Three Essays on Labor Market Entry

by

Brad J. Hershbein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2012

Doctoral Committee:

Professor John Bound, Chair
Professor Brian A. Jacob
Professor Jeffrey A. Smith
Assistant Professor Martha J. Bailey

Dissertations should not be undertaken lightly, as the writing of one almost inevitably causes hardship not just for the writer but for the writer's friends, family, advisors, and anyone else patient enough to put up with such a silly person. To all of you: thanks for putting up with me.

ACKNOWLEDGEMENTS

Chapter 1: 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.

Chapter 2: I thank Martha Bailey, John Bound, Melissa Kearney, Justin McCrary, Ryan Michaels, Jeff Smith, Gary Solon, three anonymous referees, and participants of the seminars at the University of Michigan, the Federal Reserve Bank of Boston, and the annual meetings of the Southern and Midwest Economic Associations and Population Association of America for helpful comments. Chapter 2 appears in *B.E. Journal of Economic Analysis and Policy*, volume 12, number 1, published in 2012.

Chapter 3: The research in this paper, joint with Martha Bailey and Amalia Miller, was conducted while the authors were Special Sworn Status researchers of the U.S. Census Bureau at the Michigan Census Research Data Center. Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Census Bureau. This paper has been screened to ensure that no confidential data are revealed. We are grateful to Jim Davis, Maggie Levenstein, Stan Sedo, and Clint Carter for extensive support with the preparation and revision of the restricted data proposal and disclosure process and to John Bound, Charlie Brown, John DiNardo, Claudia Goldin, David Lam, Bob Margo, Paul Rhode, Jeff Smith, Gary Solon, Sarah Turner, and three anonymous referees for helpful comments and suggestions. Chapter 3 will appear in *American Economic Journal: Applied Economics* in 2012.

The research was generously supported by a grant from the Marshall Weinberg Research Fellowship, a core grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) to the Population Studies Center at the University of Michigan (R24 HD041028), which provided generous financial support for work in the University of Michigan Research Data Center (RDC), and an NICHD Population Studies Center Trainee grant (T32 HD0007339). All errors are mine.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Three Essays on Labor Market Entry
by
Brad J. Hershbein

Chair: John Bound

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, in Chapter I 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.

Chapter II explores how high school graduate men and women vary in their behavioral responses to beginning labor market entry during a recession. In contrast with previous related literature that found a substantial negative wage impact but minimal employment impact in samples of highly educated men, the empirical evidence presented here suggests a different outcome for the less well educated, and between the sexes. Women, but not men, who graduate high school in an adverse labor market are less likely to be in the workforce for the next four years, but longer-term effects are minimal. Further, while men increase their enrollment as a short-run response to weak labor demand, women do not; instead, they appear temporarily to substitute into home production. Women's wages are less affected than men's, and both groups' wages are less affected than the college graduates previously studied.

Decades of research on the U.S. gender gap in wages describes its correlates, but little is known about why women changed their career paths in the 1960s and 1970s. Chapter III, joint with Martha J. Bailey and Amalia R. Miller, investigates the role of "the Pill" in altering womens human capital investments and its ultimate implications for life-cycle wages. Using state-by-birth-cohort variation

in legal access, we show that younger access to the Pill conferred an 8 percent hourly wage premium by age fifty. Our estimates imply that the Pill can account for 10 percent of the convergence the gender gap in the 1980s and 30 percent in the 1990s.

CHAPTER I

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

1.1 Introduction

Recently, there has been a sizable interest in the return to attending a more selective or prestigious college. Several studies have tried to identify empirically the private returns to going to a selective school, with most finding that attending a more prestigious school does indeed have a causal, positive impact on lifetime earnings. However, there has been little attention as to *why*. 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] The goal of this paper is to propose a specific mechanism for the college selectivity premium—a model of signaling—that can rationalize observed behavior.

Several factors make signaling in particular a compelling explanation for the premium. First, the relatively few studies that have attempted to measure student learning in college have found little difference across types of colleges once pre-college characteristics are controlled for (Pascarella and Terenzini 2005; Arum and Roksa 2011). While it is not clear how the "learning" measured in these studies relates to productivity on the job, this evidence suggests colleges may boost the wages of their graduates in ways other than through value added. Second, the growing literature on how employers learn about worker productivity has emphasized that this process is not immediate but occurs over time, with employers often attempting to learn about an applicant's latent ability through measures that are immediately observable, such as education or race. In this context, as the share of the labor force that are college graduates has risen, it seems reasonable that firms would

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

sort workers not just through quantity of education but through perceived measures of quality of education, as well. Finally, and related, human resources and cognitive psychology surveys have documented that recruiters looking to hire new college graduates not only actively screen applicants by college attended and grade point average, but that these measures positively correlate with on-the-job performance (McKinney and Miles 2009). Together, these findings point to the importance of examining how college selectivity and college grades are jointly determined and how employers use these measures in wage setting.

This paper makes two 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 graduating 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.

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 1.3 and 1.4 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 1.5 describes the data sets and empirical methodology that are used to explore and test the implications of the model, while Section 1.6 presents the results of these tests. The last section concludes.

1.2 A Literature Review

The earliest studies attempting to measure the return to college selectivity or quality in the U.S. context 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. Perhaps the most credible identification comes from Hoekstra (2009), who employs regression discontinuity designs based on a test cutoff for admission to a (specific) selective college. He finds a larger premium than in previous work. Dale and Krueger (2002) are 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 employers learn about the underlying productivity of workers relatively rapidly, especially in the case of 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, as credible identification is elusive.

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.

1.3 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.[2] 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). Below, I lay out a model that illustrates how signaling can affect the interactions of these college graduates and firms.

In order to focus on the behavior of students, I assume that firms are homogeneous. Prospective workers (i.e., students), on the other hand, vary in their ability, $\eta \sim N(0,1)$, and this trait affects the worker's productivity to firms.[3] While students can observe their own ability, the firms cannot. Instead, in the spirit of Spence (1973), the firms observe imperfect signals of ability that are chosen

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

[3] "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.

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.[4]

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.

1.3.1 Firm's Problem

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

$$\ln y_{it} = a_{it} + \rho_t \eta_{it} + \varepsilon_{it}, \tag{1.1}$$

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 firm but not the econometrician.

The objective of the firm is to set a wage policy in order to maximize expected profits from a

[4] 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.

new college graduate worker. Competition among firms, however, ensures that profits are zero in expectation, and so

$$w_{it}(GPA_{it}, SEL_{it}) = a_{it} + \rho_t E[\eta_{it} \mid GPA_{it}, SEL_{it}] + \varepsilon_{it}, \tag{1.2}$$

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.

1.3.2 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:

$$P(e_1) = \begin{cases} \epsilon & \text{if } e_1 < \tilde{e}_1 \\ f(e_1);\ f'(e_1) > 0,\ f''(e_1) < 0,\ \lim_{e_1 \to \infty} f(e_1) = 1 & \text{if } e_1 \geq \tilde{e}_1. \end{cases} \tag{1.3}$$

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.[5] 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).

