An Empirical Examination of the Option Value of College Enrollment

Kevin Stange[1]

University of California Berkeley

Job Market Paper

September 2007

Abstract

This paper quantifies the option value that arises when human capital investment is made sequentially in the presence of uncertainty. If enrollment in college reveals information about the relative desirability of schooling and labor market entry, and if individuals have the opportunity to abandon investment after acquiring this information, then enrollment has option value. This provides an ex-ante incentive to make investments (such as attending college for only one year) that may turn out to be suboptimal ex-post. This paper is the first to estimate the magnitude of this option value, which is also shown to have implications for welfare and education policy. The parameters of a dynamic structural model of college enrollment and completion are estimated using detailed transcript data on a panel of US youth. Option value is computed by comparing the welfare predicted by the dynamic model with that predicted by the counterfactual scenario where individuals are forced to commit to an educational outcome before enrolling in college. Estimates suggest that option value is substantial, particularly for moderate-ability individuals with the greatest uncertainty about the value of schooling. This flexibility is valued at about \$15,000 on average in 1992 dollars and up to \$27,000 for students closer to the enrollment margin. This represents 13% of the total value of the college enrollment opportunity for the average high school graduate and up to 34% for moderate-ability students. Consequently, option value is pivotal to many individuals' enrollment decisions and important to welfare. More than half of this value is due to information learned in the first year of college. Policy simulations reveal that structural changes - introducing community colleges or improving academic preparation - generally have a greater impact on educational outcomes than traditional financial incentives. Collectively, these findings suggest that uncertainty and option value are central features of educational investment and should be considered in future estimates of the returns to schooling and when evaluating education policy.

[1]Department of Economics, University of California Berkeley. Email: kstange@econ.berkeley.edu.
Website: https://webfiles.berkeley.edu/~kstange/

1 Introduction

"The most valuable of all capital is that invested in human beings."

- Alfred Marshall, *Principles of Economics*, 1890 p. 592

"I wanted to know myself before my parents were spending $30,000 on
an education...I wanted to know what I wanted to do with my life, and I didn't."

- Student explaining why she switched to community college
after starting at a private university
(reported in *The New York Times*, April 23, 2006)

It has long been recognized that education is the most important investment many people make in their lifetime. Beginning with the pioneering work of Becker, Mincer, and others, there is an enormous literature that applies the concepts and tools from investment theory to the study of individuals' education decisions. One area in which the education literature continues to lag behind the investment literature is in the treatment of uncertainty.[2] Like developing a new drug or drilling an oil well, investment in education occurs in the face of much uncertainty about its costs and benefits. Consequently, educational expectations seldom match eventual outcomes. For instance, only 51% of 1982 high school seniors who intended to earn a Bachelor's degree had done so by 1992, while 16% of those planning not to earn a four-year degree eventually do.[3] These discrepancies are not surprising when one considers just how much high school and college environments can differ. Expectations about the difficulty of college coursework, tastes for various subjects, and career ambitions may be very different from the reality students discover once they arrive on campus. This was certainly the experience of the student quoted above. She saw community college as a low-cost environment in which to resolve uncertainty about her educational ambitions.

The standard model of educational investment used in most economic research abstracts from this uncertainty. Educational investment is typically modeled as a static optimization problem where people select the lifetime utility-maximizing level with certainty. The process whereby educational intentions are converted into actual educational outcomes is typically treated as a black box. Furthermore, the

[2] Most relavent to this paper is the investment theory related to real options collected in Dixit and Pindyck (1994).
[3] National Center for Educational Statistics, *2004 Digest of Educational Statistics* Table 307.

static model suggests that few people should dropout if psychic costs are smooth and wage returns are nonlinear, as they appear to be. High rates of college dropout, combined with anecdotes such as those above, motivate a schooling determination model where uncertainty takes center stage.

While there has been some attention paid to the role of labor market uncertainty in educational decisions, relatively few studies have incorporated uncertainty about schooling completion and the option value it creates.[4] Weisbrod (1962) was the first to point out that education has option value due to uncertainty and the sequential nature of investment. If parents overlook this option value, he concluded, "there would be a tendency towards under-investment in education." Comay, Melnick, and Pollatschek (1973, 1976) extend this analysis by contrasting the optimal decision rules derived in the presence of uncertainty (and option value) with those derived in a static framework. They conclude that drop-out risk should be an important factor in educational decisions, a fact the static model ignores. Dothan and Williams (1981) develop a general valuation equation for the option value of education in the presence of uncertainty about the labor market, schooling costs, and preferences. Comparative statics imply that education is more valuable with greater uncertainty about ability, preferences, and future employment opportunities. They conclude that models that focus only on "the mean returns to education are unlikely to measure accurately the true value of education to the individual." Manski (1989) discusses one important normative implication of uncertainty: dropout may be socially desirable if it encourages marginal students to experiment with college. Though each of these studies present a sequential model of educational investment in the presence of uncertainty, none attempt to quantify their empirical importance using micro data.

The present study is most closely related to the empirical work of Altonji (1993), Cunha, Heckman, and Navarro (2005), and Arcidiacono (2004, 2005), as well as the dynamic structural models of Keane and Wolpin (1997). Altonji (1993) finds large differences between ex-ante and ex-post returns to starting college, suggesting that option value may be important. Cunha, Heckman, and Navarro (2005) decompose wage variability into heterogeneity, which is known when schooling decisions are made, and uncertainty, which is not. They conclude that uncertainty is empirically important - approximately thirty percent of people would change their schooling decisions if they had perfect information. Arcidiacono (2005) has a similar dynamic set-up to the present study, but is primarily concerned with

[4] The theoretical implications of labor market uncertainty for educational decisions were first derived in Levhari and Weiss (1974), Olson, White, and Shefrin (1979), and Williams (1979). Several notable empirical studies include Buchinsky and Leslie (2000), Chen (2007), and Wiswall (2005).

policy simulation. He demonstrates that many educational policies - such as affirmative action - have important dynamic effects that are difficult to model in a static setting.

This paper is also similar to the dynamic structural models of Keane and Wolpin (1997) and Cameron and Heckman (2001), but with one key difference. These models permit individuals to move freely between schooling and the labor market in response to new information about the relative desirability of schooling. The present study differs by restricting labor market entry to be irreversible and by assuming that individuals do not learn new information about the desirability of attending school unless they are actually enrolled. These restrictions create option value in the present specification that does not arise in these earlier studies. To date, there have been no attempts to quantify the aggregate importance to educational attainment or welfare of this option value or to examine how the effects of various policies depend on it. The present paper addresses these two gaps in the literature.

I model educational attainment as resulting from a series of sequential decisions made in the presence of uncertainty about the desirability of continuation. This model is analogous to Pindyck's (1993) model of irreversible investment with "technical" cost uncertainty, where the cost of completing a long-term project is revealed only as investment proceeds. Here enrollment reveals three pieces of information related to the three sources of uncertainty incorporated in the model. The first is uncertainty about collegiate aptitude, which influences the psychic costs (or benefits) from school attendance. Enrollment provides information in the form of course grades which are used to predict the future desirability of school.[5] The second source of uncertainty is short-term (nonpersistent) shocks to the relative costs (or benefits) of schooling. These shocks combine many factors - getting ill, having a parent lose a job, having a winning football team - that are not expected to persist over time.[6] The final source of uncertainty is about labor market opportunities associated with higher education. Individuals receive a new job offer after each year of college. On average, lifetime income increases with education but the specific realization is unknown ex-ante. Individuals learn of these opportunities only if they actually enroll. For tractability, I assume risk neutrality so that outcome variance has no direct influence on utility.

When combined with the opportunity to abandon investment after acquiring new information, this uncertainty creates an option value to college enrollment. In contrast to the standard view that uncertainty makes investment less desirable and reduces welfare, here uncertainty has the opposite effect.

[5] This specification of learning is similar to that included in Arcidiocono (2004).

[6] Such shocks are common features in the dynamic structural models of Keane and Wolpin (1997) and others.

Increased uncertainty about college completion increases welfare and makes college enrollment (and continuation) more attractive. The intuition for this result is that if dropout is allowed, then the expected utility gain from continuing in college is truncated at zero. Following enrollment, people can simply choose to drop out if continuation turns out to be undesirable. The expected value of this truncated unknown gain, which determines the magnitude of the option value, increases in its variance. A simple theoretical model is used to characterize the properties of the option value and to show how educational outcomes depend on its presence.

Option value is computed through counterfactual simulations of a structural dynamic model, which is estimated using postsecondary transcript data on a recent cohort of US youth, the National Educational Longitudinal Study. The model encompasses enrollment decisions and grade outcomes over four years as well as the decision to start at a two-year (community) or four-year college. I simulate educational outcomes and welfare using the dynamic model and compare this to the counterfactual scenario where individuals are forced to commit to an educational outcome before enrolling in college. The welfare difference between these two scenarios is the value of the option to dropout in response to the information received during college. As a basis of comparison, I also simulate outcomes and welfare in the first-best scenario, where individuals can maximize welfare ex-post.

The estimated model fits the actual data well, especially considering its relative simplicity. The fraction of students that attend four-year vs. two-year college, the fraction that drop-out after each year, and the relationship between course grades and dropout behavior are similar in the simulated and actual samples. The preferred specification, which matches the data most closely, allows for an arbitrary correlation between preference for school and academic aptitude using a mass point distribution with three points of support. Interestingly, the fitted model captures much of the actual relationship between background characteristics and educational outcomes, despite the restriction that background influences taste for school exclusively through expected collegiate performance. For instance, family income does not enter the estimation at all but large educational differences by family income are still predicted by the model due to the correlation between family income and factors that influence college performance. This suggests that much of the intergenerational transmission of education can be captured through parents' influence on predicted performance in school, rather than through differential access to credit or returns to school.

Estimates suggest that uncertainty about college completion is an important feature in postsecondary schooling decisions and outcomes; unanticipated taste shocks are half as large as the returns to the final year of college and dwarf direct tuition fees at public colleges. There is also evidence of learning about ability - over time people put increased weight on course grades in their continuation decisions. Together these sources of uncertainty create considerable option value to the initial enrollment decision. On average, students would be willing to pay \$15,000 (1992 dollars) to maintain this flexibility, with moderate-ability students (for whom educational outcomes are most uncertain) willing to pay even more (up to \$27,000). This represents 13% of the total value of the opportunity to enroll in college among all high school graduates and 34% for those closer to the enrollment margin. The ability to make decisions sequentially closes a quarter of the welfare gap between the first best outcome - where individuals can maximize welfare ex-post - and the static model where individuals must commit to educational outcomes ex-ante.

The dynamic model has option value for three reasons. First, the opportunity to drop out encourages more people to enroll, who may not want to if forced to commit ex-ante. Second, it permits those who would commit to graduate ex-ante to dropout if graduation is found to be undesirable. In aggregate, the former is greater than the latter. The ability to dropout increases enrollment by eight percentage points, decreases college completion by five percentage points, and increases average years of college attendance by three percent. Though this effect on average years of schooling is modest, significant welfare gains arise from better matching of individuals to their optimal level of schooling. New labor market opportunities are the third source of option value. Individuals with a low labor market draw after high school are able to obtain a new draw by enrolling. Schooling delays labor market entry until a favorable job offer is received. Approximately 60% of the aggregate option value comes from the information received in the first year of college.

In the final section of the paper, I use the estimated model to examine the effects of various policies aimed at increasing college enrollment and graduation, in light of the importance of option value. I find that enrollment and graduation decisions are relatively insensitive to large tuition subsidies. In contrast, the presence of community colleges have a large effect on the fraction of students who receive some postsecondary education, but the effect on college graduation is negligible. Low-income students are particularly influenced by the presence of community colleges.

Improvements to the K-12 system, which bring the least-prepared high school graduates up to the median high school GPA have large effects on both enrollment and completion. Enrollment and completion effects are particularly strong for low-income students, who are often the least prepared for college. Finally, providing a $10,000 bonus to low-income students who graduate from college would increase the graduation rate of this group slightly, but have little impact on enrollment.

Simulations also reveal that inclusion of option value is important to these predictions. Simulations using a static framework would overpredict the graduation consequences of across-the-board subsidies and improvements in academic preparedness, while underpredicting the effect of community colleges on enrollment. The consideration of option value is particularly important when evaluating policies that have different temporal characteristics, such as community colleges (which alter the tuition gradient) or across-the-board tuition reductions (which do not).

The rest of the paper proceeds as follows. The next section discusses the traditional static model of human capital investment and its short-comings, before developing a simple dynamic model which incorporates uncertainty. This section uses a simple model to demonstrate how uncertainty influences the value of college enrollment and continuation by creating option value. Section three introduces the full empirical model and discusses issues related to its estimation. The empirical model is a much-enhanced version of the conceptual one presented in section two. Estimation results are presented in Section four, which also includes a discussion of model fit. Section five uses the estimated model to calculate the option value created by the flexibility inherent in the US postsecondary educational system. Section six examines the effects of various policy interventions, to which uncertainty may be important. Section seven concludes by identifying directions for future work as well as other applications.

2 Models of Educational Investment

This section examines a simple two-period theoretical model of college enrollment and completion which incorporates completion uncertainty. The traditional model used in much economic research abstracts from uncertainty; educational investment is typically modeled as a static optimization problem. It is shown that high levels of dropout are inconsistent with the static model if degree wage effects are large, as the data suggest they are. This inconsistency motivates the development of the dynamic model.

In the dynamic model, it is assumed that the first half of college directly benefits some individuals but also provides information about the desirability of college completion, and thus has option value for everyone. Option value makes enrollment optimal for some people for whom the expected return to college is negative.

After giving a usable definition to option value, this section then explores some of its comparative statics and properties. Option value is always non-negative, increasing in the level of uncertainty, and greatest for those at the enrollment margin. Option value also considerably increases welfare. This theoretical discussion is merely used to motivate the empirical work, which estimates the structural parameters of a much-enhanced version of the model.

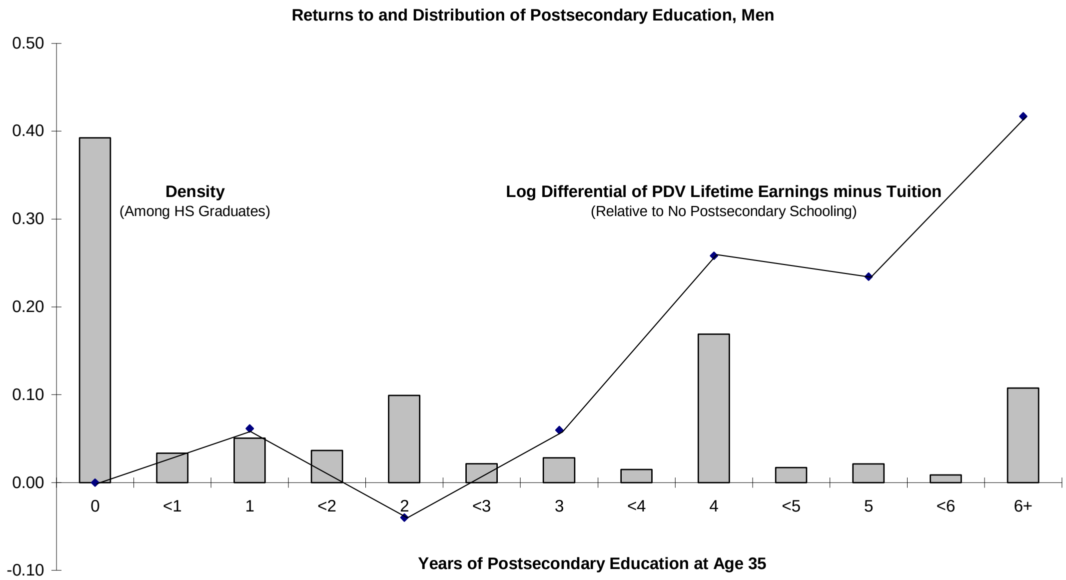
2.1 The Static Model and Its Limitations

The dynamic model developed in this paper is motivated by the observation that the traditional human capital model cannot explain the dropout behavior of many students. A simple version of the traditional model is developed in Card (1999). Individuals are assumed to maximize lifetime utility, which is a function of lifetime earnings and the cost of schooling, $U = \ln y(S) - c(S)$, where schooling cost is some increasing and convex function of years of schooling. If $y(S)$ and $c(S)$ are continuous, then the optimal schooling level satisfies the first order condition $\frac{dy_i(S_i^*)}{dS}\frac{1}{y_i(S_i^*)} = \frac{dc_i(S_i^*)}{dS}$. The benefit of an additional year of schooling (higher earnings) just offsets the additional costs (delayed earnings and psychic costs) at the optimum. Substantial non-linearities (for example, if $y(S)$ were a step function at $S = \overline{S}$) will cause the model to collapse into a discrete choice where individuals choose to enroll and graduate or choose not to enroll at all, with no individuals choosing schooling levels between zero and $\overline{S}$.

The experience of members of the National Longitudinal Survey of Youth 1979 (NLSY79) suggests that the returns to college are non-linear.[7] Figure 1 presents estimates of the earnings production function and educational distribution of male high school graduates. The solid lines plot the coefficients from a linear regression of log lifetime earnings (minus tuition) on a set of schooling level dummies and control variables.[8] From this data, it appears that the present discounted value of lifetime earnings minus tuition jumps discretely at four years of college, but is otherwise unrelated to schooling attainment.

[7]There is a substantial literature that documents the existance of nonlinearities in the returns to education. See Hungerford and Solon (1987), Jaeger and Page (1996), and Park (1999).

[8]The present discounted value of lifetime earnings are computed by summing discounted real annual income from age 18 to 62, assuming that real income is constant from 38 to 62. Discount rate is assumed to be 5%. By including several baseline characteristics in the regressions, these estimates only partially address the endogeneity and selection problems which complicate earnings comparisons by schooling level.



Notes: Density is from IPUMS-CPS years 1985 to 1990 restricted to 35 year old male high school graduates, weighted by person weights. Returns are estimated from NLSY. PDV of lifetime earnings are computed from age 18 to 62 assuming that real income is constant from age 38 to 62. Future values are discounted using a discount rate of 5%. Linear controls include dummies for ethnicity, four regions, urban, parents' education, high school GPA, AFQT percentile score, and the pairwise interactions between these last three variables. Average yearly tuition at public four-year colleges in region was subtracted for each year completed.

Figure 1

The traditional human capital model predicts that individuals should bunch at this discontinuity in the earnings production function and that very few people should fall in the intermediate ranges. Figure 1 also plots the distribution of postsecondary schooling attainment for men aged 35, who have presumably all completed their schooling. Consistent with the model, the most frequent schooling outcomes are zero (39% of the sample) and four years (17%) of college. Ten percent attend college for two years, which partially reflects Associate's degree attainment. Contrary to the theory, however, there are many people whose schooling level puts them on the flat part of the earnings production function. Fully 28% of high school graduates drop out before finishing their fourth year of college. From the perspective of traditional human capital theory where individuals optimally choose their schooling level to equate the known marginal costs and benefits of an additional year, these individuals present an unexplained puzzle.

This paper shows that dropout can be rationalized as one outcome in a dynamic model of schooling choice where the feasibility and desirability of degree completion is unknown ex-ante. As pointed out by Altonji (1993) and others, nonlinearities in the ex-post returns to schooling can create an option value to college enrollment if the difficulty of graduating is uncertain. Students with schooling outcomes on the flat part of the earnings curve may therefore be people for whom option value made enrollment worthwhile, even though the return was negative ex-post. The rest of this paper develops and examines a dynamic model that explicitly incorporates uncertainty and option value.

2.2 A Dynamic Model of College Enrollment and Completion

There are two time periods, which correspond to the first and second half of college.[9] Utility is in units of dollars and individuals are assumed to be risk-neutral. At period one, individuals decide whether or not enroll in college. Entering the labor market immediately provides zero utility while enrollment provides immediate utility $\varepsilon_{i,1}$, where $\varepsilon_{i,1}$ is the individual-specific return to the first half of college. At period two, those who have enrolled must decide whether or not to continue to graduation. Dropouts receive no further utility while completion provides additional utility of $\varepsilon_{i,1} + \varepsilon_{i,2}$. Individual-specific returns in the second half have a component that is known when the enrollment decision is being made ($\varepsilon_{i,1}$) and one that is only learned after enrollment ($\varepsilon_{i,2}$).[10] For expositional simplicity, I assume that $E[\varepsilon_{i,1}] = E[\varepsilon_{i,2}] = 0$ and $Cov[\varepsilon_{i,1}, \varepsilon_{i,2}] = 0$. Figure 2 illustrates the structure and payoffs of the model.

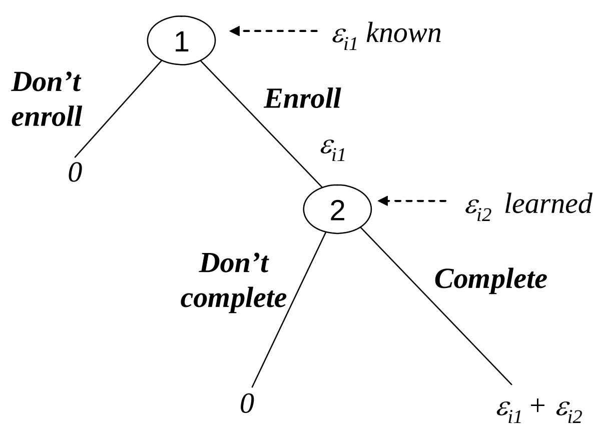


Figure 2: Structure of Choices and Payoffs - Theoretical Model

At each period, individuals make decisions to maximize the expected value of lifetime utility. The model is solved starting with the completion decision in period two. At period two, all parameters are known and the completion rule of individual i is thus:

$$Complete\ if : \varepsilon_{i,1} + \varepsilon_{i,2} > 0$$

[9] A similar set up was used by Manski (1989), Altonji(1993), Taber(2000), and Arcidiocono (2004).

