Bumpy Rides: School to Work Transitions in South Africa*

Todd Pugatch
University of Michigan

Job Market Paper

December 26, 2010

Abstract

Re-enrollment in school following a period of dropout is a common feature of the South African school to work transition that has been largely ignored in both the literature on South Africa and the wider literature on sequential schooling choice. In this paper, I quantify the importance of the option to re-enroll in the school to work transition of South African youth. To do so, I estimate a structural model of schooling choice in South Africa using a panel dataset that contains the entire schooling and labor market histories of sampled youth. Estimates of the model's structural parameters confirm the hypothesis that enrollment choices reflect dynamic updating of the relative returns to schooling versus labor market participation. In a policy simulation under which re-enrollment prior to high school completion is completely restricted, the proportion completing 12 years of schooling rises 8 percentage points, as youth who would have dropped out under unrestricted re-enrollment reconsider the long-term consequences of doing so. The results suggest that the option to re-enroll is an important component of the incentives South African youth face when making schooling decisions.

JEL classification: I21; J24; O12. *Keywords:* human capital investment; labor supply; youth unemployment; dynamic discrete choice; South Africa.

*Department of Economics, Lorch Hall, 611 Tappan Street, Ann Arbor MI 48109; tpugatch@umich.edu.
I thank Cally Ardington, Emily Beam, Rob Garlick, Jessica Goldberg, Osborne Jackson, Erik Johnson, David Lam, Murray Leibbrandt, James Levinsohn, Brian McCall, Zoë McLaren, Francesc Ortega, Christina Plerhoples, David Ratner, Matthew Rutledge, Kevin Stange, Elias Walsh, Martin Wittenberg, Dean Yang and seminar participants at the University of Michigan and University of Cape Town for helpful comments. Cally Ardington, Nicola Branson, Jessica Goldberg, and David Lam oriented me to CAPS and graciously shared their programs for formatting the data.

1 Introduction

More than a decade after the fall of apartheid in South Africa, economic outcomes for previously disenfranchised groups remain bleak, particularly for youth: 52% of 20-24 year olds were unemployed in 2005 (Bannerjee et al., 2007). South Africa's youth unemployment is severe even by the standards of its region, with the lowest employment/population ratio among its neighbors and other large African economies shown in Figure 1. Behind South Africa's poor youth labor market outcomes lies an especially slow and bumpy transition from school to work: a remarkable 22% of 20 year-olds were enrolled in grades K-10 in 2001, reflecting frequent grade repetition and re-enrollment; Figure 2 shows that only Brazil has a higher rate among a group of comparison countries. Numerous studies have examined the mechanisms generating youth economic outcomes in South Africa, but most rely on static theories that assume students have perfect foresight about future job opportunities when making schooling decisions. Even the small number of dynamic sequential schooling models in the wider literature dismiss re-enrollment in school after spells of withdrawal as a "rare event" (Keane and Wolpin, 1997: pp. 487) or ignore it entirely. Yet in South Africa, where lack of labor market opportunities and high rates of school failure make enrollment decisions difficult for many youth, re-enrollment is a key feature of the school to work transition: in my sample, one third re-enroll at some point in their schooling careers, including 20% before completing high school.[1] While static models are incapable of explaining this irregular enrollment behavior, re-enrollment becomes sensible in a dynamic context in which youth are uncertain about future outcomes and the returns associated with their choices.

This paper aims to fill this gap in the literature by quantifying the importance of the

[1] Comparable rates in the NLSY79 for the United States are 20% and 5%, respectively (Keane and Wolpin, 1997).

option to re-enroll in the South African school to work transition. Specifically, I develop and estimate a dynamic model of school advancement and job search that allows for uncertain outcomes and re-enrollment following dropout during the schooling career. As youth learn the results of their enrollment and job search choices, they update their expectations about the relative returns to enrollment versus labor market participation, leading some who dropped out of school to re-enroll. In addition to explaining observed patterns of school enrollment, completion, and labor market participation, I quantify the importance of re-enrollment by conducting simulations in which the option to re-enroll before completing high school is restricted. I find that the re-enrollment restriction increases the proportion completing 12 years of schooling by 8 percentage points, as youth who would have dropped out under unrestricted re-enrollment reconsider the long-term consequences of doing so. To my knowledge, this is the first dynamic, sequential schooling model applied to South African data, and the first in the wider development literature to focus on the importance of school re-enrollment after periods of dropout.

The difficult school to work transition of South African youth in the post-apartheid era has been well documented. Racial disparities in school quality and student outcomes found under apartheid (Case and Deaton, 1999; Case and Yogo, 1999) persisted in the post-apartheid era (Yamauchi, 2005). Fiske and Ladd (2004) and Lam, Ardington and Leibbrandt (2008) find that high rates of grade repetition lead to lengthy school careers for black and coloured youth, while Lam, Leibbrandt and Mlatsheni (2008) find slow transitions from school to employment, with only 37% of black males aged 21-22 reporting ever obtaining employment. When facing such poor labor market outcomes, forward-looking agents may find it optimal to remain in school despite relatively low probabilities of advancement. Others may drop out and later re-enroll if unsuccessful in the labor market. Static human capital investment models, however, will not capture this behavior if agents adapt their expectations and alter their choices due to acquisition of new information, making a dynamic model appropriate. Eckstein and Wolpin (1999), Arcidiacono (2004), Stange (2007) and Joensen (2008), among others, estimate dynamic schooling models in which agents face uncertainty over academic advancement and labor market opportunities.

Like each of those studies, the model in this paper will allow agents to update their expectations about the relative returns to enrolling and labor market participation in a dynamic discrete choice framework. I extend the previous work by allowing for an option to re-enroll in school after a spell of labor market participation, an important feature of school to work transitions in South Africa that prevailing dynamic schooling models have overlooked.[2] The pervasiveness of adverse shocks to South African youth – grade repetition while in school, high and lengthy unemployment when in the labor market – lead to the frequently rocky transitions from school to work observed in the data. Dropouts who update their expectations about future labor market opportunities during labor market spells may choose to re-enroll as a result. In this paper, I model these transitions and estimate their underlying structural parameters.

Estimates of the model's structural parameters confirm the hypothesis that enrollment choices reflect dynamic updating of the relative returns to schooling versus labor market participation. The model replicates basic patterns of enrollment, grade advancement and employment observed in the data, according to characteristics such as completed schooling and recent and cumulative school failures. I use my estimates of the underlying structural parameters relevant to enrollment to conduct policy simulations. To quantify the importance of re-enrollment in enrollment choices and schooling outcomes, I simulate the model

[2] There are important exceptions. Keane and Wolpin (1997) do not restrict re-enrollment in their model despite referring to it as a "rare event" (pp. 16), while Eckstein and Wolpin (1999) allow re-enrollment within 5 years of entering high school. Belzil and Hansen (2002) estimate a dynamic discrete choice model of school enrollment that allows for re-enrollment after periods of "interruption," but they assume that such interruptions occur exogenously. Light (1995) estimates a hazard model of re-enrollment. Each of these papers uses U.S. data.

under removal of the re-enrollment option. Enrollment rates rise sharply throughout the schooling distribution after restricting re-enrollment before high school completion, relative to both the data and results from the unrestricted model. I find that the re-enrollment restriction increases the proportion completing 12 years of schooling by 8 percentage points: by increasing the opportunity cost of dropout, restricted re-enrollment prolongs enrollment spells. Additional policy simulations, such as enforcement and extension of compulsory schooling and increasing school pass rates, lead to qualitatively similar results as the restricted re-enrollment simulation. The results suggest that the option to re-enroll is an important component of the incentives South African youth face when making schooling decisions.

I organize the paper as follows. The next section explains the motivation for the paper in greater depth. Section 3 develops the formal model and explains the estimation method. Section 4 describes the data. Section 5 presents results, model fit, and robustness checks. Section 6 presents results of policy simulations. Section 7 concludes.

2 Motivation for a dynamic schooling model with re-enrollment

2.1 Weaknesses in existing static and dynamic human capital models

In this section, I develop intuition for the dynamic model of school to work transitions that I formulate and estimate in this paper. I argue that both the traditional, static human capital model and prevailing dynamic schooling models fail to capture salient features of the South African educational and labor market context. A model that allows agents to update their expectations about future outcomes and alter their decisions accordingly is therefore appropriate. To illuminate these features, I focus on the choice to re-enroll in school after a spell in the labor market, a common behavior among the youth in my sample but one that is largely ignored in both static and dynamic human capital models in the literature. With a simple stylized model, I argue that dynamic updating of expectations based on past outcomes is essential to understand the choice to re-enroll in school.

First, consider a simple version of the classic Becker (1964) model of human capital investment, as discussed by Card (1999). The agent maximizes an indirect utility function based on the present discount value of (log) earnings net of schooling costs:

$$U(y(s)) = \ln y(s) - c(s) \tag{1}$$

where y is the annual earnings function, assumed concave in schooling s; and c is the cost of schooling, assumed convex in s. The agent chooses s to satisfy the first-order condition $\frac{y'(s)}{y(s)} = c'(s)$, where the marginal benefit of an additional year of schooling equals its marginal cost. This static model assumes the agent faces no uncertainty over schooling outcomes or the shapes of the earnings and cost functions, and therefore need not update her choice of s as new information about the returns to schooling (due to say, imperfect knowledge of ability or labor market opportunities) arrives. Yet such dynamic considerations are likely to be quite important to schooling choices, particularly for young adults. As Card (1999) acknowledges, "In fact the transition from school to work is often a bumpy one, as young adults move back and forth between full-time or part-time enrollment and part-time or full-time work...individuals do not necessarily know the parameters of their earnings functions when they make their schooling choices."[3]

To address these concerns, a number of studies have developed and estimated dynamic human capital models, in which agents may update their schooling decisions based on arrival of new information about academic ability (Stange, 2007; Eckstein and Wolpin,

[3]Card (1999), p. 1810-1812.

1999) and labor market opportunities (Keane and Wolpin, 1997; Joensen, 2008). However, even in most dynamic models (such as Heckman and Navarro, 2007; and those of Stange, 2007 and Joensen, 2008), exiting school to enter the labor market is a terminal action. While this may be appropriate for developed country settings where re-enrollment in school after a spell in the labor market is rare, South African youth face high levels of grade repetition and unemployment that increase the uncertainty associated with their enrollment choices and make them especially vulnerable to adverse shocks as they transition from school to work.[4] The corresponding prevalence of re-enrollment in South Africa (in my sample, nearly one third of youth re-enroll at some point before age 24) make re-enrollment essential to incorporate in a dynamic human capital model.

2.2 Why agents re-enroll

In examining the importance of the option to re-enroll, first note that agents with this choice are unambiguously better off *ex ante* than those without, by virtue of the option value of re-enrollment.[5]

Under what circumstances will an agent with the option to re-enroll choose to do so? Consider a stylized dynamic setting in which labor market outcomes and academic

[4] Figure 2, showing K-10 enrollment among 20 year-olds in selected countries, suggests that grade repetition and re-enrollment in South Africa are high in an international context.

[5] To see how the re-enrollment option increases welfare, let $V^j(s)$ be the value function for an agent currently in the labor market, who receives flow payoff $w(s)$ from working (as a function of schooling s, the state variable) and discounts the future by discount factor β. Let $V^e(s)$ be the value function for enrollment, and s' the future level of schooling. Consider the Bellman equations for two young people in the labor market who differ only in that the second enjoys the option to re-enroll in school (in order to increase his or her stock of human capital and earn higher future wages) while the first does not:

$$V^j(s) = w(s) + \beta \mathbb{E}(V^j(s)) \tag{2}$$

$$V^j(s) = w(s) + \beta \mathbb{E}(\max\{V^j(s), V^e(s')\}) \tag{3}$$

The continuation values show that the agent with the re-enrollment option is better off, because $\mathbb{E}(\max\{V^j(s), V^e(s')\}) \geq \mathbb{E}(V^j(s))$.

success are stochastic. The agent is risk neutral and maximizes expected lifetime utility. For simplicity, assume no discounting of future payoffs. The model has three periods: in period one, the agent enters the labor market with zero skill units and finds employment with probability p. If she finds employment, she earns wage w_0, otherwise her payoff is 0. In period two, she may choose to remain in the labor market or re-enroll in school. If she remains in the labor market, she earns w_0 with certainty if she worked in period one, or else she must search again with probability p of success (hence expected payoff pw_0). If she re-enrolls in school, she earns a flow payoff of 0 and passes the grade with probability q, earning one skill unit. In period three, she must enter the labor market, but this time finds employment with certainty regardless of her employment status in previous periods, and earns a wage based on her skills.[6] If she never re-enrolled or re-enrolled and failed, she earns w_0. If she re-enrolled and passed, she earns $w_1 > w_0$. Figure 3 summarizes the model; Appendix A extends the model to consider the initial choice to drop out of school.

Now consider her decision rule for re-enrollment. Since her only choice occurs in period two, we need to consider only her expected payoffs from period two forward, conditional on the period one outcome. If she found a job in period one, then she earns w_0 in each of periods two and three if she remains in the labor market, whereas she earns an expected wage of $\mathbb{E}(w) = qw_1 + (1-q)w_0$ if she re-enrolls. Hence an agent who is successful in period one job search will re-enroll if:

$$\begin{aligned}
\mathbb{E}(U(\text{re-enroll})) &\geq \mathbb{E}(U(\text{work})) \Leftrightarrow \\
qw_1 + (1-q)w_0 &\geq 2w_0 \Leftrightarrow \\
q &\geq \frac{w_0}{w_1 - w_0} \equiv \underline{q}_j \qquad (4)
\end{aligned}$$

[6] Assuming certain employment in period three simplifies the exposition without affecting my conclusions.

where the subscript j denotes that the agent obtained a job in period one. Re-enrollment thus follows a threshold rule, where the agent re-enrolls if the grade advancement probability exceeds $\underline{q}$, and remains in the labor market otherwise.

An agent who did not find work in period one and who remains in the labor market in period two will earn expected payoffs pw_0 in period two and w_0 in period three. If she re-enrolls, she earns the expected wage defined above. Therefore the re-enrollment rule for an agent who did not work in period one is:

$$\begin{aligned} \mathbb{E}(U(\text{re-enroll})) &\geq \mathbb{E}(U(\text{work})) \Leftrightarrow \\ qw_1 + (1-q)w_0 &\geq pw_0 + w_0 \Leftrightarrow \\ q &\geq \frac{pw_0}{w_1 - w_0} \equiv \underline{q}_u \end{aligned} \tag{5}$$

where the subscript u denotes that the agent was unemployed in period one. Note that because $\underline{q}_j \equiv \frac{w_0}{w_1 - w_0} > \frac{pw_0}{w_1 - w_0} \equiv \underline{q}_u$, the threshold probability of passing required to re-enroll is greater for an agent who obtained work in period one than for an agent who was unemployed. This means that all else equal, adverse labor market shocks make an agent more likely to re-enroll. These heterogeneous propensities to (re-)enroll based on information acquired from past outcomes will form the basis for the structural estimation in this paper. The result squares with economic intuition: an agent with bleaker labor market prospects will be more willing to risk failure in order to increase her human capital and earn a higher subsequent wage. The comparative statics of the model are also quite intuitive: the threshold probability to re-enroll $\underline{q}$ is decreasing in the return to schooling $w_1 - w_0$, so that agents are more likely to invest in uncertain human capital acquisition the greater the returns; and $\underline{q}$ is increasing in the probability of finding employment p, so that skill acquisition must be more assured the more likely one is to succeed in job search in the absence of such skills.

Most notable about the result, however, is that agents with identical expected lifetime incomes upon entering the labor market in period one nonetheless may exhibit different subsequent behaviors after their uncertainty about the initial job search outcome is resolved. This dynamic updating in response to labor market outcomes, and the associated behavioral sorting based on past outcomes to which it leads, is difficult to reconcile with standard human capital models assuming perfect foresight.[7]

3 A model of the school to work transition

3.1 Timing and preferences

The model of the preceding section provided intuition about how one feature of the dynamic environment, the uncertainty of labor market outcomes, may lead to divergent choices and outcomes among otherwise identical agents. Yet this model is too simple to capture the observed behavior of agents making schooling decisions over their youth. In particular, the stylized model assumed homogeneity in the school advancement probability, no financial or psychic benefits or costs of schooling, and no wage return to labor market experience. In this section, I build on the intuition of the previous stylized model to develop a more complete model of enrollment choice in the presence of uncertain schooling and labor market outcomes. I relax each of the assumptions mentioned above, and explain how I estimate the structural parameters of the model using panel data on South African youth.

Consider the discrete-time, dynamic environment of a finitely-lived agent seeking to

[7] The notion that dynamic updating could drive re-enrollment accords with Heckman and Navarro (2007). Their sequential schooling model does not allow re-enrollment, but they nonetheless speculate, "In a general model, different persons could drop out and return to school at different times as information sets are revised" (pp. 364).

maximize lifetime utility by making discrete choices about school enrollment and labor market participation. In each period, the agent observes the state vector S_t, which summarizes all known information relevant to her choice at time t. She then chooses whether to enroll in school or enter the labor market; let $d_t = \mathbb{I}(\text{enroll at time } t)$. The agent may enroll in periods $t = t_0, \ldots, T_d$, but in all subsequent periods $t = T_d + 1, \ldots, T$, the agent must participate in the labor market. In each period, the agent receives choice-specific flow payoffs $U_t^d(S_t)$. Thus the agent's optimization problem at any time in the decision period $t \leq T_d$ is:

$$\max_d \mathbb{E}\left[\sum_{t=\tau}^{T_d} \beta^{t-\tau}\left\{d_t U_t^e(S_t) + (1-d_t)U_t^j(S_t)\right\} + \sum_{t=T_d+1}^{T} \beta^{t-\tau} U_t^j(S_t)\right] \tag{6}$$

where U^e and U^j are the utility functions for enrollment and labor market participation, respectively; β is the agent's discount factor; and the expectation is taken with respect to the evolution of the state space. The (indirect) utility functions are expressed in monetary units (South African rand per year), and capture the agent's preferences for enrollment and labor market participation as a function of S_t. I assume the agent is risk neutral. I also assume that the utility functions are additively separable in observable (to the econometrician) components X_t and choice-specific shocks observed only by the agent, ϵ_t^d, so that $U_t^d(S_t) = u_t^d(X_t) + \epsilon_t^d$. The agent learns the shocks $\epsilon_t = (\epsilon_t^j, \epsilon_t^e)$ at the beginning of each period, prior to the enrollment decision.

