Worker Signals Among New College Graduates: The Role of Selectivity and GPA*

Brad Hershbein - The University of Michigan

October, 2011

Job Market Paper

Abstract

Recent studies have found a large earnings premium to attending a more selective college, but the mechanisms underlying this premium have received little attention and remain unclear. In order to shed light on this question, I develop a multi-dimensional signaling model relying on college grades and selectivity that rationalizes students' choices of effort and firms' wage-setting behavior. The model is then used to produce predictions of how the interaction of the signals should be related to wages. Using five data sets that span the early 1960s through the late 2000s, I show that the data support the predictions of the signaling model, with support growing stronger over time. I also discuss alternative explanations, including different types of human capital models; provide robustness checks; and relate the findings to both the returns-to-college-quality and employer learning literatures.

JEL Classification: I20, I21, J31

Keywords: college graduates, signaling, school quality, grade point average

*email: bjhersh@umich.edu. I thank Martha Bailey, John Bound, Charlie Brown, Michael Elsby, Brian Jacob, Dmitry Lubensky, Brian McCall, Jeff Smith, Kevin Stange, and participants at the University of Michigan labor workshop and Southern and Western Economic Association Annual Meetings for helpful comments. I also thank Michael Bastedo, Nora Dillon, Ozan Jaquette, and Jeff Smith for data on college selectivity. All errors are my own.

1 Introduction

Recently, there has been a sizable interest in the return to attending a more selective or prestigious college. Students who attend more prestigious schools earn more over their lifetime, on average, than those who attend less selective schools, but the mechanism underlying this premium is not well understood. In particular, there is disagreement over whether the earnings difference is primarily due to the college itself or whether it is driven by unobserved student characteristics. The first of these channels is consistent with human capital theory—attending the more selective school actually makes the worker more productive—and the second more closely accords with models of signaling—more innately productive workers are more likely to attend more selective schools.

Given that annual U.S. higher education expenditures are over \$460 billion, but per-student expenditures increase dramatically with college selectivity, understanding why students who attend selective colleges earn more over their lifetimes has dramatic implications for how those dollars are optimally allocated.[1] Recent theoretical work seeking to explain why students increasingly sort by ability across college selectivities suggests a positive complementarity in human capital acquisition between students' ability and the greater resources available at selective colleges, but these models have received little empirical attention.[2] On the other hand, the relatively few studies that have attempted to measure student learning in college have found little difference across different types of colleges once pre-college characteristics are controlled for (Pascarella and Terenzini 2005; Arum and Roksa 2011). While it is not clear whether the "learning" measured in these studies is of the type that firms would care about, this evidence suggests that the return to selectivity is unlikely to be due to human capital alone and that the signaling mechanism is worth a more careful investigation.

The paper makes several substantive contributions toward understanding the college selectivity premium. First, it develops a novel, multi-dimensional signaling model of ability between college graduate workers and prospective employers. In equilibrium, the utility-maximizing behavior of these agents leads to a specific—and empirically testable—relationship between the two dimensions of the signal, college selectivity and grade point average (GPA), and starting wages. While the full model is elaborate, the crux is intuitive. Students sort into different colleges by ability, and this means that college selectivity is a valuable signal of ability to employers. If grad-

[1] *Digest of Education Statistics*, 2010 edition, table 29; Hoxby (2009).

[2] Courant, Resch, and Sallee (2008) offer one of the most detailed of these types of models and review earlier ones.

uating from a more selective school sends a more precise signal of ability than graduating from a less selective school, the marginal informational benefit of an additional signal, such as GPA, is reduced. When it comes to wage setting, we would expect the relative weight firms place on the GPA signal to be lower at more selective colleges. Consequently, the change in log wages with respect to a change in GPA should be smaller the higher is selectivity. Furthermore, the ability sorting across college types also implies that the selectivity premium should fall as GPA rises. The intuition here is high-GPA students benefit less from attending a selective school because they have demonstrated their ability through their GPA; but for a lower-GPA student at a selective college, firms will discount the noisier signal and place more weight on the college type.

Second, the paper empirically tests the implications of the model. Employing five nationally representative data sets that span five decades, I consistently find strong support for the predictions of my signaling model. The return on GPA is lower at selective colleges and falls as the threshold of selectivity rises. The selectivity premium is highest for those with lower GPAs and declines as GPA rises. Moreover, both of these phenomena have become more pronounced over time as ability sorting across colleges has increased.

Finally, the current study discusses and provides some evaluation of human capital models in explaining the returns to college selectivity. I show that the existing complementarity-based models have opposite predictions from my signaling model and are thus not supported by the data. While the observed empirical patterns are possible under more general human capital models, I delineate the specific conditions that must be met for this to be the case, and I argue that they do not accord with several stylized facts and time trends. Nonetheless, investigation of these types of human capital models in the college quality context bears further research.

The paper proceeds as follows. In the next section, I review some of the recent literature on the returns to college selectivity and employers learning about workers. Sections 3 and 4 develop, characterize solutions, and derive predictions for a multi-dimensional signaling model in the context of college graduate workers whose productivity firms cannot perfectly observe. Section 5 describes the data sets and empirical methodology that are used to explore and test the implications of the model, while Section 6 presents the results of these tests. Section 7 discusses alternative explanations, including various types of human capital models, and relates the findings to the existing literature on college quality and employer learning. The last section concludes.

2 A Literature Review

The earliest studies attempting to measure the return to college selectivity or quality date to the early 1970s and are primarily based on a non-representative sample of skilled (male) World War II military veterans (Wales 1973, Psacharopoulos 1974). Conditioning on observables (including measures of cognitive ability), these early papers find a sizable wage premium in mid-career among respondents who attended colleges in the top fifth of the quality distribution. While Wales discusses several possible explanations for the premium, the data do not allow him to identify which of the explanations drive the results. More recent work has taken advantage of more representative data and advances in identification methods. Brewer, Eide, and Ehrenberg (1999) and Hoxby (2001) attempt to correct for selection on unobservables using nationally representative data, and find a selectivity premium that appears to have grown over time. Black and Smith (2006) use NLSY79 data and several approaches for identification, with their preferred GMM method yielding a selectivity premium that is smaller than the earlier studies, but still statistically significant. Saavedra (2008) and Hoekstra (2009) employ regression discontinuity designs based on a test cutoff for admission to (a specific) selective college, but only Hoekstra's paper is set in a U.S. context. He finds a larger premium than in previous work. Dale and Krueger (2002) is unusual in employing a data set only of students at selective colleges and controlling for the schools to which an individual was accepted; perhaps as a result, theirs is the only paper to find no wage premium from attending a more selective college.

Each of these papers tacitly assumes a world of perfect information in which productivity is directly known by employers, and the objective is to isolate the return to college quality from the return to latent individual ability. However, there is a growing body of work that suggests productivity is not immediately known but must be learned over time. This employer learning literature was begun by Farber and Gibbons (1996) and applied in the (quantity of) education context by Lange and Topel (2006), Lange (2007), and Arcidiacono, Bayer, and Hizmo (2010). These latter papers conclude that employer learning is relatively rapid, especially about college graduates. However, their findings suggest it is possible that, by examining earnings several years if not decades after graduation, the returns-to-college-quality studies conflate the initial premium with revelation of ability or productivity over time.

The existing theoretical work on the returns to college quality makes similar assumptions of

perfect information. In particular, several papers argue that the concomitant increases in ability sorting and school resources experienced by higher ability students can be explained by positive complementarities in student ability and resources in human capital acquisition (Rothschild and White 1995; Epple, Romano, and Sieg 2006; Courant, Resch, and Sallee 2008). The basic line of thinking in these models is that the learning of high ability students is enhanced when they are around other high-ability students and resources (better faculty, libraries, etc.), and firms observe this greater human capital acquisition and pay the students for it. There has been little empirical evaluation of this class of hypotheses, however.

More recently, there is a single paper to my knowledge that investigates a signaling mechanism empirically. Lang and Siniver (2011) investigate the returns to attending the more selective of two universities in Israel that have courses taught by common faculty and that share resources. Using a regression discontinuity design, they find a significant premium to attending the more selective institution and, given the common faculty and other resources, argue that the result is consistent with a quality signal framework. However, they cannot fully control for the possibility of peer effects, and it is unclear whether their results generalize when there is a larger set of schools or apply in the U.S. context, which has a far greater number of institutions of higher education. Thus, there is ample room for further work in exploring signaling in the college selectivity context.

3 A Multi-dimensional Signaling Model of Latent Ability

Consider the labor market between firms and new college graduates they wish to employ. In the United States, this labor market is large, with over 1.5 million graduates annually, more than 75 percent of whom are working full-time one year after graduation.[3] The market is also well-developed and competitive, as evidenced by the popularity of career fairs at colleges and geographical mobility of recent graduates (Malamud and Wozniak, 2008).

In this market, suppose that firms are homogeneous, but prospective workers (i.e., students) vary in their ability, $\eta \sim N(0, 1)$, and this trait affects the worker's productivity to firms.[4] While students can observe their own ability, the firms cannot. Instead, in the spirit of Spence (1973),

[3] *Digest of Education Statistics*, tables 268 and 391.

[4] "Ability" as used here need not be thought of purely as cognitive ability, but a combination of cognitive and noncognitive abilities mapped to a single dimension. Heckman, Stixrud, and Urzua (2006) show in their Table S3 that measures of cognitive and noncognitive ability are positively correlated.

the firms observe imperfect signals of ability that are chosen by the students. These signals, for example, might appear on a potential worker's résumé, be transmitted during a job interview, or appear in the form of references or letters of recommendation. While there may be many such signals, two of note are the undergraduate grade point average (GPA), and the prestige, reputation, or selectivity (SEL) of the degree-granting college. Because most new college graduates have limited prior working experience, both of these measures tend to feature prominently in their résumés, which often serve as the first set of information observed by firms when hiring new workers.[5]

Employers care about these signals because they can be used to form expectations about a worker's productivity. Using this information set, the firm offers a wage to the worker based on its beliefs. From the perspective of a student, increasing the value of these signals is costly—and more costly for those of lower ability—but doing so makes the individual look more productive to prospective employers, and thus can increase the anticipated wage offer. The behaviors of these agents are described more formally below.

3.1 Firm's Problem

Let the production function of a new worker i at time t be given by

$$\ln y_{it} = a_{it} + \rho_t \eta_{it} + \varepsilon_{it}, \tag{1}$$

where $\ln y$ is the natural logarithm of output. The individual-specific intercept a_{it} represents characteristics about worker i other than ability that affect productivity (e.g., through type of job), that may vary over time due to changes in technology or discrimination, and that are observable to both the firm and the econometrician. These characteristics include features such as the major or field of study at college, race, and sex. The scaling factor ρ_t is a positive parameter that measures how closely ability, η_{it}, is related to productivity and which may also vary over time as the importance of skill (or ability) in production changes. Finally, ε_{it} is a normally-distributed random disturbance term that is meant to capture other individual characteristics independent of ability that influence productivity (e.g., luck, random match quality) that are observable to the

[5] McKinney and Miles (2009) review several studies that validate the use of these signals by recruiters at colleges. Indeed, college career office web sites highlight the importance of these two pieces of information by suggesting they feature most prominently on the résumé (`http://www.careercenter.umich.edu/students/resume/sectionexplanations.html`). This is consistent with most hiring comprising a multi-stage process, with the first stage consisting of an initial screening of the résumé.

firm but not the econometrician.

The objective of the firm is to set a wage policy in order to maximize expected profits from a new college graduate worker. Competition among firms, however, ensures that profits are zero in expectation, and so

$$w_{it}(GPA_{it}, SEL_{it}) = a_{it} + \rho_t E[\eta_{it} \mid GPA_{it}, SEL_{it}] + \varepsilon_{it}, \tag{2}$$

where w_{it} represents log wages. The firm's wage schedule depends on how it forms an expectation of a worker's ability given both the GPA and selectivity signals, and this will be a function of optimal student behavior.

3.2 Student's Problem

The student faces a two-stage problem. In the first stage, which occurs during high school, she is concerned with the type, or selectivity, of college she will attend. (As the labor market of interest is new college graduate workers, the effective student population includes only those who graduate from college and then enter the workforce.) For simplicity, suppose there are two types of colleges, indexed by j and denoted selective ($j = 1$) and less selective ($j = 0$), respectively. While admission to the less selective type is guaranteed, entrance to selective schools is competitive and requires effort, $e_1 \in [0, \infty)$, from the student.

Let $P(e_1)$ equal the probability of getting into college type $j = 1$ given effort level e_1. The function $P(\cdot)$ is described by:

$$P(e_1) = \begin{cases} \epsilon & \text{if } e_1 < \tilde{e}_1 \\ f(e_1);\ f'(e_1) > 0,\ f''(e_1) < 0,\ \lim_{e_1 \to \infty} f(e_1) = 1 & \text{if } e_1 \geq \tilde{e}_1. \end{cases} \tag{3}$$

For effort levels below some threshold $\tilde{e}_1$, the probability of admittance into the selective tier of colleges is fixed at ϵ, which is assumed to be close to zero.[6] Only for effort levels above $\tilde{e}_1$ does the likelihood of admittance begin to increase, and in a concave fashion. The probability function thus allows for non-smooth returns to effort, as might be the case under certain admit/reject rules at selective colleges (Toor, 2001).

[6]The ϵ term is a simplification meant to capture students who may gain entry to selective schools through non-academically competitive means, such as legacies and scholarship athletes.

Effort, which here can be thought of as the time and energy put into studying during high school, is costly. However, students find exerting a given amount of effort less costly the greater is their ability. The cost of high school effort is given by

$$C_1(e_1) = \frac{\alpha_2}{\eta + \alpha_1} e_1 + \frac{\alpha_3}{2(\eta + \alpha_1)} e_1^2, \tag{4}$$

where α_1, α_2, and α_3 are each positive constants.[7]

In the second stage, the student has observed the admission outcome and knows what type of college she will attend.[8] At the chosen college type, she must again decide how hard to work, $e_2 \in [0, \infty)$, but this time the outcome of interest is her grade point average (GPA), a summary measure of academic performance. GPA is an affine function of effort, but there is a random noise additive component as well. This error term is independent of effort (and ability) and may reflect personality matches between the student and the professor, arbitrary grading, or simple luck. Thus,

$$GPA(e_2) = \gamma_1 + \gamma_2 e_2 + \nu; \qquad \nu \sim N(0, \sigma_\nu^2), \tag{5}$$

where γ_1 and γ_2 are positive constants. In writing the GPA-effort relationship this way I have made two assumptions. First, GPA is related linearly to effort. This is problematic in the sense that GPA is typically measured on a bounded 4-point scale and equation (5) allows for an unbounded GPA. However, as long as optimal effort levels are in a suitably restricted range, the unboundedness issue should not be a major concern.[9] Second, the GPA function is independent of college type. If the relationship does vary across selectivity type, it is not clear how, *à priori*. For example, it could be argued that classes are more difficult at more selective schools, which could imply a lower γ_1 at these schools if more effort is required to obtain the same expected grade. On the other hand, it has also been argued that grade inflation is more prevalent at selective schools (Kuh and Hu 1999), which could suggest a higher γ_1 and lower γ_2. I examine both of these cases in Appendix B, but proceed for now under (5).

[7] The value of α_1 is such that $\eta + \alpha_1 > 0$ for all but a trivially small range of η.

[8] In equilibrium, there is a wage premium from attending the selective type, and students' beliefs behave accordingly.

[9] Related is that the boundedness of GPA implies ν is not strictly independent of effort. Empirically, this seems to be trivial, however, with approximately 1 percent of individuals recording the maximum 4.0 GPA. As such, I treat this issue as ignorable.

The effort cost function in this stage is similar to that in the first stage:

$$C_2(e_2) = \frac{\delta_2}{2(\eta + \delta_1)} e_2^2, \tag{6}$$

where δ_1 and δ_2 are each positive constants.[10]

Combining both stages, the student's objective can be written

$$Max_{e_1,e_2} \quad U_i = w\left(SEL(e_1), GPA(e_2)\right) - C_1(e_1;\eta) - C_2(e_2;\eta), \tag{7}$$

where w is the log wage earned conditional on GPA and SEL, an indicator variable for whether $j = 1$, and the η subscripts in the cost functions reflect their dependence on a student's ability.[11]

3.3 Solution Characteristics

The student's problem can be solved with backward induction, beginning with the second stage. At the chosen school type j, the first-order condition implies:

$$e_{2j}^* = \frac{(\eta + \delta_1)\gamma_2}{\delta_2} \cdot \frac{\partial w(\cdot)}{\partial GPA}\Big|_{SEL=j}. \tag{8}$$

The student equates the marginal cost of exerting effort with the marginal benefit of higher wages resulting from a higher grade point average. The student's *belief* of how the wage offer changes with GPA, and how this relationship may differ by college selectivity, is key to determining optimal effort. If the belief is that wage changes linearly with GPA, then $\frac{\partial w(\cdot)}{\partial GPA}\Big|_{SEL=j}$ is a constant (which may differ for $j = \{0, 1\}$), and optimal effort rises linearly with a student's ability.[12] This leads to the common-sense prediction that, within a school type, average GPA should be higher among the higher ability students.

Substitution of optimal effort into equation (5) yields:

$$GPA_{ij}\left(e_{2j}^*(\eta_i)\right) = \gamma_1 + \left(\frac{(\eta_i + \delta_1)\gamma_2^2}{\delta_2} \cdot \frac{\partial w(\cdot)}{\partial GPA}\Big|_{SEL=j}\right) + \nu, \quad \text{or} \tag{9}$$

[10] The value of δ_1 is such that $\eta + \delta_1 > 0$ for all but a trivially small range of η.

[11] Equation 7 assumes students are risk neutral. In Appendix C, I briefly sketch how behavior changes when agents are risk-averse.

[12] Optimal effort e_2^* is rising in η as long as $\frac{\partial w(\cdot)}{\partial GPA} > 0$, although the relationship will cease to be linear if $\frac{\partial w(\cdot)}{\partial GPA}$ is not a constant.

$$GPA_{ij}\left(e^*_{2j}(\eta_i)\right) = \gamma_1 + \left(\frac{(\eta_i+\delta_1)\gamma_2^2 k_j}{\delta_2}\right) + \nu,$$

under the assumption that $\left.\frac{\partial w(\cdot)}{\partial GPA}\right|_{SEL=j}$ is a constant k_j. (I discuss the empirical validity of this assumption, as well as the linearity of GPA in ability, in Appendix C.)

Returning to the first stage, although the GPA function is unrelated to college type, there may be complementarity between the two stages if $k_0 \neq k_1$. Suppose, for example, that $k_0 > k_1$. Then an individual with ability η_i will expend more effort in the second stage at a less selective college than at a selective one, and earn a higher expected GPA. The situation would be reversed if $k_1 > k_0$. Acknowledging this possible complementarity, the first-order condition for the first stage is:

$$(w(E[GPA_{j=1,\eta}], SEL_{j=1}) - w(E[GPA_{j=0,\eta}], SEL_{j=0})) \cdot \frac{dP}{de_1^*} \leq \frac{dC_1}{de_1^*}, \quad \text{or} \tag{10}$$

$$e_1^* = \begin{cases} 0 \quad \text{if } \frac{\alpha_2}{\eta+\alpha_1} + \frac{\alpha_3 \tilde{e}_1}{\eta+\alpha_1} > f'(\tilde{e}_1)\left(w(E[GPA_{j=1,\eta}], SEL_{j=1}) - w(E[GPA_{j=0,\eta}], SEL_{j=0})\right) \\ e_1^* \;\middle|\; \frac{\alpha_2}{\eta+\alpha_1} + \frac{\alpha_3 e_1^*}{\eta+\alpha_1} = f'(e_1^*)\left(w(E[GPA_{j=1,\eta}], SEL_{j=1}) - w(E[GPA_{j=0,\eta}], SEL_{j=0})\right), \text{ else.} \end{cases}$$

Because the transition to a different selectivity college is possibly associated with a change in expected GPA, the return to moving from a less selective to more selective institution is not simply the partial derivative (technically, discrete change) of log wages with respect to selectivity but must include the expected change in GPA as well. In the first-order condition, this return is expressed as the discrete change in the wage as both arguments change, and it is multiplied by the change in probability of admission that comes with increased effort. For a (unique) interior solution to exist, this probability-weighted return must be at least equal to the marginal cost of effort at the threshold $\tilde{e}_1$, where the likelihood of admission begins to rise.