[10] This formulation assumes that people do not learn new information about the desirability of schooling while in the labor market. A richer model that allows this form of learning would permit individuals to delay school entry.

At period one, people will enroll if the expected utility from doing so is greater than zero. Expectations are taken over the distribution of the unknown returns $\varepsilon_{i,2}$. The enrollment rule of individual i is thus:

$$Enroll\ if : \varepsilon_{i,1} + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}] > 0$$

The enrollment decision incorporates not only the immediate payoffs ($\varepsilon_{i,1}$) but also the expectation of future ones ($E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}]$).[11] This model embeds the traditional static human capital model with no uncertainty. If there is no uncertainty, then $\varepsilon_{i,2} = 0$ for all i and the decision rules become:

$$\begin{aligned} Complete\ if &: \varepsilon_{i,1} > 0 \\ Enroll\ if &: \varepsilon_{i,1} + 1(\varepsilon_{i,1} > 0) \cdot (\varepsilon_{i,1}) > 0 \end{aligned}$$

where $1(\cdot)$ equals one if the statement is true. In the specific case outlined above, the static model predicts no dropouts; anyone for whom enrollment is desirable will also want to complete college. More generally, the static model predicts that all dropouts will have positive ex-post returns to the first year ($\varepsilon_{i,1} > 0$). The model can be generalized such that some people will want to drop out ex-ante, but the general findings and intuition will still hold.

2.3 The Option Value of College Enrollment

A key feature of the dynamic model where dropout is endogenous is that the expected net utility gain from completing college is truncated at zero. If $\varepsilon_{i,2}$ is sufficiently adverse, then individuals will choose to dropout rather than assume this adverse shock. By providing information about the desirability of completion, enrollment thus has value beyond the utility provided in the first period directly. This section defines the option value created by uncertainty and discusses the implications of option value for educational outcomes and welfare.[12]

[11] Here I ignore any discounting that occurs between the first and second periods. Assuming a distribution for $\varepsilon_{i,2}$, one can derive an expression for $E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}]$. For instance, if $\varepsilon_{i,2}$ is drawn from a normal distribution with mean zero and variance σ^2, then $E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}] = \Pr(\varepsilon_{i,2} > -\varepsilon_{i,1}) \cdot (\varepsilon_{i,1} + E[\varepsilon_{i,2} | \varepsilon_{i,2} > -\varepsilon_{i,1}])$ where $\Pr(\varepsilon_{i,2} > -\varepsilon_{i,1}) = \Phi\left(\frac{\varepsilon_{i,1}}{\sigma}\right)$ and $E[\varepsilon_{i,2} | \varepsilon_{i,2} > -\varepsilon_{i,1}] = \sigma \frac{\phi\left(\frac{-\varepsilon_{i,1}}{\sigma}\right)}{1-\Phi\left(\frac{-\varepsilon_{i,1}}{\sigma}\right)}$

[12] The term "option value" is sometimes used synonomously with "continuation value," though this is not exactly correct. If completing the first year of college is required in order to enter the second year, then the first year can have continuation value even if all returns are known with certainty. People with negative first year returns will decide to enroll only if second year returns are sufficiently high. In contrast, option value arises only if second period returns are uncertain and

Enrollment is valuable because it leads to outcomes people may want to commit to ex-ante and because it provides information about the desirability of completion. The value of the opportunity to enroll can be decomposed into these two parts.

$$V_d(\varepsilon_{i,1}) = V_s(\varepsilon_{i,1}) + OptionValue(\varepsilon_{i,1})$$

$V_d(\varepsilon_{i,1})$ is the value of the opportunity to enroll for individual i, in the dynamic setting where individuals can drop out if continuation ends up being undesirable. $V_s(\varepsilon_{i,1})$ is the value of the enrollment opportunity in the static case - where individuals commit to an educational outcome ex-ante. Define $\overline{\varepsilon}_{d,1}$ as the critical value above which enrollment is optimal in the dynamic setting and $\overline{\varepsilon}_{s,1}$ analogously in the static setting.[13] From above we have $V_d(\varepsilon_{i,1}) = \max\left(0, \varepsilon_{i,1} + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}]\right)$ and $V_s(\varepsilon_{i,1}) = \max\left(0, \varepsilon_{i,1} + \max\{0, E\left[\varepsilon_{i,1} + \varepsilon_{i,2}\right]\}\right)$. Thus the value of the option to abandon schooling if completion ends up being undesirable is:

$$OptionValue(\varepsilon_{i,1}) = \max\left(0, \varepsilon_{i,1} + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}]\right) - \max\left(0, \varepsilon_{i,1} + \max\{0, E\left[\varepsilon_{i,1} + \varepsilon_{i,2}\right]\}\right)$$

Option value has several important features outlined below. I discuss the intuition for these features but leave formal proofs for future work.

1. *$OptionValue(\varepsilon_{i,1})$ is non-negative for all $\varepsilon_{i,1}$*

   To see this, consider three groups of individuals which together span the space of $\varepsilon_{i,1}$.

   Group A ($\varepsilon_{i,1} < \overline{\varepsilon}_{d,1}$) does not enroll under either the static or dynamic settings. Since they do not enroll, they get zero value from the option to drop out.

   Group B ($\overline{\varepsilon}_{d,1} < \varepsilon_{i,1} < \overline{\varepsilon}_{s,1}$) enrolls in the dynamic setting but would not if they were forced to commit to their educational decision ex-ante. For these individuals, the option value is pivotal to enrollment. This option value is equal to $\varepsilon_{i,1} + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2}\}]$. In this region, this expression is positive (by definition of $\overline{\varepsilon}_{d,1}$) and increasing in $\varepsilon_{i,1}$.

   Group C ($\overline{\varepsilon}_{s,1} < \varepsilon_{i,1}$) enrolls in both the dynamic and static settings. Their option value equals

future decisions can be conditioned on new information. With enough uncertainty, even individuals with negative first period returns and expected negative second period returns may find it optimal to enroll. Throughout I adopt this latter conceptualization of "option value."

[13] $\overline{\varepsilon}_{d,1}$ solves $V_d(\varepsilon_{i,1}) = 0$ and $\overline{\varepsilon}_{s,1} = 0$

$\varepsilon_{i,1}+E[\max\{0,\varepsilon_{i,1}+\varepsilon_{i,2}\}]-\varepsilon_{i,1}-\max\{0,E[\varepsilon_{i,1}+\varepsilon_{i,2}]\}$ which reduces to $E[\max\{-\varepsilon_{i,1},\varepsilon_{i,2}\}]$. This expression is decreasing in $\varepsilon_{i,1}$ and positive since $E[\varepsilon_{i,2}]=0$ and $E[\varepsilon_{i,2}|Z]>E[\varepsilon_{i,2}]$ for any value Z.

2. *$OptionValue(\varepsilon_{i,1})$ is greatest for individuals at the enrollment margin in the static model.*

   To see this, note that the option value of individuals in Group B is maximized at $\varepsilon_{i,1}=\overline{\varepsilon}_{s,1}=0$ where the option value equals $E[\max\{0,\varepsilon_{i,2}\}]$. This is greater than the option value of any individuals in the other two groups.

3. *$OptionValue(\varepsilon_{i,1})$ is increasing in the variance of $\varepsilon_{i,2}$.*

   Like a financial option, the value of the dropout option increases in the variance of the value of the underlying asset ($\varepsilon_{i,1}$). Since the truncation point is fixed, increased variance increases the truncated conditional expectation of $\varepsilon_{i,2}$. I show this through simulations in Figure 3.[14] Panel A plots the value of the enrollment opportunity for a range of values of $\varepsilon_{i,1}$ and for different levels of uncertainty about $\varepsilon_{i,2}$. The dotted line is the value of the enrollment opportunity in the static case, $V_s(\varepsilon_{i,1})$. This value is zero for those who choose not to enroll ($\varepsilon_{i,1}<0$) and then increases linearly with $\varepsilon_{i,1}$. The other lines plot the value of the enrollment opportunity in the dynamic situation where $\varepsilon_{i,2}$ is uncertain, $V_d(\varepsilon_{i,1})$. The vertical distance between these lines represents $OptionValue(\varepsilon_{i,1})$. The figure confirms that $OptionValue(\varepsilon_{i,1})$ is increasing in σ, as well as claims 1 and 2 above. In contrast to the standard view that uncertainty reduces welfare if agents are risk averse, here uncertainty combined with the ability to respond optimally actually increases welfare.

4. *The critical value $\overline{\varepsilon}_{d,1}$ is decreasing in the variance of $\varepsilon_{i,2}$.*

   This is a corollary of claim 3. As $OptionValue(\varepsilon_{i,1})$ increases due to increased uncertainty about $\varepsilon_{i,2}$, enrollment becomes desirable to more people. This can also be seen in Panel A of Figure 3. $\overline{\varepsilon}_{d,1}$ is where the dashed lines intersect the horizontal axis. Option value will make enrollment desirable even to people for whom the first half of college is unproductive ($\varepsilon_{i,1}<0$).

[14] In Figures 3 and 4 I take 10,000 random draws of $\varepsilon_{i,2}$ from a normal distribution with mean zero and variance σ^2 for each value of $\varepsilon_{i,1}$. The figures report the average welfare and schooling outcomes across these 10,000 draws.

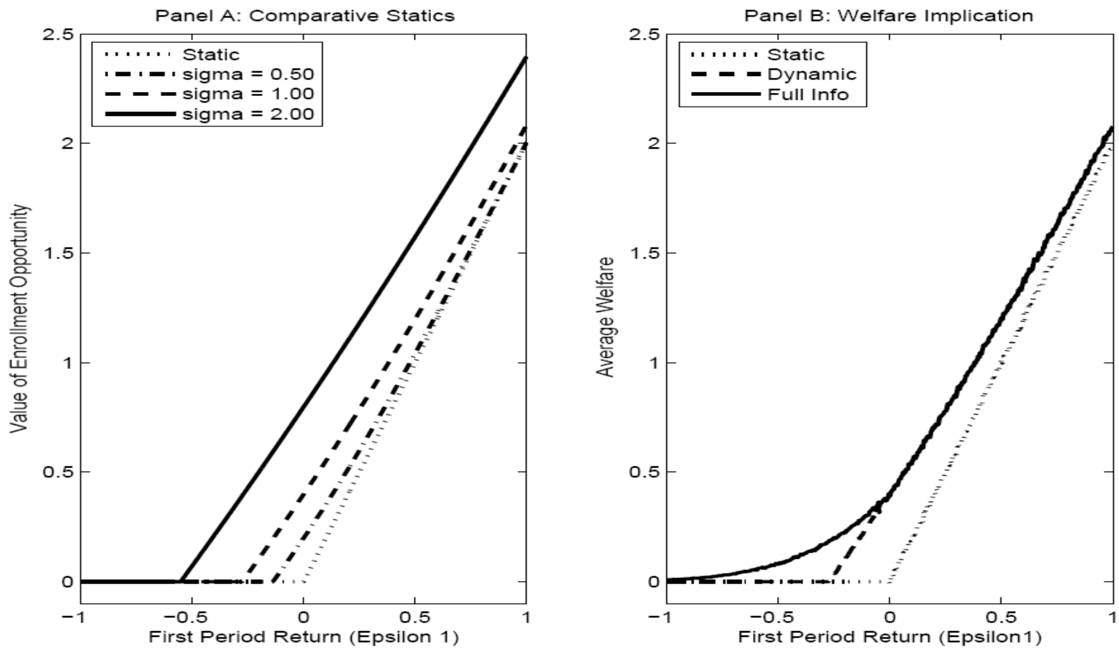


Figure 3

5. *The option to dropout can considerably improve welfare.*

Panel B of Figure 3 reports average ex-post welfare for a range of values of $\varepsilon_{i,1}$ assuming $\sigma = 1$. The dashed and dotted lines plot average ex-post welfare when dropout is and is not permitted. With enough simulations, average ex-post welfare will be identical to the expected value of the enrollment opportunity reported in Panel A. As a baseline, the solid line plots the average welfare in the counterfactual scenario where individuals can make education decisions to maximize welfare ex-post, after learning $\varepsilon_{i,2}$. The vertical distance between this line and the others represents the welfare loss resulting from incomplete information about $\varepsilon_{i,2}$. The ability to dropout after learning $\varepsilon_{i,2}$ (the dashed line) closes much of this welfare gap.

The sources of the welfare gains coming from the ability to dropout can be seen more clearly by looking at educational outcomes under the three scenarios. Figure 4 plots the average years of schooling (Panel A) and fraction enrolling (Panel B) and completing college (Panel C) under the static, dynamic, and full-information scenarios described above.[15] Individuals in Group A receive no schooling in either the static or dynamic settings, though some (with high $\varepsilon_{i,2}$) would enroll and graduate if they knew $\varepsilon_{i,2}$ with certainty. Individuals in Group B are compelled to enroll

[15] Since this is only a two-period model, the maximum years of schooling is two.

despite their negative first period returns because of the informational value. Though many will eventually drop out, others will graduate and the costs of experimenting are not too high. This group receives considerably more education in the dynamic setting. Interestingly, a small subset of these individuals actually continue to graduation due to the sunk-cost nature of their period 1 investment, despite this being suboptimal ex-post. Group C benefits from the dynamic setting because they have the option to dropout if continuation is undesirable. In the static model, all commit to graduating, even if it is undesirable ex-post. Option value increases welfare by reducing educational attainment of this group.

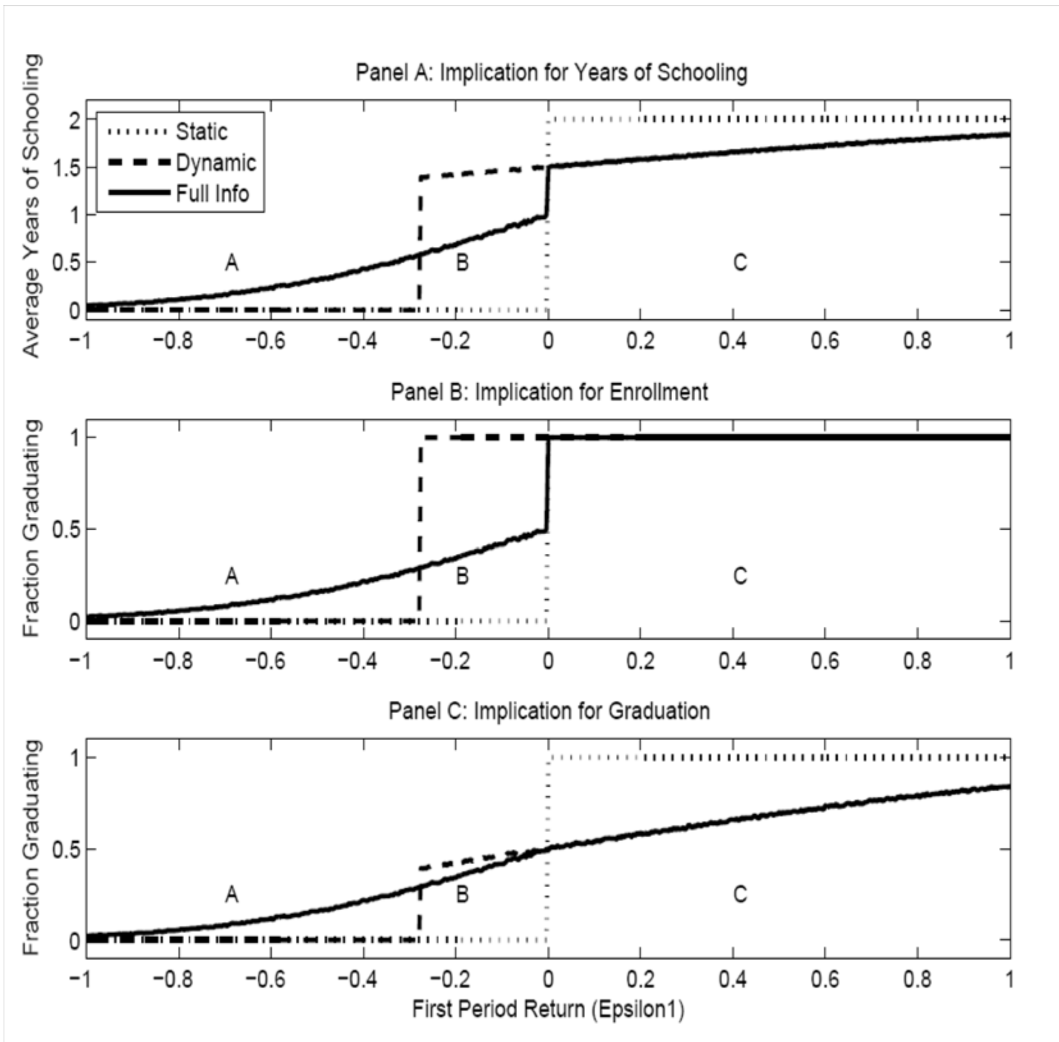


Figure 4

2.4 Summary and Implications for Empirical Work

This section has introduced a simple dynamic theoretical model of college enrollment and completion. The dynamic model was motivated by the failure of the static model to explain high rates of college dropout. In a dynamic setting, dropout occurs when new information reveals that continuation is not desirable. The opportunity to dropout in response to this information creates option value, which was shown to have important consequences for educational outcomes and welfare. Specifically, option value increases the incentive to enroll, particularly for those at the enrollment margin in the static model. Any model that ignores this value will necessarily understate the incentive to enroll and mischaracterize the social desirability of college dropout.

The model above is useful for presenting ideas and intuition, but is too simple to provide useful guidance about specific policies. The remainder of the paper turns to a much more elaborate version of the dynamic model, which I then estimate in order to quantify the importance of option value and uncertainty. The estimated model is also used to examine the effects of several policy interventions to which the magnitude of uncertainty and option value may be important.

3 Empirical Implementation

To characterize schooling uncertainty quantitatively, I estimate an empirical model that is a much richer version of the basic model presented above. The empirical model covers the enrollment decisions and grade outcomes at four time periods and allows individuals to start at either a two-year and four-year college. The model includes several sources of uncertainty. Like many dynamic models, I include unanticipated shocks to the relative desirability of school and labor market entry at each point in time. For example, receiving an unusually favorable outside job offer or getting ill influences the relative desirability of schooling and work at a single period. These shocks are assumed to be serially uncorrelated. The second source of uncertainty is about academic aptitude, which influences taste for schooling throughout college. Students do not know for certain whether they are a "B" or "C" college student until they enroll. Grades following enrollment provide a signal of this unobserved ability - students learn about their aptitude through their grades. This section presents the full empirical model and discusses issues related to its estimation.

3.1 Model and Solution

3.1.1 Structure of Choices and Preferences

I model the college enrollment and continuation decisions at four periods in time, corresponding to the four academic years after high school graduation. During the first period individuals decide whether to start at a four-year or two-year college, which I refer to pathway choice, or to not enroll in college. The pathway chosen affects the level and timing of direct schooling costs (which may differ across individuals) and unmodeled college amenities. At each time period t an individual chooses whether to enter the labor market (receiving payoff $u^{w}_{i,t}$) or continue in school for another year, receiving payoff $u^{sj}_{i,t}$ in period t and the option to make an analogous work-school decision in period $t+1$, where $j = 2, 4$ denotes pathway choice. After period four, there are no more decisions to make and all individuals enter the labor market.[16]

Utility is in dollars. The utility from discontinuing school and entering the labor market at time t equals the expected present discounted value of lifetime income from year t to age 62 ($Income_{i,t}$), conditional on school enrollment up to year t, plus a random component $\varepsilon^{w}_{i,t}$.

$$u^{w}_{i,t} = Income_{i,t} + \varepsilon^{w}_{i,t}$$

Utility derived from attending school during year t depends linearly on a heterogeneous intercept (α^{sj}_{i}, to be specified later), the difficulty of completing year t ($A_{i,t}$), direct tuition and commuting costs, and a random component $\varepsilon^{sj}_{i,t}$. $Distance_{i,t}$ and $Tuition_{i,t}$ vary by the type of school currently attending (2-year or 4-year), so individuals that start at a two-year school will pay community college tuition for the first two years then four-year college tuition for their third and fourth years.

$$u^{sj}_{i,t} = \alpha^{sj}_{i} + \alpha_A A_{i,t} - (\alpha_d Distance_{i,t} + Tuition_{i,t}) + \varepsilon^{sj}_{i,t}$$

The random shocks ($\varepsilon^{sj}_{i,t}$,$\varepsilon^{w}_{i,t}$) are learned by the individual prior to making the period t decision. The term $\alpha_A A_{i,t}$ captures the preference for school (in dollar terms) that covaries with its difficulty. I

[16] The model does not currently permit two-year and four-year colleges to affect earnings differently. Kane and Rouse (1995) find that the return to education received at two- and four-year institutions is comparable. They estimate that the average college student earned about 5% more than similar high school graduates for every year of credits completed, regardless of where those credits were earned.

assume that the difficulty of year t is distributed around an individual-specific mean, so $A_{i,t} = A_i + \varepsilon_{i,t}^A$. A_i is fixed and unobserved both to the individual and econometrician throughout. If $\varepsilon_{i,t}^A$ is mean zero and serially uncorrelated, then $E_t[A_{i,t}] = E_t[A_i]$. Individuals consider the expected utility of schooling given information available at period t when making their period t decisions:

$$E_t[u_{i,t}^{sj}] = \alpha_i^{sj} + \alpha_A E_t[A_i] - (\alpha_d Distance_{i,t} + Tuition_{i,t}) + \varepsilon_{i,t}^{sj}$$

Figure 5 depicts the structure of choices and payoffs in the full empirical model.

