College attrition and the dynamics of information revelation *

Preliminary and Incomplete

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

November 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, a dynamic structural model of schooling and work decisions, where high school graduates choose a bundle of education and work combinations. We take into account the heterogeneity in schooling investments by distinguishing between two-, four-year colleges and graduate school, as well as science and non-science majors for four-year colleges. Individuals may also choose whether to work full-time, part-time, or not at all. 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. We use our results to quantify the importance of informational frictions in explaining the observed school-to-work transitions and to examine sorting patterns.

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, the 2012 North American Summer Meeting of the Econometric Society (Northwestern), as well as seminar participants at Arizona State, Carnegie Mellon, CREST, University of Colorado-Boulder, University of North Carolina at Chapel Hill, University of Pennsylvania and Yale University for helpful 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, IZA and NBER. E-mail address: apm16@duke.edu

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

1 Introduction

About half of students entering college in the United States do not earn a bachelor's degree within five years, a proportion that has been stable since the 1970's (Bound & Turner 2011). Given that there is a large wage premium to completing a bachelor's degree, this suggests that the arrival of new information about costs and benefits of schooling is likely to be an important driving force behind this pattern. In this paper, we focus on the role of learning about academic ability and labor market productivity as an explanation for the observed rate of 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 high college attrition rates are considered a major issue, doing so is important to understand (*i*) whether these attrition rates *should* be a concern, and (*ii*) which type of policies would be effective in reducing attrition rates.

In order to quantify the importance of information frictions in the decision to leave or return to college, we estimate a dynamic model of schooling and work decisions where such decisions depend on the arrival of information on schooling ability and work productivity. A key feature of the model is that students have imperfect information about their ability and productivity. After graduating from high school, individuals decide among some combination of postsecondary schooling and work options. When entering college, individuals have some beliefs about their ability and productivity. At the end of each school year, they learn about their ability, using 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 beliefs about their labor market productivity. Likewise, in the same vein as Miller (1984), employed individuals update both their productivity and ability beliefs after receiving a wage.

We estimate a richer model than previously possible by making use of recent innovations in the computation of dynamic models of correlated learning. James (2011) shows that (*i*) integrating out over actual abilities as opposed to the signals and (*ii*) 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. Importantly, it also allows us to take into account heterogeneity in schooling investments by distinguishing between two-, four-year colleges and graduate school, as well as Science and non-Science majors for four-year colleges.[1]

We then use the estimates of our model 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. Our results suggest that learning about schooling ability and labor market productivity plays an important role in the decision to dropout or re-enter college. 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, dropouts perform poorly academically and stay out of school because their productivity in the unskilled sector is sufficiently high.

Our analysis builds on seminal research by Manski & Wise (1983) and Manski (1989), which 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 dropout lies in the fact that, after entering college, students get new information and thus learn about their ability. Altonji (1993) shows that that this experimentation also applies to college major, a key determinant of future wage benefits. 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., Light & Strayer 2000, Arcidiacono 2004, Stratton et al. 2008, and Stange 2012). 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.

Much of the learning literature assumes that the labor market is an absorbing state, implying that the decision to leave college is irreversible. By relaxing this assumption, we are able to predict the substantial college re-entry rates of over 36% 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

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

[2]Note that because of right-censoring, this underestimates the actual re-entry rate.

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.

The remainder of the paper is organized as follows. Section 2 presents the data. Section 3 describes 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 4 informally discusses the identification of the model, with Section 5 detailing the estimation procedure. Section 6 presents our preliminary estimation results. Section 7 concludes.

2 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 corresponds to the most recent data used in 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 (either in a two- or a four-year college). For four-year colleges, our definition of "Science" majors includes majors in Sciences, Technology, Engineering, Mathematics or Economics.

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.[3]

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

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

The other dependent variables in the analysis are college GPA and wages. College GPA is measured on a four-point scale and calculated as the average GPA across all semesters in the calendar year. Wages are calculated as follows:

1. We compute the median of hourly wages across weeks in October, converted to 1996 dollars.

2. If a person worked multiple jobs in a certain week, we use the hourly earnings for the job with the most hours.

3. If a person does not report earnings for jobs worked in October, we impute earnings as annual income divided by annual hours worked (4.7% of our sample).

4. Finally, we top- and bottom-code the resulting earnings distribution at the 99.5 percentile and 2.5 percentile.

It is worth noting that GPA are missing quite frequently in our data (19% of college GPAs are missing in our sample, disproportionately in the first two years of college enrollment). We address this issue by imputing the GPAs which are missing, for the first time, in the first two periods of college. Specifically, we impute those missing observations as a function of the AFQT, high school GPA, major (for four-year college) and previous GPA interacted with type of college attended in previous year, separately for each type of college and college enrollment status in the following period.[5] We further drop all current and future observations for any respondents missing wage observations while choosing a work activity (respectively missing GPA after the first two periods of college enrollment).[6]

Tables 1 through 10 below present some descriptives for our subsample, by college enrollment, major and completion status. Table 1 shows that individuals who attend college at some point and start at a four-year institution have, on average, higher AFQT

[4]Following this criterion, any individual who is unemployed in October is classified in the home production sector. Our results do not appear to be sensitive to the inclusion of the unemployment status in the work, rather than in the home production alternative.

[5]The validity of this method relies on the assumption that GPA is missing at random, once conditioning on the aforementioned set of characteristics.

[6]For further details on our sample selection, see Table 1 in Appendix section A.

scores, with science majors having higher scores than other majors.[7] 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.[8] 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.

