

College attrition and the dynamics of information revelation ^{*}

Preliminary and Incomplete

Peter Arcidiacono [†] Esteban Aucejo[‡] Arnaud Maurel[§]
Tyler Ransom [¶]

March 16, 2013

Abstract

This paper investigates the determinants of college attrition, in a setting where individuals have imperfect information about their schooling ability and labor market productivity. We estimate, using data from the National Longitudinal Survey of Youth 1997 (NLSY97), a dynamic structural model of schooling and work decisions, where students who graduated from high school (or obtained a GED) decide at each period whether to attend college, work part-time or full-time, or engage in home production. A key feature of our approach is to account for correlated learning through college grades and wages, thus implying that individuals may leave or re-enter college as a result of the arrival of new information on their ability and productivity. Counterfactual simulations allow us to quantify the importance of informational frictions in explaining the observed school-to-work transitions, and to evaluate the value of information in this context.

JEL Classification: C35; D83; J24

^{*}We thank the attendees of the 2012 Conference in Honor of John Kennan, the 2012 Cowles Foundation Conference on Structural Microeconomics and the 2012 North American Summer Meeting of the Econometric Society (Northwestern), as well as seminar participants at Arizona State and CREST for useful comments and suggestions.

[†]Duke University and NBER. E-mail address: psarcidi@econ.duke.edu

[‡]London School of Economics. E-mail address: E.M.Aucejo@lse.ac.uk

[§]Duke University and IZA. E-mail address: apm16@duke.edu

[¶]Duke University. E-mail address: tyler.ransom@duke.edu

1 Introduction

The determinants of college attrition have been the object of a large number of papers in the economics of education field. This literature was stimulated by the very substantial rate of college attendees in the United States entering the labor market without any post-secondary diploma. About half of students entering college in the United States fail to receive a bachelor's degree within five years, this proportion being fairly stable since the 1970's (see, e.g., Bound & Turner, 2011). From the viewpoint of the human capital theory of schooling investment (Becker, 1964), this stylized fact is puzzling since there are substantial sheepskin effects for obtaining a college degree (see, e.g., Heckman et al., 2006 and Bound & Turner, 2011). It would therefore not be rational for a student to decide *ex ante* to drop-out from college, thus suggesting that these decisions are the result of the arrival of new information or stem from the consumption value of college attendance. In this paper, we focus on the first type of explanation, and investigate the role played by learning about academic ability as well as labor market productivity on college attrition and re-entry. We are particularly interested in quantifying the importance of informational frictions in explaining the observed school-to-work transitions, and to evaluate the value of information in this context. Noteworthy, in the current environment where very high college attrition rates are considered as a major issue, this is important to understand (i) whether these attrition rates *should* be a concern, and (ii) which type of policy would be effective in reducing the current attrition rates.

Seminal research by Manski & Wise (1983) and Manski (1989) argued that college entry can be seen as an experiment that may not lead to a college degree. According to these authors, an important determinant of college drop-out lies in the fact that, after entering college, students get new information and thus learn about their ability. More recently, several other papers in the literature on college completion stress the importance of learning about schooling ability to account for college attrition (see, e.g., Altonji, 1993, Light & Strayer, 2000, Arcidiacono, 2004, and Stratton et al., 2008). Of particular relevance to us is the work by Stinebrickner & Stinebrickner(2012), who provide direct evidence, using subjective expectations data from Berea College, that learning about schooling ability is a major determinant of the college drop-out decision.

We contribute to the literature by specifying and estimating, using data from the National Longitudinal Survey of Youth 1997 (NLSY97), a dynamic structural model of schooling and work decisions. After graduating from high school or receiving a GED, individuals decide at each period whether to attend college, work part-time or full-time, or engage in home production. We take into account the heterogeneity in schooling investments by distinguishing between two- and four-year colleges, as well as Science and non-Science majors for four-year colleges.¹ A central feature of our model lies in the fact that students may leave college, either temporarily or permanently, as a result of the arrival of new information on their schooling ability and labor market productivity. A key difference with the standard dynamic structural schooling choice literature, initiated by Keane & Wolpin (1997), is that we allow students to have imperfect information about their ability. Namely, when entering college, individuals have some beliefs about their ability and productivity. At the end of each school year, they learn about their ability and use their grades to update their beliefs. Since schooling ability and productivity will in general be correlated, individuals will also use their grades to update their productivity belief. Similarly, along the vein of Miller (1984), employed individuals update their productivity, as well as ability beliefs, after receiving a wage.

Stange (2012) and Stinebrickner & Stinebrickner (2013) also estimates a dynamic model of college attendance decisions allowing for ability learning. However, unlike our model, the labor market is assumed to be an absorbing state, implying that the decision to leave college is irreversible. By relaxing this assumption, and allowing individuals to learn about their productivity through their wages, our model is able to predict the substantial college re-entry rates which are observed in the data.² This is an important step towards a comprehensive analysis of school-to-work transitions, building on the insights of Pugatch (2012) who provides evidence from South African data that the option to re-enroll in high school is a key determinant of the decisions to leave school and enter the labor market. In a work in progress, Heckman & Urzua (2009) estimate a dynamic model of schooling decisions where individuals learn about their ability, as well as their preferences. Their model differs from ours in several important dimensions;

¹See the recent survey by Altonji et al. (2012), who discuss the importance of heterogeneity in human capital investments.

²In our sample, 36% of the individuals leaving college before graduation are observed to re-enroll at some point. Note that because of right-censoring, this underestimates the actual re-entry rate.

notably, *ex ante* unknown ability components are assumed to be immediately revealed upon occupying the corresponding states, while in our setting students gradually learn about their ability and productivity.

Estimating dynamic models of correlated learning has been, until recently, computationally very costly. However, recent work by James (2011) shows that 1) integrating out over actual abilities as opposed to the signals and 2) using the EM algorithm where at the maximization step ability is treated as known, results in models that are computationally very fast. James (2011) builds on the results from Arcidiacono & Miller (2011) to show that estimation is computationally simple even in the presence of unobserved heterogeneity that is known to the individual. Using this approach in our current context makes estimation of our correlated learning model both feasible and fast.

We then use our model to simulate what would be the effect, in terms of college attrition rates, of providing more information to the students regarding their ability. Our preliminary results suggests that learning about schooling ability and labor market productivity does influence decisions. In particular, those who have stopped-out of college have learned that their academic abilities are relatively low but that their productivity in the unskilled sector is also low. In contrast, dropout perform poorly academically and stay out of school because their productivity in the unskilled sector is sufficiently high.

The remainder of the paper is organized as follows. Section 2 presents a dynamic model of schooling and work decisions, where individuals have imperfect information about their schooling ability and labor market productivity, and update their beliefs through the observation of grades and wages. Section 3 informally discusses the identification of the model, and then details the estimation procedure, before describing the data in Section 4. Section 5 presents our preliminary estimation results, and Section 6 concludes.