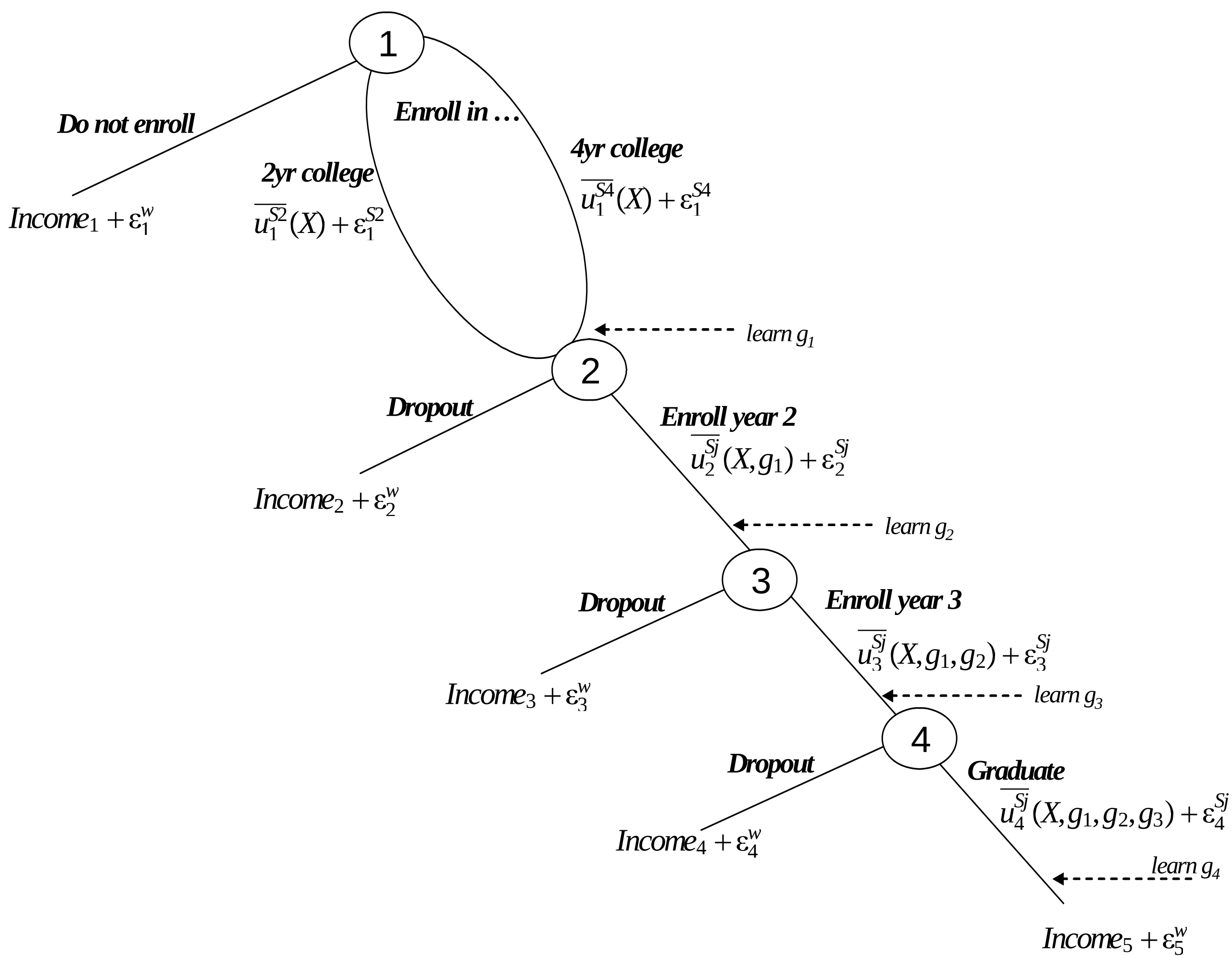


Figure 5: Structure of Choices and Payoffs - Empirical Model

3.1.2 Academic Performance

At the end of each year, students enrolled in college learn their performance during that year. Academic performance is measured by the college grade point average (on a four-point scale) during period t. I assume that grades provide a noisy signal of A_i:

$$g_{i,t} = A_i + \varepsilon_{i,t}^g$$

The $\varepsilon_{i,t}^g$ is the component of grade outcomes that is not serially correlated. This represents idiosyncratic determinants of academic performance that do not persist across time. The conditional expectation of A_i (on X_i) is given by the heterogeneous term γ_i, which is specified in the next subsection.

$$E[A_i|X_i] = \gamma_i$$

3.1.3 Heterogeneity

The variables α_i^{sj} and γ_i represent persistent preferences for school and persistent determinants of academic aptitude, respectively, which may be correlated in the population. α_i^{sj} varies with school type (j) so that individuals may have different tastes for attending a two- or four-year school. To permit a general structure of correlation between unobservable preferences and ability, I specify that α_i^{sj} and γ_i come from a mass point distribution which describe the ability and schooling preference of m different types of individuals. Type is known to the individual throughout, but is unknown to the econometrician. I also make the parametric assumption that the conditional expectation of A_i (on X_i) is linear in high school grade point average ($HSgpa_i$), percentile score on the AFQT, and whether a parent has a college degree ($ParBA_i$).

$$\begin{aligned} \alpha_i^{sj} &= \alpha_0^{sj} + \alpha_m^{sj} \\ \gamma_i &= \gamma_0 + \gamma_m + \gamma_g HSgpa_i + \gamma_t AFQT_i + \gamma_p ParBA_i \end{aligned}$$

where γ_m measures the unobserved academic aptitude of people of "type" m and α_m^{sj} is their preference for school of type j. I estimate models permitting up to three types. For Type I individuals, γ_m and α_m^{sj} are normalized to zero. As a special case, I will also estimate models with no unobserved heterogeneity, which assumes that all correlation between preference for school and academic aptitude are captured linearly through $\alpha_A E_t[A_i]$.

3.1.4 Solution

At each time t , the individual maximizes the expected discounted value of lifetime utility by choosing whether to discontinue schooling and receive $u^w_{i,t}$ or continue school for at least one more year. The decision problem can be solved for each individual by backwards recursion and by assuming a distribution for the preference and grade shocks ($\varepsilon^{sj}_{i,t}$, $\varepsilon^{w}_{i,t}$, $\varepsilon^{g}_{i,t}$). Throughout I assume that $\varepsilon^{sj}_{i,t}$ and $\varepsilon^{w}_{i,t}$ are drawn from an Extreme Value Type I distribution with location and scale parameters zero and τ, respectively.[17] Grade shocks are assumed to be normally distributed with $\varepsilon^{g}_{i,t} \sim N(0, \sigma_g)$.

With learning, individuals update their belief about A_i in response to new information received through grades. I make the parametric assumption that the conditional expectation of A_i is a weighted average of the unconditional expectation and students' cumulative grade point average.

$$\begin{aligned} E_t[A_i] &= E[A_i|X_i] \text{ if } t = 1 \\ &= \gamma_{xt} E[A_i|X_i] + (1-\gamma_{xt}) \sum_{q=1}^{q=t-1} \frac{g_{i,q}}{t-1} \text{ if } t > 1 \end{aligned}$$

This specification is a generalized case of the normal learning model. The normal learning model imposes that $\gamma_{xt} = \left(\frac{1/\sigma_a^2}{1/\sigma_a^2 + (t-1)/\sigma_g^2}\right)$, where σ_a^2 is the variance of A_i and σ_g^2 is the variance of $(g_{i,t} - A_i)$. I have not imposed that the timing of learning follow the behavior implied by the normal learning model.

At period 4 the final enrollment decision is made by comparing the lifetime utility of entering the labor market without graduating to that of continuing for one more year. In periods 2 through 4, I omit the j subscripts.

$$\begin{aligned} V^w_{i,4} &= Income_{i,4} + \varepsilon^w_{i,4} \\ V^s_{i,4} &= \alpha_i + \alpha_A E_4[A_i] - Cost_{i,4} + \beta E_4[V_{i,5}] + \varepsilon^s_{i,4} \end{aligned}$$

where $Cost_{i,4} = \alpha_d Distance_{i,4} + Tuition_{i,4}$. At period 4, expectations are taken over the distribution

[17] The assumption that labor market and schooling shocks have the same time-invariant variance will be relaxed in future work.

of labor market shocks in period 5 ($\varepsilon^w_{i,5}$) and grade shocks in period 4 ($g_{i,4}$). Since all individuals enter the workforce upon reaching period 5, $V_{i,5} = V^w_{i,5} = Income_{i,5} + \varepsilon^w_{i,5}$ and $E_4[V_{i,5}] = Income_{i,5} + \tau\gamma$ from the extreme value assumption [$\gamma = 0.577$ is Euler's constant]. Future utility is discounted at the rate β. If individuals learn about unobserved ability through grades, then $E_4[A_i]$ is a weighted average of the unconditional expectation and previous grade realizations:

$$V^s_{i,4} = \alpha_i + \alpha_A \left[\gamma_{x4} E[A_i|X_i] + (1 - \gamma_{x4}) \sum_{q=1}^{q=3} \frac{g_{i,q}}{3} \right] - Cost_{i,4} + \beta[Income_{i,5} + \tau\gamma] + \varepsilon^s_{i,4}$$

Individuals will continue to graduation if $V^s_{i,4} > V^w_{i,4}$.

At periods 2 and 3, the enrollment and continuation decisions are made by comparing the lifetime utility of entering the labor market immediately to that of continuing school for one more year.

$$\begin{aligned} V^w_{i,t} &= Income_{i,t} + \varepsilon^w_{i,t} \\ V^s_{i,t} &= \alpha_i + \alpha_A E_t[A_i] - Cost_{i,t} + \beta E_t[V_{i,t+1}] + \varepsilon^s_{i,t} \end{aligned}$$

where $V_{i,t+1} = \max(V^w_{i,t+1}, V^s_{i,t+1})$. Expectations are again taken over the distribution of all future preference shocks ($\varepsilon^w_{i,q}, \varepsilon^s_{i,q}$ for $q > t$) and grade shocks ($g_{i,q}$ for $q \geq t$), but now both of these influence future educational decisions. Integrating out the grade shocks, the $E\max$ term can be written as:

$$E_t \left[\max(V^w_{i,t+1}, V^s_{i,t+1}) \right] = \int E_t \left[\max(V^w_{i,t+1}, V^s_{i,t+1}) | g_{i,t} \right] \cdot \Pi(dg_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})$$

where $\Pi(dg_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})$ is the pdf of the t-period grade outcome conditional on information available at time t. The conditional expectation is taken only over the future preference shocks ($\varepsilon^w_{i,q}, \varepsilon^s_{i,q}$ for $q > t$). Again with the assumption that the preference shocks are not serially correlated and are drawn from an extreme value distribution, this expectation has a closed-form representation:

$$\begin{aligned} & E_t \left[\max(V^w_{i,t+1}, V^s_{i,t+1}) \right] \\ = & \int \left[\tau\gamma + \tau \log \left\{ \exp\left(\frac{1}{\tau} \overline{V}^s_{i,t+1}(g_{i,t}) \right) + \exp\left(\frac{1}{\tau} \overline{V}^w_{i,t+1} \right) \right\} \right] \cdot \Pi(dg_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\}) \end{aligned}$$

In order to actually solve and estimate the model, I discretize $g_{i,t}$ into K values and approximate

$\Pi(dg_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})$ with a discretized version $p(g^k_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})$[18]. The Emax term can then be written as

$$\begin{aligned}
& E_t\left[\max(V^w_{i,t+1}, V^s_{i,t+1})\right] \\
= & \sum_{k=1}^{K}\left[\tau\gamma + \tau\log\left\{\exp\left(\frac{1}{\tau}\overline{V}^s_{i,t+1}(g^k_{i,t})\right) + \exp\left(\frac{1}{\tau}\overline{V}^w_{i,t+1}\right)\right\}\right]\cdot p(g^k_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})
\end{aligned}$$

And the indirect utility function becomes:

$$\begin{aligned}
V^s_{i,t} &= \alpha_i + \alpha_A\left[\gamma_{xt}E[A_i|X_i] + \gamma_{gt}\sum_{q=1}^{q=t-1}\frac{g_{i,q}}{t-1}\right] - Cost_{i,t} \\
&+\beta\left[\sum_{k=1}^{K}\left[\tau\gamma + \tau\log\left\{\exp\left(\frac{1}{\tau}\overline{V}^s_{i,t+1}(g^k_{i,t})\right) + \exp\left(\frac{1}{\tau}\overline{V}^w_{i,t+1}\right)\right\}\right]\cdot p(g^k_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\})\right] + \varepsilon^s_{i,t}
\end{aligned}$$

Individuals will continue their education if $V^s_{i,t} > V^w_{i,t}$.

At period 1, the value of the two enrollment options takes a similar form:

$$\begin{aligned}
V^{sj}_{i,1} &= \alpha^{sj}_i + \alpha_A E[A_i|X_i] - Cost^j_{i,t} \\
&+\beta\left[\sum_{k=1}^{K}\left[\tau\gamma + \tau\log\left\{\exp\left(\frac{1}{\tau}\overline{V}^{sj}_{i,2}(g^k_{i,1})\right) + \exp\left(\frac{1}{\tau}\overline{V}^w_{i,2}\right)\right\}\right]\cdot p(g^k_{i,1}|X_i,)\right] + \varepsilon^{sj}_{i,1}
\end{aligned}$$

At period 1, individuals maximize expected lifetime utility by choosing between $V^{S2}_{i,1}, V^{S4}_{i,1}$, and $V^w_{i,1}$.

3.2 Implications of the model

The indirect utility functions $\left\{V^s_{i,t}, V^w_{i,t}\right\}_{t=1}^{t=4}$ provide expressions for the relative desirability of entering the labor market or continuing in school at time t. This relative value depends on a number of primitive parameters. The direct and opportunity costs as well as financial returns are captured in the terms $Cost_{i,t}$ and $Income_{i,t}$. Their importance to educational decisions have been the topic of much examination. Less frequently studied is the contribution of academic ability to continuation decisions. This is captured by α_A and the parameters of the grade function. I have modeled family background

[18] Since grades are distributed normally, the transition probabilities can be computed directly using the standard normal cumulative distribution function. $p(g^k_{i,t}|X_i, \{g_{i,1}...g_{i,t-1}\}) = \Phi\left(\frac{g^k_{i,t}+(0.5)*kstep-E_t[g_{i,t}]}{\sigma_{t,g}}\right) - \Phi\left(\frac{g^k_{i,t}-(0.5)*kstep-E_t[g_{i,t}]}{\sigma_{t,g}}\right)$ where kstep is the distance between the points in the discretized grade space.

and innate ability as influencing educational decisions primarily through expected scholastic aptitude (grades). This model can be used to quantify the contribution of family background to educational outcomes that operates through college academic performance. Family background influences academic performance, which in turn influences educational decisions.

The value of enrollment is also influenced by the amount of uncertainty and the speed at which it is revealed, as parameterized by τ and $\{\gamma_{xt}, \sigma_{gt}\}_{t=1}^{t=4}$. If τ is high, then preference shocks have a high variance, which increases the value of college enrollment and continuation. Future decisions take these preference shocks into account, so a greater variance increases the likelihood that either the schooling or work shock will be high, thus increasing the option value.

Option value decreases with the variance of grade shocks (σ_{gt}). Since grades provide a noisy signal of unobserved ability (which influences utility through academic performance), greater variance decreases the signal value of grade realizations and thus the option value created by the ability to learn about aptitude through grades. If grades provided no signal value (either because they were completely random or because there is no uncertainty about ability), the value of enrollment would be diminished.

The temporal nature of learning about ability is parameterized by $\{\gamma_{xt}\}_{t=1}^{t=4}$. If academic ability is learned quickly, then γ_{xt} should decline rapidly at first then level off. If subsequent grade shocks continue to provide new information about ability, γ_{xt} should continue to decline throughout college. The normal learning model imposes that γ_{xt} follow a specific decreasing pattern over time.

3.3 Estimation

The parameters of the model are estimated with maximum likelihood using data on the enrollment decisions, academic performance, and baseline characteristics of a panel of individuals. With no unobserved heterogeneity, individual $i's$ contribution to the likelihood function is

$$L_i = \Pr(S_{i,1}^2 = 1)^{S_{i,t}^2} \Pr(S_{i,1}^4 = 1)^{S_{i,t}^4} \Pr(S_{i,1} = 0)^{1-S_{i,1}} \prod_{t=2}^{4} \Pr(S_{i,t} = 1)^{S_{i,t}} \Pr(S_{i,t} = 0)^{1-S_{i,t}} \prod_{t=1}^{4} \Pr(g_{i,t})$$

where $S_{i,1}^2$ and $S_{i,1}^4$ indicate pathway choice in period 1 and $S_{i,t}$ is an indicator for enrollment in either type of school during period t. With the extreme value assumption on the preference shocks (which are unobserved to the econometrician), choice probabilities take the familiar logit form:

$$\begin{aligned}
\Pr(S_{i,1}^{j} &= 1) = \frac{\exp\left(\frac{1}{\tau}\overline{V}_{i,t}^{sj}\right)}{\exp\left(\frac{1}{\tau}\overline{V}_{i,1}^{S4}\right) + \exp\left(\frac{1}{\tau}\overline{V}_{i,1}^{S2}\right) + \exp\left(\frac{1}{\tau}\overline{V}_{i,1}^{w}\right)} \\
\Pr(S_{i,t} &= 1) = \frac{\exp\left(\frac{1}{\tau}\overline{V}_{i,t}^{s}\right)}{\exp\left(\frac{1}{\tau}\overline{V}_{i,t}^{s}\right) + \exp\left(\frac{1}{\tau}\overline{V}_{i,t}^{w}\right)} \quad (t = 2, 3, 4)
\end{aligned}$$

The likelihood of grade outcomes is given by the normal probability density function:

$$\Pr(g_{i,t}) = \phi\left(\frac{g_{i,t} - E_t[g_{i,t}]}{\sigma_{t,g}}\right)$$

When unobserved (to the econometrician) heterogeneity is included, the likelihood contribution of individual i must be integrated over the joint distribution of γ_m and α_m^{sj}. Since this distribution is assumed to have M mass points, the type-specific likelihood contribution must be summed over the M possible types, weighted by the probability of being each type.

$$L_i = \sum_{m=1}^{M} p_m L_{im}$$

where p_m is the probability of being "type" m, which is a parameter to be estimated.

With no heterogeneity, there are 16 parameters to estimate: five in the utility function (α_0^{S2}, α_0^{S4}, α_A, α_d, τ) and eleven in the grade equations (γ_0, γ_g, γ_t, γ_p, σ_{g1}, γ_{x2}, σ_{g2}, γ_{x3}, σ_{g3}, γ_{x4}, σ_{g4}). Unobserved heterogeneity adds four parameters ($\alpha_m^{S2}, \alpha_m^{S4}, \gamma_m, p_m$) for each additional type.

3.4 Data

The model is estimated on a panel of 1773 men participating in the National Educational Longitudinal Study (NELS).[19] NELS participants were first interviewed in 1988, while in 8th grade, then again in 1990, 1992, 1994, and 2000. Complete college transcripts were obtained in 2000 for most participants. The NELS transcript and survey data are used to construct the main variables used in the analysis - the college enrollment indicators, grade outcomes, and baseline characteristics. I supplemented the NELS dataset with institutional characteristics obtained from The College Board's 1992 Annual Survey of Colleges and the Integrated Postsecondary Education Data System 1992 Institutional Characteristics

[19] In future work I will also estimate the model on women.

survey. For each NELS individual, I merged on tuition and institutional performance (e.g. graduation rates and transfer rates) of the nearest two-year college, distance to the nearest two-year and four-year college (in miles), and average tuition levels at public two-year and four-year colleges in each state.

I define a time period as one academic year and classify individuals by years of continuous college enrollment following high school graduation. Students are considered enrolled during year t if they attempted at least six course units (approximately part-time status) in both Fall and Spring of the academic year. Since conditional expectations of lifetime income do not appear in the NELS dataset, I estimate them using data from an earlier cohort, the National Longitudinal Survey of Youth 1979 (NLSY79). Using variables that are common in both the NLSY79 and the NELS (such as high school GPA, parental education, AFQT, ethnicity, urban and region), I estimate the parameters of a lifetime income equation separately by sex using OLS and predict counterfactual lifetime income for individuals in the NELS sample. Essentially, I assume that individuals in my sample look at the experience of "similar" individuals twelve years older to form their income expectations. This approach is similar to the "reference group expectations" referred to by Manski (1991).

I restrict the dataset to on-time high school graduates with complete information on key baseline variables (sex, high school GPA, parents' education, AFQT, distance to nearest colleges) and complete college transcripts (unless no claim of college attendance). I also exclude residents of Alaska, Hawaii, and the District of Columbia. After these restrictions the final dataset contains 1773 men.

The data appendix contains more details on how the dataset was constructed.

3.5 Identification

The parameters in the utility function ($\alpha_0^{S2}, \alpha_0^{S4}, \alpha_A, \alpha_d$) are identified from the educational choices up to the scale parameter τ. For example, the difference in enrollment rates between individuals with high expected grades and low expected grades but all else equal identifies the ratio α_A/τ. Since utility is in dollar units, τ is identified from variation in $Tuition_{i,t}$ and $Income_{i,t}$ across individuals and over time. The estimate of τ is the magnitude of preference shocks that is needed to rationalize the proportions of people dropping out in each year, given the financial costs and benefits from doing so (holding all other variables constant). For instance, if the financial return to completing a fourth year of college is much higher than completing the third year, then more people should drop out before the third year than the fourth. The magnitude of this enrollment difference identifies τ - if the dropout rates are similar then the variance of preference shocks must be high (τ must be large) to rationalize the data. Cross-state tuition differences contribute to the identification of τ in the same way.

The parameters of the grade function are identified primarily from the grade outcomes in the typical manner, though the educational choices also help identify these parameters.