If the agent chooses to enter the labor market, she must search for employment, with probability of success conditional on current information contained in S. Her (flow) payoff in the labor market is her expected wage, i.e., the product of the probability that she will find employment and the wage she would earn if successful, all conditional on information known at the beginning of the period:[8]

[8]Note that I implicitly normalize the value of unemployment to zero, consistent with the absence of a

$$
\begin{aligned}
U_t^j(S_t) &= \mathbb{E}[w(X_t)] + \epsilon_t^j \\
&= \Pr(\text{work}|X_t) \times w(X_t) + \epsilon_t^j
\end{aligned} \tag{7}
$$

I model wages as a linear function the observable state variables, i.e., $w(X_t) = X_t\gamma$. The probability of working is estimated as a logit that is linear in the same covariates as the wage equation.[9]

If the agent chooses to enroll in school, she passes the grade level and accumulates associated human capital with probability that depends on the current state. Her payoffs depend on her net psychic benefits of schooling (b) based on all information known at period t, less the school fee. Thus we have:

$$
U_t^e(S_t) = b(X_t) - fee(X_t) + \epsilon_t^e \tag{8}
$$

The net psychic benefits of schooling include the agent's self-assessment of her ability based on both pre-determined characteristics and her schooling outcomes up to year t, as well as other characteristics that may influence the non-monetary benefits and costs of school.[10] I parameterize b as a linear function of X, i.e., $b(X_t) = X_t\alpha$.

Table 1 describes the elements of X, with further details in the Data section and

widely accessible unemployment insurance system in South Africa. Although other household recipients may receive public transfer payments, such as pensions, and share their resources with youth, such transfers do not depend on whether the youth is enrolled in school or in the labor market.

[9]This treatment of job search nests the approach of a companion paper that analyzes youth's first post-school labor market experiences using the same dataset (Levinsohn and Pugatch, 2010). That paper estimates $\Pr(\text{work}|X_t) = q(X_t) \times (1 - F_{w|X}(w^*))$, where q is the arrival rate of job offers, F is the wage offer distribution, and w^* is the reservation wage. Whereas that paper focused on disentangling q from the parameters of F under alternative measures of the reservation wage, my focus here is on the relative returns of enrollment versus job search, for which the distinction between the arrival rate and distribution of wage offers is less important. Therefore, I choose to forego estimation of these parameters, which simplifies estimation of the model.

[10]The psychic benefits of schooling may be negative, if an agent derives disutility from schooling.

Appendix C. The table also notes the exclusion restrictions for the utility functions (7) and (8) and the transition equations for enrollment and passing (described in the next subsection).[11] The state space X includes race and gender dummies, indicators for ability quartiles based on results from a standardized test administered in Wave 1 of the panel, and a dummy for completion of grade 6 or higher at age 12,[12] when the model begins. In addition to the state variables in Table 1, I also assume that expected wages evolve with a quadratic trend in age beginning at the close of the decision horizon ($t = T_d + 1$) until the end of the model at $t = T$.[13] The model assumes that the returns to schooling are linear, but allows for changes in slope and intercept at high school completion.[14] By including time-varying covariates such as work experience and an indicator for employment in the last relevant period, the model allows agents to update their expectations of labor market outcomes based on the results of past choices (pass or fail for enrollment, work or search for job search).

In the enrollment utility function, I account for (possible) credit constraints through inclusion of indicators for lower quintiles of household incomes (y_q for quintiles $q = 1, 2$) and their interaction with HSG, reflecting the potential difficulties youth from lower-income households face in financing post-secondary education.[15] The completed schooling

[11] These exclusion restrictions are theoretically motivated, and not necessary for identification.

[12] Completing grade 6 by age 12 represents "on-time" school completion for students who enter at age 6, the modal school entry age in South Africa.

[13] I restrict the age-expected wage profile to begin at $t = T_d + 1$ for several reasons. First, because no individual in the data is observed after age 26 ($t = T_d + 3$), I prefer to control for actual work experience, rather than age, prior to T_d. Second, estimating age coefficients with wage data on such young agents is likely to be biased, so I use auxiliary data from the 2001 South African Census to estimate the age-expected wage profile. Details may be found in Appendix C.

[14] Case and Yogo (1999) also find evidence for a slope change in the returns to schooling at high school completion, although their outcome measure is log wages, whereas my outcome is expected wages.

[15] This method of allowing household income to proxy for credit constraints follows Cameron and Heckman (2001), Carneiro and Heckman (2002) and Lam et al. (2010), although the resulting estimates may reflect the accumulation of educational disadvantage in low income households, rather than credit constraints. Formally incorporating credit constraints would require a richer model that incorporates asset accumulation and savings, as in Keane and Wolpin (2001) and Cameron and Taber (2004).

terms $(s, HSG, (s-12)*HSG)$ capture how the net psychic benefit of schooling evolves over time. The failure terms (f, f_{tot}) capture how the agent dynamically updates her self-assessed academic ability; including separate terms for recent and cumulative school failure allows the agent to place different weight on recent versus past outcomes.[16] Figure 4 summarizes timing, choices, and payoffs in the model.

3.2 State transitions

Because the choice environment is dynamic, the agent considers not only current payoffs when deciding to enroll in school, but also how the current period's choice affects expected future payoffs. Following standard practice in dynamic discrete choice models, I assume that transitions of the state variable follow a first-order Markov process, and that transitions of the unobserved shocks are conditionally independent from those of the observed state.[17]

In the present model, an agent choosing to enter the labor market considers the probability of finding employment in the current period, which enters the labor market utility

[16] The process of dynamic updating of the net psychic benefits of schooling in the model may be viewed as an approximation to Bayesian updating. Partition X_t into $[\mathcal{X}_0, \mathcal{X}_t]$, representing its time-invariant and time-varying elements, respectively. If both the prior and posterior distributions for b are normal, then the posterior mean $\mathbb{E}[b|X_t]$ is a linear combination of the prior mean $\mathbb{E}[b|\mathcal{X}_0]$ and the mean of $\mathcal{X}_t$:

$$\mathbb{E}[b|X_t] = a_1 \mathbb{E}[b|\mathcal{X}_0] + a_2 \bar{\mathcal{X}}_t$$

where $a_1 = \frac{1/\sigma_b^2}{(1/\sigma_b^2)+t\sigma_X^2}$ and $a_2 = \frac{t\sigma_b^2}{(1/\sigma_b^2)+t\sigma_X^2}$. Here, σ_b^2 and σ_X^2 are the variances of b and $\mathcal{X}_t$, respectively. Note that as t increases, the agent places relatively more weight on $\mathcal{X}_t$ when updating $\mathbb{E}[b|X_t]$, reflecting the importance of new information. Stange (2007) gives a similar interpretation of the agent's updating rule, noting that the process is similar to the normal learning model.

[17] These assumptions allow me to estimate transitions of next period's observed state X_{t+1} as a function of the current observed state X_t only. This simplification comes at a cost, as it assumes that the transition probability for X_{t+1} depends on unobserved shocks only through their influence on the current state variable. This property would not hold if, for instance, an unobserved negative household shock that does not cause a youth to fail the current grade nonetheless persists and affects the following period's academic performance, leading her to fail the subsequent grade. However, assuming that persistent shocks affect the observed state variable upon impact, thereby allowing the current observed state to serve as a sufficient statistic for the distribution of next period's observed state, seems reasonable given the large gain in computational tractability.

function, as well as the effect that current success or failure in job search will have on future job opportunities, which enters the state transition function. Because work experience variables are the only elements of the state space that evolve during periods in which the agent enters the labor market, the corresponding state transition simplifies to the probability that the agent finds work:

$$\begin{aligned}\Pr(X_{t+1}|X_t, d_t = 0) &= \Pr(x_{t+1} = 1, x_{tot,t+1} = x_{tot,t} + 1|X_t, d_t = 0) \\ &= \Pr(\text{work}|X_t, d_t = 0) \end{aligned} \tag{9}$$

Similarly, for periods in which the agent enrolls in school, the state transition simplifies to the probability that the agent will pass the grade level:[18]

$$\begin{aligned}\Pr(X_{t+1}|X_t, d_t = 1) &= \Pr(s_{t+1} = s_t + 1, f_{t+1} = 0, f_{tot,t+1} = f_{tot,t}|X_t, d_t = 1) \\ &= \Pr(\text{pass}|X_t, d_t = 1) \end{aligned} \tag{10}$$

I estimate both state transitions (9) and (10) as logits that are linear in X_t, using data on labor market and schooling outcomes, respectively.[19] The labor market utility function (7) and employment probability (9) functions also include the macro environment indicator ζ, but for simplicity agents forecast next period's macro state to be identical to the present. Denote the logit parameters governing the labor market transition (9) and

[18]This treatment of the state transition while enrolled follows Lam, Ardington and Leibbrandt (2008), who treat school progression as the outcome of a threshold advancement rule with stochastic shocks.

[19]The state transitions may also be estimated nonparametrically. Since X is discrete, a simple bin estimator would make nonparametric estimation particularly straightforward, for example. However, because many covariate cells would have few observations, consistent estimation of the state transitions using nonparametric methods would be difficult, so I opt instead for parametric estimation using logits.

enrollment transition (10) as ϕ_j and ϕ_e, respectively.

3.3 Model Solution and Estimation

Because the model has a finite horizon, it may be solved by backward recursion. Via Bellman's principle of optimality, rewrite the agent's problem as a dynamic programming problem, where the value function V_t is defined as the maximal expected present value of utility at time t, conditional on the state S_t:

$$V_t(S_t) = U_t(S_t) + \beta \mathbb{E}\left[V_{t+1}(S_{t+1})\right] \tag{11}$$

Because the value function assumes optimizing behavior by the agent from period t forward, it may also be expressed as the maximum over alternative-specific value functions:

$$\begin{aligned} V_t(S_t) &= \max_d \{V_t^d\} \\ &= \max_d \left\{ U_t^d(S_t) + \beta \mathbb{E}\left[\max_d \{V_{t+1}^d(S_{t+1})\} | S_t, d_t \right] \right\} \end{aligned} \tag{12}$$

The assumptions of additive separability and conditional independence of the state space allow me to treat the $\mathbb{E}[\max]$ term in (12) as a double integral over the marginal distributions of X and ϵ. I discussed the transitions of X in the previous section. However, to make the problem tractable for estimation, I must assume a parametric form for the distribution of ϵ. I assume that the unobserved shocks are independently and identically distributed type I extreme value, i.e., $F(\epsilon_t^d) = \exp\left(-\exp\left(-\epsilon_t^d\right)\right)$. The i.i.d. assumption on ϵ is admittedly restrictive, as it implies that unobserved shocks affecting enrollment utility occur independently from those affecting labor market utility. Although it is easy to think of unmodeled shocks that could affect enrollment and labor force participation simultaneously (illness, household job loss, etc.), many types of shocks, such as the transfer of a talented teacher or the unexpected destruction of a job due to demand-side factors, will affect the utility of just one choice. Assuming that ϵ is drawn from an i.i.d. Type I EV distribution allows me to write $\mathbb{E}[\max]$ in closed form, a considerable computational savings, though I may relax this assumption in future work.

The assumptions of additively separable utility and conditionally independent and i.i.d. Type I EV shocks greatly simplify the form of the $\mathbb{E}[\max]$ terms of the Bellman equation, which we may rewrite as:

$$\begin{aligned}\mathbb{E}\left[V_{t+1}(S_{t+1})|S_t, d_t\right] &= \mathbb{E}\left[\max_d\{V_{t+1}^d(S_{t+1})\}|S_t, d_t\right] \\ &= \mathbb{E}\left[\ln\left(\sum_d \exp\left[V_{t+1}^d(X_{t+1})\right]\right)|X_t, d_t\right] + \nu \qquad (13)\end{aligned}$$

where $\nu \cong .577$ is Euler's constant. Substituting the above closed-form expression for $\mathbb{E}[\max]$ into (12) serves to calculate the agent's expectations over the unobserved state variables ϵ. Combined with my assumptions on the transitions of the discretized observable state variables X, we have:

$$\mathbb{E}\left[V_{t+1}(S_{t+1})|S_t, d_t\right] = \sum_X \left(\left[\ln\left(\sum_d \exp\left[V_{t+1}^d(X_{t+1})\right]\right)|X_t, d_t\right] + \nu\right)\Pr(X_{t+1}|X_t) \quad (14)$$

The model's structural parameters are $\theta = (\phi_j, \phi_e, \gamma, \alpha)$, where ϕ_j and ϕ_e are the labor market and enrollment transition parameters; γ is the parameter vector for the wage equation; and α is the parameter vector describing enrollment preferences. For any value of θ, we may solve the model recursively by using the terminal condition that following the close of the decision horizon at T_d, the agent must enter the labor market. Therefore, I set $U_t^e(X_t) = U_t^j(X_t) = \mathbb{E}(w|X_t) + \epsilon_t^j$ for all $t > T_d$, which allows me to solve for the value functions in periods $t = T_d + 1, \ldots, T$ by backward induction. The Type I extreme value assumption on the distribution of utility shocks and their conditional independence over choice alternatives allow me to express the agent's conditional enrollment probability as:

$$\Pr(d_t = 1|X_t) = \frac{\exp\left(V_t^e(X_t)\right)}{\exp\left(V_t^e(X_t)\right) + \exp\left(V_t^j(X_t)\right)} \tag{15}$$

With data on the sequence of enrollment choices $\{d_t\}_{t=1}^{T_d}$, I decompose an individual's contribution to the likelihood function as:

$$l_e(\phi_e) = \prod_{t=1}^{T_d} \Pr(\text{pass}_t|X_t, d_t = 1)^{d_t} \tag{16}$$

$$l_j(\phi_j) = \prod_{t=1}^{T_d} \Pr(\text{work}_t|X_t, d_t = 0)^{1-d_t} \tag{17}$$

$$l_w(\gamma) = \prod_{t=1}^{T_d} f\left(w(X_t)^{\mathbb{I}(\text{work}_t)}\right) \tag{18}$$

$$l_d(\alpha) = \prod_{t=1}^{T_d} \Pr(d_t = 1|X_t)^{d_t} \Pr(d_t = 0|X_t)^{1-d_t} \tag{19}$$

where the components are the grade transition (16), the labor market transition (17), the wage equation (18)[20] and the enrollment choice equation (19). The individual's likelihood contribution is then:

$$l(\theta|X) = \prod_{t=1}^{T_d} l_e \cdot l_j \cdot l_w \cdot l_d \tag{20}$$

[20] Here, $f(\cdot)$ is the density of the wage residuals.

With panel data $\{d_{it}, X_{it}\}_{t=1}^{T_d}$ for $i = 1, \ldots, n$ individuals, the likelihood function becomes:

$$L(\theta|X) = \sum_{i=1}^{n} \ln l_i(\theta|X_i) \tag{21}$$

Estimation of θ may proceed by a nested fixed point (NFXP) algorithm, as in Rust (1987): in the "inner loop," the current guess of θ is used to solve the model recursively, as described above, while in the "outer loop," each guess of θ is updated through a numerical optimization procedure. The process repeats until convergence.

Because the likelihood function is additively separable in the contributions of enrollment ($d = 1$) and labor market ($d = 0$) components, I am able to estimate these parts sequentially. First, I use data on employment and school advancement to estimate (ϕ_e, ϕ_j) using (16) and (17); employment and accepted wage data are used to estimate γ using (18). The resulting estimates $(\hat{\phi}_j, \hat{\phi}_e, \hat{\gamma})$ are then substituted into (19) to estimate α. I set the discount factor β to 0.95. The likelihood function converges to identical parameter estimates from several different arbitrary starting values.

3.4 Parameter Identification

This subsection discusses identification of the model's structural parameters $\theta = (\phi_j, \phi_e, \gamma, \alpha)$. Employment and grade advancement probabilities conditional on observable characteristics identify the transition parameters ϕ_j and ϕ_e, respectively, as in standard logistic regressions for binary outcomes. Variation in state variables across individuals and over time among those in the labor market identifies γ, the parameters of the expected wage function. Although γ is identified from the selected sample that chooses to participate in the labor market and therefore may not be interpreted as causal, the discrete choice model fully specifies the selection process. Differences in enrollment rates by net psychic benefits of schooling (b), both across individuals and over time, identifies the enrollment utility function parameters α. For example, the difference in enrollment rates between those who failed their last grade enrolled and those who passed identifies the coefficient on the failure indicator in (8).

In the model, the transition parameters ϕ_e and ϕ_j capture youth expectations about future schooling and labor market outcomes, respectively. The parameters α and γ capture expectations about the utility associated with enrollment and labor force participation, respectively. Precise estimation and sensible signs of coefficients on time-varying characteristics within each structural parameter would provide evidence in favor of my hypothesis that youth dynamically update their expectations about the relative returns to enrollment versus labor market participation. The most notable coefficients are those on school failure and work experience (recent and cumulative), as such coefficients reflect updating based on previously uncertain schooling and labor market outcomes. Evidence of dynamic updating may also be found in the schooling coefficients (schooling, the indicator for high school graduate, and their interaction), because schooling also evolves based on past enrollment choices and their outcomes.

4 Data

4.1 The Cape Area Panel Study

I estimate the model with data from the Cape Area Panel Study (CAPS), a longitudinal study of youth in metropolitan Cape Town, South Africa (Lam et al., 2008c). CAPS sampled about 4,800 youth aged 14-22 in Wave 1 (2002) and currently includes four publicly available waves, the most recent conducted in 2006. In Wave 1, retrospective life histories were collected for each year stretching back to birth, and include information on school enrollment and advancement, job search, and employment. I update this retrospective life history data with information from Waves 2-4 to construct the panel used in this paper. I make several sample restrictions. I keep only those youth observed until at least age 18. Those who report advancing two or more grades in a year, or without continuous information on enrollment, are dropped from the sample. I drop those who report entering school prior to age 4 or exiting school after age 24 (which effectively sets $T_d = 24$ as the decision horizon).[21] I also drop those whose educational histories, by my definition of school advancement (described below), place them with more than 16 years of completed schooling. This leaves $N = 3,806$ individuals in the sample, comprising 68,898 person-years from ages 4-26. Table A1 shows the panel balance at selected ages. The sample size drops sharply at later ages due to both the young ages of entry into the panel (i.e., right-censoring) and attrition in later waves. Although whites are more likely than blacks or coloureds to be out of the sample after age 18, panel imbalance by race is not severe.

At each age, the data contain information on the youth's enrollment status. If enrolled, the data report the school outcome (pass or fail); if in the labor market, the data report whether employed.[22] Enrollment in school supersedes labor market participation when these are reported to occur simultaneously at a given age, in order to keep these choices consistent with their treatment as mutually exclusive in the model.[23] For post-secondary education (grades 13-16), I expand the definition of grade advancement to include a re-

[21] Reporting school entry prior to age 4 is more likely to reflect measurement error than childhood precociousness, in my view. Few observations are available in the data after age 24 (less than 20% of the sample is observed beyond 24, due to starting ages in Wave 1 and attrition), raising concern that estimating the model with these enrollment choices will result in severe finite-sample bias.

[22] Although search behavior is also reported in the sample, all non-enrolled youth are assumed to be participating in the labor market. In principle, the model could be extended to include non-participation in the labor market as a discrete choice. However, a companion paper using CAPS (Levinsohn and Pugatch, 2010) finds that 64% of those who permanently exit school and never report searching nonetheless find employment by the end of the sample, suggesting that labor market participation is not synonymous with active search for these youth.

