Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers*

Joseph Altonji, Erica Blom, and Costas Meghir

Yale University

September 22, 2011

Abstract

Motivated by the large differences in labor market outcomes across college majors, we survey the literature on the demand for and return to high school and post-secondary education by field of study. Drawing on several papers, we provide a dynamic model of education and occupation choice that stresses the roles of specificity of human capital and uncertainty about preferences, ability, education outcomes, and labor market returns. The model implies an important distinction between the ex ante and ex post returns to education decisions. We also discuss some of the econometric difficulties in estimating the causal effects of field of study on wages in the context of a sequential choice model with learning. Finally, we review the empirical literature on choice of curriculum and the effects of high school courses and college major on labor market outcomes.

*We thank Sarah Amanda Levis for excellent research assistance and Seth Zimmerman for helpful comments.

Contents

1 Introduction **4**

2 Lifecycle decision making: a theoretical model and some empirical implications **5**

2.1 Preliminaries and notation . . . 6

2.1.1 The state space . . . 6

2.1.2 Updating beliefs about ability and preferences . . . 6

2.1.3 The human capital production function . . . 7

2.2 Labor market . . . 7

2.2.1 Work and occupation choices . . . 8

2.3 Education choices . . . 9

2.3.1 Utility . . . 9

2.3.2 College education . . . 10

2.3.3 High school education . . . 11

2.4 Parental influence and transfers . . . 12

2.5 Implications of the model . . . 13

2.5.1 Theoretical implications . . . 13

2.5.2 The returns to education paths . . . 14

2.5.3 Econometric implications . . . 15

3 Empirical evidence on choice of high school curriculum **16**

3.1 School-level influences . . . 18

4 The effects of high school curriculum on educational attainment and wages **18**

4.1 Conclusion about the effects of high school curriculum . . . 23

5 Empirical evidence on the determinants of college major **24**

5.1 Trends in college major . . . 24

5.2 Expected earnings . . . 25

5.3 Preferences . . . 27

5.4 Ability and preparation . . . 27

6 Returns to college major: empirical evidence **28**

6.1 Estimation strategies in the existing literature . . . 28

6.2 Possible alternative strategies . . . 29

6.3 Estimates of the return to major . . . 30

6.4 Descriptive evidence on the returns to college major from the ACS 32
6.5 College major and occupational choice 33

7 Conclusion **34**

1 Introduction

A huge literature studies educational attainment and "the return to education," with the more recent work focussing on heterogeneity in the return.[1] There is much less work on why indviduals choose different *types* of education or on the career consequences of those choices. In some ways this is surprising, because the question of what kind of education should be provided receives an enormous amount of attention in policy discussion and is a decision that every student must make, particularly at the post-secondary level. Periodically, there are calls for reform of the high school curriculum, often in response to concern about the readiness of students for the work force. An important example is *A Nation at Risk* (Gardner (1983)). As we discuss below, the report advocated a focus on the "new basics," with an increase in the number of courses in key academic subjects such as math and science. But despite the perennial debate about what students should study, there is surprisingly little hard evidence about the labor market consequences of specific high school courses.[2]

At the college level, differences by field of study have received much less attention than the average return to post-secondary education. This is true even though the difference in returns across college majors rivals the gap between the average college graduate and the average high school graduate. For example, using the data from the 2009 American Community Survey (ACS), we find that after adjusting for basic demographics, potential labor market experience, and graduate education, the gap in log wages rates between men who majored in electrical engineering and men who majored in general education is 0.561 (0.016). This is nearly as large as the 0.577 (0.003) difference between college graduates and high school graduates. Furthermore, we find that the standard deviation of the return to various majors is 0.177 for men–about double the typical estimate of the value of the year of school. The extent to which these differences reflect unobserved differences in high school preparation and worker ability or represent compensating differentials for nonpecuniary aspects of jobs is not well-established. However, understanding these differences is important for understanding the role of education in economic success and even for providing advice to students who are deciding on a course of study.

In keeping with the treatment of schooling as homogeneous, most of the literature in labor economics on career earnings focuses on the return to general experience, job tenure, and job mobility. However, there is an important and growing literature concerned with the implications of occupation-specific human capital for wages and mobility. Field of education conditions occupational paths.

In this paper, we provide a survey of the theoretical and empirical literature on the heterogeneous nature of education and its link to particular occupation paths. On the theoretical side, rather than providing a short summary of each of the existing papers, we combine elements from a few of them to provide a theoretical model of education choice, occupation choice, and wages. The model has five key features. First, preferences, innate ability, and the initial vector of skills and knowledge early in high school shape the feasibility and the desirability of particular education

[1] See Card (1999) and Meghir and Rivkin (2011) for surveys.
[2] Due to space constraints, we do not discuss evidence on the effects of curriculum on test scores.

programs. Second, individuals only learn gradually about their preferences and ability and are also uncertain about wages. Third, the type of education program shapes what one learns during a schooling period. Fourth, education programs and occupations have different skill and knowledge prerequisites that influence learning and job performance. Switching fields in the face of new information about ability, preferences, and returns is costly. Finally, knowledge accumulation is stochastic—students cannot simply decide that they are going to complete a program of study. We draw out a number of implications of the model that have been discussed in the literature, including an important distinction between ex ante and ex post return to an education choice. We also use the model to discuss some of the econometric difficulties that sequential choice in the presence of learning poses for estimating the causal effect of field of study on wages.

We then turn to the empirical literature on the demand for and return to particular types of education, beginning with high school curriculum and then turning to college major. We place particular emphasis on the determinants of college major and the effects of high school courses and college major on wage rates. We supplement the existing literature with an analysis of the returns to college major using the ACS, which is large enough for us to consider narrowly defined majors and 5-digit occupation categories. Due to space constraints, we focus primarily on the literature for the US.

We conclude with some theoretical and econometric implications of the model as well as suggestions for future research.

2 Lifecycle decision making: a theoretical model and some empirical implications

The aim of the following section is to describe a theoretical model that brings together the educational choices with the labor market careers that follow them. This sets the scene for interpreting the review of empirical results that follows. From an econometric point of view this unified framework can help identify the assumptions required for empirical analysis. As a setup, this model is not new and brings together elements of Altonji (1993) and Arcidiacono (2004), who emphasize the dynamic decisions about type of education, as well as Keane and Wolpin (1997), Eckstein and Wolpin (1989) and Gallipoli, Meghir and Violante (2005), who also include risk considerations and examine the role of parents. We conclude this section with a discussion of the implications of the model and the important empirical challenges that this field faces. Almost all of the empirical literature we describe is either based on OLS regression or on instrumental variables strategies and focuses on specific aspects of educational choice. This work is important, but the model implies that when agents are making a sequence of choices among multiple education options, the use of "single equation" methods to estimate ex ante and ex post returns to education choices requires very strong assumptions. These methods also lead to incomplete measures of such returns, since costs are ignored. Readers who are primarily interested in the empirical work may wish to skip to Section 3 on a first pass through the paper.

We now describe the sequence of decision phases for the individual. We start from the end point where individuals make labor market decisions; we then go back through the education choices and end up at the beginning, when parents can make transfers and influence their children's decisions.

2.1 Preliminaries and notation

2.1.1 The state space

We define the state space, i.e. the set of relevant variables describing the history of events and choices made by the individual up until the current point, by Ω. This includes past grades and educational attainment, and possibly past choices to the extent they have persistent effects on behavior either directly or through expectations. Ω_t also includes assets. Once in the labor market, Ω will track labor market history, including wages and the occupational choices that help determine the vector of skills K, which drives achievement in school and wages. We use the subscript ς to denote educational stages, i.e. high school, first part of college and second part of college. When the individual enters the labor market, we switch to a subscript t which denotes years in the labor market. In general we denote labor market and education choices made in period τ by x_τ, and other information, including grades, wages, assets, etc., by g_τ. The state space is updated by $\Omega_\tau = \{\Omega_{\tau-1}, x_{\tau-1}, g_{\tau-1}\}$, where $\tau = \varsigma$ during education and $\tau = t$ during the labor market period.[3] The initial information Ω_I will contain a vector of background variables, including parental characteristics, grades, and childhood developmental characteristics. In practice, much of the state space will be unobserved by the researcher. When implementing such a model empirically, the relevant information set would have to be suitably restricted so as to make the computational problem feasible. However, here we ignore such (important) specification issues.

2.1.2 Updating beliefs about ability and preferences

It is useful to think of individuals as possessing ability a (possibly multidimensional) and preferences θ. Both are unknown to the individual and are a source of uncertainty. The individual learns about ability and preferences through environment and through experience, such as trying a set of courses or switching occupations in the labor market. This is helpful in explaining switches between fields of study and between occupations, as individuals learn what they can and cannot do well and what they like. We start with prior beliefs summarized by the distribution $F(a, \theta)$. These are updated according to Bayes' rule, i.e.:

$$F^\tau(a, \theta|\Omega_\tau) =\propto G(\Omega_\tau|a, \theta)F^{\tau-1}(a, \theta|\Omega_{\tau-1})$$

[3]We make the notational distinction between the labor market periods (t) and the educational periods to emphasize that they last a different amount of time. In the labor market we will think of periods as one year (or perhaps a quarter). In education we will have a high school period, a first period of college, and a second period of college, each lasting different amounts of time. It also helps distinguish the nature of the state space or choice variables.

where τ denotes stage (such as high school, college, or labor market period). $G(\cdot)$ is the likelihood function of the information set Ω_τ. At the initial stage, which is the end of statutory schooling, $\tau - 1$ is taken to be I, the information set is Ω_I, and $F^I(a, \theta|\Omega_I) = F(a, \theta)$, i.e. the prior distribution.

In what follows, whenever ability and preferences need to be integrated out, it is done so based on the updated distribution.

2.1.3 The human capital production function

Productivity in the labor market and hence wages will be driven by a vector of human capital indices (skills) $K_{it} = \{K_{it}^r\}$, $r = 1, ..., R$, where t measures time from entry in the labor market and r indexes skills.

Human capital evolves as

$$K_{it} = f(K_{it-1}, D_{it}, a_i, \epsilon_{it}|\Omega_s) \tag{1}$$

where Ω_s contains the completed educational history of the individual, including choices of subjects, fields of study and majors, as well as grades. This emphasizes that accumulated skills may depend on the particular type of education followed. Skills are produced by the set of past skills and by occupational experience reflected in $D_{it} = \{D_{it}^l\}$, an $L \times 1$ vector of experiences in the L possible occupations, including unemployment. Note that skills are not equivalent to occupations. However, the dependence on D_{it} reflects the fact that what one learns depends on the occupation. Before entry in the labor market, D_{it} is empty and skills are produced through education; field of study influences what one learns. ϵ_{it} is an idiosyncratic error term that reflects the fact that knowledge accumulation is stochastic; it may depend, for example, on unforeseen illness or the quality of instructors. f is increasing in K_{it} and ability, allowing possibly for self-productivity and ability-skill complementarity. Finally, the dependence separately on educational history reflects the fact that the level and type of education may lead to a permanently different type of human capital which is not perfectly substitutable with other types and which commands its own market price. For example, medical training allows one to become a surgeon, a skill not obviously substitutable with skills of individuals of a different educational background.

2.2 Labor market

The labor market is competitive and includes L different occupations, each paying an occupation-specific wage α_{lt} per efficiency unit of human capital. Human capital required for the sector is a combination of the R skills K_{it}. It is defined by the occupational production function

$$Q^l = Q^l(K, a) \tag{2}$$

Thus the set of skills an individual i accumulates defines the efficiency units she can supply in each of the sectors l. Assuming symmetric information between workers and employers, human capital is priced based on expectations of productivity from the distribution of belief about ability $F^t(a,\theta|\Omega_t)$. The resulting individual wage is

$$\ln w_{ilt} = \alpha_{lt} + lnE_{F^t}(Q^l_{it}) + u_{ilt} \tag{3}$$

where u_{ilt} is an idiosyncratic shock to efficiency units for the lth occupation. In practice, one would have to find ways to restrict and refine the information set. However, this implies that factors affecting learning from individual experience will determine wages despite the fact that they have no influence on productivity itself. For example, math grades in high school may affect lawyers' wages. (See Farber and Gibbons (1996), Altonji and Pierret (2001) and Lange (2007) for work on employer learning about productivity with only general education.) Wage equations may differ for men and women, although this has not been made explicit in the above equation.

One traditional way of parametrizing the above wage equation is as follows:

$$\ln w_{islt} = \alpha_{0lt} + \alpha_{1sl}\hat{a}_{it} + \alpha_{2sl}X_{it} + \alpha_{3sl}X^2_{it} + \alpha_{4sl}T^l_{it} + v_{ilt}$$

where T^l_{it} is person i's current tenure in occupation l and X_{it} is total labor market experience. We use the loose notation $\hat{a}_{it}$ to denote the expectation about ability at t. The subscript s indicates that the coefficients of the wage equation depend on education history including the sequence of field choices and highest degree. Finally, v_{islt} is an idiosyncratic stochastic shock. This equation is an approximation obtained by substituting out for the vector K using (1), (2), and (3). Tenure in other occupations would also belong in the equation to the extent that productivity in occupation l depends on skills learned in other occupations.

2.2.1 Work and occupation choices

We now formally describe the lifecycle decision process relating to education and the labor market. It is presentationally and analytically more convenient to start from the end and move backwards, the order in which one would actually solve such a model.

Following the completion of education the individual enters the labor market and chooses between a number of different occupations, including inactivity or home production. The labor market choices are repeated in each period and the value functions are indexed by t, which denotes length of time in the labor market and emphasizes that this is a non-stationary lifecycle problem. The terminal condition is left unspecified here.

An individual possesses a utility function, which expresses her preference over consumption and occupations. This is denoted by $u^l(c_t|\xi^l_t, a, \theta)$, where c_t is consumption and the index l denotes occupation. The preference for a particular occupation is determined by a set of parameters θ, by ability and by a random shock ξ^l (with $l = B$ for unemployment). As described above, the vector θ may not be known. Beliefs about it depend on past experiences, and indeed the updating of beliefs about preferences or ability may be an important source of switching occupations.[4]

[4]See Papageorgiou (2010) for a concise overview of the literature on learning and occupational mobility.

Now define the value of working in occupation l as $V_t^l(\Omega_t, \xi_t^l)$, $l = 1, ..., L-1$ and the value of unemployment as $V_t^B(\Omega_t, \xi_t^B)$. We have that

$$V_t^l(\Omega_t, \xi_t^l) = \max_{c, A_{t+1}} \left\{ E_{a,\theta} u^l(c_t | \xi_t^l, a, \theta) + \beta E_t [V_{t+1}^{\mathcal{L}}(\Omega_{t+1}^l, \xi_{t+1})] \right\}$$

where A_{t+1} are assets in period $t+1$, included in Ω_{t+1}^l (Ω_t includes A_t respectively). The superscript l on Ω_{t+1}^l indicates that each occupational choice will lead to a different point in the state space because of different labor market earnings (w_{ilt}), different savings behavior (A_{t+1}) and most pertinently different labor market experiences. Assets evolve based on the usual difference equation

$$A_{t+1} = (1+r)(A_t - c_t + w_{ilt})$$

The expectations operator E_t is taken with respect to information in period t. The value of unemployment is similarly defined, except that w_{ilt} is replaced by unemployment benefits.

The optimized value of the labor market in period t is then given by

$$V_t^{\mathcal{L}}(\Omega_t, \xi_t) = \max_l \left\{ \left[V_t^l(\Omega_t, \xi_t^l), l = 1, ..., L-1 \right], V_t^B(\Omega_t, \xi_t^B) \right\}$$

This value depends on the entire set of shocks ξ^l, summarized in the vector ξ as well as on other factors unknown in earlier periods, including shocks to wages. Thus from the perspective of earlier periods, this is a source of uncertainty, which will matter given risk aversion and incomplete insurance markets.

Important factors in practice are frictions and occupational shocks, such as exogenous job destruction. From a substantive point of view these can be important particularly if the extent of frictions differ from sector to sector. Given that individuals are risk averse, different risk characteristics may affect occupational and by implication educational choice. As we have set it up here there is still occupational risk, which is reflected in different volatility of wages.

2.3 Education choices

2.3.1 Utility

Let j denote the field of study followed by the individual and $\varsigma = H, C_1, C_2$ denote the level (high school, first period of college, second period of college). The flow utility of schooling choice (j, ς) depends on effort exerted, on a stochastic preference shock denoted ζ_ς^j and on consumption. It also depends on the individual's ability a and on preferences that are known up to some person specific vector of parameters θ. Ability and preferences influence net enjoyment of pursuing (j, ς), including the cost of effort. We denote this utility by $u_\varsigma^j(c_\varsigma | \zeta_\varsigma^j, a, \theta)$, which does not differ notationally from the utility conditional on occupation. We often leave the school level subscript on u and c implicit. Since ability and preferences are not known, they need to be integrated out with respect to the current distribution of beliefs and preferences $F^\varsigma(a, \theta | \Omega_{j\varsigma})$ defined earlier.

2.3.2 College education

We split college into two decision periods: C_1 represents the first period of college, wherein a student can choose between a number of general courses; C_2 is a period of specialization wherein the student chooses major. Ω_{C_1} is the resulting information and knowledge set at the end of the first period of college, including the mix of all past courses. Grade progression is contingent on meeting certain threshold requirements, i.e. $(g_{C_1} > C_1^*)$, in order to continue in major j.[5]

In the second period of college, a major is chosen out of J_{C_2} possible options, subject to having qualified. The value of a specific major j is given by

$$V_{C_2}^j(\Omega_{C_1}, \zeta_{C_2}^j) =$$

$$max_{c,A_{C_2}} \left\{ E_{a,\theta} u^j(c|\zeta_{C_2}^j, a, \theta) + \beta^{C_2} E_{C_2}[V_{t=0}^{\mathcal{L}}(\Omega_{t=0}^j, \xi_{t=0})] \right\}$$

where ζ^j are utility shocks for major j and ξ is the relevant set of shocks for the labor market that will follow completion of the major. The superscript on the discount factor here and below reflects the fact that the education periods last longer than a year. Following major choice j and the second period of college, the updated information set is $\Omega_{t=0}^j$ where, like before, the superscript j denotes that the updating follows choice j. $\Omega_{t=0}^j$ reflects whether the student met graduation requirements for a degree in j. A_{C_1} are assets at the point of decision and A_{C_2} are assets at the end of the second period in college and are contained within the set $\Omega_{t=0}^j$. They are related by

$$A_{C_2} = (1 + r^{C_1})(A_{C_1} - c - F_{C_2}^j)$$

where $F_{C_2}^j$ represents costs of education for major j. Note that differential costs of education by subject matter would be very useful empirically for the purposes of identification. The function $E[V_{t=0}(\Omega_{t=0}^j, \xi)]$ is the expected value in the labor market (which starts at $t = 0$) following major choice j, as defined above. The dependence on time t reflects the non-stationary nature of the life-cycle problem.