Parameters associated with unobserved heterogeneity are identified by common behavior which is contrary to the model. For instance, there may be individuals with poor academic performance but who still persist to graduation due to unmodeled parental pressure. If there are a sufficient number of similar individuals, then a model that permits for this type of behavior will fit the data better (i.e., have a higher likelihood). In practice, it is difficult to identify the discount rate β separately from τ. In the current specification, I fix β at 0.95. Future work will permit heterogeneity in β.

4 Estimation Results

4.1 Parameter Estimates

Table 1 provides estimates of the structural parameters for the male subsample. Columns (1) to (3) provide estimates from a base model with no learning about academic aptitude while columns (4) to (6) provide estimates from the full model described above. Both of these models are estimated with and without allowing for up to three points of unobserved heterogeneity.

In the model without learning, expectations about grade realizations are based exclusively on baseline characteristics and type, so $E_t[A_i] = E[A_i|X_i, Type]$ for all t. The parameter estimates all have the expected signs and are statistically significant. Since utility is in units of dollars, these estimates are immediately interpretable as the dollar value (in \$100,000) associated with a one-unit change in the independent variable. With no unobserved heterogeneity or learning (column (1)), the estimates imply that four-year colleges have amenities valued at \$32,000 over two-year colleges. Expecting to do well in school is also valuable. Each additional grade point (e.g. going from a C-student to a B-student) is equivalent to \$70,000. Living 100 miles from a college is equivalent to \$11,000 in tuition. A key parameter is τ, which parameterizes the variance of the preference shocks. At the estimated parameters, the preference shocks have a standard deviation of \$64,000 $\left(= \tau \frac{\Pi}{\sqrt{6}}\right)$. As expected, the grade parameter estimates show a strong positive correlation between academic performance and baseline characteristics such as academic performance in high school, AFQT test scores, and parent's education.

The estimate of α_A in column (1) could be biased if people with high academic ability also have stronger preference for attending school, independent of the causal effect of aptitude on schooling ease. Columns (2) and (3) address this concern by allowing for several different "unobserved types," each with an arbitrary correlation between schooling preference and academic aptitude. Permitting unobserved heterogeneity improves model fit considerably. Relative to type 1 individuals, type 2 individuals (20% of sample) are higher ability ($\gamma_{type2} > 0$), but have a stronger dislike of 4-year colleges ($\alpha^{S4}_{type2} < 0$) and are neutral to two-year colleges. These individuals can be thought of as good students from disadvantaged families. By contrast, type 3 individuals (34%) are lower ability ($\gamma_{type3} < 0$), have a stronger preference for 4-year colleges ($\alpha^{S4}_{type3} > 0$) and dislike two-year colleges ($\alpha^{S2}_{type2} < 0$), though this latter effect is not statistically significant. Incorporating unobserved heterogeneity does not qualitatively change the other parameter estimates. However, the estimated deviation of the preference shocks increases to $95,000. Consequently, the magnitude of the other parameter estimates also increases. Interestingly, the relationship between expected academic ability and enrollment probabilities (α_A/τ) changed little, increasing from 1.4 to 1.6 when unobserved heterogeneity is permitted. The estimated variance of the grade shocks decreases because a greater share of the performance variance is captured by baseline characteristics (including type).

Columns (4) to (6) present estimates from the full learning model presented in Section 3. The parameter estimates are very similar to estimates from the no-learning model, both qualitatively and quantitatively. With learning, individuals estimate future academic performance by calculating a weighted average of performance predicted with baseline characteristics (including type) and cumulative grade point average, where the weights (γ_{x2}, γ_{x3}, and γ_{x4}) are parameters to be estimated. The normal learning model predicts that the weight placed on baseline characteristics should decrease with t (γ_{x1} is normalized to one), as should the residual grade variance (σ_{gt}). The estimates in column (4), which do not control for unobserved heterogeneity, support this implication of the normal learning model. The best predictor of year-two grades weighs baseline characteristics and first-year grades approximately equally (45% vs. 55%). Fourth-year grades, however, are best predicted by placing only 19% of the weight on baseline characteristics and 81% on three-year cumulative grade point average.

	No Learning			Learning		
	One type	Two types	Three types	One type	Two types	Three types
	(1)	(2)	(3)	(4)	(5)	(6)
Utility parameters						
alpha_20	-2.88	-4.39	-4.21	-2.53	-6.28	-3.25
	(0.10)	(0.18)	(0.17)	(0.07)	n.a.	(0.21)
alpha_40	-2.56	-3.66	-3.66	-2.18	-3.05	-2.91
	(0.10)	(0.17)	(0.19)	(0.07)	(0.12)	(0.19)
alpha_a	0.70	1.22	1.21	0.59	0.97	1.04
	(0.04)	(0.08)	(0.08)	(0.03)	(0.06)	n.a.
alpha_d	0.11	0.21	0.25	0.12	0.20	0.22
	(0.04)	(0.06)	(0.07)	(0.04)	(0.06)	n.a.
tau	0.50	0.74	0.75	0.51	0.70	0.65
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	n.a.
Grade parameters						
gamma_0	1.21	0.35	0.88	0.80	-0.03	0.69
	(0.06)	(0.08)	(0.10)	(0.09)	(0.12)	n.a.
gamma_g	0.38	0.42	0.39	0.43	0.46	0.51
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	n.a.
gamma_t	0.43	0.61	0.67	0.61	0.69	0.72
	(0.04)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)
gamma_p	0.19	0.23	0.26	0.27	0.26	0.32
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)
gamma_x2				0.45	0.54	0.49
				(0.03)	(0.03)	n.a.
gamma_x3				0.31	0.46	0.33
				(0.04)	(0.05)	(0.03)
gamma_x4				0.19	0.36	0.22
				(0.05)	(0.06)	(0.05)
sd_gpa	0.64	0.53	0.47			
	n.a.	n.a.	n.a.			
sd_gpa1				0.65	0.59	0.62
				n.a.	n.a.	n.a.
sd_gpa2				0.52	0.50	0.51
				n.a.	n.a.	n.a.
sd_gpa3				0.52	0.50	0.52
				n.a.	n.a.	n.a.
sd_gpa4				0.55	0.53	0.54
				n.a.	n.a.	n.a.
Type-specific parameters						
gamma_Mtype2		0.84	0.62		0.89	0.22
		(0.03)	(0.02)		(0.09)	(0.09)
alpha_2Mtype2		0.25	0.05		2.82	0.55
		(0.11)	(0.14)		n.a.	n.a.
alpha_4Mtype2		-0.12	-0.25		-0.07	-2.61
		(0.08)	(0.11)		(0.08)	n.a.
probMtype2		0.44	0.20		0.61	0.07
		n.a.	n.a.		n.a.	n.a.
gamma_Mtype3			-0.92			-0.55
			(0.04)			(0.07)
alpha_2Mtype3			-0.09			-1.87
			(0.24)			(0.54)
alpha_4Mtype3			0.32			-0.44
			(0.14)			n.a.
probMtype3			0.34			0.59
			n.a.			n.a.
Observations	1773	1773	1773	1773	1773	1773
lnL (total)	5624	5297	5172	5198	5129	5105
lnL (choices)						

Table 1: Estimates of Structural Parameters for Men

Due to unobserved heterogeneity, however, these estimates can overstate the amount of learning taking place. $E[A_i|X_i]$ may not fully capture all information about future academic performance available to individuals, so the increasing weight placed on cumulative academic performance may simply capture the revelation of private information to the econometrician. Columns (5) and (6) address this concern (and the potential bias of α_A/τ discussed earlier) by allowing for several different unobserved types, each with different levels of academic aptitude, known ex-ante, and preferences for two- and four-year school. The estimates in column (6), which allow for three different types, imply that learning about academic ability continues to occur through the end of college. Controlling for unobserved heterogeneity does not change the learning parameters much.

The types identified in the learning model are slightly different than those revealed in the non-learning model. Relative to type 1, type 2 individuals (7% of the sample) have higher academic aptitude, greater-than-average preference for two-year colleges, and less preference for four-year colleges. Type 3 individuals (59%) reflect students with poor academic aptitude who have lower than expected preference for two- and four-year schools. Accounting for unobserved heterogeneity again increases τ and the scale of most other parameters, though the relationship between expected academic ability and enrollment probabilities (α_A/τ) changes little. The overall model fit also improves when unobserved heterogeneity is permitted. I now discuss model fit more directly.

4.2 Model Fit

To examine model fit, I simulate the grade outcomes and educational choices of individuals in my estimation sample 50 times and compare the predicted outcomes to the actual observed outcomes. I use the estimated parameter values in models (1), (3), (4), and (6) from Table 1. I examine the fit of these models in two ways. First, I compare actual to predicted enrollment outcomes, including initial pathway choice, dropout, and college completion. I then examine the relationship between grade outcomes and subsequent enrollment decisions. It should be noted that if the model contained utility intercepts that differ over time, by school, and by academic performance, then the moments presented below would not constitute a true test of "fit." Such a fully saturated and calibrated model would fit the data perfectly. The model I employ is much more parsimonious.

4.2.1 Enrollment decisions

Figure 6 compares the predicted enrollment decisions to the actual decisions made by individuals in the estimation sample. Overall, the model predictions fit the distribution of actual enrollment decisions reasonably well considering how unsaturated the model is. Predictions from the preferred model, which permits unobserved heterogeneity and learning about academic aptitude, tends to fit the actual data most closely. Forty-two percent of individuals are predicted not to enroll, three percentage points below the actual share. Consequently, enrollment in four-year colleges is over-predicted by three percent. The fraction of individuals enrolling in two-year colleges is identical between actual and predicted. The goodness of initial enrollment decision fit is not surprising since the model includes separate constants for two- and four-year schools in the utility function (α_0^{sj}). If the parameters were estimated using only the initial enrollment decision, these shares would fit exactly.

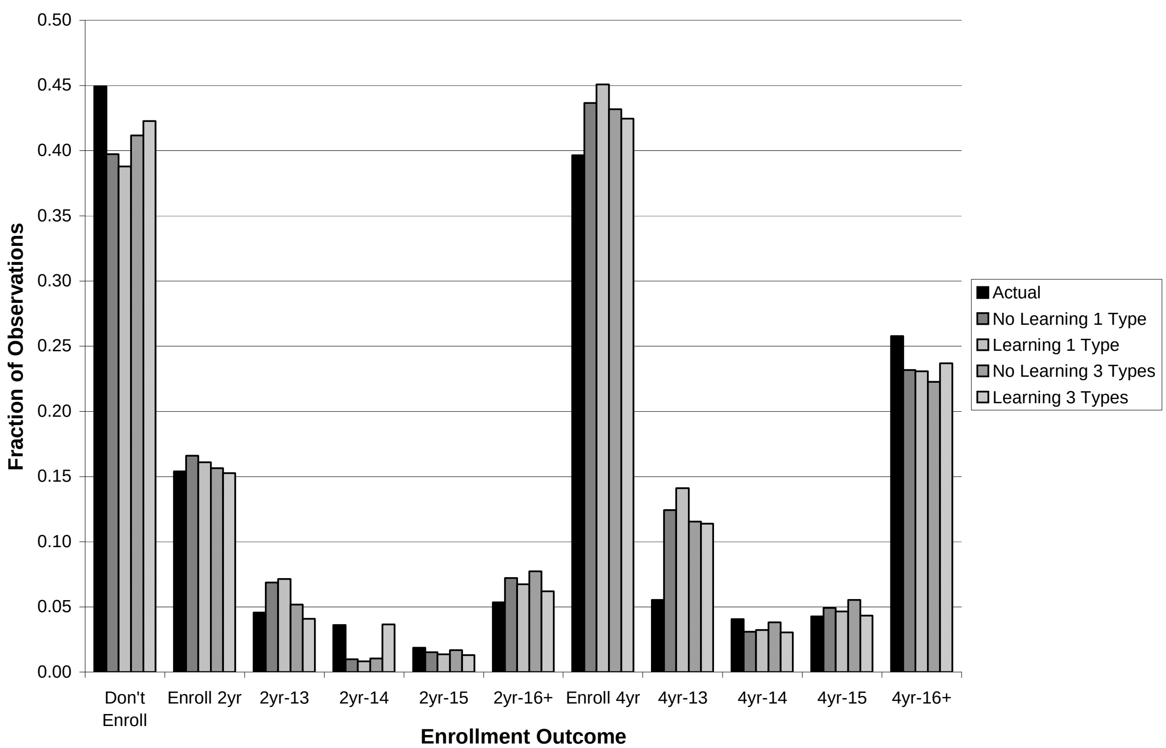


Figure 6

The fit of dropout behavior following initial enrollment decision is a much better test of the ability of the model to predict actual behavior. Since the utility intercepts do not vary over time, predicted differential dropout between different periods is driven entirely by between-period differences in the financial returns (lifetime earnings gain minus costs) and changes in expected academic performance ($E_t[A_i]$).

Figure 7 depicts the fraction of two- and four-year enrollments who drop out in each year or graduate. There are two primary discrepancies between the model predictions and actual outcomes. First, the model slightly underpredicts the fraction of people beginning at community college that drop out after one or three years and consequently over-predicts completion. The second discrepancy is that the model over-predicts dropout after the first year among people that start at a four-year college and consequently underpredicts college graduation.

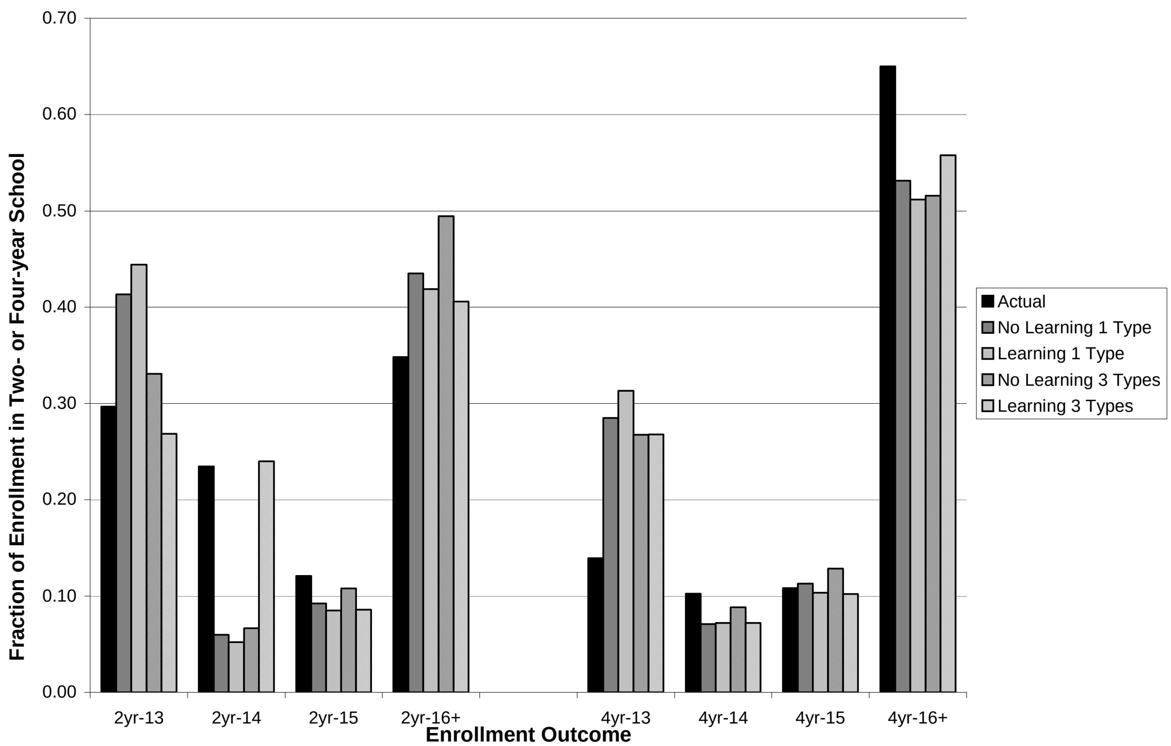


Figure 7

Figures 8 and 9 characterize the accuracy of enrollment predictions by family background. Figure 8 compares actual and predicted enrollment shares by parents' education. Recall that parent's education does not enter individuals' preferences for school directly. Rather, higher parental education increases academic aptitude, which in turn makes schooling more desirable. Also, higher parental education increases predicted lifetime income, which reduces individuals' sensitivity to schooling costs. Despite this restriction, the model captures several important features of the intergenerational correlation of education. Students whose parents have college degrees are much more likely to attend four-year schools and also to graduate from college. Students whose parents do not have a college degree are much less likely to enroll in college or to start at a four-year school, and are much less likely to graduate from college. The model replicates these basic patterns.

Figure 9 presents similar graphs by whether students come from a high- or low-income family. Family income does not enter the model at all, so this is a pretty strong test of model fit. Any correlation between family income and enrollment outcomes must operate through the correlation between family income and the three other background characteristics (high school performance, AFQT, and parental education). Nonetheless, the model still captures several important features of the data, namely the strong positive correlation between family income, college enrollment, and degree completion.

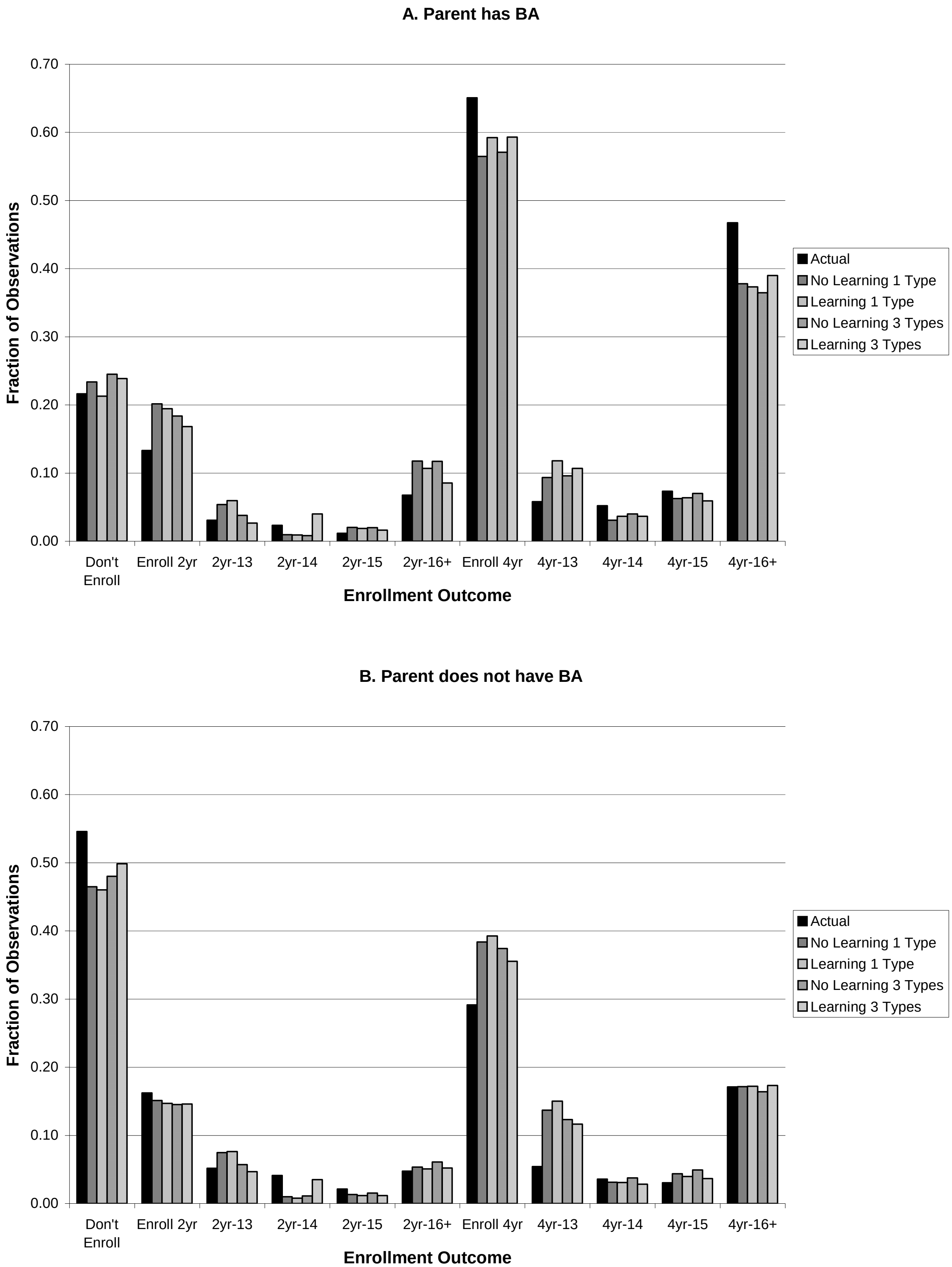

A. Parent has BA
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Fraction of Observations
Don't Enroll
Enroll 2yr
2yr-13
2yr-14
2yr-15
2yr-16+
Enroll 4yr
4yr-13
4yr-14
4yr-15
4yr-16+
Enrollment Outcome
Actual
No Learning 1 Type
Learning 1 Type
No Learning 3 Types
Learning 3 Types
B. Parent does not have BA
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Fraction of Observations
Don't Enroll
Enroll 2yr
2yr-13
2yr-14
2yr-15
2yr-16+
Enroll 4yr
4yr-13
4yr-14
4yr-15
4yr-16+
Enrollment Outcome
Actual
No Learning 1 Type
Learning 1 Type
No Learning 3 Types
Learning 3 Types