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

[5] 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.

ability. The cost of high school effort is given by

$$C_1(e_1) = \frac{\alpha_2}{\eta + \alpha_1} e_1 + \frac{\alpha_3}{2(\eta + \alpha_1)} e_1^2, \tag{1.4}$$

where α_1, α_2, and α_3 are each positive constants.[6]

In the second stage, the student has observed the admission outcome and knows what type of college she will attend.[7] 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,

$$GPA(e_2) = \gamma_1 + \gamma_2 e_2 + \nu; \qquad \nu \sim N(0, \sigma_\nu^2), \tag{1.5}$$

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 (1.5) 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.[8] Second, the GPA function is independent of college type. It turns out the qualitative implications of the model are not affected by this restriction (see Appendix 1.B), and so I proceed for now under (1.5).

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

$$C_2(e_2) = \frac{\delta_2}{2(\eta + \delta_1)} e_2^2, \tag{1.6}$$

where δ_1 and δ_2 are each positive constants.[9]

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

$$Max_{e_1, e_2} \quad U_i = w\left(SEL(e_1), GPA(e_2)\right) - C_1(e_1; \eta) - C_2(e_2; \eta), \tag{1.7}$$

where w is the log wage earned conditional on GPA and SEL, an indicator variable for whether

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

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

[8]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.

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

$j = 1$, and the η subscripts in the cost functions reflect their dependence on a student's ability.[10]

1.3.3 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:

$$e^*_{2j} = \frac{(\eta + \delta_1)\gamma_2}{\delta_2} \cdot \frac{\partial w(\cdot)}{\partial GPA}\bigg|_{SEL=j}. \tag{1.8}$$

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.[11] 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 (1.5) yields:

$$GPA_{ij}\left(e^*_{2j}(\eta_i)\right) = \gamma_1 + \left(\frac{(\eta_i + \delta_1)\gamma_2^2}{\delta_2} \cdot \frac{\partial w(\cdot)}{\partial GPA}\bigg|_{SEL=j}\right) + \nu, \quad \text{or} \tag{1.9}$$

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

under the assumption that $\frac{\partial w(\cdot)}{\partial GPA}\Big|_{SEL=j}$ is a constant k_j. (I discuss the empirical validity of this assumption, as well as the linearity of GPA in ability, in Appendix 1.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:

$$(w(E[GPA_{j=1,\eta}], SEL_{j=1}) - w(E[GPA_{j=0,\eta}], SEL_{j=0})) \cdot \frac{dP}{de^*_1} \leq \frac{dC_1}{de^*_1}, \quad \text{or} \tag{1.10}$$

[10] Equation 1.7 assumes students are risk neutral. In Appendix 1.C, I briefly sketch how behavior changes when agents are risk-averse.

[11] 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.

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

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.

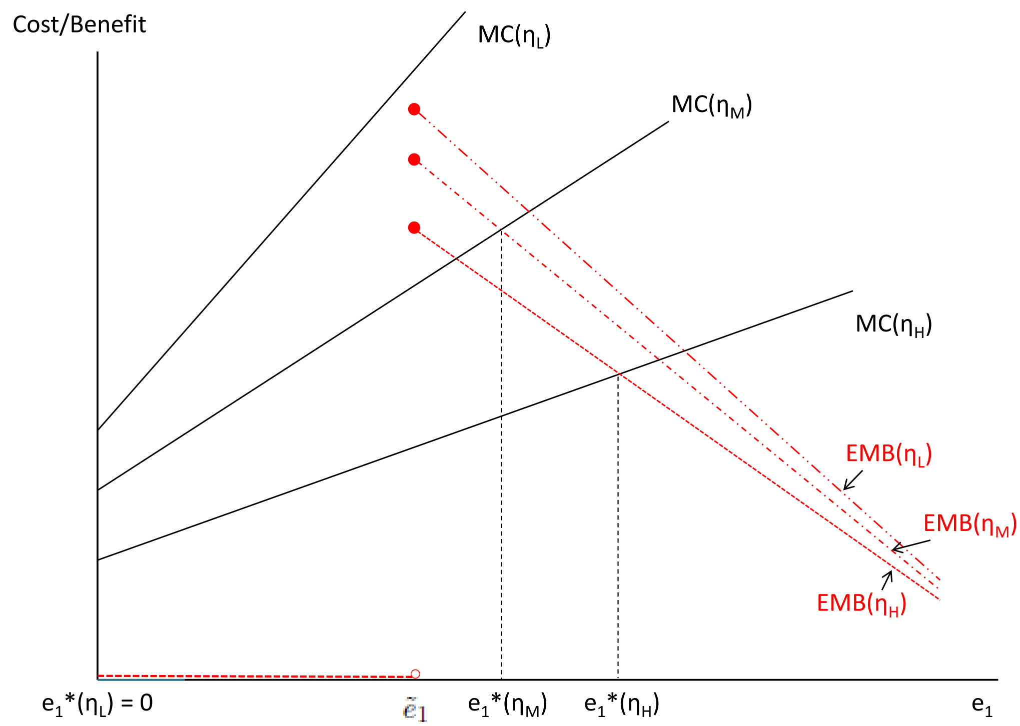
The solution can perhaps best be explained graphically, as in Figure 1.1. For the sake of exposition, the figure plots marginal cost and benefit curves for three ability types: high (η_H), medium (η_M), and low (η_L). Equation (1.4) implies that that marginal cost of effort has both the slope and intercept 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,[12] 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

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

Figure 1.1: Student's First State Solution



implicitly defined by (1.10)) 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.[13] Third, the threshold $\tilde{\eta}$ is rising in $\tilde{e}$. (Appendix 1.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.

1.4 Firm Expectations of Student Ability and Predictions

1.4.1 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] Marginal cost must decline in ability faster than does the wage premium from the endogenous 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.[14]

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 1.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:

$$E[\eta_i \mid GPA_{ij}, SEL_{ij}] = E[\eta_i \mid SEL] + \frac{Cov(\eta, GPA_j)}{\sigma^2_{GPA_j}} \left(GPA_{ij} - \mu_{GPA_j}\right). \tag{1.11}$$

The conditional expectation of ability given both selectivity and GPA is linear in GPA, with both the slope and intercept varying by selectivity tier.[15]

[14]This assumes the factors that brought about the change in $\tilde{\eta}$ were exogenous; see Hoxby (2009) and section 1.4.3 below for evidence to this effect.

