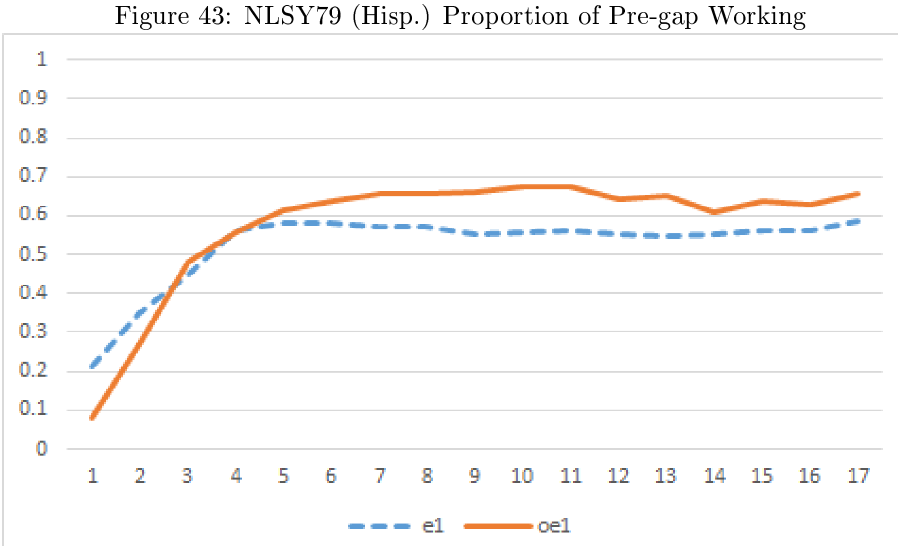


Figure 44: NLSY79 (Hisp.) Proportion of Post-gap Working

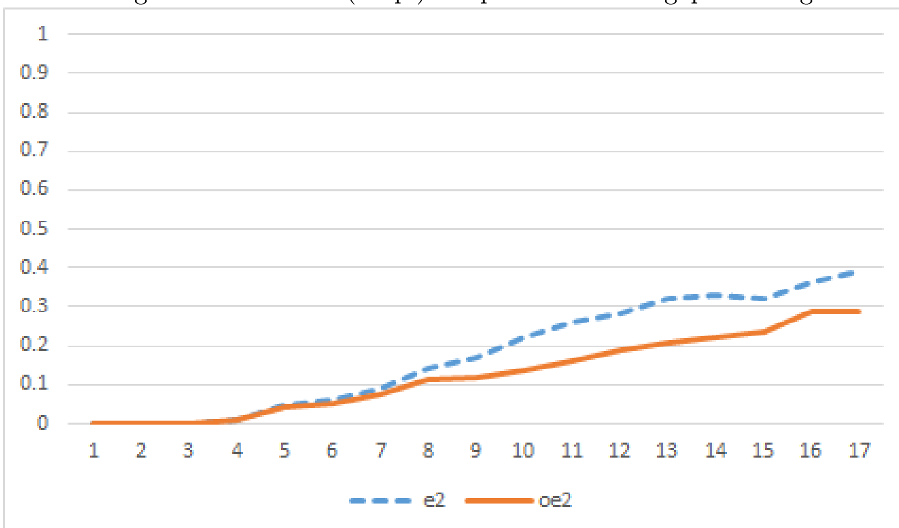


Figure 45: NLSY79 (Hisp.) Proportion of NEET

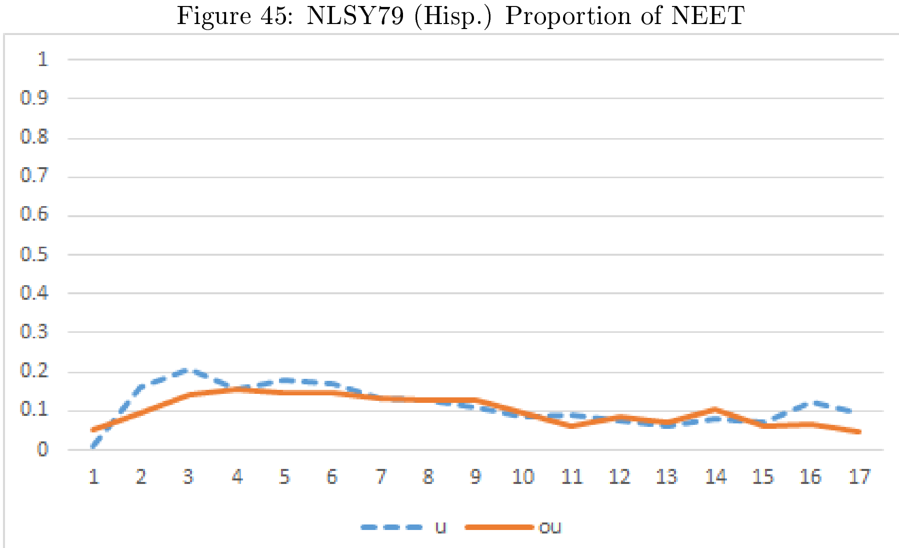


Figure 46: NLSY79 (Hisp.) Rate of Schooling Interruption

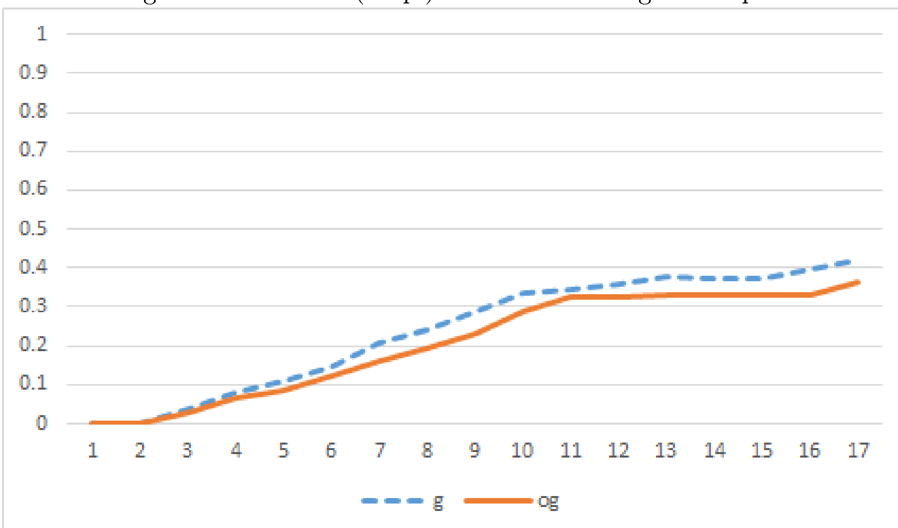


Figure 47: NLSY79 (Hisp.) Hourly Wage (1997 US Dollar)