Figure 8

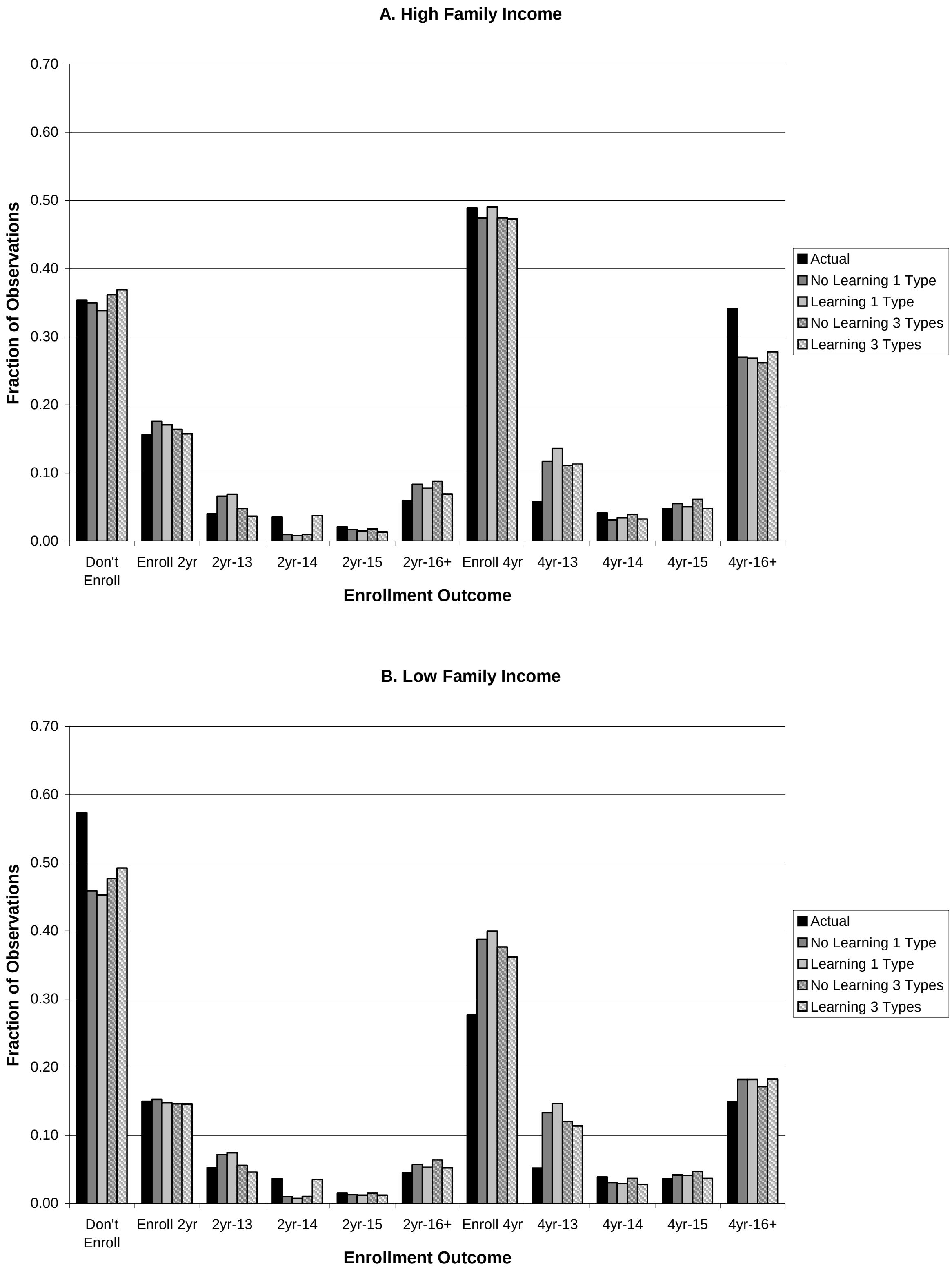

A. High Family Income
Fraction of Observations
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Don't Enroll
Enroll 2yr
2yr-13
2yr-14
2yr-15
2yr-16+
Enroll 4yr
4yr-13
4yr-14
4yr-15
4yr-16+
Enrollment Outcome
Actual
No Learning 1 Type
Learning 1 Type
No Learning 3 Types
Learning 3 Types
B. Low Family Income
Fraction of Observations
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Don't Enroll
Enroll 2yr
2yr-13
2yr-14
2yr-15
2yr-16+
Enroll 4yr
4yr-13
4yr-14
4yr-15
4yr-16+
Enrollment Outcome
Actual
No Learning 1 Type
Learning 1 Type
No Learning 3 Types
Learning 3 Types

Figure 9

4.2.2 Relationship between grades outcomes and enrollment decisions

Enrollment decisions and grade outcomes are related for a number of reasons. First, students with adverse baseline characteristics (e.g., poor grades in high school) have low expected college aptitude, which increases the disutility of school. Consequently, students with low expected academic performance will be less likely to enroll and more likely to dropout if they do enroll. Second, if students learn about the desirability of college through their grades, then students who persist to graduation will have consistently received high (and increasing) grades while those who dropped out will have received low (and decreasing) grades.

Figures 10 and 11 examine the ability of the model to replicate these features of the data. Figure 10 displays the fraction of students that complete their fourth year by their first-year grade point average. The dark line (filled circles) is the actual data, while the others correspond to predictions from the models with and without learning and accounting for unobserved heterogeneity. The preferred model with learning and unobserved heterogeneity (open circles) fits the relationship closest. The overall slope and curvature of the grade-graduation relationship is matched very closely.

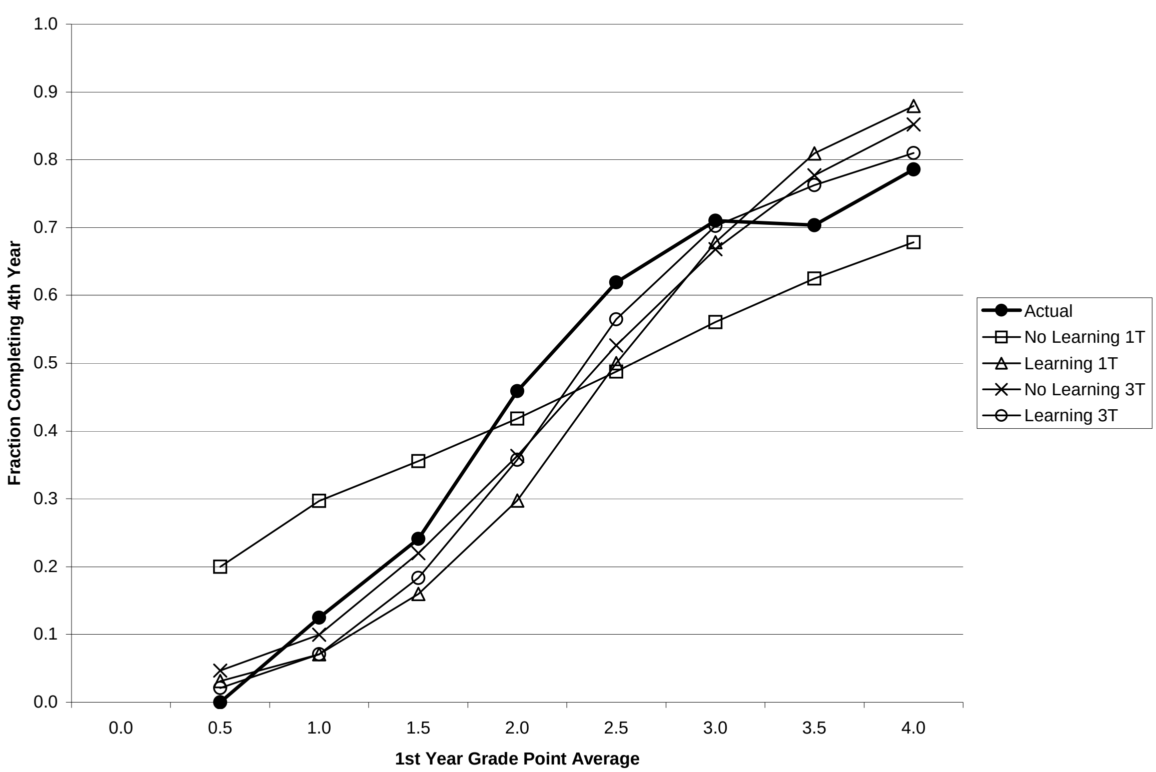


Figure 10

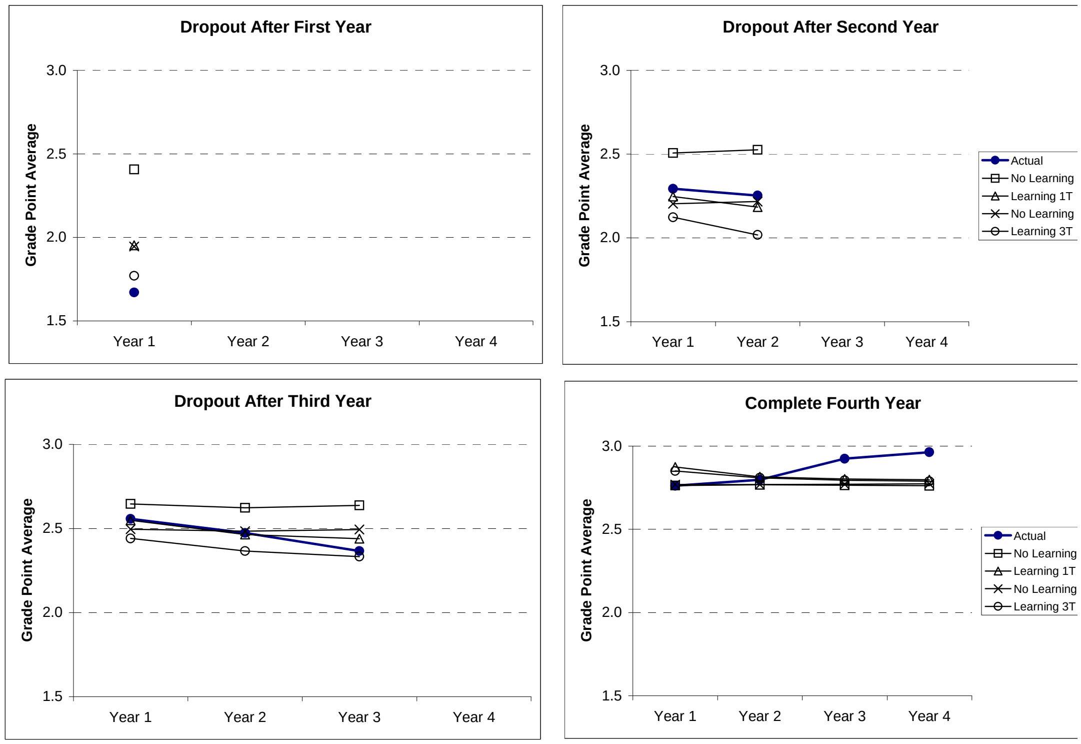
Figure 11 examines the fit of the temporal relationship between grades and education decisions by initial pathway. The figure displays the grade point average in each year by educational outcome (dropout vs completion), separately for people who begin at two- and four-year colleges. The base model with no learning or unobserved heterogeneity provides a poor fit of the data, particularly for students beginning at four-year schools. This model over-predicts grades for those who dropout in the first two years and does not allow grades to evolve over time for those who persist past the first year. Incorporating unobserved heterogeneity makes the levels of predictions more accurate, but does not aid in fitting the temporal pattern of grades. Allowing the expectation of future grades to evolve with past grade realizations through learning addresses this. The preferred model, with three points of unobserved heterogeneity and learning, seems to fit the data best. There are a few characteristics of the data that are not replicated by the model. These include (1) the inverse V among two-year entrants who drop out after their third year; and (2) the extent to which average grades rise over time for those who start at a four-year school and graduate.

4.3 Discussion

To summarize, the parameter estimates suggest that uncertainty is an important feature of postsecondary schooling outcomes. The preferred estimates (column (6) from Table 1) indicate that the deviation of unanticipated shocks to the relative preference for enrollment and labor market entry is equivalent to $83,000 in lifetime earnings. These shocks have the same order of magnitude as the incremental gain from completing a college degree. Thus, unanticipated preference shocks are an important determinant of educational outcomes. The estimates also suggest that students learn about their ex-ante unknown academic aptitude through college grades. Lastly, the estimates suggest that academic aptitude does predict enrollment outcomes and that much of the relationship between family background and schooling outcomes can be captured through the effect of background on academic performance.

Predictions from simulations using the estimated model parameters do match many features of the actual data on enrollments and grade outcomes. The overall distributions of predicted and actual outcomes is roughly similar and the model captures several main features of the relationship between grade outcomes and enrollment decisions. A few limitations of model fit will be addressed in future versions of the model.

A. Four-year College Entrants



B. Two-year College Entrants

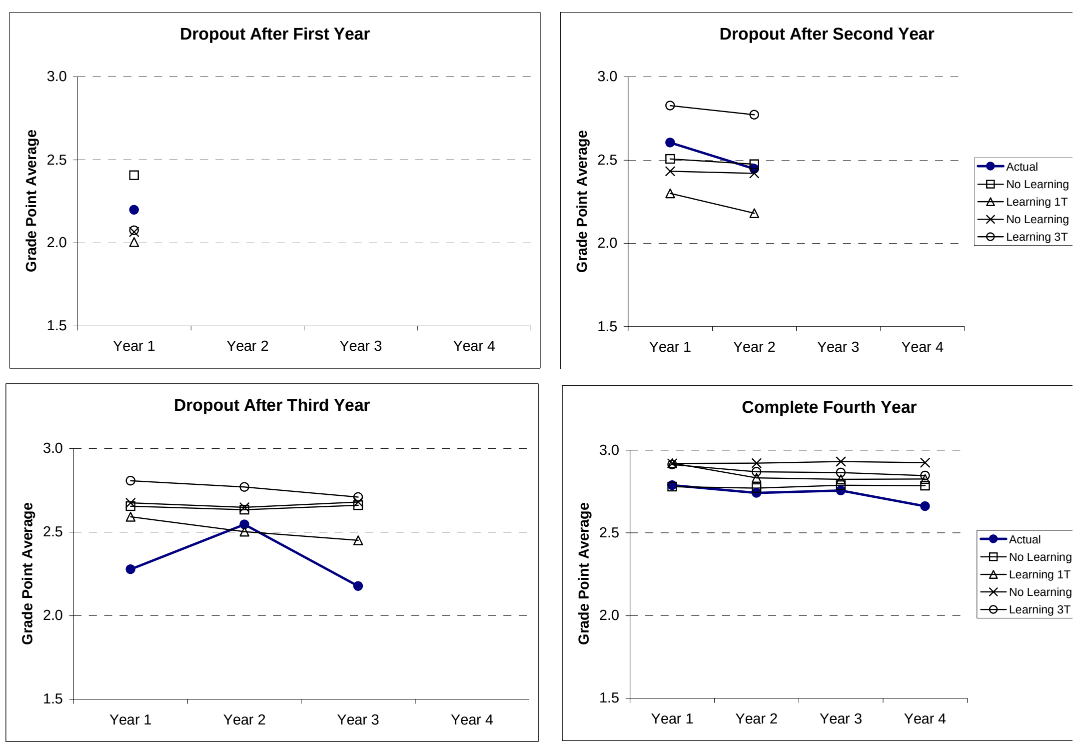


Figure 11

5 The Importance of Option Value

In this section, I estimate the option value created by the ability of students to make educational decisions sequentially and in response to new information. To do this, I treat the estimated structural model as the actual data generating process and simulate educational choices and welfare under alternative assumptions about individuals' information set.[20] In the limited-information static model, I simulate outcomes when individuals are restricted to commit to educational choices before enrolling in college. They base their decision only on information available before college enrollment. This includes baseline characteristics (high school GPA, AFQT, parent education, and type), predicted lifetime earnings, direct tuition and commuting costs, and first-period shocks ($\varepsilon_{i,1}^{S2}, \varepsilon_{i,1}^{S4}, \varepsilon_{i,1}^{w}$). As a basis of comparison, I also simulate the choices and welfare in the first-best scenario, where individuals make decisions with perfect knowledge of all future shocks.

Figure 12 summarizes the importance of option value to educational decisions. The top panel plots the average number of years of college by expected academic ability, separately for the first-best (solid), baseline dynamic (dashed), and limited-information static (dotted) models. This figure is the empirical analog to Panel A of Figure 4, where $E[A_i|X_i, Type]$ is analogous to $\varepsilon_{i,1}$. The static model predicts that education would be much more bifurcated if students were forced to commit ex-ante with limited information. People with low expected ability would get very little education while high ability students would get much more. Compared to the first-best outcome, this bifurcation reduces welfare because some ex-ante low-ability students should go to or graduate from college, while some higher ability students should not. Sequential decision-making permits individuals to come closer to the first-best outcome.

This can be seen more clearly in the middle and bottom panels, which plot the simulated enrollment and graduation rates by expected ability. Option value increases the enrollment rates of all individuals, particularly those in the middle who are on the enrollment margin in the static model. Many of these individuals would choose to enroll if they knew their shocks with certainty but would not if they were forced to commit ex-ante. For low- to moderate-ability students, option value only slightly increases college completion.

[20] To implement the simulations, I first replicate each observation 50 times. For each of these simulated observations, I then draw preference and grade shocks from the appropriately scaled EV(1) and normal distributions, respectively. The optimal choices for each individual are then computed by utility comparisons, incorporating these shocks.

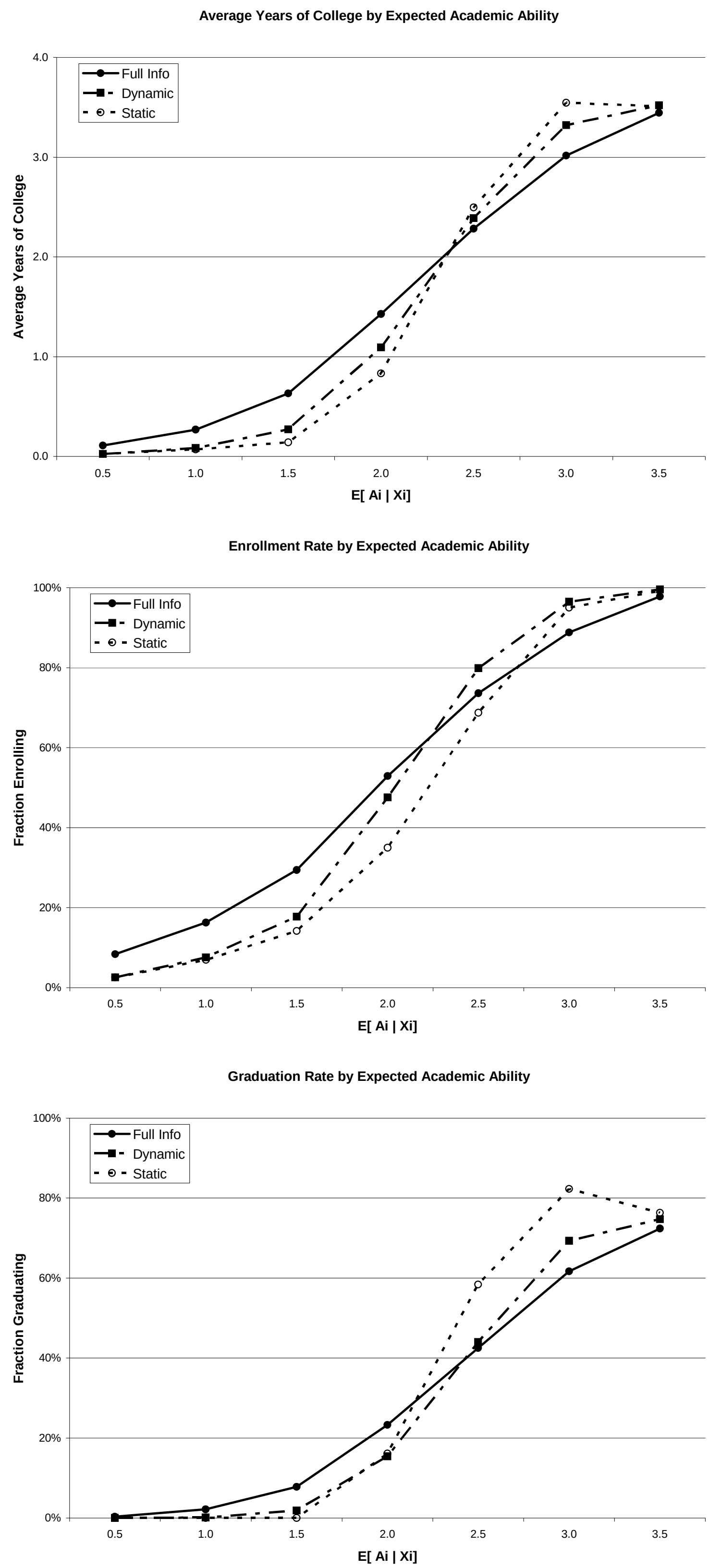

Average Years of College by Expected Academic Ability
Average Years of College
E[Ai | Xi]
Full Info
Dynamic
Static
4.0
3.0
2.0
1.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Enrollment Rate by Expected Academic Ability
Fraction Enrolling
E[Ai | Xi]
100%
80%
60%
40%
20%
0%
Graduation Rate by Expected Academic Ability
Fraction Graduating
E[Ai | Xi]

Figure 12

The biggest effect of option value on completion is to reduce it for high ability students. Some high-ability students expect to graduate - so would commit to doing so ex-ante - but then learn that completion is undesirable and would prefer to drop out. Allowing them to do so reduces completion rates but improves their welfare.

Figure 13 quantifies the value of the flexibility enabled by sequential decision-making. The figure plots the average value of the opportunity to enroll in college by expected academic ability for the same three scenarios and is the empirical analog of Panel B in Figure 3. This value is zero for those who do not enroll. The value of the opportunity to enroll is increasing in expected ability both because enrollment increases with ability and because school is less costly for high ability people, so value conditional on enrollment is also increasing. The vertical distance between the solid and dotted lines represents individuals' total welfare loss from being forced to commit to an educational outcome ex-ante, compared to the first-best situation. This loss is greatest for moderate-ability individuals. Since sequential decision making helps more individuals obtain their optimal level of education, it partially closes this welfare gap, as indicated by the dashed line. The difference between the dashed and dotted lines thus represents the value of the option to dropout whenever continuation turns out to be undesirable.