[15]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

It follows from equation (1.2) that log wages at a given time (t subscript suppressed) are given by:

$$w_{ij}(GPA_{ij}, SEL_{ij}) = a_i + \rho\left(\psi_j + \frac{(\gamma_2^2\delta_2^{-1}k_j)\sigma_{\eta_j}^2}{(\gamma_2^4\delta_2^{-2}k_j^2)\sigma_{\eta_j}^2 + \sigma_\nu^2}GPA_{ij}\right) + \varepsilon_i, \tag{1.12}$$

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.[16] The return to GPA on log wages is thus:

$$\frac{\partial w_{ij}}{\partial GPA_{ij}} = \frac{\rho\gamma_2^2\delta_2^{-1}k_j\sigma_{\eta_j}^2}{\gamma_2^4\delta_2^{-2}k_j^2\sigma_{\eta_j}^2 + \sigma_\nu^2}. \tag{1.13}$$

It was assumed earlier that, according to students' beliefs, $\frac{\partial w_{ij}}{\partial GPA_{ij}} = k_j$. In the context of (1.13), a Nash equilibrium in which beliefs are accurate means that the following should hold:

$$\frac{\partial w_{ij}}{\partial GPA_{ij}} = \frac{\rho\gamma_2^2\delta_2^{-1}k_j\sigma_{\eta_j}^2}{\gamma_2^4\delta_2^{-2}k_j^2\sigma_{\eta_j}^2 + \sigma_\nu^2} \equiv h(k_j) = k_j. \tag{1.14}$$

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.

1.4.2 Cross-sectional Predictions

How does k_1 relate to k_0? Since $\sigma_{\eta_1}^2 < \sigma_{\eta_0}^2$, by (1.14) $k_1 \neq k_0$. 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 1.1.) 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.[17] Thus the only surviving equilibrium has $k_1 < k_0$. This leads to the following prediction.

---

both selectivity tiers, then GPA at each tier should be approximately normal as well. In Appendix Figures 1 through 4 and Appendix 1.C, I show that this assumption holds up quite well empirically.

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

[17] 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 (1.11). This is strongly rejected in every data set.

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.[18] 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 (1.12) 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 $k_1 < k_0$.

1.4.3 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.[19] 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 (1.13) with respect to $\sigma^2_{\eta_1}$ yields:

$$\frac{\partial^2 w_{i1}}{\partial GPA_{i1} \partial \sigma^2_{\eta_1}} = \frac{\rho \gamma_2^2 \delta_2^{-1} k_j \sigma^2_{\nu}}{\left[\gamma_2^4 \delta_2^{-2} k_j^2 \sigma^2_{\eta_1} + \sigma^2_{\nu}\right]^2} > 0. \tag{1.15}$$

[18] See Appendix 1.C.3, "Bounding the variance of ν" for an exercise that relates the magnitudes of $\sigma^2_{GPA_1}$ and σ^2_{ν}.

[19] 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.

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 (1.14) 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.

1.5 Data and Empirical Strategy

1.5.1 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.[20] 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

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

analyses across cohorts.[21] 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.[22] 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, college GPA and the bachelor-degree-granting institution must be identifiable for the respondent.[23] Appendix Table 1.1 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 1.3 and 1.4 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 1.2 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.[24]

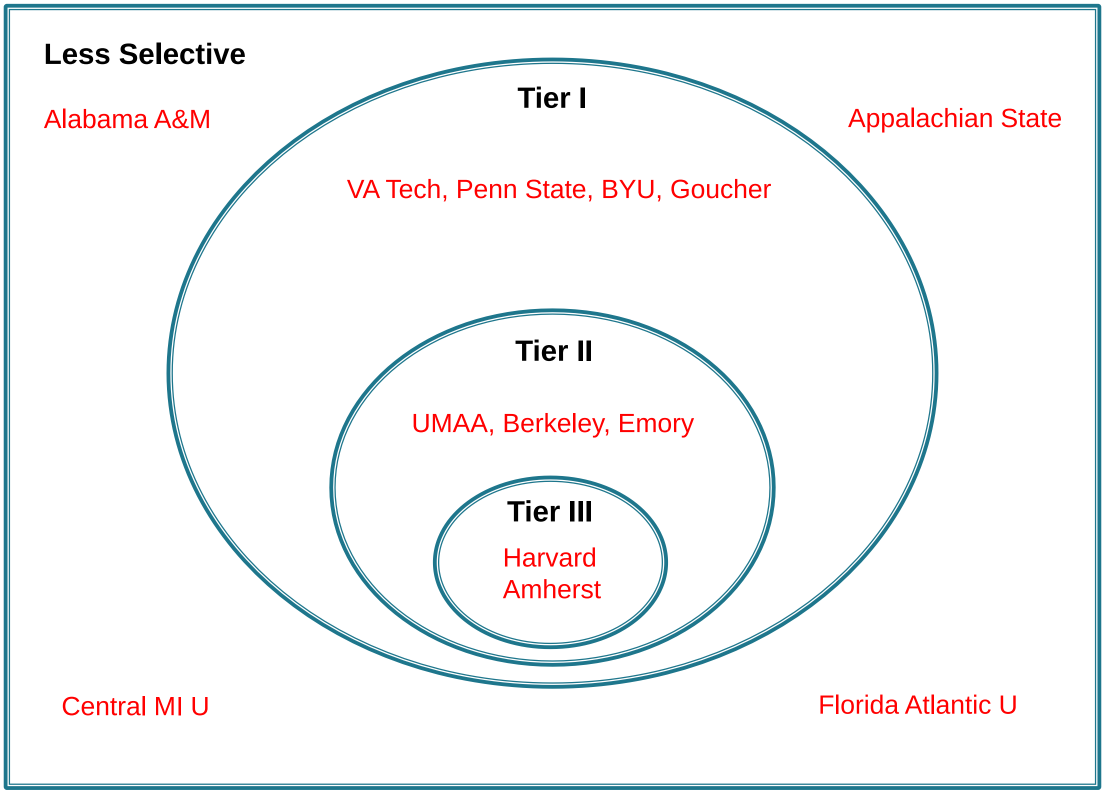
[21] I have also performed cross-sectional analysis separately for each cohort. Point estimates are qualitatively similar to those reported in this paper, although they are less precise.

[22] 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.

[23] College GPA is generally taken directly from transcripts and from self-reports when not transcripts were not available. See the data appendix for details.

[24] 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.

Figure 1.2: Examples of Colleges in Selectivity Structure



Some summary statistics of the estimation samples from each data set can be found in Table 1.1. 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.

1.5.2 Methodology

In order to test Predictions 1 through 3, I estimate the following reduced-form of equation (1.12) separately for each selectivity tier threshold using the pooled data:

$$w_{id} = \theta_0 + \theta_1 S_{id} + \theta_2 GPA_{id}(1 - S_{id}) + \theta_3 GPA_{id}(S_{id}) + \sum_d \lambda_d D_d + \sum_d \lambda_{\mathbf{X}} D_d \mathbf{X_{id}} + \epsilon_{id}, \tag{1.16}$$

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 (1.1).[25] 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 (1.16) 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).

[25] 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.

Table 1.1: 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	

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 data set. 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.

The return to selectivity in equation (1.16) 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.

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

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

where $Late_{id}$ equals 1 if the individual is from the HSB, NELS, or NLSY97 data sets, 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.

1.6 Estimation Results

1.6.1 Pooled Model

Table 1.2 presents the results from estimating equation (1.16) 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.