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 colleges and majors in our model. Due to the ongoing nature of the

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

[8] Note that these periods of college enrollment may not be consecutive.

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 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.

Table 9 further describes 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 10 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 provides suggestive evidence that learning on wages contributes to the decision to return to college.

3 The model

3.1 Overview

After graduating from high school, individuals in each period make a joint schooling and work decision. For those who have not graduated from a four-year college, their schooling options include whether to attend a two-year, a four-year institution as a science major, or a four-year institution as a non-science major. After graduating from a four-year college, the schooling option includes whether or not to enroll in graduate school. individuals choose whether to enroll in graduate school, and whether to work part-time or full-time.

In addition to choosing among the different schooling options, individuals also choose whether to work full-time, part-time, or not at all. All three of these decisions are available regardless of their schooling choice.[9] 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 incorporates this tradeoff.

Individuals only have imperfect information about (i) their schooling ability and (ii) their labor market productivity. 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 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 beliefs about their labor market

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

productivity as well as their beliefs about their schooling ability after receiving their wage.

Individuals are forward-looking and choose the sequence of actions yielding the highest value of expected lifetime utility. Hence, when making their schooling and labor market decisions, individuals take into account the option value associated with the new information acquired on different choice paths. 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 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.

3.2 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 $\mathcal{N}(0, \sigma^2_{Aj})$. Grades also depend on a set of covariates, X_{ict}, that are known to the individual and include observed ability measures and past decisions.

Denoting by t calendar time and τ the period of college enrollment, grades in two-year colleges and in the first two years of four-year colleges are given by:

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

The idiosyncratic shocks $\varepsilon_{ij\tau}$ are distributed $\mathcal{N}(0, \sigma^2_{j\tau})$ 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:[10]

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

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

3.3 A two-sector labor market

Individuals who choose one of the work options (either full-time or part-time) receives an hourly wage that 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 and individuals with a graduate school degree) 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).

Wages in sector l depend on productivity A_{il}, a set of observed characteristics X_{ilt}, time dummies δ_{lt}, and idiosyncratic shocks ε_{st}:

$$\ln(W_{ist}) = \delta_{lt} + X_{ilt}\gamma_{1l} + A_{il} + \varepsilon_{ilt}$$

We account for nonstationarity in wages by including calendar year dummies, δ_{lt}, thus incorporating business cycle effects. The time dummies at t are observed in period t but individuals must form expectations over this variable for periods $t+1$ and beyond. The idiosyncratic shocks, ε_{lt}, are assumed to be distributed $\mathcal{N}(0, \sigma_l^2)$ and are independent over time and independent of the other state variables.

3.4 Flow utilities

We denote in the following by $d_{it} = (j, k)$ the choice for individual i at time t over school, where $j \in \{a, bs, bn, 0\}$ (respectively $j \in \{gs, 0\}$) before (resp. after) graduation

[10] See Arcidiacono (2004) for a similar ability index specification.

from a four-year college, and work $k \in \{p, f, 0\}$, where gs refers to graduate school, and p and f refer to part-time and full-time work. The choice $d_{it} = (0, 0)$ then indicates the home production option: no work and no school.

Up to an intercept term and an idiosyncratic preference shock, 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:

$$\begin{aligned} U_{jkt}(Z_{it}, \varepsilon_{ijk}) &= \alpha_{jk} + Z_{1it}\alpha_j + Z_{2it}\alpha_k + \varepsilon_{ijkt} && (3.1) \\ &= u_{jk}(Z_{it}) + \varepsilon_{ijkt} && (3.2) \end{aligned}$$

where Z_{it} includes characteristics such as AFQT, race, 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_{ijkt} are assumed to follow a (standard) Type-I extreme value distribution. Embedded in Z_{1it} are expected abilities in sector j, with the exception of graduate school for which the flow payoff depends on the expected abilities in both four-year college Science and non-Science majors. Embedded in Z_{2it} are expected log-wages in sector k.[11]

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 this alternative.

3.5 The optimization problem

Individuals are forward-looking, choosing the sequence of college enrollment and labor market participation decisions yielding the highest present value of expected lifetime utility. The individual chooses d_{it} to sequentially maximize the discounted sum of payoffs:

$$E\left[\sum_{t=1}^{T} \beta^{t-1} \sum_j \sum_k (u_{jkt}(Z_{it}) + \varepsilon_{ijkt}) 1\{d_{it} = (j, k)\}\right]$$

[11] Both of these covariates will vary given the choices. However, to conserve on notation we do not put jk subscripts on the Z's.

where $\beta \in (0,1)$ is the discount factor. The expectation is taken with respect to the distribution of the future idiosyncratic shocks as well as the signals associated with the different choice paths.

Let $V_t(Z_{it})$ denote the ex ante value function at the beginning of period t, the expected discounted sum of current and future payoffs just before ε_t is revealed. Denote the conditional value function v_{ijkt} as the value of 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.

3.6 Beliefs

Individuals are uncertain about i) their future preference shocks, ii) their schooling ability and labor market productivity, and iii) the evolution of the market shocks (the δ_{lt}'s). The first we have discussed as expectations over future preference shocks are encompassed in the ex ante value function. We next describe beliefs over abilities and productivities as well as the market shocks.

3.6.1 Beliefs over schooling ability and labor market productivity

We denote A_i as the five dimensional ability vector, $A_i \equiv (A_{ia}, A_{ibs}, A_{ibn}, A_{is}, A_{iu})'$ (simply referred to as *ability* in the following). Individuals update their beliefs in a 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.[12]

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

[12] See also Antonovics & Golan (2012), James (2011) and Sanders (2010) who estimate occupational choice models with correlated learning.