The solution can perhaps best be explained graphically, as in Figure 1. For the sake of exposition, the figure plots marginal cost and benefit curves for three ability types: high (η_H), medium (η_M), and low (η_L). Equation (4) implies that that marginal cost of effort has both the slope and intercept decreasing in ability. The marginal benefit curves (dashed) capture the expected return to moving from a less to more selective institution, weighted by the change in admission probability from increased effort. For effort levels less than $\tilde{e}_1$, there is no change in admission probability from increasing effort, and so the marginal benefit curve has a value of zero. For higher effort levels, the concavity of $f(\cdot)$, the probability of admission to the selective tier, ensures that the

marginal benefit curves are downward sloping. It remains, though, to characterize the net return from moving from a less selective to more selective college.

Notably, for a fixed ability level, the expected return from switching selectivity levels is a constant, since the expected GPA arguments in the wage equation are themselves constants by second stage optimization. However, across ability levels, this expected return will vary. Since the difference in expected GPA between selectivity tiers is larger the higher is ability,[13] higher ability types experience a larger change in the net return from the GPA component when switching selectivity tiers. If $k_1 < k_0$, this means higher ability types enjoy a smaller expected wage gain when moving to the more selective tier. This effectively lowers the slope of the marginal benefit curve, as shown in the figure. (If, instead, $k_1 = k_0$, the marginal benefit curves would be identical across ability, and if $k_1 > k_0$, the slope of the marginal benefit curve would become steeper as ability rises.)

Three things bear mentioning. First, students below some ability threshold (denoted $\tilde{\eta}$ and implicitly defined by (10)) do not find it worthwhile to expend any effort in the first stage. (This characterization is shown for η_L in the figure.) Only a trivial fraction of these students (ϵ of them) will be admitted and attend the selective tier of colleges. Second, for students above this threshold, optimal effort is rising in ability under relatively weak conditions.[14] Third, the threshold $\tilde{\eta}$ is rising in $\tilde{e}$. (Appendix A provides proofs.) The first two features together imply that the likelihood of gaining admission (and attending) selective schools is rising in ability. The third feature implies, sensibly, that when more effort is required to increase the probability of gaining admittance to selective schools, only increasingly higher ability students will find it worthwhile to do so.

4 Firm Expectations of Student Ability and Predictions

4.1 Moment Expectations

For a given $\tilde{\eta}$ the features described above lead to the following propositions:

PROPOSITION 1: Mean ability is higher at more selective schools.

[13] $E[GPA_{j=1,\eta}] - E[GPA_{j=0,\eta}] = \left(\frac{\gamma_2^2(\delta_1+\eta)}{\delta_2}\right)(k_1 - k_0)$.

[14] Marginal cost must decline in ability faster than does the selective-school wage premium from the reduction in expected GPA.

Proposition 2: A higher ability threshold, $\tilde{\eta}$, leads to a larger difference in mean ability between more and less selective schools.

Proposition 3: A higher ability threshold, $\tilde{\eta}$, leads to a lower variance in ability at more selective schools.

Proposition 4: Variance in ability is lower at more selective schools when $\tilde{\eta} > 0$.

Proofs: Appendix A.2

Intuitively, because students who attempt selective entry are of higher average ability than those who do not, selective colleges will have higher ability students on average. Furthermore, raising the ability threshold for applying must amplify the average ability gap, as the applicant pool for selective colleges will shrink proportionately more than the less selective pool will grow.

It also follows that the variance of ability, conditional on the student having graduated from the selective tier, is falling in $\tilde{\eta}$. This occurs nearly mechanically; a higher minimum threshold reduces the fraction of the student population who find it worthwhile to exert effort in the first stage, and so the conditional variance falls as a result. More generally, it is not necessarily the case that the variance of ability at the selective tier is smaller than at the less-selective tier for all values of $\tilde{\eta}$. When $\tilde{\eta} > 0$ this will necessarily be true, as less than half the ability distribution "applies" to the selective schools and not all of them will get in. When $\tilde{\eta} < 0$, whether the conditional variance is smaller at the selective tier will depend on the shape of $f(\cdot)$, which will affect the skewness of ability distributions across school types. However, in the data used in this study far fewer than half of the eventual college graduates reported applying to the selective tier, so it seems reasonable that $\tilde{\eta} > 0$ and the variance of ability is smaller at the selective tier.

How do firms incorporate both selectivity and GPA into their expectations? Recall that an optimizing student's GPA is linear in η plus a normally distributed, independent error term. If η is normally distributed, *conditional on selectivity*, then GPA, as the sum of two independent normal random variables, is normally distributed as well, and GPA and η are jointly distributed as bivariate normal. As documented by Aigner and Cain (1977), among others, this would imply that the conditional expectation of ability given selectivity and GPA is of the form:

$$E[\eta_i \mid GPA_{ij}, SEL_{ij}] = E[\eta_i \mid SEL] + \frac{Cov(\eta, GPA_j)}{\sigma^2_{GPA_j}} \left(GPA_{ij} - \mu_{GPA_j}\right). \tag{11}$$

The conditional expectation of ability given both selectivity and GPA is linear in GPA, with both the slope and intercept varying by selectivity tier. Of course, bivariate normality is unlikely to hold exactly, as the necessary sorting by ability would occur only under a specific $f(\cdot)$.

Yet this assumption may not be a poor one. If the distribution of η is reasonably close to normal at both selectivity tiers, then GPA at each tier should be approximately normal as well. In Appendix C.2, I show that this assumption holds up quite well empirically. (Specifically, I examine the conditional distributions of GPA and measures of ability across selectivity tiers in the data; I also attempt to bound the variance of ν, the GPA disturbance term, using moments from the data.)

Proceeding under the bivariate normal assumption, it follows from equation (2) that log wages at a given time (t subscript suppressed) are given by:

$$w_{ij}(GPA_{ij}, SEL_{ij}) = a_i + \rho\left(\psi_j + \frac{(\gamma_2^2\delta_2^{-1}k_j)\sigma_{\eta_j}^2}{(\gamma_2^4\delta_2^{-2}k_j^2)\sigma_{\eta_j}^2 + \sigma_\nu^2}GPA_{ij}\right) + \varepsilon_i, \tag{12}$$

where ψ_j is a function of the structural parameters that depends on j, and $\sigma_{\eta_j}^2$ is the variance in ability for college type j.[15] The return to GPA on log wages is thus:

$$\frac{\partial w_{ij}}{\partial GPA_{ij}} = \frac{\rho\gamma_2^2\delta_2^{-1}k_j\sigma_{\eta_j}^2}{\gamma_2^4\delta_2^{-2}k_j^2\sigma_{\eta_j}^2 + \sigma_\nu^2}. \tag{13}$$

It was assumed earlier that, according to students' beliefs, $\frac{\partial w_{ij}}{\partial GPA_{ij}} = k_j$. In the context of (13), a Nash equilibrium in which beliefs are accurate means that the following should hold:

$$\frac{\partial w_{ij}}{\partial GPA_{ij}} = \frac{\rho\gamma_2^2\delta_2^{-1}k_j\sigma_{\eta_j}^2}{\gamma_2^4\delta_2^{-2}k_j^2\sigma_{\eta_j}^2 + \sigma_\nu^2} \equiv h(k_j) = k_j. \tag{14}$$

Since $h(\cdot)$ is continuous in k_j, is plausibly bounded on a closed interval, and maps to its own domain by assumption, k_j exists by Brouwer's fixed point theorem.

[15] $\psi_j \equiv \mu_{\eta_j}\left(1 - \frac{\zeta_j\sigma_{\eta_j}^2}{\zeta_j\sigma_{\eta_j}^2+\sigma_\nu^2}\right) - (\gamma_1\zeta^{-\frac{1}{2}} + \gamma_2)\left(\frac{\zeta_j\sigma_{\eta_j}^2}{\zeta_j\sigma_{\eta_j}^2+\sigma_\nu^2}\right)$, with $\zeta_j \equiv \gamma_2^4k_j^2\delta_2^{-2}$.

4.2 Cross-sectional Predictions

How does k_1 relate to k_0? Since $\sigma^2_{\eta_1} < \sigma^2_{\eta_0}$, by (14) $k_1 \neq k_0$. Yet, the same equation makes it possible, for certain parameter values, for either $k_1 > k_0$ or $k_1 < k_0$. It turns out, however, that any possible equilibrium with $k_1 > k_0$ cannot be supported as a (perfect Bayesian) Nash equilibrium. Suppose $k_1 > k_0$, such that the return on GPA is higher at selective colleges. Then firms must believe that, on average, the increase in ability from a one-point rise in GPA is higher at selective colleges than at less selective colleges. But it has already been shown that the variance in ability is smaller at selective colleges. (Indeed, this is verified empirically in Table 1.) With a smaller variance in ability, but a fixed GPA range, it is not rational to believe that a unit change in GPA corresponds to a larger increase in ability at selective colleges. Therefore, $k_1 > k_0$ is not a valid equilibrium.[16] Thus the only surviving equilibrium has $k_1 < k_0$. This leads to the following prediction.

PREDICTION 1: The return on GPA should be higher at less selective schools than at more selective schools.

Moreover, if the threshold $\tilde{\eta}$ is increased, the resulting variance in ability at selective schools, $\sigma^2_{\eta_1}$, will be smaller. As σ^2_{ν} and other parameters remain unchanged, however, the strength of the GPA signal at selective schools will decline further, and thus so will k_1 relative to k_0.[17] Thus, there exists the next prediction.

PREDICTION 2: As the selectivity threshold becomes more restrictive ($\tilde{\eta}$ increases), the difference in the GPA returns between less selective and more selective schools should increase.

By taking equation (12) and differencing between selective and less selective colleges and then taking the derivative with respect to GPA, one can show that the selectivity premium is a linear function of GPA with slope $k_1 - k_0$. Since it has been argued that $k_1 < k_0$, there is another prediction:

PREDICTION 3: The selectivity premium is falling in GPA whenever $k_1 < k_0$.

[16] For k_1 to be greater than k_0, the necessary condition is that the ratio of the ability-GPA covariance to the variance of GPA is larger at more selective schools (see (11). This is strongly rejected in every data set.

[17] See Appendix C.3, "Bounding the variance of ν" for an exercise that relates the magnitudes of $\sigma^2_{GPA_1}$ and σ^2_{ν}.

4.3 Trend Predictions

In addition to generating these predictions in a cross-section, the model can also be used to investigate the integration of the college market over the past 40 years that has been thoroughly documented by Hoxby (2009). In effect, reductions in communication, transportation, and information costs have nationalized (or even globalized) the college market in a way that has allowed selective colleges to become more discriminating in which students they accept. In the context of the model, the measure of the student population has increased faster than the supply of slots at selective colleges. For the market to clear, the "price" of admission also needs to have risen, or, put differently, the minimum first-stage effort threshold, $\tilde{e}$, has increased.[18] But, as was shown earlier, a rise in $\tilde{e}$ leads to a higher $\tilde{\eta}$, and this in turn yields a higher conditional expectation and lower conditional variance of ability at selective schools.

Taking the derivative of (13) with respect to $\sigma^2_{\eta_1}$ yields:

$$\frac{\partial^2 w_{i1}}{\partial GPA_{i1} \partial \sigma^2_{\eta_1}} = \frac{\rho \gamma_2^2 \delta_2^{-1} k_j \sigma_\nu^2}{\left[\gamma_2^4 \delta_2^{-2} k_j^2 \sigma^2_{\eta_1} + \sigma^2_\nu\right]^2} > 0. \tag{15}$$

Since $\sigma^2_{\eta_1}$ should be falling, this implies that the return on GPA at more selective colleges should decline as ability sorting increases.

Additionally, Murnane *et al.* (1995) and Heckman and Vytlacil (2001), among others, have documented a rising return to skill or ability since the 1980s. In the context of the model, this corresponds to a rise in ρ_t, the association between ability and productivity. While equation (14) clearly shows that the return on GPA is rising in ρ, it should be noted that the effect is more pronounced the larger is k_j. It follows that the return on GPA should have increased faster at less selective schools than at more selective schools. Combining the changes in $\sigma^2_{\eta_1}$ and ρ produces prediction 4:

PREDICTION 4: The difference in the return on GPA at less selective and more selective schools should grow larger over time.

[18] Bound, Hershbein, and Long (2009) discuss these changes in more detail and provide extensive evidence that measures of high school effort have increased greatly among those who attend and apply to selective colleges. They also show that in the absence of this increased effort, the probability of admission to selective colleges would have fallen over time.

5 Data and Empirical Strategy

5.1 Data

To test the implications derived above, I use three panel surveys of students conducted by the National Center for Education Statistics: the National Longitudinal Study of the High School Class of 1972 (NLS72), the High School and Beyond (HSB), and the National Education Longitudinal Study (NELS). These data are supplemented by two additional data sets: Project Talent (PT) and the National Longitudinal Survey of Youth, 1997 (NLSY97). Each of these nationally representative data sets tracks students beginning in secondary school, following them through postsecondary education and the transition into the workforce. They contain detailed information on postsecondary schools attended, degrees earned, course grades, and job characteristics. They also contain the results of an aptitude test battery administered to the students during adolescence, typically the senior year of high school; this score can be used as a measure of ability.[19] Importantly, the restricted-access versions of these data sets, used in this paper, allow the identification of all post-secondary institutions attended and, for the NCES data, have complete post-secondary transcript data for students who reported attending a post-secondary institution. Each survey is similar in scope and types of questions asked but covers cohorts roughly 10 years apart—college graduates in the mid 1960s (PT), late 1970s (NLS72), late 1980s (HSB), late 1990s (NELS), and mid-to-late 2000s (NLSY97). They thus facilitate analyses for pooled cohorts that span 40 years and longitudinal analyses across cohorts.[20] The Data Appendix discusses the sampling frame of these surveys in more detail.

As the focus of analysis is new college graduate workers, in each data set the sample is restricted to individuals who had earned their bachelor's degree at U.S. institutions within 6 years of high school graduation and began a job after earning their bachelor's degree.[21] Furthermore, at the time of beginning their post-college graduation job, they must have earned no additional (graduate) degree, not have been enrolled in school, been working for pay with real (year 2005) hourly earnings between 5 and 100 dollars, and have been neither self-employed nor in the military. Last,

[19] As these were low-stakes tests, the ability measure picks up both non-cognitive as well as cognitive abilities.

[20] I have also performed cross-sectional analysis separately for each cohort. Point estimates are qualitatively similar to those reported in this paper, although they are less precise.

[21] For students who transfer colleges, the bachelor degree-granting institution is used. Gill and Leigh (2003) find no wage differences among bachelor degree recipients who began at two- or four-year colleges.

college GPA and the bachelor-degree-granting institution must be identifiable for the respondent.[22] Appendix Table 1 contains more detailed information on how the restrictions affect the sample size for each data set.

Empirical analysis of the theoretical model described in Sections 2 and 3 rests on a practical measure of college selectivity. The primary measure of college selectivity used in this paper is drawn from the competitiveness index from *Barron's Profile of American Colleges*. Each year, *Barron's* classifies nearly all four-year colleges and universities in the country into six categories according to their admissions selectivity. The criteria used to classify colleges includes median ACT or SAT scores for the most recent freshman class, minimum grade point averages and high school class rank required for admission, and the acceptance rate for applicants to the most recent freshman class. Using an electronic data set of the *Barron's* rankings for the years 1972, 1982, 1992, and 2004 that was created by Bastedo and Jaquette (2009), I create three different binary indicators for college selectivity for each of the five data sets. The first of these indicator variables is coded as 1 if the college is ranked in *Barron's* top three categories and 0 otherwise (Tier I); the second is coded as 1 if the college is ranked in *Barron's* top two categories and 0 otherwise (Tier II); and the third is coded as 1 if the college is ranked in *Barron's* top category and 0 otherwise (Tier III). Note that these three tiers are nested; Figure 6 provides examples of colleges in each selectivity tier. The 1972 rankings are used for Project Talent and NLS72 (or 1974 when 1972 rankings are unavailable), the 1982 rankings for HSB, the 1992 rankings for NELS, and the 2004 rankings for NLSY97.[23]

Some summary statistics of the estimation samples from each data set can be found in Table 1. A detailed description of these variables is found in the Data Appendix. In each data set, average log wages of the post-graduation job typically rise with the selectivity of the institution attended, with this gradient getting steeper over time. Average grades also consistently rise with selectivity, but by much less than does either proxy for ability (SAT/ACT percentile or senior test score), which is consistent with $k_1 < k_0$. Additionally, not only is the variance in either ability measure falling as selectivity rises, but, consistent with the predictions of the model and the empirical argument of Hoxby (2009), this becomes more pronounced over time.

[22] College GPA is generally taken directly from transcripts and from self-reports when not transcripts were not available. See the Data Appendix for details.

[23] The rankings tend to be fairly consistent over time. The Data Appendix describes an alternative college selectivity measure that does not vary over time, and results using this measure are discussed later as a robustness check.

5.2 Methodology

In order to test Predictions 1 through 3, I estimate the following reduced-form of equation (12) separately for each selectivity tier threshold using the pooled data:

$$w_{id} = \theta_0 + \theta_1 S_{id} + \theta_2 GPA_{id}(1 - S_{id}) + \theta_3 GPA_{id}(S_{id}) + \sum_d \lambda_d D_d + \sum_d \lambda_{\mathbf{X}} D_d \mathbf{X_{id}} + \epsilon_{id}, \quad (16)$$

where w_{id} is the logarithm of the hourly wage of worker i from data set d, GPA is the college grade point average, S is an indicator that takes the value of 1 if the individual graduated from a selective college and 0 if she did not, D_d is a dummy for each data set, and $\mathbf{X_{id}}$ is a vector of dummies for sex, race, and college major. The interaction between D_d and $\mathbf{X_{id}}$ allow the effect of sex, race, and college major to vary across each data set and capture the a_{it} term in equation (1).[24] Because graduates of the same college presumably had access to similar resources in searching for their post-graduate job (e.g., the same career office on campus), the idiosyncratic error ϵ_{id} may be correlated among these students; variance estimation thus allows for this arbitrary within-college correlation.

Except for the addition of the GPA variables, equation (16) appears similar to many of the estimating equations used in the returns-to-college-quality literature. The parameter θ_2 represents the (approximate) percent increase in wages resulting from a one-point increase in GPA at a less-selective college, and θ_3 represents the same at a selective college. According to Prediction 1, $\theta_2 > \theta_3$. Moreover, as the threshold for selectivity grows higher, Prediction 2 posits that the difference between θ_2 and θ_3 should be larger. In practice, this means that we would expect $\hat{\theta}_2 - \hat{\theta}_3$ to be larger when estimated for Tier II than for Tier I (and similarly for Tier III than for Tier II).

The return to selectivity in equation (16) can vary by GPA, something that earlier work in the return to college quality did not allow. Specifically, the return to selectivity is given by $\theta_1 - (\theta_2 - \theta_3)GPA$. Prediction 3 implies that, since $\theta_2 - \theta_3 > 0$, the return to selectivity falls as GPA rises, but that it should remain weakly positive at the maximum GPA. Again, resource-ability complementarity predicts the opposite.

Furthermore, Prediction 4 argued that increasing ability-sorting across colleges and returns

[24] For consistency across data sets, race is coded as "white", "black", or "other", and college major consists of 11 categories: humanities, social sciences, psychology, life sciences, physical sciences and mathematics, engineering, education, business, arts, health, and other.

to skill should intensify the first three predictions. To test this hypothesis, I divide the data into an "early" period consisting of the data sets from the 1960s and 1970s and a "late" period consisting of the data from the 1980s, 1990s, and 2000s. (This division accords with the findings of growing returns to skill that began in the 1980s and also balances sample sizes.) I then estimate:

$$\begin{aligned} w_{id} =&\theta_0 + \theta_{11}S_{id} + \theta_{12}S_{id}Late_{id} + \theta_{21}GPA_{id}(1-S_{id}) + \theta_{22}GPA_{id}(1-S_{id})Late_{id} \\ &+ \theta_{31}GPA_{id}S_{id} + \theta_{32}GPA_{id}S_{id}Late_{id} + \sum_d \lambda_d D_d + \sum_d \lambda_{\mathbf{X}} D_d \mathbf{X_{id}} + \epsilon_{id}, \end{aligned} \tag{17}$$

where $Late_{id}$ equals 1 if the individual is from the HSB, NELS, or NLSY97 datasets, and 0 otherwise. In this equation, θ_{21} gives the return on GPA at less selective schools in the early period, $\theta_{21} + \theta_{22}$ gives the return on GPA at less selective schools in the late period, θ_{31} gives the return on GPA at more selective schools in the early period, and $\theta_{31} + \theta_{32}$ gives the return on GPA at more selective schools in the late period. The return to selectivity is given by $\theta_{11} - (\theta_{21} - \theta_{31})GPA$ in the early period, and by $\theta_{11} + \theta_{12} - ((\theta_{21} + \theta_{22}) - (\theta_{31} + \theta_{32}))\,GPA$ in the late period. Prediction 4 asserts that $\theta_{22} > \theta_{32}$, which implies that the return on GPA has grown faster at less selective schools *and* that the return on selectivity, while higher on average, has declined more rapidly with GPA.

