

Preliminary – Please do not cite

The Old College Try:
Estimating the Role of Uncertainty, Costs and Rewards
in the Experiment called College

Adalbert Mayer
Texas A&M University

October 2008

Abstract

Almost half of all high-school graduates enroll in a four year college but less than two thirds of them eventually graduate. I estimate a model that frames college enrollment as an experiment, which can result in drop-out or graduation. The decision to participate in this experiment depends on the cost of failing the experiment (enrolling and dropping out) relative to the benefit of succeeding (graduating). It also depends on the probability of graduating. This probably depends on characteristics that make it feasible and desirable to graduate and on uncertainty about these characteristics. I find that 8% of the high-school graduates from whom it is feasible and desirable to graduate do not enroll in college. Dropout rates differ across subgroups of the population. Minorities and children of parents without college education or low income have higher drop out rates. I find that this is in part explained by their lower costs of failing relative to the benefits of succeeding. This increases the incentive to enroll in college conditional on the ex-ante probability of graduating. The paper shows how uncertainty in a choice process can be studied by considering sequential choices.

1. Introduction

Almost half of all high-school graduates enroll in a four year college but less than two thirds of them eventually graduate. This suggests that young adults face a sizable amount of uncertainty about their future careers when deciding whether to enroll in college. Yet, most studies of the choice of educational attainment in general, or the choice of obtaining a college degree specifically, do not allow for uncertainty.¹ A number of recent papers (Carneiro, Hansen, and Heckman (2003), Cunha, Heckman, and Navarro (2005), and Navarro (2008)) model the choice between working after high-school and graduating from college, while allowing both choices to be associated with uncertain outcomes. However, they do not explicitly model uncertainty about college completion and treat the decision to enroll in college as synonymous to the decision to graduate from college.

The literature on college completion usually does not explicitly model uncertainty but isolates individual features that are associated with the probability of graduating or dropping out of college (Light and Strayer (2000, 2002) , Jepsen (2008), Stinebrickner and Stinebrickner (forthcoming)). Two papers that deal explicitly with the idea that some uncertainty is resolved during the process of obtaining a college education are Altonji (1993) and Manski (1989). Manski (1989) frames college enrollment as an experiment that reveals the ex-ante unknown feasibility or desirability of college graduation. I built on his framework and develop a two period model of college enrollment and completion.²

My framework offers a novel way to quantify uncertainty in a choice process. I consider a choice under uncertainty (enrollment), which is followed by a second choice (dropout) made after some of the initial uncertainty is resolved. The model presented here allows me to make inferences about the initial uncertainty by using the information contained in the second choice. This makes it possible to infer unobserved values like the share of those who do not enroll in college due to uncertainty. It clarifies the quantitative differences in the elements of the choice process that are needed to generate the

¹ Some examples are Ben-Porath (1967), Willis and Rosen (1979), or Keane and Wolpin (2001).

² Light and Strayer (2000, 2002) also model the decisions to enroll and graduate as a two stage process. They study how attendance and completion is related to race and the match of college qualities and student ability. They identify features that are associated with attendance and completion rates but do not explicitly model the role of uncertainty.

observed differences in enrollment and graduation rates across subgroups of the population. Moreover, the model can be used to assess effects of policy changes on enrollment and graduation rates.

In my model a high-school graduate can forgo a college education and start working³ which is associated with an expected utility. Alternatively, she can engage in the experiment college. The experiment can be successful and the student graduates, or it can fail and the student drops out. A successful experiment promises an increase in expected lifetime utility, the benefit of the experiment. A failed experiment may lead to financial or emotional costs and to a reduction in lifetime utility, the cost of the experiment.

The experiment succeeds if a college student realizes that it is feasible and desirable for her to pursue a college degree and graduates. I capture the individual specific desirability and feasibility of college graduation in a one dimensional variable, desirability of completing college – or short desirability. This variable depends on individual specific characteristics. For example, the scholastic ability needed to fulfill course requirements, motivation, or the ability to reap a financial return from graduation. Desirability may also depend on individual characteristics that are not directly related to scholastic ability, such as tastes associated with the college environment or the lifestyle subsequent to graduation. All of these characteristics may be affected by the amount and quality of prior schooling. An individual's background may also affect desirability, for example wealthy parents can reduce the need to work while in college, college educated parents might be able to provide advice or support that parents without a college experience can not offer.

At the time of enrollment individuals are uncertain about their individual specific desirability of graduating. Students receive a noisy signal of desirability. The probability of succeeding in college depends on the value of this signal and quality of the signal.⁴

The enrollment rates and completion rates implied by the model depend on three values, the distribution of desirability in the population, the quality of the signal, and the costs and benefits of the

³ The working option includes the possibility to enroll in a two year college or a trade school.

⁴ Manski (1989) does not distinguish between signal value and signal quality.

experiment college. A student enrolls in college if the probability of succeeding is sufficiently high, given the costs and benefits associated with failing and succeeding. A high drop out rate can be generated by a low cost to benefit ratio that induces students to enroll even if the odds of success are low. It can also be the result of a lot of uncertainty about the desirability of graduation which leads to 'wrong' enrollment decisions.

I estimate the model with data from the from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). The data contains information on enrollment and graduation as well as a number of demographic characteristics. To identify the three model mechanisms I impose an exclusion restriction. I make the identifying assumption that the performance in an aptitude test (AFQT) is related to the desirability of graduation and to the cost to benefit ratio of the experiment. It is not related to the quality of the signal received prior to college enrollment.

The parameterized model is able to generate the observed enrollment and completion rates conditional on individual characteristics. The average utility cost associated with a failure of the experiment college is estimated to be 30% of the utility benefit associated with a successful experiment. My estimates suggest that 8% of those suitable for college do not engage in the experiment college.

Drop out rates differ across subgroups of the population. Minorities and children of parents without college education or low income have higher drop out rates. This is explained by two factors. First, lower opportunity cost of enrolling in college relative to the benefits of graduating college. This increases the incentive to enroll in college conditional on the ex-ante probability of graduating. Second, lower mean desirability of graduation reflected in lower AFQT scores. This leads to higher dropout rates and to lower enrollment rates for these groups. Differences in the quality of the signal of desirability seem to explain only a small part of the differences in enrollment and dropout rates.

I find that even after controlling for AFQT score both parental education and parental income are positively associated with desirability of college graduation. This is consistent with the presence of credit constraints. However, the magnitudes are small compared to the difference between students with high and low AFQT scores. If AFQT score is viewed as a reflection of prior investment in human

capital my results suggest that improving prior schooling is a way to reduce dropout rates. Policies that reduce the cost of college attendance can trigger enrollment of students with a low ex-ante graduation probability and thus result in higher drop out rates.

In the next section, I present the model and explain the empirical strategy. In Section 3, I describe the data. In Section 4, I present and discuss the results. Section 5 concludes.

2. Model

In this section, I present a two period model of college enrollment and completion. The model is illustrated in Figure 1. In period zero high-school graduates can start working, which results in an expected lifetime utility of V_H . Alternatively, they can enroll in college. I frame college as an experiment with two possible outcomes. In period 1, individuals who attend college either graduate (the experiment is successful) or drop out (the experiment failed). The expected utility of graduating from college is V_G .⁵ The expected utility of attending college and dropping out is given by V_D . The problem is only interesting if it is ex-post not optimal to attend college and drop out, (i.e. $V_H > V_D$) and if attending college and graduating is preferable to not attending college ($V_G > V_H$).