Table 1.2: Log Hourly Wages on GPA by Selectivity

	(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

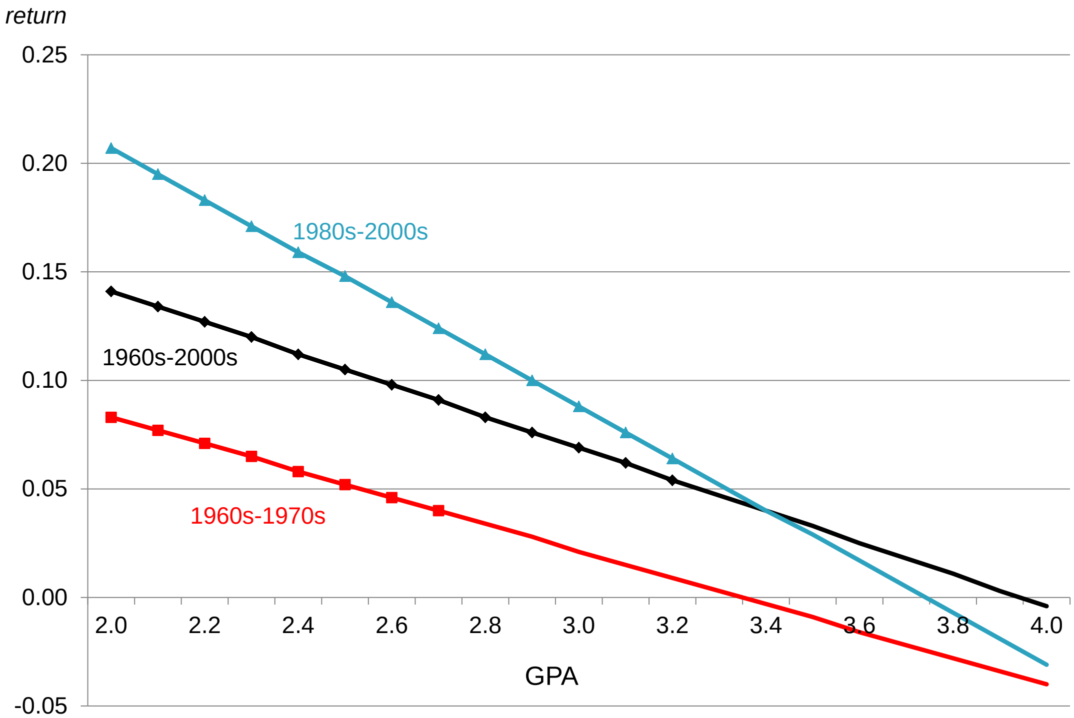
Notes: Estimates shown are for OLS regressions using sampling weights and data pooled across all data sets. The dependent variable in each column is the real log hourly wage. Standard errors (in brackets) are robust to heteroskedasticity and allow for arbitrary correlation of the error term within college. Asterisks indicate statistical significance (* p<0.10, ** p<0.05, *** p<0.01)

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 1.2, for example, repeats Table 1.2 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.[26]

[26] I have also estimated variants of (1.16) 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.

Figure 1.3: 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.

Although Table 1.2 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 (1.16) is best shown graphically. Figure 1.3 plots the selectivity return (in log points) against GPA for full-time workers using the tier II definition (column 5 of Table 1.2), 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 3.0, it is also different from the premium of 0.13 at a GPA of 2.2.[27] This confirms prediction 3 and

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

provides further support for the signaling model.

1.6.2 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 1.15). This is tested formally in Table 1.3, which is similar to Table 1.2 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.[28] Moreover, these results are robust to using the alternative quality index definition of selectivity, as shown in Appendix Table 1.2, panels B and C.

[28] 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.

Table 1.3: Log Hourly Wages on GPA by Selectivity and Time Period

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. The dependent variable in each column is the real log hourly wage. 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).

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 1.3 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 (1.14)). On net, then, it is perhaps not surprising that the GPA return has changed so little at selective colleges.

1.7 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 are consistent with 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.

Of course, that the patterns observed in the data are consistent with signaling cannot conclusively rule out alternative explanations, including variants of the human capital model. More specifically, while I have assumed a production function where the signals of GPA and selectivity provide information about the unknown ability parameter η, the production function could could include a value-added component, $f(\eta_i, SEL_i, GPA_i(e))$, where $f(\cdot)$ 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). While existing data do not allow the examination of this productive value added, it is interesting that 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. Their finding, along with the varying "returns" to GPA by college selectivity found in this paper, imply restrictions on any generalized value-added model that seeks to explain the college selectivity premium.

None of this is meant to imply that institutions of higher education should be thought of primarily as signaling devices for students. Indeed, nothing in the model or the empirical results is inconsistent with college-going providing human capital to students. Rather, the intent of this chapter is to show that signaling provides a compelling alternative mechanism underlying the college selectivity premium.

Moreover, the signaling model is 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 have 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. Finally, the model also suggests why employers appear to learn about the productivity of college graduate workers much faster than that of high school graduate workers (Arciadocono *et al.* 2008): the signals that college graduates can send to employers are more revelatory of ability than those from high school graduates, so there is less to to be learned.[29] Human capital models that seek to explain the selective college premium should also reconcile these stylized facts in order to be persuasive.

It is also worth emphasizing that the evidence in favor of signaling is *not* at odds with the findings of (ability-adjusted) returns to college selectivity in mid-career. Although employer-learning papers typically assume that the the role of the signal generally diminishes over time as the underlying characteristic that firms care about is revealed through experience (Altonji and Pierret 2001, Lange 2007, Arciadocono *et al.* 2008), this need not be true in the presence of job frictions where the initial signal can affect the productivity profile. In fact, Heisz and Oreopoulos (2006) find empirical

[29] I present evidence of this phenomenon in Appendix 1.D.

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. In turn, Bose and Lang (2011) provide a microtheoretical foundation for this friction as firms' try to match specific tasks to the workers they think are best able to handle them; as the initial task assignments are based on what the firms observe *ex ante* about the workers, the signals play a role in further training and the chance of promotion. In the presence of career ladders, first jobs matter 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.

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 attractive 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 (if not farther) for the forward-looking student.

1.8 Appendices

1.8.1 Appendix 1.A: Proofs

1.A.1: Section 3

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

PROOF: Follows immediately from first-order conditions in (1.10) and definition of $\tilde{\eta}$:

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

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 (1.10) 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. ■

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. ■

1.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:

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

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:

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

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

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

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:

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

An application of Leibnitz's rule shows that:

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

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:

$$1-\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2+\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right] \quad < \quad 1-\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]^2-\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right], \quad \text{or}$$

$$\left[\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]^2-\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}\right]^2-\tilde{\eta}\left[\frac{\phi(\tilde{\eta})}{\Phi(\tilde{\eta})}+\frac{\phi(\tilde{\eta})}{1-\Phi(\tilde{\eta})}\right]>0.$$

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

1.8.2 Appendix 1.B: Relaxing functional form on the GPA-effort function

In Section 1.3, the relationship between effort and GPA, given by equation (1.5), assumed the same linear function for all college tiers. 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.