[23] Work or search while enrolled never exceeds 3% of the sample at any grade level, and never exceeds 2% during grades 1-12.

sponse of "no grade/continuing" to the survey question on school result, as this is consistent with continuation in a multi-year degree program, and is the modal response for these grades.[24] Table A2 shows, using monthly calendar data collected during CAPS Waves 1-4, that most school enrollment and labor force spells last at least one year, regardless of whether I allow a spell to include other concurrent activities or restrict the definition to full-time spells (although removal of censored spells reduces mean duration somewhat, because the longest spells tend to be censored). Moreover, because most school enrollment spells tend to follow the January-December academic year (as shown in Figure 5), my assumption in the model that time periods last one year seems reasonable.

Data on accepted wages and school fees are available only at each panel interview rather than at each age. To overcome this restriction, I predict wages and school fees by regressing observed values on the state variables in Table 1.[25] Appendix B contains additional information on the South African education system, while Appendix C contains additional information on the sample and variable definitions.

For the purposes of the model, the first decision period t_0 occurs at age 12, as no youth enter the labor market and obtain work prior to this age, so extending the model to an earlier age would be superfluous. Of course, by age 12 substantial, non-random differences have already arisen among youth due to unobserved characteristics related to family background and motivation. Therefore, I include a dummy for completion of at least 6 years of schooling at age 12 as a state variable ("on-time" completion for a student

[24]Unfortunately I am unable to distinguish reliably whether those in post-secondary education are making satisfactory progress to degree completion, because few students report academic failure in these grades.

[25]I predict school fees using only race dummies, schooling, and a dummy for high school graduate, to be consistent with racial disparities in school expenditure and the discontinuous jump in fees in post-secondary schooling. I am concerned that if additional information on, say, cumulative grades failed were used to predict school fees, results could reflect unmodeled endogenous school choice and therefore be biased. For both expected wages and school fees, I replace predicted values below the minimum value observed in the sample with the first percentile from the data, in order to avoid non-positive predicted values and extreme outliers.

who enters at age 6), which proxies for the effect of such early life differences. The decision period ends at $T_d = 24$. The model ends at $T = 60$, the age of public pension eligibility in South Africa.[26] I top-code grades failed and work experience since age 12 at 3 and 4, respectively, which reduces the dimension of the state space while still accurately capturing more than 99% of the person-year observations in the sample.

4.2 Summary statistics and stylized facts

Table 2 presents summary statistics for the sample. The racial distribution reflects the unique racial composition of Cape Town, where coloureds (mixed racial heritage) are prominent. Schooling careers range on average from ages 6 to 18, with a mean of 1.2 grades failed.[27] There is wide variation (standard deviation 1.2 years) in completed schooling even by age 12. Only slightly more than half of the sample has worked, though this in part reflects right-censoring of school careers rather than failure in job search. As mentioned in Section 1, one third of the sample has re-enrolled in school (after disenrolling for at least one year) at some point in the sample, including 20% who re-enroll before completing secondary school.[28] As explained in Section 2, this re-enrollment behavior is difficult to reconcile with standard human capital theory, but is a prominent feature of the dynamic model I estimate.[29]

[26] Although individuals must also pass a means test to receive a public pension, in practice about 80% of elderly blacks and coloureds receive the pension. The pension is quite generous relative to the median incomes of these groups (Case, 2001).

[27] "Failure" refers to any type of failure to complete a grade level while enrolled, and includes withdrawal from school as well as academic failure.

[28] Comparable rates in the NLSY79 are 20 percent ever re-enrolled and 5 percent re-enrolled before high school completion (Keane and Wolpin, 1997).

[29] Moreover, it is unlikely that measurement error, in the form of recall bias, leads to a gross overstatement of re-enrollment rates. If recall bias were severe, we would expect re-enrollment rates to be greater for observations from ages before youth entered the panel and were subject to frequent interviews. Yet in a regression of the re-enrollment indicator on a dummy for whether the observation overlapped with the time of the panel, the coefficient is positive and significant (coefficient 0.06, t-statistic 20.0), which is the opposite of the result one would find in the presence of recall bias. The regression includes a full set of age and schooling dummies, controls for work experience and school failure (in both the last relevant period

Table 3 shows the means of selected variables across various demographic and educational categories. Racial disparities in schooling and labor market outcomes are substantial: blacks enter school one year later and last enroll nearly one year earlier than whites, on average, and fail 1.7 grades, compared to 0.5 for whites. Black wages are less than one third those earned by whites. Coloureds occupy an intermediate group between these extremes. Gender differences are not nearly as pronounced, however. Schooling and labor market outcomes follow the expected patterns when splitting the sample by ability or completed schooling. However, re-enrollment is not correlated with schooling or labor market disadvantage in a simple way. For instance, although similar proportions of black and whites re-enroll at some point in the sample (39% and 43%, respectively), 31% of black youth re-enroll before completing grade 12, compared to only 14% of whites. Similarly, high ability (those scoring above the age-adjusted, within-sample median on the standardized literacy and numeracy evaluation) youth re-enroll at higher rates than those with low ability (39% to 28%), but both groups re-enroll at similar rates before grade 12 (18% and 22%). This suggests that the option to re-enroll may be taken for different reasons according to background characteristics and current levels of completed schooling. The model accounts for such heterogeneity through its demographic controls and by allowing enrollment preferences and labor market returns to change discontinuously following high school completion.

High rates of failure, dropout and re-enrollment contribute to lengthy school careers among Cape Town youth, even among those who do not reach high levels of schooling. Among those who complete 8-9 years of schooling (for whom "on-time" completion would be no later than age 15 for a student entering at age 6 who enrolled continuously with no failure), nearly 45% remain enrolled till at least age 17, as shown in Table 4, panel (a).

and cumulative), and individual fixed effects, with standard errors clustered at the individual level.

Among blacks, this figure exceeds 75%, including more than 29% who remain enrolled till at least age 19. School careers in this group are even longer among the subset who have dropped out and re-enrolled in school at least once: Table 4, panel (b) shows that 83% of re-enrollees who complete 8-9 years of schooling remain in school until at least age 17.

Figure 6, panel (a) depicts patterns of enrollment and school failure by level of completed schooling. Enrollment rates are relatively high among those in the primary school grades (1-6), though not universal,[30] falling steadily to a trough around 20% at high school completion. Those with some post-secondary education (grades 13 and above) tend to remain enrolled, however. Enrollment and failure rates are negatively correlated, with failure highest at 11 years of completed schooling (when students must pass the "matric," or secondary school completion exam). Table 5, Panel (a) shows that such failure stems largely from academic failures, rather than withdrawal from school or continuation in an educational program.[31,32]

The discrete choice model of this paper treats (expected) wages and school fees as the directly measurable economic factors influencing enrollment decisions; Figure 6, panels (b) and (c) show employment, wages and fees as functions of completed schooling. Employment rates and wages appear to increase convexly in completed schooling, with large returns to post-secondary schooling. The apparent non-linearity of returns at these levels of schooling motivates the inclusion of terms (HSG and $HSG * (s - 12)$, described in Table 1) allowing the intercept and slope of the expected wage equation to change in post-secondary school-

[30]Recall that the sample begins at age 12, when many youth will already be close to primary school completion. Extending the sample to age 4 makes primary school enrollment nearly universal.

[31]Of course, school withdrawal could be another form of academic failure in which a failing student stops attending rather than face outright failure.

[32]Note that beginning in post-secondary (completed schooling 12 or above), I alter the definition of failure so that "no grade/continued" counts as passing the level, which explains why this response is zero for this group. Relatively high rates of "no grade/continued" prior to secondary school completion may be due to enrollment in vocational training or equivalency programs, although enrollees in National Training Certificate (NTC) programs report similar "no grade/continued" rates as the full sample.

ing. School fees also jump sharply in post-secondary education, suggesting that tuition costs, rather than failure rates (which are relatively low in these grades), are the main disincentive to enrollment in higher education. Yet despite the large increase in fees for higher education, their (mean) reported level never exceeds 15,000 rand per year, compared to annualized wages of more than 60,000 rand among those with some higher education (Table 2).[33]

4.3 Evidence consistent with dynamic updating of expected returns

Although the descriptive evidence presented thus far in this section helps to illustrate the relative costs and benefits of enrollment over the school career in the South African context, relatively straightforward modifications to the standard, static model of human capital investment (incorporating school failure and employment rates to modify expected returns, for instance) could still capture observed behavior quite well.[34] To justify the specification of a dynamic model, we must also observe behavior that could not be accounted for in the static case, such as enrollment (and re-enrollment) decisions reflecting dynamic updating of expected returns following acquisition of new information. I will argue that the next set of figures are consistent with such dynamic updating. In the following subsection, I consider alternative explanations for the observed enrollment and dropout patterns.

Histograms of the duration of the first two spells of enrollment and labor market entry (Figure 7) provide the first pieces of suggestive evidence on dynamic updating. Recall that I define enrollment and labor market participation as mutually exclusive and exhaustive states in the model, so that an enrollment spell must be followed by a labor market spell,

[33]These values, as with all currency-denominated units in the paper (unless otherwise noted), are in real South African rand per year, base year 2002. The South African rand traded at 10.3 per US dollar in August 2002, when CAPS Wave 1 began.

[34]For instance, Altonji (1993) develops a model of sequential schooling choice with uncertain completion, but ultimately estimates reduced-form regressions for wages and completion probabilities.

and vice versa. The histogram of the first enrollment spell (panel [a]) shows a largely symmetric bell shape, which would be consistent with a static human capital model with smooth returns to schooling. However, the first labor market spell (panel [c]) tends to be short, with a mode of just one year. The second enrollment spell (panel [b]) then lasts for 1-3 years, followed by a second labor market spell whose histogram has a long right tail (panel [d]). This pattern – a short initial experience in the labor market, followed by a period of additional human capital investment and subsequently stronger attachment to the labor market – suggests that youth are experimenting in the labor market, and choosing to return to school if their initial experience is dissatisfying.[35] Figure 8 shows that this experimentation, as reflected in re-enrollment and dropout rates, gathers steam in secondary school and peaks at the end of secondary school and the beginning of post-secondary education (grades 12-13).[36]

I have hypothesized that such experimentation in labor market entry and exit occurs because youth dynamically update their expectations about the relative returns to enrollment and job search based on the outcomes of their past choices. For instance, a student who fails a grade may be more likely to fail in the future, as she learns that she is less able academically than previously thought. Similarly, a youth who succeeds in her job search may be more likely to succeed in the future, as she builds a professional network and firm-specific human capital or forms a long-term contract with her employer. Figure 9 shows that these patterns exist in the data.[37] If youth know these conditional probabilities of passing and employment, they will change their behavior as they update their expectations based on the outcomes of their past choices.

[35]Figure 7 includes both completed and censored spells. All panels of the figure appear similar when restricting attention to completed spells, although the right tails are not as long.

[36]Re-enrollment means enrollment following at least one period of non-enrollment.

[37]Of course, these patterns may also simply reflect selection, as the less able and employable are more likely to fail and be unemployed, respectively, both now and in the past. The model fully specifies the dynamic selection process, however, and includes a rich set of controls to account for such selection.

Figures 10 and 11 suggest that enrollment decisions reflect such dynamic updating about returns. Panel (a) of Figure 10 shows enormous differences in enrollment rates based on the last school result: among those with less than 12 years of schooling, those who passed their last grade are more than 60 percentage points more likely to enroll than those who failed. Similarly, panel (b) shows that enrollment is sharply decreasing in total grades failed since age 12. Figure 11 shows that re-enrollment is correlated with labor market outcomes: in panel (a), we see that youth with less than high school qualifications are less likely to re-enroll in school if they were employed in their last period of non-enrollment (particularly in the 9-11 years of schooling category), while panel (b) shows that re-enrollment tends to decrease in work experience. Together, the figures are consistent with agents dynamically reassessing the returns to enrollment as they acquire new information based on past outcomes. The importance of such updating to the decision-making process makes a dynamic human capital investment model appropriate.

4.4 Alternative explanations of observed enrollment behavior

Several alternatives to my preferred "dynamic updating" hypothesis could explain the observed enrollment patterns, among them: 1) responses to household shocks; 2) preferences for leisure; 3) credit constraints and the need to accumulate savings; and 4) behavioral explanations. I address each of these in turn.

Responses to household shocks. Youth experiencing household shocks, such as a parent's job loss or the severe illness of a household member, may choose to change their enrollment behavior abruptly as a result. Unfortunately, the available life history data are not rich enough for me to model or otherwise rule out the role of such shocks. The most detailed data on household shocks are available only during the period of the survey waves (2002-2006), and the life history data on reasons for dropout are largely uninformative: youth citing "other/don't know/no response" as the reason for dropout swamp all other factors, as shown in panel (b) of Table 5. Data on household characteristics available throughout youth life histories include marriage, pregnancy, and co-residence with parents or grandparents, which are more properly modeled as endogenous choices rather than external shocks, and would therefore require a richer model. However, the model in its current form is consistent with the presence of household shocks, either through their effects on observable outcomes (such as grade advancement and employment) or through the unobservable (to the econometrician) state variable ϵ. Finally, Lam et al. (2010) find using CAPS data that household shocks (measured in Wave 3) do not significantly affect post-secondary enrollment, making me more confident that the exclusion of such shocks would not alter the conclusions I arrive at using more limited data.

Preferences for leisure. Another alternative explanation for the observed dropout and re-enrollment patterns is that youth prefer to take some time off from school occasionally and enjoy a period of leisure. Such a preference for leisure would be particularly likely at or near the completion of secondary school, when youth might wish to enjoy a "gap year" before continuing their studies. Yet the proportion of youth who cite economic factors (rather than other reasons more likely to signal a preference for leisure) as the reason for dropout *increases* to 23% among those with 12 years of schooling (compared to 12% for those with less than 12 years of schooling), precisely when we would expect youth to be taking a "gap year." Moreover, a sizable fraction (20%) of the sample drops out and re-enrolls prior to completing high school (see Table 2), before the typical "gap year" ages. This suggests that preferences for leisure, in the form of an extended period of non-enrollment, are not driving dropout and re-enrollment rates.

Credit constraints. If youth are credit constrained, they may need to drop out of school temporarily in order to accumulate savings, and re-enroll when they have saved enough to pay school fees. Increases in re-enrollment rates after high school completion (Figure 8) and the relatively high rates of youth from the poorest households citing school fees as their reason for dropout (Table 5, panel [c]) are consistent with credit constraints. Yet credit constraints do not appear to bind for the majority of youth: Table 5, panel (c) also shows that for the full sample, the proportion of dropouts reporting that they can not afford to attend school *falls* from 7% to 6% among high school graduates, the group most likely to face prohibitively high education expenses. Formally incorporating credit constraints would require a richer model that incorporates asset accumulation and savings, as in Keane and Wolpin (2001) and Cameron and Taber (2004). However, in the model I follow the existing literature (such as Cameron and Heckman, 2001; Carneiro and Heckman, 2002; and Lam et al., 2010) and allow household income to proxy for credit constraints. Specifically, I include indicators for lower household income quintiles and their interaction with an indicator for high school graduation; if youth from the poorest households face credit constraints that affect their enrollment choices, the coefficients on these terms will be negative. It must be noted, however, that the resulting estimates can not prove the existence of credit constraints, but may instead reflect the accumulation of educational disadvantage in low income households.

Behavioral explanations. Many behavioral explanations are possible for the observed enrollment patterns, such as myopia or asymmetric information about the relative returns to schooling versus labor market participation. Although I do not test such alternative behavioral hypotheses directly, if the model provides a good fit to the data, it would show that a model based on rational, forward-thinking, and well-informed agents is consistent with observed behavior. I turn to the results and fit of the model in the next section.

5 Results

5.1 Parameter estimates

Table 6 presents estimates of the model's structural parameters $\theta = (\phi_e, \phi_j, \alpha, \gamma)$.[38] The transition parameters for school advancement and employment (ϕ_e and ϕ_j, respectively) in columns (1) and (2) are mostly as expected, though with a few surprises. In the school advancement equation of column (1), I find that blacks, coloureds and males are significantly less likely to pass to the next grade level. The probability of passing is increasing in ability and for those who completed at least 6 years of schooling by age 12. The probability of passing declines with the level of schooling, though the separate slope and intercept terms for high school graduates are positive, reflecting the lower failure rates in post-secondary education reported in Table 5. Perhaps surprisingly, those who failed their last period enrolled are more likely to pass, though this effect is not statistically significant. Unsurprisingly, however, the probability of advancement declines with the cumulative number of grades failed since age 12.

Column (2) shows results for the labor market transition equation (9), corresponding to the parameter ϕ_j. Most coefficients have the expected sign: blacks and females are significantly less likely to find work than white males; the probability of employment is increasing in schooling, with a higher intercept (though no steeper gradient) for high school graduates. As hypothesized, employment is path dependent, with the coefficient on employment during the last period of non-enrollment quite large and precise. Surprisingly, those with at least 6 years of schooling at age 12 and more total work experience are less likely to be employed. Employment is quite sensitive to macro conditions, with a precisely estimated negative coefficient on the bad macro environment dummy.

Column (3) presents parameter estimates from the enrollment utility function of the

[38] The model converges to identical parameter estimates from arbitrary starting values.

dynamic discrete choice model, i.e., the coefficients from (8).[39] The coefficient for black shows that blacks are no more likely to enroll in school than whites, conditional on all other covariates. For coloureds, however, the coefficient is negative and significant, consistent with their shorter school careers. Coefficients on schooling at age 12 and the highest ability group are positive, consistent with greater consumption value of schooling for the more academically able. The schooling coefficient is negative, reflecting increases in both academic difficulty and opportunity costs as youth progress through school. Coefficients on recent and cumulative failure are negative, consistent with dropout after learning about one's academic ability through school outcomes. These school failure coefficients are strong evidence in favor of the hypothesis that youth dynamically update their enrollment behavior based on past schooling outcomes.

Since the school fee term in the enrollment utility function (8) accounts for the direct costs of schooling, the negative sign on the indicator for high school graduate may reflect any of several impediments to higher education: greater psychic costs, difficulty in obtaining admission, or credit constraints. The coefficients on all household income terms are negative, which may reflect the accumulated disadvantage of poverty, credit constraints in school financing, or both. The argument for credit constraints strengthens when considering the negative interaction terms on high school graduate and indicators for the first two household income quintiles, which means that conditional on high school graduation, youth from poorer households are less likely to enroll in post-secondary education, as also found by Lam et al. (2010). However, neither of these interaction terms is significant at conventional levels, weakening the case for credit constraints. Ultimately, the model can

[39] A note on the interpretation of coefficients from the enrollment equation in Table 6: as is well known, logit coefficients such as these are identified only up to scale. Because the units of the labor market utility function (as well as the school fee term in the enrollment utility function) are in South African rand per year (ten thousands, base year 2002), so too are the units of the enrollment utility coefficients, provided the scale parameter of the i.i.d. Type I extreme value shocks is unity, as assumed.

not determine whether lower post-secondary enrollment by poorer youth stems from credit constraints or accumulated disadvantage.