6 Estimation Results

6.1 Pooled Model

Table 2 presents the results from estimating equation (16) on the pooled data. Columns 1 through 3 use selectivity tier I, II, and III, respectively, on the entire eligible sample, while columns 4 through 6 repeat the analysis on the full-time worker sample. At less selective colleges, the return on GPA is highly significant at about 9 percent for the whole sample, regardless of the selectivity threshold. However, these returns are uniformly smaller at selective colleges, and for tier II and tier III colleges, the returns are statistically indistinguishable from zero. Of course, the standard errors tend to be much larger for the selective college GPA estimates, especially at the higher tiers, because the effective sample sizes are so much smaller. Consequently, the null hypothesis that the returns on GPA are the same across selectivities cannot be rejected at conventional levels in columns 1 through 3. Nonetheless, the point estimates are fairly close to 0 for selective colleges in

columns 2 and 3.

For the full-time sample, the patterns are remarkably similar. Less selective college graduates earn a GPA return of 10 to 11 percent, but graduates from selective colleges do not enjoy the same benefit from a higher GPA. A graduate of a tier I (or higher) school earns only 0.073 log points per point increase in GPA, and this return is statistically less than that at non-tier I schools at the 10 percent level. The GPA returns fall monotonically as the selectivity threshold increases to tier II and tier III. The return at tier II is one third the size of less selective schools', and the difference is statistically significant at 5 percent. The tier III gap is even more dramatic, although it is not as precisely estimated.

The pattern of these coefficients and the magnitude of their differences are striking. Furthermore, these results are reasonably robust to the specific definition of selectivity. Panel A of Appendix Table 2, for example, repeats Table 2 using the an alternative measure of selectivity suggested by Black and Smith (2006) that is based on college inputs. The table shows similar, if noisier, patterns. The data therefore appear to confirm predictions 1 and 2.[25]

Although Table 2 shows that the selectivity premium estimate is positive and statistically significant at the mean GPA of 3.0, the return on selectivity implied by equation (16) is best shown graphically. Figure 7 plots the selectivity return (in log points) against GPA for full-time workers using the tier II definition (column 5 of Table 2), although using the sample of all workers or other selectivity thresholds does not appreciably change the picture. Since $\hat{\theta}_2 > \hat{\theta}_3$, the selectivity return slopes downward. Looking at the pooled 1960s through 2000s sample, students with a GPA of 2.0, around the 5th percentile of the pooled sample, earn 0.14 log points more at their first job if they graduated from a selective college, and the marker at this point indicates that this premium is statistically significant at the 5 percent level. The premium is reduced to about 7 percent at the sample mean GPA of 2.97, and although it remains positive for the rest of the GPA distribution, it ceases to be statistically different from zero at GPAs above 3.2. Perhaps more important, one can reject that the selectivity premium is the same for any two different GPA points; thus, the 0.14 log point premium at a GPA of 2.0 is not only different from the 0.07 log point premium at a GPA of

[25] I have also estimated variants of (16) that interact selectivity with the controls for sex, race, and major. These interaction coefficients typically are small and statistically insignificant for sex and race, although the returns to social sciences, physical sciences, and engineering (relative to humanities) are larger at selective schools. Allowing these interactions, however, has minimal effect on the GPA estimates presented above.

3.0, it is also different from the premium of 0.13 at a GPA of 2.2.[26] This confirms prediction 3and provides further support for the signaling model.

6.2 The Model Over Time

Both the rising return to ability and increased ability sorting at colleges should serve to widen the gap in GPA returns between selective and less selective colleges (equation (15)). This is tested formally in Table 3, which is similar to Table 2 but provides estimates separately for the 1960s-1970s and 1980s-2000s periods.

Panel A shows that in the early period, graduates of less selective colleges earned a statistically significant return on GPA of between 5 and 7 percent. Their counterparts at selective colleges earned a much lower premium: at tier I colleges, the return is marginally significant at 3 to 4 percent; at the more selective tier II and tier III colleges, the point estimates are essentially zero. However, these gaps are small enough in magnitude (and the selective college GPA coefficients are too noisily measured) that a null of no difference between the groups cannot be rejected.

Switching to panel B and the late period, the coefficient estimates for graduates of less selective schools are about 0.13 for the whole sample and 0.14 to 0.15 for full-time workers. At tier I colleges, the GPA return, while statistically significant, is about half this size. For the full sample, the gap in GPA returns widens from 0.018 in the early period to 0.059 in the late period, roughly tripling, though the latter difference just fails statistical significance. For full-time workers, however, the gap rises from 0.030 to 0.071 and is significant at the 10 percent level.

At tier II and III schools, the growth in the gap is more pronounced, largely because the returns on GPA at these selective schools did not change at all. Among all workers, the tier II gap grows from 0.046 to a statistically significant 0.136, and the tier III gap increases from 0.078 to 0.137. For full-time workers, these gaps jump from 0.062 to 0.119 and 0.062 to 0.130. Only the last of these, owing to the small sample size of tier III grads, fails to be statistically significant.[27] Moreover, these results are robust to using the alternative quality index definition of selectivity, as shown in Appendix Table 2, panels B and C.

[26]The linearity of GPA results in all Wald statistics of selectivity differences across GPA having the same value.

[27]The GPA estimates for less selective colleges are lower when the selectivity threshold is higher because the less selective group includes the tier I colleges that are not tier II (columns 2 and 5) or the tier II colleges that are not tier III (columns 3 and 6). If the tier III selective college estimate in panel B is compared with the less selective estimate from column 4, the two are statistically different at 10 percent.

When one attempts to measure whether the *growth* in the GPA return gaps is statistically significant, this difference-in-difference, while of a non-trivial magnitude, comes up short. Despite this growth averaging (across selectivity tiers) about 0.06 log points, greater than the GPA returns at less selectives in the early period, the estimates at selective schools are too noisily measured for a double difference to have sufficient precision for these data. While a null of no growth in the gap cannot technically be rejected, the size of the point estimates is suggestive.

Returning to Figure 7 and the selectivity premium by GPA, we find that not only has the selectivity premium risen throughout the GPA range between the 1960s-1970s and 1980s-2000s periods, but, as a consequence of $\hat{\theta}_{22}$ being larger than $\hat{\theta}_{32}$, the premium's decline with GPA has become more pronounced. The selectivity premium at a GPA of 2.0 increased from 0.083 log points in the early period to 0.207 log points in the late period, for a gain of 0.124. At a GPA of 3.0, closer to the mean, the selectivity premium rose from a statistically insignificant 2 percent to 9 percent. While this growth is considerable, it is much smaller than the gain at a 2.0 GPA, and growth in the selectivity premium at higher levels of GPA is smaller still and generally not statistically significant. Furthermore, while one can easily reject that the selectivity premium does not vary with GPA in the later period, this hypothesis cannot be rejected in the early period, where both the level and slope are smaller.

These results support prediction 4, that the GPA return gap between more and less selective schools has widened over time and, consequently, that the selectivity premium has become more dependent on GPA. Moreover, the specific mechanisms underlying the prediction are supported. The GPA return at less selective schools has unambiguously risen as ρ has increased. The GPA return at more selective colleges has barely changed over time: not only is the effect of ρ on these GPA returns weaker than at less selective colleges, but the shrinking ability variance would have served to reduce the GPA return (equation (14)). On net, then, it is perhaps not surprising that the GPA return has changed so little at selective colleges.

7 Alternative Explanations and Discussion

The findings described above match the signaling predictions outlined in section 4, but they do not conclusively rule out alternative explanations that could produce the patterns observed in the data. In this section, I discuss two of these alternatives: (a) variants of a human capital model,

and (b) differential measurement error. While there is considerable evidence against latter of these explanations driving the observed outcomes, it is not possible to rule out all possible human capital models. Nevertheless, the empirical findings presented above are inconsistent with the most common human capital model in the college quality literature, the resource-ability complementarity model, and specific necessary conditions must be met in order to rationalize the findings under more general human capital models.

Aside from these other potential explanations, one may wonder how representative jobs soon after graduation are of the jobs in mid-career typically studied in college quality papers. Using the NLSY79, I demonstrate that these early jobs are, in fact, highly predictive of later career wages. Finally, I relate the empirical support for signaling found here back to the motivation of the broader employer-learning and returns-to-college-quality literatures.

7.1 Human Capital Models

Might the empirical findings stem from a variant of the human capital model? In a general human capital framework with effort, the individual productivity function can be expressed as

$$ln\ y_i = a_i + \rho\,(\eta_i + f(\eta_i, SEL_i, GPA_i(e))) + \varepsilon_i, \tag{18}$$

where the notation is as in (1). Here, $f(\cdot)$ represents the productive value added by graduating from college, and may depend on the individual's initial ability, the selectivity or prestige of the college attended, and the effort exerted (as passed through the GPA function).[28]

7.1.1 Resource-Ability Complementarity Models

The resource-ability complementarity models postulated by Rothschild and White (1995) and thoroughly explored by Courant *et al.* (2008) were developed to rationalize as an efficient process the sorting of students by measured aptitude across colleges with varying levels of resources. In these models, $f(\cdot)$ is increasing in all of its arguments and each of its cross-partial derivatives is positive. Thus, students who are naturally bright or have well-educated parents will get more out of going to selective colleges than those with weaker endowments; furthermore, for a given student, the return on effort will be greater at selective schools because the resources there can be used

[28] In the pure signaling model, $f(\cdot)$ is equal to 0, and SEL and GPA provide information about the unobserved η.

more intensively. This complementarity between GPA and selectivity in human capital production should cause firms to more greatly reward increases in GPA for students at selective colleges. In the context of equation (16), this would imply $\theta_2 \leq \theta_3$, the opposite of the signaling model.[29] This would further imply that the selectivity premium itself, $\theta_1 - (\theta_2 - \theta_3)GPA$, is, in fact, *rising* in GPA. The estimates presented in Table 2 and Figure 7, however, easily reject these hypotheses and suggest that resource-ability complementarity models, at least by themselves, are not driving student and firm behavior.

7.1.2 More General Human Capital Models

In more general human capital models, however, the form of $f(\cdot)$ may be more flexible. Given equation (18), in order to observe a lower return on GPA at selective colleges, it must be the case that $\frac{\partial f}{\partial GPA}$ is smaller at selective colleges than at less selective colleges. Put differently, this condition states value added increases less with GPA at selective colleges than it does at less selective colleges (which, as noted earlier, is precisely the opposite of the resource-ability complementarity models). However, the evidence does not favor this hypothesis: in their survey of learning during college, Arum and Roksa (2011) do not find significant differences in the correlations of GPA with learning (as measured by the Collegiate Learning Assessment) by school selectivity.

Moreover, even granting that GPA is less correlated with value added at selective schools is insufficient for human capital to be the driving force behind the empirical results. For the selectivity premium to decline as GPA rises (Figure 7), one of two things must also be true in equation (18): either (1) ε is (more) negatively correlated with GPA at selective colleges than at less selective colleges, or (2) $f_{\eta,SEL} < 0$. The first of these is equivalent to a form of omitted variable bias—low GPA students receive large, positive productivity shocks from attending selective colleges, as could be the case if athletes and legacies (who are of lower average ability and GPA) have increased job opportunities due to teamwork skills and expanded social networks. While this cannot be fully discounted, the trend goes the wrong way in explaining the data: the share of athletes and legacies at selective colleges has been shrinking, not expanding, as these schools have increasingly sought

[29] This is the case under a pure HC model, which assumes η and the output of $f(\cdot)$ are known to the firm. It is also possible to have a hybrid human capital-signaling model in which the firm must form an expectation of $\eta + f(\cdot)$ using the observables selectivity and GPA. Under this framework, the greater return to GPA at selectives from complementarity is undercut by the opposite effect from signaling, and the results imply that the signaling effect dominates.

to admit historically disadvantaged but high ability students (Breland *et al.* 2002). The second possibility is that lower ability students actually benefit more, in the sense of greater learning, than higher ability students from the resources at selective colleges. This is diametrically opposed to the ability-resource complementarity human capital models and, if it were true, would suggest that there are tremendous net social welfare gains to be had by reallocating weaker students from less selective to more selective colleges, and stronger students from more selective to less selective colleges. As the increase in ability sorting has grown concomitantly with the information that students know about colleges and colleges know about students, such an inefficiency would represent a rather large market failure.

To summarize, while it is possible to rationalize the empirical findings under a form of the human capital model, several tenuous restrictions must be imposed, some of which are exactly counter to the prevailing human capital models in the college quality literature. In contrast, the findings flow naturally from the signaling model.

7.2 Differential Measurement Error

One concern is that there is a form of non-classical measurement error in the dependent variable (wages). Suppose the true wage is y^* but what is observed is the mismeasured wage $y = y^* + \xi$, with $\xi = -\lambda(y^* - \mu_{y^*}) + \nu$ (where ν is white noise). For $0 < \lambda \leq 1$, $cov(y^*, \xi) < 0$, or there is bias toward the mean.[30] Such a scenario might arise if survey respondents (and particularly graduates of more selective colleges) with high actual wages underreport their labor earnings (or overreport their hours), and those with low actual wages overreport their labor earnings (or underreport their hours). Bound, Brown, and Mathiowetz (2001) document that this is frequently the case in many surveys of labor market data.

Assuming that ξ is uncorrelated with the other covariates in equation (16), including ϵ, it follows that

$$plim\ \hat{\beta} = \frac{\text{cov}(y, x)}{\text{var}(x)} = \frac{\text{cov}(y^* + \xi, x)}{\text{var}(x)} = \frac{\text{cov}(y^*, x)}{\text{var}(x)} + \frac{\text{cov}(\xi, x)}{\text{var}(x)} = (1 - \lambda)\frac{\text{cov}(y^*, x)}{\text{var}(x)} = (1 - \lambda)\beta.$$

Whenever $0 < \lambda \leq 1$, this measurement error will cause attenuation bias. If this error is more

[30] Classical measurement error in the dependent variable would have $cov(y^*, \varepsilon) = 0$. It is well known that this results in consistent but less precise estimates than the case without measurement error.

severe for graduates of selective colleges, then $\lambda_{SEL} < \lambda_{NONSEL}$ and $\hat{\beta}_3$ can be less than $\hat{\beta}_2$ even if, in truth, $\beta_3 = \beta_2$. However, this situation is unlikely for at least two reasons. First, relatively greater reversion-to-mean error in reported wages among selective-college graduates should lead to relatively less variance in their observed wages. But, as shown in Table 1, the wage dispersion of this group is, if anything, larger than that of their peers who attended less selective colleges. While this does not completely rule out greater reversion-to-mean error, it implies that "true" wage dispersion among graduates of selective schools would have to be very large indeed to support the moments in the data. Furthermore, there is a second, practical reason to doubt this explanation: most of the literature on survey reporting error on income-related questions finds minimal error in reporting wage and salary income, and the measurement error that is found is almost entirely concentrated in less-educated and low-income populations, who are speculated to have lower average cognitive skills (see the reviews of Bound *et al.* 2001 and Moore, Stinson, and Welniak 2000.) It seems strange to hypothesize significant measurement error in the wages of a sample of college graduates and conjecture further that this measurement error is more severe among the graduates of selective schools, who have higher average incomes and cognitive abilities.

What of the possibility of the more familiar form of (greater) classical measurement error in the right-hand-side variable of GPA for selective college graduates? While this, too, would attenuate the estimates of β_3, it seems equally unlikely for the same reasons mentioned above. Additionally, the analysis is robust to using only observations with data sets that have transcript-derived GPAs, which should exhibit negligible measurement error differences across selectivity tiers. In short, the lower return of GPA for graduates of selective colleges does not appear to be driven by issues of measurement error.

7.3 Representativeness of Initial Jobs and Persistence of Signaling

While most prior work on the returns to college selectivity focuses on mid-career earnings, I look at the first job after college graduation, and there may be some concern that initial jobs are not representative of career earnings. While it is true that jobs change frequently in early career (Bureau of Labor Statistics 2010), the wage of the first post-graduate job still has predictive power of career earnings. For example, several studies have documented that graduating in a recession has negative effects on wages that persist for a decade or more (Kahn 2010, Oreopoulos *et al.* 2005).

But long-run correlations between early and later wage outcomes are not due solely to business cycle variation. The NLSY, 1979 cohort, shows that the log real hourly wage of the first post-graduation job is a significant predictor of the same wage measure at age 35, controlling for major, race, sex, and observation year (but not experience, which is endogenous). In fact, the partial correlation of the two wage measures is 0.20, and the partial correlation of the initial log wage with *cumulative* wage and salary income from college graduation through one's early 40s is 0.27.[31] Thus, looking at the first job after graduation can be highly informative about later career outcomes.

Second, most of the studies examining returns to college quality focus on wages in mid-career, often ten to twenty years after college graduation (see the introduction for a list). The fact that they still find earnings premia that late in worker careers, controlling for ability, would seem to be at odds with college selectivity serving as a signal. However, the persistence of a college selectivity wage premium through mid-career does not necessarily obviate the role of signaling. This persistence is consistent with labor frictions or job ladders, where entry into subsequent jobs is influenced by existing and past jobs. In such an environment, a signal of high ability may increase the chance of obtaining a high-paying initial job, and this high-paying initial job increases the chance of having a high-paying job later in one's career. The signal then, even conditional on ability, would be associated with a wage premium mid-career. In fact, Heisz and Oreopoulos (2006) find empirical support for this exact type of labor friction using data on Canadian college graduates and the types of training they receive as a function of their initial job placements. My own analysis of NLSY79 data, described in Appendix ?, finds additional evidence suggestive of job-ladder-style frictions.

7.4 Employer Learning

The signaling model in this paper can also help explain why employers appear to learn about the productivity of college graduate workers much faster than that of high school graduate workers. Arcidiacono *et al.* (2008), for example, show that while ability (AFQT) is only weakly correlated with log wages among recent high school graduates, with this correlation growing with worker experience, the ability-wage correlation among college graduates shows up immediately, with negligible growth over the career. In the context of ability signaling, this is precisely the result one would expect to find if the signals that college graduates can send are more revelatory of ability than

[31] The point estimates on first-job wages in these regressions have t-statistics between 5 and 10.

those from high school graduates. Curiously, the authors' attempt to demonstrate this supposition is relegated to a brief section in an appendix, where they regress AFQT on college entrance scores and college major and find a high R^2 (0.57 to 0.73). However, these regressions do not actually show that college graduates can better signal their ability to employers: as mentioned earlier, it is not at all clear that college entrance scores are visible to potential employers, and there is no attempt to compare signals with those of high school graduates.

I undertake such an exercise here. Specifically, using a regression similar to (16), I calculate how well the signals of college selectivity and GPA (along with college major, race and sex) can predict the standardized measure of aptitude in the pooled data. For comparison, I construct a sample of (exact) high school graduates who take wage jobs within a year of high school graduation and aren't self-employed or in the military. While college selectivity does not have a direct analogue at the high school level, high school GPA replaces college GPA as the relevant signal in this sample. Because other characteristics of the high school record may serve as signals, I include some specifications that also include quartile indicators for each of sports, leadership, and prior work experience, and the number of semesters (and their square) taken in each academic, business, and vocational subject.[32]

As the interest is in the variance of the prediction error, the relevant statistic is $\frac{1}{n}\sum \hat{\sigma}^2$, the mean squared error (or average variance of the residuals), and not R^2, which normalizes by the variance in ability. Table 4 shows the calculated mean squared error of the prediction, as well as the total variance of ability, for both the high school graduate and college graduate samples.[33] The MSE is substantially lower (about 30 percent less) among the college sample (column 1) than among the high school sample (column 2), and this difference is similar in size in both the early and late periods (panels B and C). Even with the additional potential high school signals (column 3), the MSE is larger for the high school graduates than for the college graduates. Furthermore, these additional signals seem to have less marginal predictive power in the late period relative to the early period, particularly among full-time workers (columns 5 and 6). These relative prediction errors help illustrate why employer learning is more rapid among college graduate workers than high school graduate workers: the initial signals available can more precisely pinpoint the worker's

[32] See the Data Appendix for details on the construction of these measures.