Continuing with a major requires one to have satisfied the grade requirement $g_{C_1} > C_1^*$. Thus the value when the student has to choose one of J_{C_2} majors is

$$V_{C_2}(\Omega_{C_1}, \zeta_{C_2}, \xi_{C_2}, g_{C_1}) =$$

$$max_{j,\mathcal{L}} \left\{ \left[V_{C_2}^j(\Omega_{C_1}, \zeta_{C_2}^j) \times 1(g_{C_1} > C_1^*),\ j = 1, ..., J_{C_2} \right], V_{t=0}^{\mathcal{L}}(\Omega_{C_1}, \xi_{C_2}) \right\}$$

with g_{C_1} representing the grades obtained in the first period and C_1^* being the grade threshold for promotion.[6] Grades

[5] The thresholds could depend on major, but we avoid further complicationg the notation.

[6] The grades belong to the state space Ω_{C_1}, but we make the dependence of V_{C_2} on g_{C_1} explicit for clarity.

are a random variable whose distribution depends on individual characteristics and effort, which we leave implicit. The choice includes the possibility of starting off in the labor market with information set $\Omega_{t=0} = \Omega_{C_1}$, i.e. with the educational history up until that moment, but without completing a major. This is reflected in the value function $V^{\mathcal{L}}_{t=0}(\Omega_{C_1}, \xi_t)$.

Upon finishing high school the individual chooses between college (if she qualifies) and the labor market.. The information set at the beginning of that stage (C_1) is summarized by Ω_H. Choosing college involves choosing a particular curriculum (for example humanities versus social science versus natural sciences). This problem is presentationally similar to the one of choosing major. The value of each of the J_{C_1} curriculum options is

$$V^j_{C_1}(\Omega_H, \zeta^j_{C_1}) = max_{c,A_{C_1}} \left\{ E_{a,\theta} u^j(c|\zeta^j_{C_1}, a, \theta) + \beta^{C_1} E_{C_1}[V_{C_2}(\Omega^j_{C_1}, \zeta_{C_2}, \xi_{C_2}, g_{C_1})] \right\}$$

where $\Omega^j_{C_1}$ is the state space resulting from the jth choice. Assets A^{C_1} are given by

$$A_{C_1} = (1 + r^H)(A_H - c - F^j_{C_1})$$

where A_H are assets following high school completion and $F^j_{C_1}$ are the monetary costs of following curriculum j in the first period of college. The value at the start of the college choice period is

$$\begin{gathered} V_{C_1}(\Omega_H, \zeta_{C_1}, \xi_{C_1}, g_H) = \\ max_{j,\mathcal{L}} \left\{ \left[V^j_{C_1}(\Omega_H, \zeta^j_{C_1}) \times 1(g_H > H^*),\ j = 1, ..., J_{C_1} \right], V^{\mathcal{L}}_{t=0}(\Omega_H, \xi_{C_1}) \right\} \end{gathered}$$

where $g_H > H^*$ signifies that the high school grades are sufficient to qualify for college.

2.3.3 High school education

The earliest choice the individual has to make is to attend high school and follow a particular curriculum j, $j = 1, ..., J_H$. The value of choosing j is

$$V^j_H(\Omega_I, \zeta^j_H) = max_{c,A_H} \{E_{a,\theta} u^j(c|\zeta^j_H, a, \theta) + \beta^H E_H[V_{C_1}(\Omega^j_H, \xi_{C_1}, g_H)]\},$$

where Ω_I represents the initial information set, including parental background and earlier school achievement, and Ω^j_H represents the updated information set given the curriculum choice j.[7] High school is funded by parents. Parents also make promises of transfers that will depend on whether the child attends high school or not. These are included in Ω_I, while at the end of this stage assets (included in Ω_H) are given by

$$A_H = (1 + r)(A^U + w^H_0 - c)(1 - HS) + (1 + r^H)(T - c)HS$$

[7] The choice that a student faces may vary across schools. School quality would influence the function as well, but we leave this implicit.

where A^U are parental transfers if the child enters the labor market, while T are transfers if the child completes high school ($HS = 1$). Any costs of school are implicit in the amount transfered. The different interest rates reflect the differences in the amounts of time for the two activities, with work lasting one year (or quarter). The resulting choice faced by the individual is between one of the possible fields of study j and entering the labor market. She earns w_0^H if she chooses to work at that point. The value in anticipation of this choice is

$$V_H(\Omega_I, \zeta_H, \xi_H) = max_{j,\mathcal{L}} \left\{ \left[V_H^j(\Omega_I, \zeta_H^j), \ j = 1, ..., J_R \right], \ V_{t=0}^{\mathcal{L}}(\Omega_I, \xi_H) \right\}$$

The notation on the state space makes explicit that the initial position may differ depending on the parental transfers, to which we now turn.

2.4 Parental influence and transfers

The starting point of decision is at the end of compulsory schooling. We suppose that while the child makes her own decisions, they can be influenced by parents. Part of the influence is through genetic and cultural factors that are not the subject of conscious choices by the parent. Part is through decisions that influence the health, educational and broader social experiences of the child prior to high school.[8] These are all implicit in the original information set Ω_I, and although some may be the result of earlier choices, we do not have much to say about them here. However, one important source of influence is finance: parents choose how much to transfer to their children, either unconditionally or conditional on desired actions by the children (such as attending high school or college). This is a crucial source of funding for children who may not have other access to financial sources. From a policy perspective, understanding such transfers lies at the heart of understanding the extent to which outcomes can be influenced through government transfers.

Parents possess assets A^P. They need to choose how much to transfer to their children, given that they will fund high school and given that they may care about the child's welfare as well as whether the child completes high school. We denote by A^U transfers offered if the child drops out of school and by T transfers given if the child continues schooling. Given these, parents solve the problem

$$\max_{T, A^U} E_0[V^P(A^P - T \times HS - A^U \times (1 - HS), HS) + \kappa[V_H(\Omega_I, \zeta_H, \xi_H)]]$$

subject to

$$HS = 1 \left\{ max_j [V_H^j(\Omega_I, \zeta_H^j), \ j = 1, ..., J_R] > V_{t=0}^{\mathcal{L}}(\Omega_I, \xi_H) \right\}, \ A^U \geq 0, \ T \geq 0$$

where V^P is the parents' value function, HS is a binary indicator for high school attendance, and A^U and T are part of Ω_I affecting the child's education decision. Note that we assume that transfers cannot be negative; for example, the

[8] Surveys of this literature include Currie (2009), Todd and Wolpin (2003), Heckman and Masterov (2007), and many others.

parents cannot fine the children to induce them to attend high school. Finally, κ is the weight attached by parents to child utility. In a standard altruistic model, transfers will increase with A^P but will be lower for higher ability children, who can achieve higher value V^H. This is a more complicated model because of the parents' preference for their child to attend high school and their resulting willingness to use transfers to "distort" the child's choice towards education. The key point here is the link that this model creates between parental wealth and education in a world with liquidity constrants: wealthier parents will make more transfers enabling childen to study, whereas equal ability children from lower income backgrounds may be unable to do so if they cannot borrow. Conditional parental tranfers can easily be incorportated into the college stage of the model.

2.5 Implications of the model

2.5.1 Theoretical implications

The model has a number of implications for the demand for education, choice of field, and the return to education, some of which we have already pointed out. The model indicates that choice of high school courses and a college program should depend on the flow of utility from a set of courses. This will depend on ability, the prior stock of knowledge, and tastes for education. Family, peers, neighborhood, school, and background influence all of these, and thus influence choice of high school courses and choice of college program for those who attend college, as well as occupational choice. The menu of courses and programs offered by the student's high school and course requirements for grade advancement and graduation will affect curriculum choice. Choice will also depend on the effects of a given high school program on future education options, and on the labor market opportunities, current and expected. Consequently, variables that influence tastes for the quantity and type of post-secondary education and for particular occupations should influence high school course selection, as should variables that influence the student's ability to succeed in college and variables that influence wage rates in particular occupations. To the extent that uncertainty is important, interventions that provide individuals with better information about ability a and wage prospects will lead to better outcomes. As in standard models of education with no uncertainty and only one type of human capital, educational attainment depends negatively on the discount rate and net tuition costs.

The model implies an interesting interplay between preferences and the purely financial return to education.[9] First, preferences and innate ability affect the ex ante *financial* return to completing high school, starting college, and choosing a particular major *even if they play no role in the wage equation.* This is because they influence the likelihood that an an individual will ultimately choose to complete a program of study and receive the ex post payoff associated with it. Second, persons with preferences for fields with high labor market payoffs, such as engineering, have higher ex ante and ex post returns to high school completion and college attendance. This is true even though preferences do not directly enter the wage equation. These two results imply, for example, that parental education and gender could affect the financial return to education even if they have no effect on wage rates. Altonji (1993) demonstrates

[9]See Altonji (1993) for proofs of the claims in this section using a stripped-down version of the model sketched in this section.

this empirically.

Sequential decision making under uncertainty about preferences, ability, and knowledge accumulation opens up some interesting possibilities concerning the effects of wages on education outcomes. One example is that an increase in expected wages of college graduates holding the wages of high school graduates and college dropouts constant boosts the high school graduation rate and increases college enrollment but has an ambiguous effect on the college dropout rate. To see this, note first that the college wage increase will increase the ex ante return to starting college, holding constant human capital at the end of high school and beliefs about preferences and ability. On one hand, this will induce some individuals with relatively low probabilities of completing college to start, raising the college dropout rate. On the other hand, the higher payoffs will induce some individuals who would otherwise have dropped out after the first period of college to continue. A second example is that an increase in the payoff to a degree in one field, say engineering, holding the ex post payoff in other fields constant, can lead to an increase in the graduation rate in the other fields. The reason is that the increase in the engineering wage raises the ex ante return to starting college. Consequently, this higher wage will increase the number of persons who spend the first period of college in a program geared toward completing a degree in engineering by more than it reduces the number who start college in a program geared toward a humanities degree. If enough of the college entrants who start in engineering ultimately conclude that they prefer the humanities major and the occupations it leads to, and/or enough conclude that they are much more likely to be able to meet graduation requirements in humanities than in engineering, then the flow into humanities following the first period of college could be enough to offset the smaller number who start college with the intention of pursuing a humanities degree.

The role of risk aversion brings forward two important empirical issues. First, individuals will also care about uncertainty in ability and the effect of this uncertainty on alternative degrees or courses of study, possibly avoiding ones where ability might matter a lot. In general we will see a risk-return trade-off induced not only from the macroeconomic environment and the possible volatility of returns but also because of uncertainty in individual ability and microeconomic uncertainty in wages within each sector.

2.5.2 The returns to education paths

The model offers a systematic way of defining and measuring returns to education. Thus, for example, we can consider the returns to an economics major from the perspective of someone having to make a choice of major. This will be defined as

$$R(\Omega_{C_1}, \zeta^j_{C_2}, \xi_{C_2}) = \frac{V^j_{C_2}(\Omega_{C_1}, \zeta^j_{C_2}) - V^{\mathcal{L}}_{t=0}(\Omega_{C_1}, \xi_{C_2})}{V^{\mathcal{L}}_{t=0}(\Omega_{C_1}, \xi_{C_2})}$$

This depends on individual heterogeneity, known to the individual but not to the econometrician. The estimated return will be an average of this function. Interestingly, this return will depend on individual history. Thus, we can document how the returns differ across individuals with different choice history of courses, grades, etc. The ex ante heterogeneity of returns is driven partly by these differences in early experiences and partly by the way that they affect individual perceptions and expectations of their ability and preferences. Moreover, they take fully into account the costs of education, including opportunity cost and effort involved in alternative course choices.

2.5.3 Econometric implications

The model we presented also brings to the fore econometric issues that will appear when reviewing the literature. First, wages across different levels of education and fields of study are best described by a Roy-type model, implying heterogeneous returns to education. In other words this is a switching regressions framework (see Quandt (1972), Heckman and Robb (1985)). Ex post the selection into any of the sectors will be endogenous to the extent that preferences about education are correlated with unobserved determinants of wages (lazy at school and lazy at work for example) and to the extent that information about future wages is known by the individual at the time education choices are made, but is unobserved by the econometrician. For example, if an individual knows they will be more productive working in an office rather than outdoors, this will influence the sequence of education choices and imply that education choice is endogenous, even if education and labor market ability are themselves independent.

This brings us to the final difficult question of identifiability and identification in practice. The model highlights the costs of education as the main source of exogenous variation: a low-cost college can trigger a chain of decisions that leads to college completion, where this may have not happened if fees were higher. Thus if there is variation in the costs of obtaining alternative combinations of education, and if such variation can be taken as exogenous (i.e. not correlated with quality of education or with the characteristics of the individuals having access to such fees), this can be an important source of exogenous variation. However, as shown in Heckman and Navarro (2007) and further discussed in Meghir and Rivkin (2010), this may be far from enough to identify such dynamic models non-parametrically. Thus, in practice, identification will be in part driven by such exogneous variables and by restrictions on the functional forms of the distribution of unobservables.

The dynamic model presents a clear approach to estimating both an education choice model and the returns to education, as well as the specific properties of wages. Any source of endogeneity is taken into account and there is transparency regarding the assumptions made. However, it is a complicated model; to include all the detail we have suggested may be almost impossible in practice. Thus, a dynamic approach will have to impose a number of simplifications. Meanwhile, most of the literature we review has taken the simpler approach of estimating relatively simple wage equations based on instrumental variables (and sometimes OLS). In general, however, IV techniques will not provide interpretable estimates except under very strong assumptions. In the simplest case of binary treatments (e.g. attend college or not) with heterogeneous effects, Imbens and Angrist (1994) have shown that IV will identify the Local Average Treatment Effect parameter (LATE). This is the effect of the binary treatment on those who were induced into treatment (college in our example) by the variation in the value of the instrument. Suppose, for example, as in their paper, that the instrument is also binary (e.g. the implementation of a policy). If the instrument is independent of all the unobservables, such as would be the case if it had been randomised, then LATE identifies the effect of those who switched into treatment as a result of the reform under the key additional assumption that no one is induced out of treatment as a result of the policy (monotonicity).

With many different treatments, as would arise in our context of multiple levels of education and curriculum choices, instrumental variables would not provide results with an obvious interpretation, even if we had enough instruments. This of course does not mean that the dynamic model is necessarily our only alternative. Consider, for example, the wage equation, and take expectations conditional on a set of instruments that have been randomized (Z) and on educational history denoted by S. Assume that the dimension l indexes the education path. (We will ignore for this example the labor market history as well as time in the labor market.)

$$E\left[\ln w_{il}|S_i = S_i^l, Z = z_i\right] = \alpha_{lt} + \beta_{s^l} S_i^l + E(u_{il}|S_i = S_i^l, Z = z_i)$$

This is a standard Roy model and its identifiability has been studied by Heckman and Honore (1990). β_{s^l} is the average ex post return to education path S_i^l. In terms of our notation the issue is whether we can construct the control functions $E(u_{il}|S_i = S_i^l, Z = z_i)$ and more specifically what assumptions would ensure that it has independent variation from the education choices S_i^l. The form of this function is driven by the dynamic selection process described earlier in the theoretical model. Constructing this control function will require a model for S. In general this is not going to be a single index model; hence the need for many instruments. For example, Cameron and Heckman (1998) have shown that education choice can be represented by an ordered choice model (which would simplify the identification and estimation problem) when there is just one unobserved factor driving educational decisions. When we depart from this very restrictive framework, we will need more than one index to represent the choice of S_i as well as many instruments. This is an issue of both sufficient amounts of exogenous variation and the correct structure. In interpreting the results of the existing empirical literature, we need to remember that important identification questions remain both when implementing the full dynamic structure and when estimating a simpler version of the model. This does question the interpretability of the various IV results we discuss, even if there were general agreement that the instruments are appropriate. If, however, there is only one heterogeneity factor and the impact of each education path is homogeneous, many of these difficult identification problems disappear and we are just left with the need for a suitable number of instruments that can explain education choice. In this case, we can do away with the model as far as estimating the wage equation is concerned.

3 Empirical evidence on choice of high school curriculum

There is relatively little empirical work in economics on how high school students choose curricula, particularly in comparison to research on choice of college major. The main margins of choice for a typical high school student are among vocational, general, and academic curricula, and, within the latter, between a focus in social sciences and humanities and a focus in mathematics and science. The level and number of courses (subject to promotion and graduation requirements) are also choice variables. Contrary to popular perception, the number and level of courses in academic subjects taken by high school students in America have risen over the past thirty years. Data on course taking by high school seniors reported in Ingels et al (2008) for the years 1982, 1992, and 2004 shows an increase from 1982, particularly in science. Supplementary Table 1 compares course-taking trends from 1990 to 2009 for high school graduates, as reported by Nord et al (2011). Course-taking overall is up, and in particular, course-taking in core academic subjects (mathematics, science, social science, and English) have all increased. Furthermore, the percentage of students taking more rigorous programs of study has increased as well. Of course, one would expect curricum choice to change over time as the occupational mix of labor demand changes. As we discuss in Section 4, part of the trend is due to changes in state level graduation requirements. To some degree, state requirements reflect perceptions of what is required for the labor market and for post-secondary education. It would be interesting to decompose trends in college attendance and major choice over the past 40 years into the contribution of changes in high school curriculum and the contribution of changes in the link between high school curriculum and major choice.

We are not aware of any study that has estimated a structural model of high school curriculum choice along the lines sketched in the previous section. Zietz and Joshi (2005) use the NLSY (1997) to estimate a two-period model of leisure maximization, subject to minimum consumption constraints. They find that "academic aptitude, pre-high school academic performance, and lifetime consumption goals as driven by peer pressure and family background are by far the most important determinants of program choice." Meer (2007), using the NELS88 data, finds that the principle of comparative advantage is at play when students choose between academic and vocational high school curricula. In the remainder of this section, we briefly summarize some of the descriptive evidence on the determinants of curriculum.