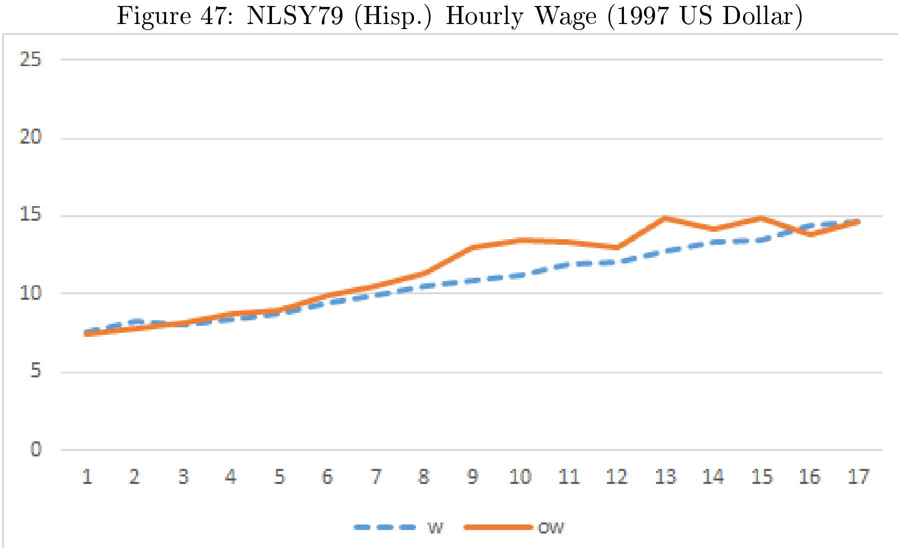


Figure 48: NLSY97 (Hisp.) Proportion of Pre-gap School

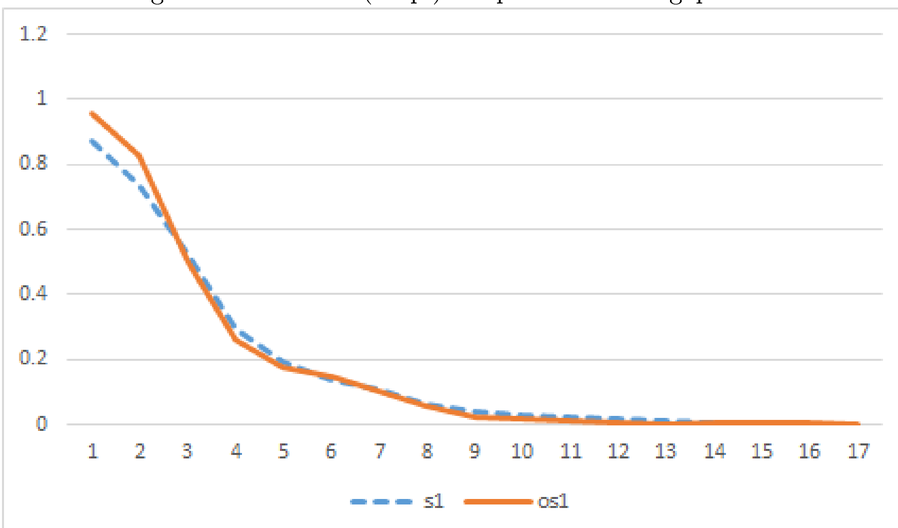


Figure 49: NLSY97 (Hisp.) Proportion of Post-gap School

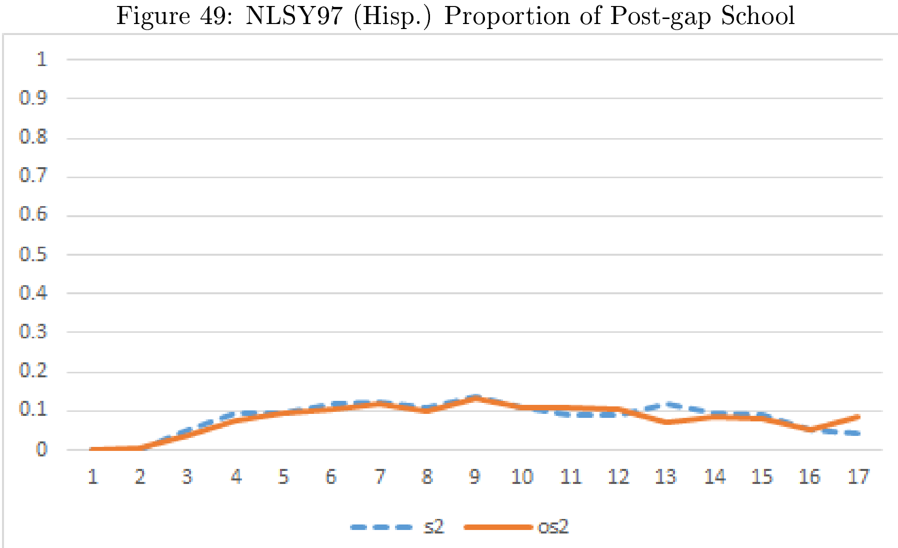