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:

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

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

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

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 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):

$$E_t(A_i) = (\Sigma_{t-1}^{-1}(A_i) + \Omega_{it})^{-1}(\Sigma_{t-1}^{-1}(A_i)E_{t-1}(A_i) + \Omega_{it}\widetilde{S}_{it}) \quad (3.3)$$

$$\Sigma_t(A_i) = (\Sigma_{t-1}^{-1}(A_i) + \Omega_{it})^{-1} \quad (3.4)$$

where 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 (multiplied by λ_{1j}^2 for four-year colleges in junior and senior years). $\widetilde{S}_{it}$ 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. Individuals then integrate out over the possible signals they could receive for each possible decision.

3.6.2 Beliefs over market shocks

We also need to specify how individuals form beliefs about the market. Individuals observe the current values of δ_{st} and δ_{ut}. Their beliefs are the process governing the δ_t's is an AR1:

$$\delta_{lt+1} = \phi_{0l} + \phi_1 \delta_{lt-1} + \zeta_{lt} \tag{3.5}$$

We assume that ζ_{lt} is distributed N(0,σ_ζ) and that individuals believe it is independent over time and across sectors. Given the realizations of the δ_{lt}'s, individuals then integrate over possible realizations of the ζ_{lt}'s when forming their expectations over the future.

4 Identification

Before turning to the estimation procedure, we discuss below the identification of the model.[13] As is common for these types of dynamic discrete choice models (see, e.g., Rust,1994, Magnac & Thesmar, 2002, and Arcidiacono & Miller, 2013), 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 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 insightful 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

[13] For 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.

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 expressed as follows:

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

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 latent heterogeneity types, we also need to tell apart the type-specific unobserved (to the econometrician only) 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 will be predicted to have a high type-specific unobserved schooling ability.

5 Estimation

We 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 flow utility parameters. The validity of this sequential approach follows from the key assumption that choices

only depend on ability through the (observed) sequence of signals. This results in the likelihood being separable in the outcome and choice contributions.

5.1 Additive separability

Specifically, we consider the case of an individual i 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. We write the individual contributions to the likelihood of the grades, log-wages and choices by integrating 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_i the grades, w_{iu} (resp. w_{is}) the unskilled (resp. skilled) log-wages and d_i the decisions, as a five-dimensional integral:

$$\begin{aligned} & l(d_{i1}, \ldots, d_{iT_d}, G_{i1}, \ldots, G_{iT_c}, w_{iu1}, \ldots, w_{iuT_u}, w_{is1}, \ldots, w_{isT_s}) \\ = & \int l(d_{i1}, \ldots, d_{iT_d}, G_{i1}, \ldots, G_{iT_c}, w_{iu1}, \ldots, w_{iuT_u}, w_{is1}, \ldots, w_{isT_s}|A) l(A) dA \end{aligned}$$

where $l(A)$ is the 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:

$$l(d_{i1}, \ldots, d_{iT_d}, G_{i1}, \ldots, G_{iT_c}, w_{iu1}, \ldots, w_{iuT_u}, w_{is1}, \ldots, w_{isT_s}) = L_{d_i} \times L_{G_i, w_{iu}, w_{is}}$$

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

$$L_{di} = l(d_{i1}) l(d_{i2}|d_{i1}, G_{i1}) \ldots l(d_{iT_d}|d_{i1}, d_{i2}, \ldots, d_{iT_d-1}, G_{i1}, G_{i2}, \ldots w_{iu1}, w_{iu2}, \ldots, w_{is1}, w_{is2}, \ldots)$$

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:

$$\begin{aligned} L_{G_i, w_{iu}, w_{is}} = & \int l(G_{i1}|d_{i1}, A) \ldots l(G_{iT_c}|d_{i1}, d_{i2}, \ldots, A) l(w_{iu1}|d_{i1}, A) \ldots l(w_{iuT_u}|d_{i1}, d_{i2}, \ldots, A) \\ & \times l(w_{is1}|d_{i1}, A) \ldots l(w_{isT_s}|d_{i1}, d_{i2}, \ldots, A) l(A) dA \end{aligned}$$

Where $l(w_{iut}|d_{i1}, \ldots, A)$, $l(w_{ist}|d_{i1}, \ldots, A)$, and $l(G_{it}|A)$ are Gaussian pdf of respectively, the unskilled, skilled log-wage and GPA distributions.

5.2 Estimation of grade and wage parameters

Estimation of the parameters 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 the EM algorithm (Dempster et al., 1977). The estimation procedure iterates over the following two steps, until convergence:

- E-step: update the posterior ability distribution from all the observed outcome data (log-wages and grades), using the outcome equation parameters obtained from the previous iteration. This follows from the bayesian updating formulas (3.3)-(3.4), for the posterior ability mean and covariance, given in Section 3.6.1. The (population) variance of the ability distribution is then updated as follows, for each iteration k of the EM estimation:

$$\Delta^k = \frac{1}{N}\sum_{i=1}^{N}\left(\Sigma_i^k(A) + E_i^k(A)E_i^k(A)'\right)$$

  where N denotes the number of individuals in the sample, $E_i^k(A)$ the posterior ability mean ($E_i^k(A)'$ its transposed) and $\Sigma_i^k(A)$ the posterior ability covariance computed at the beginning of the E-step.

- 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 across sectors (two-year college, four-year college Science major, four-year college non-Science major, skilled or unskilled labor).