There is a substantial literature on the role of gender, race/ethnicity, and socio-economic status in determining high school curriculum, which we touch on briefly here. Historically, girls tended to take less math and science than boys, despite equal (or greater) opportunities or prior achievement (Oakes (1990), Catsambis (1994), Ayalon (1995)). These differences begin to emerge in middle school (Ma & Willms (1999), Catsambis (1994)). Girls tend to have less positive attitudes toward or fewer aspirations for careers in math or science; they are less interested in math and less confident about their mathematics abilities (Dick & Rallis (1991)). However, Goldin et al (2006) report that among graduating seniors the male/female ratio of mean number of high school courses in math, science courses, and chemistry declined from between 1.3 and 1.4 in 1957 to between 0.9 and 1.0 in 2000. In physics, the male/female ratio declined from 3.1 to 1.21.

Students from high SES backgrounds tend to be streamed into more academic tracks; this can be explained by a variety of factors including higher intrinsic ability (cognitive or non-cognitive), better preparation in primary school, peer effects, or parental lobbying (Vanfossen, Jones & Spade (1987)). Interestingly, African-American and Latino students have positive attitudes toward math, despite low achievement (Catsambis (1994)). That said, minorities enroll in math-intensive courses at lower rates than whites; however, this is mostly explained by SES and prior achievement (as measured, for example, by GPA) (Ferguson (2009)).

Parental influence manifests itself in many ways, both direct and indirect. Firstly, parents provide a genetic endowment to their children as well influential early childcare with important outcomes in terms of cognitive and non-cognitive abilities. Secondly, they choose neighborhoods which in many cases determine the high school their child attends, which in turn affects course selection directly (through the course offerings and tracking regulations at the school in question) or indirectly (through the implicit selection of the student's peers). Finally, parents can influence course choice directly, either by incentivizing or constraining their children to make particular choices, or by lobbying schools and teachers (Useen (1992)). Parental influence varies by SES and race.

3.1 School-level influences

School policies, particularly course requirements, scope of offerings, and tracking guidelines, are an important influence on curriculum choice. These policies are in turn shaped in part by state and school district regulations, as we document below. A substantial fraction of the variance in curriculum choice is across high schools (see note 9).

To the extent that school behavior can be taken as independent of the unobserved characteristics of the students that attend the school, such variation can be (and has been) used to identify the effects of particular high school curricula. In practice, however, school choice and cross school competition and specialization (attracting particular parts of the student market) will undermine the credibility of such a strategy. In Section 4, we discuss three studies that use curriculum reforms as a source of exogenous variation.

Peers, teachers and facilities (e.g., availability of science labs or a theater) may also influence curriculum choice, but attempts to identify the causal effect of these factors are subject to the same endogeneity problem.

4 The effects of high school curriculum on educational attainment and wages

As we noted in the introduction, there is surprisingly little hard evidence about the causal effects of specific high school courses on educational attainment and labor market outcomes.[10] Historically, there has been a lack of data sets that contain detailed student data on course taking, test scores, and post-secondary outcomes. That problem has been remedied, particularly in the US, although lack of data on adult labor market outcomes continues to be a serious limitation. However, data are not the only problem. As is clear from the model in Section 2, student course selection is not random given the available options. Furthermore, student curriculum choices are shaped by school requirements, tracking policy, and guidance, and these reflect to some extent the qualifications and interests of the student body. Even with excellent data, it is difficult to identify the causal effects of courses on educational attainment, choice of college major and occupation, and wage rates.

[10] Due to space constraints, we do not discuss evidence on the effects of curriculum on test scores.

In this section we review the limited evidence on the effects of high school curriculum. We emphasize wage effects but also touch upon educational attainment and choice of major. We discuss the approaches to estimation that are used and briefly summarize the results. None of the studies we reviewed model the endogeneity of curriculum choice by allowing for dynamics. The implication is a lack of clarity of the underlying determinants of both wages and educational choice, as we discussed in Section 2. The existing empirical studies estimate equations of the form

$$S_i = C_i G_s + X_i B_s + C_i g_{si} + e_{si} \tag{4}$$

for educational attainment S and

$$\ln Wage_i = C_i G_w + X_i B_w + \rho S_i + C_i g_{wi} + e_{wi} \tag{5}$$

for wages where here C_i denotes high school curriculum, X_i denotes background and other characteristics, e_{wi} and e_{si} are unobserved random intercepts, and g_{si} and g_{wi} are unobserved random coefficients. C_i is typically a count of the number of year-long courses taken in various subjects. It is likely to be correlated with the composite error terms in the equations, which will lead to bias in OLS estimates of G_s and G_w. While many studies recognise this endogeneity of curriculum choice, they differ in a number of ways including in the assumed exogenous source of variation for the curriculum, in whether or not all courses are examined at the same time rather than one by one or in subsets, in whether courses in a given subject area are differentiated by level, and in whether the analysis is conditional on high school graduation.

An important issue is whether one should control for post-secondary education S in the wage equation. Doing so affects the interpretation of the coefficient on C. From a theoretical point of view, the entire history and type of educational attainment will affect wages to the extent that each path leads to a different set of skills. As specified in (5), G_w is the average effect of courses on $lnWage$ holding years of post secondary education constant. For example, high school math and science courses may influence the types of jobs obtained by students who enter the workforce after high school or facilitate completion of a BS in engineering as opposed to a lower-paying college major. Controlling for post-secondary education in some fashion makes sense because post-secondary schooling is costly. However, if the return ρ exceeds the interest rate, then part of the return to C_{ih} is through facilitating profitable investments in S. It raises the option value of investment in S. Altonji (1995) reports a set of estimates based on including S in the wage equation with ρ set to 0.04, which he assumes is the real discount rate. Most studies report estimates with and without S. None of the papers that include S addresses the fact that it is endogenous in the wage equation for many of the same reasons that C_i is endogenous.

We now to turn to the studies. Altonji (1995) is the first comprehensive study of the effects of curriculum on post-secondary educational attainment and wages.[11]He uses data from the National Longitudinal Study of 1972 (NLS72),

[11]Rumsberger and Daymont (1984) is a noteworthy early study of the effects of curriculum on wages shortly after high school. They divide courses into academic, vocational, and "other" courses rather than look at specific subjects. They find a positive relationship between

which follows samples of 12th grade students from a large number of high schools. His specification of (4) and (5) includes counts of courses in eight subject areas: math, science, English, social studies, foreign languages, commercial studies, industrial arts and fine arts. His main identification strategy uses the substantial amount of variation across high schools in the average value C_h of C_i as an instrumental variable for C_i, where h is the high school i attends.[12] If the variation in C_h reflects high school or school district policies that are unrelated to the distribution of aptitude, ability, and prior preparation of the student body, then use of high school averages as excluded instruments would yield consistent estimates of the average treatment effect of an extra semester of math, science, etc. But casual observation, as well as the evidence from the NLS72 and other similar data sets, indicates that variation across schools in the quantity of academic courses is positively related to both the level and quality of the courses and to the quality of the students. This association would imply a positive bias in G_w and G_s, because course quality and level is not accounted for and the controls for family background, primary school preparation, and high school quality are imperfect.

Altonji's IV estimates indicate that the effects of additional courses in academic subjects are small. Even when controls for family background and ability are excluded, the combined effect of an extra year of science, math, foreign language, English, and social studies is only 0.3 percent—far less than the value of a year of high school. The combined effect of an extra year of science, math, foreign language is 0.030 when only basic controls are included, but falls to 0.017 (0.012) when family background controls for the student and the school are added. The point estimate for math is actually negative, -0.007 (0.015), although one cannot rule out a substantial positive effect given the standard error. The IV estimates indicate that students who do not go to college benefit from vocational courses.

Altonji also reports OLS estimates. These are somewhat larger than the IV estimates, but nevertheless also suggest that the value of additional courses is too small to account for the value of a year of high school.[13] The conclusion is not very sensitive to how post-secondary education is treated, in part because the courses have only have a modest effect on post-secondary education. Altonji's conclusion is not that courses do not matter, but rather that the results, which are not easy to dismiss with an appeal to unobserved heterogeneity, pose a challenge for researchers. He raises the possibility that the estimates for particular courses are affected by the interaction of biases in estimates of the curriculum variables and perhaps by the control variables that are in the model.

Levine and Zimmerman (1995) use a framework and methods similar to Altonji's. However, they use different data sets (NLSY79 and High School and Beyond (HS&B)), focus on math and science classes, stress differences by gender, and look at a broader set of outcomes. They rely primarily upon OLS because their estimates using cross high school variation in course taking are noisy. Controlling for post-secondary education, their OLS estimates suggest that a year of math raises wages by between 2.8 and 5.6% for men and about 4.4% for women. The point estimates vary within education group and are largest for female college graduates. The estimates of the effects of science are mixed in sign

total credits and hourly earnings, but the link is weak among high school graduates.

[12] For example, he reports (Table 1) that cross-school variation accounts for 25.9%, 26.9% and 26.4% of the variation across students in science, foreign language, and math, respectively.

[13] The simple OLS regression coefficients relating the log wage to a year of science, foreign language, and math are 0.054, 0.040, and 0.072, respectively.

and not statistically significant.

Levine and Zimmerman find that both math and science courses boost educational attainment. There is evidence that additional math and science courses increase the probability of choosing a technical college major. This evidence relates directly to the question of how the type of education one chooses at the high school level conditions the optimal choice of education and occupation later. Additional math and science courses move women toward jobs that involve more mathematical reasoning, although the evidence is stronger in HS&B than in NLSY79.

As the authors emphasize, the OLS estimates may be biased upward by selection in course-taking and the other issues discussed in more detail in Altonji (1995). Nevertheless, the findings of the positive effect of math and the evidence pointing to an effect of math and science on the probability that the individual majors in a technical field are interesting.

Rose and Betts (2004) use the 1982 senior cohort from HS&B to provide a more nuanced study of the return to high school math. Their primary dependent variable is log earnings in 1991. They control for demographic characteristics and family background variables in most of their analysis, include high school dropouts, and control for highest degree. Their main innovation is to take advantage of detailed transcript information and differentiate math courses by level—vocational math, pre-algebra, algebra/geometry, intermediate algebra, advanced algebra, and calculus. They find that a year of math substantially raises wage rates even when math GPA and math test scores are controlled for. The returns are larger for advanced math courses, particularly algebra and geometry. They obtain positive estimates using OLS and an IV approach based on cross-school variation, although the OLS estimates are stronger, perhaps in part because they are more precise.

The math results are robust to including English, science, and foreign language. English courses enter positively, and the OLS estimate suggests the return of 2.6% to a semester of upper-level English, while the IV estimate is 0.071. The coefficients on science courses are negative for lower-level courses and positive but not significant for upper-level physics, chemistry, and AP biology. The OLS estimates indicate that a student who takes a year of calculus, English, chemistry, and foreign language would earn about 8.6% more than a 12th grade student who did not take any courses. Consequently, Rose and Betts come closer than Altonji in accounting for the value of a year of high school. On the other hand, the return to a year of low-level courses in the 11th grade and in the 12th grade is small. Rose and Betts attempt to reconcile their OLS and IV results with Altonji's smaller estimates. They conclude that the difference stems from their disaggregation by course level in case of IV and disaggregation by course level and their use of math GPA rather than test scores as a control in the case of OLS. Consequently, the type of math one takes appears to make a difference. On the other hand, Rose and Betts' IV estimates are large and negative for several of the advanced math courses. Some of the differences between the two studies may reflect sampling error. Furthermore, concerns remain about positive selection in who is taking advanced courses, as Rose and Betts are aware. But overall, their study significantly advances our understanding of the payoff to different types of courses.

Joensen and Neilson (2008) estimate the return to advanced high school math by exploiting an educational reform in Denmark. Denmark requires students to choose among discrete packages of courses. The packages are standard across high schools. Until 1984, students could only take advanced math in combination with advanced physics. This was changed in 1988 to allow students to combine advanced math with chemistry. The reform was piloted at some high schools during the years 1984 through 1987. Joensen and Neilson make the case that the reforms were exogenous for students who had chosen their school prior to when the pilot was introduced. They find that taking the advanced math course in combination with the advanced chemistry course increases earnings by 0.20 log points. The estimate rises to 0.25 when they exclude schools that are able to unilaterally decide whether to implement the pilot program because there is evidence of negative selection for that group. Interestingly, the OLS estimate is similar (0.23) when detailed controls are included, as is the IV estimate based on the cross high school means C_h. There is not much evidence of selection bias.[14] Controlling for post-secondary educational attainment eliminates most of the effect, in contrast to the US studies.

In interpreting their estimates, Joensen and Neilson make the monotonicity assumption that students who had chosen advanced math when only the advanced math and physics combination was available would continue to take an advanced math package following the introduction of the advanced math-chemistry option. This is reasonable, but it is important to point out that even if this assumption is true, their estimator does not have a LATE interpretation as the effect of advanced math. This is because their IV strategy precludes controlling for whether the student took math in combination with physics or math in combination with chemistry. Some of the students who were induced to choose the advanced math and chemistry combination would have chosen a curriculum involving less math, while some would have chosen advanced math with advanced physics. Despite this ambiguity in the interpretation of the results, this study is a valuable contribution that illustrates the potential for research that exploits sharp curriculum reforms.

Goodman (2009) takes advantage of curriculum reforms inspired by *A Nation at Risk* (1983). In the years following the report, a number of states established course requirements in core academic subjects or increased existing ones. Goodman's instrument is the indicator $MathReform$ for whether students in a given high school class from a given state were subject to an increased math requirement relative to what prevailed for the high school class of 1982. Because the state laws are not very powerful instruments, a large sample could help in improving precision if the within state error correlation is limited. Goodman employs a Two Sample Instrumental Variables (TSIV) estimator that uses micro data sets on student high school transcripts to estimate the first stage and wage and educational outcome data from the 2000 Decennial Census to estimate the second stage. Controlling for state and high school cohort, $MathReform$ is associated with an increase of about 0.184 years of basic math and 0.214 years of advanced math for black males. The impact for black females is somewhat smaller, and the impact for whites is about half as large and is concentrated in basic math. The corresponding TSIV estimates of the effect of a year of math are 0.083 (0.023) for black males, 0.078 (0.028) for black females, 0.089 (0.061) for white males, and -0.030 (0.048) for white

[14] Note that the OLS results are weighted toward students who took math in combination with advanced physics because the advanced math and chemistry combination was only available in the pilot schools. Consequently, OLS and IV estimators identify different parameters.

females. About 40% of this effect is due to an increase in educational attainment.[15] When education policy controls, economic controls, and census division level trends are added, the TSIV estimates become 0.052 (0.018), 0.035 (0.030), 0.033 (0.030) and 0.005 (0.041) respectively. The effects tend to be larger for disadvantaged groups, but one should not make too much of the differences given the sampling errors. It should also be kept in mind that the specification does not include controls for other courses. (The reform instruments are not powerful enough to identify the effects of different courses.)

4.1 Conclusion about the effects of high school curriculum

Over the past 15 years, some progress has been made toward the goal of providing hard evidence on the effects of high school curriculum on wages and educational attainment. Breaking out courses by level, as in the Rose and Betts study, indicates that taking more advanced courses has a substantial economic return. While there are clear concerns about selection bias, their result is reinforced by Joensen and Neilson's clever IV study, although it is worrisome that they find that most of their large effect of taking advanced math and chemistry is through post-secondary education independent of the field of post-secondary education. These results, along with Goodman's TSIV estimates and Levine and Zimmerman's results, suggest that additional math courses have substantial value. There is also some evidence that vocational courses have value for students who did not intend to go to college. (See, for example, Altonji (1995) and Mane (1998) and additional references that she provides.) However, there is still considerable uncertainty about the value of particular courses of study, about how the courses affect post-secondary education, and about heterogeneity in these effects.

We have two suggestions for research. The first is to make use of the huge student record data sets that have become available in several US states and other countries. This information, combined with post-secondary education records and earnings (as is possible, for example, for the states of Florida and Texas) and in conjunction with information about specific reforms to curriculum at the high school or school district level, provides a way to address the problem of endogeneity of curriculum. Furthermore, the data sets can be used to study how curriculum affects post-secondary field, which has been difficult to do because of sample size considerations in the panel data sets discussed above. The second is to integrate the study of curriculum choice into a dynamic model of educational attainment and labor market outcomes, along with sources of exogenous of variation in curriculum. This would help elucidate how courses are chosen and how education choices today influence future opportunities. Furthermore, it will provide a systematic way of addressing selection in course-taking in a dynamic setting when estimating the return to curriculum.

[15] We arrived at the 40% figure based on the fact that the reduced form coefficients relating log earnings to *MathReform* fall from 0.028 to 0.019 (0.008) for black males, 0.026 to 0.016 (0.008) for black females, and 0.020 to 0.011 (0.010) for white males when education/graduating class interactions are added to the baseline specification. The estimates are negative but not significant for white females regardless of whether education is in the model.

5 Empirical evidence on the determinants of college major

The theoretical framework presented above implies that college major choice is influenced by expectations of future earnings, preferences, ability, and preparation. We discuss some of the evidence on the role of each of these factors, drawing on both the small number of structural models and on the more extensive reduced-form approaches. First, however, we document some trends.

5.1 Trends in college major

Patterns of college major choice have been relatively steady over the past twenty years. Supplementary Figure 1 illustrates the trends across college graduation cohorts in the fractions of degrees accounted for by the five most popular aggregated college majors. The results are shown separately for men and women.[16] In keeping with the high school course-taking trends, the fraction of science majors has shown a slight uptick in recent years. Taking a longer view, the fraction of education majors has decreased substantially over the past forty years, while the proportion of business and economics majors peaked about twenty-five years ago, following a period of rapid growth in the 1980s. Later in this section we estimate the contribution of this change in composition to the growth of the college wage premium in the 1980s.

Not surprisingly, women are far more prevalent in education, and men in engineering and in business and economics. However, this figure disguises interesting within-category trends. While women have caught up with men in the sciences overall, Turner & Bowen (1999) point out that the disaggregated trends show a surge of women in biology, for example, but not in the more mathematical sciences. Similarly, the oft-used "social sciences" category includes majors as dissimilar as economics and sociology. To address this, Figure 1 reports the trend over time in the fraction of women in selected disaggregated majors relative to the fraction of women with a degree in any field. The trend in the relative fraction is not affected by the rise in the share of women among all college graduates. Remarkably, although the fraction of women in most science-related majors has increased over time, the relative fraction female among computer science majors has actually dropped.[17]

Table 1 describes various characteristics associated with some of the more common college majors: share enrolled, fraction female, math and science course content, SAT scores, wages, and the share attending graduate school.[18] There is a great deal of variation across majors: engineers have among the highest SAT math scores, for example, while elementary education majors have among the lowest. Similarly, wages tend to be high for engineers and low for elementary education majors, suggesting that perhaps much of the wage differences between majors are due to

[16] Data are from the 2009 ACS; respondents are assumed to graduate at age 22. Supplementary Figure 2 reports trends from 1984 to 2009 based on data from the Integrated Postsecondary Education System (IPEDS). There are some minor differences that could be due to recall bias in ACS, changes in IPEDS coverage, the fact that the ACS includes immigrants who received degrees abroad and IPEDS includes degrees of foreign students who later return home, the fact that some individuals obtain more than one bachelors degree, and differences in aggregation of majors.