2 The model

2.1 The setting

Time is supposed to be discrete, with horizon T . We denote by $\beta \in (0, 1)$ the discount factor. After graduating from high school or receiving a GED, individuals make at the beginning of each period t a joint schooling and work decision. Namely, they decide whether to attend college, either in a two-year or four-year institution, as well as whether to enter the labor market and work part-time or full-time. Students enrolled in a four-year college also get to choose between a Science and non-Science major.³ Importantly, none of these decisions lead to an absorbing state. In particular, in contrast with, e.g., Stange (2012), our model does allow for college re-entry. It is also worth noting that this model allows for college employment.⁴ Working while in college may be detrimental to academic performance (see, e.g., Stinebrickner & Stinebrickner, 2003) but is also likely to be a channel through which individuals learn about their productivity. Our framework will incorporate this tradeoff. Finally, individuals are forward-looking and choose the sequence of actions yielding the highest value of expected lifetime utility.

The general idea underlying the information structure of the model is the following. Aside from the future state-specific idiosyncratic shocks, individuals only have imperfect information about (i) their schooling ability and (ii) their labor market productivity. Specifically, before enrolling in college, individuals are supposed to have a belief about their schooling ability. If they attend college, they learn about their ability by observing their schooling performance, as measured by their Grade Point Average (GPA) at the end of the academic year. The gap between the observed and predicted GPA then provides a noisy signal for their ability, which is used to update their belief in a Bayesian

³It follows that individuals effectively choose at each period among twelve (mutually exclusive) alternatives, namely (1) attending a two-year college without working, (2) attending a four-year college, Science major, without working, (3) attending a four-year college, non-Science major, without working, (4) entering the labor market and working part-time or (5) full-time, (6) working part-time while in two-year or (7) four-year college, Science major, (8) four-year college, non-Science major, (9) working full-time while in two-year or (10) four-year college, Science major, (11) four-year college, non-Science major, and finally (12) entering the home production sector.

⁴See also Joensen (2009) who estimates a dynamic structural model of schooling and work decisions allowing for work while in college.

fashion. Since schooling ability and productivity will in general be correlated, the GPA also provides some information about labor market productivity. Individuals will therefore use their GPA to update their productivity belief. Similarly, those who participate to the labor market update their productivity as well as their ability beliefs after receiving their wage.

In this model, individuals who choose to work while in college will get two signals, through their GPA and their wage, on their ability and productivity. Interestingly, without the need to invoke a credit constraint argument, the value of information implies that working while in college may be optimal for some students in spite of a detrimental impact on their academic performance.

We shall now detail the main components of the model, namely the grade equations and the labor market, together with the learning process, and the flow utility functions for each alternative.

2.2 Grades

We denote by $j \in \{a, bs, bn\}$ the type of college and major attended, where a (for *Associate*) denotes a two-year college, bs (for *Bachelor, Science*) a four-year college Science major, and bn (for *Bachelor, non-Science*) a four-year college non-Science major. We assume that grades depend on A_{ij} where A_{ij} is the unobserved schooling ability about which individuals have some beliefs initially given by the prior distribution $\mathcal{N}(0, \sigma_{A_j}^2)$. Grades also depend on X_{ict} , a vector of observed covariates that includes an observed measure for ability (AFQT) and a set of demographic characteristics (gender, race, etc), as well as indicators for whether the individual is returning to college after a break, switching colleges or majors.

Grades in two-year colleges and in the first two years of four-year colleges are given by, denoting by τ the period of college enrollment :

$$G_{ij\tau} = \gamma_{0j} + X_{ict}\gamma_{1j} + A_{ij} + \varepsilon_{ij\tau}$$

The idiosyncratic shocks $\varepsilon_{ij\tau}$ are distributed $\mathcal{N}(0, \sigma_{j\tau}^2)$ and are independent from the other state variables. Define the type- j (college, major) academic index of i at time t , AI_{ijt} , as:

$$AI_{ijt} = \gamma_{0j} + X_{ict}\gamma_{1j} + A_{ij}$$

The academic index AI_{ijt} gives expected grades conditional on knowing A_{ij} but not the idiosyncratic shock $\varepsilon_{ij\tau}$.

For four-year colleges and periods $\tau > 2$, we express grades relative to AI_{ijt} as follows:⁵

$$G_{ij\tau} = \lambda_{0j} + \lambda_{1j}AI_{ijt} + \varepsilon_{ij\tau}$$

Hence, the return to the academic index varies over period of college enrollment and across majors. In particular, consistently with Hansen et al. (2004), our specification allows the effect of latent ability on grades to decrease with the number of years spent in college.

2.3 A two-sector labor market

Upon entering the labor force (either full-time or part-time) each individual receives at each period t an hourly wage W_{ikt} which depends on his graduation status. We assume that there are two sectors in the labor market, which are referred to in the following as *skilled* (four-year college graduates labor force only, $l = s$) and *unskilled* (all the others labor market participants, including high school graduates or GED recipients, college dropouts and stopouts as well as two-year college graduates, $l = u$).

We suppose that in the skilled sector, wages depend on productivity A_{is} as well as on a set of observed characteristics X_{ist} :

$$\ln(W_{ist}) = \gamma_{0s} + X_{ist}\gamma_{1s} + A_{is} + \varepsilon_{ist}$$

X_{ist} is then a vector of observed covariates such as gender, race, AFQT as well as time-varying characteristics such as age, labor market experience and a dummy for work part-time. Noteworthy, we account for nonstationarity in the wages by including calendar year dummies in X_{ist} , thus incorporating business cycle effects. Finally, the idiosyncratic shocks ε_{st} are distributed $\mathcal{N}(0, \sigma_s^2)$ and are independent over time and independent of the other state variables.

Wages in the unskilled sector follow a similar pattern. Wages depend on productivity A_{iu} and observed covariates X_{iut} :

$$\ln(W_{iut}) = \gamma_{0u} + X_{iut}\gamma_{1u} + A_{iu} + \varepsilon_{iut}$$

⁵See Arcidiacono (2004) for a similar ability index specification.

where ε_{ut} are distributed $\mathcal{N}(0, \sigma_u^2)$ and again independent over time and independent from the other state variables. Finally, X_{iut} includes the set of characteristics discussed above as well as dummies for the number of completed years of college.