Namely, at the M-step of each iteration k of the EM estimation, denoting by $l_{ik}(A)$ the posterior ability distribution computed at the E-step, we maximize the expected complete log-likelihood El_{ik}:

$$\begin{aligned} El_{ik} &= \int \ln(l(G_{i1}|d_{i1},A)\dots l(G_{iT_c}|d_{i1},d_{i2},\dots,A)l(w_{iu1}|d_{i1},A)\dots l(w_{uiT_u}|d_{i1},d_{i2},\dots,A))l_{ik}(A)dA \\ &= El_{ik,a} + El_{ik,bs} + El_{ik,bn} + El_{ik,s} + El_{ik,u} \end{aligned}$$

For instance, the parameters of the unskilled wage equation are updated by maximizing the contribution $El_{ik,u}$, which writes, denoting by $l_{ik}(A_u)$ the marginal posterior distribution of A_u:

$$El_{ik,u} = \int (\ln(l(w_{iu1}|d_{i1},A_u)) + \dots + \ln(l(w_{iuT_u}|d_{i1},d_{i2},\dots,A_u)))l_{ik}(A_u)dA_u$$

Note that this term is additively separable over time. For any given period τ of unskilled labor market participation, it follows from the normality assumptions on the idiosyncratic productivity shocks and the unobserved ability that:

$$\int \ln(l(w_{iu\tau}|d_{i1}, d_{i2}, \ldots, A_u))l_{ik}(A_u)dA_u \quad = \\ -\frac{1}{2}\ln(2\pi\sigma_u^2) - \frac{1}{2\sigma_u^2}\left(\Sigma_{iuu}^k(A) + (w_{iu\tau} - X_{iut}\gamma_{1u} - \delta_{ut} - E_{iu}^k(A))^2\right)$$

where t refers to calendar time (which should be understood as individual-specific here), $E_{iu}^k(A)$ and $\Sigma_{iuu}^k(A)$ denote respectively the posterior mean and variance of the ability in the unskilled sector (computed at the E-step). This equality implies that the wage equation parameters $(\gamma_{1u}, \delta_{ut})$ can be simply updated by regressing (via OLS) the log-wages in the unskilled sector on the set of observed characteristics, calendar time dummies, and the posterior (unskilled) ability mean which plays the role of a selection correction term. Besides, the idiosyncratic shock variance (σ_u^2) is updated as follows:

$$\sigma_{u,k+1}^2 \quad = \quad \frac{\sum_{i,\tau}\left(\Sigma_{iuu}^k(A) + (w_{iu\tau} - X_{iut}\gamma_{1u} - \delta_{ut} - E_{iu}^k(A))^2\right)}{N_u^{obs}}$$

where N_u^{obs} is the total number of wage observations in the unskilled sector. Skilled wage equation parameters are updated similarly.

The updating rule above needs to be adjusted to account for the ability index specification of the grade equations along with the time-varying variances of the idiosyncratic shocks. For instance, for four-year colleges (period of enrollment $\tau > 2$), the contribution to the log-likelihood writes:

$$\int \ln(l(G_{ij\tau}|d_{i1}, d_{i2}, \ldots, A_j))l_{ik}(A_j)dA_j \quad = \\ -\frac{1}{2}\ln(2\pi\sigma_{j\tau}^2) - \frac{1}{2\sigma_{j\tau}^2}\left(\lambda_{1j}^2\Sigma_{ijj}^k(A) + (G_{ij\tau} - \lambda_{0j} - \lambda_{1j}AI_{ijt}^k)^2\right)$$

where $j \in \{bs, bn\}$, $\Sigma_{ijj}^k(A)$ denotes the posterior variance of the college-j ability (computed at the E-step), and $AI_{ijt}^k = \gamma_{0j} + X_{ict}\gamma_{1j} + E_{ij}^k(A)$ is the posterior mean of the ability index in college j. It follows that the parameters $(\gamma_{0j}, \gamma_{1j}, \lambda_{0j}, \lambda_{1j}, (\sigma_{j\tau}^2)_\tau)$ are updated by solving the following minimization problem:

$$\min \sum_{i,\tau}\left(\ln(\sigma_{j\tau}^2) + \frac{1}{\sigma_{j\tau}^2}\left(\lambda_{1j\tau}^2\Sigma_{ijj}^k(A) + (G_{ij\tau} - \lambda_{0j\tau} - \lambda_{1j\tau}AI_{ijt}^k)^2\right)\right)$$

where $(\lambda_{0j\tau}, \lambda_{1j\tau}) = (0, 1)$ for $\tau \leq 2$, and $(\lambda_{0j\tau}, \lambda_{1j\tau}) = (\lambda_{0j}, \lambda_{1j})$ otherwise.

5.3 Estimation of the flow payoffs

With the estimates of the grade, graduation, and wage transitions taken as given, we estimate the flow payoffs in a second state. Following Arcidiacono & Miller (2011), we express the future payoffs in such a way that avoids solving the full backwards recursion problem. Namely, the expected value function at time $t+1$ can be expressed relative to the conditional value function for one of the choices plus a function of the conditional choice probabilities. With the assumption that the preference shocks are distributed Type 1 extreme value, the expected value function can be expressed as:

$$E_t\left[V_{t+1}(Z_{it+1}|d_{it}=(j,k)\right] \;=\; E_t\left[v_{j'k't+1}(Z_{it+1}) - \ln(p_{j'k't+1}(Z_{it+1}))|d_{it}=(j,k)\right]$$

for any choice (j',k'), where $p_{j'k't+1}(Z_{it+1})$ is the conditional choice probability (CCP) of choosing $d_{it+1}=(j',k')$.