Coefficients for the wage equation (parameter γ) presented in column (4) generally show the expected pattern with respect to demographic, schooling and ability variables, with a large wage return to higher education evident in the high school graduate slope coefficient. A bad macro environment dampens wages, also as expected. The school fee regressions reported in Table A3 show that fees are less expensive for blacks and coloureds, likely reflecting lower school quality. School fees increase with grade level, with a large and discontinuous jump in post-secondary schooling.

5.2 Model fit

Figure 12 assesses model fit by comparing observed versus predicted grade advancement (panel [a]) and employment (panel [b]), corresponding to the transition equations (10) and (9), respectively. The model fits the data well for both types of transitions. Figure 13 plots enrollment probabilities for various state variables. The model does well in predicting enrollment by completed schooling (panel [a]), except in the final levels of post-secondary schooling. The model also does well in predicting enrollment as a function of previous school failure, reflecting students' learning about their academic ability: in panel (b), agents in the data and the model are less likely to enroll following a failed grade, which is also true in the case of cumulative failures, as in panel (c). Panels (b) and (c) show that the model captures how youth dynamically update their behavior based on past schooling outcomes.

Another important dimension of fit to consider is whether the model can replicate the patterns of dropout and re-enrollment observed in the data. Figure 14 presents observed and predicted dropout and re-enrollment rates by completed schooling. To calculate predicted values, I simulate 50 enrollment histories for each observation, using the state variable at age 12 (the first decision period) as the initial condition. The model predicts the stylized pattern of dropout and re-enrollment well, accurately capturing the rise of dropout through the end of secondary school and its subsequent fall in post-secondary. However, the predicted magnitudes are sometimes quite different from the data. The model grossly overestimates dropout and re-enrollment rates throughout the schooling distribution, likely because the model does not adequately capture the switching costs (both monetary and psychic) youth face when considering dropout or re-enrollment. Despite the relatively high re-enrollment rates observed in the data, youth exhibit more continuity in their enrollment choices than the model predicts.

5.3 Robustness checks

In considering the school to work transition of South African youth throughout their adolescence, the model of this paper treats transitions between high school and post-secondary education and transitions between schooling and labor force participation in the same framework. Although the model allows for enrollment preferences and the returns to schooling to change discontinuously at high school completion, one might still worry that school to work transitions for those with post-secondary schooling are fundamentally different in nature than for those without. To explore this possibility, I re-estimate the model using only observations from youth who do not enroll in post-secondary schooling. Comparing the results from this selected sample to those from the full sample allows one to see the extent to which the inclusion of those pursuing post-secondary schooling drives my results.

Table A4 shows results from the sample with no post-secondary schooling. Results are qualitatively similar to those from the full sample presented in Table 6. In particular, youth who never enroll in post-secondary schooling are more likely to drop out of school if they have failed their most recent grade and with cumulative grade failures, just as in the full sample. This result suggests that the process of dynamic updating of the expected relative returns to enrollment versus labor force participation is the same for these youth.

One may also be concerned that my assumptions about how forward-looking youth are in the model drives my results. In particular, one might worry that results will change dramatically depending on my choice of the discount factor β or the time horizon T. Table A5 shows that this is not the case. The table shows estimates of the enrollment utility parameter (γ) under alternative scenarios for β and T.[40] In column (1), β is set to 0.9 (from 0.95 in the main estimates). In columns (2) and (3), T is set to ages 52 and 65 (from 60 in the main estimates), respectively; $T = 52$ corresponds to South African life expectancy in 2006, while $T = 65$ corresponds to the age of public pension eligibility for males prior to a recent change in the law (and therefore may have been the retirement age males had in mind when making enrollment choices). Across all columns, coefficients are qualitatively similar to those from the main results in column (3) of Table 6.

6 Policy Simulations

In this section, I return to the motivating question of the paper: how important is the opportunity to re-enroll in the school to work transition of South African youth? To answer this question, I simulate a counterfactual scenario in which the option to re-enroll in school following dropout is restricted, and explore how enrollment decisions and outcomes change as a result. I also consider several other policy simulations that potentially alter the incentives for enrollment versus labor market participation: compulsory schooling, increased pass rates, and a subsidy for post-secondary school fees. I conclude the section by comparing results from the restricted re-enrollment simulation with those from the other

[40]Estimates of the enrollment utility parameter γ only are shown because this is the only structural parameter that depends explicitly on the discount factor β and the time horizon T in the model. All other structural parameters remain unchanged.

policies.

6.1 Restricted re-enrollment

Given the prevalence of re-enrollment among youth in the sample, it is natural to ask what role the option to re-enroll plays in decisions about schooling and labor market participation. Youth who know that they may re-enroll in school after a labor market spell may make different choices about human capital investment than those for whom labor market participation is effectively a terminal action. The model developed in this paper is well-suited to explore the ramifications of the option to re-enroll. By fully specifying the enrollment choice environment for youth and recovering its structural parameters, the model may be adapted to consider a counterfactual scenario in which the option to re-enroll is restricted.

To quantify the importance of re-enrollment, I modify the model by restricting re-enrollment prior to completion of high school, making labor market participation before completing grade 12 an absorbing state. I still allow for re-enrollment after high school completion, because restricting re-enrollment in the transition between secondary and post-secondary education seems particularly unrealistic. Although this scenario of restricted re-enrollment is admittedly extreme (and possibly unenforceable under the existing educational infrastructure in South Africa), it nonetheless provides a useful thought experiment, allowing me to assess empirically the importance of the re-enrollment option. The exercise is similar in spirit to that of Heckman and Urzua (2008), who simulate the elimination of the General Educational Development certificate (GED), a high school equivalency certificate available in the U.S. that is typically earned by high school dropouts who later re-enroll.

Specifically, I simulate 50 enrollment histories for each observation, using the state variable at age 12 as the initial condition, and eliminating the option to re-enroll before completing grade 12. Comparing simulated enrollment probabilities and completed schooling to analogous simulation results from the unrestricted model provides an assessment of how the option to re-enroll influences youth schooling choices and outcomes. The expected effects of such a policy are ambiguous: although restricting the freedom to enroll in school among a subpopulation (in this case, dropouts) ought to depress enrollment rates, potential dropouts facing no possibility of returning to school may choose to remain in school as a result, causing overall enrollment rates to rise.

Figure 15 shows enrollment and dropout probabilities from the data, the unrestricted model, and the restricted re-enrollment simulation. Enrollment rates in the restricted re-enrollment scenario exceed observed and predicted (from the unrestricted model) enrollment rates throughout most of the schooling distribution (panel [a]). The magnitude of these increases are striking: for example, 86% of those with 7 years of schooling enroll in school under the restricted re-enrollment scenario, compared to 76% in the data and 71% in the unrestricted model. Dropout rates are correspondingly lower in the restricted re-enrollment simulation (panel [b]). The intuition for these results is that youth facing restricted re-enrollment will be less willing to experiment in the labor market if the opportunity cost includes not only foregone human capital during a short labor market spell, but also foregone human capital in all future periods.

Table 7 shows the distribution of completed schooling at age 20 from the data and the restricted re-enrollment simulation. Column (1) reports the percentage of the sample in each schooling category under simulation of the unrestricted model; column (2) reports percentages under the restricted re-enrollment simulation; and column (3) reports the difference (standard errors in parentheses). The results show a clear rightward shift in the schooling distribution. The shares of the sample completing high school or attaining some post-secondary education climb by 8.3 percentage points each, with corresponding declines in the lower schooling categories. The results are qualitatively similar to those of Heckman and Urzua (2008), who use a structural schooling model to find that elimination of the GED would increase secondary completion by 2.1 percentage points in the U.S.[41] In both cases, restricting the option to re-enroll reduces dropout. While the magnitudes of the effects I find may be surprisingly large, the large share of my sample re-enrolling in school prior to completing grade 12 (20%) suggests that the restricted re-enrollment policy should indeed have a dramatic effect on the incentives faced by many youth.

A final note is in order on the welfare implications of the restricted re-enrollment simulation considered here. Restricted re-enrollment imposes costs on youth, even if it induces some to remain in school longer and complete more schooling than they otherwise would. In the unrestricted model, the possibility of re-enrollment confers option value on agents who have dropped out. Youth who choose to disenroll and re-enroll do so rationally. As forward-looking agents who compare the present value of enrollment and labor market participation in each period, their dropout and re-enrollment decisions are rational responses to shocks such as school failure and unemployment. Any restricted re-enrollment policy reduces youth welfare by shrinking their choice set (as noted in Section 2.2). From a social welfare perspective, restricted re-enrollment is efficient only if the resulting reduction in school dropout generates sufficiently high social returns, which I do not model in this paper. This same caveat about welfare implications applies to other policy simulations as well, in particular to the compulsory schooling simulation considered below.

[41] Similarly, Heckman, LaFontaine and Rodriguez (2008) find, using a panel of U.S. states, that an increase in the difficulty of passing the GED test reduced high school dropout rates.

6.2 Compulsory schooling

Schooling is compulsory in South Africa from ages 7 to 15 (or until completion of grade 9, if this occurs before age 15). Child labor laws also prohibit employment by those under age 15. In the data, however, truancy rates in ages 12-15 reach as high as 8%, with child labor rates as high as 21% among truants, as shown in Table A6. Therefore, the model of this paper assumes that neither the compulsory schooling nor child labor laws are enforced. I use the model to simulate the enforcement or extension of these laws by restricting youth to be enrolled in school until a certain age; this restriction models enforcement of both the compulsory schooling and child labor laws because youth in the model may only work if they are not enrolled in school. As before, I simulate 50 enrollment histories for each observation under both the unrestricted model and the compulsory schooling policy, using the state variable at age 12 as the initial condition. The results of the simulation, in addition to being interesting in their own right, also help to shed additional light on the importance of the option to re-enroll: making schooling compulsory over a period of time also removes the option to drop out and re-enroll during that period.

Table 8 presents results of simulations in which I make schooling compulsory through ages 15 and 16 (columns [2] and [3], respectively). Panel (a) shows the percentage of the sample in each schooling category at age 20, while panel (b) reports differences between the compulsory schooling simulation and the unrestricted model. Compulsory schooling leads to a redistribution of youth from the lower schooling categories (less than 12 years) to the higher schooling categories (12 or more years). Enforcing compulsory schooling until age 15 (the existing law), for instance, leads to a decrease of 2.5 percentage points in those completing less than 9 years of schooling, but a 0.5 percentage point increase in those with exactly 12 years of schooling and a 2.6 percentage point increase in those with more than 12 years. The effects are generally larger in magnitude when extending compulsory schooling to age 16.

The results of the compulsory schooling simulation are qualitatively similar to those of the restricted re-enrollment simulation of the preceding section: in each case, the shares of youth in the two highest schooling categories increase. These results are consistent with each other because each policy increases the cost of early dropout (or completely restricts it, in the case of compulsory schooling), changing the incentives for youth who might otherwise end their schooling careers. For both simulations, however, the restrictions come at the cost of reducing youth's choice sets, and therefore their individual welfare.

6.3 Increased pass rates

As noted in section 4, South African youth face high rates of school failure and grade repetition. Lam, Ardington and Leibbrandt (2008) document how grade failure and repetition lead to lengthy school careers in South Africa. One might therefore argue that passing thresholds are too high in South African schools, and that relaxing standards would lead to improved outcomes for youth. The model of this paper is well suited to explore such a policy, because it makes pass rates a key factor that youth consider when making enrollment choices. Using the same methodology as the previous policy simulations, I consider the effects of a 10 percentage point increase in pass rates at all levels of schooling.[42]

Table 9 presents results of the increased pass rate simulation. Not surprisingly, the distribution of completed schooling among 20 year-olds shifts to the right relative to the unrestricted model, with those completing 12 years increasing 4.3 percentage points and those with more than 12 years increasing 14 percentage points. The results are qualitatively similar to those from both the restricted re-enrollment and compulsory schooling simulations, showing that increasing the benefits of enrollment (through an increase in

[42] Pass rates are set to 100% when the increase would result in rates exceeding 100%.

school pass rates) has similar effects as increasing the costs of dropout. The increased pass rate policy has important limitations, however, because increasing pass rates without commensurate increases in (unmodeled) school quality may not translate to gains in labor market outcomes.

6.4 Post-secondary schooling subsidy

Given the sharp increase in school fees for post-secondary education (Figure 6, panel (c)), another policy simulation I consider is a subsidy for post-secondary education. Specifically, I simulate a 25% reduction in school fees for post-secondary education. As with the previous policy simulations, I report results from simulating 50 enrollment histories for each observation, using the state variable at age 12 as the initial condition, and modifying the school fee term in the enrollment utility function (8) to reflect the subsidy.

The post-secondary fee subsidy has negligible effects on the proportion of the sample enrolling in post-secondary school by age 22. Table 10 shows that in the full sample, the subsidy increases post-secondary enrollment by 0.3 percentage points, but the increase is not statistically significant. The subsidy has identical effects on youth from the first two quintiles of household incomes, the group that we would expect to respond most to the subsidy. The results suggest that fees do not explain much about post-secondary enrollment patterns, though I must note that because mine is not a formal model of asset accumulation and borrowing, the simulation does not allow me to conclude anything about the presence or absence of credit constraints.

6.5 Summary of policy simulations

In this section, I have considered a variety of policies that alter the enrollment incentives of youth: restricted re-enrollment prior to high school completion, enforcement and extension of compulsory schooling, increased pass rates, and a subsidy for post-secondary schooling. Table A7 presents a summary of results of these policy simulations (omitting the post-secondary fee subsidy, where the focus was on a different outcome). The results from each policy simulation may be compared directly to the others because I use the same random numbers to simulate the model under each policy. The policies have qualitatively similar effects: each policy increases the proportion of 20 year-olds in the sample who complete 12 or more years of schooling, while decreasing the proportion completing less than 12 years, as shown in panel (b). The restricted re-enrollment simulation has the largest effect among the policies considered on those just below high school completion (9-11 years) and those at exactly 12 years of schooling. Restricted re-enrollment also has effects that are generally larger in magnitude across the schooling distribution than both compulsory schooling simulations, suggesting that the option to re-enroll has more influence on enrollment decisions than enforcement and extension of compulsory schooling. The dramatic effect of the re-enrollment restriction demonstrates the importance of the option to re-enroll in the schooling decisions of South African youth.

7 Conclusion

In this paper, I quantify the importance of school re-enrollment in the school to work transition of South African youth. Specifically, I formulate a dynamic discrete choice model of the transition between school and work, and estimate it using a panel of young South Africans. The model accounts for uncertainty in schooling and labor market outcomes, and allows for dynamic updating of expected returns based on the outcomes of past choices. The model also allows students to re-enroll in school after a spell in the labor market, a frequent choice by South African youth, but one that is largely ignored in both the literature on South Africa and in prevailing dynamic human capital models. Each of these features of the model – uncertain outcomes, dynamic updating of expectations, and the option to re-enroll – would be missing from standard static human capital models, but are essential to understand the difficult transitions between school and work faced by South African youth.

Structural parameter estimates confirm the hypothesis that youth dynamically update their expectations about the relative returns to enrollment versus labor market participation based on schooling and labor market outcomes. Evidence of such dynamic updating is a key finding of this paper. The model replicates basic patterns of grade advancement, employment and enrollment observed in the data, according to characteristics such as completed schooling and recent and cumulative school failures. However, the model performs less well in predicting dropout and re-enrollment rates because it does not adequately account for switching costs between enrollment and labor force participation. Nonetheless, the model matches the stylized pattern of dropout and re-enrollment throughout the schooling distribution, making the model an appropriate basis for policy simulation.

I use the estimated structural parameters of the model to conduct several policy simulations that alter the incentives for enrollment versus labor market participation. In the first simulation, I restrict the option to re-enroll in school before completing grade 12, making the labor market an absorbing state for those who drop out prior to completing high school. Enrollment rates rise sharply under this restriction on re-enrollment, relative to both the data and results from the unrestricted model. The re-enrollment restriction increases the proportion of the sample completing 12 years of schooling by 8 percentage points. The effects of restricted re-enrollment are similar for all racial groups. The results show that the option to re-enroll is an important component of the incentives South African youth face when making schooling decisions.

Simulations of the enforcement (and extension) of compulsory schooling and an increase in school pass rates show qualitatively similar effects as the restricted re-enrollment policy: in all cases, the proportion of the sample completing less than 12 years of schooling falls, while the proportion completing 12 or more years increases. Simulation of a 25% fee subsidy for post-secondary education shows no significant effects on the proportion of the sample enrolling in post-secondary schooling, however. The results suggest increasing the opportunity cost of school dropout (as in the restricted re-enrollment and compulsory schooling simulations) or raising the expected benefits of enrollment (as in the increased pass rate simulation) can have dramatic effects on youth enrollment decisions. Youth are less responsive to changes in the direct costs of schooling, however.

The model of this paper is quite general, and allows for straightforward extensions such as job search or employment while enrolled, or choice of educational path (e.g., academic versus vocational), that I may explore in future work. Its emphasis on uncertainty, academic and labor market shocks, and the option to re-enroll is appropriate for the South African context, where high grade repetition and unemployment make the school to work transition especially difficult. Nonetheless, the model may be relevant for other contexts as well, not only for other developing countries where youth face similarly difficult circumstances, but also developed countries in which educational choices for dropouts are an increasingly important part of human capital investment. In the United States, for instance, the prevalence of high school dropouts obtaining the GED or mid-career workers enrolling in community or for-profit colleges makes the option to re-enroll an important consideration for students.

The key findings of the model – that youth update their expected returns and enrollment decisions based on past schooling and labor market outcomes; that youth alter their enrollment behavior when the option to re-enroll after dropout is restricted; and that youth respond similarly to other changes in the costs and benefits of enrollment – suggest that a dynamic framework is essential for understanding the school to work transition in South Africa.

References

Altonji, Joseph G. "The Demand for and Return to Education When Education Outcomes are Uncertain." *Journal of Labor Economics*, Vol. 11, No. 1, Part 1: Essays in Honor of Jacob Mincer (Jan., 1993), pp. 48-83.

Arcidiacono, Peter. "Ability sorting and the returns to college major." *Journal of Econometrics* 121 (2004): 343375.

Banerjee, Abhijit, Sebastian Galiani, James Levinsohn, Zoë McLaren and Ingrid Woolard. "Why Has Unemployment Risen in the New South Africa?" NBER Working Paper Series No. 13167, June 2007.

Becker, Gary S. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: Columbia University Press, 1964.

Belzil, Christian and Jorgen Hansen. "Unobserved Ability and the Return to Schooling." *Econometrica*, Vol. 70, No. 5 (Sep., 2002), pp. 2075-2091.