2.4 Learning

Individuals are uncertain about their schooling ability and labor market productivity $A_i = (A_{ia}, A_{ibs}, A_{ibn}, A_{is}, A_{iu})'$ (simply referred to as *ability* in the following), and are assumed to learn about them in Bayesian fashion. Their initial ability beliefs are given by the population distribution of A , which is supposed to be multivariate normal with mean zero and covariance matrix Δ . Importantly, we do not restrict Δ to be diagonal, thus allowing for correlated learning across the five different ability components.⁶

Namely, at each period t of college attendance, individuals use their GPA to update their belief about their schooling ability in all college options (A_{ia}, A_{ibs}, A_{ibn}), as well as their labor market productivity in both sectors (A_{is}, A_{iu}). The GPA provides a noisy signal for their ability, which is denoted by $S_{ij\tau}$ for type- j college option and period of enrollment τ . For two-year colleges and the first two years of four-year colleges, the signal is given by:

$$S_{ij\tau} = G_{ij\tau} - \gamma_{0j} - X_{ict}\gamma_{1j}$$

For four-year colleges and subsequent periods ($\tau > 2$), the index specification yields:

$$S_{ij\tau} = \frac{G_{ij\tau} - \lambda_{0j} - \lambda_{1j}(\gamma_{0j} + X_{ict}\gamma_{1j})}{\lambda_{1j}}$$

Similarly, individuals who participate to the labor market update their ability beliefs after receiving their wages. The signal is given by, for sector l and period t :

$$S_{ilt} = \ln(W_{ilt}) - \gamma_{0l} - X_{ilt}\gamma_{1l}$$

Finally, individuals may choose to work while in college, in which case they will receive two ability signals ($S_{ij\tau}, S_{ilt}$).

It follows from the normality assumptions on the initial prior ability distribution and on the idiosyncratic shocks that the posterior ability distributions are also normally

⁶See also Antonovics & Golan (2012), James (2011) and Sanders (2010) who estimate occupational choice models with correlated learning across occupations.

distributed. Specifically, denoting by $E_t(A_i)$ and $\Sigma_t(A_i)$ the posterior ability mean and covariance at the end of period t , we have (see DeGroot, 1970):

$$E_t(A_i) = (\Sigma_{t-1}^{-1}(A_i) + \Omega_{it})^{-1}(\Sigma_{t-1}^{-1}(A_i)E_{t-1}(A_i) + \Omega_{it}\tilde{S}_{it}) \quad (2.1)$$

$$\Sigma_t(A_i) = (\Sigma_{t-1}^{-1}(A_i) + \Omega_{it})^{-1} \quad (2.2)$$

Ω_{it} is a (5×5) matrix with zeros everywhere except for the diagonal terms corresponding to the occupations of the individual in period t (namely two-year college, four-year college Science major, four-year college non-Science major, skilled or unskilled labor), which are given by the inverse of the idiosyncratic shock variances. \tilde{S}_{it} a (5×1) vector with zeros everywhere except for the elements corresponding to the occupations of the individual in period t , which are given by the ability signals received in this period.

2.5 Flow utilities

We denote in the following by $d_{it} = (j, k)$ the choice for individual i at time t over school, $j \in \{a, bs, bn, 0\}$, and work $k \in \{p, f, 0\}$, where p and f refer to full and part-time work. The choice $d_{it} = (0, 0)$ then indicates the home option: no work and no school. Should the individual be a four-year college graduate, the individual only chooses among the work alternatives, implying $j = 0$.

Up to an intercept term, we assume that the utility of the choice (j, k) is additively separable. Let Z_{1it} denote variables that affect the utility of school and Z_{2it} denote the variables that affect the utility of working. The flow payoff for choice (j, k) is given by:

$$u_{jkt}(Z_{it}, \varepsilon_{ijk}) = \alpha_{jk} + Z_{1it}\alpha_j + Z_{2it}\alpha_k + \varepsilon_{ijk}$$

where Z_{it} includes characteristics such as AFQT, race, gender, as well as the previous choice. Controlling for the previous choice allows for switching costs, in a similar spirit as in Keane & Wolpin (1997). The idiosyncratic preference shocks ε_{ijk} are assumed to follow a (standard) Type-I extreme value distribution.

Embedded in Z_{1it} are expected grades which in turn depend on expected ability. Embedded in Z_{2it} are expected wages. Both of these will vary given the choices. However, to conserve on notation we do not put jk subscripts on the Z 's.

Finally, the home production sector ($d_{it} = (0, 0)$) is chosen as a reference alternative, and we normalize accordingly the corresponding flow utility to zero. The flow utility parameters therefore need to be interpreted relative to the home production alternative.

2.6 The optimization problem

Individuals are forward-looking, and choose the sequence of college enrollment and labor market participation decisions yielding the highest present value of expected lifetime utility. Formally, the corresponding value function writes, using the notations introduced above:

$$V = \max_{d_{it}} E \left[\sum_{t=0}^T \beta^t \sum_{j \in \{a, bs, bn, 0\}} \sum_{k \in \{p, f, 0\}} u_{jkt} 1\{d_{it} = (j, k)\} \middle| \Omega_{t_0} \right]$$

Ω_{t_0} corresponds to the initial vector of state variables, which includes the age, accumulated work experience, AFQT, race, gender as well as the vector of idiosyncratic shocks, all measured at time $t = 0$. The expectation is taken with respect to the distribution of the future idiosyncratic shocks, keeping in mind that the flow utility functions u_{jkt} have already been integrated with respect to the prior ability distribution. The latter integration step is absent from most of the dynamic structural schooling choice literature, which typically makes the assumption that individuals have perfect information about their ability.⁷

The conditional value function is given by:

$$v_{jkt}(Z_{it}) = u_{jkt}(Z_{it}) + \beta E_t(V_{t+1}(Z_{t+1}) | Z_{it}, d_{it} = (j, k))$$

Given the assumption that the ε 's are i.i.d. Type 1 extreme value,

$$v_{jkt}(Z_{it}) = u_{jkt}(Z_{it}) + \beta E_t \left[\ln \left(\sum_j \sum_k \exp(v_{jkt+1}(Z_{it+1})) \right) \middle| Z_{it}, d_{it} = (j, k) \right] + \beta\gamma$$

where γ denotes Euler's constant.

⁷Notable exceptions include Arcidiacono (2004), Heckman & Urzua (2009), Stange (2012) and Stinebrickner & Stinebrickner (2013).

2.7 Finite dependence

Conditional on the one-period-ahead state variables besides ε , the future utility term can be expressed as:

$$\ln \left(\sum_j \sum_k \exp(v_{jkt+1}(Z_{it+1})) \right) = v_{j'k't+1}(Z_{it+1}) - \ln(p_{j'k't+1}(Z_{it+1}))$$

for any choice (j', k') , where $p_{j'k't+1}(Z_{it+1})$ is the conditional choice probability (CCP) of choosing $d_{it+1} = (j', k')$. Consider any choice (j', k') as well as the choice $(0, 0)$ (home). Given these initial choices, there exists a sequence choice such that, in expectation, individuals will be in the same state three periods ahead, implying:

$$\begin{aligned} E_t [V_{t+3}(Z_{it+3}) | d_{it} = (0, 0), d_{it+1} = (j', k'), d_{it+2} = (0, 0)] &= \\ E_t [V_{t+3}(Z_{it+3}) | d_{it} = (j', k'), d_{it+1} = (0, 0), d_{it+2} = (0, 0)] & \end{aligned}$$

Since in estimation we use differences of conditional value functions, we can reformulate the problem in terms of two-period ahead flow payoffs and conditional choice probabilities and then estimate the conditional choice probabilities in a first stage.