Recall that in estimation it is difference in the conditional value functions that are relevant, not the conditional value functions themselves. Consider any choice (j',k') as well as the choice $(0,0)$ (home). Given these initial choices, it is straightforward to show that there exists a sequence of choice such that, in expectation, individuals will be in the same state three periods ahead, namely:

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

We can then reformulate the problem in terms of two-period ahead flow payoffs and conditional choice probabilities and then estimate the conditional choice probabilities (CCPs) in a first stage. The differenced conditional value function is then:

$$v_{jkt}(Z_{it}) - v_{00t}(Z_{it}) = \begin{pmatrix} u_{jk}(Z_{it}) - \beta E_t\left(\ln\left[p_{00t+1}(Z_{t+1})\right]|Z_{it},d_{it}=(j,k)\right) \\ +\beta E_t\left(\ln\left[p_{jkt+1}(Z_{t+1})\right] - u_{jk}(Z_{t+1})|Z_{it},d_{it}=(0,0)\right) \\ +\beta^2 E_t\left(\ln\left[p_{00t+2}(Z_{t+2})\right]|Z_{it},d_{it}=(0,0),d_{it+1}=(j,k)\right) \\ -\beta^2 E_t\left(\ln\left[p_{00t+2}(Z_{t+2})\right]|Z_{it},d_{it}=(j,k),d_{it+1}=(0,0)\right)\end{pmatrix}$$

Estimation of the flow utility parameters then involves the following steps:

1. Estimate the CCPs via a flexible multinomial logit model.[14]

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

2. For, the expected differenced future value terms along the finite depends paths.

3. Estimate 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 a multinomial logit with an offset term.

5.4 Estimation with permanent unobserved heterogeneity

We account for permanent unobserved unobserved heterogeneity by assuming that individuals are one of R types where type is orthogonal to the covariates at $t = 1$. Accounting for type-specific unobserved heterogeneity breaks down the separability between the choice and outcome components of the likelihood described above as our full likelihood function is:

$$\sum_i \ln \left[\sum_{r=1}^{R} \pi_r (L_{d,i|r} L_{G,w_u,w_s,i|r}) \right] \tag{5.1}$$

But note that we were already using the EM algorithm in estimation which, as illustrated in Arcidiacono & Jones (2003), restores the additive separability of the likelihood function.

Following Arcidiacono & Miller (2011), we then use an adaptation of the EM algorithm where, rather than updating the structural parameters of the decision process at each step, we use their two stage approach and approximate the decision process with a reduced form. Let $L^*_{d,i|r}$ give the reduced form likelihood conditional on being of type r. The probability of i being the rth type follows from Bayes rule:

$$q_{ir} = \frac{\pi_r (L^*_{d,i|r} L_{G,w_u,w_s,i|r})}{\sum_{r'=1}^{R} \pi_{r'} (L^*_{d,i|r'} L_{G,w_u,w_s,i|r'})} \tag{5.2}$$

In the first stage we recover the parameters of the grade and wage processes, the (type-specific) CCPs, and the conditional probabilities of being each type.

The second stage boils down to a weighted multinomial logit with an offset term. This is identical to the case without unobserved heterogeneity except that now the q_{ir}'s are used as weights. 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.

5.5 Missing college majors

In our data, four-year college majors are missing at a fairly high rate. This is especially true for the first period of college enrollment, where close to 30% of majors are missing. We take this issue into account within our estimation procedure, by treating the unobserved major in the first year of college enrollment as another unobserved latent variable. The estimation procedure discussed above can be easily adjusted to allow for this additional latent variable.

Specifically, along with the type-specific unobserved heterogeneity distribution, the distribution of (unobserved) majors, conditional on each heterogeneity type, is going to be estimated within the first stage of our estimation procedure. The distribution of the unobserved majors is then taken as given in the second stage of the estimation, which still corresponds to a weighted multinomial logit where the weights are given by $Pr(Type, Major|data)$ instead of $Pr(Type|data)$, and the log-likelihood is conditional on both the heterogeneity type and the major.

6 Preliminary results

In this section, we discuss a set of preliminary results, which were obtained for the subsample of males. All the estimation results discussed below were obtained assuming the existence of $R = 2$ unobserved heterogeneity types. Type 1 (respectively Type 2) individuals account for 60.3% (resp. 39.7%) of the population.

6.1 Grade parameters

The parameter estimates for the grade equations are presented in Table 11. Conditional on observables, blacks are found to have lower GPA than whites across the board, in particular in 4-year colleges science majors as well as in 2-year colleges. While both grades in high school and AFQT score are significant predictors of grades at both schools and majors, it is worth noting that the former plays a major role, especially in 4-year colleges science majors. Working generally seems to have little effect on grades, regardless of time spent working. Interestingly, returns to ability are found to

be larger after junior year for humanities majors in 4-year colleges, a pattern which does not show up significantly for science majors. Finally, type-specific unobserved ability (known to the agent) plays an important role, with students in the Type 1 group having significantly lower GPA for both types of schools and majors.

6.2 Wage parameters

Estimates of the wage equations are given in Table 12. All else being equal, blacks have lower wages in both sectors, with a larger wage gap in the unskilled sector. The opposite is actually true for Hispanics in the skilled sector. Returns to experience 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 and significant, even though they are quantitatively fairly small. However, working while in school results in a substantial wage loss, particularly while at a four-year college. We also find positive returns to graduate schooling in the skilled sector (with the exception of the first year of graduate school). Finally, our results point to the existence of heterogeneity in productivity across the two unobserved types, with type 1 individuals being receiving significantly higher wages in both sectors.