Cameron, Stephen V. and James J. Heckman. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *The Journal of Political Economy*, Vol. 109, No. 3 (Jun., 2001), pp. 455-499.

Cameron, Stephen V. and Christopher Taber. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *Journal of Political Economy*, 2004, vol. 112, no. 1: 132-182.

Card, David. "The Causal Effect of Education on Earnings." *Handbook of Labor Economics*, Volume 3, Chapter 30. Edited by O. Ashenfelter and D. Card. Elsevier, 1999.

Carneiro, Pedro, and James J. Heckman, 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal* 112(482), 705-734.

Case, Anne. "Does Money Protect Health Status? Evidence from South African Pensions." Mimeo, Princeton University 2001.

Case, Anne and Angus Deaton. "School Inputs and Educational Outcomes in South Africa." *Quarterly Journal of Economics*, Vol. 114, No. 3 (Aug., 1999), pp. 1047-1084.

Case, Anne and Motohiro Yogo. "Does School Quality Matter? Returns to Education and the Characteristics of Schools in South Africa." National Bureau of Economic Research

Working Paper No. 7399, October 1999.

Eckstein, Zvi and Kenneth I. Wolpin. "Academic and Labor Market Success: The Impact of Student Employment, Abilities, and Preferences." *Econometrica*, Vol. 67, No. 6, (Nov., 1999), pp. 1295-1339.

Fiske, Edward B. and Helen F. Ladd. *Elusive Equity: Education Reform in Post-Apartheid South Africa.* Washington DC: Brookings Institution Press, 2004.

Heckman, James J., Paul A. LaFontaine and Pedro L. Rodriguez. "Taking the Easy Way Out: How the GED Testing Program Induces Students to Drop Out." IZA Discussion Paper No. 3495, May 2008.

Heckman, James J. and Salvador Navarro. "Dynamic discrete choice and dynamic treatment effects." *Journal of Econometrics* 136 (2007): 341396.

Heckman, James J. and Sergio Urzua. "The Option Value of Educational Choices and the Rate of Return to Educational Choices." Institute for Computational Economics Lecture, University of Chicago, August 1, 2008.

International Labour Organization. *Global Employment Trends for Youth.* Geneva: International Labour Office, 2008.

Joensen, Juanna Schrøter. "Academic and Labor Market Success: The Impact of Student Employment, Abilities, and Preferences." Mimeo, 2008.

Keane, Michael P. and Kenneth I. Wolpin. "The Career Decisions of Young Men." *Journal of Political Economy*, Vol. 105, No. 3, (Jun., 1997), pp. 473-522.

Keane, Michael P. and Kenneth I. Wolpin. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, Vol. 42, No. 4 (Nov., 2001), pp. 1051-1103.

Lam, David, Murray Leibbrandt and Cecil Mlatsheni. (2008a). "Education and Youth Unemployment in South Africa." Mimeo, 2008.

Lam, David, Cally Ardington and Murray Leibbrandt. (2008b). "Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa." University of Michigan Population Studies Center Research Report 08-632, February 2008.

Lam, David, Cally Ardington, Nicola Branson, Anne Case, Murray Leibbrandt, Alicia

Menendez, Jeremy Seekings and Meredith Sparks. (2008c). "The Cape Area Panel Study: A Very Short Introduction to the Integrated Waves 1-2-3-4 Data." The University of Cape Town, October 2008.

Lam, David, Cally Ardington, Nicola Branson, Kendra Goostrey, and Murray Leibbrandt. "Credit Constraints and the Racial Gap in Post-Secondary Education in South Africa." Mimeo, 2010.

Levinsohn, James and Todd Pugatch. "The Role of Reservation Wages in Youth Unemployment in South Africa: A Structural Approach." University of Michigan, mimeo, 2010.

Light, Audrey. "Hazard model estimates of the decision to reenroll in school." *Labour Economics* 2 (1996) 381-406.

Rust, John. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica*, Vol. 55, No. 5 (Sep., 1987), pp. 999-1033.

Stange, Kevin. "An Empirical Examination of the Option Value of College Enrollment." Mimeo, 2007.

Yamauchi, Futoshi. "Race, equity, and public schools in post-Apartheid South Africa: Equal opportunity for all kids." *Economics of Education Review* 24 (2005): 213233.

A Why students drop out (in context of re-enrollment model of Section 2.2)

The re-enrollment model of Section 2.2 conditions on agents already participating in the labor market. Here, I consider how such initial dropout and labor market participation was a rational choice in the context of the model. As in the model, ignore time discounting, assume that all agents have zero skill units ($s = 0$) at $t = 0$, and that one skill unit is the maximum possible. Further assume that agents with $s = 1$ obtain employment at wage w_1 with certainty. Agents with 0 skill units who were not employed in the previous period find a job with probability $p \in (0, 1)$. Let q denote the probability of passing the grade and obtaining one skill unit.

At $t = 0$, an agent will drop out and enter the labor market if:

$$\begin{aligned} \mathbb{E}_0[U(\text{labor market})] &> \mathbb{E}_0[U(\text{enroll})] \Leftrightarrow \\ pw_0 + \mathbb{E}_1(\max\{\text{labor market,re-enroll}\}|s = 0) &> q\mathbb{E}_1(\text{work}|s = 1) \\ &\quad + (1 - q)\mathbb{E}_1(\max\{\text{labor market,re-enroll}\}|s = 0) \end{aligned} \tag{22}$$

Denote $\mathbb{E}_1(\max\{\text{labor market,re-enroll}\}|s = 0) \equiv \mathbb{E}max$ to conserve notation. Noting that $\mathbb{E}_1(\text{work}|s = 1) = 2w_1$, we have:

$$\begin{aligned} \mathbb{E}_0[U(\text{labor market})] &> \mathbb{E}_0[U(\text{enroll})] \Leftrightarrow \\ pw_0 + \mathbb{E}max &> q(2w_1) + (1 - q)\mathbb{E}max \Leftrightarrow \\ pw_0 &> q(2w_1 - \mathbb{E}max) \Leftrightarrow \\ \frac{pw_0}{2w_1 - \mathbb{E}max} &\equiv \underline{q}_d > q \end{aligned} \tag{23}$$

The decision to drop out may be summarized by a threshold value of q, denoted $\underline{q}_d$, such that academic success is sufficiently unlikely relative to the expected gains from working. (This is analogous to analysis of the re-enrollment decision, which also depended on threshold values for q.) Here, the option value of re-enrollment $\mathbb{E}max$ also plays a role: the greater this option value, the greater must be the probability of advancement q to justify remaining enrolled in school.

B Education in South Africa

The South African education system consists of General Education (grades 1-9), Further Education (grades 10-12) and Higher Education (grades 13 and above). Education is compulsory for youth aged 7-15 or through completion of grade 9 (whichever comes first), but I treat this regulation as non-binding due to the presence of non-enrolled youth in this age range in the data. To enter post-secondary schooling, students must pass a nationally standardized "matric" exam at the completion of grade 12. Post-secondary education includes both academic and vocational programs ("Technikons"). Additionally, students at the secondary level may enroll in vocational "National Training Certificate" (NTC) programs.[43] For simplicity, I do not distinguish between academic and vocational education in the model. Private schools serve less than 3% of the student population in South Africa, according to government statistics; in the Western Cape province the figure is 3.1% (Fiske and Ladd, 2004).

The government subsidizes public education, but students must pay fees to attend, and such fees vary considerably among schools. Public schools are self-governing, and are free to set their own admissions policies and fees (subject to provincial government approval). Although admissions can not discriminate based on race, test scores, or ability to pay fees, prevailing patterns of residential segregation serve to maintain quality differences among schools. Moreover, despite legal prohibitions, Fiske and Ladd (2004) conclude that "there is little doubt that many schools consider a family's likely ability to pay their fee when making admissions policy" (p. 143). Although low-income families may qualify for fee exemptions under a policy adopted in 1998, only 2.5% of primary school students and 3.7% of secondary school students receive the exemption (these figures rise to 4.1% and 5.7%, respectively, in historically white schools; Fiske and Ladd, 2004). At the post-secondary level, the South African government offers subsidized loans to qualified students who pass a means test through the National Student Financial Aid Scheme (NSFAS), and individual institutions also offer financing. Private banks also offer student loans at market interest rates.

C Data Definitions

This section discusses variable definitions and sample selection criteria. The data comes from retrospective life history data collected in Wave 1 of CAPS, augmented with life events recorded in Waves 2-4.[44] The Wave 1 retrospective life histories record events by youth's age, where age refers to the age at which the event occurred in the case of living arrangements and marriage, and to age at the beginning of the calendar year in the case of enrollment and progression through school, labor force participation, and pregnancy. I

[43]See Appendix C for information on mapping NTC programs to grade levels.

[44]The Cape Area Panel Study Waves 1-2-3 were collected between 2002 and 2005 by the University of Cape Town and the University of Michigan, with funding provided by the US National Institute for Child Health and Human Development and the Andrew W. Mellon Foundation. Wave 4 was collected in 2006 by the University of Cape Town, University of Michigan and Princeton University. Major funding for Wave 4 was provided by the National Institute on Aging through a grant to Princeton University, in addition to funding provided by NICHD through the University of Michigan.

follow this convention in mapping Wave 2-4 responses to youth's age.

Schooling level covers grades 1-16, with National Training Certificate (NTC) I, II and III mapped to grades 10, 11, and 12, respectively.[45] Students enrolled in university or Technikon programs that include grade 12 are considered enrolled in grade 12. Reporting successful completion of the grade level or reporting enrollment in a higher grade level in a subsequent year is considered passing the level for grades 1-12. Beginning at grade 13, reporting successful completion of the grade level or "no grade/continuing" are considered passing the level, up to a maximum of 16 years completed schooling. This distinction is made because "no grade/continuing" is the modal response for those enrolled in the post-secondary education sector, indicating that most youth are continuing in their programs of higher education, whereas "passing" reports at these levels drop considerably. Unfortunately, I am unable to determine whether students are making satisfactory progress towards degree completion. All other results while enrolled are considered failure. I define "dropout" as disenrollment following a year of enrollment, and "re-enrollment" as enrollment following a year of non-enrollment. Schooling histories in which levels regress with age are re-coded so that such regression can not occur. Grades failed represent the accumulation of periods of enrollment in which the agent did not pass the grade, and therefore may include events such as withdrawal, illness or residential moves rather than outright academic failure.

Labor force participation variables (i.e., work and search) and wages are conditional on non-enrollment at a given age, where reporting enrollment supersedes reports of labor market participation. School fees are conditional on enrollment, and include total household expenditure on fees and other educational expenses, in real rand per year (base year 2002). Wages are full-time annual equivalent based on 160 working hours per month (those reporting monthly hours above 160 are considered full-time and do not receive an adjustment). Wages and school fees are available only at the time of the interview, rather than as retrospective histories; predicted values are imputed based on observed characteristics for purposes of estimation. Work experience includes only those periods of simultaneous work and non-enrollment; I exclude work experience while enrolled in school.[46]

I make several sample restrictions. I keep only those observed until at least age 18. Those who report advancing two or more grades in a year, or without continuous information on enrollment, are dropped from the sample. I drop those who report entering school prior to age 4 or exiting school after age 24 (which effectively sets $T_d = 24$ as the decision horizon). I also drop those whose educational histories, by our definitions above, place them with more than 16 years of completed schooling. The restriction on observing a person until at least age 18 accounts for more than 60% of those dropped from the sample.

Other covariates are largely self-explanatory. Ability quartiles refers to in-sample rank of age-adjusted score on the literacy and numeracy evaluation (LNE) administered to all

[45] NTC conversion based on coding in CAPS, derived variable $w1h\_higrd$.

[46] Work or search while enrolled never exceeds 3% of the sample at any grade level, and never exceeds 2% during grades 1-12.

CAPS respondents in Wave 1. Unfortunately, because this ability measure was taken in Wave 1, when the sampled youth were at least age 14, it is not predetermined with respect to enrollment choices in the model, which begins at age 12. I include it, however, because it is a measure common to all in the sample, and therefore helps to distinguish between the role of ability and human capital investment in labor market returns. To mitigate bias in the LNE score due to age differences in Wave 1, I adjust for age as follows: using the estimation sample, I regress the standardized literacy and numeracy evaluation (LNE) score on age and age squared at Wave 1 (when the test was administered) and get predicted residuals. I then sort observations into quartiles based on these residuals. Household income quintiles are derived from the distribution of household per capita income reported in Wave 1 of CAPS.[47]

I calibrate the wage-age profile for years following the decision horizon (i.e., from periods $T_d + 1$ to T) using the 10% public use micro-sample of the 2001 South African Census. First, I define the estimation sample as native-born residents of urban areas in Western Cape province (which includes metropolitan Cape Town, from which CAPS respondents are drawn) who are ages 25-64, in the labor force, and classify themselves as one of the three major racial groups (white, black, or coloured). I also exclude the self-employed and unpaid workers, leaving a sample of $n = 111,772$. I then predict employment and income by running logit and OLS regressions, respectively, using as controls race and gender dummies, years of schooling, a high school graduate dummy, race-schooling and race-high school graduate interactions, age and age squared.[48] I then create an expected income variable for each observation as the product of these predicted values (i.e., $\hat{\mathbb{E}}(\text{income}) = \hat{\Pr}(\text{work}) \times \hat{\text{income}}$), and regress expected income on the same set of controls. I save the coefficients on age and age squared from this regression for use in the wage equation of the structural model; the coefficients on age and age squared are 3,695.3 and -45.9, respectively.[49] The macro environment variable is based on the South African employment/population ratio for 15-24 year olds, from the World Bank Africa Development Indicators.

All variables measured in monetary values used in this paper are in real South African rand per year (base year 2002), unless otherwise noted. The South African rand traded at 10.3 per US dollar in August, 2002 when CAPS Wave 1 began.

[47] Due to non-response, 7% of the sample uses imputed values for household income, based on multiple imputation conducted by CAPS.

[48] The included controls are the largest subset (other than age variables) of the controls used in the structural model that are available in the Census. Note that interactions of schooling and the high school graduate dummy are not included because the maximum years of schooling reported in the Census is 13. I use income rather than wages because the latter are unavailable in the Census. I adjust income to 2002 South African rand to be consistent with the base year of CAPS.

[49] Since I care only about the coefficients, inconsistent estimation of their standard errors due to the generated outcome variable is not problematic in this context.

Table 1: Elements of observable state space **X** and exclusion restrictions

Variable	**Description**	**enrollment**	**passing**	**wage/ employment**
Time-invariant (X_0):				
$race$	$\mathbb{I}$(race=r) for r=black,coloured	x	x	x
g	$\mathbb{I}$(female)	x	x	x
s_0	$\mathbb{I}$(schooling$\geq$ 6 at $t=0$)	x	x	x
z_q	$\mathbb{I}$(ability quartile=q)	x	x	x
y_q	$\mathbb{I}$(household income quintile=q) for q =1,2	x		
Time-varying (X_t):				
s	schooling	x	x	x
f_{tot}	total grades failed since $t=0$	x	x	
f	$\mathbb{I}$(failed last grade enrolled)	x	x	
x_{tot}	work experience			x
x	$\mathbb{I}$(worked last period not enrolled)			x
HSG	$\mathbb{I}$(high school graduate)	x	x	
$(s-12)*HSG$	(schooling-12)*$\mathbb{I}$(high school graduate)	x	x	x
y_q*HSG	interactions between y_q and HSG	x		
ζ	$\mathbb{I}$(bad macro environment)†			x

†A bad macro environment is a year in which the employment/population ratio for 15-24 year olds is less than 25%. Between 1991-2006, this occurred in 2002-2003.
Note: "x" denotes inclusion in equation, by column. Enrollment refers to equation (8); passing refers to (10); wage/employment refers to (7) and (9).

Table 2: Summary statistics

Variable	**Obs.**	**Mean**	**Std. Dev.**
female	3,806	0.54	0.50
black	3,806	0.28	0.45
coloured	3,806	0.55	0.50
white	3,806	0.17	0.38
age of school entry	3,806	6.2	1.1
age last enrolled	3,806	17.9	2.3
completed schooling	3,806	10.9	2.2
grades failed	3,806	1.2	1.3
schooling at age 12	3,806	5.4	1.2
LNE score	3,806	0.23	0.98
ever re-enrolled	3,806	0.33	0.47
ever re-enrolled, grade 1-12	3,806	0.20	0.40
passed matric (high school)	3,806	0.50	0.50
ever enrolled in post-secondary	3,806	0.27	0.44
ever worked	3,806	0.51	0.50
work experience	3,806	1.1	1.3
wage (FTE, max.)	2,708	40,287	45,002

LNE score is from literacy and numeracy evaluation administered in CAPS Wave 1, normalized to mean zero and standard deviation one. Wage is annual full-time equivalent based on 160 monthly hours, maximum over all individual's observations in panel, in South African rand (real, base year 2002). Survey weights applied.

Table 3: Education and labor market outcomes, by selected characteristics

	school entry age	**age last enrolled**	**completed schooling**	**grades failed**	**ever worked**	**ever re-enrolled**	**ever re-enrolled (pre-12)**	**wage (FTE, max.)**
race								
black	6.9	18.5	10.3	1.7	0.31	0.39	0.31	22,453
coloured	6.0	17.2	10.6	1.2	0.64	0.27	0.16	36,125
white	5.9	19.3	12.8	0.5	0.40	0.43	0.14	70,412
gender								
male	6.3	17.8	10.6	1.4	0.55	0.33	0.20	42,551
female	6.2	18.0	11.1	1.1	0.47	0.34	0.20	38,067
ability								
low	6.5	17.3	9.8	1.7	0.49	0.28	0.22	27,616
high	5.9	18.5	11.9	0.8	0.53	0.39	0.18	51,082
schooling								
< 9	6.9	15.3	7.0	2.2	0.48	0.08	0.07	24,604
9-11	6.4	17.6	10.2	1.8	0.47	0.30	0.30	29,497
12	5.9	18.1	12.0	0.6	0.63	0.37	0.22	42,094
> 12	5.8	20.3	13.7	0.3	0.44	0.54	0.07	65,453

All values are survey-weighted means. Ability based on age-adjusted median from literacy and numeracy evaluation administered in CAPS Wave 1. Ever re-enrolled pre-12 refers to re-enrollment before completing grade 12. Wage is annual full-time equivalent based on 160 monthly hours, maximum over all individual's observations in panel, in South African rand (real, base year 2002).

Table 4: Age last enrolled if completed 8-9 years schooling, by race and re-enrollment status

Age last enrolled	**Full sample**	**Black**	**Coloured**	**White**
Panel (a): full sample				
≤ 16	54.7	23.5	71.0	58.9
17-18	34.4	47.3	27.5	41.1
19+	11.0	29.2	1.5	0.0
Total	100.0	100.0	100.0	100.0
Panel (b): ever re-enrolled				
≤ 16	16.3	4.9	33.5	N/A
17-18	56.7	51.7	64.2	N/A
19+	27.0	43.4	2.3	N/A
Total	100.0	100.0	100.0	N/A

Cells report survey-weighted fraction of racial group last enrolled in each age category, among those with 8-9 years of completed schooling. Panel (a) is full sample, Panel (b) restricts sample to those who ever dropped out and re-enrolled in school.