The decision to graduate or to drop out depends on the individual desirability of college completion, a_{1i} .⁶ This variable captures the effects of all the features correlated with the likelihood of succeeding in (or benefiting from) college. Examples for such features are tastes, intelligence, drive, or financial resources. A student prefers graduating over dropping out if the desirability of graduating is above a cut-off level, \underline{a}_1 :

$$\text{graduate} \quad \text{if } a_{1i} > \underline{a}_1,$$

⁵ It is possible to rewrite these utility expressions as a function of discounted expected utilities in multiple periods. E.g. $V_H = \sum_{t=0}^T \beta^t u(H)$, where T stands for the number of periods in the remaining life, β is the discount rate, and $u(H)$ the utility in a given period. I consider expected utilities, for risk-averse individuals these depend on the uncertainty about events after the choices about college enrollment and completion.

⁶ In the empirical part of the paper I allow for a correlation between a_{1i} and V_D or V_G .

and

drop out if $a_{1i} < \underline{a}_1$.

The individual desirability, a_{1i} , is observed in period 1. It is only partially known when the initial enrollment decision is made. In period 0, each individual observes an informative but noisy signal of the desirability of completing college. After receiving her signal, a_{0i} , each individual chooses between the expected utility V_H associated with working and the experiment college. The choice depends on the likelihood that the experiment will be successful, $\Pr(a_{1i} > \underline{a}_1 | a_{0i})$, and on the potential gain of a successful experiment, $V_G - V_H$, relative to the potential loss, $V_D - V_H$, from a failed experiment. Individual i enrolls in college if:

$$\Pr(a_{1i} > \underline{a}_1 | a_{0i})V_G + (1 - \Pr(a_{1i} > \underline{a}_1 | a_{0i}))V_D > V_H,$$

or if

$$\Pr(a_{1i} > \underline{a}_1 | a_{0i}) > \frac{V_H - V_D}{V_G - V_D}. \quad (1)$$

The probability that it is desirable to complete college depends on the signal a_{0i} . The signal is informative and $\frac{d \Pr(a_{1i} > \underline{a}_1)}{da_{0i}} > 0$. This implies a cutoff level, \underline{a}_0 , exist. If $a_{0i} \leq \underline{a}_0$ individual i does not enroll in college. If $a_{0i} > \underline{a}_0$ individual i attends college.

I normalize $V_H = 0$ and $V_G = 1$. The value of V_D measures the loss in expected utility due to the failed experiment (dropping out of college) relative to the possible gain from the successful experiment (graduating from college). I normalize $\underline{a}_1 = 0$ and assume that the distribution of a_{1i} is given by:

$a_{1i} \sim N(\mu_a, 1)$.⁷ The noisy signal received in period one is given by: $a_{0i} = a_{1i} + \varepsilon_i$. The noise term, ε_i , is independent of a_{1i} . I assume that $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. The parameter σ_ε^2 captures the quality of the signal.

⁷ Assuming a variance of one is a normalization without loss of generality.

A high value is associated with a noisy (low quality) signal. A low value implies that most of the variation in a_{1i} is known to individuals in period zero.

Illustration of the model

There are four possible outcomes for each individual (illustrated in Figure 2). Individuals with a low signal a_{0i} do not enroll in college. For some of these individuals college completion is not desirable, they made the ex-post right decision (Outcome 1). For some completion would be desirable ($a_{1i} > 0$) and they made the ex-post wrong decision (Outcome 2). Individuals with $a_{0i} > \underline{a}_0$ enroll in college. Some college attendees realize that college completion is not desirable and drop out (Outcome 3). Those who enroll in college and find it desirable to graduate (Outcome 4).

Figure 3 illustrates the mechanics of the model. Part a) shows the distribution of the desirability of completion. Completion is desirable for the fraction $\Phi(\mu_a)$ of the population.⁸ For example, if $\mu_a = 0$ completion is desirable for half of the individuals. However, since the true value of a_{1i} is unknown at $t=0$, some individuals will make an ex-post ‘wrong’ enrollment choice. Part b) of Figure 3 displays the distribution of a_{0i} for individuals with $a_{1i} = a_{1H} > 0$ and therefore should enroll in college. Most of these individuals obtain a signal, a_{0i} , that is bigger than \underline{a}_0 and enroll in college. But some individuals receive a low signal and decide to forgo a college education. A misleading signal can also lead to enrollment of individuals who do not desire to complete college. This can be seen in part c) of Figure 3. It shows the distribution of the signal for individuals with $a_{1L} < 0$. Some individuals receive a signal above \underline{a}_0 . They enroll and drop out subsequently. A noisier signal – i.e. a higher σ_ε^2 – leads to

⁸ $\Pr(a_{1i} > \underline{a}_1) = 1 - \Phi(-\mu_a) = \Phi(\mu_a)$.

a more spread out distribution of a_{0i} conditional on a_{1i} . This makes both kinds of ex-post suboptimal choices more likely.⁹

The likelihood of ex-post suboptimal choices is also affected by the cost of the experiment relative to the benefits. If the gains associated with graduating are high and the costs of dropping out are low it is worth while to engage in the experiment even if a successful outcome is not likely. As seen in equation (1), the cutoff level \underline{a}_0 depends on the cost of the experiment relative to its benefits, V_D . A higher value of V_D (closer to zero) makes the experiment less costly and lowers the value of \underline{a}_0 . It leads to a higher enrollment rate and to a higher drop out rate. At the same time it reduces the likelihood that individuals with a positive desirability do not enroll in college.

If a high dropout rate is due to a high σ_ε^2 it is associated with high rate of individuals who should have attended college but did not. If the high drop out rate is due to the cost benefit ratio few individuals with a positive desirability do not enroll.

Empirical Strategy

The model is characterized by 3 parameters: μ_a , V_D , and σ_ε^2 . Each of these parameters can depend on a number of observable characteristics captured by X_i . Hence for each value of the vector X_i there are three parameters to be estimated: $\mu_a(X_i)$, $\sigma_\varepsilon^2(X_i)$, and $V_D(X_i)$. However we only observe two data moments. The fraction of individuals that attends college and the fraction that graduates from college:

$$P(\text{College}) = p_C(\mu_a(X_i), \sigma_\varepsilon^2(X_i), V_D(X_i)),$$

and

$$P(\text{Graduate}) = p_G(\mu_a(X_i), \sigma_\varepsilon^2(X_i), V_D(X_i)).$$

⁹ A higher value of σ_ε^2 can increase or decrease enrollment. To see this, consider the two extreme cases: Under complete certainty, $\sigma_\varepsilon^2(X_i) = 0$ only those suitable enroll, $P(\text{Coll}) = \Phi(\mu_a)$. With a completely uninformative signal, $\sigma_\varepsilon^2(X_i) = \infty$, everyone (nobody) enrolls if $\Phi(\mu_a)V_G - (1 - \Phi(\mu_a))V_D > (<)0$.

The model conditional on X_i only is not identified.

I overcome this problem by imposing an exclusion restriction. I assume that AFQT score (a measure of aptitude)¹⁰ is correlated with the expected desirability of completion and the cost / benefit ratio of enrolling in college, but it is not correlated with the uncertainty about the return to college: $\mu_a(X_i, AFQT_i)$ and $V_D(X_i, AFQT_i)$ are correlated with both X_i and $AFQT_i$;¹¹ while $\sigma_\varepsilon^2(X_i)$ is correlated with X_i , but not with $AFQT_i$.