2.8 Unobserved heterogeneity

So far, we have assumed that the idiosyncratic shocks entering the GPA, log-wage and flow utility equations are mutually independent and serially uncorrelated. Although we control for unobserved ability through the (correlated) unobserved schooling ability and labor market productivity factors $A_i = (A_{ia}, A_{ibs}, A_{ibn}, A_{is}, A_{iu})'$, it is desirable to allow for a more general form of unobserved heterogeneity, which would in particular allow for some correlation between the unobserved preference terms. We deal with this issue using a finite mixture approach a la Heckman & Singer (1984), as is standard in the dynamic structural literature (see, e.g., Keane & Wolpin, 1997). Namely, we assume that there are R types of individuals in the population, and type-specific intercepts are included in the ability indexes, in the log-wages as well as in the flow utilities.

3 Identification and estimation

3.1 Identification

Before turning to the estimation procedure, let us now discuss the identification of the model.⁸ As is common for these types of dynamic discrete choice models (see, e.g., Rust,1994, and Magnac & Thesmar, 2002), identification of the flow utility parameters hinges on the distributional assumptions imposed on the idiosyncratic shocks, the normalization of the home production utility and the discount factor β , which is set equal to 0.9 in the following.

Let us now consider the identification of the outcome equations (grades and log-wages). The GPA $G_{ij\tau}$ is only observed for the individuals who are enrolled in a type- j (college, major) in their τ -th period of college enrollment. To the extent that college enrollment decisions depend on the ability (A_i), this raises a selection issue. We show the identification of the grade equation parameters by using, for each period τ , the prior ability at the beginning of the period ($E_{t-1}(A_{ij})$) as a control function in the grade equation (see Navarro, 2008, for an excellent review of the control function approach). Specifically, we consider the following augmented regression for $j \in bs, bn$ and $\tau > 2$:

$$G_{ij\tau} = \lambda_{0j} + \lambda_{1j}(\gamma_{0j} + X_{ict}\gamma_{1j}) + \lambda_{1j}E_{t-1}(A_{ij}) + \nu_{ij\tau}$$

where it follows from the bayesian updating rule (see Equation (2.1), p.8) that $E_{t-1}(A_{ij})$ can be expressed as a weighted sum of all the past ability signals. Under the key assumption, consistent with the specification of the flow utilities in Subsection 2.5, that college enrollment decisions only depend on ability through the ability beliefs, application of ordinary least squares to this equation identifies the parameters $(\lambda_{0j}, \lambda_{1j})$, with the ability index coefficients $(\gamma_{0j}, \gamma_{1j})$ being identified from the first and second period grades.

Identification of the ability index coefficients also follows from the assumption that enrollment decisions only depend on ability through the past ability signals. Specifically, grades in the first two years of four-year college as well as in two-year colleges can be

⁸For the sake of exposition, we first consider the case of the model without type-specific unobserved heterogeneity, before discussing the identification of the unobserved heterogeneity parameters.

expressed as follows:

$$G_{ij\tau} = \gamma_{0j} + X_{ict}\gamma_{1j} + E_{t-1}(A_{ij}) + \nu_{ij\tau}$$

Application of ordinary least squares therefore directly identifies $(\gamma_{0j}, \gamma_{1j})$. Similar arguments can be used for the identification of the log-wage equations in each sector.

Finally, the signal-to-noise ratios as well as the ability covariance matrix Δ are identified from the past ability signal coefficients. Of particular interest here are the correlations between the different ability components, which are identified from individuals switching occupations.

In our specification with R heterogeneity types, we also need to tell apart the type-specific unobserved (to the econometrician only) persistent heterogeneity components from the ability beliefs. For instance, low AFQT individuals who choose to enroll in a four-year college right after high school graduation should have a high unobserved preference for four-year college. Furthermore, individuals with low AFQT who are enrolled in college may decide to leave college after getting a high GPA. It follows from this type of behavior that these individuals would have a high type-specific unobserved schooling ability.

3.2 Estimation procedure

Let us first detail the estimation procedure for the specification without type-specific unobserved heterogeneity. Assuming that the idiosyncratic shocks are mutually and serially uncorrelated, estimation proceeds in two stages, which consists of (i) estimation of the grade and log-wage equations and (ii) estimation of the dynamic discrete choice model. The validity of this sequential approach follows from the key assumption that individual choices only depend on their ability through the (observed) sequence of signals. This results in the likelihood being separable in the outcome and choice contributions.

Specifically, let us consider the case of an individual attending college during T_c periods, who participate to the unskilled (resp. skilled) labor market during T_u (resp. T_s) periods and for whom we observe a sequence of T_d decisions.⁹ We write the individual contributions to the likelihood of the grades, log-wages and choices by integrating

⁹Individual subscripts are omitted in the following for notational convenience.

out the unobserved ability terms $A = (A_a, A_{bs}, A_{bn}, A_s, A_u)'$, which breaks down the dependence across the grades, log-wages, choices and between all of these variables. The contribution to the likelihood writes, denoting by G the grades, w_u (resp. w_s) the unskilled (resp. skilled) log-wages and d the decisions, as a four-dimensional integral:

$$\begin{aligned} & l(d_1, \dots, d_{T_d}, G_1, \dots, G_{T_c}, w_{u1}, \dots, w_{uT_u}, w_{s1}, \dots, w_{sT_s}) \\ &= \int \int \int \int \int l(d_1, \dots, d_{T_d}, G_1, \dots, G_{T_c}, w_{u1}, \dots, w_{uT_u}, w_{s1}, \dots, w_{sT_s} | A) l(A) dA \end{aligned}$$

where $l(A)$ is the Gaussian pdf of the ability distribution $\mathcal{N}(0, \Delta)$.

From the law of successive conditioning, and using the fact that choices depend on A only through the signals, we obtain the following partially separable expression:

$$l(d_1, \dots, d_{T_d}, G_1, \dots, G_{T_c}, w_{u1}, \dots, w_{uT_u}, w_{s1}, \dots, w_{sT_s}) = L_d \times L_{G, w_u, w_s}$$

Where the contribution of the sequence of decisions is given by:

$$L_d = l(d_1) l(d_2 | d_1, G_1) \dots l(d_{T_d} | d_1, d_2, \dots, d_{T_d-1}, G_1, G_2, \dots, w_{u1}, w_{u2}, \dots, w_{s1}, w_{s2}, \dots)$$

This simply corresponds to the product over T_d periods of the type-1 extreme value choice probabilities obtained from the dynamic discrete choice model.