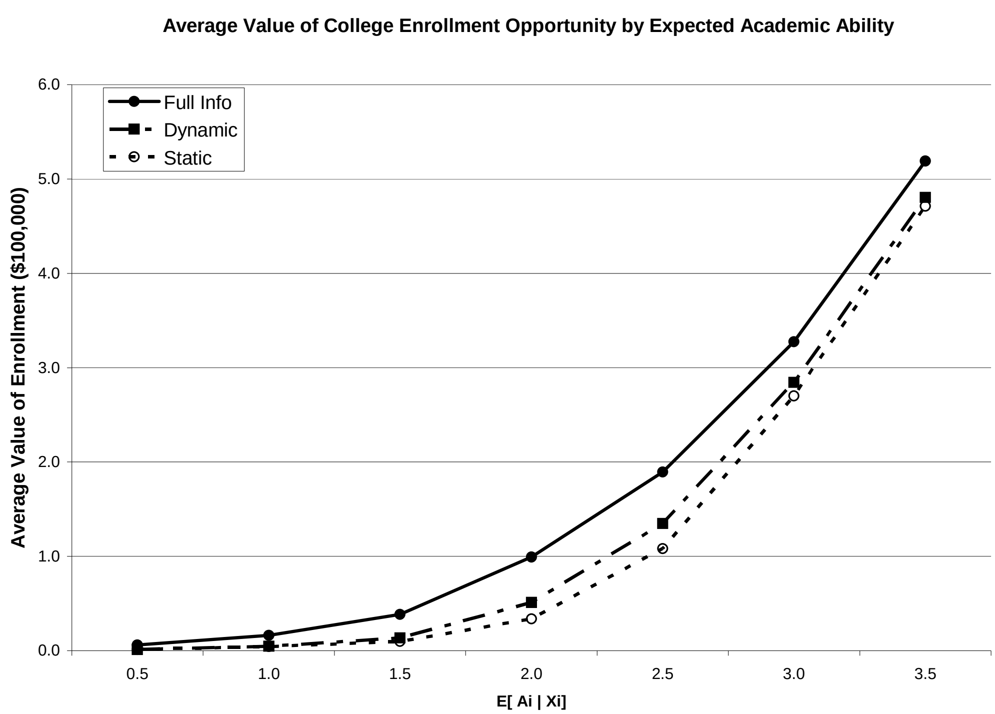


Figure 13

Table 2 presents calculations of the option value by expected ability categories. On average, students would be willing to pay \$15,000 (in 1992 dollars) for the option to dropout. This represents approximately one quarter of the welfare difference between the full-information and limited-information models. This value varies considerably with ability. Moderate-ability students, for whom educational outcomes are most uncertain, are willing to pay up to \$27,000, while the lowest ability students derive virtually no value from the option. The option is worth less to higher ability students because their enrollment decisions do not depend on it.

	Value of Enrollment Opportunity ($1,000)			Welfare Loss	Option Value		
$E[A_i \mid X_i]$	Full Info	Dynamic	Static	(F.I. - Static)	$	% Total	% Loss
0.5	6.2	1.3	1.3	4.9	0.0	0%	0%
1.0	16.5	4.8	4.4	12.2	0.4	8%	3%
1.5	38.5	13.5	9.9	28.6	3.6	27%	13%
2.0	99.3	51.4	33.9	65.4	17.5	34%	27%
2.5	189.7	134.9	108.2	81.5	26.7	20%	33%
3.0	327.7	284.4	270.2	57.5	14.2	5%	25%
3.5	519.2	480.4	471.3	47.9	9.1	2%	19%
All	159.5	117.6	102.3	57.2	15.3	13%	27%

Notes: Value of enrollment opportunity is averaged across 50 simulations

Table 2: Option Value, by Expected Academic Ability

Additional simulations are used to allocate the total option value into the years in which new information is learned. The first three years of college each provide new information about academic ability (in the form of grade signals) and the relative desirability of schooling and work $(\varepsilon^{S2}_{i,t}, \varepsilon^{S4}_{i,t}, \varepsilon^{w}_{i,t})$. To do this decomposition, I simulate educational choices and welfare when individuals are restricted to commit to educational choices before enrolling in college (the static model discussed above), after the first year, after the second year, and after the third year (the baseline dynamic model). Figure 14 summarizes this decomposition. For moderate-ability students, the most valuable information is that which is learned in the first year of college, when the wisdom of their enrollment decision is most uncertain. Higher ability students derive relatively more value from information received later, when graduation decisions are made. Approximately 60% of the total option value derives from information learned in the first year, while the other two years account for about 20% each.

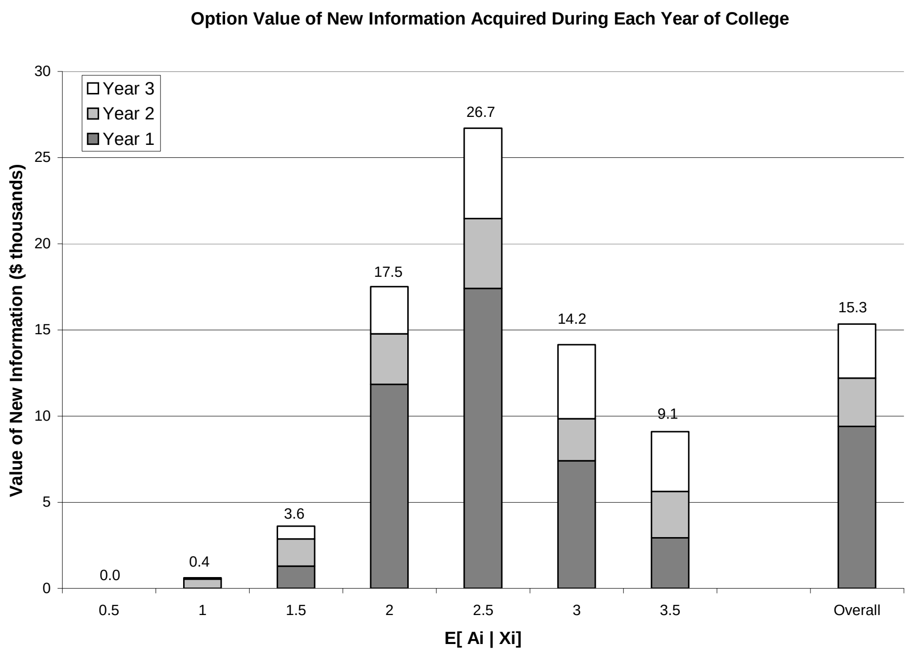


Figure 14

To summarize, the value of the option to dropout is considerable, particularly for moderate ability students who have the most uncertainty about their net benefit from schooling. The option to dropout has value both because it encourages more people to enroll, who may not want to if forced to commit ex-ante, and it because it permits dropout if graduation is undesirable among those who would commit to graduate ex-ante. In aggregate, the former is greater than the latter. Furthermore, the majority of the aggregate option value comes from the information received in the first year of college.

6 Policy Simulations

In this section, I use the estimated model to simulate the effects of four different types of policies on college enrollment, and graduation. The structural approach to policy evaluation provides several benefits over the reduced-form approaches which are more common. First, one can examine the effects of regime-changing policies, such as if community colleges were not to exist. Second, the approach enables the evaluation of many different types of policies simultaneously and on the same sample of individuals. Estimates derived from instrumental variables and natural experiments are well-identified, though are sometimes difficult to compare because samples and policy details differ across studies. Lastly, the structural approach can shed light on the mechanisms through which policy interventions operate. For instance, if a tuition subsidy is found not to affect college graduation, the structural model can determine whether this is due to a limited effect on enrollment, completion, or both.

First I examine reductions in the direct cost of schooling, through lower tuition and reduced commuting costs. Then I quantify the contribution of community colleges to educational outcomes by examining outcomes if community colleges were eliminated as an option. Third I examine the effect of increasing high school preparedness by increasing high school grades up to a 2.5 GPA. Finally, I examine the effect of a $10,000 graduation bonus for students from families with below-average income. Since the structural model is dynamic, option value is explicitly incorporated when these policies are evaluated. To examine the importance of this inclusion, I conclude by contrasting the policy effects predicted using the dynamic and static models.

6.1 Reduced Tuition

Tuition fee levels are the most obvious and debated higher-education policy. Since the vast majority of college students attend public two- and four-year colleges, state policy-makers (legislatures, higher education boards, university regents) have significant influence over the out-of-pocket costs of most college students.[21]

One implication of the theoretical model is that the responsiveness of enrollment and completion to direct tuition fees depends on the magnitude of preference shocks. The more that decisions are driven by these shocks, the less sensitive they will be to policies that affect tuition. My estimates imply that direct costs are generally small relative to the magnitude of these shocks and relative to the lifetime wage differentials between education outcomes. Therefore, it is not surprising that policy simulations suggest that educational outcomes are relatively unaffected by changes in direct costs.

Figure 15 contrasts the estimated educational outcomes for two policies that substantially reduce direct costs. The first policy sets tuition fees at public two- and four-year colleges to zero. Enrollment and completion both increase by less than one percentage point. On a percentage basis, this effect is slightly greater for students from low income families. As summarized in Kane (2006), previous studies typically find that a $1,000 change in college costs ($1990) is associated with an approximately 5 percentage point difference in college enrollment rates, though some studies find much smaller effects.

[21] For a survey of estimates of the impact of tuition subsidies on enrollment, see Kane (2006).

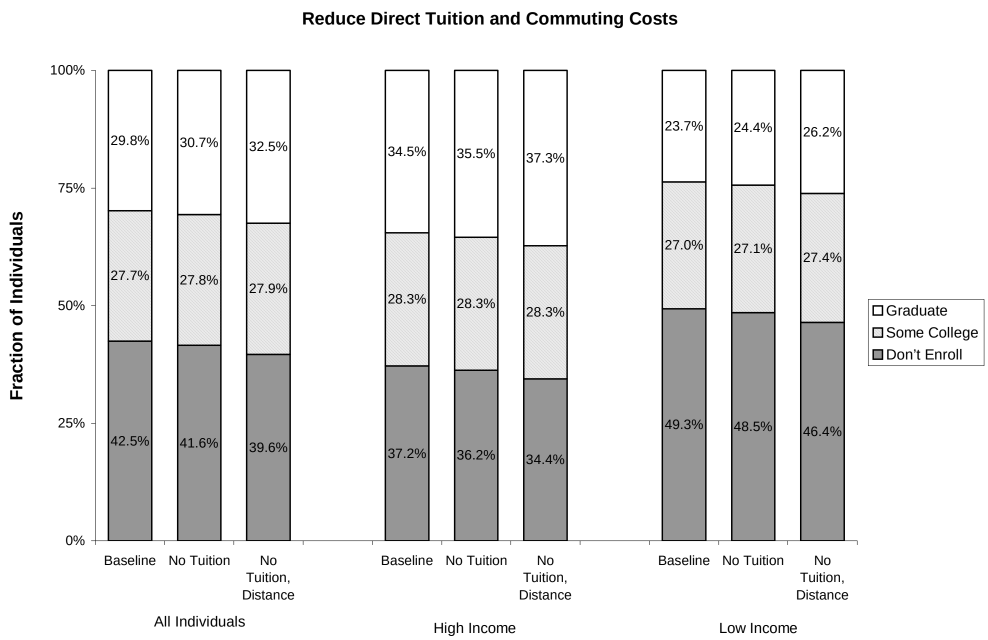


Figure 15

The parameter estimates suggest that commuting costs can be very significant. For instance, living 100 miles from a college is equivalent to having to pay an additional $22,000 per year for the typical student. Compensating students for this inconvenience (and ignoring family's location decisions) would increase enrollment and completion by an additional two percentage points. Again, the effect is slightly stronger for students from low income families.

6.2 Community Colleges

The past few decades have witnessed a considerable expansion of community colleges: community colleges absorbed nearly half of the increase in college enrollment between 1980 and 1994, accounting for 38% of all postsecondary enrollments by 1995 (Kane and Rouse, 1999). The net effect of community colleges on educational attainment is theoretically ambiguous due to offsetting democratization and diversion effects. On the one hand, the accessibility of community colleges provides opportunities to students that otherwise would not attend college, expanding educational attainment. On the other hand, some critics of community colleges argue that community colleges do not adequately prepare or facilitate the transfer of high-ability students, diverting them from obtaining a Bachelor's degree.

The empirical work at the individual level generally finds that community colleges increase overall educational attainment by making postsecondary schooling available to those who otherwise would not attend college, but the educational attainment of some students who would attend anyway is also lowered by the difficulty of transferring from community colleges.[22] At the state level, Rouse (1998) concludes that community colleges potentially provide an cost efficient way of increasing overall educational attainment, but have little effect on degree completion.

To quantify the effect of community colleges on educational outcomes, I compare the counterfactual scenario where the community college option is eliminated to the status quo, where individuals can attend community college. Figure 16 presents the educational outcomes under these two scenarios. Similar to previous research, I find that community colleges greatly expand postsecondary participation, but have limited effect on college completion. Community colleges increase overall enrollment by more than seven percentage points - a fifteen percent increase - but increase college completion by only one half of a percentage point. The effects are proportionately larger for students from low-income families. To the extent that college enrollment - and not just completion - has value, community colleges can significantly increase welfare.

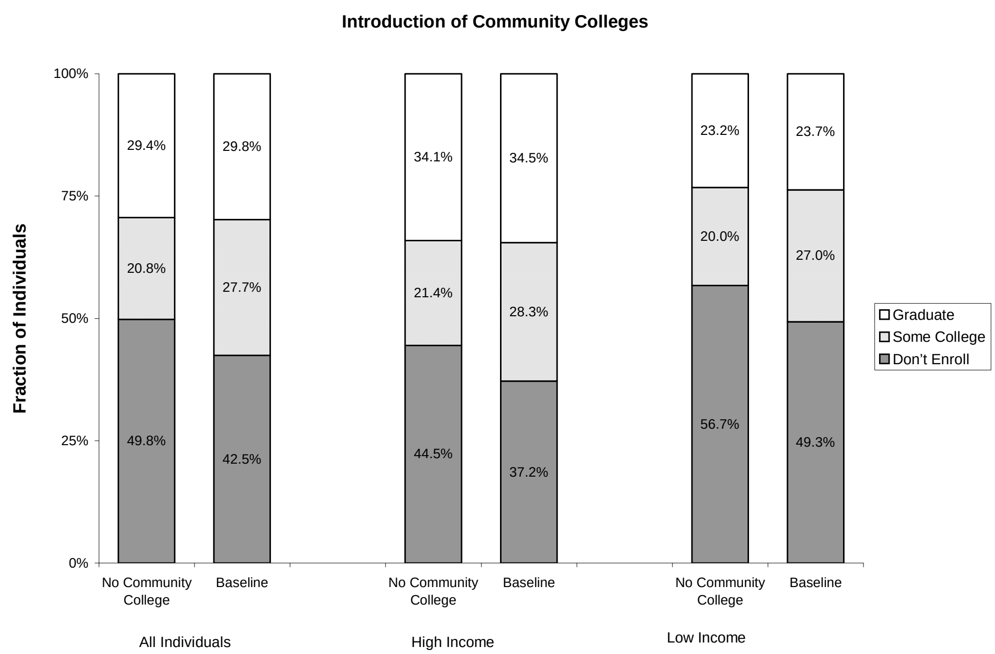


Figure 16

[22] Rouse (1995) estimates that starting at a community college reduces the educational attainment of some students by 0.5 to 1.0 total years of schooling, but does not affect Bachelor's degree attainment. Leigh and Gill (2003) reach conclusions when controlling for initial educational aspirations. In follow-up work, Leigh and Gill (2004) conclude that this diversion effect operates partially through reduced educational aspirations. Gonzalez and Hilmer (2006) find that 2-year colleges have an unambiguously positive effect on the educational attainment of Hispanics. On the other hand, Anderson (1981) and Velez (1985) note that students at community colleges are much less likely to persist to a 4-year degree and transfers from community colleges have lower graduation rates than do students that attend 4-year colleges directly.

6.3 Improve Academic Preparedness

There is widespread concern about the preparedness of American youth for college. In 2000, 28 percent of all entering college freshmen enrolled in at least one remedial mathematics, reading, or writing class. Lack of preparation is particularly acute for students who begin college at community colleges, where remedial participation exceeds forty percent [NCES, 2004]. Lack of student preparedness is even more problematic if college remediation is ineffective, as suggested by Martorell and McFarlin (2007).

To examine the importance of academic preparedness, I simulate the consequence of a policy that increases the high school grade point average of low-achieving students up to a 2.5 (approximately the median). In the model, an increase in high school achievement raises the mean of collegiate grades, which in turn increases the desirability of college attendance. The results are presented in Figure 17.

Improved academic achievement in high school increases both college enrollment and completion. Overall, enrollment increases by 3.8 percentage points and two-thirds of this gain is reflected in increased completion. Effects are much larger, both on an absolute and relative basis, for low-income students, who are much more likely to be ill-prepared for college. Among low income students, this policy would increase enrollment by 4.8 percentage points (nine percent) and the fraction who complete college by 3.2 percentage points - a thirteen percent increase.

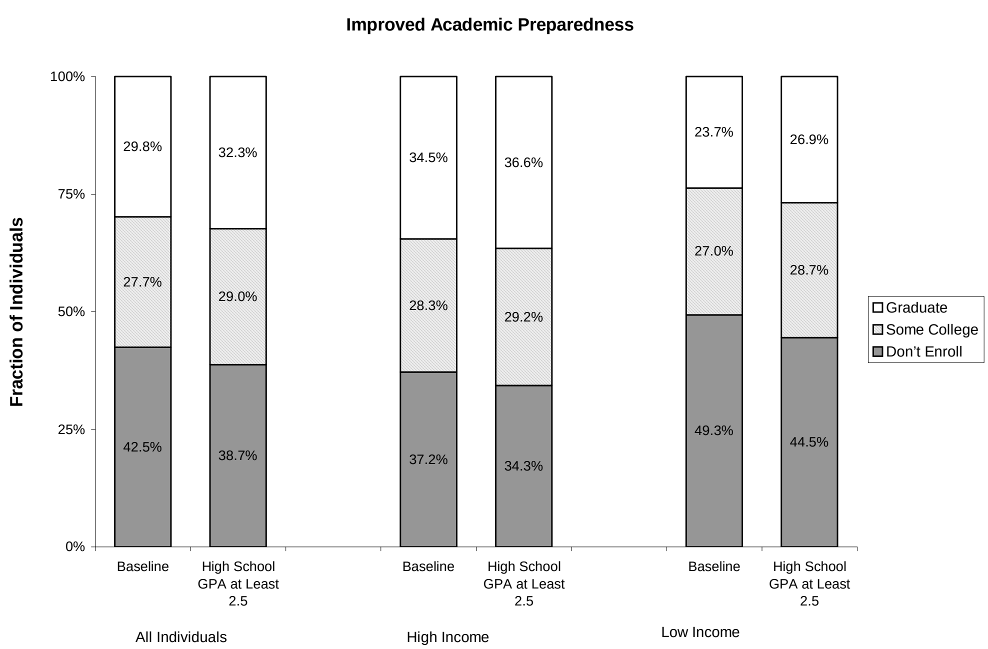


Figure 17

6.4 Graduation-Contingent Bonus

As an alternative to subsidized tuition, governments could offer explicit incentives to graduate. Graduation-contingent bonuses have been discussed primarily as a means for reducing high school dropout, but have occasionally been proposed with regards to postsecondary student aid. For instance, Fischer (1987) proposed giving "unsubsidized loans while the student is in school, with subsidies provided in the form of partial loan cancellation only after degree attainment." Completion subsidies increase the marginal return to graduation. This directly reduces dropout after the third year of college, but also increases the option value of college enrollment and continuation after the first and second years. Keane and Wolpin (2000) found that such a high school graduation bonus would equalize schooling outcomes by race.

In Figure 18, I report the results of a $10,000 graduation bonus targeted towards students from low-income families. Such a policy would have negligible effect on enrollment but would increase the fraction of low-income students who graduate college by about one percentage point.

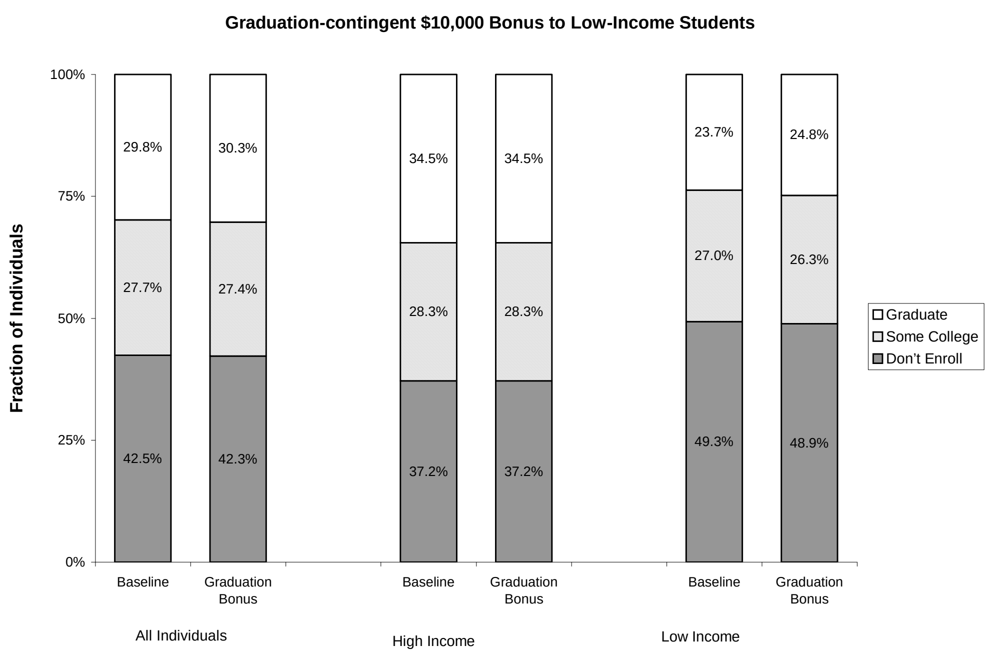


Figure 18

6.5 Dynamic vs. Static Policy Predictions

All of the above education policies have a temporal dimension. For instance, the graduation bonus directly alters the financial gain to the final year of college but not the first three. Community colleges explicitly alter the tuition gradient by making the first half of college cheaper than the last half. If students are forward-looking these changes will be considered during enrollment decisions and they will alter the continuation value associated with each year of college. An extension to the theoretical section, included in the Appendix, suggests that predicted policy effects may depend on whether decisions are assumed to be made in a static or dynamic fashion. This section examines whether the use of a static framework provides misleading predictions about policy effects. I simulate educational choices under the four policies described above, but assuming that people must commit to their decision before enrolling in college. Table 3 compares the estimated program impact on enrollment rates, graduation rates, and average years of school between the static and dynamic models.