[33] In the college regressions, the partial correlation of GPA on ability is always lower, and often statistically significantly so, at selective colleges than at less selective colleges, consistent with equation (11).

ability, so there is less to be revealed through experience.

8 Conclusion

This paper formalizes and tests a model of ability signaling to explain the return to college quality that has been documented in the literature. Notably, it is the first work to both theoretically rationalize and empirically test a specific mechanism for this return. Based on data that span four decades of students, the empirical results strongly support the signaling model. Not only is the return on GPA smaller at selective schools than at less prestigious institutions, the return on selectivity itself declines as GPA, and average ability, rise. Moreover, these characteristics have strengthened over time as changes in the college and labor markets have caused ability sorting and the importance of ability in wage-setting to increase. Each of these findings is in accord with the signaling model, but not with the complementarity-based human capital models common in the theoretical college quality literature.

This evidence in favor of signaling is *not* at odds with findings of (ability-adjusted) returns to selectivity in mid-career or rapid employer learning about college-educated employee ability. First jobs matter in the presence of career ladders because they open doors; as a consequence, a medium ability student who graduated from a selective college can have better career opportunities than a high ability student who graduated from a less selective college. The literature's findings of rapid employer learning about college graduates can be explained by the role of signals, which lead employers to pre-condition on ability in their hiring decision, so that little learning through experience is expected.

The model is also appealing in that it can aid in understanding other stylized facts in the literature. For example, Bound, Hershbein, and Long (2009) document the increase in competitive behavior among high school students trying to get admitted into selective colleges, while Babcock and Marks (2010) show that study and class time among college students has declined sharply over the past 40 years. The rising return to selectivity partially brought about by increased ability sorting may help explain this apparent shift in effort from college (second stage) to high school (first stage). Because the greater degree of sorting leads to less variance in ability at selective schools and makes GPA a noisier signal there, students have less incentive to work as hard as they did previously. As the top students increasingly attend the selective colleges, the average aptitude at

less selective colleges falls, and thus so does the average effort. We would therefore expect study time to decline across the selectivity spectrum, as Babcock and Marks (2010) find.

More generally, the two-dimensional signaling framework presented here is relevant to settings other than the new college graduate labor market. For example, it could also be applied to an experienced labor market where a worker sends signals of her productivity both through the last company she worked for (the "selectivity" indicator) and her list of accomplishments while she worked there (the "GPA" measure). The general idea in this context is that a prospective employer can better infer the worker's innate productivity from where she has worked than it can from a series of bullet points playing up her contributions. This context is also appealing because it ties directly into the one described in this paper through a career ladder mechanism, magnifying the incentives faced as far back as high school for the forward-looking student.

References

[1] Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review* 30(2) January: 175-187.

[2] Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2(4) 76-104.

[3] Arum, Richard, and Josipa Roksa. 2011. *Academically Adrift: Limited Learning on College Campuses.* Chicago: University of Chicago Press.

[4] Babcock, Philip and Mindy Marks. 2010. "The Falling Time Cost of College: Evidence from Half a Century of Time Use Data." *Review of Economics and Statistics.* Forthcoming.

[5] *Barron's Profile of American Colleges, 14th ed.* 1984. Woodbury: Barron's Educational Series.

[6] *Barron's Profile of American Colleges, 19th ed.* 1992. Woodbury: Barron's Educational Series.

[7] Bedard, Kelly. 2001. "Human Capital versus Signaling Models: University Access and High School Dropouts." *Journal of Political Economy* 109(4): 749-775.

[8] Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24(3): 701-728.

[9] Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics*, vol. 5, eds. James J. Heckman and Edward Leamer. Amsterdam: North Holland.

[10] Bound, John, Brad Hershbein, and Bridget T. Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." *Journal of Economic Perspectives* 29(4): 119-146.

[11] Breland, Hunter, James Maxey, Renee Gernand, Tammie Cumming, and Catharine Trapani. 2002. *Trends in College Admission 2000.* Association for Institutional Research. Available at: `http://www.airweb.org/images/trendsreport.pdf`. Accessed on 1 October 2011.

[12] Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg. 1999. "Does it Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *Journal of Human Resources* 34(1) Winter: 104-123.

[13] Bureau of Labor Statistics. 2010. *Number of Jobs, Labor Market Experience, and Earnings Growth: Results From A Longitudinal Survey.* Available at: `http://www.bls.gov/news.release/nlsoy.toc.htm`. Accessed on 10 August 2011.

[14] Courant, Paul N., Alexandra M. Resch, and James M. Sallee. 2008. "On the Optimal Allocation of Students and Resources in a System of Higher Education." *B.E. Journal of Economic Analysis & Policy* 8(1) (Advances).

[15] Dale, Stacy B., and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4) November: 1491-1527.

[16] *Digest of Education Statistics, 2008.* 2009. Washington, D.C.: U.S. Department of Education.

[17] Epple, Dennis, Richard Romano, and Holger Stieg. 2006. "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education." *Econometrica* 74(4): 885928.

[18] Gill, Andrew M., and Duane E. Leigh. 2003. "Do the Returns to Community Colleges Differ Between Academic and Vocational Programs?" *Journal of Human Resources* 38(1): 134-155.

[19] Heckman, James, and Edward Vytlacil. 2001. "Identifying The Role Of Cognitive Ability In Explaining The Level Of And Change In The Return To Schooling." *Review of Economics and Statistics* 83(1): 1-12.

[20] Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3): 411-482.

[21] Heisz, Andrew, and Philip Oreopoulos. 2006. "The Importance of Signalling in Job Placement and Promotion." Statistics Canada Analytical Studies - Research Paper Series 11F0019MIE 236.

[22] Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Review of Economics and Statistics* 91(4): 717-724.

[23] Hoxby, Caroline M. 2001. "The Return to Attending a More Selective College: 1960 to the Present." I*n Forum Futures: Exploring the Future of Higher Education, 2000 Papers*, eds. Maureen Devlin and Joel Meyerson. Jossey-Bass. 13-42.

[24] Hoxby, Caroline M. 2009. "The Changing Selectivity of American Colleges." *Journal of Economic Perspectives* 29(4): 95-118.

[25] Kahn, Lisa B. 2010. "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy." *Labour Economics* 17(2): 303-316.

[26] Kuh, George, and Shouping Hu. 1999. "Unraveling the Complexity of the Increase in College Grades from the Mid-1980s to the Mid-1990s. *Educational Evaluation and Policy Analysis* 21(3): 297-320.

[27] Lang, Kevin, and Erez Siniver. 2011. "Why is an Elite Undergraduate Education Valuable? Evidence from Israel." NBER Working Paper 16730.

[28] Lange, Fabian. 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25(1): 1-35.

[29] Lange, Fabian, and Robert Topel. 2006. "The Social Value of Education and Human Capital." *Handbook of the Economics of Education*, Volume 1, eds. Eric Hanushek and Finis Welch. Amsterdam: North Holland.

[30] Malamud, Ofer, and Abigail Wozniak. 2008. "The Impact of College Graduation on Geographic Mobility: Identifying Education Using Multiple Components of Vietnam Draft Risk." *IZA Working Paper Series* No. 3432.

[31] McKinney, Arlise P., and Angela Miles. 2009. "Gender Differences In U.S. Performance Measures for Personnel Selection." *Equal Opportunities International* 28(2): 121-134.

[32] Moore, Jonathan C., Linda L. Stinson, and Edward J. Welniak, Jr. 2000. "Income Measurement Error in Surveys: A Review." *Journal of Official Statistics* 66(4): 331-361.

[33] Murnane, Robert, John Willett, and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77(2): 251-266.

[34] National Center for Education Statistics. 1988. *1987-88 Directory of Postsecondary Institutions: Volume 1, 4 year and 2 year*. Washington: Department of Education.

[35] Oreopoulos, Phil, Till von Wachter, and Andrew Heisz. 2006. "The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Hetereogeneity in the Market for College Graduates." Columbia University mimeo.

[36] Pascarella, Ernest T., and Patrick T. Terenzini. 2005. *How College Affects Students: A Third Decade of Research*. San Francisco: Jossey-Bass.

[37] Psacharopoulos, George. 1974. "College Quality as a Screening Device." *Journal of Human Resources* 9(4) Autumn: 556-558.

[38] Rothschild, Michael, and Lawrence J. White. 1995. "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs." *Journal of Political Economy* 10(3): 573-586.

[39] Saavedra, Juan. 2008. "The Returns to College Quality: A Regression Discontinuity Analysis." Mimeo, Harvard University.

[40] Spence, Michael. 1973 "Job Market Signaling." *Quarterly Journal of Economics* 87(3) 355-374.

[41] Toor, Rachel. 2001. *Admissions Confidential: An Insider's Account of the Elite College Selection Process* (1st ed.). New York: St. Martin's Press.

[42] Wales, Terence. 1973. "The Effect of College Quality on Earnings: Results from the NBER-Thorndike Data." *Journal of Human Resources* 8(3) 306-317.

Appendix A: Proofs

A.1: Section 3

CLAIM: Below $\tilde{\eta}$, students exert no effort in first stage.

PROOF: Follows immediately from first-order conditions in (10) and definition of $\tilde{\eta}$:

$$\frac{\alpha_2}{\tilde{\eta}+\alpha_1} + \frac{\alpha_3 \tilde{e}_1}{\tilde{\eta}+\alpha_1} = f'(\tilde{e}_1)\left(w(E[GPA_{j=1,\tilde{\eta}}], SEL_{j=1}) - w(E[GPA_{j=0,\tilde{\eta}}], SEL_{j=0})\right). \qquad \blacksquare$$

CLAIM: Above $\tilde{\eta}$, e_1^* is rising in η if marginal cost falls in ability faster than does expected marginal benefit.

PROOF: Totally differentiating (10) yields:

$$\left(\frac{-\alpha_2 - \alpha_3 e_1^*}{(\eta+\alpha_1)^2}\right) d\eta + \left(\frac{\alpha_3}{\eta+\alpha_1}\right) de_1^* = f''(e_1^*)\cdot w(\cdot)\cdot de_1^* + f'(e_1^*)\frac{\partial w(\cdot)}{\partial \eta}.$$

Rearranging and evaluating $\frac{\partial w(\cdot)}{\partial \eta}$:

$$\left[\frac{-\alpha_2 - \alpha_3 e_1^*}{(\eta+\alpha_1)^2} - f'(e_1^*)(k_1^2 - k_0^2)\frac{\gamma_2^2}{\delta_2}\right] d\eta = \left[\frac{-\alpha_3}{\eta+\alpha_1} + f''(e_1^*)\cdot w(\cdot)\right] de_1^*, \quad \text{or}$$

$$\frac{de_1^*}{d\eta} = \frac{\frac{-\alpha_2-\alpha_3 e_1^*}{(\eta+\alpha_1)^2} - f'(e_1^*)(k_1^2 - k_0^2)\frac{\gamma_2^2}{\delta_2}}{\frac{-\alpha_3}{\eta+\alpha_1} + f''(e_1^*)\cdot w(\cdot)}.$$

The denominator is strictly negative. The numerator will be negative (and the quotient positive) if and only if $-f'(e_1^*)(k_1^2 - k_0^2)\frac{\gamma_2^2}{\delta_2} < \frac{\alpha_2+\alpha_3 e_1^*}{(\eta+\alpha_1)^2}$. Note that the second term is strictly positive and $-f'(e_1^*)$ is negative. If $k_1 \geq k_0$, the quotient will always be positive. If $k_1 < k_0$, the condition binds, with the left-hand side of the inequality representing the slope of expected marginal benefit and the right-hand side the slope of marginal cost. $\blacksquare$

CLAIM: $\tilde{\eta}$ is rising in $\tilde{e}$.

PROOF: This follows from the previous claim by replacing e_1^* with $\tilde{e}$ and η with $\tilde{\eta}$. However, as $w(\cdot)$ is a function of η and not $\tilde{\eta}$, $\frac{\partial w(\cdot)}{\partial \tilde{\eta}} = 0$. The quotient is thus unambiguously positive. $\blacksquare$

A.2: Section 4

PROPOSITION 1: $E[\eta \mid j=1] - E[\eta \mid j=0] > 0$

Proof: A firm's expectation of the ability of a student who graduated from a selective college is:

$$
\begin{aligned}
E[\eta \mid j=1] &= \int_{-\infty}^{\infty} \eta P(e_1(\eta))\phi(\eta)d\eta \bigg| \; j=1 \\
&= \frac{\epsilon\Phi(\tilde{\eta})}{\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \frac{\int_{-\infty}^{\tilde{\eta}} \eta\phi(\eta)d\eta}{\int_{-\infty}^{\tilde{\eta}} \phi(\eta)d\eta} \quad + \quad \frac{\int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta}{\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \frac{\int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{\int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \\
&= \frac{-\epsilon\phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta}, \qquad \text{(A.2.1)}
\end{aligned}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and cumulative distribution functions, respectively. Similarly, the firm's expectation of ability if the student had graduated from a less selective college is:

$$
\begin{aligned}
E[\eta \mid j=0] &= \int_{-\infty}^{\infty} \eta P(e_1(\eta))\phi(\eta)d\eta \bigg| \; j=0 \\
&= \frac{(1-\epsilon)\Phi(\tilde{\eta})}{(1-\epsilon)\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty}[1-f(e_1^*(\eta))]\phi(\eta)d\eta} \frac{\int_{-\infty}^{\tilde{\eta}} \eta(1-\epsilon)\phi(\eta)d\eta}{\int_{-\infty}^{\tilde{\eta}} \phi(\eta)d\eta} \\
&\quad + \frac{\int_{\tilde{\eta}}^{\infty}[1-f(e_1^*(\eta))]\phi(\eta)d\eta}{(1-\epsilon)\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty}[1-f(e_1^*(\eta))]\phi(\eta)d\eta} \frac{\int_{\tilde{\eta}}^{\infty} \eta[1-f(e_1^*(\eta))]\phi(\eta)d\eta}{\int_{\tilde{\eta}}^{\infty}[1-f(e_1^*(\eta))]\phi(\eta)d\eta} \\
&= \frac{-(1-\epsilon)^2\phi(\tilde{\eta}) - \phi(\tilde{\eta}) - \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{(1-\epsilon)\Phi(\tilde{\eta}) + 1 - \Phi(\tilde{\eta}) - \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \\
&= \frac{-(1-\epsilon)^2\phi(\tilde{\eta}) - \phi(\tilde{\eta}) - \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{1 - [\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta]}. \qquad \text{(A.2.2)}
\end{aligned}
$$

The *difference* in expected ability from attending a more versus less selective college can be expressed as:

$$
E[\eta_1] - E[\eta_0] = \frac{-\epsilon\phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \quad + \quad \frac{(1-\epsilon)^2\phi(\tilde{\eta}) + \phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{1 - [\epsilon\Phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta]}. \qquad \text{(A.2.3)}
$$

Note that both denominators are positive by construction and that $\int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta > 0$, since $f(\cdot)$ is increasing in its argument. Thus every term in both numerators is positive, except for $-\epsilon\phi(\tilde{\eta})$; however, it was assumed that ϵ is close to zero. It therefore follows that $E[\eta_1] - E[\eta_0] > 0$. ■

Proposition 2: $\frac{\partial(E[\eta \mid j=1] - E[\eta \mid j=0])}{\partial\tilde{\eta}} > 0$

Proof: For $\epsilon \to 0$, we have:

$$
E[\eta_1] - E[\eta_0] \quad \approx \quad \frac{\int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{\int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta} \quad + \quad \frac{2\phi(\tilde{\eta}) + \int_{\tilde{\eta}}^{\infty} \eta f(e_1^*(\eta))\phi(\eta)d\eta}{1 - \int_{\tilde{\eta}}^{\infty} f(e_1^*(\eta))\phi(\eta)d\eta}. \qquad \text{(A.2.4)}
$$

An application of Leibnitz's rule shows that:

$$\frac{\partial\left(E[\eta_1]-E[\eta_0]\right)}{\partial\tilde{\eta}} = \frac{f(e_1^*(\tilde{\eta}))\phi(\tilde{\eta})\left[\int_{\tilde{\eta}}^{\infty}\eta f(e_1^*(\eta))\phi(\eta)d\eta - \tilde{\eta}\int_{\tilde{\eta}}^{\infty}f(e_1^*(\eta))\phi(\eta)d\eta\right]}{\left[\int_{\tilde{\eta}}^{\infty}f(e_1^*(\eta))\phi(\eta)d\eta\right]^2} - \frac{2\tilde{\eta}\phi(\tilde{\eta})\left(1-\int_{\tilde{\eta}}^{\infty}f(e_1^*(\eta))\phi(\eta)d\eta\right)+\tilde{\eta}f(e_1^*(\tilde{\eta}))\phi(\tilde{\eta})(1-2\phi(\tilde{\eta}))}{\left[1-\int_{\tilde{\eta}}^{\infty}f(e_1^*(\eta))\phi(\eta)d\eta\right]^2}. \tag{A.2.5}$$

The first term is unambiguously positive. Suppose $\tilde{\eta}<0$. Then the second term is unambiguously negative, and the whole expression is positive. If $\tilde{\eta}=0$, then the second term equals zero, and the whole expression is again positive. If $\tilde{\eta}>0$. ■

Proposition 3: $\frac{\partial V(\eta_{j=1})}{\partial\tilde{\eta}}<0$

Proof: For a standard normally distributed random variable η and constant $\tilde{\eta}$, $V(\eta|\eta>\tilde{\eta}) = 1-\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2+\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]$. $V(\eta_{j=1})$ is actually $k\cdot V(\eta f(\eta)|\eta>\tilde{\eta})$, where k is a positive constant that adjusts for the renormalization of the distribution of $\eta f(\eta)$ on the interval from $\tilde{\eta}$ to infinity. Since k is a constant and $f(\eta)$ is a positive-valued increasing function, the derivative of $V(\eta|\eta>\tilde{\eta})$ with respect to $\tilde{\eta}$ will have the same sign as the derivative of $k\cdot V(\eta f(\eta)|\eta>\tilde{\eta})$ with respect to $\tilde{\eta}$. It thus suffices to show that the derivative of the first variance is negative.

$$\frac{\partial V(\eta|\eta>\tilde{\eta})}{\partial\tilde{\eta}} = -2IMR^3(\tilde{\eta})+3\tilde{\eta}IMR^2(\tilde{\eta})+(1-\tilde{\eta}^2)IMR(\tilde{\eta}),$$

where $IMR(\tilde{\eta})$ is the inverse Mills ratio, $\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}$. Graphing this function reveals it to be negative for all values of $\tilde{\eta}$. ■

Proposition 4: $V(\eta_{j=1})<V(\eta_{j=0})$ if $\tilde{\eta}>0$.

Proof: First note that, because $f(\cdot)$ is increasing and maps between 0 and 1, it follows that $V(\eta_{j=1}) = V[f(e_1^*(\eta))\phi(\eta)|\eta>\tilde{\eta}] < V[\phi(\eta)|\eta>\tilde{\eta}] = 1-\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2+\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]$. Next, because some individuals with $\eta>\tilde{\eta}$ do not get admitted to the selective college and instead attend the less selective college, $V(\eta_{j=0}) = V[[(1-f(e_1^*(\eta)))\phi(\eta)|\eta>\tilde{\eta}]+[\phi(\eta)|\eta<\tilde{\eta}]] > V[\phi(\eta)|\eta<\tilde{\eta}] = 1-\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]^2-\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]$. It thus suffices to show that:

$$\begin{aligned} 1-\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2+\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right] \quad &< \quad 1-\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]^2-\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right], \quad \text{or} \\ \left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2-\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]^2-\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}+\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right] &> 0. \end{aligned}$$

Graphing this function reveals it to be positive for all values of $\tilde{\eta}>0$. ■

Appendix B: Relaxing functional form on the GPA-effort function

In Section 3, the relationship between effort and GPA, given by equation (5), assumed the same linear function for all college tiers. Here I relax that assumption by allowing the linear relationship to vary by college tier and sketch how the solution characteristics change from the canonical setup. Suppose that the GPA function is now

$$GPA_j(e_2) = \gamma_{1j} + \gamma_{2j}e_2 + \nu,$$

where the j subscript indicates that the coefficients are specific to college type. Because there exists a well-defined maximum GPA in the data (4.0), the functions should converge as effort increases, leaving two cases of interest.