The contribution of the observed sequence of grades, unskilled and skilled log-wages is given by:

$$\begin{aligned} L_{G, w_u, w_s} &= \int \int \int \int \int l(G_1 | d_1, A) \dots l(G_{T_c} | d_1, d_2, \dots, A) l(w_{u1} | d_1, A) \dots l(w_{uT_u} | d_1, d_2, \dots, A) \\ &\quad \times l(w_{s1} | d_1, A) \dots l(w_{sT_s} | d_1, d_2, \dots, A) l(A) dA \end{aligned}$$

Where $l(w_{ut} | d_1, \dots, A)$, $l(w_{st} | d_1, \dots, A)$, and $l(G_t | A)$ are Gaussian pdf of respectively, the unskilled and skilled log-wage distribution, and the GPA.

The estimation of the outcome equations proceeds as follows. Instead of maximizing directly the likelihood of the outcomes, which would be computationally costly because of the ability integration, we compute the parameter estimates using an EM algorithm (Dempster et al., 1977). The estimation procedure iterates over the following two steps:

- E-step: compute the posterior distribution of A from the outcome data (wages and grades), using the outcome equation parameters obtained from the previous iteration.

- M-step: given the posterior ability distribution obtained at the E-step, maximize the expected complete log-likelihood of the outcome data, which is separable in each occupation (two-year college, four-year college Science major, four-year college non-Science major, skilled or unskilled labor).

Namely, at each iteration k of the EM estimation, denoting by $l_k(A)$ the posterior ability distribution computed at the E-step, we maximize the expected complete log-likelihood El_k :

$$\begin{aligned} El_k &= \int \int \int \int \int \ln(l(G_1|d_1, A) \dots l(G_{T_c}|d_1, d_2, \dots, A)l(w_{u1}|d_1, A) \dots l(w_{uT_u}|d_1, d_2, \dots, A))l_k(A)dA \\ &= El_{k,a} + El_{k,bs} + El_{k,bn} + El_{k,s} + El_{k,u} \end{aligned}$$

For instance, the contribution of unskilled wages ($El_{k,u}$) writes, denoting by $l_k(A_u)$ the marginal posterior distribution of A_u :

$$El_{k,u} = \int (\ln(l(w_{u1}|d_1, A_u)) + \dots + \ln(l(w_{uT_u}|d_1, d_2, \dots, A_u)))l_k(A_u)dA_u$$

The dynamic discrete choice component of the model is then estimated using a CCP approach. The key idea is to use the existing mapping, under the Type-I extreme value assumption on the idiosyncratic shocks, between the differences in future value functions and the CCPs (Hotz & Miller, 1993). In practice, estimation of the flow utility parameters involves the following steps:¹⁰

1. Estimation of the CCPs. Notice that the vector of prior ability $E_{t-1}(A)$ and prior variances $\Sigma_{t-1}(A)$ is a *predicted* state variable in this setting, which have been estimated from the previous stage. In practice we estimate twelve flexible binary logit models (one for each alternative) to predict these CCPs.¹¹
2. Estimation of the flow utility parameters after expressing the future value function as a function of the CCPs. Having estimated the CCPs in a first step, this simply amounts to estimating (*via* maximum likelihood) a multinomial logit, with a predicted offset term.

¹⁰Standard errors are estimated using bootstrap with 200 replications.

¹¹The CCPs are identified from the data and could in principle be estimated nonparametrically. However, we choose to estimate them using a logit specification to avoid the curse of dimensionality.

Accounting for type-specific persistent unobserved heterogeneity breaks down the separability between the choice and outcome components of the likelihood described above. However, following Arcidiacono & Miller (2011), we can use here also an adaptation of the EM algorithm to reinstate the separability between the two stages at the maximization step, allowing in particular to update the choice parameters separately from the wages, grades and learning parameters. This results in an iterative estimation procedure, in which the flow utility parameters are estimated after conditioning on the heterogeneity type *and* the CCPs, which are now also unobserved to the econometrician. Relative to full solution methods, this estimation procedure yields very substantial computational savings, and only uses the CCPs two periods ahead. Thanks to the latter feature, our estimates do not hinge on any behavioral assumptions of the model far into the future.

4 Data

The model is estimated using data from the National Longitudinal Survey of Youth 1997 (NLSY97). The NLSY97 is a longitudinal, nationally representative survey of 8,984 American youth who were born between January 1, 1980 and December 31, 1984. Respondents were first interviewed in 1997 and have continued to be interviewed annually (for a total of 14 Rounds as of 2010, which is the most recent data available at the time of the current version of the paper) on such topics as labor force activities, education, and marriage and fertility, among many others.

Of particular importance for our analysis is the choice variable, d_t , which is constructed at each period as follows:

1. Any individual attending a college in the month of October is classified as being in college for this year.
2. Any individual reporting college attendance who also reports working at least one week in October and at least 10 hours per week is classified as working part-time while in school, with full-time work requiring at least 35 hours per week and four weeks worked in October.
3. Any individual not in college (according to the criterion above) is classified as

working part-time or full-time according to the criteria above.¹²

4. Finally, all other cases are classified as home production.¹³

Tables 1 through 11 below present some descriptives for our subsample, by college enrollment, major and completion status (as of Round 14 of the survey). Table 1 shows that individuals who attend college at some point and start at a four-year institution have, on average, higher AFQT scores, with science majors having higher scores than other majors.¹⁴ The proportion of blacks and Hispanics is also lower among four-year college attendees, with white males disproportionately choosing science majors. Conversely, it is worth noting that those starting at a two-year college tend to have a lower AFQT score, and disproportionately come from minorities. Overall, this difference in composition between two and four-year colleges (and between majors in four-year colleges) stresses the need to distinguish between college and major type when modeling college enrollment decisions.

Table 2 reports the mean GPA (on a scale between 0 and 4) by type of college attended, major and period of enrollment.¹⁵ Looking at the individuals enrolled in a four-year college with a science major, the evolution of the GPA provides clear evidence of selection over time. Individuals who leave college or switch to a two-year institution or other type of major have lower GPA than those who stay enrolled in a four-year college science major. We find a similar pattern for two-year college enrollees, with the GPA being on average lower for these students than for either type of four-year college enrollees. Overall, these descriptive findings are consistent with two stories, which may not be mutually exclusive: (*i*) individuals decide to leave or switch college/major as they learn about their schooling ability, or (*ii*) those who leave or switch college/major tend to have a lower ability, that they observe perfectly even before starting college. Telling apart these two explanations is a key objective of our structural estimations, which will be discussed in the following section.