	Estimated Change in Educational Outcomes					
	Overall		High Income		Low Income	
Community College	Dynamic	Static	Dynamic	Static	Dynamic	Static
Fraction Enrolling	7.4%	7.0%	7.3%	7.0%	7.4%	6.9%
Fraction Graduating	0.4%	0.4%	0.4%	0.3%	0.5%	0.4%
Years of School	0.13	0.13				
Elimination of Direct Costs						
Fraction Enrolling	2.8%	2.8%	2.8%	2.8%	2.9%	2.9%
Fraction Graduating	2.6%	3.2%	2.8%	3.2%	2.5%	3.0%
Years of School	0.11	0.12				
Improve Preparedness						
Fraction Enrolling	3.7%	3.8%	2.9%	3.0%	4.8%	4.9%
Fraction Graduating	2.5%	4.3%	2.0%	3.1%	3.1%	5.9%
Years of School	0.13	0.18				
Graduation Bonus						
Fraction Enrolling	0.2%	0.3%	0.0%	0.0%	0.4%	0.6%
Fraction Graduating	0.5%	0.4%	0.0%	0.0%	1.1%	1.0%
Years of School	0.01	0.01				

Table 3: Simulated Policy Effects: Dynamic vs. Static Model

While predictions are qualitatively similar, a static model may misstate the magnitude of some policy effects. Compared to the dynamic model, the static model slightly under-predicts the effect that community colleges have on expanding enrollment. In addition to making college less expensive, community colleges increase the option value of enrollment because dropout is less costly so more people experiment. The static model does not fully incorporate this added benefit. In contrast, the static model over-predicts the effect of across-the-board tuition reduction on college completion. The static model predicts a more bimodal education distribution, with many non-enrollees and graduates, but few dropouts. Consequently, more enrollees are predicted to continue through to graduation. The static model also over-predicts the graduation consequences of increasing academic preparedness in high school. Expected performance in college - which depends heavily on high school GPA - has much more influence on educational outcomes in the static model. If decision-making is dynamic, then less weight is placed on ex-ante expected performance as new information is acquired during each year.

6.6 Discussion

Simulations suggest that the effects of four broad classes of policies, all of which aim to increase postsecondary participation and completion, differ substantially in nature and in magnitude. Extensive tuition subsidies have modest effects on enrollment which translate one-for-one to modest effects on college graduation. The effects are mostly broad-based, impacting low-income students only slightly more than high-income students. Consistent with earlier research, community colleges have a large effect on the fraction of students who receive some postsecondary education, but the effect on college graduation is negligible. Low-income students are particularly influenced by the presence of community colleges.

Improvements to the K-12 system, which bring the least-prepared high school graduates up to the median have large effects on both enrollment and completion. Enrollment and completion effects are particularly strong for low-income students, who are often the least prepared for college. This finding is consistent with Cameron and Heckman (2001), who conclude that family background has a tremendous influence on educational attainment through college readiness, rather than through credit constraints. Finally, providing a $10,000 bonus to low-income students who graduate from college would increase the graduation rate of this group slightly, but have little impact on enrollment.

Two implications of the estimated model are worth noting. First, uncertainty is an important feature of educational decisions and failing to account for it may provide misleading estimates of policy effects. This is particularly true when evaluating policies that have different temporal characteristics, such as community colleges (which alter the tuition gradient) or across-the-board tuition reductions (which do not). A second implication is that academic performance is one major channel through which background characteristics - such as parents' education or income - influence educational outcomes. Policies which directly affect students' academic aptitude - such as improving college preparation - can substantially reduce socioeconomic gaps in educational outcomes.

7 Summary and Conclusions

This paper examines the empirical importance of uncertainty and option value to college enrollment. It is the first to quantify the magnitude of the option value that arises when individuals make decisions to invest in a college education sequentially and when the desirability of doing so is uncertain. Estimates suggest that this value is substantial. In contrast to a scenario where individuals must commit to an educational outcome ex-ante, the current flexible system increases welfare by $15,000 on average. This represents 13% of the overall value of the opportunity to enroll in college. Moderate-ability students, who are more reluctant to enroll, derive even more value from this flexibility. The traditional human capital model ignores this value.

The finding that flexibility substantially improves welfare has direct implications for the potential costs of student "tracking." This paper suggests that, at least in the U.S. postsecondary context, students learn quite a bit about their ability in the first few years of college. Forcing students to commit ex-ante will make educational outcomes more polarized by background characteristics and reduce welfare. This welfare loss must be weighed against any efficiency gains resulting from tracking.

The general framework developed herein could also be used in a number of different contexts in which decisions are partially irreversible and made in the presence of uncertainty. One potential application is the use of "exploding" job offers. Firms hiring many law or business school graduates force students to commit to a job early in the fall, possibly before their industry or locational preferences are finalized. The model implies that firms would have to compensate individuals for this loss of flexibility, through a signing bonus or higher salary. Investments in health can also be understood as motivated by option value considerations. Since many health conditions (e.g. weight gain, diabetes onset, lung cancer) are partially irreversible, forward-looking individuals should make costly health investments when young in order to preserve the option of being healthy when old. Subsidies for preventative care, a healthy diet, and exercise among the young can be rationalized by this option value if individuals are not completely forward-looking.

On the specific topic of college enrollment, the current specification could also be extended in a number of directions. First, the model could be modified to allow for discount rate heterogeneity by introducing a fourth unobserved "type" with a very large discount rate (say 50%). In a separate paper, I show that front-loaded tuition subsidies are optimal if students are not fully forward-looking and do not consider the informational value of enrollment. Consequently, the presence of individuals with high discount rates may influence the outcomes and welfare consequences of the examined policies, as well as the option value estimates.

A second extension is to isolate the contribution of the various sources of option value. In the current form of the model, enrollment has option value because it (1) reveals information about college aptitude and the persistent psychic costs (or benefits) of school attendance; (2) reveals information about the short-term (nonpersistent) relative desirability of school attendance; and (3) allows individuals to delay labor market entry until receiving a high labor market offer. The current specification does not permit the estimation of the contribution of each of these components. Allowing individuals to receive labor market offers while not in school would separate the third source from the first two.

Lastly, the model could be extended to the high school and graduate school decisions, as well as modified to permit delayed college entry and stop out. High school and college graduation both have option value because they enable college and graduate school enrollment, respectively. The ability to drop out and return to college (referred to as "stop out") or to delay college entry are also valuable. The current analysis does not consider these additional sources of option value but future work should.

References

[1] Altonji, J., 1993. "The Demand for and Return to Education When Education Outcomes Are Uncertain." *Journal of Labor Economics*, Volume 11: 48-83.

[2] Anderson, Kristine L., 1981. "Post-High School Experiences and College Attrition." *Sociology of Education*, Volume 54 (January): 1-15.

[3] Arcidiacono, Peter, 2004. "Ability Sorting and the Returns to College Major." *Journal of Econometrics*, Volume 121 (2004): 343-375.

[4] Arcidiacono, Peter, 2005. "Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?" *Econometrica*, Volume 73 (5): 1477-1524.

[5] Buchinsky, Moshe and Phillip Leslie, 2000. "Educational Attainment and The Changing U.S. Wage Structure: Dynamic Implications Without Rational Expectations." mimeo, University of California Los Angeles.

[6] Cameron, Stephen and James J. Heckman, 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy*, Volume 109 (3): 455-499.

[7] Card, David, 1999. "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Volume 3A, ed. by O. Ashenfelter and D. Card. New York: Elsevier, 1801-1864.

[8] Chen, Stacey, 2007. "Estimating the Variance of Wages in the Presence of Selection and Unobservable Heterogeneity" *Review of Economics and Statistics*, forthcoming

[9] Comay, Y., A. Melnik, M.A. Pollatschek, 1973. "The Option Value of Education andthe Optimal Path for Investment in Human Capital," *International Economic Review*. Vol. 14 (2). p 421-35.

[10] Cunha, Flavio, James Heckman, and Salvador Navarro, 2005. "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, Oxford University Press, vol. 57(2), pages 191-261, April

[11] Dee, Thomas, 2004. "Are There Civic Returns to Education?" *Journal of Public Economics* 88(9), August 2004, pages 1697-1720.

[12] Dixit, A.K., R.S. Pindyck, 1994. *Investment under Uncertainty*, Princeton University Press, Princeton, N.J.

[13] Dothan, Uri and Joseph Williams, 1981. "Education as an Option," *Journal of Business,* Vol. 51 No. 1.

[14] Eckstein, Z. and K.Wolpin (1999): "Why Youths Drop out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, Vol. 67 (6). p 1295-1339.

[15] Fischer, F., 1987. "Graduation-Contingent Student Aid." *Change*, p 40-47.

[16] Gonzalez, Arturo and Michael J. Hilmer, 2006. "The Role of 2-year Colleges in the Improving Situation of Hispanic Postsecondary Education." *Economics of Education Review.* forthcoming.

[17] Hungerford, T. and Solon, G. 1987. "Sheepskin Effects in the Returns to Education," in *Review of Economics and Statistics*, Vol. 69, 175-177.

[18] Jaeger, D. and M. Page, 1996. "Degrees Matter: New Evidence on Sheepskin Effects in Returns to Education" in *Review of Economics and Statistics*, Vol. 78.

[19] Kane, Thomas J. 2006. "Public Intervention in Postsecondary Education" in Eric Hanushek and Finis Welch (eds.) *Handbook on the Economics of Education* (Amsterdam: Elsevier/North Holland, 2006).

[20] Kane, Thomas J. and Cecilia Elena Rouse, 1995. "Labor Market Returns to Two- and Four-Year College," *The American Economic Review*, Volume 85 (3): 600-614.

[21] Kane, Thomas J. and Cecilia Elena Rouse, 1999. "The Community College: Educating Students at the Margin Between College and Work," *The Journal of Economic Perspectives*, Volume 13 (1): 63-84.

[22] Keane, Michael P. and Kenneth I. Wolpin, 2000. "Eliminating Race Differences in School Attainment and Labor Market Success" *Journal of Labor Economics*, Volume 18 (4): 614-652.

[23] Kilburn, M. Rebecca, Lawrence M. Hanser, and Jacob A. Klerman, 1998. "Estimating AFQT Scores for NELS Respondents," RAND, MR-818-OSD/A, 1998.

[24] Levhari, David, and Yoram Weiss, 1974. "The Effect of Risk on the Investment in Human Capital," *American Economic Review* 64 (Dec. 1974), 950-963.

[25] Leigh, D.E. and A.M. Gill, 2003. "Do Community Colleges Really Divert Students from Earning Bachelor's Degrees?" *Economics of Education Review*, Volume 22: 23-30.

[26] Leigh, D.E. and A.M. Gill, 2004. "The Effect of Community Colleges on Changing Students' Educational Aspirations" *Economics of Education Review*, Volume 23: 95-102.

[27] Manski, Charles, 1989, "Schooling as Experimentation: A Re-Appraisal of the Postsecondary Dropout Phenomenon." *Economics of Education Review,* Volume 8 (4): 305-312.

[28] Manski, Charles, 1991. "Nonparametric Estimation of Expectations in the Analysis of Discrete Choice Under Uncertainty" in W. Barnett, J. Powell, and G. Tauchen (editors), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, New York: Cambridge University Press, 1991, pp. 259-275.

[29] Martorell, Paco and Isaac McFarlin, 2007. "Help or Hindrance? The Effects of College Remediation on Academic and Labor Market Outcomes." preliminary working paper April 2007.

[30] Milligan, Moretti, and Oreopoulos Does Education Improve Citizenship? Evidence from the U.S. and the U.K." *Journal of Public Economics* 88(9-10), 2004

[31] Moretti 2004 "Estimating the Social Return to Higher Education: Evidence From Longitudinal and Repeated Cross-Sectional Data" *Journal of Econometrics* 121(1-2), 2004

[32] National Center for Educational Statistics (NCES). 2003. *Remedial Education at Degree Granting Postsecondary Institutions in Fall 2000*, NCES 2004-010, by Basmat Parsad and Laurie Lewis. Project Officer: Bernard Greene. Washington, DC.

[33] National Center for Educational Statistics, *2004 Digest of Educational Statistics,* Table 307.

[34] Olson, Lawrence; White, Halbert; and Shefrin, H.M. "Optimal Investment in Schooling When Income Are Risky." *Journal of Political Economy*, 87(3), 1979, pp. 522-39.

[35] Park, Jin Heum, 1999. "Estimation of Sheepskin Effects Using the Old and the New Measures of Educational Attainment in the Current Population Survey," *Economics Letters*, Volume 62: 237-240.

[36] Rouse, Cecilia, 1995. “Democratization or Diversion? The Effect of Community Colleges on Educational Attainment.” *Journal of Business and Economic Statistics*, Volume 13 (2): 217-224.

[37] Rouse, Cecilia, 1998. “Do Two-Year Colleges Increase Overall Educational Attainment? Evidence from the States.” *Journal of Policy Analysis and Management,* Volume 17 (4): 595-620.

[38] Stange, 2007 "Optimal College Tuition When Completion is Uncertain." mimeo. UC Berkeley.

[39] Taber 2000 “Semiparametric Identification and Heterogeneity in Dynamic Programming Discrete Choice Models,” *Journal of Econometrics*, June 2000

[40] *The New York Times* Education Life Supplement, Page 24, April 23, 2006.

[41] Velez, William, 1985. “Finishing College: The Effects of College Type,” *Sociology of Education*, Volume 58 (July): 191-200.

[42] Weisbrod, B. (1962): “Education and Investment in Human Capital,” *The Journal of Political Economy*, Vol. 70, No. 5.

[43] Williams. Joseph T., 1979. "Uncertainty, and the Accumulation of Human Capital over the Life Cycle," *Journal of Busrness* 52 (4), 521-548.

[44] Wiswall, Matthew, 2005. "Balancing Training and Recruitment: The Effects of Teacher Certification Laws on Teacher Quality." mimeo, University of California Los Angeles.

8 Appendix

8.1 Dataset Construction

The dataset used in estimation and simulation was constructed from several sources. Table A1 provides an overview of the main variables used in the analysis. The sample of individuals comes from the National Educational Longitudinal Study (NELS). The NELS is a longitudinal survey of a representative sample of U.S. 8th graders in 1988. Interviews were conducted in 1988, 1990, 1992, 1994, and 2000 and complete college transcripts were obtained for most individuals in 2000. The core schooling outcome variables, including yearly grade point average and indicators for enrollment were constructed directly from the college transcripts. The transcripts consist of course-specific records, including student ID, college IPEDS ID number, subject, month and year, credits, letter grade, and standardized numeric grade on a four-point scale. Course-level records were aggregated up to the student x college x term level to identify the primary school enrolled in, and then to the student x year level. The final transcript data contains student x year records of credits attempted, credits earned, grade point average, and several other variables. Individuals were considered enrolled during academic year t if they attempted at least six course credits (the traditional definition of part-time enrollment) during both the Fall and Spring semesters of year t. The model is a model of college dropout, so I categorize people according to their number of years of continuous enrollment. Students who "stop-out," but eventually return and possibly graduate are grouped with students who dropout permanently in the same year. From the 1992 NELS surveys I utilize high school grade point average, standardized test scores, parents' highest education level, and family income during high school. I convert NELS senior year test scores into AFQT percentile scores using the cross-walk developed by RAND researchers in Kilburn, Hanser, and Klerman (1998).

I supplemented the NELS dataset with institutional characteristics obtained from The College Board's 1992 Annual Survey of Colleges (ASC) and the Department of Education's 1992 Integrated Postsecondary Education Data System (IPEDS) Institutional Characteristics survey. Both survey the universe of public and private two- and four-year colleges in the United States. IPEDS has a higher response rate, but a much more limited number of data fields. To minimize missing values, I merged the two datasets by IPEDS ID number. From the IPEDS, I calculated average tuition levels at public two-year and four-year colleges in each state and merged this data onto the NELS. Latitude/longitude coordinates were then assigned to each college in IPEDS/ASC and high school in the NELS by zip code from the US Census 1990 Gazetteer Files (http://www.census.gov/geo/www/gazetteer/gazette.html). From this, I calculated distance from each NELS high school to the nearest public two-year and four-year college (in miles). I also assigned performance measures (e.g. graduation rates and transfer rates) of the nearest two-year college to each NELS high school. Table A1 describes the main variables used in the analysis and Table A2 displays summary statistics.

Table A1: Variable Descriptions and Sources

Variable	Description	Source
1. hsgpa	Cumulative grade point average in high school on 4.0 scale	NELS.
2. afqt	Armed Forces Qualifying Test percentile score	Constructed from NELS test score variables using method developed by RAND (see text).
3. pareduc	Years of school attended by most educated parent	NELS.
4. parba	Indicator for whether at least one parent earned a BA degree	NELS. Constructed from pareduc variable.
5. lowincome	Indicator for whether family income during high school was below $35,000 (approximately the median)	NELS. Constructed from faminc variable.
6. urban	Attended urban high school	NELS. Constructed from phsurban variable.
7. regionne	High school in Northeast	NELS. NLSY categorization.
8. regionnc	High school in Northcentral	NELS. NLSY categorization.
9. regionso	High school in South	NELS. NLSY categorization.
10. regionwe	High school in West	NELS. NLSY categorization.
11. white	Ethnicity white	NELS
12. black	Ethnicity black	NELS
13. latino	Ethnicity latino	NELS
14. distance2	Distance from high school to nearest public two-year college.	Computed from lat/long coordinates of high school (NELS) and each public 2-year college in state (ACS and IPEDS)
15. distance4	Distance from high school to nearest public four-year college.	Computed from lat/long coordinates of high school (NELS) and all public 4-year college in state (ACS and IPEDS)
16. tuition2	Average tuition ($1992) of public two-year colleges in high school state	IPEDS
17. tuition4	Average tuition ($1992) of public four-year colleges in high school state	IPEDS
18. income1	Expected present discounted value of lifetime income if do not enter college in first year after high school. (thousands of $1992)	Estimated using out-of-sample prediction from NLSY (see text).
19. income2	Expected present discounted value of lifetime income if exit college after first year (thousands of $1992)	Estimated using out-of-sample prediction from NLSY (see text).
20. income3	Expected present discounted value of lifetime income if exit college after second year (thousands of $1992)	Estimated using out-of-sample prediction from NLSY (see text).
21. income4	Expected present discounted value of lifetime income if exit college after third year (thousands of $1992)	Estimated using out-of-sample prediction from NLSY (see text).
22. income5	Expected present discounted value of lifetime income if complete four years of college (thousands of $1992)	Estimated using out-of-sample prediction from NLSY (see text).
23. gpa(t)	Grade point average during year (t) of college	Computed from NELS college transcripts for all courses taken for credit (including failures).
24. enroll(t)	Indicator for enrollment in college during year (t)	Computed from NELS college transcripts. Individual must have attempted at least six units of college credit (approx part-time) in each semester during year (t).
25. contenroll	Years of continuous enrollment in college after high school graduation.	Constructed from enroll(t).
26. fouryear(t)	Indicator for enrollment in four-year college during year (t)	Constructed from enroll(t) and college type from IPEDS. Equals one if enroll(t) = 1 and enrolled in a four-year school in either semester
27. twoyear(t)	Indicator for enrollment in two-year college during year (t)	Constructed from enroll(t) and fouryear(t)

Table A2: Descriptive Statistics

Variable	Mean	Standard Deviation	Min	Max
Baseline Variables				
hsgpa	2.71	0.68	0.14	4.00
afqt	47.7	26.8	1	99
pareduc	14.2	2.2	10	19
parba	0.29	0.45	0	1
lowincome	0.57	0.50	0	1
urban	0.61	0.49	0	1
regionne	0.16	0.37	0	1
regionnc	0.32	0.47	0	1
regionso	0.32	0.47	0	1
regionwe	0.20	0.40	0	1
white	0.74	0.44	0	1
black	0.09	0.28	0	1
latino				
distance2	15.4	20.4	0	162
distance4	23.6	26.8	0	234
tuition2	1496	868	280	3476
tuition4	2309	771	1251	4265
Educational Outcomes				
Enroll year 1	0.55	0.50	0	1
year 2	0.53	0.50	0	1
year 3	0.46	0.50	0	1
year 4	0.42	0.49	0	1
twoyear1	0.15	0.36	0	1
fouryear1	0.40	0.49	0	1
gpa year 1	2.42	0.86	0.00	4.00
year 2	2.48	0.89	0.00	4.00
year 3	2.63	0.87	0.00	4.00
year 4	2.75	0.84	0.00	4.00
contenroll	13.91	2.13	12	19
contenroll = 12	0.45	0.50	0	1
contenroll = 13	0.10	0.30	0	1
contenroll = 14	0.08	0.27	0	1
contenroll = 15	0.06	0.24	0	1
contenroll = 16+	0.31	0.46	0	1

Notes: All variables have 1773 observations, with the exception of GPA variables which are restricted to those enrolled in each year

One limitation of the NELS dataset is that respondents are relatively young (approximately 26 years old) at the time of the final survey year. Since income at this age is a poor indicator of ultimate lifetime income due to job instability and graduate school attendance, I instead estimate individuals' expectation of lifetime income using data from an earlier cohort. This procedure is described in the next section. I restrict the dataset to on-time high school graduates with complete information on key baseline variables (sex, high school GPA, parents' education, AFQT score, distance to nearest colleges) and complete college transcripts (unless no claim of college attendance). I also exclude residents of Alaska, Hawaii, and the District of Columbia. After these restrictions the final dataset contains 1773 men.