[17] See Goldin et al (2006) for an excellent overview of long-run trends in the educational attainment and high school curriculum and college major choices of women.

[18] The wage and share data are from the ACS, and the other data are from the 1993/2003 Baccalaureate & Beyond survey.

differences in mathematical ability and high school course work.

The table also reports the 10th and 90th percentiles of the hourly wage distribution. The size of the gap relative to the mean varies substantially across majors. For example, mean earnings in economics is about 10 percent higher than in mechanical engineering, but the 90-10 differential is twice as large. Part of the variation in the 90th percentile reflects differences across fields in the contribution of graduate school, but the basic pattern is present for those who do not go to graduate school.[19]

5.2 Expected earnings

The work on the impact of expected earnings on major choice uses three main approaches. The first uses a rational expectations-type framework in which expected future earnings are based upon a statistical model of earnings. Berger (1988) is an early example of this. He models the utility from major j as the consumption value of the major plus the present value of expected lifetime earnings in major j. In doing so he improves upon previous myopic models that incorporate only first-year post-college earnings. The coefficient on expected earnings in the conditional logit model has the expected positive sign.

However, Berger's model does not account for uncertainty about preferences, ability, or academic progress. Changes in these factors will lead some students to leave school or change fields, as emphasized in the theoretical section. In Altonji's (1993) model, uncertainty regarding ability and preferences leads to probabilistic major completion even if ex post monetary payoffs are known. He emphasizes the distinction between the ex ante and ex post return to a particular course of study. (Note that the estimates in the next section are all ex post returns; that is, returns contingent on completion of a major.) The ex ante return is the return associated with starting a particular major, and includes the possibility of dropping out entirely, switching majors, or proceeding on to a graduate degree. Altonji does not estimate his theoretical model, but Arcidiacono's (2004) econometric model has some of the same features.[20] Montmarquette et al (2002) break down expected (ex ante) returns into the probability of completing a major, and earnings contingent on completing a major. They find that expected earnings (the product of the two) has more predictive power than either on its own.

While many papers include the expected wages associated with a terminal college degree in the choice equation, Eide & Waehrer (1998) also incorporate the option value of graduate school. For both men and women the option

[19] For example, among individuals without graduate degrees the mean, 10th percentile and 90th percentile values are 43.7, 14.3, and 80.6 for economics and 41.2, 20.1, and 61.9 for mechanical engineering. See Supplementary Table 2.

[20] Altonji estimates probit models for the probability of 19 post-secondary education outcomes as a function of demographic characteristics, family background, high school curriculum, and a vector of 12th grade aptitude and achievement measures. He also estimates a model of the wages associated with each of these outcomes. He uses the education outcomes in the wage payoffs to compute the ex ante return to starting college as a function of student characteristics. The ex post payoffs matter in proportion to the probability that they will be realized. He provides estimates for men and women of the effects of family background, high school curriculum, and aptitude and achievement on the ex ante return to starting college. He shows that students with higher ex ante returns are more likely to start college and that 12th graders who expect to attend college have a higher ex ante return. Stange (2011) estimates a dynamic structural model of education choice which shows that the option value of early choices can be substantial. See also Heckman, Lochner and Todd (2008). Neither of these papers consider field of study. See Altonji (1993) for references to a few earlier papers that consider the implications of uncertainty about completion probabilities in examining the demand for and return to education, including Manski (1989). Weisbrod (1962) introduced the concept of the financial option value of a year of education.

value of graduate school increases the likelihood of majoring in science or liberal arts relative to business, although the magnitudes differ significantly by gender.

Although most papers that have examined the role of expected wage rates on education choices have used a statistical model of wages as the basis for expectations, a few recent papers directly measure expectations. These papers make use of specialized surveys (usually confined to one school) to assess students' subjective expectations about wages. As we note below, they also assess the relative importance of monetary returns versus preferences for particular majors and occupations. Betts (1996) interviews students at the University of California, San Diego; Zafar (2009a,b) interviews sophomores at Northwestern; Arcidiacono, Hotz, & Kang (2010) interview male students at Duke; and Stinebrickner & Stinebrickner (2011) conduct a 12-wave panel survey of Berea College students. A more representative sample of colleges would be preferable, but these studies provide unique insights into college students' decision making.

Betts (1996) finds that seniors are much better informed about wages than freshmen, suggesting that students may wait to learn about their abilities and preferences before investing in information-gathering. This has implications for the timing and informativeness of information shocks about the labor market in our model, relative to the preference and learning shocks ζ and ϵ. Given that it is difficult to major in science and engineering if one does not lay the foundations in freshman and sophomore year, this raises the question of whether students obtain the labor market information after it is too late. Betts also shows that students from lower-income families systematically underestimate earnings for college-educated workers.

Arcidiacono, Hotz & Kang (2010) find that major choice is based on comparative advantage, in that students expect to have higher earnings in the major in which they are currently enrolled than in other majors (with the exception of economics, which all students believe would lead to higher earnings). Students also tend to be more accurate about future earnings in their own major than in other majors. Zafar (2009b) and Betts (1996) find similar results. Reassuringly, students' earnings expectations are fairly accurate.

The literature on choice of major also considers risk aversion. With concave preferences, students should consider both the variance as well as the mean of earnings associated with particular programs of study. (See table 2 of the descriptive section for 10th and 90th percentiles of wages by major.) Saks & Shore (2005) find that students from wealthier families are more likely to choose "riskier" majors, as is implied by a model in which agents have decreasing absolute risk aversion. Christiansen et al (2007) investigate the risk-return trade-off in major choice and conclude that many majors are not at the efficient frontier. However, one would want to take into account costs of education in a more comprehensive fashion, including monetary and effort costs as well as opportunity costs. The frontier interpretation may be very restrictive in this context.

5.3 Preferences

Researchers have of course recognized that non-pecuniary aspects of majors and their associated occupations may play important roles as well. While many researchers relegate preferences to the error term in the choice equation, others have endeavored to be more explicit about the role of preferences. Daymont & Andrisani (1984) make use of survey questions regarding the importance of various job characteristics (being a leader, working with people, making lots of money, helping others) and run regressions of realized majors on these characteristics; they also find important gender differences in these characteristics. Nearly a third of the gender difference in choosing a business major is explained by these variables, whereas for other majors these variables play only a very minor role. However, this paper does not include ability controls or test scores. Using data for a large university, Leuwerke et al (2004) find both mathematical ability and congruence between the student's stated task preferences and the requirements of engineering predicts persistence in engineering.

Blakemore & Low (1984) find that women tend to choose majors that are subject to less atrophy, so that time away from the labor market for child care will be relatively less costly. Turner and Bowen (1999) also outline stark gender differences in major choice but point out this could be due to differences in preferences or due to the "chilling" effect of past labor market discrimination. They also highlight the importance of finer degrees of disaggregation in major choice, pointing out that biology and physics, while often grouped together, in fact require very different skill sets.

Zafar (2009a) is principally interested in decomposing gender differences into differences in abilities and preferences. He finds that preferences play a strong role. For example, in addition to expected future earnings, students care about enjoying coursework, parental approval, and the social status of future occupations. Moreover, while the choices of men depend roughly equally on pecuniary and non-pecuniary attributes, women's choices depend roughly twice as much on non-pecuniary attributes than pecuniary ones. Stinebrickner and Stinebrickner find that student preferences and ability play a crucial role in field of study.[21]

5.4 Ability and preparation

We have not yet said anything about how high school preparation and innate ability conditions choice of major. Turner & Bowen (1999) show that the effects of SAT scores (math and verbal) are non-linear and, moreover, that the effects differ by gender. These scores account for 45% of the gender gap in math/physical sciences, but only 8% in psychology, leaving a great deal of room for gender differences in preferences. Zafar (2009a) also rules out gender differences in beliefs about ability as a driver of gender differences in choices.

Stinebrickner & Stinebrickner (2011), having the advantage of panel data, are able to track student major choices over time. Rather than asking students to state their expected major, they ask the students to assign probabilities

[21] Other papers in the large literature in economics and other fields on gender differences in major choice include Dickson (2010) and Canes and Rosen (1995).

to completing various majors (see also Manski (1993) and Manski (2004)). They find that although many students begin their college careers assigning a high probability to finishing a science major, many learn that their abilities are not adequate. In our model, we incorporate this new information about one's abilities through performance shocks in college. Ost (2010) echoes Stinebrickner & Stinebrickner (2011) and Arcidiacono (2004) in documenting a positive relationship between grades and persistence in a particular major. However, Ost points that physical science majors get higher grades in non-science courses than in science courses and thus may be tempted away from science. This effect is exacerbated by increasing grade inflation in non-science fields relative to science fields. He finds additionally that the presence of high-achieving peers in the physical sciences positively influences one's own persistence. It is somewhat artificial to separate ability and preparation from preferences in determining major choice. Ability to pursue a particular course of study with a reasonable level of effort and to perform well in related jobs influences the utility associated with the activity, holding wages constant.

6 Returns to college major: empirical evidence

Unlike the literature on the returns to high school curriculum, the literature on the returns to college curriculum is substantial. In this section we discuss a selected sample thereof. We also present descriptive estimates of returns to different majors using data from the ACS, supplemented with characteristics associated with different majors (average SAT scores, numbers of math and science courses, etc) extracted from the Baccalaureate and Beyond 1993 dataset. We begin with a discussion of the estimation strategies used and possible alternative ones. We then turn to the results.

6.1 Estimation strategies in the existing literature

Estimating the returns to college major is fraught with many of the same difficulties that plague estimations of the average return to schooling more generally. The main difficulties are omitted variables that influence both choice and earnings, and selection bias based on heterogeneity in returns. To the extent that students select into particular majors on the basis of their anticipated future returns, OLS estimates of the returns will be biased and will not reflect only the causal impacts of major choice. As such they will represent neither the average treatment effect, nor the effect of treatment on the treated.[22] Most papers nevertheless use OLS and hope that controls are adequate.

Berger (1988) is one of the few papers that attempts to address this kind of selection. He uses a conditional logit with five major categories and finds that students choose majors on the basis of the present value of expected lifetime earnings. To obtain identification, he allows family background variables (father's occupation, parental education, race, etc) to affect the choice of major without affecting earnings and allows ability measures (IQ, Knowledge of World of Work) and cohort to affect major choice only through the earnings equation. Both of these assumptions are questionable. There is clear evidence that the psychic costs and benefits of college depend on ability and performance—

[22]Sample selection bias associated with labor force participation is also an issue, as Hamermesh and Donald (2008) show (see below). Few studies address it.

see for example Arcidiacono (2004). Furthermore, family background might easily affect general skills and may thus influence earnings as well as the payoff to particular majors.[23]

The coefficient on the selection correction term is found to be significant for only two majors within each specification (education and liberal arts in the case of log wages; education and science in the case of log earnings). While Berger interprets the signs on these coefficients as indicative of positive selection, the lack of consistency across specifications and lack of significance makes this less compelling.

The final approach is to use a dynamic discrete choice model along the lines of Section 2. Arcidiacono (2004) uses this approach.

6.2 Possible alternative strategies

Other methods to control for selection in major choice do not appear to have been used. If there were sufficient regional variation in offerings across colleges, such as the availability of an engineering school, a variant of Card's (1995) "distance to college" instrument might be tried. Freeman (1976) and Siow (1984), among others, show that the supply to specific occupations depends upon market conditions at the time that students choose a field of study. More work could be done using variation across time and place in the demand for particular types of majors. However, the work of Lisa Kahn (2010), Paul Oyer (2006) and others indicates that market conditions at the time of labor market entry have long-term effects on earnings prospects. Consequently, one might question whether labor market conditions early in college can be excluded from wage equations even several years after graduation. Perhaps parental occupation or college major could be used if parental earnings, assets, highest degree level, and ability measure are controlled for.

A natural approach would be to use a regression discontinuity (RD) design in situations where certain majors have a test score cut-off or enrollment cap for entry. Some US universities use GPA cut-offs as a way to ration access to some programs. The opportunity for RD approaches is greater in countries that admit students to particular programs and particular colleges based primarily on test scores.[24]However, because students are in fact making a sequence of decisions about whether to remain in school and about which field to choose (particularly in the US context), there is an important distinction between ex ante and ex post returns to schooling decisions. Both of these returns to school choice depend on when the choice is made. The ex post return parameter to graduating in engineering identified by a regression discontinuity design based on grades or test scores at the time of college entry is different from the corresponding parameter identified using a grade cutoff after the second year of college.

Another selection issue is selection into the labor force. Employment rates differ substantially across majors both because of differences in labor force attachment and differences in the unemployment probability. Insofar as this problem is restricted to women, one "solution" is to estimate wage equations for men only, as is standard in the returns

[23]See also Willis and Rosen (1979).

[24]Depending on the institutional details, one must address selection in the decision to particular universities and programs. See Carvalho and Magnac (2010) for an example of such a study of college and major choice for Brazil. With additional data, their analysis could be extended to labor market outcomes. Bertrand et al (2010) apply an RD strategy based on admissions cut-off scores to engineering colleges in India that vary across castes, although large sampling errors limit what they can conclude about earnings effects.

to schooling literature more generally. However, this sidesteps many interesting and important questions regarding gender differences in choice of, and returns to, college major, and it does not address differential unemployment risk. Hamermesh and Donald (2008) is the only paper we reviewed that explicitly corrects for selection into employment.[25] They find that accounting for selection into the labor force reduces earnings differentials across major by 10-20% . Furthermore, this bias is most important for education majors, which is not surprising given the much larger proportion of female education majors and the fact that Hamermesh and Donald pool men and women. This points to the importance of an integrated approach where both employment and earnings prospects play a role in choosing major and employment selection is controlled for.

6.3 Estimates of the return to major

As we already noted, most studies use OLS with control variables rather than addressing endogenous selection into college major with an IV strategy or a selection model with exclusion restrictions. The control variables range from simple demographics, a small set of family background measures, and perhaps test scores, to detailed high school transcript information. The case for controlling for high school grades, tests, and courses prior to college seems clear. However, omitted variables bias in estimates of the effects of these variables could spill over into bias in estimates of returns to major. Some studies control for college transcript information and college quality measures. Some also include performance measures in college, arguing that insofar as ability differs by major, the returns to college major may be conflated with returns to ability. Whether one includes college variables, such as "semesters of math," depends on whether one wishes to measure the total effect of a particular college major (including human capital accumulation in the form of coursework and grades) or the effect of the title of the degree, net of substantive skill differences. Hamermesh and Donald (2008) show that there is a substantial return to upper-level math and science credits and grades holding eleven major categories constant. (Explaining differences in the returns to majors with differences in course content and grades, as opposed to the credential effect of the field of degree, is an interesting challenge for research.) Controlling for occupation is hard to defend other than as part of a strategy to identify why majors pay differently.

Table 2 presents certain estimates from selected papers. Aside from methodological concerns and differences in control variables, caution is urged in comparing results across studies for two main reasons. First, the level of aggregation of majors differs widely from study to study, ranging from four to over ten. Second, the time periods in question vary somewhat, although many use the same data set (NLS72). The estimates shown are the coefficients on major dummies in OLS regressions with the specified control variables.[26] Coefficients for men and women are presented separately when available. For ease of comparison, coefficients have been re-calculated, when necessary, relative to the education major; significance levels are not reported unless education was the original excluded category in the study.

[25] They also model survey non-response, but find that this has a negligible effect on results.

[26] The exception is Berger (1988), who runs separate regressions for each major category. The estimates shown are the premiums (over education) in predicted log wages.

As has been well-documented, the return to skill has increased substantially over the past few decades, particularly in the 1980s. Grogger and Eide (1995) decompose this into a change in returns and a change in the composition of majors.[27] In particular, they find that the return to math ability increased substantially for women, while a trend toward more technical subjects accounted for much of the increase for men. The coefficients presented in the table are the coefficients on the major; however the authors also include major interacted with experience and experience-squared, whose coefficients are not shown.[28]

Overall, however, the relative results have remained remarkably consistent over time. In particular, engineering consistently commands a high premium (around 0.40 relative to education), usually followed by business and science. Humanities, social sciences, and education are further behind. Controlling for pre-college test scores and grades reduces earnings differentials substantially. Hamermesh and Donald (2008), for example, report a standard deviation of 0.305 points of log earnings differences across twelve majors; this is halved when GPA, upper division math courses, upper division math grades, annual hours worked and a few additional controls are included. Controlling for selection into the labor force and non-response bias reduces the standard deviation somewhat further, to 0.139.

There has been speculation that the difference in returns is due to difference in math ability. Paglin and Rufolo (1990), for example, explain 82% of the variance across college majors in entry-level wages on the basis of GRE-math scores.[29] Moreover, Grogger and Eide show that the return to math ability has increased over time: the effect of a one standard deviation increase in math ability grew from a 2% increase in wages in 1978 to 5% in 1986, and from 3% to 7.5% for women.

There are also differences across majors in hours worked. Hamermesh and Donald (2008) find that controlling for hours worked and selection into the workforce reduces earnings differentials.

One might also be interested in the effects of college major on occupation, and in particular, the extent to which human capital is major-specific. Robst (2007) shows that students who are employed in a field unrelated to their field of study suffer a wage penalty, suggesting that this is at least partially the case; however, the wage penalty varies by field. The most specific fields, such as engineering, suffer harsher penalties than fields which develop more general skills, such as liberal arts. Malamud (2011), however, shows that the British system that requires students to specialize early leads to greater field-switching upon labor market entry than the Scottish system, which allows later specialization. This suggests that delaying specialization to learn about one's preferences and comparative advantage

[27] These changes in composition are documented in Section 5.