¹²These criteria for labor force participation resemble those of Keane and Wolpin (1997).

¹³Following this criterion, any individual who is unemployed during the whole year is classified in the home production sector. Our preliminary results do not appear to be sensitive to the inclusion of the unemployment status in the work, rather than in the home production alternative.

¹⁴AFQT is standardized to be zero-mean and standard deviation 1 for our estimation subsample.

¹⁵For instance, period 2 corresponds to the second year of college enrollment. These periods of college enrollment may not be consecutive.

Table 3 lists the frequencies of continuous enrollment until graduation (either in four or two-year colleges), stopping-out (i.e. leaving college before graduation and returning to school at some point) and dropping-out (i.e. permanently leaving college, before graduation) in the NLSY97 full sample, our estimation subsample, and type of college/major first enrolled in. Our subsample slightly understates dropping-out because we discarded observations in right-censored missing interview spells, and missing an interview is positively correlated with dropping-out of college. Also evident from Table 3 is the fact that dropping-out and stopping-out are more common in two-year colleges relative to four-year colleges. Four-year science majors have the lowest proportions of dropping-out and stopping-out. This again points at the need to distinguish between these two types of college/major in our model. Due to the ongoing nature of the survey and the fact that some respondents are still in college, Table 4 aims to identify the lower bound of the stopout rate. For example, of those who had graduated with a four-year college degree by round 14 of the survey, 12.3% were stopouts. For those beginning college in a four-year university science major, this number is 5.3%, compared with 10.9% for humanities majors. For those originating in a two-year college, the figure is 29.2%.

Table 5 shows that those who continuously complete college (CC) have higher AFQT scores, higher high school GPA, and come from families with higher income and mothers who are more educated. It is also interesting to note that stopouts on average straddle the continuous completion and dropout categories. This highlights the importance of studying stopping out as a third category of college completion. The descriptive evidence presented in Table 5 also points to the fact that family background variables are important to include in an analysis of college completion.

Table 6 breaks out Table 2 by college completion status. Similar to Table 2, there is evidence of selection over time and over (eventual) completion status. This further supports the idea that those who leave college may do so because of a bad signal on ability in the form of low grades. To illustrate this more fully, Table 7 expounds this point by presenting differences between actual and expected grades in the first period of college (when dropout rates are highest), where expected grades are taken as a function of race, gender, AFQT, work status, and age. Interestingly, this shows that the selection patterns discussed above still hold after controlling for this set of observed characteristics.

Since we incorporate school and work combinations in the choice model, Tables 8 through 10 describe the evolution of GPA over time by work-in-school status. For both major types in four-year colleges, average GPA is roughly decreasing in work intensity, and increasing over time within each work intensity category. For two-year colleges, GPA is decreasing in work intensity only in the first period. By periods 3 and 4, the opposite is true. This illustrates a substitution effect between school and work intensity—those working hardest take longer than two years to complete a two-year degree.

Finally, to illustrate learning on wages as a reason for stopouts to return to college, Table 11 lists the difference between actual and expected log wages for those who have stopped out, broken out by next-period decision. Those who have left college for the labor force and choose to return to school have 8% lower wages on average the year before returning to school, even when conditioning on a rich set of individual, family background, and labor force experience characteristics. This is quantifiable evidence that learning on wages contributes to the decision of college dropouts to return to college and become stopouts.

Table 1: AFQT, gender and race, broken down by college enrollment status

Variable	Full Sample			College, start in two-year		
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
AFQT	4238	0.000	1.000	857	-0.017	0.898
male	4238	0.511	0.500	857	0.471	0.499
black	4238	0.253	0.435	857	0.219	0.414
hispanic	4238	0.185	0.389	857	0.233	0.423
	College, start in four-year sci			College, start in four-year hum		
AFQT	300	0.939	0.805	864	0.645	0.866
male	300	0.607	0.489	864	0.390	0.488
black	300	0.183	0.388	864	0.226	0.418
hispanic	300	0.107	0.309	864	0.102	0.303

Table 2: GPA over time by type of college attended

Period	Four-year college Sci		Four-year college Hum		Two-year college	
	Obs	Mean	Obs	Mean	Obs	Mean
year 1	300	3.049	864	3.046	857	3.017
year 2	254	3.140	724	3.117	545	3.040
year 3	202	3.255	694	3.188	293	3.138
year 4	182	3.297	644	3.242	119	3.066

Table 3: Outcomes of college enrollees

	Estimation			College Type	
	Full-Sample	Subsample	Two-Year	Four-Year Sci	Four-year Hum
Continuous college (CC)	41.10%	45.87%	29.15%	62.33%	56.48%
Stopped-out (SO)	23.81%	23.90%	31.75%	16.33%	18.87%
Dropped-out (DO)	35.10%	30.23%	39.10%	21.33%	24.65%
Total N	5,217	2,008	844	300	864

Note: College Type refers to the *first* type of college enrolled in.

Table 4: Outcomes of college enrollees, conditional on having graduated by Round 14 of NLSY97

	Started in any college		Started in 2yr college	
	N	%	N	%
CC	578	87.7	68	70.8
SO	81	12.3	28	29.2
Total	659	100	96	100

	Started in 4yr college & Sci		Started in 4yr college & Hum	
	N	%	N	%
CC	144	94.7	366	89.1
SO	8	5.3	45	10.9
Total	152	100	563	100

Table 5: Background characteristics of college enrollees

	AFQT	High School GPA	Mother with BA	1996 Family Income (\$1996)
CC	0.68	3.37	34.53%	68,847
SO	0.28	2.98	19.79%	52,580
DO	0.12	2.93	18.40%	50,427

Table 6: Average GPA over time by college completion status

Period	CC			DO/SO		
	Four-Year Sci	Four-Year Hum	Two-Year	Four-Year Sci	Four-Year Hum	Two-Year
1	3.17	3.13	3.21	2.85	2.93	2.94
2	3.23	3.16	3.17	2.94	3.05	2.97
3	3.31	3.22	3.21	3.10	3.12	3.10
4	3.34	3.29	3.18	3.16	3.13	3.02

5 Preliminary results

In this section, we discuss a set of (preliminary) results, which were obtained without type-specific unobserved heterogeneity. The parameter estimates for the grade equa-

Table 7: Difference between actual and expected period- t grades (by $t + 1$ period college decision)

	Grades residual	Std Dev	N	Mean diff T (p-val)
Leave 4-year college & science	-0.277	0.673	71	4.24
Stay in 4-year college (or switch)	0.019	0.562	1010	(0.00)
Leave 4-year college & humanities	-0.159	0.647	347	6.11
Stay in 4-year college (or switch)	0.018	0.495	3109	(0.00)
Leave 4-year college (any major)	-0.180	0.653	418	7.34
Stay in 4-year college (or switch)	0.018	0.512	4119	(0.00)
Leave 2-year college	-0.127	0.774	525	5.24
Stay in 2-year college (or switch)	0.048	0.597	1398	(0.00)