The model parameters define the likelihood of each of the four possible outcomes (see Figure 2):

- (1) An individual does not enroll in college and has no regrets ex-post if $a_{0i} < \underline{a}_0$ and $a_{1i} < \underline{a}_1$. The fraction of the population that falls into this category is given by:

$$P(\text{NoCollege, NoRegret}) = \int_{-\infty}^{\underline{a}_0} \int_{-\infty}^{\underline{a}_1} f(a_{0i}, a_{1i}) da_{1i} da_{0i}.$$

- (2) An individual does not enroll in college and regrets the decision if $a_{0i} < \underline{a}_0$ and $a_{1i} > \underline{a}_1$. The fraction of the population that falls into this category is given by:

$$P(\text{NoCollege, Regret}) = \int_{-\infty}^{\underline{a}_0} \int_{\underline{a}_1}^{\infty} f(a_{0i}, a_{1i}) da_{1i} da_{0i}.$$

- (3) An individual goes to college and drops out if: $a_{0i} > \underline{a}_0$ and $a_{1i} < \underline{a}_1$. The fraction of the population that falls into this category is given by:

$$P(\text{College, Drop}) = \int_{\underline{a}_0}^{\infty} \int_{-\infty}^{\underline{a}_1} f(a_{0i}, a_{1i}) da_{1i} da_{0i}.$$

- (4) An individual goes to college and graduates if: $a_{0i} > \underline{a}_0$ and $a_{1i} > \underline{a}_1$. The fraction of the population that falls into this category is given by:

$$P(\text{College, Graduate}) = \int_{\underline{a}_0}^{\infty} \int_{\underline{a}_1}^{\infty} f(a_{0i}, a_{1i}) da_{1i} da_{0i}.$$

¹⁰ I explain this measure in Section 3.

¹¹ I assume that $\mu_a(X_i, AFQT_i)$ and $V_D(X_i, AFQT_i)$ are both continuous and monotonous functions of $AFQT_i$.

The fractions $P(\text{College,Drop})$ and $P(\text{College,Graduate})$ are observed. The fractions

$P(\text{NoCollege,NoRegret})$ and $P(\text{NoCollege,Regret})$ are not observed separately, only their sum

$P(\text{NoCollege})$ is observed. Hence, I combine these probabilities to derive the likelihood function:

$$L(\mu_a(X_i, AFQT_i), V_D(X_i, AFQT_i), \sigma_\varepsilon^2(X_i)) = \\ I(\text{HighSchool}) \int_{-\infty}^{\underline{a}_0} f(a_{0i}, a_{1i}) da_{0i} + I(\text{DropOut}) \int_{\underline{a}_0}^{\underline{a}_1} \int_{-\infty}^{\underline{a}_1} f(a_{0i}, a_{1i}) da_{1i} da_{0i} + I(\text{Graduate}) \int_{\underline{a}_0}^{\infty} \int_{\underline{a}_1}^{\infty} f(a_{0i}, a_{1i}) da_{1i} da_{0i}$$

It is possible to derive a closed form expression for this likelihood function. The choices of individuals are characterized by two cutoff values, \underline{a}_0 and \underline{a}_1 , and by the joint distribution of a_{0i} and a_{1i} , $f(a_{0i}, a_{1i})$.

The cutoff level for a_{1i} is normalized to zero, $\underline{a}_1 = 0$. The joint distribution of a_{0i} and a_{1i} is given by:

$$\begin{pmatrix} a_{0i} \\ a_{1i} \end{pmatrix} \sim N \left(\begin{bmatrix} \mu_a \\ \mu_a \end{bmatrix}, \begin{bmatrix} 1 + \sigma_\varepsilon^2 & 1 \\ 1 & 1 \end{bmatrix} \right).$$

The probability of a positive desirability conditional on the signal a_{0i} is given by:¹²

$$\Pr(a_{1i} > 0 | a_{0i}) = 1 - \Phi \left(\frac{\underline{a}_1 - \left[\mu_{ai} + \frac{1}{1 + \sigma_{\varepsilon i}^2} (a_{0i} - \mu_{ai}) \right]}{\sqrt{\frac{\sigma_{\varepsilon i}^2}{1 + \sigma_{\varepsilon i}^2}}} \right).$$

Together with expression (1) this implies a cutoff level for a_{0i} , \underline{a}_0 :

$$\frac{V_H - V_D}{V_G - V_D} = 1 - \Phi \left(\frac{\underline{a}_1 - \left[\mu_{ai} + \frac{1}{1 + \sigma_{\varepsilon i}^2} (\underline{a}_0 - \mu_{ai}) \right]}{\sqrt{\frac{\sigma_{\varepsilon i}^2}{1 + \sigma_{\varepsilon i}^2}}} \right).$$

¹² The distribution of a_{1i} conditional on a_{0i} is given by: $(a_{1i} | a_{0i}) \sim N(\mu_{0i}, \sigma_{0i}^2)$. The mean of this distribution is

$E(a_{1i} | a_{0i}) = \mu_{0i} = \mu_{ai} + \frac{1}{1 + \sigma_{\varepsilon i}^2} (a_{0i} - \mu_{ai})$, the variance is $Var(a_{1i} | a_{0i}) = \sigma_0^2 = \frac{\sigma_{\varepsilon i}^2}{1 + \sigma_{\varepsilon i}^2}$.

3. Data

I use data from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). The NLSY79 covers individuals born between 1957 and 1964. It is made up of three sub-samples; a nationally representative sample of 6,111, and a supplemental sample of 5295 that oversamples minorities and youth from economically disadvantaged backgrounds, and a sample of individuals that were enlisted in the military in 1978. I construct two samples; the first (full sample) consists of the individuals in both the national representative sample and the sample of minorities and disadvantaged youth. The second sample (representative sample) is based only on the national representative NLSY79 sample. The score on the Armed Forces Qualifying Test (AFQT) is available for all individuals in the sample. The AFQT is an aptitude test that captures skills valued in college and the workplace. It is commonly used in the literature as a proxy for scholastic ability.

I only consider individuals who graduated from high-school by 1994 and who participated in the survey for at least 6 years after their graduation from high-school and for at least 6 years after their first enrollment in a 4-year college. That leaves 7331 individuals for the full sample and 4618 for the representative sample. In the full (representative) sample information on parental income at ages 17 and/or 18 is available for 4197 (2653) individuals. Table 1 displays the number of observations and the population share for various subgroups. On the left the numbers for the full sample and on the right the numbers for the representative sample.

Table 2 displays the enrollment and graduation rates for 4-year colleges by demographic characteristics and AQFT quartile. I also calculate the completion rate, the probability of graduating conditional on enrolling. For the cohort captured by the NLSY79 the differences in college attendance and graduation rates between genders are small. Minorities are less likely to attend and graduate from college than non-minority students. Children of college educated parents are more likely to attend and graduate from college. The same is true for children with above average parental income. AFQT score is strongly related to enrollment and completion rates. Only 17% (5%) of students in the lowest AFQT

quartile enroll (graduate from) college. For students in the highest AFQT quartile these numbers are 79% and 61%.

All these data patterns are consistent with a number of previous studies that offer additional descriptive details on the rates of college attendance and graduation conditional on individual characteristics. For example, Light and Strayer (2000, 2002) who use NLSY79 data, Belley and Lochner (2007) who analyze college attendance using data from both the NLSY79 and the NLSY97, or Jepsen (2008) who uses NELS88 data.

4. Estimation and Results

The model is characterized by three structural parameters, $\mu_a(X_i, AFQT_i)$, $V_D(X_i, AFQT_i)$, and $\sigma_\varepsilon^2(X_i)$. The first two parameters may depend on the AFQT score of an individual, while $\sigma_\varepsilon^2(X_i)$ is not correlated with AFQT score. All three parameters may depend on a number of individual characteristics captured by X_i . This relationship is described by the following equations:

$$\mu_a(X_i, Z_i) = X_i\beta_\mu + \gamma_{\mu 1}AFQT_i + \gamma_{\mu 2}AFQT_i^2 + \gamma_{\mu 3}AFQT_i^3 + u_i \quad (2)$$

$$\sigma_\varepsilon(X_i) = X_i\beta_\sigma \quad (3)$$

The variables contained in the vector X_i are categorical. The only continuous variable in the model is AFQT score. Use a polynomial of degree three in both equations with AFQT score. The model is only meaningful if $V_D < 0$, and for mechanical reasons the likelihood function is not defined if $V_D > 0$.