Case 1: $\boldsymbol{\gamma_{11} > \gamma_{10}}$ **;** $\boldsymbol{\gamma_{21} < \gamma_{20}}$, or there is a higher intercept but smaller slope at the more selective tier. This case could correspond with greater grade inflation/compression at selective schools, as the return on effort to GPA is diminished. As indicated by equation (8), the lower slope implies a contraction of effort across the ability distribution at selective schools. On the other hand, $\frac{\partial w}{\partial GPA}$ may rise, since for a fixed change in expected GPA, there is now a larger variation in ability.[34] Thus the difference in effort distribution from the original setup is uncertain, but higher ability students still exert more effort at each school type. Functionally, this should lead to a smaller difference in the returns to GPA at the different tiers relative to the homogeneous case.

Case 2: $\boldsymbol{\gamma_{11} < \gamma_{10}}$ **;** $\boldsymbol{\gamma_{21} > \gamma_{20}}$, or there is a lower intercept but steeper slope at the more selective tier. This case could correspond with harder classes (or smarter peers) at selective schools, with more effort required to achieve the same expected grade as at less selective schools. As indicated by equation (8), the steeper slope implies an increase of effort across the ability distribution at selective schools. On the other hand, $\frac{\partial w}{\partial GPA}$ may fall, since for a fixed change in expected GPA, there is now a smaller variation in ability. Thus the difference in effort distribution from the original setup is again uncertain, but higher ability students still exert more effort at each school type. Functionally, this should lead to a larger difference in the returns to GPA at the different tiers relative to the homogeneous case.

Appendix C: Empirical Support for Model Assumptions

C.1: Linearity of GPA in effort and ability

The model in Section 3 makes a strong functional form assumption that expected GPA is linear in effort (equation (5)). With the additional assumption of normally distributed ability, optimization implies that (1) average GPA is a linear function of ability and (2) average wages are a linear function of GPA. (Both of these slopes can, and generally will, vary across selectivity tiers.) This appendix section provides empirical support for these assumptions using both graphs and statistical tests.

To demonstrate the validity of (1), Appendix Figures 1 and 2 present nonparametric estimates of GPA on the normalized senior test score for less selective colleges and for selectivity tier II.[35] Each

[34] In the absence of the error term ν, grade inflation/compression can makes grades more important to employers, since average ability levels vary more across grades. This effect will be mitigated, however, the larger is the variance of ν.

[35] The specific procedure is a local linear regression using an Epanechnikov kernel with the bandwidth that minimizes

figure has six panels: one that pools all cohorts, and one for each cohort separately. The relationship in the first panel of Appendix Figure 1, which pools all the data from less selective colleges, shows a distinct linear pattern between ability and GPA. The only appearance of strong curvature occurs at the endpoints of the ability distribution, where there are few observations and large standard errors, as shown by the shaded 95 percent confidence bands. The other panels of the figure show this pattern holds across each data set individually except for Project Talent in the 1960s, which shows a slight convex shape. Notably, this is the sole data set for which only categorical self-reported GPA is available, and aggregation effects may overly influence the nonparametric estimates. For selectivity tier II in Appendix Figure 2, the relationships are noisy, but it is easy to see that a straight line lies within each panel's confidence band. Furthermore, higher-order global polynomial specifications (beyond linear) are rejected empirically. Taken together, there seems little evidence from these graphs to call into question the assumption of linearity of GPA in ability.

While it follows from this assumption that average wages should be linear in GPA, I test this, too. I modify equations (16) and (17) to allow for selectivity-specific quadratics or cubics in GPA. Wald tests are then performed on the higher-order polynomial terms against a null of zero; a rejection would suggest that wages are not, in fact, linear in GPA. Appendix Table 3 shows the F-statistics and p-values from these Wald tests. Panel A presents pooled data, while panels B and C perform tests separately for the "early" and "late" periods.

Panel A shows that while nonlinearity does not seem to present among the sample of all workers (columns 1 through 3), there is some evidence in favor of a quadratic specification among full-time workers who graduated from less selective colleges. Specifically, the Wald tests in columns 4 and 5 can reject the null at 10 percent, though not at 5 percent. The quadratic pattern suggested by the data is convex, such that the return on GPA is rising in GPA. Tracing out the estimates, the return on GPA at less selective colleges exceeds the return at more selective colleges once GPA reaches 2.6, about half a standard deviation below the mean. Thus, even allowing this nonlinearity would not alter the conclusion that GPA returns are larger at less selective colleges.

Panels B and C show that the nonlinear GPA returns are driven entirely by the early period and actually prefer a cubic specification. (Interestingly, it is in Project Talent in the early period where evidence of a nonlinear GPA-ability relationship was found.) Tracing out the estimates in this case reveals that GPA returns are higher at less selective colleges except at the highest portion of the GPA distribution (GPA ≥ 3.5), which is relatively sparse in the early period. Therefore, this does not seem a major threat to the model assumptions, either. In sum, the linearity assumptions are empirically plausible.

C.2: Empirical densities of GPA and ability

Figures 2 and 3 show kernel density estimates of GPA across selectivity tiers for each of the five data sets used in the paper.[36] At less selective institutions, in each time period, the estimated densities appear approximately normal upon visual inspection, with a single peak, minimal skewness, and only slight truncation at the upper bound of 4. While the densities at the selective tiers are not quite as well behaved, this is somewhat expected due to their much smaller sample sizes. Still, even these densities tend to be unimodal and reasonably symmetric, the more so the larger the number of observations used to construct them.

integrated squared error. Nonparametric estimate for the other selectivity tiers are not shown for brevity but are available on request.

[36]Bandwidth is chosen according to the Sheather-Jones plug-in method with the Epanechnikov kernel.

Figures 4 and 5 show similar kernel density estimates of the senior test score measure of student ability. (I have rescaled this ability measure to have a mean of 0 and variance of 1 among the full estimation sample to better reflect the model.) As expected, dispersion in ability falls sharply as selectivity rises, and this is even more prevalent in the more recent periods, except for the NLSY97, which uses a different testing scheme (see Data Appendix). These densities, moreover, also exhibit an approximately normal distribution, even more so than the GPA densities in most cases. They are all single-peaked, show little excess kurtosis, and exhibit relatively little skewness. (The NELS densities do have slightly more pronounced left skewness, but this is at least partially an artifact of the testing instrument, which exhibited a greater degree of upper-level censoring than in earlier periods.[37])

Nonetheless, I simulated data to resemble these empirical distributions in order to examine whether the implications of bivariate normality shown in equation (11) are robust to departures from exact normality. The resulting biases in the slope and intercept terms were minimal, on the order of 2 percent, and the true parameters could not be statistically rejected. While it would be unreasonable to expect the densities of GPA and ability to be precisely normal in the data, treating them as approximately normal does not seem unreasonable.

C.3: Bounding the variance of ν

A minimum bound of the variance of ν can be estimated by using equation (9) with bounds on GPA of 1 to 4 (assuming a minimum graduation threshold of GPA equal to 1). Then the expression $\left(\frac{(\eta_i+\delta_1)\gamma_2^2 k_j}{\delta_2}\right)$ is effectively bounded between 0 and 3. With $\eta \sim N(0,1)$, fewer than 1 out of 10000 observations will take on an (absolute) value greater than 4, so with $\delta_1 = 4$, the expression $\eta_i + \delta_1$ is approximately bounded between 0 and 8. This implies that $\frac{\gamma_2^2 k_j}{\delta_2}$ has an effective upper bound of 0.375. The variance of GPA as given by (9) is:

$$V(GPA_{ij}) = \frac{\gamma_2^4 k_j^2}{\delta_2^2}\sigma_{\eta_j}^2 + \sigma_\nu^2,$$

and, in the data, this variance is approximately 0.256 at less selective schools and 0.235 at tier II schools. If $\frac{\gamma_2^2 k_j}{\delta_2} = 0.375$, then $\frac{\gamma_2^4 k_j^2}{\delta_2^2} = 0.1406$. Thus, even assuming that the variance in ability conditional on selectivity is as large as the unconditional variance of 1, the deterministic component of GPA can account for at most $\frac{0.1406*1}{0.235}$, or about three-fifths, of the overall variance, leaving at least two-fifths due to the noise term, ν. In practice, however, the fraction of variance in GPA due to the stochastic component is probably higher. For example, the observed empirical support of GPA seems to have a lower bound closer to 1.5 than 1, and there appears to be relatively minor censoring at a GPA of 4 (see Figures 2 and 3); together, these suggest that $\frac{\gamma_2^2 k_j}{\delta_2}$ has an upper bound less than 0.375 and perhaps closer to 0.25. The fraction of variance due to ν would then be on the order of 70 percent. Additionally, if $\sigma_{\eta_j}^2 < 1$, the relevance of ν rises further. The importance of the random component in explaining the variance of GPA is therefore likely substantial.

[37] This censoring does not result from the sample restriction used in this paper but is rather symptomatic of all respondents with this metric in the NELS.

C.5: Job frictions in the NLSY79

The NLSY79 provides additional evidence consistent with job-ladder-style frictions. By separately regressing wages one to two years after college graduation on major, race, sex, and observation year, I create wage residuals and divide them into five quintiles. I repeat this process for wages at nine to ten years after college graduation. (The samples are restricted to contain the same individuals in both regressions.) These residualized wage quintiles allow me to examine wage mobility by college selectivity and ability. Appendix Table 4, for example, shows two earnings matrices: one for tier II graduates from the lower three quartiles of the ability (AFQT) distribution, and one for non-tier II graduates from the top quartile of the ability distribution. It is thus possible to compare wage mobility between lower-ability students who graduated from selective colleges with higher-ability students who graduated from less selective colleges.

Among the former group, two-thirds of individuals are in the two highest residualized earnings quintiles in the first two years after graduation. Half of these (one-third of the total) are still in the top two earnings quintiles nine to ten years after graduation. For the high-ability, less-selective college grads, about half start out in the top two earnings quintiles, and 62 percent of these (31 percent of the total) are still there by nine to ten years out. While the retention rate is larger for this group than the previous one (62 percent versus 50 percent), suggesting some learning about ability is taking place, nearly identical fractions of each total group end up in the top two earnings quintiles, despite the large differences in their measured ability. For selective college grads, even if they are not of top ability, it is possible for the initial signal to still pay dividends 10 years after graduation.[38]

C.6: A comment on risk-averse agents

The model assumes students are risk neutral, but if they are uniformly risk averse, qualitatively nothing changes except effort distributions (by ability) will be compressed. Intuitively, this occurs because higher wages—and thus effort—exhibit diminishing marginal returns to utility. If risk aversion is positively correlated with ability, outcomes become ambiguous: college sorting by ability is mitigated by risk aversion in the first stage, and the GPA-ability correlation is mitigated in the second stage at less selective colleges. (Greater mixing by ability at selective colleges due to varying risk aversion makes the GPA return there ambiguous). This would generally bias against finding a selectivity premium or differences in GPA return by selectivity. On the other hand, if risk aversion is negatively correlated with ability (Shapiro et al. ????), then outcomes are qualitatively as in the risk neutral case: sorting by ability is strengthened in the first stage, and effort distribution widens in the second stage but is ability-rank preserving.

C.7: A comment on worker sorting across firms

The model assumes that all firms are homogeneous and distinguish workers by paying them different amounts based on their signals of productivity. More realistically, firms are heterogeneous and are willing to hire only workers whose expected productivity is within some band, with variations in pay of new workers quite small within a given firm (controlling for job type). Put differently, a

[38] Of course, this simple exercise does not rule out that lower ability students graduating from selective colleges systematically have something else about them that yields high wages. In the context of equation (2), this is equivalent to ε being negatively correlated with η at selective colleges. The evidence here is meant to illustrate that it is *possible* for the impact of signals to persist, not that they definitively do.

higher value of a signal does not raise a worker's pay at some fixed firm; rather, it qualifies the worker to get hired at a different company that hires higher ability workers at a higher wage. While this distinction is worth mentioning, as the treatment is imprecise in this regard, it is not important for empirical analysis. As long as workers can costlessly sort across firms, then the implications continue to hold, and firm heterogeneity of this sort is unimportant.

C.8: A comment on GPA differences between men and women

Finally, it is well-documented that there are substantial differences in GPA between men and women (Pascarella and Terenzini 2005), and this is empirically true in each of the data sets used in this study, with women averaging a 0.1 to 0.2 point advantage over men. Moreover, this advantage is roughly constant throughout the distribution except in the extreme tails. In the context of the model, this would be consistent with women and men having different intercepts but the same slope in equation 5, which would not affect their optimization. Employers presumably build this into their expectations of productivity, and this can be controlled empirically by using dummies for sex in the regressions. Of course, this assumes the same ability distribution for men and women, and this seems reasonable using senior year ability scores (although not SAT/ACTs, which are known to exhibit differences by sex).

Data Appendix

The National Center of Education Statistics (NCES) has conducted four nationally-representative, large-scale, longitudinal surveys of secondary students since 1972. Each of these surveys originally sampled between 12,000 and 25,000 students in a given grade cohort, with follow-up survey waves over the next several years. Designed to shed light on the secondary school to post-secondary school and school-to-work transitions, the surveys ask questions about demographic background, school experiences, education and work expectations, and labor market outcomes. Additionally, each survey cohort was administered a cognitive test battery. In most cases, the data variables are directly comparable across the four different surveys. Central to the analysis presented here, the restricted-access versions of these data sets allow the identification of all post-secondary institutions attended and have complete post-secondary transcript data for most students who reported attending a post-secondary institution. Because the most recent of these four surveys is too new to have data on respondents' post college-graduation transitions, I use the first three surveys, described below.

I supplement the NCES data with two additional, nationally-representative data sets that allow analysis of the new college graduate labor market in the 1960s—Project Talent—and the 2000s—the National Longitudinal Survey of Youth, 1997. These surveys cover much of the same sets of questions as do the NCES surveys, including specific colleges attended and cognitive test batteries. While self-reported cumulative GPA is available in these latter data sets, transcript data, unfortunately, are not.

NLS72

The National Longitudinal Study of the High School Class of 1972 queried approximately 17,000 high school seniors in the spring of 1972, with follow-up waves in 1973, 1974, 1976, 1979, and

1986.[39] I focus on respondents from the 1976 and 1979 waves, by which time most respondents have completed their undergraduate post-secondary education.

HSB
The High School and Beyond survey consists of two cohorts: sophomores in 1980 and seniors in 1980 (approximately 14,000 students of each). Each cohort had follow-ups in 1982, 1984, and 1986; the sophomore cohort alone had an additional follow-up in 1992. Because the 1992 follow-up is several years after the sophomore cohort was on track to graduate from college (1986), I use the senior cohort and focus on the 1986 wave.

NELS
The National Educational Longitudinal Survey began following nearly 25,000 8th graders in 1988, with follow-ups in 1990, 1992, 1994, and 2000. As these students were on track to graduate high school in 1992 and college in 1996 (under normal progression), I focus on respondents in the 2000 wave.

Project Talent
Project Talent surveyed approximately 100,000 each of 9th, 10th, 11th, and 12th graders in 1960, with follow-ups one, five, and 11 years after anticipated high school graduation.[40] I use the recently available ICPSR 1-in-4 sample of the senior cohort, as the other cohorts do not have the required job timing information necessary for analysis, and focus on the 5-year follow-up.

NLSY97
The National Longitudinal Survey of Youth, 1997 surveyed 8,984 12 to 17 year-olds beginning in 1997, with annual follow-ups. By 2009, the last data year available, respondents are aged 25 through 29. I use the geocoded version, available with application from the Bureau of Labor Statistics, and information from all available waves.

Sample Restrictions and Variable Construction

Because the five data sets differ in the timing of their follow-up interviews, care was taken to make them as consistent as possible. In each survey, the estimation sample was restricted to individuals who had earned their bachelors degree at U.S. institutions within 6 years of high school graduation, and at the time of observation had earned no additional (graduate) degree, were not currently enrolled in school, were working for pay with real (year 2005) hourly earnings between 5 and 100 dollars, and were neither self-employed nor in the military. After imposing these conditions, the final sample size consists of 2,803 individuals for NLS72; 1,078 individuals for HSB; 1,902 individuals for NELS; 2,025 in Project Talent; and 829 in NLSY97. Appendix Table 1 contains more detailed information on how the restrictions affect the sample size for each data set.

[39] As in all of the NCES surveys here, new individuals were often added in some of the later waves.
[40] Based on normal progression. Respondents were followed regardless of actual high school graduation.

College Information

College major, GPA, date of graduation, and college itself are taken from the institution from which the respondent graduated. When available, these measures come directly from the post-secondary transcript (90.5% of cases in the NLS72, 55.0% of cases in the HSB, and 94.9% in the NELS); otherwise, they are taken from self-reported information in the survey.[41] For students who attended more than one post-secondary institution before earning a bachelor's degree, GPA is based on courses taken at the degree-granting school.

While detailed college major is provided in the data, I collapse these into 11 categories that are consistent across data sets: humanities, social sciences, psychology, life sciences, physical sciences and mathematics, engineering, education, business, arts, health, and other.

When transcript data are available, GPA is calculated as the credit-weighted average of all course grades (on the standard 4 point scale) earned at the institution of graduation up to the date of degree receipt. Courses that do not receive grades (e.g., pass/fail, audits, drops, and withdrawals) are ignored in the GPA calculation. When transcript data are unavailable, self-reported GPA is used. (For observations with both measures available, the correlation between the two is 0.84 for NLS72, 0.87 for HSB, and 0.75 for NELS.). In the NLS72 and HSB, GPA is self-reported categorically (A, A-/B+, B, B-/C+, etc.) for all post-secondary courses to date (not just at the degree-granting institution). Project Talent also uses a categorical scale, although it is finer than NLS72 and HSB (A, A-, B+, B, etc.). These categories are converted to a 4 point numeric scale. NELS and NLSY97 ask respondents to report cumulative GPA as a numeric variable; NELS converts these self-reports to a 4 point scale internally, while NLSY97 provides the institution-specific grading scale; in this latter case, I performed the 4-point conversion manually.

College selectivity indicators are matched to the degree-granting college of each sample respondent using either the FICE code (NLS72, HSB, and Project Talent) or UNITID code (NELS and NLSY97) of the institution.

Alternative Selectivity Measures

While the *Barron's* rankings constitute the preferred selectivity metric due to their construction from attributes based entirely on students, as another measure of college selectivity I adopt the strategy of a quality index advocated by Black and Smith (2006). The quality index is created by applying factor analysis on five characteristics of each college: the faculty-student ratio, the rejection rate of applicants, the freshman retention rate, mean SAT/ACT score of entering freshmen, and mean faculty salaries. The factor analysis produces weights, or factor loadings, for each of these characteristics under the assumption they are each composites of some latent underlying "factors." Calling the first and most important of these factors "quality", the factor loadings allow construction of a quality index, a linear combination of the characteristics that accounts for their correlation. Using data on colleges from 1991 provided by Smith, I create the quality index for each college that has sufficient data and then compute percentiles.[42] Again, three different binary indicators for selectivity are calculated. The first of these is coded 1 if the quality index percentile is at or

[41] The much lower transcript data rate in the HSB is due to post-secondary transcripts being collected earlier in that survey (in 1984, four years after high school) relative to the others. Consequently, students who earned their degrees more than four years after high school graduation do not have complete transcript data.

[42] Data for each characteristic from before 1991 are not readily available for many colleges, which prevents it from being the preferred quality measure. However, as student characteristics evolve slowly (Black and Smith 2006), using 1991 data should still be a reasonable proxy for earlier cohorts.

above 80, and 0 otherwise (QI I); the second is coded 1 if the percentile is at or above 90, and 0 otherwise (QI II), and the third is coded 1 if the percentile is at or above 95, and 0 otherwise (QI III).[43] Of the ten highest ranked colleges by the quality index, all ten are considered to be in *Barron's* highest category in 1992, nine are in the highest category in 1982, and eight are in the highest category in 1972. (The top ten not in *Barron's* highest category 1982 or 1972 are ranked in the second-highest category.) More generally, the quality index approach is less discriminating between selectivity levels than is the *Barron's* system, but the effect is minor. Complete summary statistics using the quality index are available on request.