Table 8: Average GPA over time for 4-year college science students by work intensity

	4-year Sci			
	1	2	3	4
Work FT	2.99	3.17	3.13	3.27
Work PT	3.01	3.09	3.20	3.24
No Work	3.07	3.17	3.33	3.36
Total	3.05	3.14	3.25	3.30

Table 9: Average GPA over time for 4-year college humanities students by work intensity

	4-year Hum			
	1	2	3	4
Work FT	3.00	3.06	3.26	3.16
Work PT	3.03	3.13	3.16	3.25
No Work	3.06	3.12	3.19	3.28
Total	3.05	3.12	3.19	3.24

Table 10: Average GPA over time for 2-year college students by work intensity

	2-year			
	1	2	3	4
Work FT	3.03	3.09	3.21	3.16
Work PT	2.95	3.08	3.11	3.06
No Work	3.07	2.94	3.10	2.96
Total	3.02	3.04	3.14	3.07

Table 11: Difference between actual and expected wages in time t for stopouts (by $t + 1$ decision)

	Wage residual	Std Dev	N	Mean diff T (p-val)
Stay in work	0.029	0.471	4,719	-2.59
Return to school	-0.052	0.335	219	(0.01)
Total	0.025	0.466	4,938	

tions are presented in Table 12. At 4-year colleges, both males and Blacks have lower grades conditional on observables. This also holds true for blacks at 2-year colleges. While both grades in high school and AFQT score are significant predictors of grades at both schools, the latter is relatively important in 4-year colleges. The major dummies in the 4-year equation indicate that grades are higher in the humanities and lower in the sciences. Working generally seems to have little effect on grades, regardless of time spent working.

Estimates of the wage equations are given in Table 13. There are positive returns to high school grades and AFQT scores in both markets, with higher returns to the former in the skilled sector. Returns to experience, at least for men, are higher in the skilled sector. Interestingly, experience in the unskilled sector does not translate into higher labor market earnings in the skilled sector. Returns to schooling in the unskilled sector are positive. However, working while in school results in a substantial wage loss, particularly while at a four-year college.

Table 14 presents the covariance matrix for the unobserved abilities. Since we do not (currently) have controls for unobserved heterogeneity, these abilities are treated as

unknown to the individual. Unobserved 2 year and 4 year ability are highly correlated as are unobserved productivity in the skilled and unskilled sector. The covariances are much smaller across schooling and work ability, with the covariance between skilled productivity and 2-year ability being particularly small. Unobserved ability is also more important in the skilled sector as reflected in the higher variance of unobserved productivity there than in the unskilled sector.

Given the estimates of the learning portion of the model, we can measure sorting by unobserved ability. Table 15 shows the mean for each unobserved ability for different choice paths. Namely, we take individuals in the year 2007 and calculate their expected abilities and then average across those who chose a particular path.

Though sorting effects are small, the signs are generally in the appropriate direction. Those who go continuously to college are relatively high in both schooling abilities. Those who work while in college are not as strong on academic ability and this may be part of the reason they chose to work, having received a relatively weaker signal on their academic ability. While working, these individuals discover that their unobserved ability in the unskilled sector is low and hence remain in school. Those who stopout but then graduate have similar productivities in the unskilled sector to those who worked while in school. However, their academic ability is slightly lower.

The last two rows consider those who stopped out but then did not graduate and those for whom dropping out was an absorbing state. The academic abilities of these two groups are very similar. However, this is not true for the productivity in the unskilled sector. Namely, those who stopout but do not graduate have particularly low productivity in the unskilled sector. These individuals bounce back and forth between schooling and work and unfortunately find that they are not particularly productive in school or in the unskilled sector.

Finally, Table 16 reports the structural parameter estimates obtained from the maximum likelihood procedure described in Subsection 3.2. Notably, the results indicate that individuals with higher prior ability have a higher utility for college (relative to home production), with a significant coefficient for four-year colleges. Similarly, individuals with higher AFQT have a higher utility for college, as well as work, though again only significantly so for four-year colleges. The coefficients on the race and gender dummies are consistent with the well-documented college enrollment gaps between males and females and minorities and whites.

Table 12: Estimates of 2 and 4 year GPA Parameters

	4 year		2 year	
	Coeff.	Std. Error	Coeff.	Std. Error
Male	-0.121	(0.018)	-0.026	(0.017)
Black	-0.082	(0.041)	-0.166	(0.023)
Hispanic	-0.003	(0.024)	-0.067	(0.023)
AFQT	0.104	(0.009)	0.052	(0.011)
HS Grades	0.301	(0.025)	0.267	(0.014)
Work FT	-0.023	(0.029)	0.014	(0.019)
Work PT	-0.018	(0.038)	-0.064	(0.022)
Switch College Type	0.078	(0.026)	-0.027	(0.033)
Don't know major	0.002	(0.036)		
Humanities	0.048	(0.019)		
Science	-0.044	(0.004)		

Note: other variables included were mother's and father's education, year in school, and age.

6 Conclusion

This paper examines the determinants of college attrition, in a situation where individuals have imperfect information about their schooling ability and labor market productivity. Using longitudinal data from the NLSY97, we estimate a dynamic model of college attendance, major choice and work decisions. A key feature of our framework is to account for correlated learning about ability and productivity through college grades and wages. Preliminary evidence suggests learning about each of these components does influence decisions. In particular, those who have stopped out of college have learned that their academic abilities are relatively low but that their productivity in the unskilled sector is also low. In contrast, dropout perform poorly academically but stay out of school because their productivity in the unskilled sector is sufficiently high.

References

- Altonji, J. (1993), ‘The demand for and return to education when education outcomes are uncertain’, *Journal of Labor Economics* **11**, 48–83.
- Altonji, J., Blom, E. & Meghir, C. (2012), ‘Heterogeneity in human capital investments: High school curriculum, college majors, and careers’, *Annual Review of Economics* **4**, 185–223.
- Antonovics, K. & Golan, L. (2012), ‘Experimentation and job choice’, *Journal of Labor Economics* **30**, 333–366.
- Arcidiacono, P. (2004), ‘Ability sorting and the returns to college major’, *Journal of Econometrics* **121**, 343–375.
- Arcidiacono, P. & Miller, R. (2011), ‘Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity’, *Econometrica* **79**, 1823–1867.
- Becker, G. (1964), *Human Capital*, New York: Columbia University Press.
- Bound, J. & Turner, S. (2011), Dropouts and diplomas: The divergence in collegiate outcomes, in E. Hanushek, S. Machin & L. Woessmann, eds, ‘Handbook of the Economics of Education’, Vol. 4, Elsevier.
- DeGroot, M. (1970), *Optimal Statistical Decisions*, New York: McGraw Hill.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), ‘Maximum likelihood from incomplete data with the em algorithm’, *Journal of the Royal Statistical Society, Series B* **39**, 1–38.
- Hansen, K., Heckman, J. & Mullen, K. (2004), ‘The effect of schooling and ability on achievement test scores’, *Journal of Econometrics* **121**, 39–98.
- Heckman, J., Lochner, L. & Todd, P. (2006), Earnings functions, rates of return and treatment effects: the Mincer equation and beyond, in E. Hanushek & F. Welch, eds, ‘Handbook of the Economics of Education’, Vol. 1, Elsevier.