Therefore, I specify a functional form than ensures a negative value:¹³

$$V_D(X_i, Z_i) = -e^{X_i\beta_{VD} + \gamma_{VD1}AFQT_i + \gamma_{VD2}AFQT_i^2 + \gamma_{VD3}AFQT_i^3} \quad (4)$$

¹³ The value for σ has to be positive as well, but the estimation algorithm never leads to negative values. Hence no transformation is necessary.

As a robustness check, I estimate the model with an additional exclusion restriction. I assume that residing in an urban or rural area at age 17 is correlated with the cost / benefit ratio of graduation or dropping out and it is also correlated with the quality of the signal $\sigma_\varepsilon^2(X_i)$. I assume that it is not correlated with the expected desirability of completion after conditioning on X_i and AFQT. This leads to 2 specifications:

- (1) Only one exclusion restriction.
- (2) Two exclusion restrictions.

The results are displayed in Table 3. The coefficients for urban in specification (2) are not significant and results for both specifications are very similar.¹⁴

The first column displays the results for equation (2). The expected value of desirability, μ_a , is positively related to AFQT score. Holding the other covariates fixed, graduation seems be slightly more desirable for females – the p-value for the female coefficient is .075. Parental college attendance and parental income exhibit a clear positive relationship with μ_a . This could be due to unobserved academic ability not captured by AFQT score. It could also be the result of help provided by parents through advice or encouragement. Other explanations for this connection are tastes taught by, or inherited from, wealthy or educated parents. It is also possible that wealthy parents reduce the need to work during college or alleviate other credit constraints. Conditional on the other covariates Blacks are more likely to desire to graduate from college. The point estimate for Hispanics is positive but not significant.

The second column displays the results for equation (4). V_D is the expected utility cost of attending college and dropping out relative to the expected gain from graduating. It is a negative number. A small relative cost (a value of V_D close to zero) can be the result of a low utility cost of dropping out compared to the expected utility of not enrolling. It can also be the result of a high utility

¹⁴ As a further robustness check, I also estimate the two specifications with a different functional form for the relationship between the covariates and VD, $V_D(X_i, Z_i) = -\Phi(X_i\beta_{VD} + Z_i\gamma_{VD})$ the results are displayed in Table A2. In general the results are similar. One of the coefficients for urban is now significant.

gain from graduating. The relative cost is higher for children of college educated parents¹⁵ and it is increasing in AFQT score. The other coefficients are not significant. However, it has to be pointed out that some of the standard errors are high – so it is not possible to rule out an effect of any of these covariates.

The third column displays the results for equation (3). The variable σ_ϵ captures the quality of the signal. The estimates for the relationship between the covariates and σ_ϵ are imprecise and it is not possible to make conclusive statements. However, there is to be some evidence that σ_ϵ is higher for minorities.

The results in Table 3 capture marginal effects – holding the other covariates fixed. To analyze differences in behavior between subgroups it is necessary to account for the distribution of covariates in each subgroup. For example, students of college educated parents tend to have above average AFQT scores. Table 4 presents the predicted values for the three structural parameters based on the estimates from specification (1) in Table 3.¹⁶

After accounting for these compositional effects it can be seen that females are predicted to have a higher desirability of college completion. The same is true for children of college educated or high income parents. A random white student is more likely to choose to graduate from college than a random minority student. Whites tend to have higher AFQT scores and are more likely to have college educated and wealthy parents. Students with higher AFQT scores are more likely to desire to complete college.

The second column of Table 4 suggests that the ratio of the expected cost and benefits of enrollment is higher for males than for females. It is also higher for students with college educated or high income parents and students with high AFQT scores. Whites have higher cost relative to benefits than for minority students. As seen in column three, the variance of the noise component, σ_ϵ^2 , is about

¹⁵ In the second specification the point estimate is higher but it is no longer significant.

¹⁶ Note, these are mean predicted values. Some of the standard errors for the coefficients in Table 3 are relatively high. This leads to a range of possible predicted values.

the same size as the variance of the unobserved component of the actual suitability, $Var(a_{1i}) = 1$, conditional on observed covariates. The variance of the signal, σ , is predicted to be higher for children of college educated parents, otherwise it does not vary much across subgroups.

Table 4 also displays the enrollment and graduation probabilities implied by the predicted structural parameters. To facilitate comparison, I show the observed probabilities on the right. It can be seen that the model is able to generate the observed probabilities conditional on demographics and AFQT score.¹⁷ The parameterized model makes it possible to calculate the unobserved outcome of not going to college and regretting it. Table 4 shows that this is a rare event. Only 2% of the population or about 4% of those not attending college would have chosen to graduate had they enrolled in college. Or in other words, 8% of those from whom it is desirable to graduate college do not enroll in college. Individuals who enroll in college and then drop out make up 20% of the population in the full sample.

The model distinguishes three channels that lead to differences in enrollment and graduation rates: differences in desirability of college completion, $\mu_a(AFQT, X_i)$; differences in quality of the signal that for this desirability, $\sigma_\varepsilon(X_i)$; and differences in the cost of enrolling and dropping out relative to the reward from graduating, $V_D(AFQT, X_i)$.

Students from less educated and low income families are less likely to enroll in college and graduate from college. They also have higher drop out rate. The same is true for minority students.¹⁸ The model explains this data pattern with lower cost of the experiment college compared to its benefits for minorities and students from poorer less educated backgrounds. For a given probability of graduating the students have a higher incentive to engage in the experiment college. One explanation for the lower cost to benefit ratio of minorities are admission processes or subsidies that are based on race (affirmative action). Another explanation are higher returns to college for minorities relative to non-minority students (see Arcidiacono et al. 2008). A low cost to benefit rate creates an incentive to

¹⁷ The model has free parameters 24, the numbers in table 4 reflect 27 data moments.

¹⁸ Conditional on AFQT and background variables minorities are more likely to enroll in college and graduate from college than non-minority students.

enroll even if the probability of graduation is relatively low. This results in enrollment of students with a lower expected desirability of graduation. This result is consistent with the results of Light and Strayer (2002). They find “that minorities are more likely than observationally equivalent whites to attend colleges” (p. 43), but minority students who enroll in college “possess fewer favorable unobserved factors”(p. 1).

5. Conclusion

Like Manski (1989) I view enrollment in college as an experiment. I extend his model and estimate three factors that lead to enrollment: desirability of college graduation, the relative costs and benefits of enrolling, and uncertainty. It is not necessary to make assumptions about the relationship between utility from expected income streams and non-pecuniary utility. The model is able to generate the observed enrollment and graduation rates for various population subgroups. The simple decision framework of the model makes it possible to infer a number of parameters of interest that are not directly observable.

The results hinge on the identification assumption that the quality of the signal of desirability is not related to AFQT score. A next step is to explore the possibility to relax the assumption that AFQT score is not related to signal quality. While it might not be possible to obtain estimates for the model parameters it may be feasible to obtain bounds for the parameter values.

A number of future research questions arise naturally from this paper. Soon the individuals covered by the NLSY97 will be old enough to allow a similar analysis based on this data set. This will make it possible to update the analysis for younger generations. Moreover, comparing the results from the two data sets makes it possible to study effects of demographic shifts like the increase in the share of college educated parents.

A natural extension is to link college enrollment and graduation data to subsequent earnings. The choice framework in this paper is based on differences in utility. It abstracts completely from monetary income associated with schooling choices. By combining the utility based approach with income data it might be possible to address issues like non-pecuniary costs and benefits of education choices or risk-aversion.