Ability Measures

For each data set, I construct two measures of cognitive ability: SAT/ACT percentile and (high school) senior year test score. The SAT/ACT percentile is calculated from the SAT or ACT score of the respondent as follows. For students with SAT scores, their verbal and math scores were adjusted to the re-centered scale using the College Board's concordance table[44], summed, and then converted to a percentile score using the 2005-2006 year distribution, also from the College Board.[45] For students with ACT scores (and without SAT scores), composite scores were converted to SAT equivalent scores using concordance table jointly developed by the College Board and the ACT[46] and then converted into percentiles as above. (Similar results are produced if ACT scores are converted directly into percentiles using the ACT score distribution.) SAT and ACT scores have relatively high item non-response, in part because not all valid sample respondents took either exam, and they are unavailable for the HSB sample, as they were not collected for the senior cohort. However, because the scores are mapped to a fixed distribution, this measure is comparable across time.

For each of the NCES data sets and Project Talent, the senior year test score is based on an aptitude test battery administered to students during their senior year of high school (and thus is available only for students who were surveyed during that wave.) The test batteries are similar but not identical across survey waves and are intended to measure reading comprehension, vocabulary, and mathematical knowledge. Scores are normalized to have a (population) mean of 0 and standard deviation of 1 among high school seniors within each cohort.

For NLSY97, I use the internally constructed ASVAB percentile score. About 80 percent of respondents completed the Armed Services Vocational Aptitude Battery, a 12-component test, in 1997. Based on four of these components—word knowledge, paragraph comprehension, mathematical knowledge, and arithmetic reasoning—NLSY staff computed percentile scores within three-month age groups. While not representative of high school seniors, these scores represent age-adjusted norms within cohorts.

While the senior year test and ASVAB scores are not strictly comparable across time, unlike college entrance exams, they are low-stake tests, the results for which had no direct impact on student outcomes. As such, the results reasonably capture both cognitive and non-cognitive aptitude

[43] As in the *Barron's* rankings, colleges without sufficient data to calculate a quality index are usually less selective ones. A virtue of using a binary measure for selectivity rather than a continuous one is that more colleges (and thus respondents) can be analyzed, and estimates can be compared across different selectivity measures without worrying about sample composition effects arising from the inability to cardinally rank every school.

[44] `http://professionals.collegeboard.com/data-reports-research/sat/equivalence-tables/sat-score`

[45] `http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2005/02_v&m_composite_percentile_ranks_0506.pdf`

[46] `http://professionals.collegeboard.com/profdownload/act-sat-concordance-tables.pdf`

(motivation, perseverance, etc.), which is more directly in line with the theoretical ability measure.

Job Information

Job information was taken from the first job that began after the respondent graduated with a bachelors degree except in NELS, where it was taken from the current job held at the year 2000 interview (the only post-graduation job information available.)

In NLS72, earnings data are provided at the weekly level, and hourly earnings are constructed by dividing weekly earnings at the first post-graduation job by the number of hours worked in an average week at that job. In HSB, there are data for the number of hours usually worked per week, the frequency at which one gets paid, and the rate of pay at this frequency. A majority of sample individuals report being paid annually (about 55 percent), but hourly, weekly, biweekly, and monthly are also options. In order to construct a comparable rate of pay variable, I transform the earnings variables into an hourly figure. The transformation is the identity function for hourly workers and is the rate of pay divided by the product of usual hours worked per week and the number of weeks in the frequency unit (with 4.3 weeks per month and 52 weeks per year). In NELS and Project Talent, the hourly rate of pay is constructed in a similar fashion as in HSB. For NLSY97, there is an internally constructed hourly wage variable already available. Hourly earnings in each data set are deflated to year 2005 dollars using the Personal Consumption Expenditures Deflator, and then logged.

High School Characteristics

High school GPA is taken from categorical student responses for each data set except for NELS, where it is constructed (within the data set) using high school transcript data. High school GPA is converted to a 4-point scale in a manner analogous to undergraduate GPA. Each data set has students report the number of semesters (or Carnegie units) of each academic subject taken (English, math, science, social science, and foreign language) during high school, and these are standardized to be in semester units. I also constructed (separately by data set) indices for participation in high school sports, leadership activities, and work experience based on student responses to a similar set of questions available in each data set except for NLSY97. From these indices, I generate dummies for being in each quartile, or separate dummies if the quartile measures cannot be made.

Job information for high school graduates was constructed from the same set of questions used for college graduates except that the relevant sample wave was the immediate one after scheduled high school graduation.

Figure 1: Student's First Stage Solution

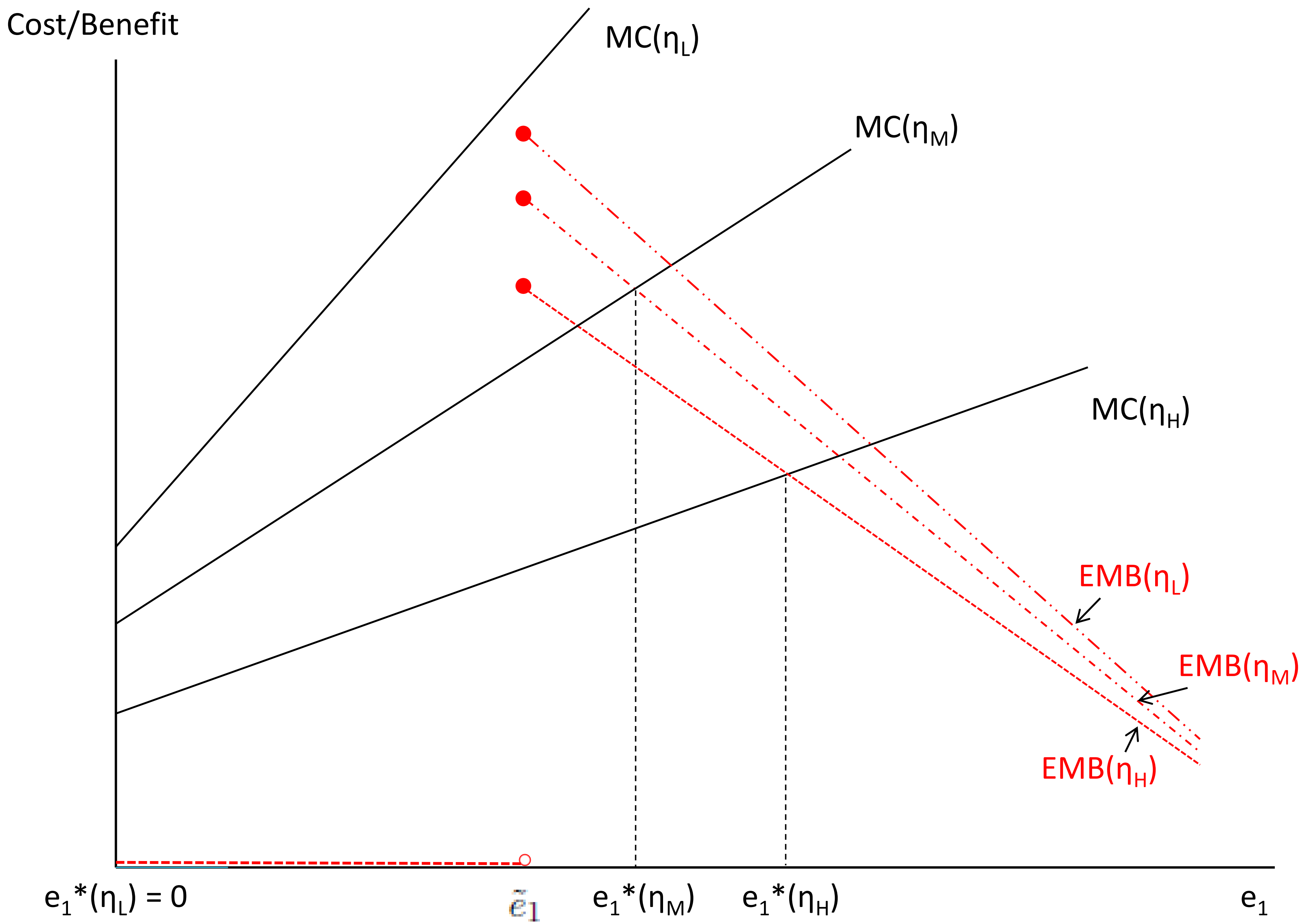
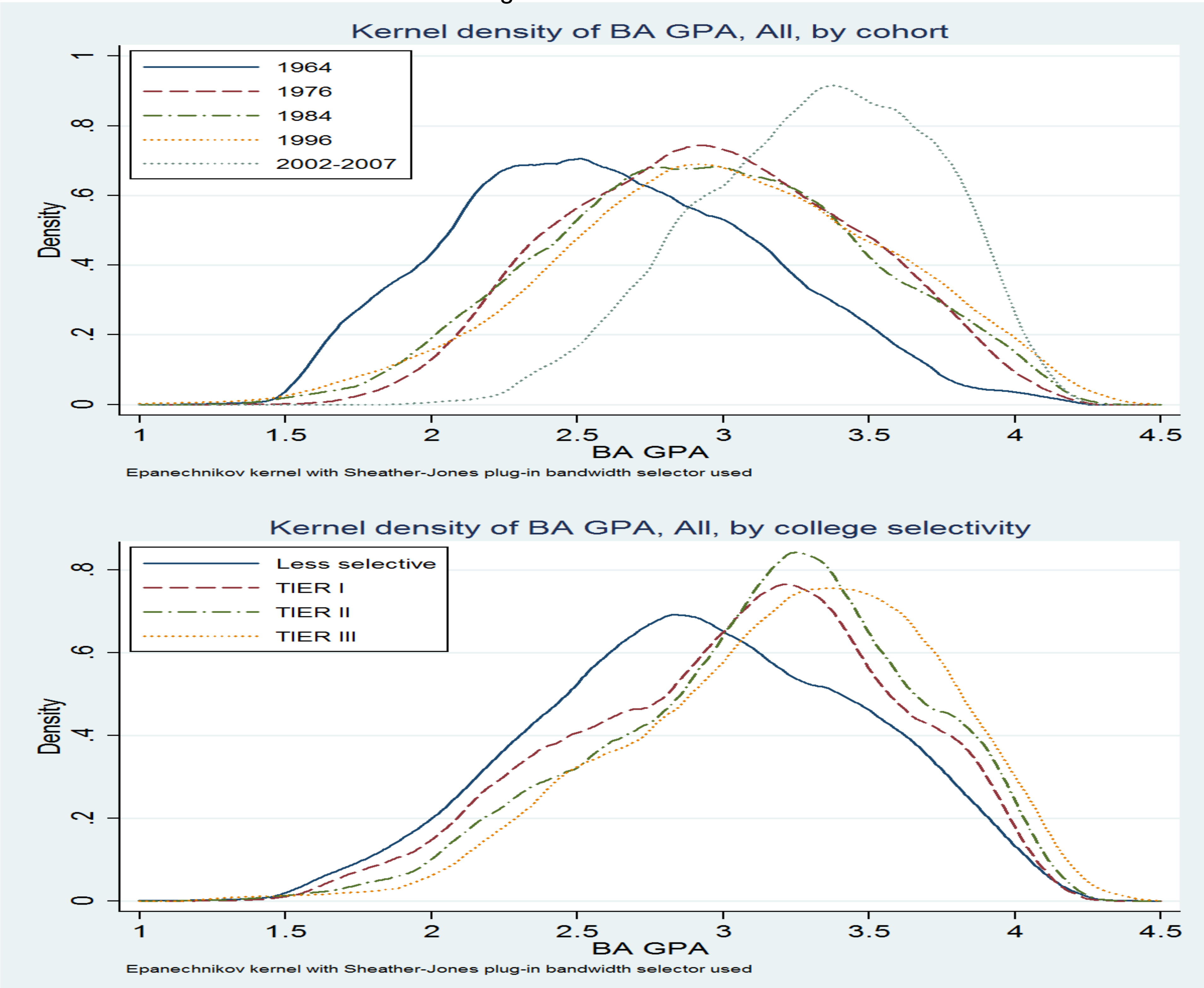


Figure 2: GPA densities



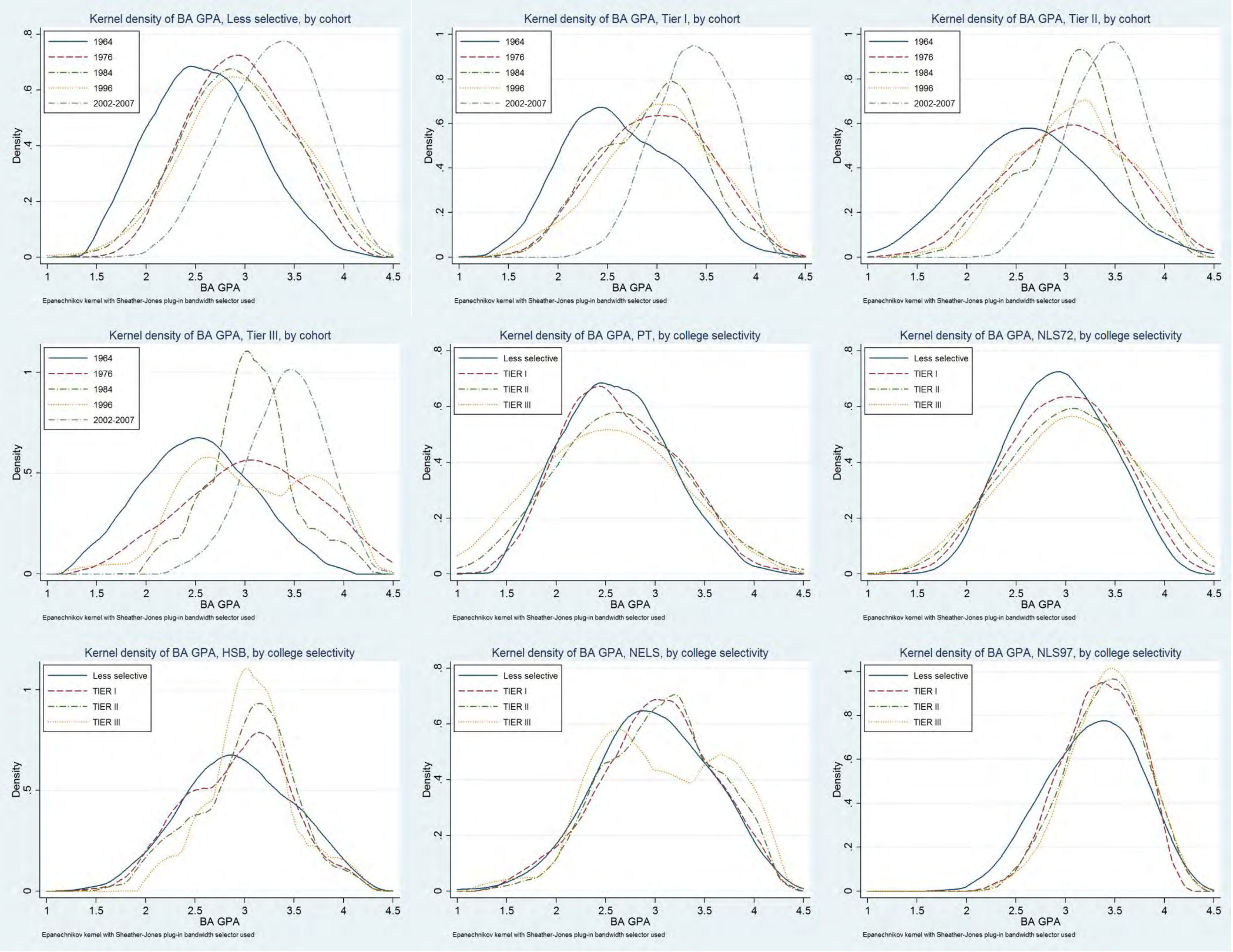
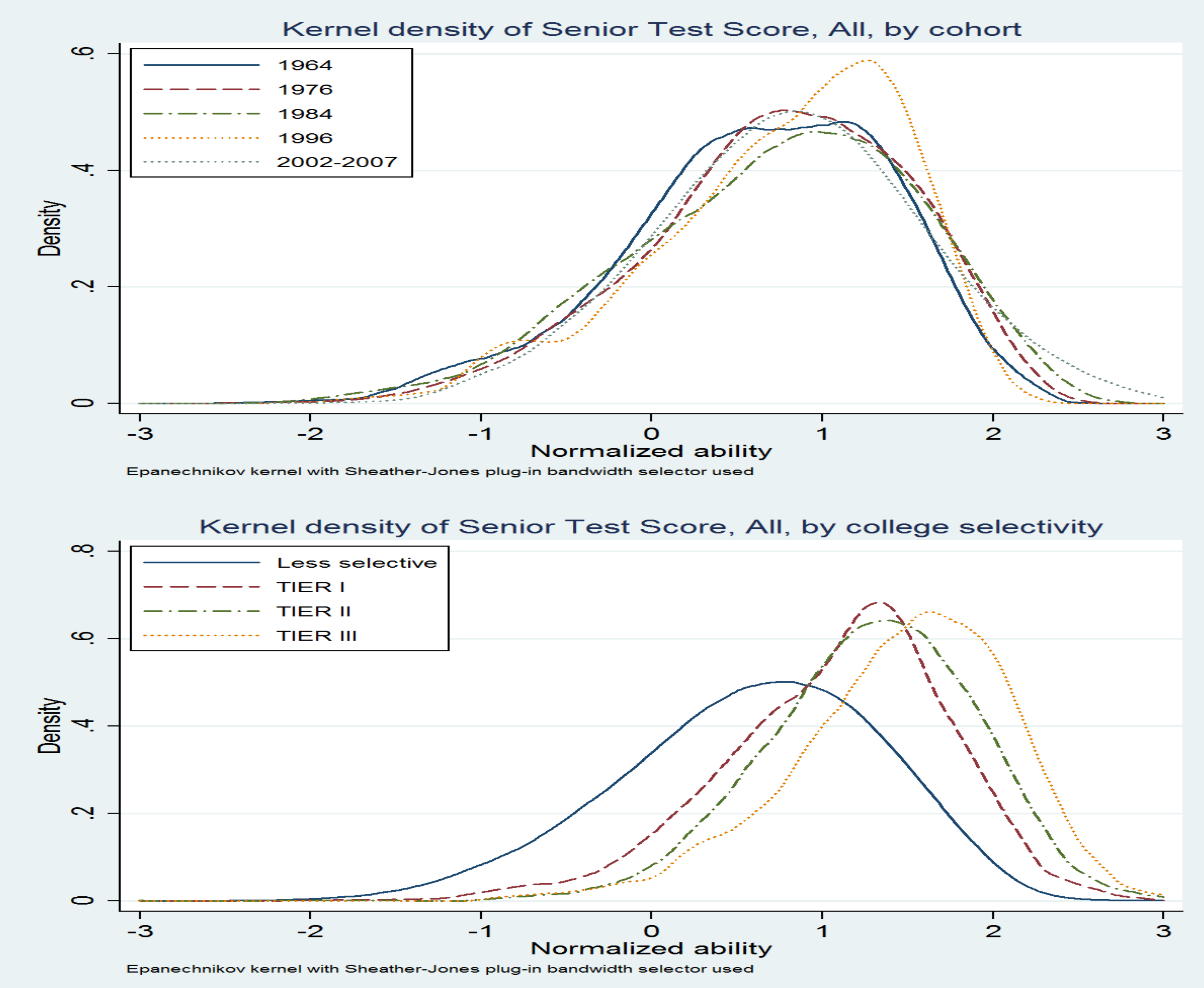