6.3 Learning

Table 13 presents the correlation matrix for the unobserved abilities (initially unknown to the individual) in each sector, along with their variances. With the exception of 4-year Sciences and 2-year colleges, schooling ability is highly correlated across college types and majors. A similar picture emerges across skilled and unskilled sectors within the labor market, which are strongly correlated. Importantly, even though the correlations are smaller, schooling abilities are generally found to be positively correlated with labor market productivity in each sector. An exception to this is the negligibly small correlation between ability in 4-year humanities and productivity in the skilled sector. Finally, it is worth noting that the unobserved ability variance is larger for 4-year sciences and the skilled sector (even after rescaling by the variance of the corresponding

outcomes), which suggests that the role played by unobserved ability is relatively more important in those sectors.

Table 14 further shows that, even though our approach allows to account for both types of unobserved ability (known and unknown to the individuals), residual variation in log-wages and GPA remains sizeable. Our estimates also suggest that grades in 2-year colleges are noisier signals of ability than in 4-year colleges, with the precision of the signals increasing over time for all types of colleges and majors.

6.4 Sorting

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 final year of our data (2010), calculate their posterior abilities and then average across those who chose a particular path.

Though sorting effects are relatively small, the signs are generally in the expected direction. Those who go continuously to college and do not work while in school are relatively high in all schooling abilities, as well as in the skilled and unskilled labor market productivity. Those who work while enrolled in college have lower ability across the board 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 also have lower schooling ability than those continuously enrolled in college and not working, but they have a higher productivity in the unskilled sector than those continuously enrolled and working.

The last two rows consider those who stopped out but then did not graduate and those for whom dropping out was an absorbing sate. Individuals in these two groups have relatively low schooling ability for all college types and majors. It is particularly interesting to note that those who stopout but do not graduate have particularly low productivity in the unskilled sector, which is much lower than that of the dropouts. 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.

6.5 Flow payoffs

Finally, Table 16 reports the structural parameter estimates obtained from the procedure described in Subsection 5.3. Notably, the results indicate that individuals with higher prior ability have a higher utility for four-year college (relative to home production), with a larger coefficient for Science majors. Similarly, individuals with higher AFQT have a higher utility for four-year college, especially again for Science majors. The same holds true for high school grades, which are positively associated with the utility for all schooling options. Overall, this pattern is consistent with a cost of effort decreasing with these ability measures. Consistent with the existence of higher monetary costs of attending a four-year college (as opposed to a two-year college), individuals whose parents went to college also have a higher utility for four-year colleges, but not for two-year colleges. As expected, individuals with higher expected log-wages have a higher utility for work. Futhermore, the estimated coefficients on previous activities point to the existence of substantial switching costs across types of colleges and majors. Finally, type-1 individuals are found to have higher preferences for four-year colleges, even though the corresponding coefficients are quite small.

7 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. Estimation results show that a sizable fraction of the dispersion in college grades as well as log-wages is attributable to the ability components which are gradually revealed to the individuals as they accumulate more signals. These ability components are highly correlated across college types and majors, skilled and unskilled labor market, and we also find clear evidence of correlation between schooling ability and labor market productivity. Finally, we document the existence of sorting on ability and labor market productivity based on college enrollment and labor market partici-

pation decisions, suggesting that school-to-work transitions are partly driven by ability learning. 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. & Jones, J. (2003), 'Finite mixture distributions, sequential likelihood and the EM algorithm', *Econometrica* **71**, 933–946.

Arcidiacono, P. & Miller, R. (2011), 'Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity', *Econometrica* **79**, 1823–1867.

Arcidiacono, P. & Miller, R. (2013), Identifying dynamic discrete choice models off short panels. Unpublished.

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.

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), 'Determinants of college completion: School quality or student ability?', *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 processess, *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.

Stratton, L., O'Toole, D. & Wetzel, J. (2008), 'A multinomial logit model of college stopout and dropout behavior', *Economics of Education Review* **27**, 319–331.

A Data

This appendix section details the criteria we use to select our estimation subsample. Table 1 outlines each of these criteria.

Table 1: Sample Selection

Selection criterion	Resultant persons	Resultant person-years
Full NLSY97 sample	8,984	125,776
Drop females	4,599	64,386
Drop missing AFQT, HS grades or Parental education	3,408	47,712
Drop HS Dropouts (or those not receiving GED)	3,027	42,378
Drop observations before HS graduation	3,027	31,772
Drop right-censored missing interview spells	3,027	30,096
Drop any who attend college at a young age or graduate college in 2 or fewer years	3,023	29,957
Drop any who are not in HS at age 15 or under or have other outlying data	2,689	27,066
Drop those who don't report a 4-year college major after the first year	2,532	24,141
Drop observations after someone has a missing wage	2,395	20,903
Drop observations after someone has a missing grade in year 3+ of college	2,395	20,562
Drop anyone reporting college attendance after college graduation	2,395	20,560
Final estimation subsample	2,395	20,560

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

	Full Descriptive Sample			College, start in two-year		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
AFQT	6,105	0.000	1.000	1,882	-0.129	0.877
male	6,105	0.500	0.500	1,882	0.468	0.499
black	6,105	0.246	0.431	1,882	0.248	0.432
hispanic	6,105	0.180	0.384	1,882	0.226	0.419
	College, start in four-year sci			College, start in four-year hum		
AFQT	370	0.811	0.835	1,102	0.491	0.895
male	370	0.616	0.487	1,102	0.391	0.488
black	370	0.189	0.392	1,102	0.240	0.427
hispanic	370	0.114	0.318	1,102	0.112	0.315