References

- Altonji Joseph G. (1993) The Demand for and Return to Education When Education Outcomes are Uncertain. *Journal of Labor Economics*, Vol. 11, No. 1, Part 1: 48-83
- Arcidiacono Peter, Patrick Bayer and Aurel Hizmo (2008) “Beyond Signaling and Human Capital: Education and the Revelation of Ability” Mimeo Duke University
- Belley Philippe and Lance Lochner (2007) The Changing Role of Family Income and Ability in Determining Educational Achievement *Journal of Human Capital*, , vol. 1, no. 1
- Ben-Porath, Y. (1967) The production of human capital and the lifecycle of earnings, *Journal of Political Economy*, **75**, 352-365.
- Cameron, S. V., and J. J. Heckman (1998): Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males, *Journal of Political Economy*, 106(2), 262-333.
- Carneiro, P., Hansen, K., and Heckman, J. (2003) Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice, *International Economic Review*, 44, 361–422.
- Cunha, F., J. J. Heckman, and S. Navarro (2005): Separating Uncertainty from Heterogeneity in Life Cycle Earnings, The 2004 Hicks Lecture, *Oxford Economic Papers*, 57(2), 191_261.
- Jepsen, Christopher (2008) Multinomial probit estimates of college completion at 2-year and 4-year schools *Economics Letters*, Volume 98, Issue 2, 155-160.
- Keane, Michael P. 2002. “Financial Aid, Borrowing Constraints, and College Attendance: Evidence from Structural Estimates.” *A.E.R.* 92 (May): 293–97.
- Keane, M. P., and K. I. Wolpin (1997): The Career Decisions of Young Men, *Journal of Political Economy*, 105(3), 473-522.

- Kodde, David A. (1986) Uncertainty and the Demand for Education. *The Review of Economics and Statistics*, Vol. 68, No. 3, 460-467.
- Light Audrey and Wayne Strayer (2000) “Determinants of College Completion: School Quality or Student Ability?”. *Journal of Human Resources* 35 (2000): 299-332.
- Light Audrey and Wayne Strayer (2002) “From Bakke to Hopwood: Does Race Affect College Attendance and Completion?” *Review of Economics and Statistics* 84: 34-44.
- Manski, C. F. (1989). Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Economics of Education Review*, 8, 305–312.
- Navarro Salvador (2008) Using Observed Choices to Infer Agent's Information: Reconsidering the Importance of Borrowing Constraints, Uncertainty and Preferences in College Attendance. Mimeo, University of Wisconsin-Madison
- Stinebrickner, Todd and Ralph Stinebrickner (forthcoming)“The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study,” *American Economic Review*
- Willis, Robert J. and Sherwin Rosen. 1979. “Education and Self-Selection. *Journal of Political Economy* Vol. 87, 7-36.

Figure 1
Decisions and Utilities

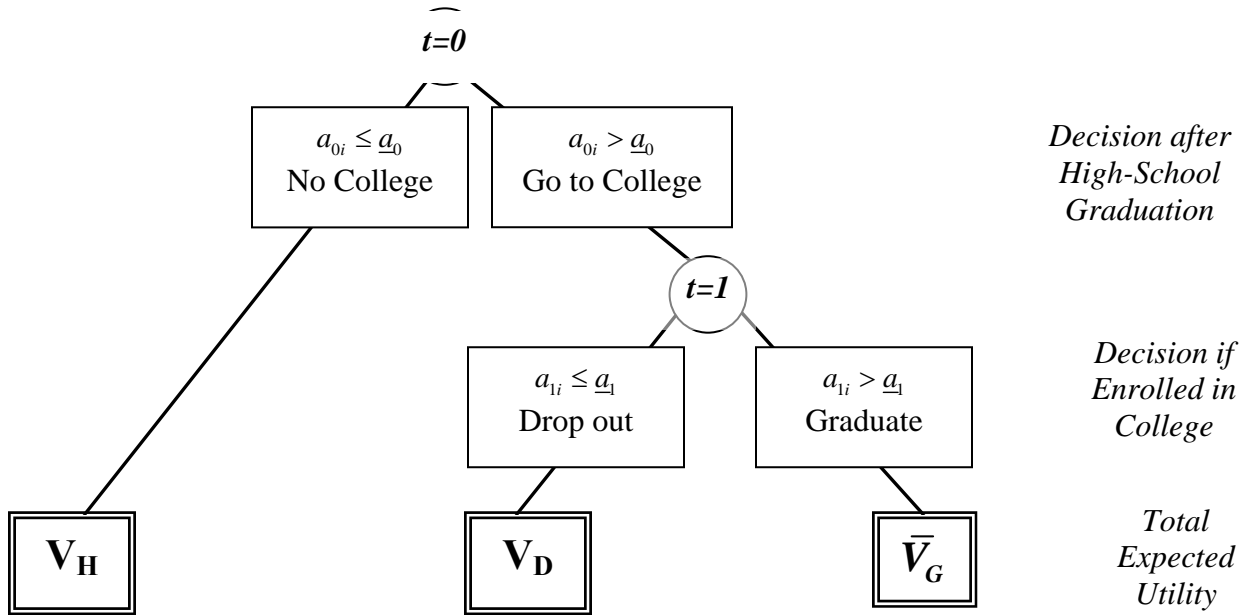


Figure 2
Possible outcomes

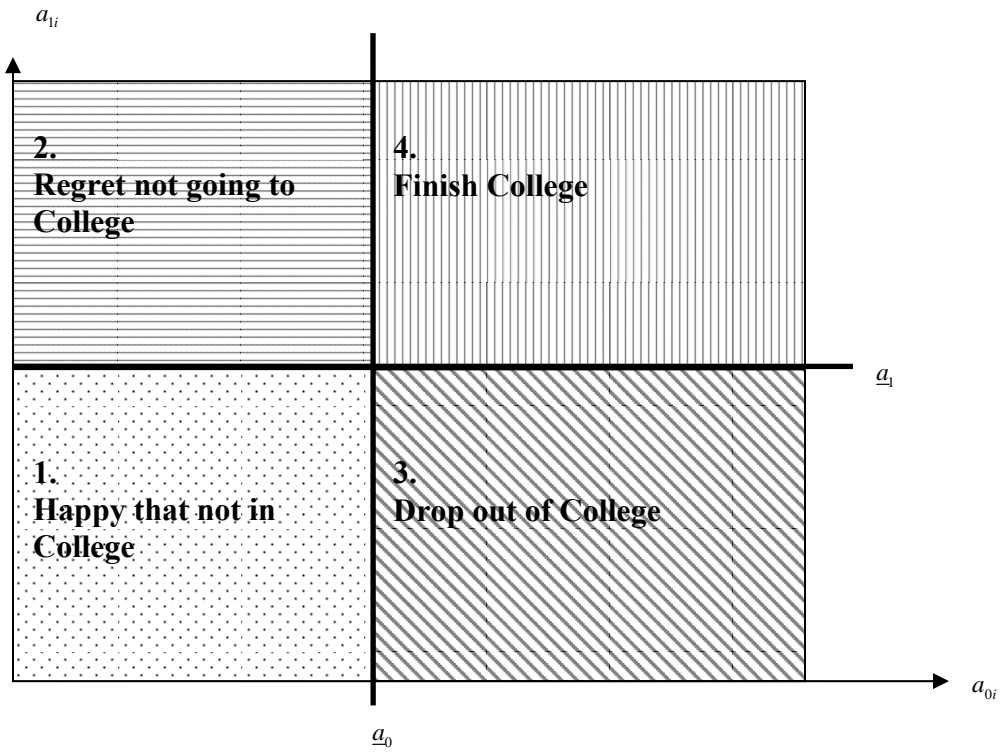
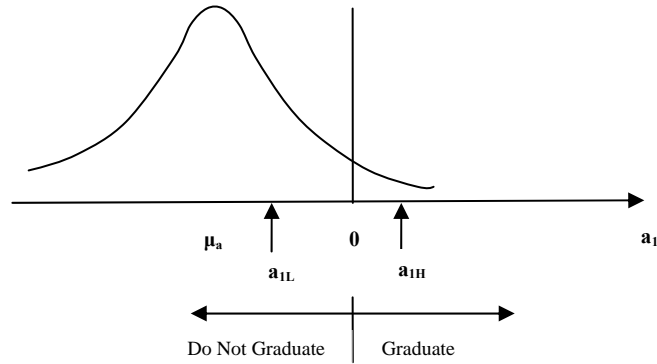
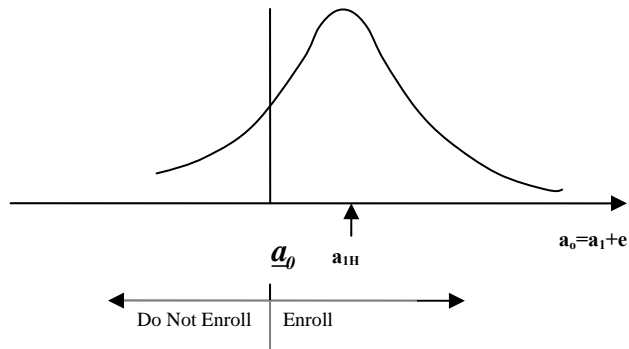