Table 5: Reasons for failure, dropout and rates of school non-affordability

	Full sample	**Completed schooling**	
		$<$ 12 years	**12+ years**
Panel (a): Reason for failure			
Fail (any reason)	0.14	0.15	0.08
academic failure	0.07	0.08	0.01
withdrew	0.03	0.03	0.03
no grade/continued	0.03	0.04	0.00
unspecified	0.01	0.00	0.05
Panel (b): Reason for dropout			
economic	0.15	0.12	0.20
academic/behavioral	0.03	0.03	0.03
health (self)	0.07	0.07	0.07
family	0.03	0.04	0.01
pregnancy or baby	0.02	0.03	0.01
pregnancy or baby (females only)	0.04	0.06	0.02
other/dk/no response	0.73	0.75	0.69
Panel (c): Can't afford school (among dropouts)			
Full sample	0.07	0.07	0.06
Household income quintile			
first (bottom)	0.14	0.12	0.22
second	0.09	0.08	0.13
third	0.06	0.05	0.07
fourth	0.04	0.04	0.03
fifth (top)	0.01	0.01	0.01

Cells report survey-weighted fraction of observations in person-year panel, among school enrollees (reasons for failure) and dropouts (reasons for dropout and can't afford school). "No grade/continued" not considered failure for completed schooling of 12+ years. Reasons for non-enrollment classified as follows: "economic" includes work, search, or can't afford school; "academic/behavioral" includes passed matric or completed all offered schooling, failed, expelled, school isn't important, or imprisoned; "family" includes married, pregnant, moved or caring for ill relative; "other/dk/no response" includes non-response and school closed. Household income quintile refers to per capita household income, Wave 1.

Table 6: Structural parameter estimates

	(1)	(2)	(3)	(4)
equation	pass	work	enrollment	wage
parameters	ϕ_e	ϕ_j	α	γ
black	-0.30	-0.99	0.21	-3.02
	(0.11)	(0.12)	(0.55)	(0.30)
coloured	-0.55	-0.14	-4.51	-2.40
	(0.10)	(0.12)	(0.49)	(0.31)
female	0.27	-0.37	0.78	-0.55
	(0.04)	(0.05)	(0.24)	(0.10)
$\mathbb{I}$(schooling$\geq$ 6, age 12)	0.58	-0.17	2.26	0.28
	(0.05)	(0.06)	(0.29)	(0.08)
ability quartile 2	0.49	-0.04	0.37	-0.04
	(0.05)	(0.07)	(0.32)	(0.07)
ability quartile 3	0.95	0.10	-0.07	0.17
	(0.06)	(0.08)	(0.35)	(0.09)
ability quartile 4	1.56	0.11	1.85	0.66
	(0.09)	(0.10)	(0.44)	(0.14)
HH income quintile 1			-1.37	
			(0.34)	
HH income quintile 2			-1.43	
			(0.34)	
schooling	-0.39	0.20	-0.23	0.13
	(0.01)	(0.02)	(0.04)	(0.02)
high school graduate	1.19	0.14	-16.31	-0.06
	(0.15)	(0.08)	(1.29)	(0.14)
(schooling-12)*	0.69	-0.09	-4.52	1.76
high school graduate	(0.16)	(0.08)	(1.17)	(0.42)
failed last grade	0.09		-6.73	
enrolled	(0.08)		(0.69)	
cumulative grades failed	-0.60		-3.04	
since age 12	(0.05)		(0.56)	
worked last period		3.08		0.07
not enrolled		(0.08)		(0.14)
work experience		-0.57		0.19
		(0.04)		(0.03)
HH income quintile 1*HSG			-1.28	
			(0.85)	
HH income quintile 2*HSG			-0.25	
			(0.17)	
$\mathbb{I}$(bad macro environment)		-1.61		-0.94
		(0.07)		(0.11)
constant	4.38	-2.33	11.32	3.70
	(0.14)	(0.22)	(0.71)	(0.38)
N	24,679	13,747	3,806	4,370
R^2				0.30
$\ln L$	-736,735.0	-532,639.9	-3,138,189.6	

Unit of observation in columns (1), (2) and (4) is the person-year, from ages 12-26. Unit of observation in column (3) is person. Columns (1), (2) and (3) are logits (column (3) includes full solution to dynamic programming problem); columns (4) is OLS regression. All estimates use survey weights. All standard errors robust. Wages are full-time equivalent, based on 160 hours of work per month, in tens of thousands of real South African rand per year (constant 2002 values). Model units for enrollment equation same as wage equation. Cumulative failure since age 12 top-coded at 3. Work experience top-coded at 4. "Bad macro environment" refers to employment/population ratio for 15-24 year olds below 25 percent.

Table 7: Distribution of completed schooling at age 20 under restricted re-enrollment (before high school completion) simulation

Completed schooling	unrestricted (1)	restricted re-enrollment (2)	difference (2)-(1)
< 9 years	15.6	11.6	-4.0
	(0.1)	(0.1)	(0.1)
$9 - 11$ years	46.8	34.2	-12.6
	(0.2)	(0.1)	(0.2)
12 years	18.4	26.7	8.3
	(0.1)	(0.1)	(0.2)
> 12 years	19.2	27.5	8.3
	(0.1)	(0.1)	(0.2)

Table shows distribution of completed schooling at age 20 in unrestricted model and under "no re-enrollment before high school completion" simulation. Percentage of sample within each completed schooling category shown (standard error in parenthesis). In simulation, 50 simulated histories are generated for each observation. Column (1) shows simulation results of unrestricted model. Column (2) shows simulation results in which re-enrollment restricted prior to high school completion. Column (3) shows difference is means, Column (1) minus Column (2). All results calculated using survey weights.

Table 8: Distribution of completed schooling at age 20 under compulsory schooling simulation

Final compulsory schooling age	**unrestricted**	**age 15**	**age 16**
	(1)	(2)	(3)
Panel (a): percent of sample in schooling category			
< 9 years	15.6	13.1	11.3
	(0.1)	(0.1)	(0.1)
9 − 11 years	46.8	46.3	46.4
	(0.2)	(0.2)	(0.2)
12 years	18.4	18.9	18.7
	(0.1)	(0.1)	(0.1)
> 12 years	19.2	21.7	23.7
	(0.1)	(0.1)	(0.1)
Panel (b): Difference (compulsory minus unrestricted)			
< 9 years		-2.5	-4.4
		(0.1)	(0.1)
9 − 11 years		-0.5	-0.4
		(0.2)	(0.2)
12 years		0.5	0.3
		(0.2)	(0.2)
> 12 years		2.6	4.5
		(0.2)	(0.2)

Table shows distribution of completed schooling at age 20 in unrestricted model and under compulsory schooling simulation. Percentage of sample within each completed schooling category (Panel [a]) or difference between compulsory schooling simulation and unrestricted model (Panel [b]) shown (standard error in parenthesis). In simulation, 50 simulated histories are generated for each observation. Column (1) shows simulation results of unrestricted model. Columns (2)-(3) shows simulation results in which schooling made compulsory through ages 15 and 16, respectively. All results calculated using survey weights.

Table 9: Distribution of completed schooling at age 20 under increased pass rate (10 percentage points) simulation

Completed schooling	**unrestricted** (1)	**increased pass rate** (2)	**difference** (2)-(1)
<9 years	15.6	5.1	-10.5
	(0.1)	(0.1)	(0.1)
$9-11$ years	46.8	39.0	-7.8
	(0.2)	(0.1)	(0.2)
12 years	18.4	22.7	4.3
	(0.1)	(0.1)	(0.2)
>12 years	19.2	33.2	14.0
	(0.1)	(0.2)	(0.2)

Table shows distribution of completed schooling at age 20 in unrestricted model and under increased pass rate simulation. Percentage of sample within each completed schooling category shown (standard error in parenthesis). In simulation, 50 simulated histories are generated for each observation. Column (1) shows simulation results of unrestricted model. Column (2) shows simulation results in which pass rates increased 10 percentage points at all grade levels (or set to 100% if initial pass rate exceeds 90%). Column (3) shows difference is means, Column (1) minus Column (2). All results calculated using survey weights.

Table 10: Proportion enrolling in post-secondary schooling by age 22 under post-secondary fee subsidy

	Predicted (no subsidy) (1)	**Predicted (25% subsidy)** (2)	**Difference** (2)-(1)
full sample	44.2	44.4	0.3
	(0.2)	(0.2)	(0.3)
household income quintile			
first (bottom)	24.0	24.4	0.3
	(0.3)	(0.3)	(0.5)
second	27.5	27.9	0.3
	(0.4)	(0.4)	(0.5)

Table shows percentage of sample enrolling in post-secondary education by age 22, by indicated characteristic, based on model simulation. In simulation, 50 simulated histories are generated for each observation. Column (1) shows simulation results of unrestricted model. Column (2) shows simulation results in which post-secondary fees reduced by 25%. Column (3) shows difference is means, Column (2) minus Column (1). All results calculated using survey weights. All results calculated using survey weights.

Table A1: Panel balance

		Proportion censored at age			
	N	18	20	22	24
full sample	3,806	0.00	0.27	0.58	0.83
black	1,714	0.00	0.26	0.55	0.81
coloured	1,660	0.00	0.27	0.57	0.82
white	432	0.00	0.31	0.65	0.92

Cells show number of observations (N) or percent of sample with missing enrollment information by age. Survey weights used in calculation.

Table A2: Mean spell duration from monthly calendar data

Type of spell	**all spells**	**full-time only**	**complete spells only**
school	20.0	14.3	10.2
labor force participation	14.1	14.4	7.2
work	13.8	13.7	7.5

Table shows mean spell duration, in months, by type of spell. Sample monthly calendar reports from estimation sample, CAPS Waves 1-4 (August 2002-December 2006). Labor force participation includes work or search. "Full-time only" refers to spells in which no other concurrent activity was reported. Survey weights used in calculation.

Table A3: Predicted school fees

black	-0.69
	(0.06)
coloured	-0.62
	(0.06)
schooling	0.06
	(0.01)
high school graduate	0.64
	(0.06)
constant	0.24
	(0.10)
N	3138
R^2	0.40

Table shows coefficients in school fee prediction equation. Unit of observation is person-year, where school fees measured in South African rand per year (ten thousands). In enrollment utility (8), school fee set to first percentile of school fee distribution if predicted school fee falls below minimum observed school fee in sample. Survey weights used in regression.

Table A4: Structural parameter estimates (no post-secondary schooling)

	(1)	(2)	(3)	(4)
equation	pass	work	enrollment	wage
parameters	ϕ_e	ϕ_j	α	γ
black	-0.11	-1.06	-3.69	-2.23
	(0.14)	(0.16)	(0.88)	(0.25)
coloured	-0.38	-0.19	-7.13	-1.63
	(0.13)	(0.15)	(0.79)	(0.25)
female	0.26	-0.37	0.19	-0.55
	(0.05)	(0.06)	(0.40)	(0.06)
$\mathbb{I}$(schooling$\geq$ 6, age 12)	0.46	-0.17	2.23	0.23
	(0.05)	(0.06)	(0.31)	(0.07)
ability quartile 2	0.40	-0.18	0.96	0.02
	(0.06)	(0.08)	(0.38)	(0.07)
ability quartile 3	0.75	-0.04	0.54	0.07
	(0.07)	(0.09)	(0.42)	(0.08)
ability quartile 4	1.27	0.10	1.00	0.41
	(0.08)	(0.09)	(0.47)	(0.11)
HH income quintile 1			-1.11	
			(0.47)	
HH income quintile 2			-1.26	
			(0.46)	
schooling	-0.43	0.20	0.05	0.15
	(0.01)	(0.02)	(0.96)	(0.02)
high school graduate		0.01		-0.04
		(0.09)		(0.11)
failed last grade	0.14		-6.59	
enrolled	(0.08)		(1.63)	
cumulative grades failed	-0.54		-3.09	
since age 12	(0.05)		(0.76)	
worked last period		3.17		-0.11
not enrolled		(0.09)		(0.08)
work experience		-0.61		0.20
		(0.04)		(0.03)
$\mathbb{I}$(bad macro environment)		-1.39		-0.84
		(0.08)		(0.07)
constant	4.33	-2.28	12.22	2.73
	(0.17)	(0.24)	(9.74)	(0.30)
N	17,845	11,963	2,979	3,728
R^2				0.21
$\ln L$	-588,726.6	-438,327.3	-1,421,917.0	

Unit of observation in columns (1), (2) and (4) is the person-year, from ages 12-26. Unit of observation in column (3) is person. Columns (1), (2) and (3) are logits (column (3) includes full solution to dynamic programming problem); columns (4) is OLS regression. All estimates use survey weights. All standard errors robust. Wages are full-time equivalent, based on 160 hours of work per month, in tens of thousands of real South African rand per year (constant 2002 values). Model units for enrollment equation same as wage equation. Cumulative failure since age 12 top-coded at 3. Work experience top-coded at 4. "Bad macro environment" refers to employment/population ratio for 15-24 year olds below 25 percent. Sample excludes those who enroll in post-secondary schooling.

Table A5: Enrollment parameter estimates (robustness checks)

	(1) $\beta = .9$	(2) $T = 52$	(3) $T = 65$
black	-0.58	0.36	0.15
	(0.32)	(0.53)	(0.67)
coloured	-3.18	-4.05	-4.72
	(0.29)	(0.46)	(0.54)
female	0.19	0.68	0.82
	(0.14)	(0.23)	(0.26)
$\mathbb{I}$(schooling$\geq$ 6, age 12)	1.08	1.99	2.38
	(0.17)	(0.28)	(0.34)
ability quartile 2	0.27	0.35	0.38
	(0.19)	(0.31)	(0.34)
ability quartile 3	0.13	-0.08	-0.06
	(0.20)	(0.33)	(0.39)
ability quartile 4	1.31	1.60	1.95
	(0.26)	(0.42)	(0.48)
HH income quintile 1	-0.89	-1.28	-1.42
	(0.20)	(0.33)	(0.37)
HH income quintile 2	-0.89	-1.34	-1.48
	(0.20)	(0.32)	(0.36)
schooling	0.03	-0.15	-0.26
	(0.03)	(0.10)	(0.23)
high school graduate	-11.34	-15.24	-16.79
	(0.70)	(0.88)	(3.60)
(schooling-12)*	-0.52	-3.88	-4.80
high school graduate	(0.71)	(1.09)	(2.26)
failed last grade	-6.01	-6.64	-6.77
enrolled	(0.41)	(0.67)	(1.16)
cumulative grades failed	-1.55	-2.76	-3.16
since age 12	(0.21)	(0.49)	(0.97)
HH income quintile 1*HSG	-0.70	-1.26	-1.29
	(0.56)	(0.77)	(1.13)
HH income quintile 2*HSG	-0.16	-0.36	-0.20
	(0.17)	(0.18)	(0.27)
constant	6.87	10.07	11.87
	(0.42)	(0.73)	(2.88)
N	3,806	3,806	3,806
$\ln L$	-2,193,919.5	-2,936,402.3	-3,229,274.8

Table shows enrollment utility parameter (γ) estimates under modifications to model, as indicated by column. All estimates use survey weights. All standard errors robust. Model units scaled to 10,000 South African rand per year (constant 2002 values). Survey weights used in all estimates.

Table A6: Truancy and child labor

Age	**Truancy**	**Employment (among truants)**
12	0.01	0.05
13	0.01	0.09
14	0.02	0.21
15	0.08	0.08

Table shows proportion of sample who are truants (not enrolled in school before completing grade 9) and employment rate among truants, by age. Survey weights used in calculation.

Table A7: Policy simulation results summary

Simulation	**unrestricted**	**no re-enrollment before high school completion**	**compulsory schooling (age 15)**	**compulsory schooling (age 16)**	**increased pass rate (10 p.p.)**
completed schooling	(1)	(2)	(3)	(4)	(5)
Panel (a): percent of sample in schooling category					
<9 years	15.6	11.6	13.1	11.3	5.1
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
$9-11$ years	46.8	34.2	46.3	46.4	39.0
	(0.2)	(0.1)	(0.2)	(0.2)	(0.1)
12 years	18.4	26.7	18.9	18.7	22.7
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
>12 years	19.2	27.5	21.7	23.7	33.2
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
Panel (b): Difference (policy minus unrestricted model)					
<9 years		-4.0	-2.5	-4.4	-10.5
		(0.1)	(0.1)	(0.1)	(0.1)
$9-11$ years		-12.6	-0.5	-0.4	-7.8
		(0.2)	(0.2)	(0.2)	(0.2)
12 years		8.3	0.5	0.3	4.3
		(0.2)	(0.2)	(0.2)	(0.2)
>12 years		8.3	2.6	4.5	14.0
		(0.2)	(0.2)	(0.2)	(0.2)

Table shows distribution of completed schooling at age 20 in unrestricted model and policy simulations. Percentage of sample within each completed schooling category (Panel [a]) or difference between indicated policy simulation and unrestricted model (Panel [b]) shown (standard error in parenthesis). In simulation, 50 simulated histories are generated for each observation. Column (1) shows simulation results of unrestricted model. Column (2)-(5) shows policy simulation results in which re-enrollment restricted before high school completion (column [2]); school enrollment compulsory until age 15 or 16 (columns [3]-[4], respectively); and grade advancement rates increased 10 percentage points (column [5]). All results calculated using survey weights.