Table 2: GPA over time by type of college attended

	Four-year college Sci			Four-year college Hum			Two-year college		
Period	Obs	Mean	Std Dev	Obs	Mean	Std Dev	Obs	Mean	Std Dev
year 1	295	3.042	0.757	855	3.058	0.655	870	3.048	0.746
year 2	290	3.141	0.594	866	3.127	0.545	580	3.045	0.651
year 3	259	3.223	0.545	960	3.198	0.492	336	3.149	0.517
year 4	237	3.280	0.493	910	3.248	0.486	154	3.093	0.612

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%	28.97%	29.15%	61.69%	56.49%
Stopped-out (SO)	23.81%	31.43%	31.75%	16.61%	18.48%
Dropped-out (DO)	35.10%	39.60%	39.10%	21.69%	25.03%
Total N	5,217	2,006	856	295	855

Notes: College Type refers to the *first* type of college enrolled in. Estimation subsample refers to the subsample over which we estimate the structural model.

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

	Type of college first attended							
	Two-year		Four-Year Sci		Four-year Hum		Any	
	N	%	N	%	N	%	N	%
CC	197	67.5%	245	91.8%	620	86.1%	1,062	83.0%
SO	95	32.5%	22	8.2%	100	13.9%	217	17.0%
Total	292	100.0%	267	100.0%	720	100.0%	1,279	100.0%

Table 5: Background characteristics of college enrollees

	AFQT	High School GPA	Mother with BA	1996 Family Income ($1996)
CC	0.67	3.28	32.52%	65,670
SO	0.29	2.94	18.97%	49,284
DO	0.11	2.88	17.27%	49,552

Table 6: Average GPA over time by college completion status

	Four-Year Sci		Four-Year Hum		Two-Year	
Period	CC	DO/SO	CC	DO/SO	CC	DO/SO
1	3.18	2.82	3.14	2.96	3.24	2.97
2	3.24	2.96	3.19	3.03	3.20	2.97
3	3.27	3.09	3.24	3.11	3.23	3.11
4	3.33	3.13	3.29	3.13	3.21	3.04

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

	Mean residual	Std Dev	N	Mean diff T (p-val)
4-year Science Majors:				
Drop out from 4-year college & science	-0.242	0.708	98	4.43
Complement	0.020	0.548	1167	(0.00)
Switch to 4-year college & humanities	-0.003	0.543	160	0.07
Complement	0.000	0.570	1105	(0.94)
Switch to 2-year college	-0.280	1.094	22	2.35
Complement	0.005	0.552	1243	(0.02)
4-year Humanities Majors:				
Drop out from 4-year college & humanities	-0.142	0.604	434	6.27
Complement	0.016	0.482	3894	(0.00)
Switch to 4-year college & science	0.023	0.542	121	0.53
Complement	-0.001	0.497	4207	(0.60)
Switch to 2-year college	-0.206	0.658	76	3.65
Complement	0.004	0.494	4252	(0.00)
2-year students:				
Drop out from 2-year college	-0.118	0.750	583	5.29
Complement	0.046	0.584	1499	(0.00)
Switch to 4-year college (any major)	0.105	0.514	228	2.63
Complement	-0.013	0.652	1854	(0.01)
Switch to 4-year college & science	0.078	0.543	29	0.66
Complement	-0.001	0.640	2053	(0.51)
Switch to 4-year college & humanities	0.076	0.514	170	1.63
Complement	-0.007	0.648	1912	(0.10)

Note: regression covariates include sex and race dummies, AFQT, age, and work intensity dummies

Table 8: Major transition matrix for stopouts

	Major when returned after stopping out				
Major before stopping out	science	humanities	2-year college	don't know / unreported	N
science	22.50%	27.50%	20.00%	30.00%	80
humanities	4.01%	54.32%	21.91%	19.75%	324
2-year college	4.44%	21.56%	49.05%	24.95%	473
don't know / unreported	4.62%	21.94%	37.64%	35.80%	433
Total	5.50%	30.15%	36.79%	27.56%	1,310

Note: Table only includes first instance of stopping out

Table 9: Average GPA over time by college type and work intensity

(a) 4-year Sciences

	Period			
	1	2	3	4
Work FT	3.02	3.09	3.11	3.24
Work PT	2.97	3.09	3.16	3.18
No Work	3.07	3.18	3.30	3.37
Total	3.04	3.14	3.22	3.28

(b) 4-year Humanities

	Period			
	1	2	3	4
Work FT	3.02	3.09	3.22	3.15
Work PT	3.04	3.13	3.19	3.26
No Work	3.07	3.14	3.20	3.27
Total	3.06	3.13	3.20	3.25

(c) 2-year college

	Period			
	1	2	3	4
Work FT	3.08	3.10	3.21	3.18
Work PT	2.98	3.08	3.15	3.10
No Work	3.10	2.95	3.07	2.96
Total	3.05	3.04	3.15	3.09

Table 10: Difference between actual and expected wages in time t for stopouts (by $t+1$ decision)

	Mean residual	Std Dev	N	Mean diff T (p-val)
Stay in work	0.026	0.491	6,583	2.77
Return to school	-0.054	0.369	297	(0.01)
Total	0.022	0.487	6,880	

Note: regression covariates include levels and interactions of the following variables: sex, race, and year dummies; AFQT; experience; age; and work intensity dummies