Here I allow 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: $\boldsymbol{\gamma_{11} > \gamma_{10}\,;\gamma_{21} < \gamma_{20}}$, 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 (1.8), 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.[30] 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: $\boldsymbol{\gamma_{11} < \gamma_{10}\,;\gamma_{21} > \gamma_{20}}$, 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 (1.8), 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.

1.8.3 Appendix 1.C: Empirical Support for Model Assumptions

1.C.1: Linearity of GPA in effort and ability

The model in Section 1.3 makes a strong functional form assumption that expected GPA is linear in effort (equation (1.5)). 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 1.5 and 1.6 present nonparametric estimates

[30] 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.

of GPA on the normalized senior test score for less selective colleges and for selectivity tier II.[31] Each figure has six panels: one that pools all cohorts, and one for each cohort separately. The relationship in the first panel of Appendix Figure 1.5, 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 1.6, 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 (1.16) and (1.17) 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 1.3 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.

1.C.2: Empirical densities of GPA and ability

Figures **??** and **??** show kernel density estimates of GPA across selectivity tiers for each of the five data sets used in the paper.[32] 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

[31] The specific procedure is a local linear regression using an Epanechnikov kernel with the bandwidth that minimizes integrated squared error. Nonparametric estimate for the other selectivity tiers are not shown for brevity but are available on request.

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

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.

Figures **??** and **??** 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.[33])

Nonetheless, I simulated data to resemble these empirical distributions in order to examine whether the implications of bivariate normality shown in equation (1.11) 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.

1.C.3: Bounding the variance of ν

A minimum bound of the variance of ν can be estimated by using equation (1.9) 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 (1.9) 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 **??** and **??**); 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.

[33] 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.

1.C.4: 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, 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.

1.C.5: 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 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.

1.C.6: 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 1.5, 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).

1.8.4 Appendix 1.D: Signaling and 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 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 (1.16), 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.[34]

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. Appendix Table 1.D.1 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.[35] 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 ability, so there is less to be revealed through experience.

[34] See the data appendix for details on the construction of these measures.

[35] 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 (1.11).

Appendix Table 1.D.1: Prediction Errors on Ability for College and High School

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	—	No	Yes	—	No	Yes
Full-time only?	No	No	No	Yes	Yes	Yes

Notes: Estimates shown are mean squared errors (MSE) from OLS regressions of normalized 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.

1.8.5 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.[36] 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.[37] 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.

[36] As in all of the NCES surveys here, new individuals were often added in some of the later waves.

[37] Based on normal progression. Respondents were followed regardless of actual high school graduation.

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 1.1 contains more detailed information on how the restrictions affect the sample size for each data set.

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.[38] 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

[38] 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.

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.[39] Again, three different binary indicators for selectivity are calculated. The first of these is coded 1 if the quality index percentile is at or 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).[40] 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[41], summed, and then converted to a percentile score using the 2005-2006 year distribution, also from the College Board.[42] 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[43] 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

[39]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.

[40]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.

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

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

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

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 (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.)

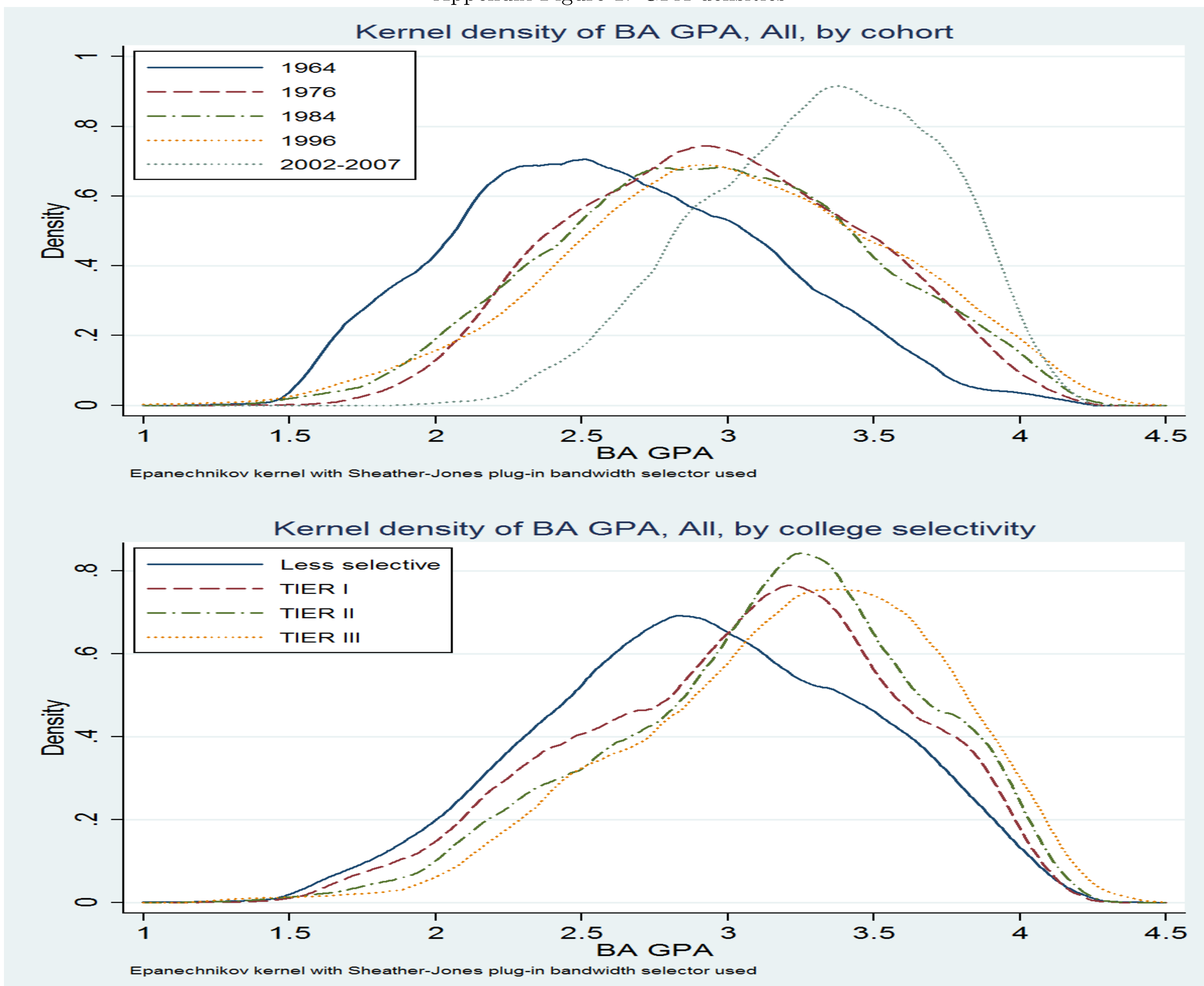
In NLS72, earnings data are provided at the weekly level, and hourly earnings are constructed by dividing weekly earnings at the first post-graduation job by the number of hours worked in an average week at that job. In HSB, there are data for the number of hours usually worked per week, the frequency 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