Figure 1: Youth employment/population in South Africa and selected countries

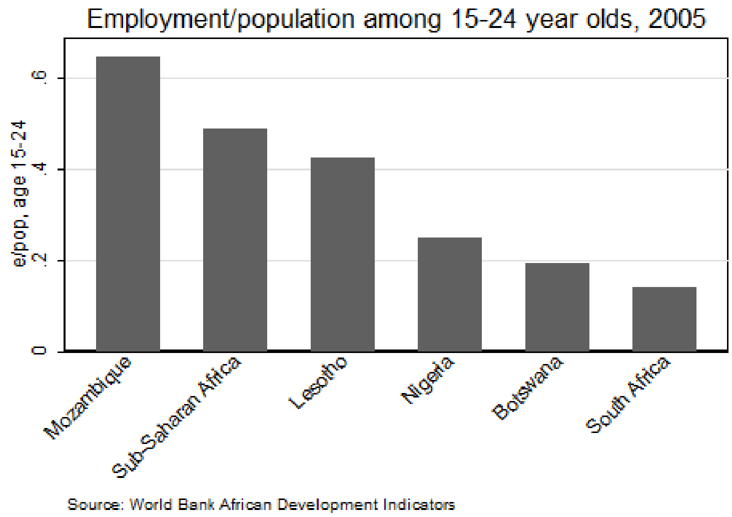


Figure 2: K-10 enrollment at age 20 in South Africa and selected countries

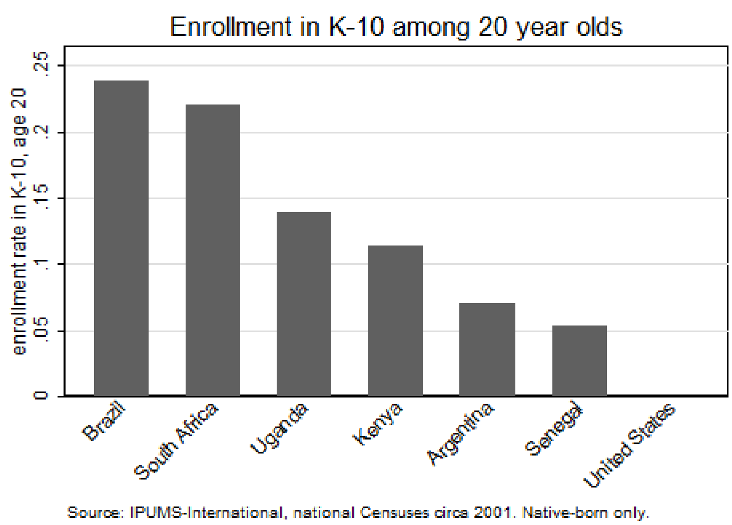
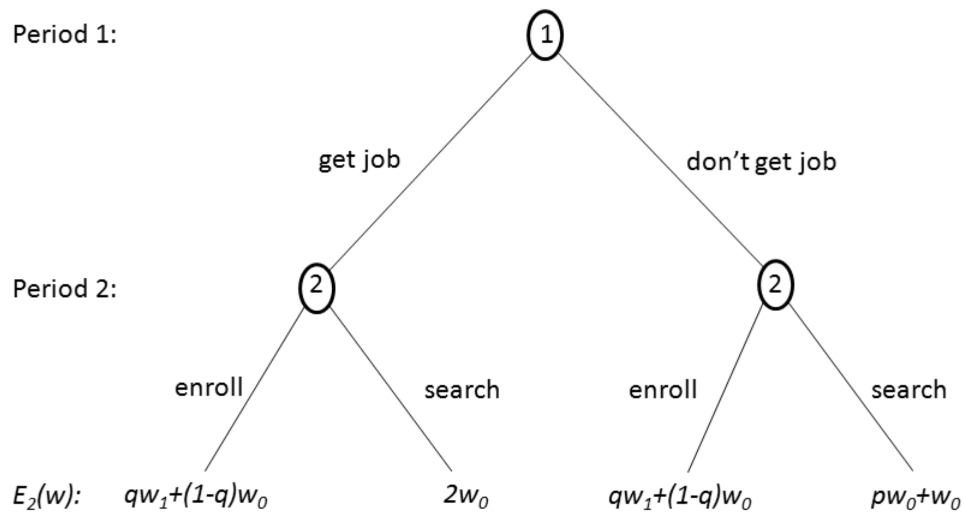


Figure 3: Simple model of re-enrollment



Numeral within each node refers to time period of model. Agents begin in labor force in period 1. Agents then choose to enroll in school or remain in labor force in period 2. All agents are in the labor force in period 3, with guaranteed employment. Payoffs at bottom refer to expected payoffs at outset of period 2.

Figure 4: Model timing, choices and payoffs

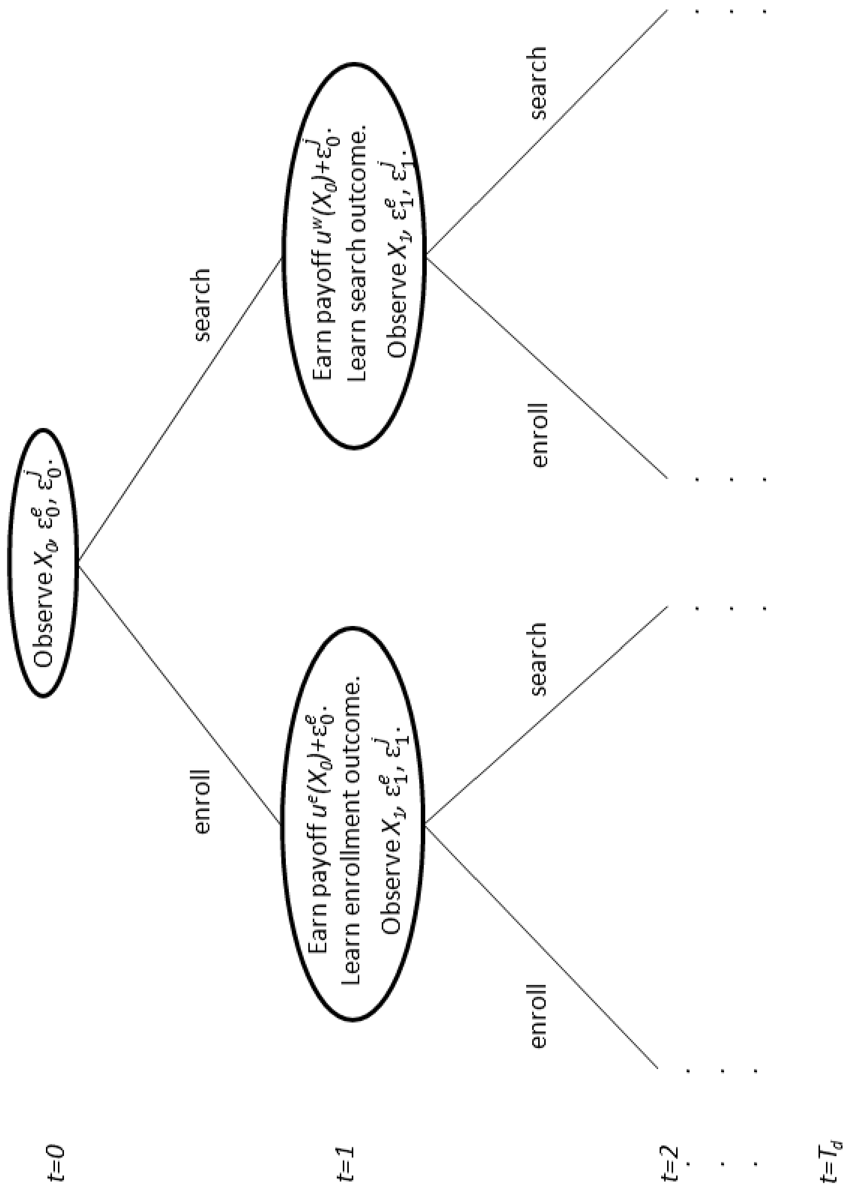


Figure 5: Schooling spell birth and death

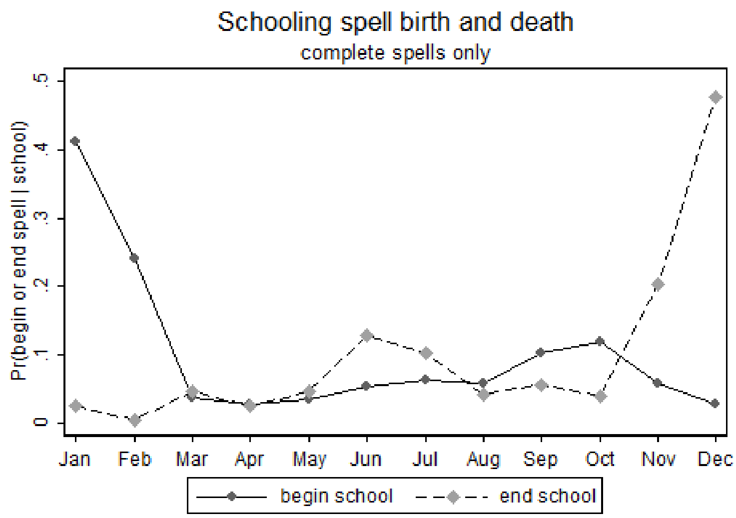
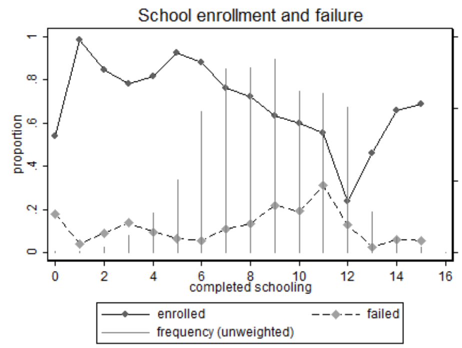
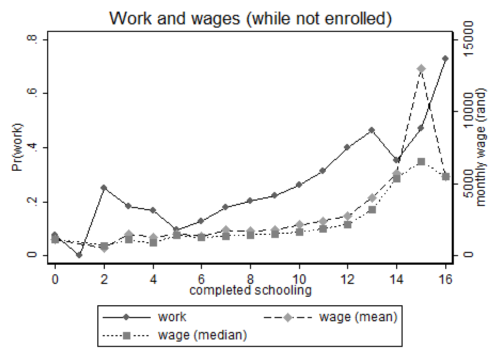


Figure shows proportion of students beginning or ending a schooling spell in each month, using CAPS monthly calendar data, Waves 1-4 (August 2002-December 2006). Censored spells excluded.

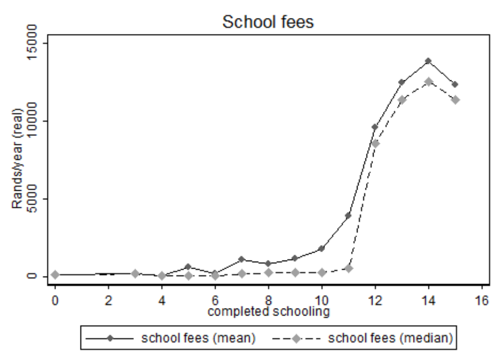
Figure 6: Enrollment, employment and school fees



(a)



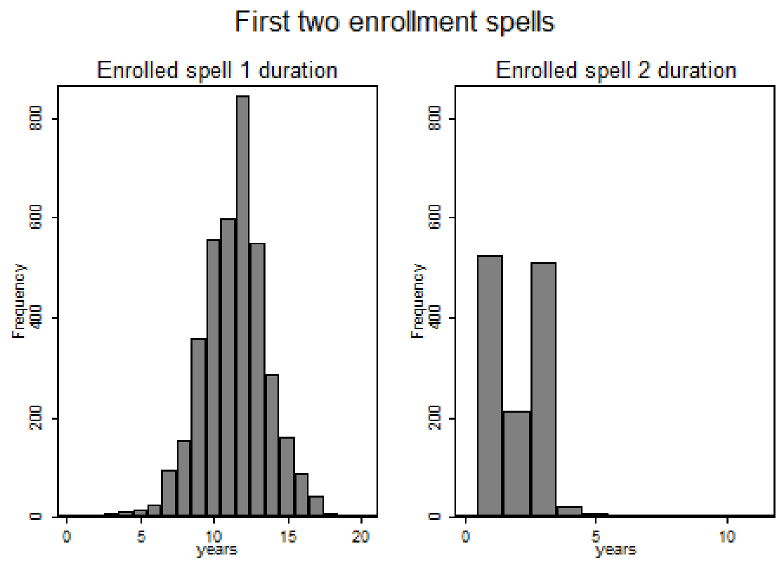
(b)



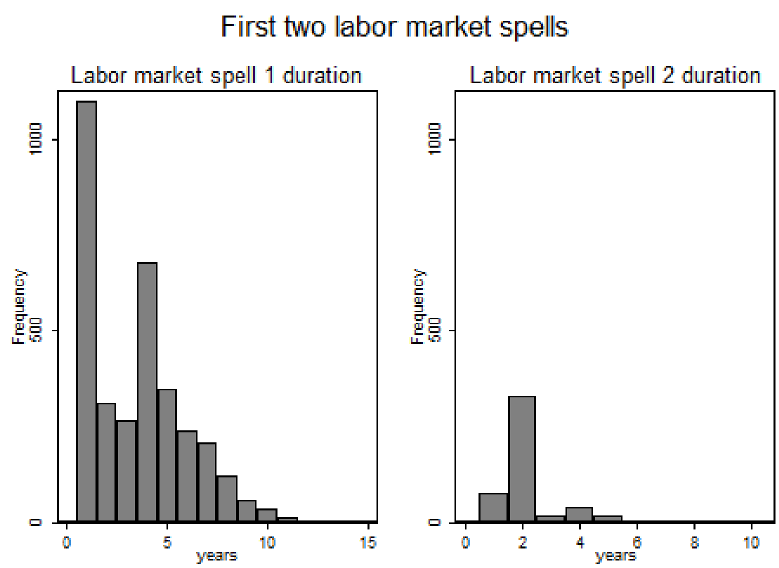
(c)

Figure shows labeled characteristics among person-year observations in sample, conditional on completed schooling, i.e., proportion enrolled in school among person-year observations with *s* years completed schooling. Failure rates and school fees condition on current school enrollment. Employment rates condition on current non-enrollment. Wages condition on current employment. Additionally, spikes in panel (a) are histogram of person-year observations by completed schooling (unweighted frequencies).

Figure 7: Duration of first two enrollment and labor market spells



(a)



(b)

Figure shows durations of first two school enrollment and labor market spells. Enrollment and labor market participation are exhaustive and mutually exclusive states, i.e., labor market spells are defined as periods of non-enrollment. Observations are considered enrolled when both enrollment and labor market participation reported.

Figure 8: Re-enrollment and dropout

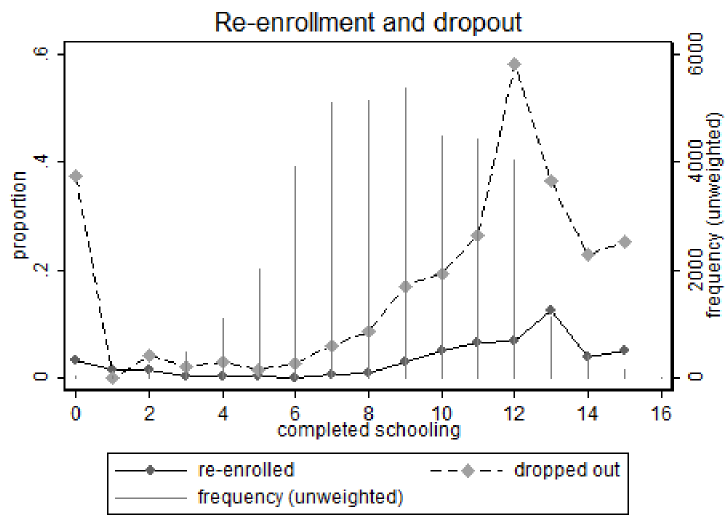
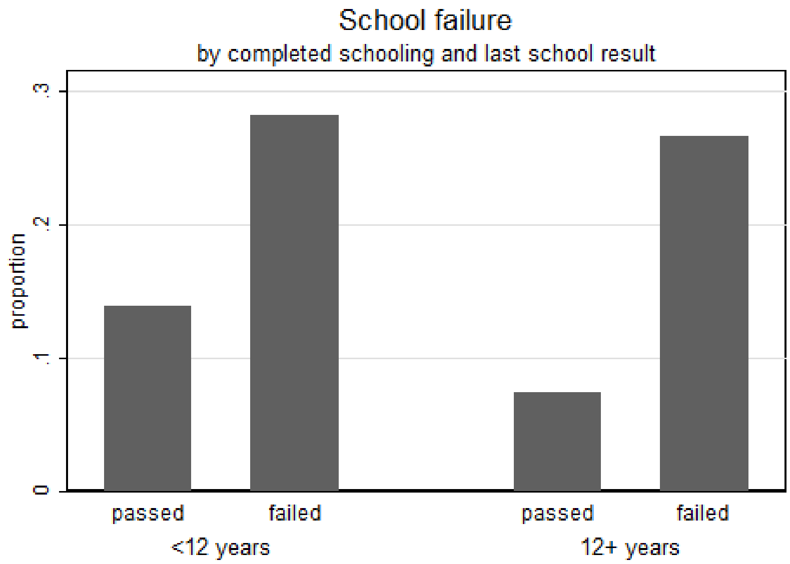
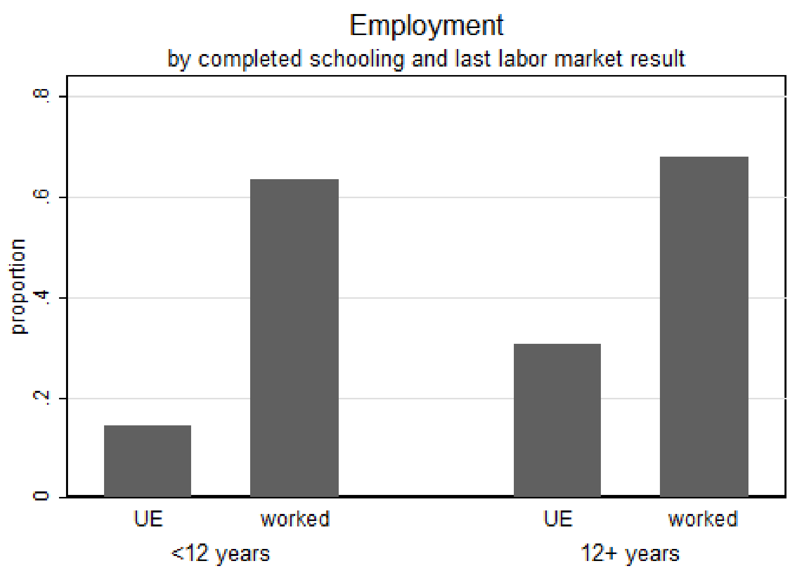


Figure shows re-enrollment and dropout among eligible person-year observations in sample, conditional on completed schooling. Re-enrollment refers to enrollment after at least one period of disenrollment, so that figure depicts Pr(enroll at t|not enrolled at $t-1$, schooling $= s$). Dropout refers to disenrollment after a period of enrollment, i.e., Pr(not enrolled at t|enrolled at $t-1$, schooling $= s$). Additionally, spikes are histogram of all person-year observations by completed schooling (unweighted frequencies).