[28] Several other studies address the contribution of major choice to the gender gap among college graduates, including the recent study by Black et al (2008), who provide additional references. Black et al use the National Survey of College Graduates, which provides large sample sizes, highly disaggregated major categories and a measure of full-time experience. Results differ somewhat across race/ethnic groups, but for whites the unexplained gender gap is -0.184 when highest degree, highest degree major field, and age are controlled for, compared to a total gap of -0.339. For workers with high labor force attachment, the unexplained gap is -0.086 and the total gap -0.297. Black et al (2006) provide a similar analysis of the contribution of differences in detailed major to race differentials in the wages of highly educated men. Among individuals who speak English at home, differences in education and language proficiency account for essentially all of the small gap between white men and Hispanic men but only a quarter of the much larger gap between white men and black men. The reduction in the gap is about the same for the two groups. However, they account for the entire gap for black men born outside the South with parents who have some education.

[29] See also Weinberger (1999). Some of this reflects the effect of major on math scores. In the ACS data we explain about 58% of the variance in the major-specific returns with SAT math and verbal scores.

may outweigh any loss in field-specific skills.

Not much is known about the effect of college major on opportunities at the graduate level. Black et al (2003) provide estimates of wage gaps by undergraduate major relative to economics for those who obtain an MBA or a graduate law degree. With the exception of chemical engineers who obtain an MBA, economists earn more than their counterparts from the other most common pre-MBA or pre-law majors. These results illustrate the importance of considering the options that an earlier education choice offers.

6.4 Descriptive evidence on the returns to college major from the ACS

In this subsection we provide additional evidence on the relative returns to college major using the ACS. The ACS is an extremely large data set, which permits us to examine detailed major categories with and without 5-digit occupation controls. Unfortunately, the data set lacks test scores and family background measures. Table 3 provides OLS estimates of major coefficients separately for men and women, with and without occupation controls. These regressions include dummy variables for advanced degrees, a cubic in potential experience, and race/ethnicity as controls. Only the 23 most popular of the 171 major categories included in the regression are reported in the table. These twenty-three account for just over half of the college-plus sample. The omitted category is General Education. Near the bottom of the table we also report the sample weighted standard deviation of the major coefficients adjusted for sampling error.[30] The differences across majors are large. Consistent with results in the literature, the estimates show that engineers have the highest returns and education majors the lowest. This is true even after controlling for occupation. In most fields the size of the premium over a general education degree is higher for men than it is for women. For men (women), a shift from a college major that is one standard deviation below the mean to one that is one standard deviation above the mean is associated with an increase of 0.354 (0.292) log points in earnings. To compare, the high school–college differential is 0.577.[31] The standard deviation is 0.177 for men and 0.146 for women. Calculating the standard deviation for women using male weights (i.e., the fraction of men in each major) increases the standard deviation to 0.171. This indicates that much of the gender difference in the standard deviation is due to men selecting into majors at the high end of the earnings distribution (e.g, engineering), rather than gender differences in the dispersion of the major coefficients.

How have trends in major choice affected the gender gap holding the major coefficients for men and women constant? Figure 2 takes the major coefficients from Table 3 and calculates their weighted average by age and gender. (General education is the omitted category.) For individuals who graduated from college in the early 1970s, the weighted average of the female coefficients using the male major weights is about 0.1 above the value using the female weights. The gap narrows for the later cohorts but does not move very much. The gap at the beginning of the period is even wider using

[30] The sample is restricted to individuals between 23 and 59 years of age who worked more than 34 hours a week for more than 40 weeks in the previous year and who had a college or advanced degree. All majors are included, not just those relating to the coefficients presented in the table. We report the full set of coefficients in Supplementary Table 3. See Carnivale et al (2011) for a descriptive analysis of the returns to college major using the ACS.

[31] This is the coefficient on a college degree dummy in a regression that excludes the college major indicators, includes education level indicators and the other controls used in Table 3, and is estimated on a sample that includes high school graduates and above.

the male weights, but it narrows during the 70s. These results indicate that over most of the period, the differences in college major choice account for about a third of the gender gap in log wages.

The figure also show how changes in the composition of majors have affected the average return to a college degree. For both men and women, the curves show an increase of about 0.05 in the weighted average of the returns between the early 1970s and about 1986. This is due to a shift toward more lucrative majors, particularly engineering and business, with a peak in the mid-80s. The figure shows a decline of about 0.02 after that. These shifts in the distribution of majors are illustrated in Supplementary Figure 1.

We have also estimated, but do not present, a set of regressions using major characteristics from the B&B and the demographic and educational variables used above. Including 51 B&B major dummies yields an R^2 of 0.261, an increase of 0.06 over the R^2 of a model with only highest-degree indicators. Replacing the major dummies with major-specific SAT scores[32] (math and verbal) or including counts of the average number of courses in 8 different disciplines (math, business, etc) leads to R^2 values of 0.234 and 0.241, respectively. SAT scores account for 54.7% of the variation in the return to college major across the B&B categories (weighted by frequency in the ACS sample). We obtain 62.5% of the variation using GRE math, verbal, and writing scores for graduate school applicants in a smaller number of major categories.

What is the role of occupation? For men, the standard deviation of the college major coefficients is 0.177 when occupation is excluded and 0.098 when it is controlled for. Thus a substantial part of the difference in pay across majors is related to the occupations that they lead to. Part of this wage difference represents compensating differentials for nonpecuniary factors, of course.

In summary, wages vary greatly across college major. The variation is large enough for the tendency for men to choose high paying majors to be an important factor in the gender gap in wages. A substantial part of the differences in returns is almost certainly due to differences in the market value of tasks that require the specific knowledge and skills particular majors develop. However, pre-college differences in skill and ability, as captured by the SAT scores, and compensating differentials for nonpecuniary attributes also play a role. Much remains to be learned about why majors pay so differently.

6.5 College major and occupational choice

In this subsection we discuss the empirical link between college major and occupation. For some majors, such as nursing or accounting or engineering, the path is clear. For others it is not. Supplementary Table 4 illustrates the proportion of graduates in various majors employed in the three most common 5-digit occupations for that major. About 70% of accounting or auditing majors aged 25-29 are employed in the top three occupations. Even among those aged 55-59, the figure is 51%. For nursing, the comparable figures are 90 and 79%. Other majors show a similar trend. Many majors, however, have high occupational dispersion at all ages: history, psychology, business, and communications, for

[32]SAT Math and Verbal scores are predicted using the combined SAT Math and Verbal scores when the math and verbal scores are missing.

example. Overall, the table suggests a "fanning-out" of occupations as individuals are promoted or switch occupations as their careers progress.

Figure 3 provides further suggestive evidence for this trend. This figure reports the probability density (across majors) of the fraction of people in a major who are in one of the top ten occupations for that major. The peak of the densities occurs when the fraction is about 0.43. However, the heavy right tail and lower peak for those aged 25 to 34 indicates that occupational concentration is substantially larger for young workers.

The flip side of this analysis is the proportion of workers in an occupation accounted for by the three most common majors in that occupation. In Supplementary Table 5 we report this statistic for a few occupations by stage of career. Not surprisingly, almost all registered nurses major in either nursing, biology, or psychology regardless of stage of career. The values for marketing and sales managers and for accountants and auditors are similar, though less extreme. Post-secondary school teachers and marketing and sales are examples of occupations that draw on a wider array of majors. With a large panel data set, one could go much further in examining the link between major and career path. We have already noted that much the variance in the returns to majors is associated with differences in the pay of the occupations that they lead to. The clear message is that specificity of skill is important.

7 Conclusion

The demand for and return to *types* of human capital investments is an exciting research area that is highly relevant for both education and labor market policy, as education and occupation choices differ greatly in the monetary and non-monetary rewards that they bring.

Rather than summarizing the paper, we close by restating the questions for empirical research. The recent theoretical work on education and occupation choice that we have synthesized here emphasizes several key factors: the sequential nature of the decision-making process; the partial irreversibility of some decisions because of the heterogeneity of human capital and the cumulative nature of human capital investment in many fields; the importance of innate ability, preferences, and pre-high school learning in shaping the feasibility and non-pecuniary costs and benefits of particular education and occupation paths; and the essential role of uncertainty about ability, preferences, knowledge accumulation, and wage rates at each stage. The theory has some very clear implications for empirical research. The first concerns the determination of preferences for schooling in general and types of schooling in particular, which we touch on briefly in Section 3. The second concerns the determination of the ability to do math, excel at science, write well, etc. The third concerns estimation of the knowledge production function. How should skills and knowledge be classified? How can they be measured using observables such as course content, grades, and tests? How much does learning in school or on the job depend on the program or the occupation? The fourth is how grades, knowledge and ability determine promotion and admission to colleges, a subject on which there is a substantial literature that we have not touched on here. The fifth is how best to model the agent's information set and learning about ability and preferences. The sixth is estimation of the ex post payoff to knowledge, ability, and degree by level and field. We summarized a number of papers that attempt this for college education and a smaller number that look at high school curriculum. Much progress has been made, but there is a long way to go on the road to credible measures of the payoff to fields of study. We suggest several approaches that might prove fruitful.

The overriding question is the choice of education and occupation at each stage in the life course and the consequences of those choices. For example, how does the current utility and expected future utility of spending the first period of college in a pre-engineering curriculum versus a fine arts curriculum depend on preferences, ability, and the stock of human capital at the start of college? The large earnings gaps across fields that attract students admitted to the same universities with similar grades and test scores strongly suggests that compensating differentials are of critical importance.

References

[1] Altonji JG. 1993. The demand for and return to education when education outcomes are uncertain. *J. Labor Econ.* 11(1):48–83

[2] Altonji JG. 1995. The effects of high school curriculum on education and labor market outcomes. *J. Hum. Resour.* 30(3):409–38

[3] Altonji JG, Pierret CR. 2001. Employer learning and statistical discrimination. *Q. J. Econ.* 116(1):313-350

[4] Arcidiacono P. 2004. Ability sorting and the returns to college major. *J. Econom.* 121(1-2):343–75

[5] Arcidiacono P, Hotz VJ, Kang S. Modeling college major choices using elicited measures of expectations and counterfactuals. *IZA Disc. Pap. 4738*

[6] Berger MC. 1988. Predicted future earnings and choice of college major. *Ind. & Labor Relat. Rev.* 41(3):418–29

[7] Bertrand M, Hanna R, Mullainathan S. 2010. Affirmative action in education: evidence from engineering college admissions. *J. Pub. Econ.* 94(1-2):16–29

[8] Betts JR. 1996. What do students know about wages? Evidence from a survey of undergraduates. *J. Hum. Resour.* 31(1):27–56

[9] Black DA, Haviland AM, Sanders SG, Taylor LJ. 2006. Why do minority men earn less? A study of wage differentials among the highly educated. *Rev. Econ. Stat.* 88(1):300-13

[10] Black DA, Haviland AM, Sanders SG, Taylor LJ. 2008. Gender wage disparities among the highly educated. *J. Hum. Resour.* 43(3):630–59

[11] Black DA, Sanders S, Taylor L. The economic reward for studying economics. *Econ. Inquiry* 41(3):365–77

[12] Cameron SV, Heckman JJ. 1998. Life cycling schooling and dynamic selection bias: models and evidence for five cohorts of American males. *J. Pol. Econ.* 106(2):262–333

[13] Canes BJ, Rosen HS. 1995. Following in her footsteps? Faculty gender composition and women's choice of college majors. *Ind. & Labor Relat. Rev.* 48(3):486–504

[14] Card DE. 1995. Using geographic variation in college proximity to estimate the return to schooling. In *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*, ed. Christofides LN, Grant EK, Swidinsky R, 201–22. Toronto: University of Toronto Press

[15] Card DE. 1999. The causal effect of education on earnings. In *Handbook of Labor Economics*, ed. Ashenfelter O, Card DE, 3A. Amsterdam: Elsevier

[16] Carnevale AP, Strohl J, Melton M. 2011. What's it worth? The economics value of college majors. Center on Education and the Workforce, Georgetown University

[17] Carvalho JR, Magnac T. 2010. College choice and entry exams. Work. Pap., CAEN, Universidade Federal do Ceara, Fortaleza, Brazil and Toulouse Sch. Econ., Toulouse, France

[18] Christiansen C, Joensen JS, Nielsen HS. 2007. The risk-return trade-off in human capital investment. *Labour Econ.* 14(6):971–86

[19] Cunha F, Heckman JJ, Schennach SM. 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica.* 78(3):883–931

[20] Currie J. 2009. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. *J. Econ. Lit.* 47(1):87–122

[21] Dickson L. 2010. Race and gender differences in college major choice. *Annals Amer. Acad. Pol. & Soc. Sci.* 627(1):108–24

[22] Eckstein Z, Wolpin KI. 1989. The specification and estimation of dynamic stochastic discrete choice models: a survey. *J. Hum. Resour.* 24(4):562–98

[23] Eide E, Waehrer G. 1998. The role of the option value of college attendance in college major choice. *Econ. Educ. Rev.* 17(1):73–82

[24] Farber HS, Gibbons R. 1996. Learning and wage dynamics. *Q. J. Econ.* 111(4):1007–47

[25] Finnie R, Frenette M. 2003. Earning differences by major field of study: Evidence from three cohorts of recent Canadian graduates. *Econ. Educ. Rev.* 22(2):179–92

[26] Freeman R. 1976. A cobweb model of the supply and starting salary of new engineers. *Ind. & Labor Relat. Re*v. 29(2):236–48

[27] Gardner DP. 1983. A nation at risk: the imperative for educational reform. Washington, DC: National Commission on Excellence in Education

[28] Goldin C, Katz LF, Kuziemko I. 2006. The homecoming of American college women: The reversal of the gender gap in college. *J. Econ. Perspect.* 20(4):133–56

[29] Goodman J. 2009. The labor of division: returns to compulsory math coursework. Work. Pap., Harvard Kennedy School

[30] Grogger J, Eide E. 1995. Changes in college skills and the rise in the college wage premium. *J. Hum. Resour.* 30(2):280–310

[31] Hamermesh DS, Donald SG. 2008. The effect of college curriculum on earnings: an affinity identifier for non-ignorable non-response bias. *J. Econom.* 144(2):479–91

[32] Heckman JJ, Honore BE. 1990. The empirical content of the Roy model. *Econom.* 58(5):1121–49

[33] Heckman JJ, Lochner LJ, Todd PE. 2008. Earnings functions and rates of return. *J. Human Capital.* 2(1):1–31

[34] Heckman JJ, Masterov DV. 2007. The productivity argument for investing in young children. *Appl. Econ. Perspect. Pol.* 29(3):446–93

[35] Heckman JJ, Navarro S. 2007. Dynamic discrete choice and dynamic treatment effects. *J. Econom.* 136(2):341–96

[36] Heckman JJ, Stixrud J, Urzua S. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *NBER Work. Pap. 12006*

[37] Imbens GW, Angrist JD. 1994. Identification and estimation of local average treatment effects. *Econom.* 62(2):467-75

[38] Ingels SJ, Dalton BW, LoGerfo L. 2008. Trends among high school seniors, 1972-2004 (NCES 2008-320). Washington, DC: US Department of Education, National Center for Education Statistics

[39] James E, Alsalam N, Conaty JC, To D. 1989. College quality and future earnings: where should you send your child to college? *Am. Econ. Rev.* 79(2):247–52

[40] Joensen JS, Nielsen HS. 2009. Is there a causal effect of high school math on labor market outcomes? *J. Hum. Resour.* 44(1):171–98

[41] Kahn LB. 2010. The long-term labor market consequences of graduating from college in a bad economy. *Labour Econ.* 17(2):303–16

[42] Keane MP, Wolpin KI. 1997. The career decisions of young men. *J. Pol. Econ.* 105(3):473–522

[43] Lange F. 2007. The speed of employer learning. *J. Labor Eco*n. 25(1):1–35

[44] Levine PB, Zimmerman DJ. 1995. The benefit of additional high-school math and science classes for young men and women. *J. Bus. & Econ. Stat.* 13(2):137–49

[45] Loury LD, Garman D. 1995. College selectivity and earnings. *J. Labor Econ.* 13(2):289–308

Malamud O. 2011. Discovering one's talent: learning from academic specialization. *Ind. & Labor Relat. Rev.* 64(2):375–405

Mane F. 1999. Trends in the payoff to academic and occupation-specific skills: the short and medium run returns to academic and vocational high school courses for non-college-bound students. *Econ. Educ. Rev.* 18(4):417–37

[46] Manski CF. 1989. Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Econ. Educ. Rev.* 8(4):305–12

[47] Manski CF. 1993. Adolescent econometricians: how do youth infer the returns to schooling? In *Studies of Supply and Demand in Higher Education,* 43–60. NBER

Manski CF. 2004. Measuring expectations. *Econometrica*. 72(5):1329–76

[48] Meghir C, Rivkin SG. 2011. Econometric methods for research in education. In *Handbook of the Economics of Education*, ed. Hanushek E, Machin S, Woessmann L, 3(3):1–87. Elsevier

[49] Montmarquette C, Cannings K, Mahseredjian S. 2002. How do young people choose college majors? *Econ. Educ. Rev.* 21(6):543–56

[50] Neal D. 1999. The complexity of job mobility among young men. *J. Labor Econ.* 17(2): 237–61

Nord C, Roey S, Perkins R, Lyons M, Lemanski N, Tamir Y, Brown J, Schuknecht J, Herrold K. 2011. America's high school graduates: results from the 2009 NAEP high school transcript study (NCES 2011-462). Washington, DC: US Department of Education, National Center for Education Statistics

Ost B. 2010. The role of peers and grades in determining major persistence in the sciences. *Econ. Educ. Rev.* 29(6):923–34

Oyer P. 2006. Initial labor market conditions and long-term outcomes for economists. *J. Econ. Perspect.* 20(3):143–160

[51] Paglin M, Rufolo AM. 1990. Heterogeneous human capital, occupation choice, and male-female earnings differences. *J. Labor Econ.* 8(1):123–44

[52] Papageorgiou T. 2010. *Learning your comparative advantages.* PhD thesis. Yale Univ.