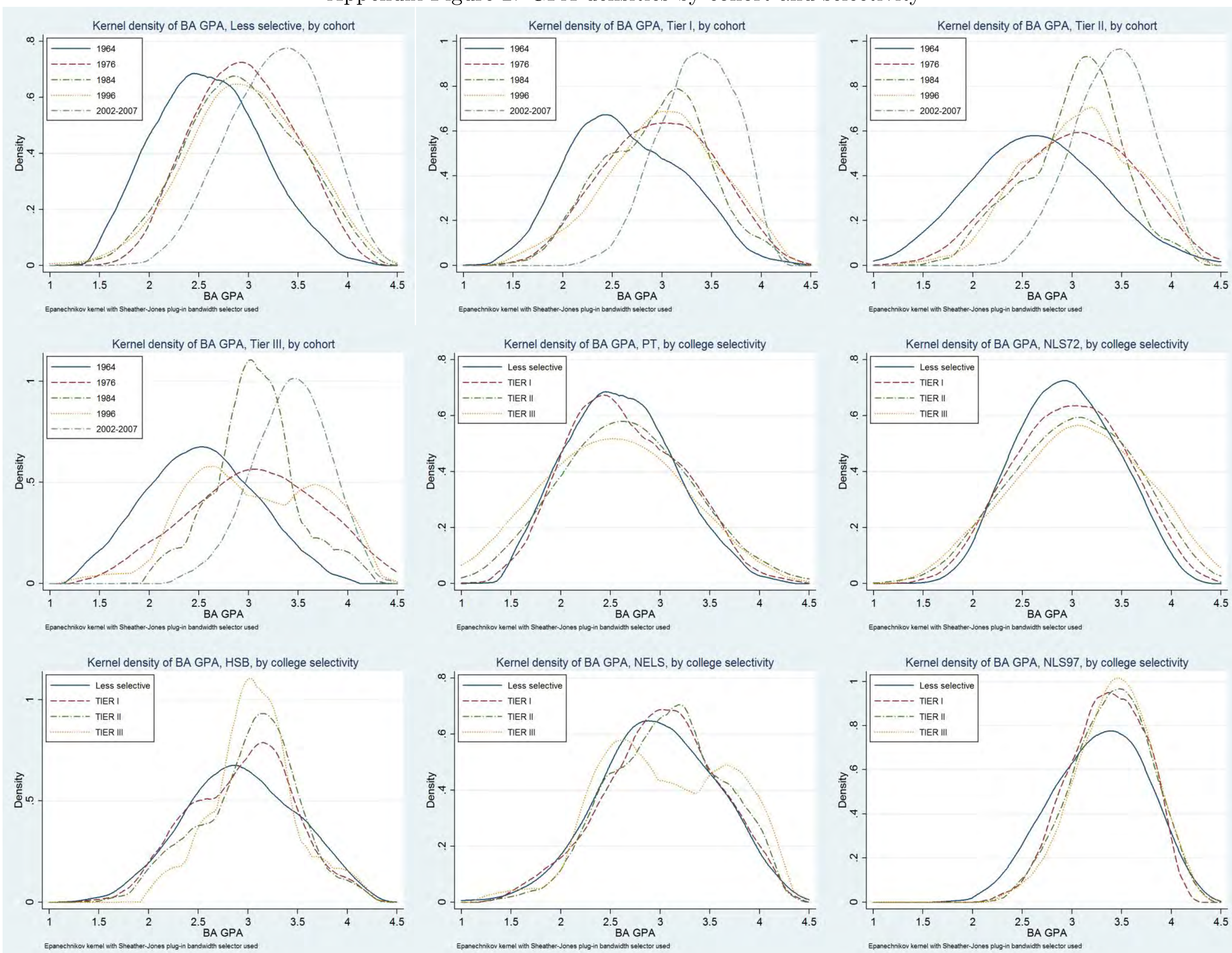
High school GPA is taken from categorical student responses for each data set except for NELS, where it is constructed (within the data set) using high school transcript data. High school GPA is converted to a 4-point scale in a manner analogous to undergraduate GPA. Each data set has students report 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.