Table 11: Estimates of 2 and 4 year GPA Parameters

	4 year Science		4 year Humanities		2 year	
	Coeff.	Std. Error	Coeff.	Std. Error	Coeff.	Std. Error
Constant	3.31	(0.164)	3.14	(0.101)	3.23	(0.050)
Black	-0.17	(0.076)	-0.03	(0.040)	-0.12	(0.025)
Hispanic	0.03	(0.071)	0.04	(0.045)	0.03	(0.023)
AFQT	0.07	(0.037)	0.09	(0.020)	0.07	(0.011)
HS Grades	0.24	(0.033)	0.18	(0.020)	0.18	(0.010)
Work FT	0.04	(0.066)	-0.02	(0.041)	0.02	(0.022)
Work PT	0.02	(0.048)	0.00	(0.031)	0.03	(0.021)
Year 2+					-0.06	(0.021)
λ_0 (ability index intercept)	-0.50	(0.296)	-0.88	(0.256)	0.00	(—)
λ_1 (ability index loading)	1.12	(0.089)	1.28	(0.082)	1.00	(—)
Unobserved type 1	-0.32	(0.045)	-0.19	(0.030)	-0.10	(0.018)
Person-years obs.	935		1,828		1,584	

Note: Controls for parental education (college dummy) and age were also included.

Table 12: Estimates of Skilled and Unskilled Wage Parameters

	Skilled		Unskilled	
	Coeff.	Std. Error	Coeff.	Std. Error
Constant	2.25	(0.064)	1.78	(0.024)
Black	-0.05	(0.017)	-0.09	(0.006)
Hispanic	0.05	(0.019)	0.00	(0.006)
Age	0.01	(0.005)	0.02	(0.002)
Unskilled Experience	-0.01	(0.004)	0.05	(0.002)
Skilled Experience	0.07	(0.004)		
PT	-0.06	(0.017)	-0.01	(0.006)
PT 2 year			-0.10	(0.012)
PT 4 year			-0.17	(0.011)
FT 2 year			-0.03	(0.012)
FT 4 year			-0.04	(0.013)
PT graduate school	-0.05	(0.033)		
FT graduate school	-0.03	(0.027)		
1 year graduate school	-0.03	(0.026)		
2 years graduate school	0.02	(0.026)		
3 years graduate school	0.13	(0.044)		
4+ years graduate school	0.07	(0.060)		
1 year college			0.04	(0.007)
2 years college			0.06	(0.008)
3 years college			0.11	(0.010)
4+ years college			0.14	(0.010)
Unobserved type 1	0.29	(0.011)	0.10	(0.005)
person-years	2,273		13,140	

Note: Controls for AFQT, High School GPA, parental education (college dummy) and calendar year dummies were also included.

Table 13: Correlation Matrix and Variances for Unobserved Abilities

	Skilled	Unskilled	4 year Science	4 year Humanities	2 year
Correlation matrix					
Skilled	1.000	0.723	0.178	0.036	0.223
Unskilled	0.723	1.000	0.267	0.215	0.251
4 year Science	0.178	0.267	1.000	0.604	0.206
4 year Humanities	0.036	0.215	0.604	1.000	0.763
2 year	0.223	0.251	0.206	0.763	1.000
Variances	0.129	0.077	0.130	0.076	0.086

Table 14: Idiosyncratic Variances

Period	Skilled	Unskilled	4 year Science	4 year Humanities	2 year
1	0.129	0.158	0.314	0.316	0.367
2			0.159	0.137	0.277
3			0.108	0.105	0.175
4			0.057	0.069	
5+			0.137	0.140	

Table 15: Average Posterior Ability in 2010 for Different Choice Paths

Choice Path	Skilled	Unskilled	4 year Science	4 year Humanities	2 year	N
Continuous College, no work	0.02	0.02	0.05	0.03	0.02	104
Continuous College, work	-0.01	-0.01	-0.01	0.00	0.00	368
Stopout, graduate	0.02	0.02	0.01	0.00	0.00	136
Stopout then Dropout	-0.05	-0.05	-0.04	-0.02	-0.02	108
Dropout	-0.01	-0.01	-0.01	-0.01	-0.01	381

Table 16: Flow Utility Estimates

	2-year	4-year Sci	4-year Hum	Work PT	Work FT	Grad School
Constant	-1.992	-7.507	-6.993	-3.678	-4.079	-6.308
AFQT	0.016	2.176	1.093	0.245	-0.066	
Black	-0.251	1.701	1.547	-0.300	-0.623	
Hispanic	0.079	-1.368	-0.781	-0.057	0.105	
HS grades	0.258	1.153	1.344	-0.185	0.053	
Parent college	-0.189	1.609	1.185	-0.362	-0.398	
Prior Academic Ability	0.002	1.376	0.301			
Expected Log Wage				1.513	1.513	
Previous HS	0.677	1.817	1.500	0.370	0.084	
Previous 2-year	2.636	1.139	0.902	0.147	0.106	
Previous 4-year Sci	1.042	4.730	2.229	0.460	0.313	0.029
Previous 4-year Hum	0.368	1.945	3.547	0.553	0.606	0.571
Previous Work PT	0.093	0.364	0.421	1.495	1.125	0.024
Previous Work FT	-0.068	0.004	0.276	0.962	1.825	0.189
Previous Grad School				0.057	0.951	4.816
Graduated 4-year college				-4.646	-1.314	
Work PT	-1.406	-7.262	-5.071			-1.160
Work FT	-1.489	-3.748	2.866			-2.292
4yr Sci weight						-0.193
4yr Hum weight						0.603
Unobserved type 1	-0.120	0.251	0.224	-0.026	-0.014	-0.651
log likelihood	-25,191					
person-years	20,560					