[53] Quandt RE. 1972. A new approach to estimating switching regressions. *J. Amer. Stat. Assoc.* 67(338):306–10

Robst J. 2007. Education and job match: the relatedness of college major and work. *Econ. Educ. Rev.* 26(4):397–407

Rose H, Betts JR. 2004. The effect of high school courses on earnings. *Rev. Econ. Stat.* 86(2):397-513

Rumberger RW, Daymont TN. 1984. The economic value of academic and vocational training acquired in high school. In *Youth and the Labor Market,* ed. M Borus, 157–92. Kalamazoo: WE Upjohn Institute for Employment Research

[54] Rumberger RW, Thomas SL. 1993. The economic returns to college major, quality and performance: a multilevel analysis of recent graduates. *Econ. Educ. Rev.* 12(1):1–19

Siow A. 1984. Occupational choice under uncertainty. *Econometrica* 52(3):631–45

[55] Stange KM. 2011. An empirical investigation of the option value of college enrollment. *Am. Econ. J.* In press

Todd PE, Wolpin KI. 2003. On the specification and estimation of the production function for cognitive achievement. *Econ. J.* 113(485):F3–F33

[56] Violante G, Gallipoli G, Meghir C. 2005. Education decisions, equilibrium policies and wages dispersion. 2005 Meeting Papers 522, Soc. Econ. Dyn.

[57] Weinberger CJ. 1999. Mathematical college majors and the gender gap in wages. *Ind. Relat.* 38(3):407–13

Weisbrod BA. 1962. Education and investment in human capital. *J. Pol. Econ.* 70(5):106–23

Willis RJ, Rosen S. 1979. Education and self-selection. *J. Pol. Econ.* 87(5):S7-36

[58] Zafar B. 2009. How do college students form expectations? *Staff Report 378,* Federal Reserve Bank, New York, NY

Zafar B. 2009. College major choice and the gender gap. *Staff Report 364,* Federal Reserve Bank, New York, NY

[59] Zietz J, Joshi P. 2005. Academic choice behavior of high school students: economic rationale and empirical evidence. *Econ. Educ. Rev.* 24(3):297–308

Table 1: College major characteristics

Major	Share	% female	Math credits	Science credits	SAT M	SAT V	Mean wages	p10 wages	p90 wages	p90 wages[a]	% MA	% PhD	% PD[b]
Mathematics	1.5	0.4	27.7	13.7	592	538	42.71	15.91	73.53	63.73	33.9	9.1	4.4
Mech. Eng.	1.5	0.1	12.3	66.1	613	566	44.12	21.03	66.18	58.82	31.4	4.1	1.9
Elec. Eng.	1.9	0.1	12.3	66.5	606	571	46.45	21.79	70.08	63.24	33.8	5.2	2.3
Chemistry	1.2	0.4	12.0	57.2	604	597	45.37	13.73	87.83	52.29	22.1	20.8	16.0
Comp. & IT	2.9	0.3	11.4	11.1	582	556	37.99	17.16	58.82	56.37	22.1	1.7	1.2
Bio. Sci.	4.5	0.5	5.8	56.1	577	575	41.27	12.75	85.29	46.08	20.2	11.3	22.0
Economics	2.1	0.3	5.6	6.3	597	573	50.60	15.25	111.76	78.43	26.6	3.4	9.0
Finance	1.9	0.3	4.9	5.1	563	534	42.34	15.69	77.30	65.36	20.4	0.7	3.7
Accounting	4.2	0.5	3.9	3.7	571	534	39.46	15.69	67.23	60.33	18.6	0.6	3.9
Marketing	2.3	0.5	4.0	5.1	526	514	34.12	13.94	58.82	57.29	11.8	0.2	1.3
Educ. (other)[b]	12.3	0.8	3.1	7.1	488	496	24.54	12.75	38.64	32.68	39.2	2.1	2.6
Bus. Mgmt & Admin.	6.8	0.4	3.7	4.3	522	510	33.42	13.73	55.56	51.96	16.5	0.6	2.1
Agr. & Agr. Sci.	1.1	0.3	3.6	21.1	546	549	29.04	11.13	49.02	45.75	15.8	5.5	5.2
Psychology	4.8	0.7	2.7	8.2	530	540	29.15	12.25	49.02	41.18	30.8	6.9	6.0
Music & Drama	1.3	0.6	2.3	6.5	539	575	25.56	10.59	42.58	37.91	27.7	4.1	3.2
Fit. & Nutr.	1.0	0.6	2.9	19.9	520	518	25.20	11.27	40.96	38.15	19.4	2.5	3.3
Comm.	2.9	0.6	1.6	5.8	512	537	29.20	12.50	49.41	49.02	14.2	1.1	2.8
Soc. Sci. (oth.)	3.2	0.6	2.0	7.2	514	526	28.79	12.66	47.06	41.67	25.7	3.7	4.9
Letters	5.2	0.6	2.3	6.1	540	592	30.63	12.35	51.93	46.35	25.8	3.9	7.4
Poli. Sci.	2.4	0.4	2.2	6.2	542	571	42.09	14.71	75.41	57.25	21.1	3.6	21.6
History	2.3	0.4	2.2	6.0	558	595	36.12	12.65	64.17	49.02	27.9	4.9	13.5
Art & Art Hist.	1.5	0.7	2.0	6.9	555	592	27.09	11.09	44.71	43.57	21.2	1.2	2.5
Social Work & HR	1.5	0.8	1.5	5.2	460	487	25.41	12.50	40.20	38.24	34.0	0.9	2.1
Phil. & Rel.	1.3	0.3	1.8	5.7	567	595	27.52	10.23	47.06	39.22	28.3	8.4	9.1
Nursing	3.6	0.9	1.0	13.0	488	497	33.51	18.82	49.02	46.08	18.2	1.2	3.3

Source: NCES B&B 1993/2003 for course counts and SAT scores. Majors with a share less than 1%, or missing SAT scores, are not shown. Shares and hourly wage data are from the 2009 ACS, top- and bottom-coded at 5 and 400 USD per hour, respectively.

Sample selection: Wage observations are included if the individual has at least a bachelor's degree, is working >34 hours per week and >40 weeks per year, and is 23-59 years old.

[a] Excluding advanced degrees.

[b] Professional degree.

[c] Includes library science.

Table 2: Summary of selected empirical studies on the returns to college major

Study	Data & Method.	Controls	Outcome variable	Majors	Outcomes	
					Men	Women
Daymont & Andrisani (1984)	NLS72; OLS	work experience; preferences; father's occupation; marital status; highest degree; type and size of community; race; region	log hourly earnings	Business	.16	.17
				Engineering	.33	.19
				Math & sci.	.18	.25
				Soci. sci.	.06	.00
				Humanities	.04	-.05
					1 yr exp.	12 yrs exp.
Berger (1988)	NLSM; conditional logit and selection correction in the wage equation	log of yrs of exp, grad yr−1900, and their interaction; IQ score and Knowledge of World of Work score; US male unemp. rate, race, health status, married, smsa, South, enrolled in school, log of annual wks worked; selection correction	log hourly wages for 1974 male college graduates in 1986 USD, corrected for selection bias	Business	.3473	.1332
				Engineering	.4130	.3637
				Science	.1237	.2226
				Lib. Arts	.1045	-.0719
					(1)[a]	(2)[b]
James et al (1989)	NLS72 (men only); WLS	family background, SAT, hs rank, acad. track, math credits, Catholic hs, various college-level variables, labor market variables; occupation and industry dummies in (2)	1985 log annual earnings	Business	.262	.1452
				Engineering	.47	.451
				Math & sci.	.1966	.1198
				Soc. sci.	.2378	.1829
				Humanities	.059	-.026
					Men	Women
Altonji (1993)	NLS72; OLS	SAT, hs grades, self-assessment of college ability; various education interactions; exp and exp-squared; gender, race, family background; hs curriculum; post-grad deg.	log of real hourly wage; coefficients on terminal majors presented not all presented)	Bus.	.1796	.2422
				Engineering	.4119	.2836
				Phys. sci.	.2432	.0743
				Math, CS	.3887	.2325
				Life sci.	.1231	.2057
				Soc. sci.	.0975	.0117
				Humanities	.0646	-.0048

[a] Without occupation or industry dummies.
[b] With occupation and industry dummies.

Study	Data & Method.	Controls	Outcome variable	Majors	Outcomes	
					Men	Women
Rumberger and Thomas (1993)	Survey of Recent College graduates (1987); hierarchical linear modeling, OLS	family background, race, GPA, private college, college selectivity, labor market variables	1987 log annual earnings	Business	.1839	.2521
				Engineering	.3913	.5135
				Sci. & math	.2552	.3036
				Health	.2984	.4433
				Soc. sci.	.0765	.1255
					Whites	Blacks
Loury & Garman (1995)	NLS72; OLS	college selectivity, years of education, parental income, GPA, SAT, weeks worked, rural dummy	ln weekly earnings, 1979 or 1986	Business	.212	.262
				Eng. & Sci.	.374	.549
				Soc. sci.	.094	-.097
				Humanities	-.164	.030
					Men	Women
Grogger & Eide (1995)	NLS72, HSB; GLS	std. tests, hs grades; family income; experience; race; educ. attainment (not shown: with occupation controls), full-time workers only	log hourly wage 1977-79, 86	Business	.155	.1057
				Engineering	.2797	.0659
				Science	.0607	.2152
				Soc. sci.	.0215	.0168
					Men	Women
Finnie & Frenette (2003)	National Graduate Survey (Stats Can); OLS	pre-programme educ. level, age, post-grad exp., self-employment status, marriage/children, region, language; adv. degree-holders excluded	log annual earnings 1995	Commerce	.06	.04
				Eng & CS	.22	.16
				Math & Sci.	.31	.09
				Health	.06	.19
				Soc. sci.	-.12	-.04
				Arts & Hum.	-.15	-.06
Hamermesh & Donald (2008)	graduates of UT-Austin, 1980-2000[c]; double selection correction (into employment and survey non-response)	high-school background, college achievement, demographic, post-grad deg., hours worked, quadratic in propensity scores for working & survey response	log earnings (selected majors only presented)	Bus. (hard)	.489***	
				Bus. (soft)	.378***	
				Engineering	.316***	
				Nat. sci.	.265***	
				Soc. sci.	.279***	
				Humanities	.086	

*** $p<0.01$, ** $p<0.05$, * $p<0.10$

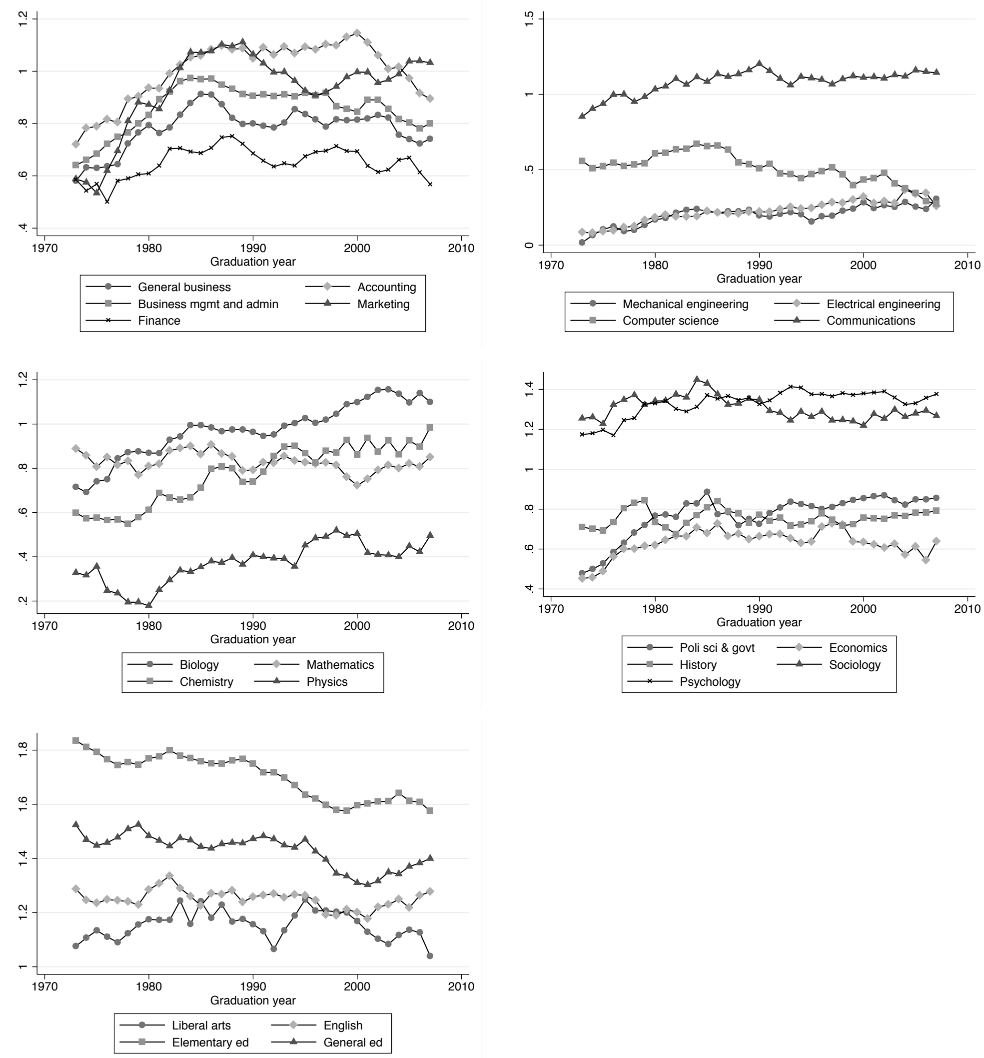
[c] Selected years.

Table 3: Effects of college major on log wages by gender, with and without occupation controls

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
Communications	0.202***	0.207***	0.063***	0.058**
Computer Science	0.441***	0.531***	0.161***	0.242***
Elementary Education	-0.024*	-0.009	-0.015	0.009
Electrical Engineering	0.556***	0.561***	0.258***	0.293***
Mechanical Engineering	0.554***	0.524***	0.265***	0.264***
English Language And Literature	0.107***	0.152***	0.026*	0.063***
Liberal Arts	0.073***	0.154***	0.021	0.055*
Biology	0.196***	0.302***	0.068***	0.114***
Mathematics	0.288***	0.426***	0.143***	0.224***
Chemistry	0.250***	0.366***	0.101***	0.193***
Criminal Justice And Fire Protection	0.076***	0.226***	-0.013	0.076***
Economics	0.400***	0.517***	0.224***	0.275***
Anthropology And Archeology	0.069**	0.135***	-0.001	0.053
Political Science And Government	0.246***	0.327***	0.112***	0.158***
Sociology	0.077***	0.165***	0.012	0.075***
Fine Arts	-0.021	0.017	-0.067**	-0.035
Nursing	0.391***	0.408***	0.172***	0.243***
General Business	0.218***	0.339***	0.077***	0.142***
Accounting	0.310***	0.431***	0.143***	0.199***
Business Management And Administration	0.199***	0.292***	0.054***	0.104***
Marketing And Marketing Research	0.256***	0.356***	0.089***	0.150***
Finance	0.342***	0.518***	0.151***	0.243***
History	0.105***	0.167***	0.033*	0.064***
R^2	0.200	0.217	0.330	0.337
SD of major coefficients	0.146	0.177	0.074	0.098
N	125794	140706	124858	139493

Notes: *** p<0.01, ** p<0.05, * p<0.10
All specifications include dummy variables for highest level of education attained, a cubic in potential experience, and race dummies. Bachelor's degrees are 4-digit; only a selected sample of the 171 are shown. Wages are top- and bottom-coded at 5 and 400 USD per hour, respectively. General Education is the excluded category. Occupation controls are 5-digit. SD is calculated over all majors using ACS weights.
Sample selection: Observations are included if the individual has at least a bachelor's degree, is working >34 hours per week and >40 weeks per year, and is 23-59 years old.

Figure 1: Relative fraction female, by major



Note: Relative fraction female is calculated by dividing the fraction female in a particular major in a particular year by the fraction of female college graduates that year, then smoothed using a three-year moving average. Data are from the ACS.

Figure 2: Average of major coefficients by age

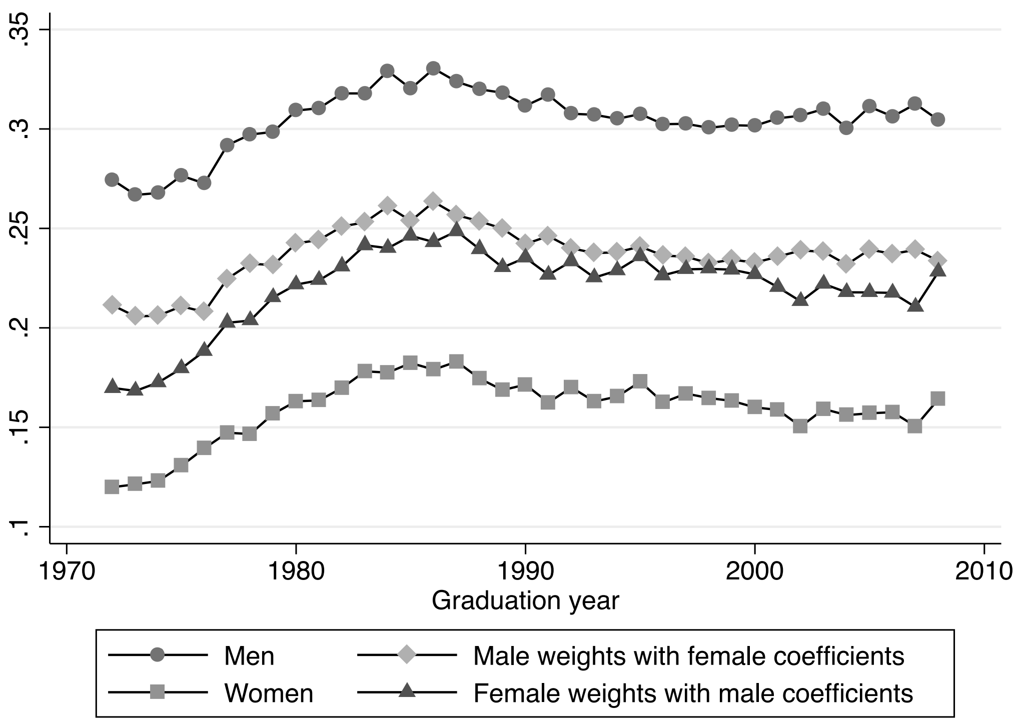
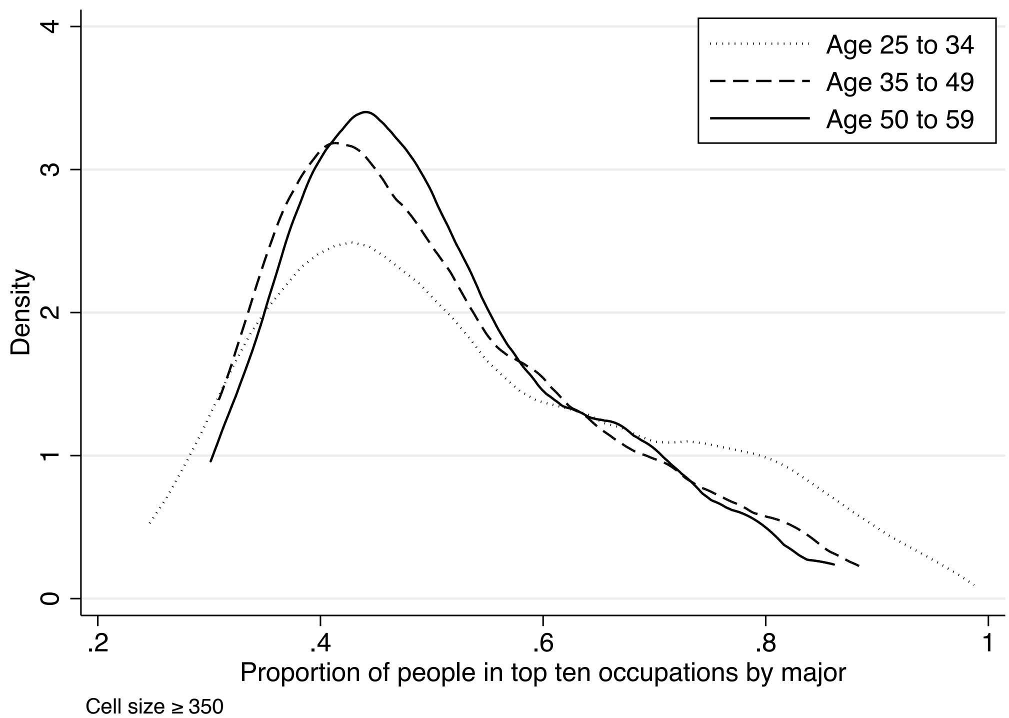


Figure 3: Occupational dispersion by age



Supplementary Tables and Figures

Supplementary Table 1: Trends in high school course-taking

	1990	2005
Credits earned, total	23.6	27.2
Core academic, total	13.7	16.0
Core science	2.8	3.5
Core math	3.2	3.9
Core social science	3.5	4.2
Core English	4.1	4.4
Percentage taking curricula that are:		
Rigorous	5	13
Midlevel	26	46
Standard	9	16
Below standard	60	25
Percentage taking STEM courses:		
Algebra II	53	76
Calculus	7	17
Advanced biology	28	45
Chemistry	45	70
Physics	24	39

From 2011 NAEP report. All differences are significant at 5%.