Figure 9: Failure and employment rates, by last outcome



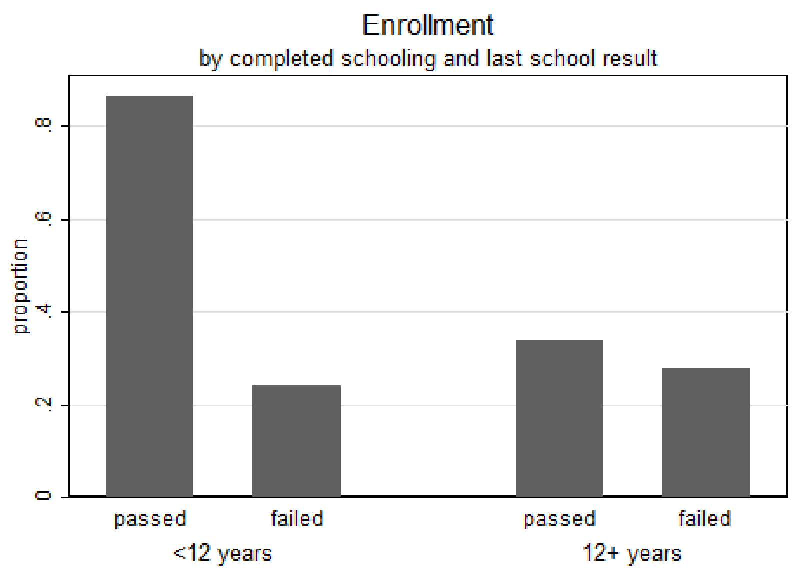
(a)



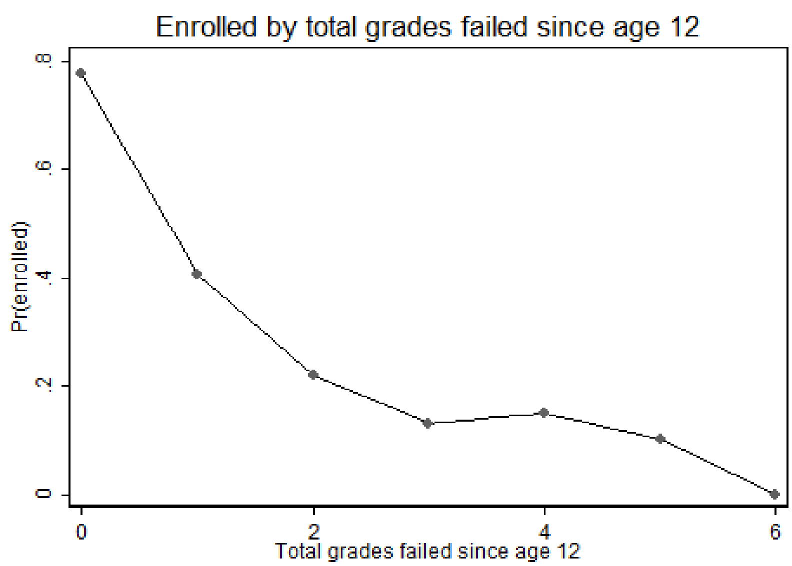
(b)

Panel (a) shows proportion of enrollees (person-year observations) who fail current grade, conditional on schooling outcome of last enrollment period and completed schooling category. Panel (b) shows employment rates among non-enrollees (person-year observations), conditional on labor market outcome of last non-enrollment period and completed schooling category.

Figure 10: Enrollment, by school outcomes



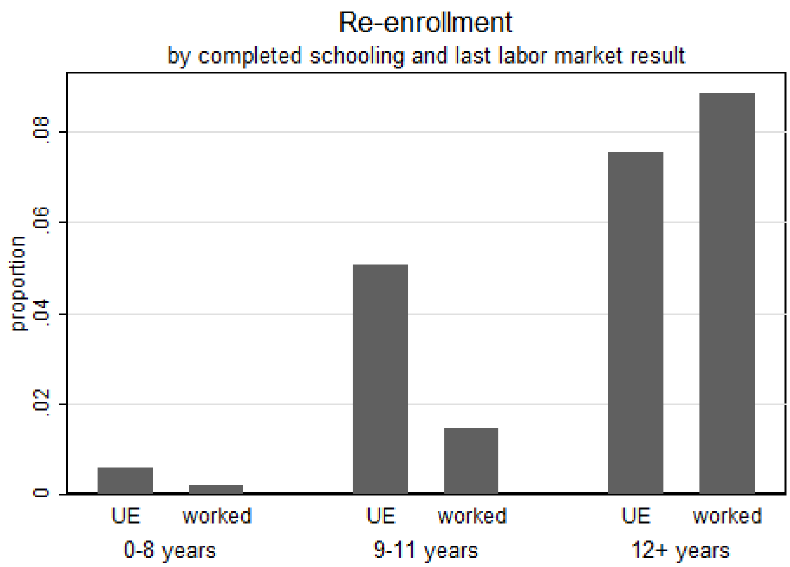
(a)



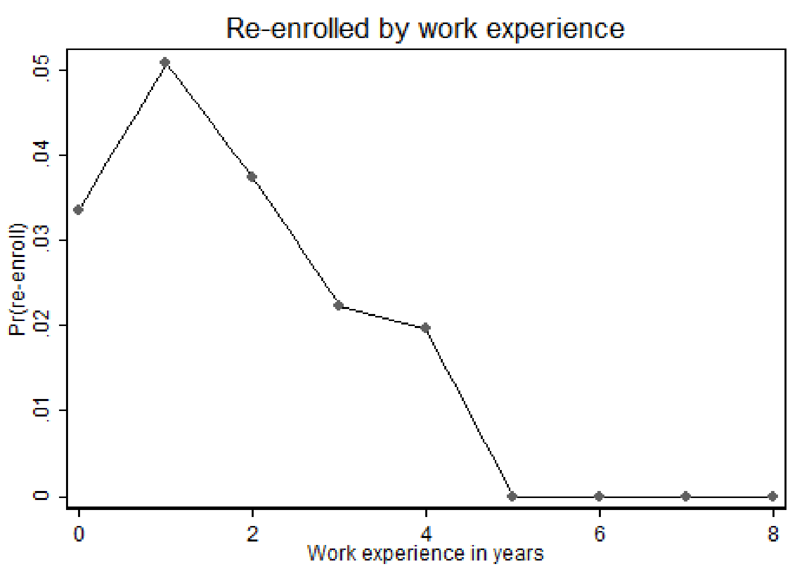
(b)

Figure shows proportion of person-year observations who enroll, conditional on schooling outcome of last enrollment period and completed schooling category (Panel [a]) or cumulative grades failed since age 12 (Panel [b]).

Figure 11: Re-enrolled, by labor market outcomes



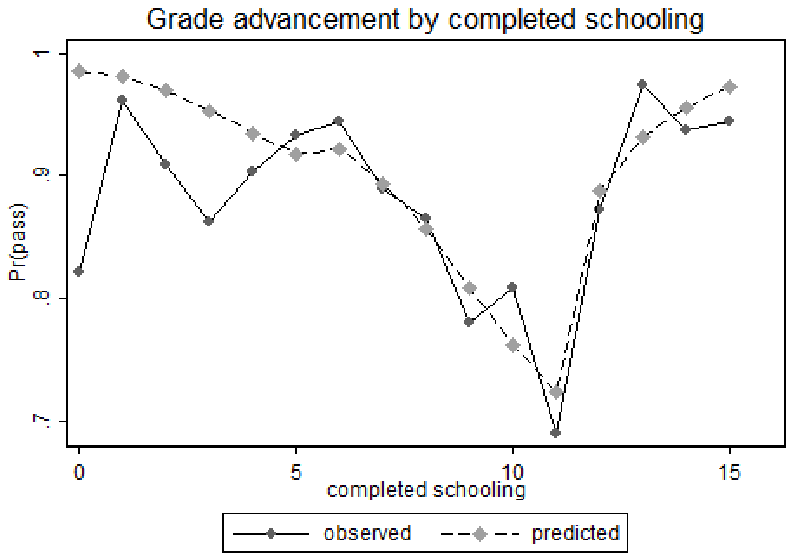
(a)



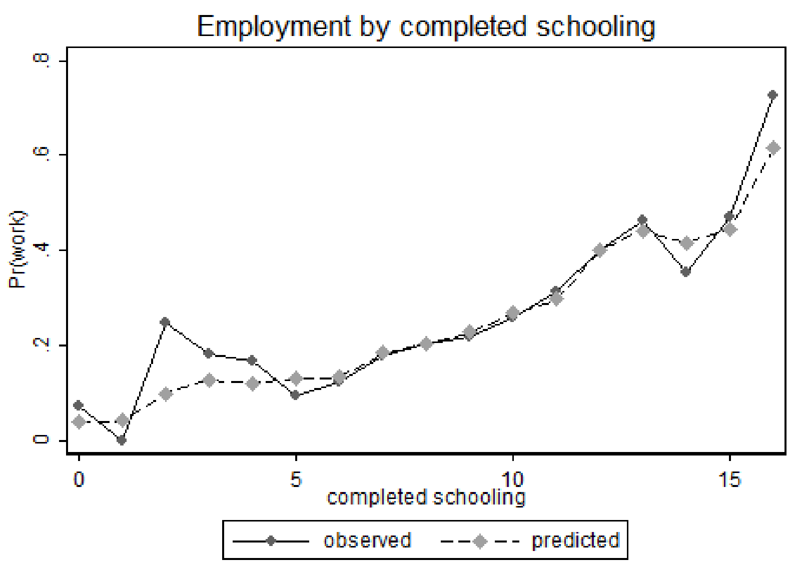
(b)

Figure shows proportion of person-year observations who re-enroll (Pr(enroll at t|not enrolled at $t-1$), conditional on labor market outcome of last non-enrollment period and completed schooling category (Panel [a]) or work experience (Panel [b]).

Figure 12: Model fit: grade advancement and employment



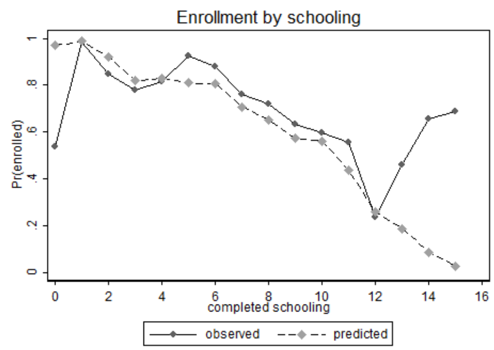
(a)



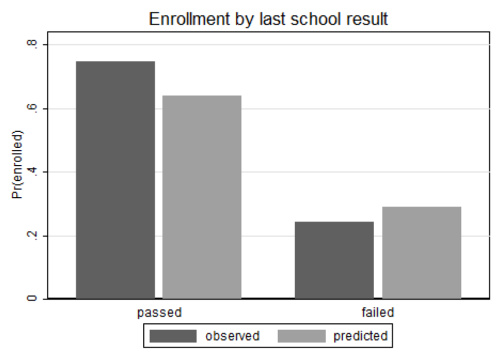
(b)

Figure shows labeled characteristics among person-year observations in sample, conditional on completed schooling, i.e., proportion advancing to next grade among person-year observations with s years completed schooling. Grade advancement rates condition on current school enrollment. Employment rates condition on current non-enrollment. Predicted grade advancement and employment calculated as $\int \Pr(y|S,\hat{\theta})f(S)dS$ for $y =$ pass,work, using observed distribution of states S and survey weights.

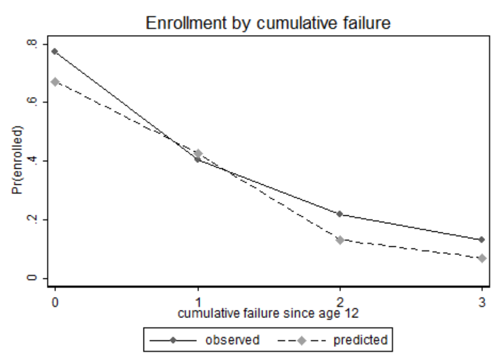
Figure 13: Model fit: enrollment



(a)



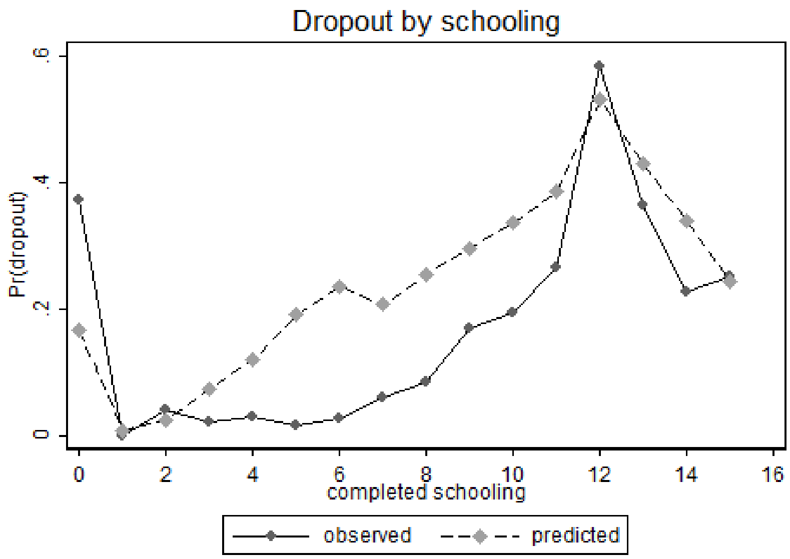
(b)



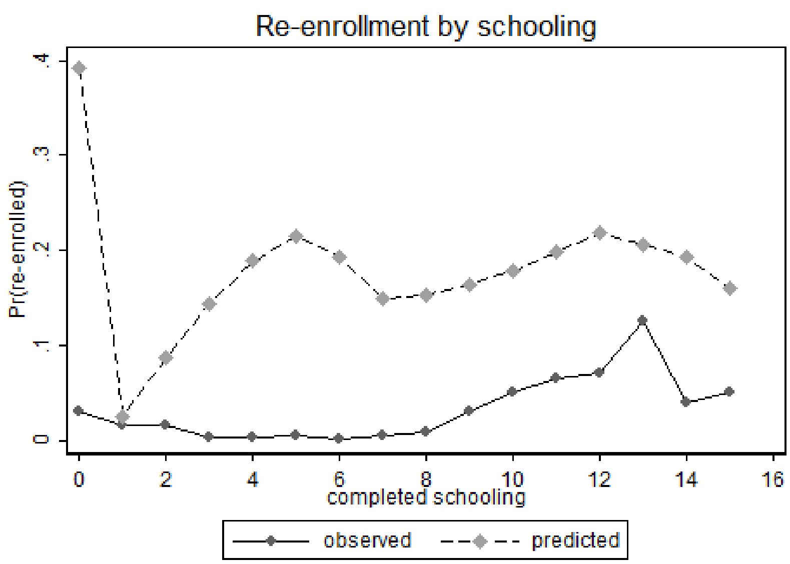
(c)

Figure shows enrollment rates among person-year observations in sample, conditional on labeled characteristics, i.e., proportion enrolled in school among person-year observations with s years completed schooling. Predicted enrollment calculated as $\int \Pr(d|S, \hat{\theta}) f(S) dS$, using observed distribution of states S and survey weights.

Figure 14: Model fit: dropout and re-enrollment



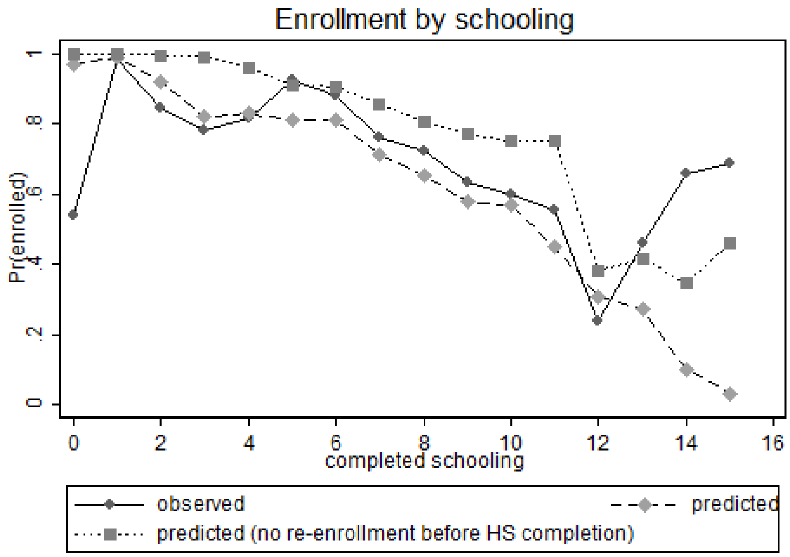
(a)



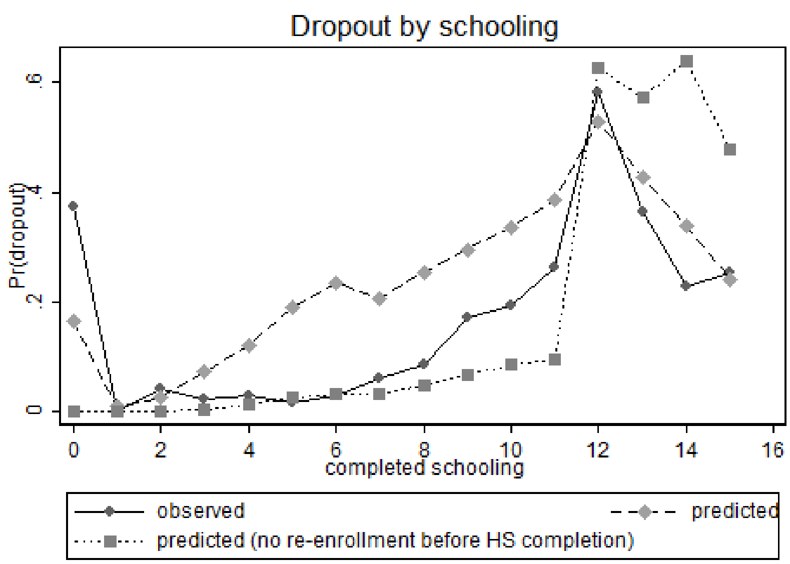
(b)

Figure shows dropout and re-enrollment among eligible person-year observations in sample, conditional on completed schooling. Dropout refers to disenrollment after a period of enrollment, i.e., Pr(not enrolled at $t|$enrolled at $t-1$, schooling $= s$). Re-enrollment refers to enrollment after at least one period of disenrollment, so that figure depicts Pr(enroll at $t|$not enrolled at $t-1$, schooling $= s$). Predicted values calculated as $\int \Pr(y|S,\hat{\theta})f(S)dS$ from simulation in which enrollment history simulated 50 times for each observation in the dataset, using the observed state at $t = 1$ (age 12) as the initial condition. Predicted dropout and re-enrollment are undefined for age 12 because no simulated prior history exists at that age.

Figure 15: Restricted re-enrollment simulation



(a)



(b)

Figure shows enrollment and dropout among person-year observations in sample, conditional on completed schooling. Dropout refers to disenrollment after a period of enrollment, i.e., Pr(not enrolled at t|enrolled at $t-1$, schooling $= s$). Predicted values calculated as $\int \Pr(y|S,\hat{\theta})f(S)dS$ from simulation in which enrollment history simulated 50 times for each observation in the dataset under "no re-enrollment before high school completion" restriction, using the observed state at $t = 1$ (age 12) as the initial condition. Predicted dropout is undefined for age 12 because no simulated prior history exists at that age.