8.2 Estimating Conditional Income Expectations

Expectations of lifetime income under different schooling outcomes is a key factor in educational choices. One limitation of the NELS dataset is that respondents are relatively young (approximately 26 years old) at the time of the final survey year. Since income at this age may be a poor indicator of ultimate lifetime income, I do not estimate expectations using individual's actual labor market outcomes. Instead I estimate individuals' expectation of lifetime income using data from a cohort about 12 years earlier, the National Longitudinal Survey of Youth 1979 (NLSY79). This approach is similar to the "reference group expectations" referred to by Manski (1991).

The NLSY79 is a Department of Labor longitudinal survey of 12,686 men and women who were 14-22 years old in 1979. They have been surveyed annually or biennially since. Using variables that are common in both the NLSY79 and NELS (such as high school GPA, parental education, AFQT, ethnicity, urban and region), I first estimate the parameters of a lifetime income equation on the NLSY79 data. Equation A1 below is estimated using OLS and used to predict counterfactual lifetime income for individuals in the NELS sample.

$$\begin{aligned} Income_i(S_i) = {} & \omega_0 + \omega_{13}1(S_i = 13) + \omega_{14}1(S_i = 14) + \omega_{15}1(S_i = 15) + \omega_{16}1(S_i \geq 16) \\ & + \omega_b Black_i + \omega_l Latino_i + \omega_c Central_i + \omega_s South_i + \omega_w West_i + \omega_u Urban_i \\ & + \omega_g HSgpa_i + \omega_t AFQT_i + \omega_p ParentEd_i \\ & + \omega_{gt} HSgpa_i * AFQT_i + \omega_{gp} HSgpa_i * ParentEd_i + \omega_{tp} AFQT_i * ParentEd_i + \epsilon_i^{\omega} \end{aligned} \tag{A1}$$

Where S_i is years of continuous enrollment in college after high school graduation and $Income_i(S_i)$ is the present discounted value of lifetime income from age $[18 + (S_i - 12)]$ to 62. NLSY79 individuals are ages 39 to 47 in 2004, the most recent year for which data is available, so I assume that earnings are constant from age 39 to 62. The base specification permits the intercept of lifetime income to vary with observable background and ability variables, but restricts the lifetime income returns to each year of college to be constant within sex. An alternative specification allows the return to some college (S = 13, 14, or 15) and a BA (S $\geq$ 16) to vary with high school gpa, AFQT, and parent's education.

Table A3 provides estimates of the parameters of the lifetime income equation for both the base and heterogeneous-returns model for different assumed values of the discount rate. Table A4 presents the estimated lifetime income by sex and for different assumptions for the NELS sample.

Table A3: Parameter Estimates from Lifetime Income Equation

Dependent variable: PDV of lifetime income post-school

	Men in NLSY Sample			
	d = 5%	d =10%	d = 5%	d =10%
	(1)	(2)	(3)	(4)
contenroll = 13	34.64	19.36	95.61	56.34
	(22.09)	(10.63)	(75.61)	(36.57)
contenroll = 14	55.54	32.10	126.74	76.04
	(24.77)	(11.35)	(82.67)	(39.35)
contenroll = 15	165.89	88.99	238.77	133.56
	(37.88)	(17.43)	(85.66)	(40.95)
contenroll > 15	328.13	183.50	-82.26	-6.75
	(30.08)	(14.17)	(154.00)	(72.66)
ParentEd	7.65	4.63	13.15	6.96
	(8.58)	(4.02)	(8.22)	(3.85)
Black	-81.17	-44.76	-80.82	-44.71
	(19.14)	(8.97)	(19.06)	(8.95)
Latino	5.93	2.64	1.23	0.71
	(23.21)	(11.01)	(22.81)	(10.84)
NorthCentral	-45.96	-24.49	-42.48	-22.52
	(24.85)	(11.74)	(24.89)	(11.75)
South	-56.99	-29.31	-54.99	-28.10
	(24.54)	(11.61)	(24.52)	(11.60)
West	-56.01	-29.65	-52.13	-27.85
	(25.08)	(11.78)	(25.12)	(11.76)
Urban	32.60	13.06	31.70	12.65
	(15.71)	(7.51)	(15.57)	(7.43)
HSgpa	42.78	24.61	74.22	39.36
	(41.40)	(19.49)	(40.77)	(19.42)
AFQT	1.52	0.99	3.03	1.64
	(1.35)	(0.64)	(1.34)	(0.64)
HSgpa*AFQT	0.25	0.04	-0.19	-0.14
	(0.43)	(0.20)	(0.42)	(0.19)
HSgpa*ParentEd	-0.55	-0.36	-2.03	-1.14
	(3.95)	(1.84)	(3.81)	(1.80)
AFQT*ParentEd	-0.03	-0.03	-0.08	-0.05
	(0.08)	(0.04)	(0.09)	(0.04)
(s13-s15)*AFQT			-0.16	-0.27
			(0.67)	(0.32)
s16*AFQT			1.91	0.65
			(1.33)	(0.64)
(s13-s15)*HSgpa			0.47	-3.39
			(28.99)	(13.57)
s16*HSgpa			46.76	29.88
			(47.09)	(22.43)
(s13-s15)*ParentEd			-4.01	-1.15
			(4.91)	(2.40)
s16*ParentEd			9.97	3.94
			(9.85)	(4.52)
Constant	223.52	112.62	134.05	72.50
	(93.60)	(44.27)	(91.50)	(43.26)
Observations	1,982	1,982	1,982	1,982
R-squared	0.30	0.33	0.30	0.34

Robust standard errors in parentheses

Table A4: Predicted Lifetime Income and Incremental Returns by Years of Continuous Enrollment

Men in NELS Sample

	d		Predicted Present Value of Lifetime Income (,000)					Predicted Incremental Income Increase (,000)			
			12	13	14	15	16	13	14	15	16
(1)	5%	mean	484	519	540	650	813	35	21	110	162
		stdev	90	90	90	90	90	0	0	0	0
(2)	10%	mean	245	265	277	334	429	19	13	57	95
		stdev	39	39	39	39	39	0	0	0	0
(3)	5%	mean	476	508	539	651	754	32	31	112	103
		stdev	72	66	66	66	152	11	0	0	94
(4)	10%	mean	242	259	279	337	403	18	20	58	67
		stdev	33	26	26	26	68	10	0	0	47

Notes: Parameters were estimated using the NLSY.

8.3 Policy Implications of Option Value

This section explores several policy implications of the presence of completion uncertainty and option value using the simple theoretical model developed in Section 2. First, I demonstrate how the predicted effects of different tuition subsidies depend on the inclusion of option value. When option value is considered, non-budget-neutral subsidies for the first half of college are predicted to increase schooling levels and welfare more than across-the-board or back-loaded subsidies. Moreover, front-loaded subsidies can be welfare improving and budget neutral if credit constraints or incomplete information cause some students to ignore the continuation value from enrollment. The magnitude of the welfare gains from such a policy depends critically on the extent to which enrollment provides new information. Second, I show how option value influences the effect of community colleges expansion on schooling outcomes. Simulations suggest that the static model which ignores option value can provide a misleading guide to the likely effects of community colleges.

Tuition Subsidies and Optimal Tuition Timing

First suppose the government wished to subsidize education due to wage spillovers (Moretti 2004) or civic externalities (Milligan, Moretti, Oreopoulos 2004 and Dee 2004) and did not require budget neutrality.[23] They may consider a subsidy to the first half of college, a similar subsidy to the second half, or equal subsidies to the first and second half. I simulate the schooling and welfare outcomes under these three different policies and report the results in Figure A1.[24] In the static model (left),

[23] The welfare calculations that follow reflect only the distribution of benefits of the subsidy. I do not model the incidence of taxes to support these subsidies. These calculations would be appropriate for a government that was committed to a tuition subsidy (regardless of the costs) and wanted to know the distribution of benefits.

[24] In Figures A1 and A2 I take 10,000 random draws of $\varepsilon_{i,2}$ from a normal distribution with mean zero and variance σ^2 for each value of $\varepsilon_{i,1}$. The figures report the average welfare and schooling outcomes across these 10,000 draws. The simulations in Figure A1 assume a subsidy of 0.50 either provided entirely in period one or period two, or split equally

the back-loaded and across-the-board subsidies have the same effect: individuals with $\varepsilon_{i,1} \in (0.25, 0)$ will now enroll and complete college. In contrast, front-loaded subsidies will cause individuals with $\varepsilon_{i,1} \in (0.50, 0)$ to enroll for the first half, but will not impact completion. The welfare consequences of the three schemes are similar, but the front-loaded scheme provides additional benefits to those who would enroll in only one year. The static model predicts that the three schemes have identical welfare impact on anyone who would enroll without the subsidy ($\varepsilon_{i,1} > 0$).

The dynamic model (right) has qualitatively different predictions about the various tuition subsidies. Most notably, the dynamic model predicts that all three schemes increase the average schooling of affected groups by more than one but less than two years. In contrast to the static model, here the front-loaded subsidy will increase college completion. After enrolling, some people learn that completion is desirable. In the static model, no such learning takes place and thus front-loaded subsidies will not affect completion. Furthermore, the distribution of benefits now depends on the subsidy timing. Front-loaded subsidies now have greater benefits than the other schemes to all individuals, even those that would enroll without the subsidy.

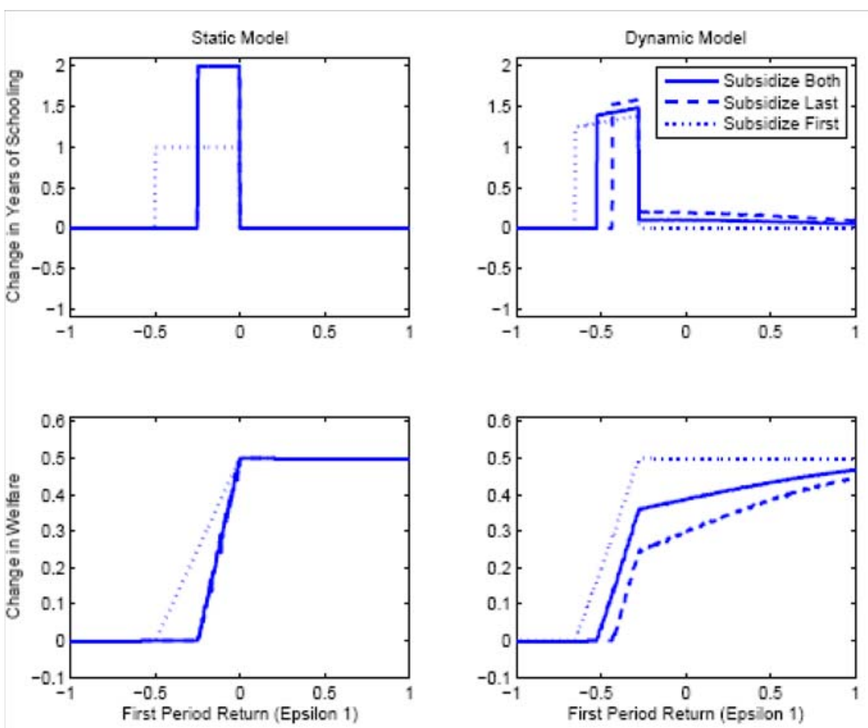


Figure A1

Now suppose the government wished to set tuition optimally, but remain budget neutral so that

between periods one and two.

any subsidies in one period must be offset by higher tuition fees in the other.[25] Also suppose the social costs of schooling are c_1 and c_2 for the first and second periods, respectively, and do not vary across individuals. The social planner sets tuition levels in each period (T_1, T_2) to maximize aggregate welfare.[26] It can be shown that if individuals consider the option value associated with college enrollment, then tuition levels should be set at their marginal social cost: $T_1^* = c_1$ and $T_2^* = c_2$.

In this case, welfare can not be increased by altering tuition prices from their social cost levels. The intuition for this result is that while reducing T_1 causes the marginal students to enroll, the welfare gain associated with this change is zero since marginal students are those for whom enrollment is not valuable. The only direct welfare gain from a reduction in T_1 is that some people will now pay lower tuition in the first period. The balanced-budget constraint necessarily means that this benefit is exactly offset by an increase in second period tuition, which generates an equally-sized welfare loss.

Now consider a situation where individuals do not consider continuation value when making enrollment decisions. Instead, individuals are concerned only about current-period payoffs, so educational outcomes result from a series of static decisions. This behavior could result from credit constraints or extremely high discount rates. Individuals enroll if $\varepsilon_{i,1} - T_1 > 0$. Individuals for whom enrollment cannot be justified by first-period payoffs alone will not enroll. Again the social planner sets (T_1, T_2) to maximize aggregate welfare subject to the budget constraint that all schooling costs must be covered by tuition.[27] It can be shown that now the optimal tuition scheme must satisfy:

$$\begin{aligned}
T_1^* \quad &: \quad T_1 = c_1 - \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{(P_2)^2}{\delta P_2/\delta T_2}\right) \\
T_2^* \quad &: \quad T_2 = c_2 + \left(\frac{1-\lambda}{\lambda}\right)\left(\frac{P_2}{\delta P_2/\delta T_2}\right) \\
\lambda \quad &= \quad 1 + \frac{f_1(T_1)}{P_1} \cdot E[\max\{0, T_1 + \varepsilon_{i,2} - T_2)\}]
\end{aligned}$$

where $f_1(\cdot)$ is density of $\varepsilon_{i,1}$. The expression $E[\max\{0, T_1 + \varepsilon_{i,2} - T_2)\}]$ represents the social gain from compelling the marginal individual (defined by $\varepsilon_{i,1} = T_1$) to enroll. This gain is scaled by the

[25] The optimal tuition results summarized in this section are derived in a separate paper. See Stange (2007), "Optimal College Tuition When Completion is Uncertain," working paper.

[26] Specifically, the social planner maximizes $E[0, \varepsilon_{i,1} - T_1 + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2} - T_2\}]]$ subject to the budget constraint that all schooling costs must be covered by tuition: $T_1P_1 + T_2P_1P_2 \geq c_1P_1 + c_2P_1P_2$, where P_1 and P_2 are the overall enrollment and completion rates, respectively. The outer expectation is taken over $\varepsilon_{i,1}$ while the inner expectation is taken over $\varepsilon_{i,2}$.

[27] Now social welfare equals $E[1(\varepsilon_{i,1} - T_1 > 0) \cdot (\varepsilon_{i,1} - T_1 + E[\max\{0, \varepsilon_{i,1} + \varepsilon_{i,2} - T_2\}])]$ where the expectation is taken over the distribution of $\varepsilon_{i,1}$ in the population.

density of individuals at the margin, given by $f_1(T_1)$. Since $f_1(T_1)$, $E\left[\max(0,\cdot)\right]$, and P_1 are all positive at $(T_1, T_2) = (c_1, c_2)$ and $\delta P_2/\delta T_2 < 0$, then $\lambda > 1$, $T_1^* < c_1$, and $T_2^* > c_2$. A budget-neutral shift in tuition from period one to period two will improve social welfare. In the case where $c_1 = c_2$, then the above conditions reduce to $T_1^* = c_1$ and $T_2^* = c_2$ as the variance of $\varepsilon_{i,2}$ goes to zero.

This result provides theoretical justification for a wide range of policies, including community colleges and tax credits, which subsidize the first half of college in order to compel more people to enroll. Community colleges, when combined with transfer to a four-year university, create a tuition schedule which is explicitly back-loaded. This may be a preferred option (welfare-enhancing) for individuals who are not completely forward-looking. The optimal tuition schedule was found to depend on the extent to which college completion is uncertain (variance of $\varepsilon_{i,2}$), the density of people at the enrollment margin, and the responsiveness of completion to second-period tuition ($\delta P_2/\delta T_2$), in addition to the more obvious factors such as the marginal social costs (c_1, c_2). These quantities have received little attention in empirical work.

Expansion of Community Colleges

Community colleges are one of the key mechanisms states use to expand postsecondary access. The past few decades have witnessed a considerable expansion of community colleges: community colleges absorbed nearly half of the increase in college enrollment between 1980 and 1994, accounting for 38% of all postsecondary enrollments by 1995 (Kane and Rouse, 1999). The net effect of community colleges on educational attainment is theoretically ambiguous due to offsetting democratization and diversion effects. On the one hand, the accessibility of community colleges provides opportunities to students that otherwise would not attend college, expanding educational attainment. On the other hand, community colleges may decrease attainment because they do not adequately facilitate the transfer of high-ability students, diverting them from obtaining a Bachelor's degree.

Since community colleges alter the timing of schooling costs and graduation probabilities, the effect of their introduction will depend on the level of schooling uncertainty and option value. To examine this, I simulate schooling outcomes when individuals can choose between two different schooling pathways. The first is just the baseline dynamic model described above with no subsidies. In the second pathway, individuals receive a subsidy in the first period but must pay a cost twice as large in the second period if they continue in school. This pathway is an approximation of community colleges: individuals pay low schooling costs while at community college but transferring to a four-year university in period two is difficult. Consequently, it is more costly to graduate having started at community college (due to costly transfer) but dropout is also less costly.

The top four panels of Figure A2 reports the enrollment and graduation rates predicted by the static and dynamic models before (solid line) and after (dashed lines) introducing community colleges. The effects can be distinguished for four distinct groups. Individuals in Group A choose not to enroll regardless of the availability of community college, while individuals in Group D will always attend four-year college directly. If $\varepsilon_{i,1}$ is high enough, starting at community college is undesirable since it increases the total cost of graduating. Individuals in these two groups are unaffected by the introduction of community colleges in both the static and dynamic settings.

Individuals in Groups B and C choose to attend community college when it is available. Group B contains new enrollees for whom community college is pivotal to the enrollment decision (the "democratization" effect). Group C contains individuals who would attend four-year college if community college were not available (the "diversion" effect). The static and dynamic models make different predictions regarding the response of these groups to the introduction of community college. In the static model, individuals in Group B now enroll for the first period, but do not graduate. In contrast, the dynamic model predicts that some individuals in Group B will graduate, after learning that graduation actually is desirable. In the static model, individuals in Group C no longer graduate once community colleges are introduced. They find it optimal to attend community college for one period and then drop out. The dynamic model predicts that the graduation rate of Group C is reduced by community college, but not nearly so sharply. These effects are summarized in the bottom two panels, which plot the change in average years of schooling. The magnitude of the diversion and democratization effects, reflected in the sizes of the various groups and their predicted outcomes, depend on whether uncertainty and option value is incorporated.

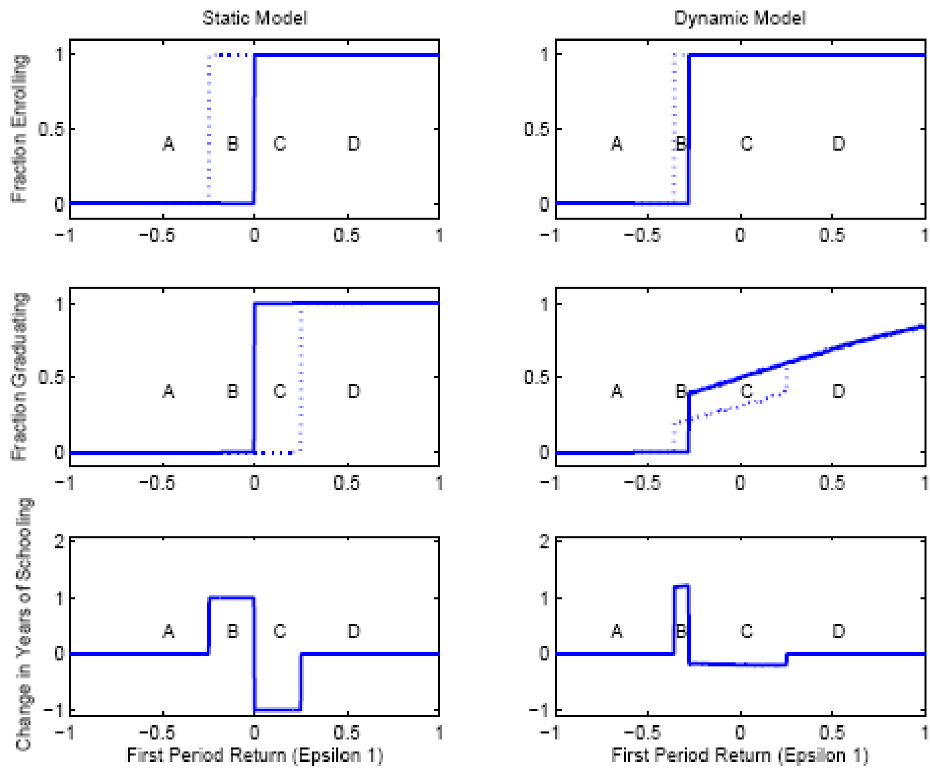

Static Model
Dynamic Model
Fraction Enrolling
Fraction Graduating
Change in Years of Schooling
A
B
C
D
First Period Return (Epsilon 1)
First Period Return (Epsilon 1)

Figure A2