Supplementary Table 2: College major characteristics, no advanced degrees

Major	Share	% female	Math credits	Science credits	SAT M	SAT V	Mean wages	p10 wages	p90 wages
Mathematics	1.2	0.4	27.7	13.7	592	538	37.76	14.01	63.73
Mech. Eng.	1.4	0.1	12.3	66.1	613	566	40.43	19.61	58.82
Elec. Eng.	1.7	0.1	12.3	66.5	606	571	41.61	20.10	63.24
Comp. & IT	3.2	0.3	11.4	11.1	582	556	35.83	16.18	56.37
Bio. Sci.	3.2	0.6	5.8	56.1	577	575	27.26	10.98	46.08
Economics	2.0	0.3	5.6	6.3	597	573	43.15	13.90	78.43
Finance	2.2	0.4	4.9	5.1	563	534	38.21	14.71	65.36
Accounting	4.9	0.5	3.9	3.7	571	534	36.88	14.71	60.33
Marketing	3.0	0.5	4.0	5.1	526	514	32.90	13.73	57.29
Educ. (other)[b]	10.5	0.8	3.1	7.1	488	496	21.17	11.27	32.68
Bus. Mgmt & Admin.	8.3	0.4	3.7	4.3	522	510	31.56	13.24	51.96
Agr. & Agr. Sci.	1.2	0.3	3.6	21.1	546	549	26.51	10.46	45.75
Comm. Arts	1.3	0.7	3.0	6.7	529	527	25.00	11.40	41.18
Psychology	4.1	0.7	2.7	8.2	530	540	24.61	11.03	41.18
Music & Drama	1.3	0.6	2.3	6.5	539	575	22.57	9.80	37.91
Fit. & Nutr.	1.2	0.6	2.9	19.9	520	518	22.91	10.62	38.15
Comm.	3.6	0.6	1.6	5.8	512	537	28.17	12.25	49.02
Soc. Sci. (oth.)	3.1	0.6	2.0	7.2	514	526	25.56	11.76	41.67
Letters	4.9	0.6	2.3	6.1	540	592	27.14	11.62	46.35
Poli. Sci.	2.0	0.4	2.2	6.2	542	571	33.32	12.60	57.25
History	1.8	0.4	2.2	6.0	558	595	29.52	11.13	49.02
Art & Art Hist.	1.7	0.7	2.0	6.9	555	592	25.57	10.37	43.57
Social Work & HR	1.4	0.8	1.5	5.2	460	487	23.19	11.76	38.24
Phil. & Rel.	1.1	0.4	1.8	5.7	567	595	23.77	9.80	39.22
Journ.	1.2	0.6	1.1	5.5	496	533	29.94	12.20	50.98
Nursing	4.2	0.9	1.0	13.0	488	497	31.12	17.71	46.08

Source: NCES B&B 1993/2003 for course counts and SAT scores. Not all majors are shown; only those which appear in Table 1 are presented. Shares and hourly wage data are from the 2009 ACS, top- and bottom-coded at 5 and 400 USD per hour, respectively.

Sample selection: Wage observations are included if the individual has a bachelor's degree, is working >34 hours per week and >40 weeks per year, and is 23-59 years old. Individuals with an advanced degree are excluded.

[a] Includes library science.

Supplementary Table 3: Effects of college major on log wages by gender, with and without occupation controls – all control variables and major coefficients reported

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
PhD	0.290***	0.220***	0.296***	0.264***
Masters	0.204***	0.173***	0.181***	0.139***
Professional degree	0.441***	0.497***	0.254***	0.259***
Potential experience	0.083***	0.097***	0.066***	0.083***
Potential experience^2	-0.003***	-0.003***	-0.002***	-0.002***
Potential experience^3	0.000***	0.000***	0.000***	0.000***
Black, non-hispanic	-0.084***	-0.215***	-0.052***	-0.137***
Native American, non-hispanic	-0.163***	-0.171***	-0.132***	-0.139***
Asian, non-hispanic	-0.022**	-0.099***	-0.000	-0.076***
Pacific Islander, non-hispanic	-0.220**	-0.159	-0.114	-0.089
Mixed raced, non-hispanic	-0.044**	-0.085***	-0.038**	-0.059***
Any race hispanic	-0.088***	-0.205***	-0.037***	-0.130***
General Agriculture	-0.043	-0.059	-0.094*	-0.068
Agriculture Production And Management	0.011	0.047	-0.101*	0.034
Agricultural Economics	0.074	0.150**	-0.010	0.027
Animal Sciences	-0.079*	-0.078*	-0.074*	-0.060
Food Science	0.255***	0.236*	0.110	0.120
Plant Science And Agronomy	-0.027	0.026	-0.041	0.029
Soil Science	-0.292	0.127	-0.280	0.103
Miscellaneous Agriculture	0.053	0.154	-0.042	0.079
Environmental Science	0.095**	0.162***	-0.032	0.036
Forestry	0.250***	0.157***	0.162***	0.048
Natural Resources Management	0.047	0.112***	-0.038	0.000
Architecture	0.238***	0.272***	0.067*	0.086***
Area Ethnic And Civilization Studies	0.155***	0.247***	0.063*	0.122**
Communications	0.202***	0.207***	0.063***	0.058**
Journalism	0.174***	0.183***	0.029	0.033
Mass Media	0.112***	0.102***	0.013	-0.011
Advertising And Public Relations	0.181***	0.228***	0.018	0.057
Communication Technologies	0.188***	0.165***	0.039	0.042
Computer And Information Systems	0.295***	0.421***	0.072**	0.160***
Computer Programming And Data Processi	-0.052	0.229***	-0.091	0.043
Computer Science	0.441***	0.531***	0.161***	0.242***
Information Sciences	0.410***	0.421***	0.167***	0.164***
Computer Administration Management And	0.290***	0.357***	0.076	0.134**
Computer Networking And Telecommunicat	0.174**	0.258***	0.036	0.074

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
Cosmetology Services And Culinary Arts	-0.163*	-0.024	-0.003	0.165*
Educational Administration And Supervi	0.106**	0.069	0.072	-0.011
School Student Counseling	-0.006	0.065	0.018	0.122
Elementary Education	-0.024*	-0.009	-0.015	0.009
Mathematics Teacher Education	0.051*	0.006	0.020	0.015
Physical And Health Education Teaching	0.059**	-0.002	0.046*	-0.002
Early Childhood Education	-0.057***	-0.230	-0.015	-0.164
Science And Computer Teacher Education	-0.017	0.018	-0.015	0.053
Secondary Teacher Education	0.031	0.010	0.010	0.013
Special Needs Education	0.081***	0.105**	0.067***	0.102**
Social Science Or History Teacher Educ	0.013	-0.014	-0.001	-0.016
Teacher Education: Multiple Levels	-0.012	0.003	-0.014	0.021
Language And Drama Education	0.031	0.013	0.016	0.035
Art And Music Education	-0.004	-0.040	-0.007	0.000
Miscellaneous Education	0.043*	0.052	0.015	-0.011
General Engineering	0.417***	0.392***	0.170***	0.169***
Aerospace Engineering	0.616***	0.548***	0.274***	0.272***
Biological Engineering	0.224**	0.135*	0.053	0.025
Architectural Engineering	0.542***	0.357***	0.262**	0.171**
Biomedical Engineering	0.468***	0.472***	0.203**	0.171**
Chemical Engineering	0.526***	0.614***	0.252***	0.346***
Civil Engineering	0.406***	0.482***	0.138***	0.240***
Computer Engineering	0.562***	0.606***	0.227***	0.293***
Electrical Engineering	0.556***	0.561***	0.258***	0.293***
Engineering Mechanics Physics And Scie	0.715***	0.429***	0.379**	0.235***
Environmental Engineering	0.400***	0.530***	0.166*	0.267***
Geological And Geophysical Engineering	0.342	0.639***	0.063	0.385***
Industrial And Manufacturing Engineeri	0.483***	0.469***	0.221***	0.227***
Materials Engineering And Materials Sc	0.341**	0.429***	0.064	0.194***
Mechanical Engineering	0.554***	0.524***	0.265***	0.264***
Metallurgical Engineering	0.374*	0.452***	0.155	0.209***
Mining And Mineral Engineering	0.771*	0.412***	0.590	0.215**
Naval Architecture And Marine Engineer	0.530**	0.360***	0.246*	0.147
Nuclear Engineering	0.600**	0.651***	0.364	0.406***
Petroleum Engineering	0.682***	0.869***	0.332*	0.590***
Miscellaneous Engineering	0.260***	0.394***	0.129	0.206***
Engineering Technologies	0.141	0.345***	0.010	0.143***
Engineering And Industrial Management	0.298**	0.374***	0.111	0.135**
Electrical Engineering Technology	0.150	0.315***	-0.075	0.140***
Industrial Production Technologies	0.282***	0.261***	0.146*	0.087**
Mechanical Engineering Related Technol	0.481***	0.262***	0.186*	0.107**

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
Miscellaneous Engineering Technologies	0.176*	0.327***	0.041	0.121***
Linguistics And Comparative Language A	0.080*	0.172*	0.001	0.047
French German Latin And Other Common F	0.128***	0.164***	0.066***	0.070*
Other Foreign Languages	0.081	0.117*	0.017	0.047
Family And Consumer Sciences	0.020	0.203***	-0.002	0.102*
Court Reporting	0.289*	-0.013	0.067	-0.014
Pre-Law And Legal Studies	0.139***	0.234***	0.017	0.117*
English Language And Literature	0.107***	0.152***	0.026*	0.063***
Composition And Speech	0.090*	0.141**	0.002	0.090*
Liberal Arts	0.073***	0.154***	0.021	0.055*
Humanities	0.113*	0.113	0.044	0.004
Library Science	-0.046	0.110	-0.036	0.073
Biology	0.196***	0.302***	0.068***	0.114***
Biochemical Sciences	0.262***	0.308***	0.096**	0.111**
Botany	-0.012	0.008	-0.062	-0.007
Molecular Biology	0.196***	0.260***	0.080	0.112*
Ecology	0.050	0.068	-0.008	0.012
Genetics	0.196**	0.178	0.080	0.025
Microbiology	0.185***	0.228***	0.063	0.071
Pharmacology	0.387**	0.168	0.135	0.022
Physiology	0.157***	0.193***	0.016	0.077
Zoology	0.088	0.328***	-0.005	0.153***
Miscellaneous Biology	0.154**	0.189***	0.075	0.057
Mathematics	0.288***	0.426***	0.143***	0.224***
Applied Mathematics	0.537***	0.641***	0.286***	0.375***
Statistics And Decision Science	0.473***	0.523***	0.228***	0.206***
Military Technologies	0.670***	0.280	0.761***	0.099
Intercultural And International Studie	0.119**	0.192***	0.019	0.061
Nutrition Sciences	0.144***	0.402***	0.081*	0.232*
Neuroscience	0.088	0.160	-0.094	-0.032
Mathematics And Computer Science	0.722**	0.638***	0.444*	0.335***
Cognitive Science And Biopsychology	0.164	0.367**	0.022	0.137
Interdisciplinary Social Sciences	0.082*	0.195***	0.019	0.081
Multi-Disciplinary Or General Science	0.116***	0.282***	0.015	0.110***
Physical Fitness Parks Recreation And	0.018	0.045*	-0.034	0.005
Philosophy And Religious Studies	0.086*	-0.003	0.028	-0.004
Theology And Religious Vocations	-0.242***	-0.304***	-0.172***	-0.142***
Physical Sciences	-0.085	0.130	-0.163*	0.016
Astronomy And Astrophysics	0.438**	0.339**	0.378*	0.212**
Atmospheric Sciences And Meteorology	0.196	0.335***	0.097	0.152**
Chemistry	0.250***	0.366***	0.101***	0.193***

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
Geology And Earth Science	0.162***	0.261***	0.018	0.117***
Geosciences	0.332*	0.422***	0.095	0.213**
Oceanography	0.025	0.244***	-0.037	0.086
Physic	0.292***	0.383***	0.113***	0.187***
Nuclear, Industrial Radiology, And Bio	0.155	0.276***	0.000	0.112
Psychology	0.076***	0.157***	0.019	0.051**
Educational Psychology	-0.015	-0.098	-0.026	-0.115
Clinical Psychology	0.149**	0.119	0.109	0.087
Counseling Psychology	-0.095**	-0.102	-0.100**	-0.111
Industrial And Organizational Psycholo	0.176**	0.466***	0.046	0.247***
Social Psychology	0.054	0.129	-0.007	-0.028
Miscellaneous Psychology	0.060	0.199**	-0.004	0.149*
Criminal Justice And Fire Protection	0.076***	0.226***	-0.013	0.076***
Public Administration	0.240***	0.292***	0.051	0.098**
Public Policy	0.204**	0.346***	0.028	0.170*
Human Services And Community Organizat	-0.077**	-0.016	-0.098***	-0.052
Social Work	-0.027	0.009	-0.034*	0.017
General Social Sciences	0.055*	0.166***	0.017	0.099**
Economics	0.400***	0.517***	0.224***	0.275***
Anthropology And Archeology	0.069**	0.135***	-0.001	0.053
Criminology	0.123**	0.191***	0.052	0.064
Geography	0.154***	0.212***	0.004	0.085***
International Relations	0.242***	0.398***	0.093**	0.229***
Political Science And Government	0.246***	0.327***	0.112***	0.158***
Sociology	0.077***	0.165***	0.012	0.075***
Miscellaneous Social Sciences	0.340***	0.364***	0.164**	0.213**
Construction Services	0.298*	0.430***	0.121	0.225***
Electrical And Mechanic Repairs And Te	-0.550*	0.145*	-0.332	0.108
Precision Production And Industrial Ar	0.122*	-0.003	0.025	0.013
Transportation Sciences And Technologi	0.292***	0.259***	0.111	0.081**
Fine Arts	-0.021	0.017	-0.067**	-0.035
Drama And Theater Arts	-0.025	-0.089*	-0.065*	-0.135***
Music	-0.109***	-0.034	-0.109***	-0.038
Visual And Performing Arts	-0.097	0.122	-0.084	0.028
Commercial Art And Graphic Design	0.093***	0.127***	-0.017	0.009
Film Video And Photographic Arts	0.014	0.082	-0.012	0.007
Art History And Criticism	0.132***	0.263**	0.064*	0.149
Studio Arts	-0.020	-0.147*	-0.031	-0.178**
General Medical And Health Services	0.163***	0.183***	0.032	0.105*
Communication Disorders Sciences And S	0.144***	0.294***	0.049*	0.138*
Health And Medical Administrative Serv	0.197***	0.242***	0.058	0.058

Major	Major dummies only		With occupation controls	
	Female	Male	Female	Male
Medical Assisting Services	0.338***	0.315***	0.177***	0.152*
Medical Technologies Technicians	0.252***	0.360***	0.137***	0.266***
Health And Medical Preparatory Program	0.481***	0.496***	0.306***	0.190**
Nursing	0.391***	0.408***	0.172***	0.243***
Pharmacy Pharmaceutical Sciences And A	0.641***	0.626***	0.253***	0.406***
Treatment Therapy Professions	0.208***	0.220***	0.101***	0.095**
Community And Public Health	0.151***	0.226***	0.030	0.076
Miscellaneous Health Medical Professio	-0.042	0.083	-0.035	0.078
General Business	0.218***	0.339***	0.077***	0.142***
Accounting	0.310***	0.431***	0.143***	0.199***
Actuarial Science	0.632***	0.764***	0.160	0.337***
Business Management And Administration	0.199***	0.292***	0.054***	0.104***
Operations Logistics And E-Commerce	0.350***	0.403***	0.169***	0.181***
Business Economics	0.432***	0.458***	0.273***	0.204***
Marketing And Marketing Research	0.256***	0.356***	0.089***	0.150***
Finance	0.342***	0.518***	0.151***	0.243***
Human Resources And Personnel Manageme	0.203***	0.258***	0.037	0.064*
International Business	0.284***	0.398***	0.090**	0.149***
Hospitality Management	0.093**	0.095**	0.049	0.060
Management Information Systems And Sta	0.406***	0.485***	0.152***	0.223***
Miscellaneous Business & Medical Admin	0.063	0.363***	-0.037	0.192***
History	0.105***	0.167***	0.033*	0.064***
United States History	0.090	0.127	-0.061	-0.031
Constant	2.217***	2.214***	2.292***	2.235***
R^2	0.200	0.217	0.330	0.337
N	125794	140706	124858	139493

Notes: *** p<0.01, ** p<0.05, * p<0.10
Bachelor's degrees are 4-digit. Wages are top- and bottom-coded at 5 and 400 USD per hour, respectively. General Education is the excluded category. Occupation controls are 5-digit; coefficients not shown.
Sample selection: Observations are included if the individual has at least a bachelor's degree, is working >34 hours per week and >40 weeks per year, and is 23-59 years old.