Figure 3: GPA densities by cohort and selectivity

Figure 4: Senior Test densities



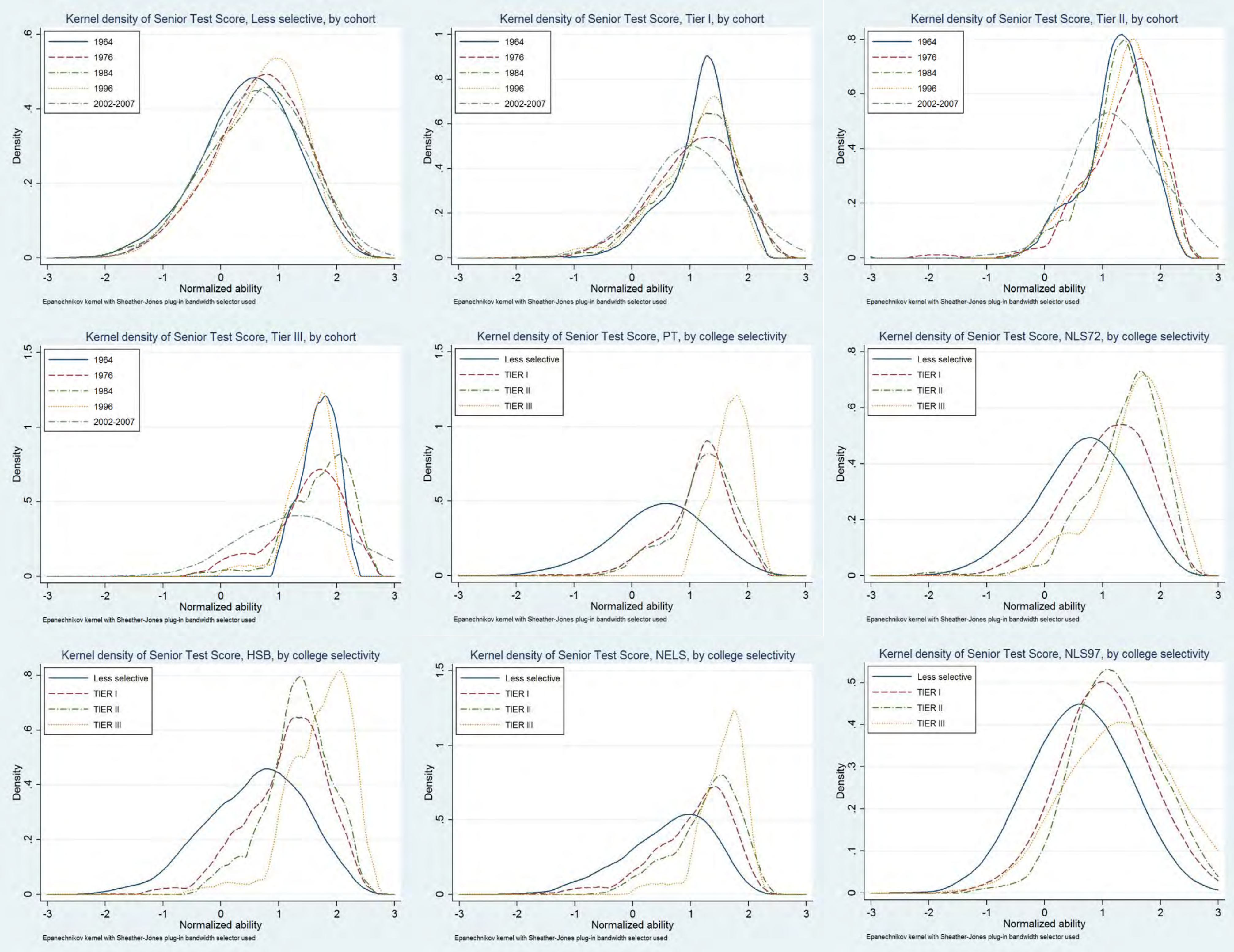


Figure 5: Senior Test densities by cohort and selectivity

Figure 6

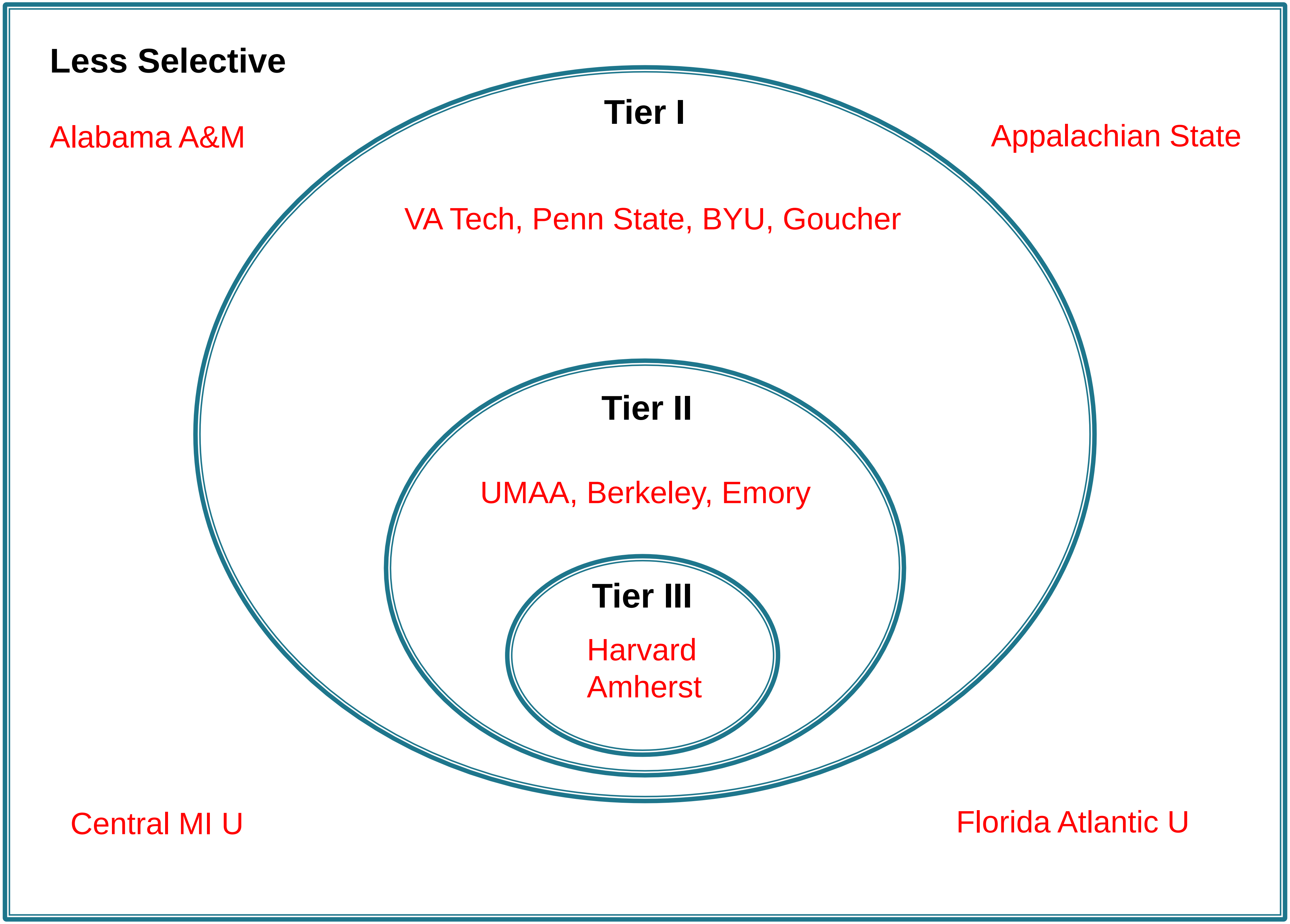
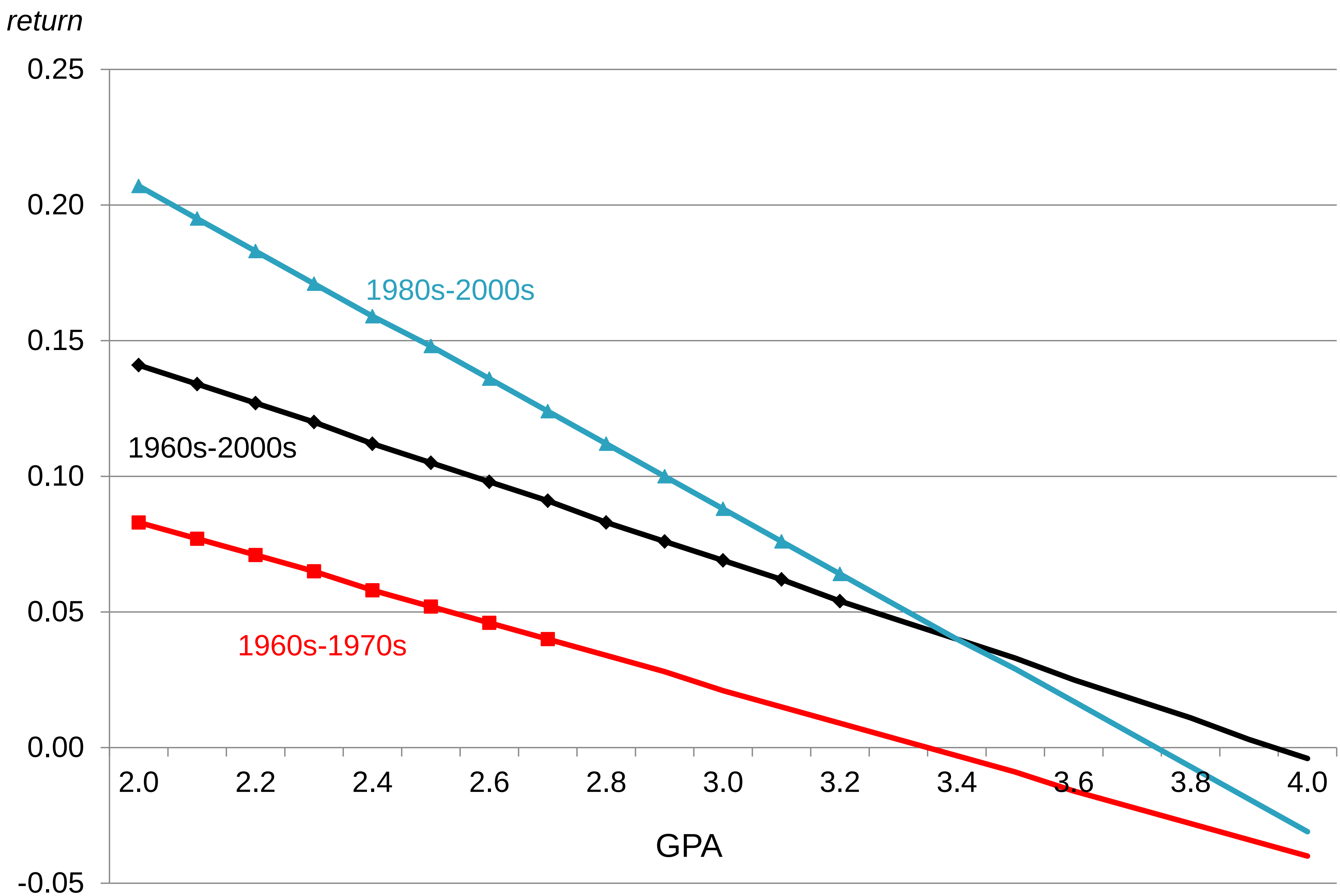

Less Selective
Alabama A&M
Appalachian State
Tier I
VA Tech, Penn State, BYU, Goucher
Tier II
UMAA, Berkeley, Emory
Tier III
Harvard
Amherst
Central MI U
Florida Atlantic U

Figure 7: Selectivity premium, by GPA
(Tier II, Full-time workers)



Notes: Line markers indicate point-wise statistical significance against a null of 0 at the 5 percent level.
The selectivity return is statistically significantly different (at 5 percent) for any two GPA values for the 1960s-2000s sample and the 1980s-2000s sample, but not the 1960s-1970s sample.

Table 1: Summary Statistics of Selected Variables

Panel A: Pooled	All		Tier I		Tier II		Tier III	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	2.966	0.509	3.051	0.505	3.134	0.485	3.232	0.437
Barron's Tier I:	0.305	0.460	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.105	0.307	0.344	0.475	1.000	0.000	1.000	0.000
Barron's Tier III:	0.031	0.174	0.103	0.304	0.299	0.458	1.000	0.000
Female	0.574	0.495	0.550	0.498	0.522	0.501	0.515	0.501
Black	0.055	0.228	0.034	0.183	0.040	0.195	0.061	0.240
Other race	0.054	0.226	0.067	0.250	0.074	0.262	0.087	0.282
Real wage ($2005)	14.48	7.204	15.58	8.360	16.40	9.810	17.20	11.010
Full-time	0.856	0.351	0.842	0.364	0.810	0.392	0.785	0.412
SAT/ACT percentile	55.6	25.3	68.0	21.5	76.4	19.0	84.2	15.6
Senior test score	0.731	0.762	1.080	0.662	1.277	0.609	1.464	0.601
Observations	8637		2404		815		231	

Panel B: Project Talent	All		Tier I		Tier II		Tier III	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	2.624	0.480	2.640	0.514	2.628	0.467	2.565	0.273
Barron's Tier I:	0.247	0.431	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.047	0.213	0.192	0.394	1.000	0.000	1.000	0.000
Barron's Tier III:	0.004	0.065	0.017	0.131	0.090	0.288	1.000	0.000
Female	0.591	0.492	0.606	0.489	0.517	0.502	0.300	0.481
Black	0.014	0.116	0.004	0.066	0.000	0.000	0.000	0.000
Other race	0.011	0.103	0.008	0.089	0.000	0.000	0.000	0.000
Real wage ($2005)	13.88	4.454	14.78	4.412	14.60	3.828	13.69	4.290
Full-time	0.924	0.265	0.930	0.255	0.911	0.286	0.715	0.260
SAT/ACT percentile	—	—	—	—	—	—	—	—
Senior test score	0.629	0.758	1.142	0.584	1.195	0.666	1.698	0.260
Observations	2025		490		122		11	

Panel C: NLS72	All		Tier I		Tier II		Tier III	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	2.955	0.478	2.981	0.502	3.012	0.525	3.043	0.503
Barron's Tier I:	0.209	0.407	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.053	0.224	0.254	0.435	1.000	0.000	1.000	0.000
Barron's Tier III:	0.009	0.094	0.043	0.203	0.170	0.377	1.000	0.000
Female	0.515	0.500	0.476	0.500	0.476	0.501	0.459	0.510
Black	0.064	0.244	0.050	0.219	0.058	0.235	0.124	0.337
Other race	0.046	0.210	0.065	0.247	0.045	0.208	0.000	0.000
Real wage ($2005)	14.42	6.857	14.71	6.776	14.94	5.779	15.15	7.219
Full-time	0.879	0.327	0.878	0.328	0.928	0.260	0.843	0.373
SAT/ACT percentile	53.9	26.2	67.6	22.4	75.9	21.9	83.2	20.0
Senior test score	0.740	0.751	1.067	0.667	1.366	0.621	1.498	0.559
Observations	2803		554		138		22	

Table 1: Summary Statistics of Selected Variables, cont'd

Panel D: HSB	*All*		*Tier I*		*Tier II*		*Tier III*	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	2.955	0.471	2.973	0.441	3.040	0.440	3.148	0.414
Barron's Tier I:	0.254	0.436	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.105	0.306	0.411	0.493	1.000	0.000	1.000	0.000
Barron's Tier III:	0.029	0.169	0.115	0.320	0.280	0.451	1.000	0.000
Female	0.575	0.495	0.559	0.497	0.580	0.496	0.414	0.500
Black	0.066	0.248	0.041	0.199	0.056	0.231	0.125	0.336
Other race	0.054	0.225	0.056	0.230	0.045	0.209	0.031	0.175
Real wage ($2005)	12.33	7.579	13.49	9.800	14.85	12.211	14.08	4.749
Full-time	0.826	0.379	0.802	0.399	0.741	0.441	0.814	0.395
SAT/ACT percentile	—	—	—	—	—	—	—	—
Senior test score	0.732	0.802	1.145	0.652	1.330	0.538	1.709	0.504
Observations	1078		264		98		33	

Panel E: NELS	*All*		*Tier I*		*Tier II*		*Tier III*	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	2.994	0.472	3.036	0.462	3.076	0.468	3.093	0.480
Barron's Tier I:	0.336	0.472	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.134	0.341	0.398	0.490	1.000	0.000	1.000	0.000
Barron's Tier III:	0.044	0.206	0.132	0.339	0.332	0.472	1.000	0.000
Female	0.576	0.494	0.515	0.500	0.463	0.499	0.490	0.502
Black	0.062	0.241	0.033	0.180	0.039	0.193	0.065	0.247
Other race	0.093	0.290	0.131	0.338	0.155	0.362	0.118	0.325
Real wage ($2005)	17.99	8.178	20.29	9.741	22.23	12.298	24.91	16.674
Full-time	0.934	0.248	0.945	0.228	0.934	0.248	0.951	0.216
SAT/ACT percentile	54.8	24.4	68.2	20.6	77.2	18.4	86.5	12.6
Senior test score	0.758	0.727	1.047	0.652	1.279	0.536	1.543	0.373
Observations	1902		717		310		109	

Panel F: NLSY97	*All*		*Tier I*		*Tier II*		*Tier III*	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GPA	3.313	0.392	3.351	0.361	3.397	0.344	3.422	0.308
Barron's Tier I:	0.483	0.500	1.000	0.000	1.000	0.000	1.000	0.000
Barron's Tier II:	0.189	0.392	0.391	0.489	1.000	0.000	1.000	0.000
Barron's Tier III:	0.071	0.258	0.148	0.355	0.378	0.487	1.000	0.000
Female	0.613	0.487	0.575	0.495	0.547	0.499	0.594	0.495
Black	0.072	0.258	0.041	0.197	0.036	0.186	0.028	0.164
Other race	0.068	0.252	0.059	0.236	0.060	0.237	0.106	0.308
Real wage ($2005)	13.74	7.128	14.15	7.399	13.90	5.736	14.04	5.006
Full-time	0.710	0.454	0.728	0.445	0.699	0.460	0.662	0.477
SAT/ACT percentile	58.6	24.9	68.2	21.6	75.9	18.3	82.2	16.9
Senior test score	0.810	0.758	1.044	0.709	1.253	0.660	1.308	0.732
Observations	829		379		147		56	

Notes: Statistics shown are weighted using sampling weights provided in the data. GPA is measured on a four point scale (0 to 4). Senior test scores follow a standard normal distribution (among high school seniors) within each dataset. The number of observations for SAT/ACT percentile and Senior test score are less than that shown, as not all sample individuals had these measures (SAT/ACT percentile unavailable in PT and HSB). See Data Appendix for variable construction.