Figure 50: NLSY97 (Hisp.) Proportion of Pre-gap Working

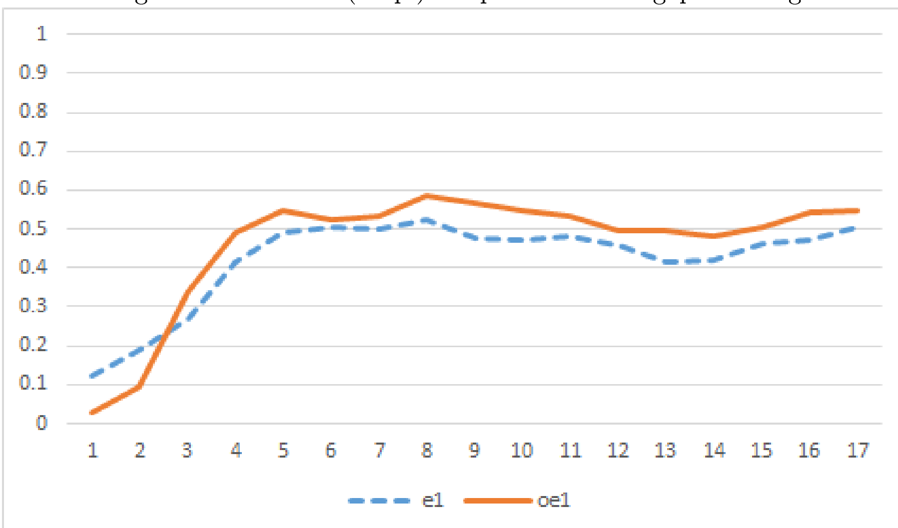


Figure 51: NLSY97 (Hisp.) Proportion of Post-gap Working

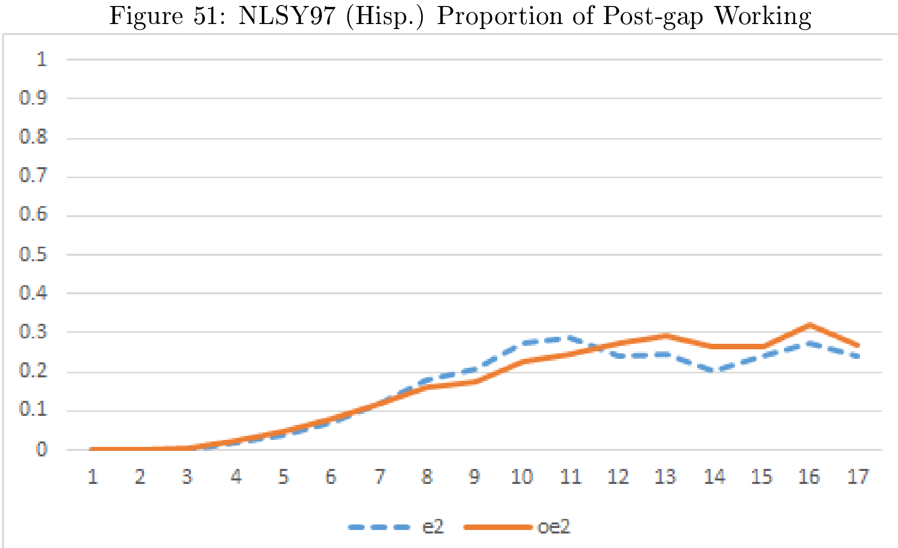