- Heckman, J. & Singer, B. (1984), ‘A method for minimizing the impact of distributional assumptions in econometric models for duration data’, *Econometrica* **52**, 271–320.
- Heckman, J. & Urzua, S. (2009), The option value of educational choices and the rate of return to educational choices. Unpublished.
- Hotz, V. & Miller, R. (1993), ‘Conditional choice probabilities and the estimation of dynamic models’, *Review of Economic Studies* **60**, 497–529.
- James, J. (2011), Ability matching and occupational choice. Unpublished.
- Joensen, J. (2009), Academic and labor market success: the impact of student employment, abilities and preferences. Unpublished.
- Keane, M. & Wolpin, K. (1997), ‘The career decisions of young men’, *The Journal of Political Economy* **105**, 473–522.
- Light, A. & Strayer, W. (2000), ‘The demand for and return to education when education outcomes are uncertain’, *Journal of Human Resources* **35**, 299–332.
- Magnac, T. & Thesmar, D. (2002), ‘Identifying dynamic discrete decision processes’, *Econometrica* **70**, 801–816.
- Manski, C. F. (1989), ‘Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon’, *Economics of Education Review* **8**, 305–312.
- Manski, C. & Wise, D. (1983), *College Choice in America*, Harvard University Press.
- Miller, R. (1984), ‘Job matching and occupational choice’, *The Journal of Political Economy* **92**, 1086–1120.
- Navarro, S. (2008), Control functions, in S. Durlauf & L. Blume, eds, ‘The New Palgrave Dictionary of Economics’, London: Palgrave Macmillan Press.
- Pugatch, T. (2012), Bumpy rides: School to work transitions in south africa. IZA Discussion Paper No. 6305.
- Rust, J. (1994), Estimation of dynamic structural models, problems and prospects: Discrete decision processes, in C. Sims, ed., ‘Advances in Econometrics: Sixth World Congress’, Vol. 2, Cambridge University Press, New York, pp. 5–33.

- Sanders, C. (2010), Skill uncertainty, skill accumulation, and occupational choice. Unpublished.
- Stange, K. (2012), ‘An empirical investigation of the option value of college enrollment’, *American Economic Journal: Applied Economics* **4**, 49–84.
- Stinebrickner, T. & Stinebrickner, R. (2003), ‘Working during school and academic performance’, *Journal of Labor Economics* **21**, 473–491.
- Stinebrickner, T. & Stinebrickner, R. (2012), ‘Learning about academic ability and the college drop-out decision’, *Journal of Labor Economics* **30**, 707–748.
- Stinebrickner, T. & Stinebrickner, R. (2013), ‘Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model’, *Forthcoming in the Journal of Labor Economics* .
- Stratton, L., O’Toole, D. & Wetzel, J. (2008), ‘Ability sorting and the returns to college major’, *Economics of Education Review* **27**, 319–331.

Table 13: Estimates of Skilled and Unskilled Wage Parameters

	Skilled		Unskilled	
	Coeff.	Std. Error	Coeff.	Std. Error
Male	0.037	(0.018)	0.088	(0.006)
Black	-0.055	(0.016)	-0.050	(0.006)
Hispanic	0.077	(0.021)	0.009	(0.006)
AFQT	0.025	(0.008)	0.024	(0.003)
HS Grades	0.060	(0.011)	0.011	(0.004)
Unskilled Experience	0.011	(0.005)	0.042	(0.002)
Male×Unskilled Exp.	-0.009	(0.008)	0.020	(0.002)
Skilled Experience	0.030	(0.008)		
Male×Skilled Exp.	0.044	(0.011)		
PT	-0.054	(0.017)	-0.062	(0.007)
PT 2 year			-0.053	(0.015)
PT 4 year			-0.090	(0.011)
FT 2 year			-0.064	(0.011)
FT 4 year			-0.088	(0.010)
1 year college			0.030	(0.008)
2 years college			0.058	(0.009)
3 years college			0.060	(0.011)
4 years college			0.104	(0.010)

Note: other variables included were mother's and father's education, year dummies, and age.

Table 14: Covariance Matrix for Unobserved Abilities

	2 year	4 year	Unskilled	Skilled
<i>Covariance matrix</i>				
2 year	0.120	0.079	0.013	0.001
4 year	0.079	0.122	0.015	0.011
Unskilled	0.013	0.015	0.063	0.054
Skilled	0.001	0.011	0.054	0.108
<i>Correlation coefficients</i>				
2 year	1.000	0.653	0.150	0.008
4 year	0.653	1.000	0.170	0.095
Unskilled	0.150	0.170	1.000	0.656
Skilled	0.008	0.095	0.656	1.000

Table 15: Average Posterior Ability in 2007 for Different Choice Paths

Choice Path	2 year	4 year	Unskilled	Skilled
Continuous college, no work	0.045	0.071	0.016	0.021
Continuous college, work	0.008	0.013	-0.023	-0.024
Stopout, graduate	-0.012	-0.015	-0.021	-0.026
Stopout, don't graduate	-0.038	-0.060	-0.044	-0.037
Dropout	-0.040	-0.054	-0.007	-0.004

Table 16: Flow Utility Estimates

	2 year	4 year	Work PT	Work FT
AFQT	1.290	6.023	0.636	0.279
Male	-0.108	-2.298	0.475	-0.154
black	-0.035	1.759	-1.369	-1.173
Hispanic	-0.451	-1.085	-0.924	-0.987
Ability	0.719	5.779		
E[ln wage]			-0.475	1.373
Experience			0.998	0.793
Previous 2 year	2.834	1.471	0.071	-0.425
Previous 4 year	1.115	3.843	0.160	-0.440
Previous Part-time	-0.114	-0.146	1.315	1.457
Previous Full-time	-0.451	-0.627	1.561	2.627

Note: Also included work part-time, full-time, age, and year dummies in the college option, and dummies for college grad and college grad times male in the work options. Bold indicates statistically significant at the 95% level.