Table 2: Log hourly wages on GPA by selectivity
(Dependent variable is real log hourly wage)

	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Sel. Dummy @ GPA=3.0	0.075***	0.097***	0.145***	0.060***	0.069***	0.128***
	[0.015]	[0.025]	[0.039]	[0.014]	[0.022]	[0.037]
GPA, less-selective	0.093***	0.093***	0.089***	0.113***	0.107***	0.103***
	[0.014]	[0.013]	[0.013]	[0.014]	[0.013]	[0.012]
GPA, selective	0.069***	0.023	0.011	0.071***	0.035	0.016
	[0.023]	[0.047]	[0.069]	[0.021]	[0.035]	[0.077]
p-val for diff	0.326	0.144	0.261	0.079	0.045	0.263
Controls for sex, race, and college major?	Yes	Yes	Yes	Yes	Yes	Yes
Full-time only?	No	No	No	Yes	Yes	Yes
Observations	8637	8637	8637	7580	7580	7580
Adjusted R-squared	0.238	0.236	0.235	0.262	0.260	0.259

Notes: Estimates shown are for OLS regressions using sampling weights and data pooled across all datasets. Standard errors (in brackets) are robust to heteroskedasticity and allow for arbitrary correlation of the error term within college. Asterisks indicate statistical significance (p<0.10, ** p<0.05, *** p<0.01).*

Table 3: Log hourly wages on GPA by selectivity
(Dependent variable is real log hourly wage)

Panel A: Pooled, early	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Sel. Dummy @ GPA=3.0	0.046**	0.021	-0.026	0.046***	0.021	0.044
	[0.020]	[0.025]	[0.048]	[0.016]	[0.025]	[0.047]
GPA, less-selective	0.051***	0.050***	0.048***	0.068***	0.064***	0.061***
	[0.016]	[0.015]	[0.015]	[0.015]	[0.014]	[0.014]
GPA, selective	0.033	0.004	-0.030	0.038*	0.002	-0.001
	[0.023]	[0.036]	[0.127]	[0.020]	[0.042]	[0.147]
p-val for diff	0.489	0.236	0.542	0.195	0.155	0.676
Panel B: Pooled, late	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Sel. Dummy @ GPA=3.0	0.094***	0.129***	0.173***	0.071***	0.088***	0.142***
	[0.021]	[0.034]	[0.048]	[0.020]	[0.029]	[0.044]
GPA, less-selective	0.135***	0.132***	0.122***	0.154***	0.146***	0.136***
	[0.022]	[0.019]	[0.020]	[0.023]	[0.020]	[0.019]
GPA, selective	0.076**	-0.004	-0.015	0.083**	0.027	0.006
	[0.036]	[0.067]	[0.075]	[0.033]	[0.044]	[0.086]
p-val for diff	0.152	0.048	0.075	0.073	0.012	0.135
p-val for diff-in-diff	0.419	0.235	0.666	0.372	0.323	0.673
Controls for sex, race, and college major?	Yes	Yes	Yes	Yes	Yes	Yes
Full-time only?	No	No	No	Yes	Yes	Yes
Observations	8637	8637	8637	7580	7580	7580
Adjusted R-squared	0.240	0.239	0.237	0.264	0.262	0.261

Notes: Estimates shown are for OLS regressions using sampling weights. Panel A shows results from the 1960s and 1970s and Panel B from the 1980s, 1990s, and 2000s. Standard errors (in brackets) are robust to heteroskedasticity and allow for arbitrary correlation of the error term within college. Asterisks indicate statistical significance (p<0.10, ** p<0.05, *** p<0.01).*

Table 4: Prediction Errors on Ability for College and High School
(Dependent variable is normalized ability measure)

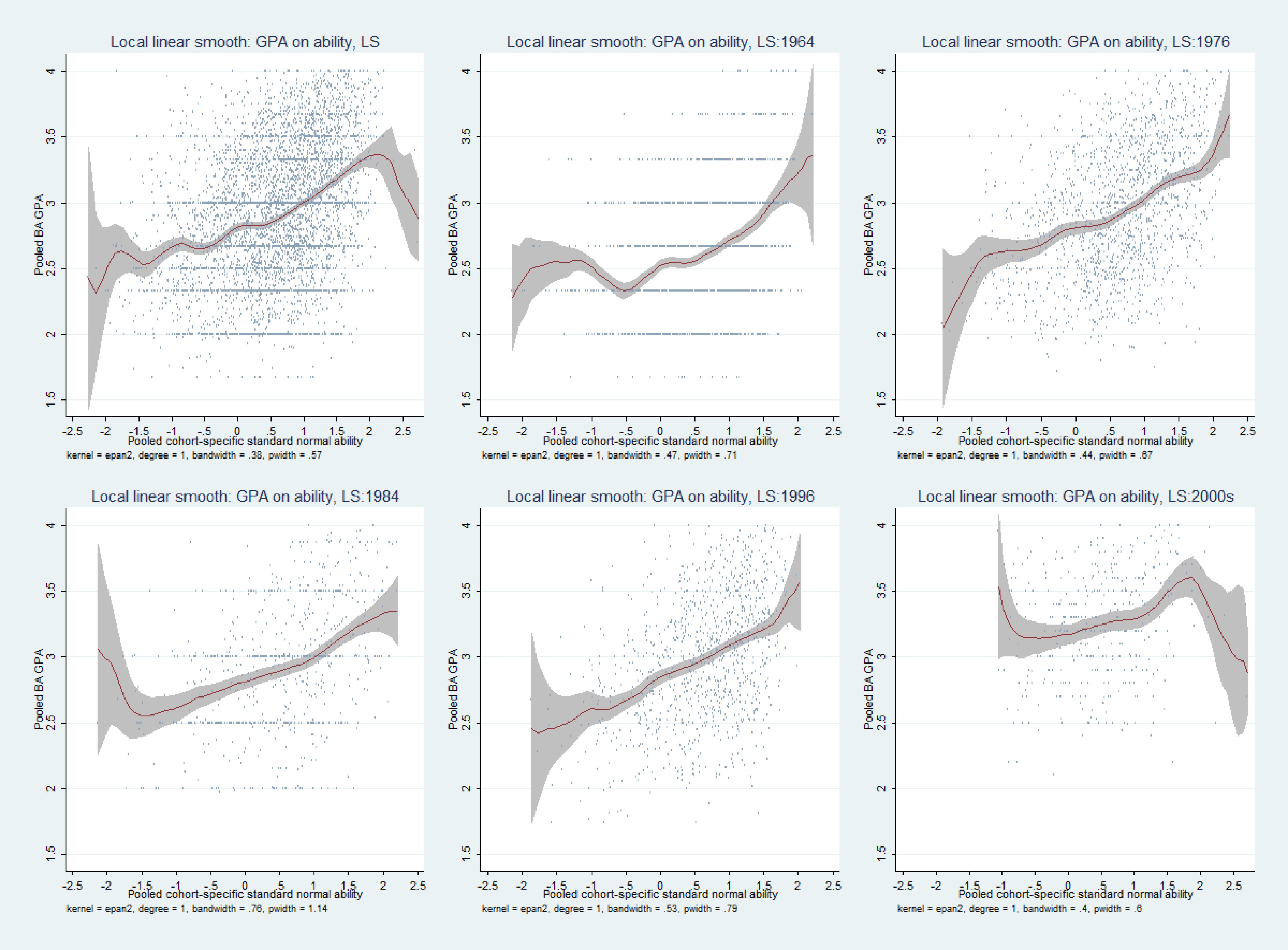
Panel A: Pooled, All	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Coll*	*HS*	*HS*	*Coll*	*HS*	*HS*
MSE	0.433	0.625	0.525	0.434	0.630	0.528
mean(ability)	0.721	-0.397	-0.397	0.724	-0.429	-0.429
var(ability)	0.579	0.821	0.821	0.574	0.789	0.789
Controls for course-taking, sports, leadership, and work	—	No	Yes	—	No	Yes
Full-time only?	No	No	No	Yes	Yes	Yes

Panel B: Pooled, early	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Coll*	*HS*	*HS*	*Coll*	*HS*	*HS*
MSE	0.431	0.624	0.505	0.435	0.608	0.496
mean(ability)	0.703	-0.465	-0.465	0.706	-0.461	-0.461
var(ability)	0.561	0.754	0.754	0.557	0.726	0.726

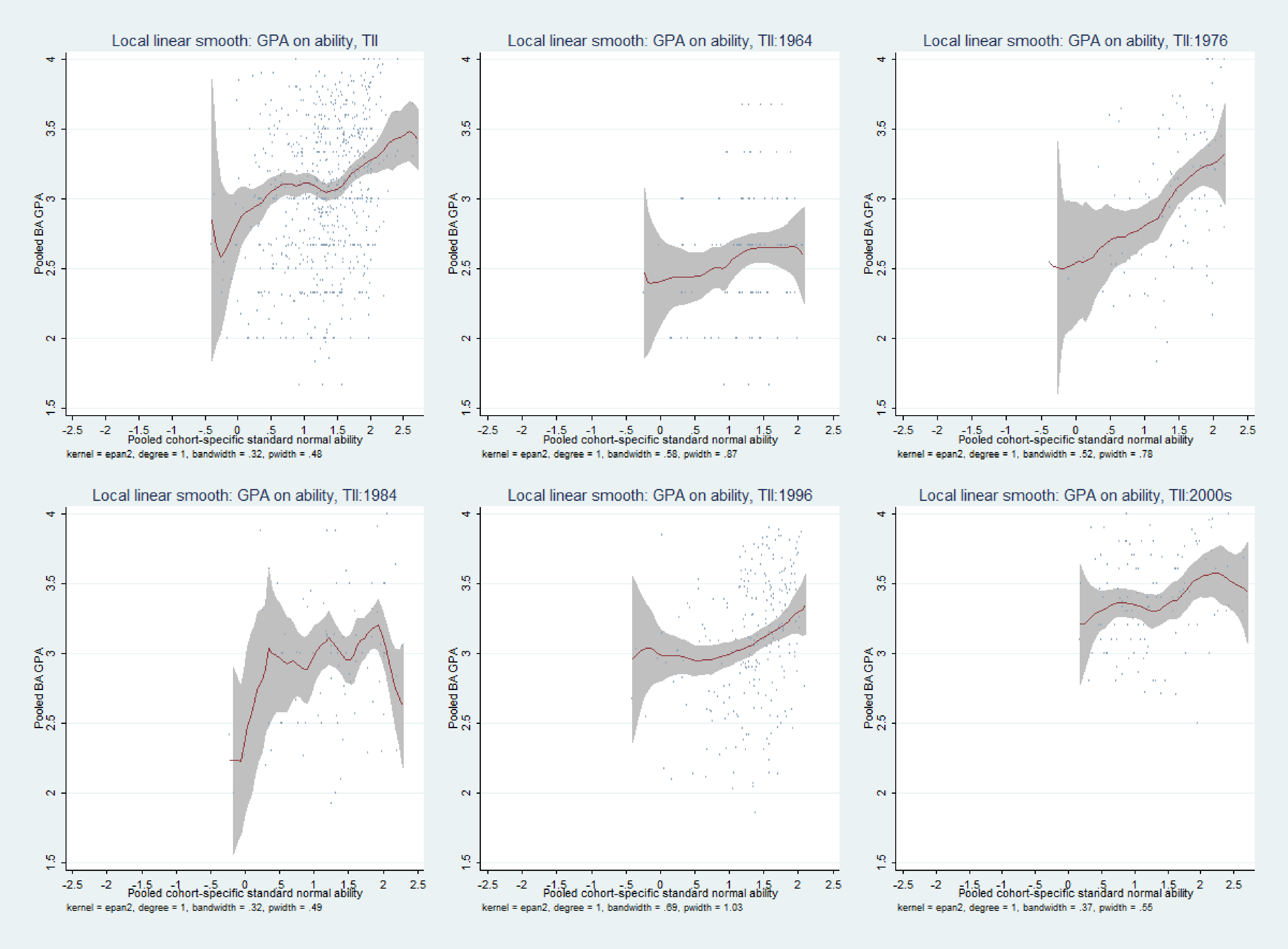
Panel C: Pooled, late	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Coll*	*HS*	*HS*	*Coll*	*HS*	*HS*
MSE	0.431	0.617	0.524	0.433	0.638	0.534
mean(ability)	0.743	-0.332	-0.332	0.747	-0.390	-0.390
var(ability)	0.600	0.877	0.877	0.594	0.865	0.865
Controls for course-taking, sports, leadership, and work	Yes	Yes	Yes	Yes	Yes	Yes
Full-time only?	No	No	No	Yes	Yes	Yes

Notes: Estimates shown are mean squared errors (MSE) and mean absolute errors (MAE) from OLS regressions of ability on signals using sampling weights. All samples are restricted to those who are working with wages. All regressions include controls for sex and race. Selectivity signals for college group also include college major, college GPA, selectivity dummy, and interactions of the selectivity dummy with college GPA. The selectivity thresholds are based on Tier II thresholds; using Tier I or Tier III thresholds produces similar results. High school signals include high school GPA and other controls as shown. Panel A shows results for all cohorts together; Panel B from the 1960s and 1970s; and Panel C from the 1980s, 1990s, and 2000s.

Appendix Figure 1: Local linear regression of GPA on ability, by cohort: Less selective



Appendix Figure 2: Local linear regression of GPA on ability, by cohort: Tier II



Appendix Table 1: Sample Sizes with Restrictions

Panel A: Unweighted	Project Talent	NLS72	HSB	NELS	NLSY97
Respondents in relevant survey wave	17,121	18,245	10,536	12,144	8,984
... who earned a BA within 6 years of HS graduation	5,364	4,362	1,874	3,676	1,610
... and who earned no post-BA degree	5,181	4,251	1,863	3,061	1,610
... and who were not enrolled in school	3,355	3,125	1,418	2,348	995
... and who were working but not self-employed or in the military	2,404	3,018	1,366	2,038	945
... and who reported real hourly wages between $5 and $100	2,100	2,818	1,118	1,913	861
... and whose GPA and graduation college were identifiable	**2,025**	**2,803**	**1,078**	**1,902**	**829**
... and who worked full-time (at least 35 hours per week)	1,835	2,464	918	1,771	592

Panel B: Weighted	Project Talent	NLS72	HSB	NELS	NLSY97
Respondents in relevant survey wave	2,509,790	3,043,599	3,024,579	3,148,608	3,875,690
... who earned a BA within 6 years of HS graduation	567,901	720,193	571,177	797,286	816,298
... and who earned no post-BA degree	548,323	700,009	567,294	671,668	816,298
... and who were not enrolled in school	357,761	515,117	443,354	529,764	509,898
... and who were working but not self-employed or in the military	254,268	497,958	423,813	456,747	481,659
... and who reported real hourly wages between $5 and $100	209,332	464,895	347,896	427,158	438,978
... and whose GPA and graduation college were identifiable	**202,695**	**462,195**	**335,178**	**425,277**	**424,140**
... and who worked full-time (at least 35 hours per week)	187,318	406,071	276,887	397,234	301,304

Sources: Author's calculations from the respective data sets

Notes: Relevant survey wave is 1965 for Project Talent, 1976 and 1979 for NLS72, 1986 for HSB, 2000 for NELS, and 2000 through 2008 for NLSY. Real wages are in year 2005 dollars, and are limited to jobs that began after graduation. The row in ***bold*** *constitutes the sample size for the main analysis. Weights are from the relevant survey wave, and for NLSY97, are averaged across five birth cohorts.*

Appendix Table 2: Log hourly wages on GPA by selectivity (*Quality Index* 1991)
(Dependent variable is real log hourly wage)

Panel A: Pooled, All	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*QI*	*QII*	*QIII*	*QI*	*QII*	*QIII*
GPA, less-selective	0.085***	0.089***	0.088***	0.107***	0.106***	0.100***
	[0.014]	[0.013]	[0.012]	[0.014]	[0.013]	[0.012]
GPA, selective	0.095***	0.068	0.052	0.087***	0.053	0.106**
	[0.024]	[0.043]	[0.084]	[0.021]	[0.033]	[0.048]
p-val for diff	0.708	0.632	0.673	0.400	0.120	0.904
Controls for sex, race, and college major?	Yes	Yes	Yes	Yes	Yes	Yes
Full-time only?	No	No	No	Yes	Yes	Yes
Observations	8637	8637	8637	7580	7580	7580
Adjusted R-squared	0.241	0.235	0.237	0.264	0.260	0.260

Panel B: Pooled, early	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*QI*	*QII*	*QIII*	*QI*	*QII*	*QIII*
GPA, less-selective	0.055***	0.052***	0.048***	0.074***	0.064***	0.061***
	[0.016]	[0.015]	[0.015]	[0.016]	[0.015]	[0.014]
GPA, selective	0.019	0.004	0.023	0.017	0.019	0.044
	[0.023]	[0.027]	[0.038]	[0.018]	[0.031]	[0.040]
p-val for diff	0.137	0.108	0.519	0.008	0.164	0.671

Panel C: Pooled, late	(1)	(2)	(3)	(4)	(5)	(6)
Selectivity Tier	*QI*	*QII*	*QIII*	*QI*	*QII*	*QIII*
GPA, less-selective	0.127***	0.132***	0.121***	0.145***	0.145***	0.131***
	[0.021]	[0.020]	[0.019]	[0.022]	[0.020]	[0.019]
GPA, selective	0.094**	-0.002	-0.031	0.103***	0.015	0.081
	[0.039]	[0.066]	[0.151]	[0.034]	[0.045]	[0.070]
p-val for diff	0.447	0.049	0.317	0.289	0.008	0.478
p-val for diff-in-diff	0.944	0.238	0.409	0.748	0.127	0.650
Controls for sex, race, and college major?	Yes	Yes	Yes	Yes	Yes	Yes
Full-time only?	No	No	No	Yes	Yes	Yes
Observations	8637	8637	8637	7580	7580	7580
Adjusted R-squared	0.245	0.240	0.241	0.268	0.264	0.263

Notes: Estimates shown are for OLS regressions using sampling weights. Panel A shows results for all cohorts together; Panel B from the 1960s and 1970s; and Panel C from the 1980s, 1990s, and 2000s. Standard errors (in brackets) are robust to heteroskedasticity and allow for arbitrary correlation of the error term within college. Asterisks indicate statistical significance (p<0.10, ** p<0.05, *** p<0.01).*

Appendix Table 3: Wald Tests of Nonlinearity of Wages in GPA
(Dependent variable is normalized ability measure)

Panel A: Pooled, All	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Less-selective, quadratic	0.86	1.14	1.06	3.07	3.03	2.11
	[0.354]	[0.284]	[0.303]	[0.080]	[0.082]	[0.147]
Selective, quadratic	0.36	0.15	0.26	0.09	0.00	0.71
	[0.549]	[0.696]	[0.610]	[0.763]	[0.992]	[0.398]
Less-selective, cubic	0.46	0.96	0.66	1.51	1.90	1.13
	[0.634]	[0.385]	[0.520]	[0.221]	[0.149]	[0.324]
Selective, cubic	0.61	0.33	0.73	0.91	0.13	0.92
	[0.545]	[0.721]	[0.483]	[0.401]	[0.879]	[0.398]
Full-time only?	No	No	No	Yes	Yes	Yes

Panel B: Pooled, early	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Less-selective, quadratic	0.43	0.12	0.33	0.79	0.55	0.93
	[0.511]	[0.725]	[0.567]	[0.376]	[0.457]	[0.335]
Selective, quadratic	0.17	0.86	0.39	0.03	1.12	0.03
	[0.680]	[0.354]	[0.534]	[0.875]	[0.290]	[0.873]
Less-selective, cubic	3.18	1.35	1.58	4.42	3.31	3.39
	[0.042]	[0.260]	[0.206]	[0.012]	[0.037]	[0.034]
Selective, cubic	0.23	0.44	0.29	0.17	0.62	0.87
	[0.791]	[0.644]	[0.751]	[0.846]	[0.538]	[0.421]

Panel C: Pooled, late	(1)	(2)	(3)	(4)	(5)	(6)
Education Group	*Tier I*	*Tier II*	*Tier III*	*Tier I*	*Tier II*	*Tier III*
Less-selective, quadratic	0.05	0.00	0.00	0.40	0.30	0.02
	[0.826]	[0.980]	[0.947]	[0.529]	[0.582]	[0.878]
Selective, quadratic	0.50	0.21	1.12	0.00	0.03	1.14
	[0.479]	[0.650]	[0.290]	[0.955]	[0.860]	[0.285]
Less-selective, cubic	0.21	0.09	0.01	1.03	0.16	0.10
	[0.808]	[0.915]	[0.985]	[0.357]	[0.853]	[0.906]
Selective, cubic	0.92	0.39	0.80	0.92	0.06	0.83
	[0.400]	[0.680]	[0.450]	[0.397]	[0.944]	[0.438]

Notes: Estimates shown are F statistics (and p-values in brackets) from Wald tests for whether the coefficients on higher-order polynomial terms in GPA are equal to a null of zero. See Table 3 for other notes.

Appendix Table 4: NLSY79 Wage Mobility, By Ability and Selectivity

Residualized earning quintiles matrix (percent): 1st-3rd AFQT quartiles, Tier II

1-2 yrs out ↓\ 9-10 yrs out →	*Q1*	*Q2*	*Q3*	*Q4*	*Q5*
Q1	4.2	0.0	0.0	4.2	4.2
Q2	0.0	0.0	0.0	4.2	4.2
Q3	0.0	4.2	0.0	8.3	0.0
Q4	0.0	12.5	4.2	4.2	0.0
Q5	4.2	8.3	4.2	4.2	25.0

Residualized earning quintiles matrix (percent), 4th AFQT quartile, NOT Tier II

1-2 yrs out ↓\ 9-10 yrs out →	*Q1*	*Q2*	*Q3*	*Q4*	*Q5*
Q1	3.1	5.2	2.1	2.6	1.0
Q2	4.2	4.7	5.2	2.6	1.6
Q3	1.6	6.3	3.1	3.7	2.6
Q4	5.2	2.1	4.7	8.3	6.8
Q5	1.0	1.0	5.2	6.3	9.9

The matrices reflect the distribution of residualized earnings, by quintiles, one to two years after college graduation (rows) and nine to ten years after graduation (columns). The top matrix presents the distribution for graduates of selective colleges from the lower three quartiles of the AFQT distribution, and the bottom matrix presents the distribution for graduates of less selective colleges from the top quartile of the AFQT distribution. Residualized earnings are created by regressing log wages of workers on dummies for sex, race, college major, and observation year among a sample of college graduates working in both specified time horizons.