Figure 52: NLSY97 (Hisp.) Proportion of NEET

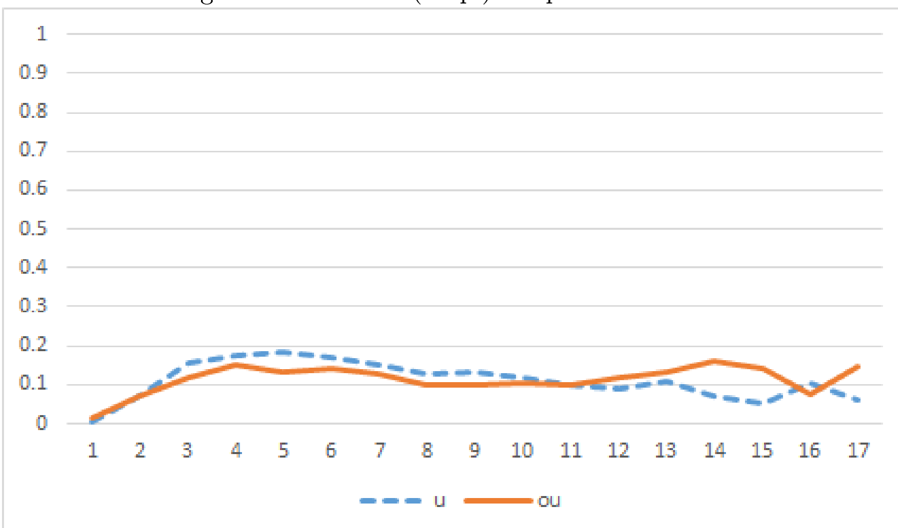


Figure 53: NLSY97 (Hisp.) Proportion of Schooling Interruption

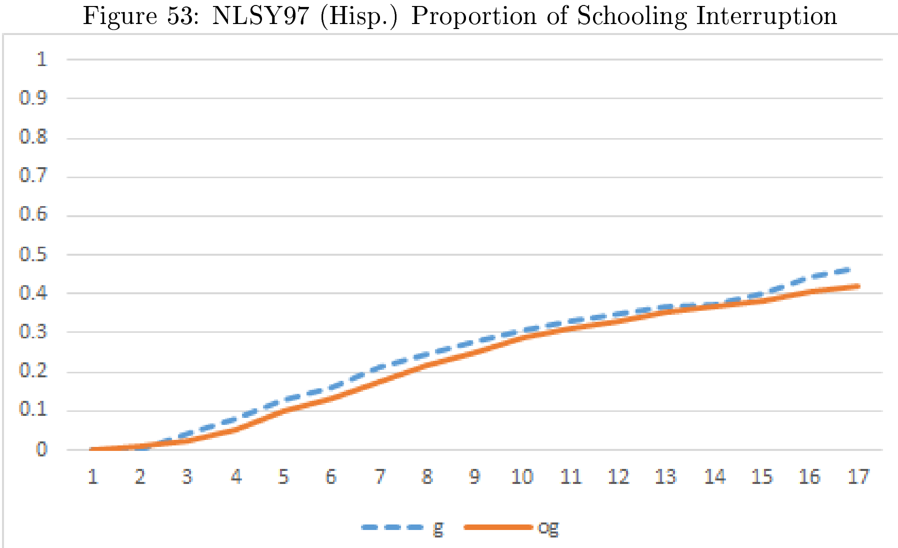


Figure 54: NLSY97 (Hisp.) Hourly Wage (1997 US Dollar)