Figure 3
Enrollment and Graduation Probabilities

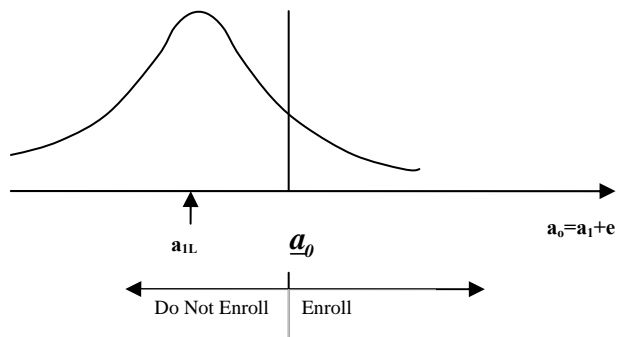


a) Distribution of desirability

a_{1L} negative return to completion, a_{1H} positive return to completion



b) Distribution of signal conditional on actual a_{1H}



c) Distribution of signal conditional on actual a_{1L}

Table 1
Sample Composition

| | Full Sample | | | | | Representative Sample | | | | |
|------------------|--------------|-------|-----------|------|-------|-----------------------|-------|-----------|------|-------|
| | Observations | Mean | Std. Dev. | Min | Max | Observations | Mean | Std. Dev. | Min | Max |
| Female | 7331 | 0.52 | 0.50 | 0 | 1 | 4618 | 0.53 | 0.50 | 0 | 1 |
| Black | 7331 | 0.30 | 0.46 | 0 | 1 | 4618 | 0.11 | 0.32 | 0 | 1 |
| Hispanic | 7331 | 0.14 | 0.35 | 0 | 1 | 4618 | 0.05 | 0.23 | 0 | 1 |
| Mother College | 7331 | 0.18 | 0.38 | 0 | 1 | 4618 | 0.22 | 0.41 | 0 | 1 |
| Father College | 7331 | 0.23 | 0.42 | 0 | 1 | 4618 | 0.29 | 0.45 | 0 | 1 |
| Parental Income | 4197 | 21275 | 14492 | 2004 | 75001 | 2653 | 24570 | 15008 | 2050 | 75001 |
| Urban | 7331 | 0.79 | 0.40 | 0 | 1 | 4618 | 0.77 | 0.42 | 0 | 1 |
| AFQT Percentile | 7331 | 42 | 28 | 1 | 99 | 4618 | 51 | 28 | 1 | 99 |
| Attend College | 7331 | 0.44 | 0.50 | 0 | 1 | 4618 | 0.46 | 0.50 | 0 | 1 |
| Graduate College | 7331 | 0.24 | 0.42 | 0 | 1 | 4618 | 0.28 | 0.45 | 0 | 1 |

Table 2
Enrollement and Graduation Rates

| | Full Sample | | | Representative Sample | | |
|----------------------|-------------------|-------------------|-------------------|-----------------------|-------------------|-------------------|
| | <u>Enrollment</u> | <u>Graduation</u> | <u>Completion</u> | <u>Enrollment</u> | <u>Graduation</u> | <u>Completion</u> |
| Male | 0.43 | 0.23 | 0.53 | 0.47 | 0.29 | 0.62 |
| Female | 0.45 | 0.24 | 0.54 | 0.46 | 0.27 | 0.59 |
| Non-Minority | 0.48 | 0.30 | 0.63 | 0.48 | 0.31 | 0.64 |
| Hispanic | 0.38 | 0.15 | 0.39 | 0.41 | 0.17 | 0.42 |
| Black | 0.40 | 0.15 | 0.38 | 0.36 | 0.13 | 0.37 |
| No Parent College | 0.33 | 0.14 | 0.42 | 0.33 | 0.16 | 0.47 |
| Parent with College | 0.70 | 0.47 | 0.67 | 0.70 | 0.51 | 0.72 |
| Parental Income low | 0.34 | 0.14 | 0.41 | 0.34 | 0.17 | 0.50 |
| Parental Income High | 0.54 | 0.34 | 0.64 | 0.55 | 0.37 | 0.68 |
| AFQT_Q1 | 0.21 | 0.05 | 0.24 | 0.17 | 0.03 | 0.20 |
| AFQT_Q2 | 0.40 | 0.15 | 0.38 | 0.34 | 0.13 | 0.39 |
| AFQT_Q3 | 0.57 | 0.34 | 0.59 | 0.54 | 0.34 | 0.62 |
| AFQT_Q4 | 0.79 | 0.61 | 0.77 | 0.79 | 0.61 | 0.78 |
| Total | 0.44 | 0.24 | 0.53 | 0.46 | 0.28 | 0.61 |

Table 3
 Estimation Results
 Full Sample

| | Specification (1) | | | Specification (2) | | |
|--------------------|------------------------|------------------------|-----------------------|------------------------|------------------------|-----------------------|
| | μ | $\ln(-VD)$ | σ | μ | $\ln(-VD)$ | σ |
| AFQT | 0.070 <i>5.72</i> | 0.077 <i>2.06</i> | - - | 0.073 <i>7.08</i> | 0.098 <i>2.38</i> | - - |
| AFQT squared | -0.001 <i>-4.05</i> | -0.001 <i>-1.43</i> | - - | -0.001 <i>-4.91</i> | -0.002 <i>-1.82</i> | - - |
| AFQT cubed | 0.000 <i>4.30</i> | 0.000 <i>1.38</i> | - - | 0.000 <i>5.05</i> | 0.00 <i>1.70</i> | - - |
| Urban | - - | - - | - - | - - | -0.007 <i>-0.06</i> | 0.08 <i>1.19</i> |
| Female | 0.11 <i>1.78</i> | -0.12 <i>-0.61</i> | -0.17 <i>-1.42</i> | 0.13 <i>2.40</i> | 0.00 <i>0.01</i> | -0.08 <i>-0.68</i> |
| Parent College | 0.68 <i>7.45</i> | 1.37 <i>2.54</i> | 0.75 <i>1.63</i> | 0.76 <i>4.19</i> | 2.31 <i>1.06</i> | 1.26 <i>0.87</i> |
| Parent Income High | 0.20 <i>3.01</i> | -0.19 <i>-0.67</i> | -0.24 <i>-1.39</i> | 0.22 <i>3.82</i> | -0.03 <i>-0.10</i> | -0.11 <i>-0.66</i> |
| Black | 0.39 <i>5.33</i> | 0.02 <i>0.08</i> | 0.21 <i>1.80</i> | 0.38 <i>5.00</i> | 0.07 <i>0.25</i> | 0.16 <i>2.31</i> |
| Hispanic | 0.13 <i>1.57</i> | -0.18 <i>-0.62</i> | 0.22 <i>1.65</i> | 0.12 <i>1.37</i> | -0.12 <i>-0.50</i> | 0.17 <i>1.77</i> |
| Constant | -2.84 <i>-13.83</i> | -3.87 <i>-4.23</i> | 0.82 <i>2.58</i> | -2.93 <i>-19.65</i> | -5.06 <i>-2.05</i> | 0.48 <i>1.26</i> |
| Log Likelihood | -3341.4 | | | -3339.7 | | |
| Observations | 4197 | | | 4197 | | |