Supplementary Table 4: Proportion of people in top three occupations by major

Major	Age 25-29 Top 3 occupations	%age	Age 40-44 Top 3 occupations	%age	Age 55-59 Top 3 occupations	%age
Comm.	Marketing and Sales Managers Secretaries and Admin. Assts Human Resources, Training, and Labor Relations Specialists	11.0	Elem. & Middle School Teachers Marketing and Sales Managers Miscellaneous Managers	15.0	Miscellaneous Managers Writers and Editors Chief Executives and Legislators	14.5
Comp. Sci.	Computer Software Engineers Computer Programmers Computer Scientists and Systems Analysts	51.7	Computer Software Engineers Computer Programmers Computer Scientists and Systems Analysts	37.4	Computer Software Engineers Computer Programmers Computer Scientists and Systems Analysts	42.8
Elem. Ed.	Elem. & Middle School Teachers Preschool and Kindergarten Teachers Special Education Teachers	77.3	Elem. & Middle School Teachers Preschool and Kindergarten Teachers Education Administrators	65.0	Elem. & Middle School Teachers Education Administrators Preschool and Kindergarten Teachers	56.0
Elec. Eng.	Electrical and Electronics Engineers Computer Software Engineers Petroleum, Mining, Geol. Engineers	39.3	Computer Software Engineers Electrical and Electronics Engineers Petroleum, Mining, Geol. Engineers	29.5	Electrical and Electronics Engineers Petroleum, Mining, Geol.Engineers Miscellaneous Managers	29.7
Biology	Physicians and Surgeons Postsecondary Teachers Clinical Laboratory Techs	19.4	Physicians and Surgeons Dentists Postsecondary Teachers	29.2	Physicians and Surgeons Postsecondary Teachers Miscellaneous Managers	26.5
Chem.	Postsecondary Teachers Chemists and Materials Scientists Physicians and Surgeons	37.8	Physicians and Surgeons Miscellaneous Managers Chemists and Materials Scientists	31.1	Physicians and Surgeons Chemists and Materials Scientists Postsecondary Teachers	23.4
Psych.	Social Workers Counselors Elem. & Middle School Teachers	20.5	Counselors Social Workers Elem. & Middle School Teachers	17.2	Psychologists Counselors Social Workers	15.8
Econ.	Accountants and Auditors Financial Analysts and Advisors Lawyers, Judges, Magistrates, etc.	18.7	Miscellaneous Managers Financial Managers Lawyers, Judges, Magistrates, etc.	19.0	Lawyers, Judges, Magistrates, etc. Miscellaneous Managers Accountants and Auditors	21.7
Poli. Sci. & Govt.	Lawyers, Judges, Magistrates, etc. Elem. & Middle School Teachers Miscellaneous Legal Support Workers	20.6	Lawyers, Judges, Magistrates, etc. Miscellaneous Managers Elem. & Middle School Teachers	25.5	Lawyers, Judges, Magistrates, etc. Miscellaneous Managers Chief Executives and Legislators	28.8
Fine Arts	Designers Waiters and Waitresses Retail Salespersons	21.9	Designers Elem. & Middle School Teachers Artists and Related Workers	25.6	Elem. & Middle School Teachers Artists and Related Workers Designers	19.2
Nursing	Registered Nurses Nursing, Psych, Home Health Aides Licensed Practical and Licensed Vocational Nurses	89.8	Registered Nurses Med. & Health Services Managers Nursing, Psych, Home Health Aides	83.4	Registered Nurses Med. & Health Services Managers Postsecondary Teachers	79.4
Acct.	Accountants and Auditors Financial Managers Bookkeeping, Accounting, and Auditing Clerks	70.2	Accountants and Auditors Financial Managers Miscellaneous Managers	52.9	Accountants and Auditors Financial Managers Chief Executives and Legislators	50.9
Business Mgmt. & Admin.	Accountants and Auditors Supervisors/Managers, Sales Workers Miscellaneous Managers	18.9	Accountants and Auditors Supervisors/Managers, Sales Workers Miscellaneous Managers	17.5	Miscellaneous Managers Accountants and Auditors Supervisors/Managers, Sales Workers	17.0
History	Elem. & Middle School Teachers Secondary School Teachers Lawyers, Judges, Magistrates, etc.	20.2	Lawyers, Judges, Magistrates, etc. Elem. & Middle School Teachers Miscellaneous Managers	23.1	Lawyers, Judges, Magistrates, etc. Elem. & Middle School Teachers Postsecondary Teachers	22.8

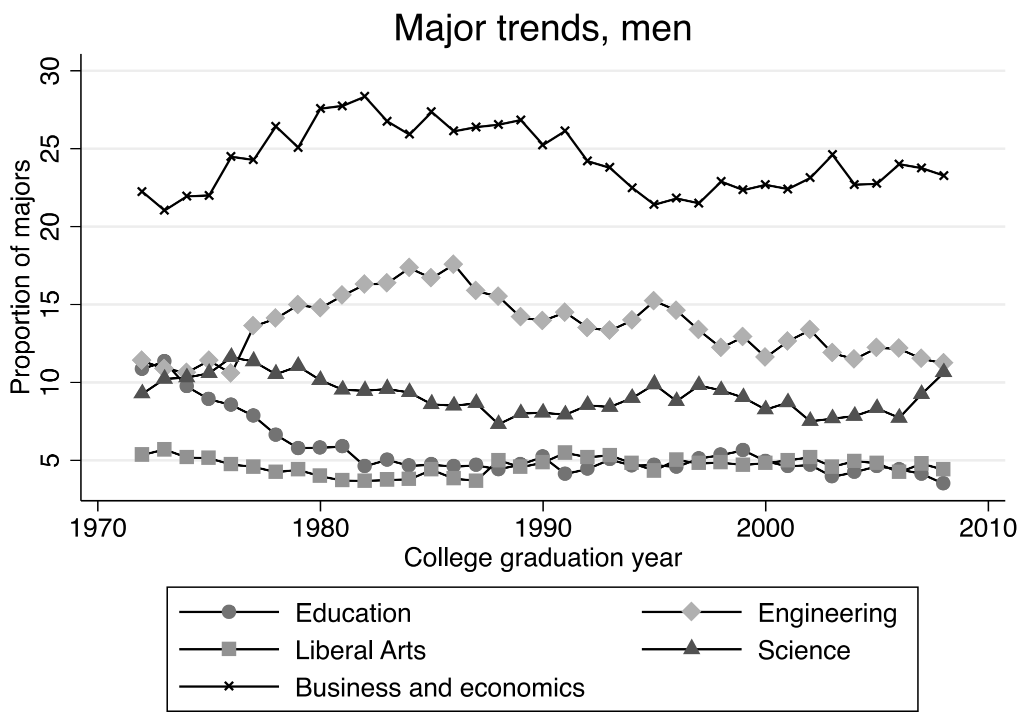
Note: Occupations are 5-digit; only a selected sample shown.

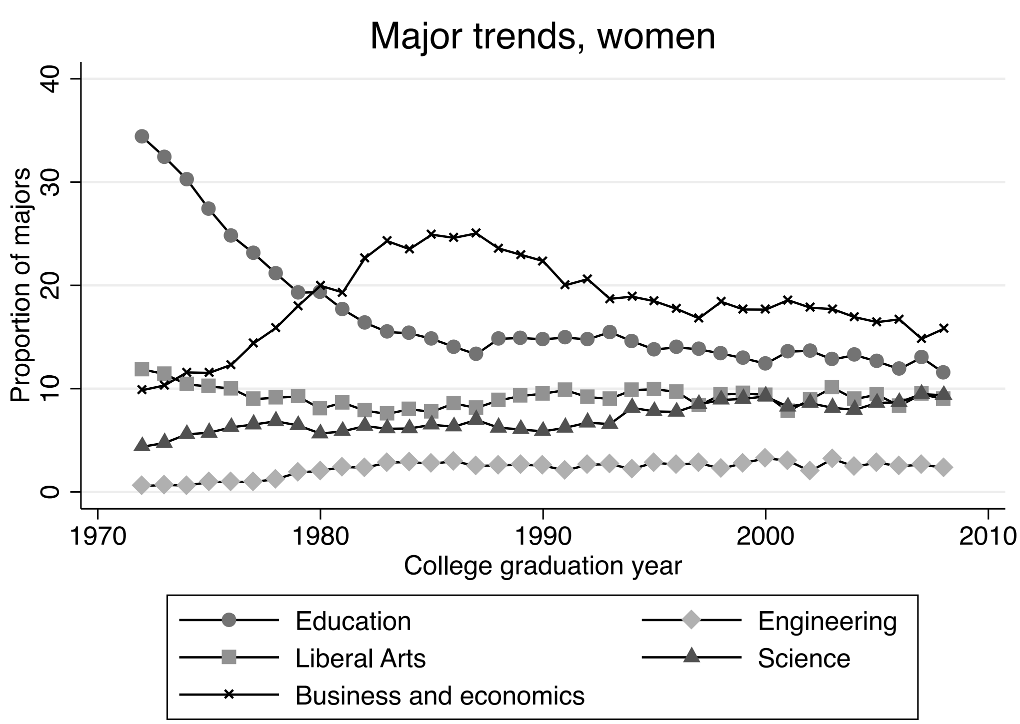
Supplementary Table 5: Proportion of people in top three majors by occupation

Major	Age 25-29		Age 40-44		Age 55-59	
	Top 3 majors	%age	Top 3 majors	%age	Top 3 majors	%age
Chief Executives and Legislators	Business Mgmt. & Admin. General Business Economics	27.8	General Business Business Mgmt. & Admin. Accounting	27.0	Business Mgmt. & Admin. General Business Accounting	24.9
Miscellaneous Managers, Including Postmasters and Mail	Business Mgmt. & Admin. General Business Marketing And Marketing Research	21.5	Business Mgmt. & Admin. General Business Electrical Engineering	19.7	Business Mgmt. & Admin. General Business Biology	19.0
Lawyers, and Judges, Magistrates, and Other Judicial Workers	Political Science And Government English Language And Literature History	36.4	Political Science And Government History English Language And Literature	38.8	Political Science And Government History English Language And Literature	40.3
Marketing and Sales Managers	Marketing And Marketing Research Business Mgmt. & Admin. Communications	41.1	Marketing And Marketing Research Business Mgmt. & Admin. General Business	32.1	Business Mgmt. & Admin. Marketing And Marketing Research General Business	28.5
Financial Managers	Finance Business Mgmt. & Admin. General Business	41.5	Accounting Business Mgmt. & Admin. Finance	46.1	Accounting Business Mgmt. & Admin. General Business	45.4
Education Administrators	Psychology General Education English Language And Literature	20.2	General Education Elementary Education Psychology	24.3	General Education Elementary Education English Language And Literature	28.9
Accountants and Auditors	Accounting Finance Business Mgmt. & Admin.	69.1	Accounting Business Mgmt. & Admin. Finance	72.1	Accounting Business Mgmt. & Admin. General Business	65.8
Computer Software Engineers	Computer Science Computer Engineering Electrical Engineering	63.7	Computer Science Electrical Engineering Computer Engineering	45.4	Computer Science Mathematics Electrical Engineering	41.7
Social Workers	Psychology Social Work Sociology	59.6	Social Work Psychology Sociology	59.3	Social Work Psychology Sociology	52.9
Postsecondary Teachers	Biology Psychology English Language And Literature	20.4	English Language And Literature Biology Psychology	20.9	English Language And Literature Nursing Biology	19.2
Elementary and Middle School Teachers	Elementary Education General Education Early Childhood Education	44.5	Elementary Education General Education English Language And Literature	45.2	Elementary Education General Education Special Needs Education	50.7
Secondary School Teachers	General Education English Language And Literature Art And Music Education	26.8	General Education English Language And Literature Physical And Health Education Teaching	24.4	General Education Secondary Teacher Education Elementary Education	26.2
Physicians and Surgeons	Biology Biochemical Sciences Psychology	53.2	Biology Chemistry Biochemical Sciences	53.0	Biology Chemistry Health & Med. Prep. Progs.	53.3
Registered Nurses	Nursing Biology Psychology	86.2	Nursing Psychology Multi-Disciplinary Or General Science	82.9	Nursing Psychology Multi-Disciplinary Or General Science	81.9
First-Line Supervisors/Managers, Sales Workers	Business Mgmt. & Admin. Marketing And Marketing Research General Business	32.3	Business Mgmt. & Admin. General Business Marketing And Marketing Research	30.7	Business Mgmt. & Admin. General Business General Education	26.1
Sales Representatives, Wholesale and Manufacturing	Marketing And Marketing Research Business Mgmt. & Admin. General Business	35.4	Business Mgmt. & Admin. General Business Marketing And Marketing Research	37.4	Business Mgmt. & Admin. General Business Marketing And Marketing Research	30.7

Note: Occupations are 5-digit; only a selected sample shown.

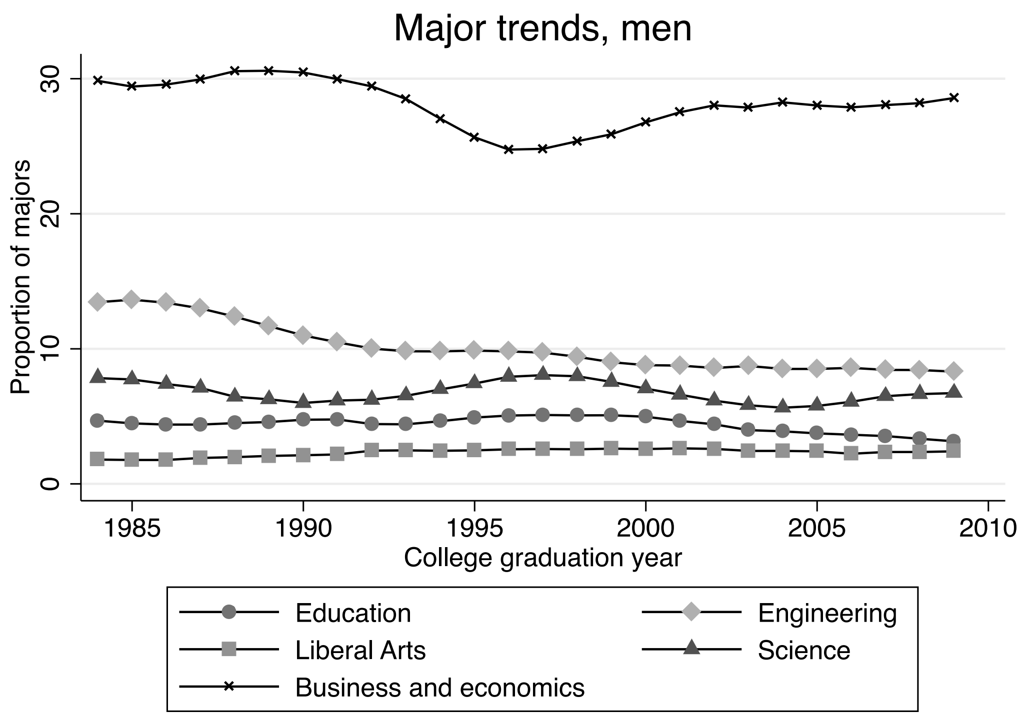
Supplementary Figure 1

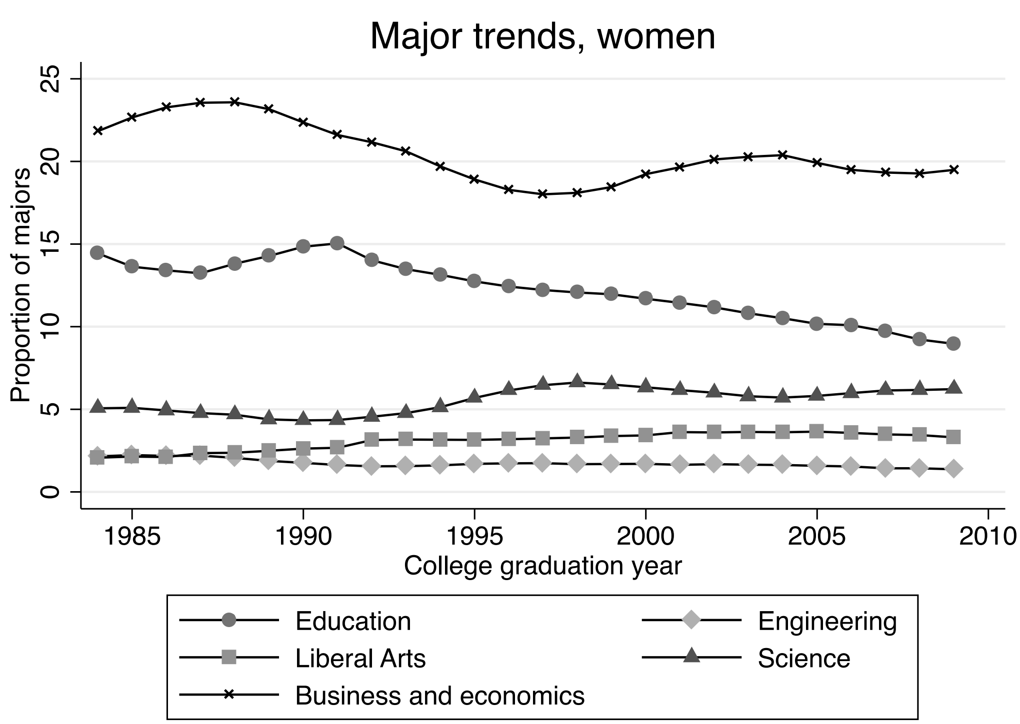




Selected majors only. Data from ACS.

Supplementary Figure 2





Selected majors only. Data from IPEDS.