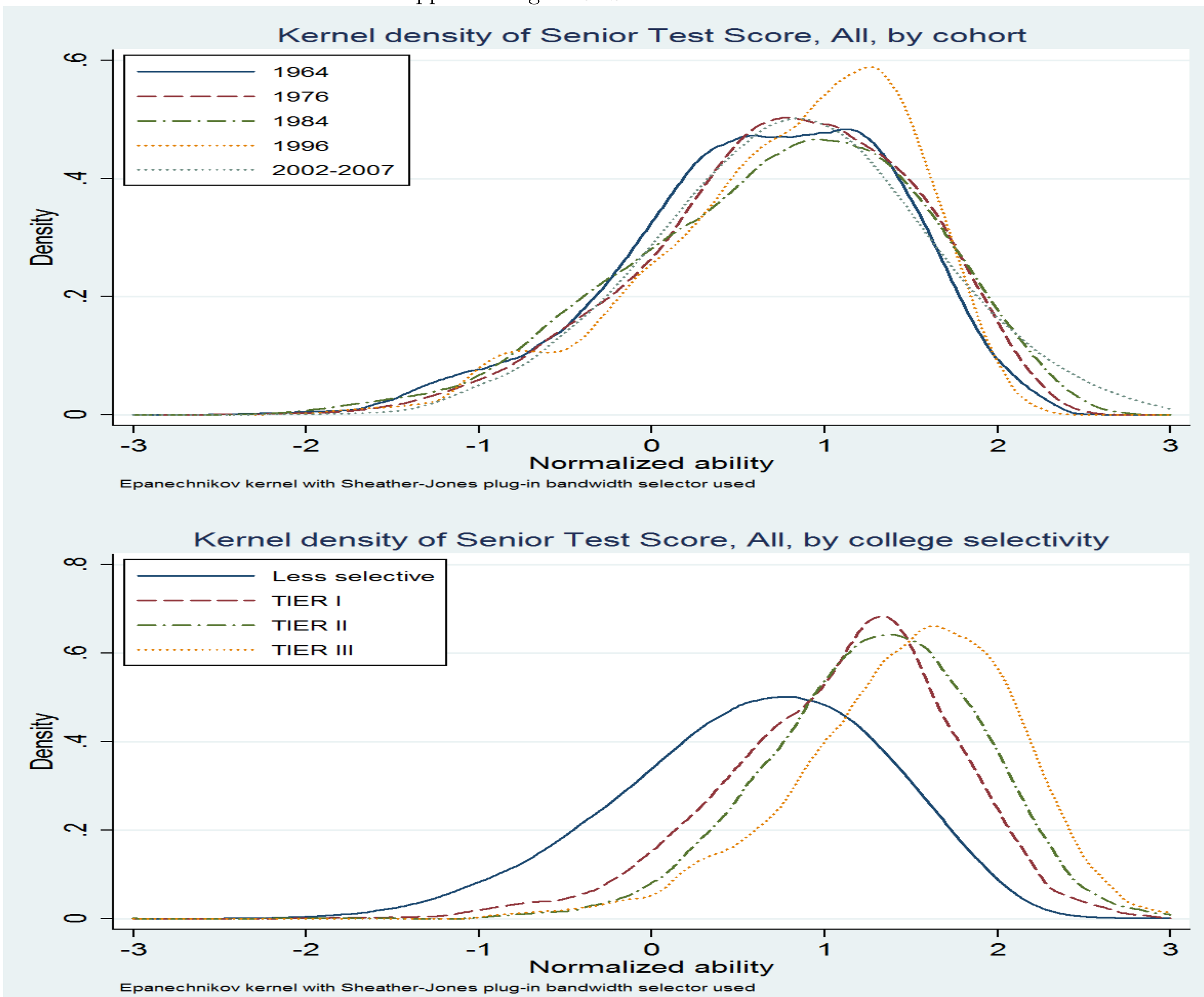
Appendix Figure 1: GPA densities



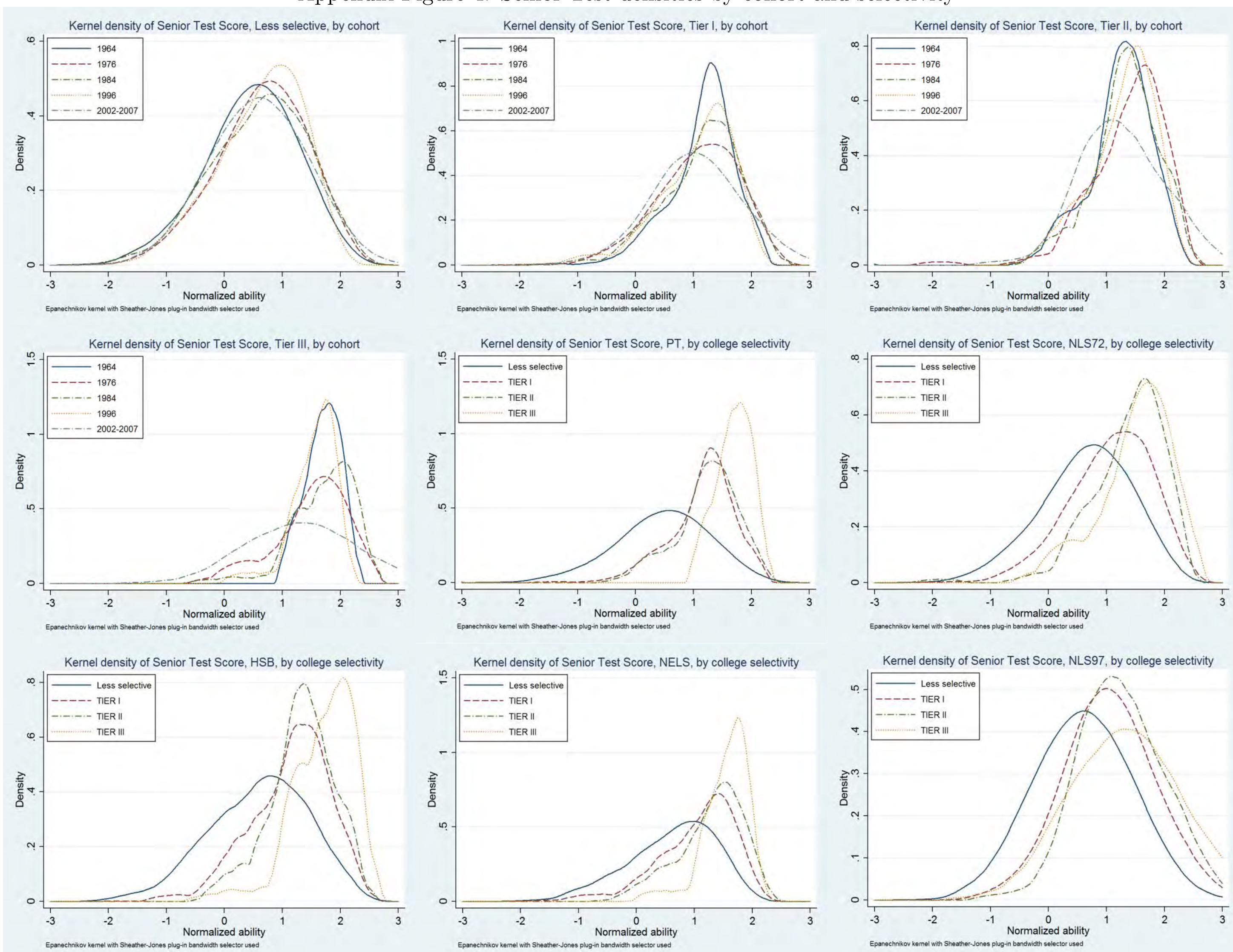
Appendix Figure 2: GPA densities by cohort and selectivity



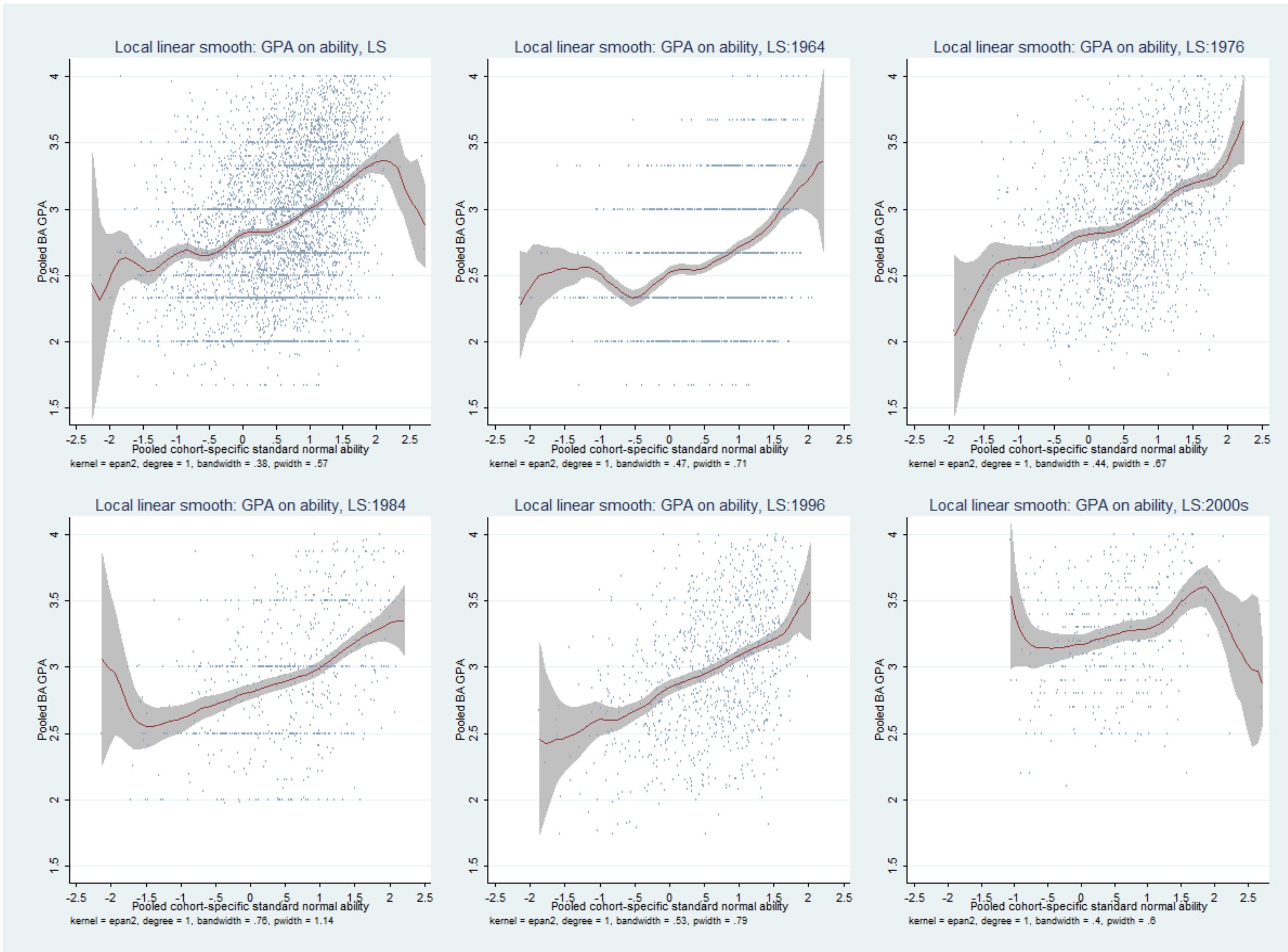
Appendix Figure 3: Senior Test densities