z-statistic in small print

Table 4
Predicted Structural Parameters & Probabilities
Based on Results from Specification (1) in Table 3

| | Predicted Values | | | Implied Probabilities | | | | Observed Probabilities | | |
|----------------------|------------------|-------|----------|-----------------------|----------|----------|-----------|------------------------|----------|----------|
| | μ | V_D | σ | Enroll | Graduate | Complete | Regret HS | Enroll | Graduate | Complete |
| Male | -0.94 | -0.27 | 1.03 | 0.42 | 0.22 | 0.52 | 0.02 | 0.43 | 0.23 | 0.53 |
| Female | -0.85 | -0.22 | 0.86 | 0.45 | 0.24 | 0.53 | 0.02 | 0.45 | 0.24 | 0.54 |
| White | -0.62 | -0.34 | 0.89 | 0.48 | 0.30 | 0.62 | 0.03 | 0.48 | 0.30 | 0.63 |
| Black | -1.23 | -0.12 | 1.04 | 0.39 | 0.15 | 0.38 | 0.01 | 0.38 | 0.15 | 0.39 |
| Hispanic | -1.27 | -0.11 | 0.98 | 0.37 | 0.14 | 0.37 | 0.01 | 0.40 | 0.15 | 0.38 |
| Parents No College | -1.26 | -0.09 | 0.76 | 0.33 | 0.14 | 0.41 | 0.01 | 0.33 | 0.14 | 0.42 |
| Parent with College | 0.03 | -0.61 | 1.39 | 0.69 | 0.46 | 0.66 | 0.05 | 0.70 | 0.47 | 0.67 |
| Parental Income Low | -1.22 | -0.17 | 0.99 | 0.35 | 0.14 | 0.41 | 0.02 | 0.34 | 0.14 | 0.41 |
| Parental Income High | -0.45 | -0.34 | 0.89 | 0.55 | 0.34 | 0.62 | 0.02 | 0.54 | 0.34 | 0.64 |
| AFQT_Q1 | -1.74 | -0.06 | 0.92 | 0.22 | 0.05 | 0.22 | 0.01 | 0.21 | 0.05 | 0.24 |
| AFQT_Q2 | -0.93 | -0.16 | 0.91 | 0.41 | 0.17 | 0.41 | 0.02 | 0.40 | 0.15 | 0.38 |
| AFQT_Q3 | -0.34 | -0.32 | 0.97 | 0.59 | 0.34 | 0.58 | 0.04 | 0.57 | 0.34 | 0.59 |
| AFQT_Q4 | 0.56 | -0.77 | 1.04 | 0.83 | 0.66 | 0.80 | 0.03 | 0.79 | 0.61 | 0.77 |
| Total | -0.89 | -0.24 | 0.94 | 0.43 | 0.23 | 0.52 | 0.02 | 0.44 | 0.24 | 0.53 |

Table A1
Numerical Illustration of Model Implications

| $V_D = -.7 \quad \sigma = .5$ | | | | $V_D = -.7 \quad \sigma = 1$ | | | |
|-------------------------------|--------|----------|----------|------------------------------|--------|----------|----------|
| μ | Enroll | Graduate | Complete | μ | Enroll | Graduate | Complete |
| -1.2 | 0.11 | 0.08 | 0.71 | -1.2 | 0.07 | 0.04 | 0.58 |
| -0.6 | 0.29 | 0.22 | 0.76 | -0.6 | 0.27 | 0.17 | 0.63 |
| 0 | 0.54 | 0.45 | 0.82 | 0 | 0.59 | 0.42 | 0.71 |
| 0.6 | 0.78 | 0.69 | 0.88 | 0.6 | 0.86 | 0.69 | 0.80 |

| $V_D = -.3 \quad \sigma = .5$ | | | | $V_D = -.3 \quad \sigma = 1$ | | | |
|-------------------------------|--------|----------|----------|------------------------------|--------|----------|----------|
| μ | Enroll | Graduate | Complete | μ | Enroll | Graduate | Complete |
| -1.2 | 0.17 | 0.09 | 0.57 | -1.2 | 0.17 | 0.07 | 0.42 |
| -0.6 | 0.38 | 0.25 | 0.65 | -0.6 | 0.46 | 0.23 | 0.50 |
| 0 | 0.64 | 0.48 | 0.74 | 0 | 0.77 | 0.47 | 0.62 |
| 0.6 | 0.85 | 0.71 | 0.84 | 0.6 | 0.94 | 0.72 | 0.76 |

Table A 2
 Estimation Results
 Full Sample

| | Specification (3) | | | Specification (4) | | |
|--------------------|------------------------|------------------------|-----------------------|------------------------|-----------------------|-----------------------|
| | μ | $\ln(-VD)$ | σ | μ | $\ln(-VD)$ | σ |
| AFQT | 0.069 <i>5.17</i> | 0.034 <i>1.18</i> | - - | 0.065 <i>5.10</i> | 0.031 <i>1.35</i> | - - |
| AFQT squared | -0.001 <i>-3.77</i> | -0.001 <i>-0.89</i> | - - | -0.001 <i>-3.62</i> | 0.000 <i>-0.94</i> | - - |
| AFQT cubed | 0.000 <i>4.14</i> | 0.000 <i>1.04</i> | - - | 0.000 <i>4.00</i> | 0.000 <i>1.09</i> | - - |
| Urban | - - | - - | - - | - - | -0.51 <i>-2.34</i> | -0.25 <i>-1.63</i> |
| Female | 0.11 <i>1.92</i> | -0.30 <i>-1.26</i> | -0.15 <i>-1.02</i> | 0.11 <i>2.05</i> | -0.07 <i>-0.34</i> | -0.09 <i>-0.86</i> |
| Parent College | 0.63 <i>10.13</i> | 1.06 <i>1.91</i> | 0.43 <i>2.42</i> | 0.60 <i>9.21</i> | 0.66 <i>2.54</i> | 0.28 <i>1.53</i> |
| Parent Income High | 0.25 <i>4.35</i> | -0.17 <i>-0.75</i> | -0.13 <i>-1.65</i> | 0.25 <i>4.25</i> | 0.08 <i>0.34</i> | -0.07 <i>-0.60</i> |
| Black | 0.40 <i>5.56</i> | -0.16 <i>-0.69</i> | 0.08 <i>1.01</i> | 0.43 <i>5.40</i> | 0.32 <i>0.99</i> | 0.39 <i>2.09</i> |
| Hispanic | 0.13 <i>1.56</i> | -0.24 <i>-0.89</i> | 0.10 <i>1.12</i> | 0.12 <i>1.35</i> | -0.16 <i>-0.54</i> | 0.15 <i>1.20</i> |
| Constant | -2.90 <i>-13.73</i> | -2.35 <i>-2.89</i> | 0.56 <i>2.22</i> | -2.83 <i>-14.70</i> | -1.89 <i>-4.99</i> | 0.78 <i>4.15</i> |
| Log Likelihood | -3341.3 | | | -3338.0 | | |
| Observations | 4197 | | | 4197 | | |

z-statistic in small print

Table A 3 a)
Predicted Structural Parameters & Probabilities
Based on Results from Specification (2) in Table 3

| | Predicted Values | | | Implied Probabilities | | | | Observed Probabilities | | |
|----------------------|------------------|-------|----------|-----------------------|----------|----------|-----------|------------------------|----------|----------|
| | μ | V_D | σ | Enroll | Graduate | Complete | Regret HS | Enroll | Graduate | Complete |
| Male | -0.97 | -0.27 | 0.94 | 0.42 | 0.22 | 0.52 | 0.02 | 0.43 | 0.23 | 0.53 |
| Female | -0.86 | -0.24 | 0.84 | 0.45 | 0.24 | 0.53 | 0.02 | 0.45 | 0.24 | 0.54 |
| White | -0.62 | -0.38 | 0.91 | 0.48 | 0.30 | 0.62 | 0.03 | 0.48 | 0.30 | 0.63 |
| Black | -1.26 | -0.11 | 0.88 | 0.39 | 0.15 | 0.38 | 0.01 | 0.38 | 0.15 | 0.39 |
| Hispanic | -1.31 | -0.09 | 0.81 | 0.37 | 0.13 | 0.36 | 0.01 | 0.40 | 0.15 | 0.38 |
| Parents No College | -1.30 | -0.04 | 0.55 | 0.33 | 0.14 | 0.41 | 0.00 | 0.33 | 0.14 | 0.42 |
| Parent with College | 0.07 | -0.80 | 1.74 | 0.69 | 0.46 | 0.66 | 0.06 | 0.70 | 0.47 | 0.67 |
| Parental Income Low | -1.26 | -0.14 | 0.81 | 0.35 | 0.14 | 0.41 | 0.02 | 0.34 | 0.14 | 0.41 |
| Parental Income High | -0.45 | -0.41 | 0.99 | 0.55 | 0.34 | 0.62 | 0.03 | 0.54 | 0.34 | 0.64 |
| AFQT_Q1 | -1.78 | -0.04 | 0.74 | 0.22 | 0.05 | 0.22 | 0.01 | 0.21 | 0.05 | 0.24 |
| AFQT_Q2 | -0.93 | -0.15 | 0.85 | 0.41 | 0.17 | 0.42 | 0.02 | 0.40 | 0.15 | 0.38 |
| AFQT_Q3 | -0.34 | -0.34 | 1.02 | 0.59 | 0.34 | 0.58 | 0.04 | 0.57 | 0.34 | 0.59 |
| AFQT_Q4 | 0.56 | -0.95 | 1.18 | 0.83 | 0.66 | 0.80 | 0.03 | 0.79 | 0.61 | 0.77 |
| Total | -0.91 | -0.26 | 0.89 | 0.43 | 0.23 | 0.52 | 0.02 | 0.44 | 0.24 | 0.53 |

Table A 3 b)
Predicted Structural Parameters & Probabilities
Based on Results from Specification (3) in Table A2

| | Predicted Values | | | Implied Probabilities | | | | Observed Probabilities | | |
|----------------------|------------------|-------|----------|-----------------------|----------|----------|-----------|------------------------|----------|----------|
| | μ | V_D | σ | Enroll | Graduate | Complete | Regret HS | Enroll | Graduate | Complete |
| Male | -1.00 | -0.14 | 0.67 | 0.42 | 0.22 | 0.52 | 0.01 | 0.43 | 0.23 | 0.53 |
| Female | -0.91 | -0.09 | 0.52 | 0.45 | 0.23 | 0.53 | 0.00 | 0.45 | 0.24 | 0.54 |
| White | -0.67 | -0.16 | 0.57 | 0.47 | 0.30 | 0.63 | 0.01 | 0.48 | 0.30 | 0.63 |
| Black | -1.29 | -0.05 | 0.62 | 0.39 | 0.15 | 0.37 | 0.00 | 0.38 | 0.15 | 0.39 |
| Hispanic | -1.34 | -0.04 | 0.60 | 0.37 | 0.13 | 0.36 | 0.00 | 0.40 | 0.15 | 0.38 |
| Parents No College | -1.31 | -0.04 | 0.49 | 0.33 | 0.14 | 0.41 | 0.00 | 0.33 | 0.14 | 0.42 |
| Parent with College | -0.06 | -0.30 | 0.86 | 0.69 | 0.46 | 0.66 | 0.02 | 0.70 | 0.47 | 0.67 |
| Parental Income Low | -1.29 | -0.07 | 0.61 | 0.35 | 0.14 | 0.41 | 0.01 | 0.34 | 0.14 | 0.41 |
| Parental Income High | -0.49 | -0.16 | 0.57 | 0.55 | 0.34 | 0.62 | 0.01 | 0.54 | 0.34 | 0.64 |
| AFQT_Q1 | -1.79 | -0.02 | 0.57 | 0.22 | 0.05 | 0.22 | 0.00 | 0.21 | 0.05 | 0.24 |
| AFQT_Q2 | -1.00 | -0.07 | 0.57 | 0.42 | 0.17 | 0.41 | 0.01 | 0.40 | 0.15 | 0.38 |
| AFQT_Q3 | -0.40 | -0.15 | 0.62 | 0.59 | 0.34 | 0.58 | 0.01 | 0.57 | 0.34 | 0.59 |
| AFQT_Q4 | 0.53 | -0.37 | 0.66 | 0.82 | 0.66 | 0.81 | 0.01 | 0.79 | 0.61 | 0.77 |
| Total | -0.95 | -0.11 | 0.59 | 0.43 | 0.23 | 0.52 | 0.01 | 0.44 | 0.24 | 0.53 |

Table A 3 c)
Predicted Structural Parameters & Probabilities
Based on Results from Specification (4) in Table A2

| | Predicted Values | | | Implied Probabilities | | | | Observed Probabilities | | |
|----------------------|------------------|-------|----------|-----------------------|----------|----------|-----------|------------------------|----------|----------|
| | μ | V_D | σ | Enroll | Graduate | Complete | Regret HS | Enroll | Graduate | Complete |
| Male | -0.97 | -0.15 | 0.77 | 0.42 | 0.22 | 0.53 | 0.01 | 0.43 | 0.23 | 0.53 |
| Female | -0.88 | -0.13 | 0.67 | 0.45 | 0.23 | 0.53 | 0.01 | 0.45 | 0.24 | 0.54 |
| White | -0.66 | -0.18 | 0.61 | 0.47 | 0.30 | 0.63 | 0.01 | 0.48 | 0.30 | 0.63 |
| Black | -1.23 | -0.10 | 0.95 | 0.39 | 0.15 | 0.37 | 0.01 | 0.38 | 0.15 | 0.39 |
| Hispanic | -1.33 | -0.05 | 0.68 | 0.37 | 0.13 | 0.36 | 0.00 | 0.40 | 0.15 | 0.38 |
| Parents No College | -1.28 | -0.08 | 0.67 | 0.33 | 0.14 | 0.41 | 0.01 | 0.33 | 0.14 | 0.42 |
| Parent with College | -0.06 | -0.29 | 0.84 | 0.69 | 0.46 | 0.66 | 0.02 | 0.70 | 0.47 | 0.67 |
| Parental Income Low | -1.27 | -0.10 | 0.76 | 0.35 | 0.14 | 0.41 | 0.01 | 0.34 | 0.14 | 0.41 |
| Parental Income High | -0.48 | -0.20 | 0.66 | 0.55 | 0.34 | 0.62 | 0.01 | 0.54 | 0.34 | 0.64 |
| AFQT_Q1 | -1.75 | -0.05 | 0.78 | 0.22 | 0.05 | 0.22 | 0.01 | 0.21 | 0.05 | 0.24 |
| AFQT_Q2 | -0.98 | -0.10 | 0.70 | 0.42 | 0.17 | 0.40 | 0.01 | 0.40 | 0.15 | 0.38 |
| AFQT_Q3 | -0.39 | -0.18 | 0.68 | 0.59 | 0.34 | 0.58 | 0.02 | 0.57 | 0.34 | 0.59 |
| AFQT_Q4 | 0.54 | -0.40 | 0.68 | 0.82 | 0.66 | 0.81 | 0.02 | 0.79 | 0.61 | 0.77 |
| Total | -0.93 | -0.14 | 0.72 | 0.43 | 0.23 | 0.53 | 0.01 | 0.44 | 0.24 | 0.53 |