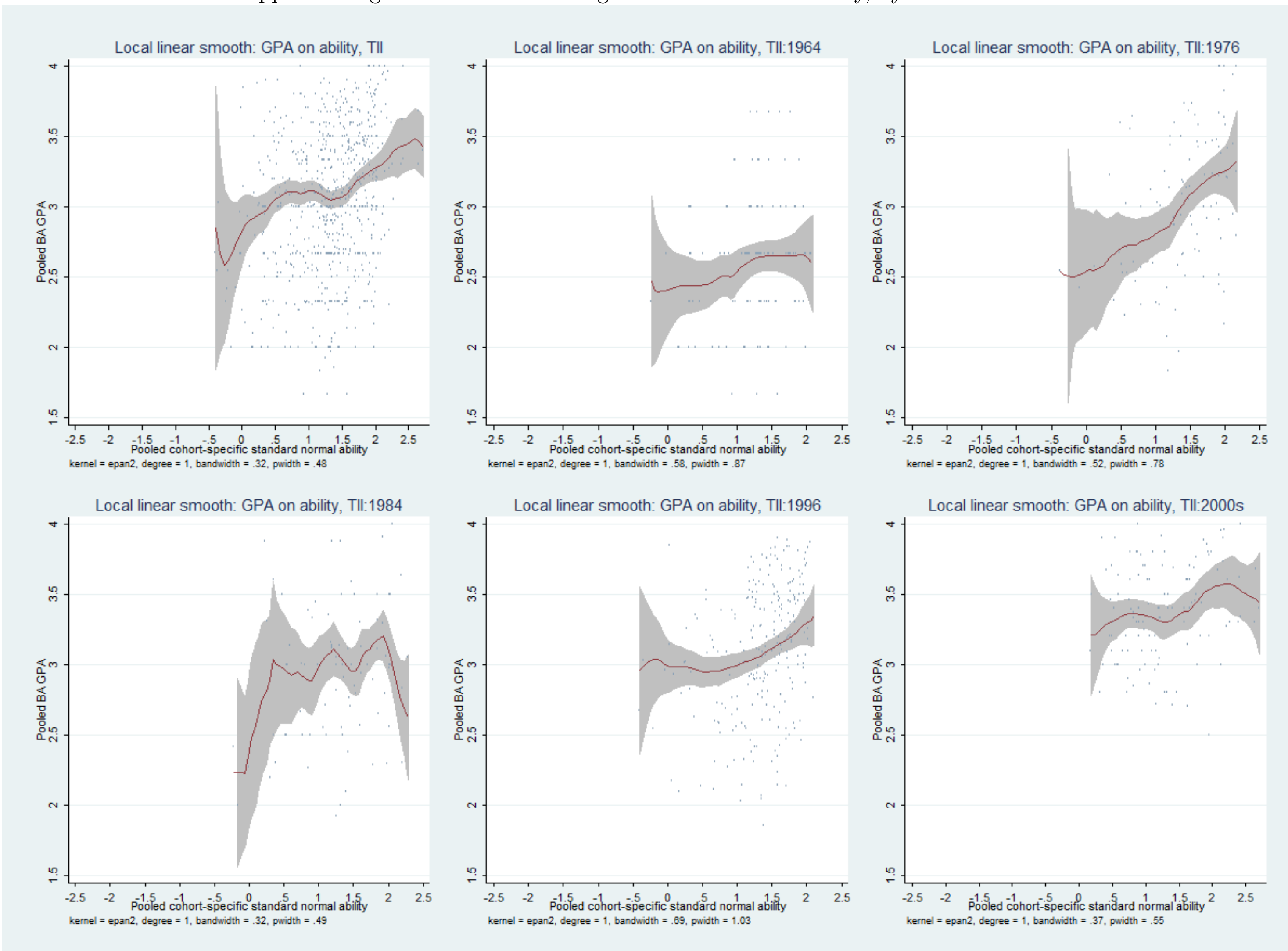
Appendix Figure 4: Senior Test densities by cohort and selectivity



Appendix Figure 5: Local linear regression of GPA on ability, by cohort: Less selective



Appendix Figure 6: Local linear regression of GPA on ability, by cohort: Tier II



Appendix Table 1.1: 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					
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

Notes: Data are from author's calculations from the respective data sets. 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 1.2: Log hourly wages on GPA by selectivity (Quality Index 1991)

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 on the real log hourly wage using sampling weights. College selectivity is based on the Quality Index from Black and Smith (2006). 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 1.3: Wald Tests of Nonlinearity of Wages in GPA

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 1.3 for other notes.

1.9 References

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.

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.

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

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.

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

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

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

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.

Bose, Gautam, and Kevin Lang. 2011. "A Theory of Monitoring and Internal Labor Markets." Boston University mimeo.

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.

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.

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.

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.

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.

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).

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.

Digest of Education Statistics, 2008. 2009. Washington, D.C.: U.S. Department of Education.

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

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.

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.

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.

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

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.

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.

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

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

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.

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

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

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.

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.

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

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.

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.

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

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.

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

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

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.

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

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

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

CHAPTER II

Graduating High School in a Recession: Work, Education, and Home Production

2.1 Introduction

The 2008-2009 recession is widely acknowledged as one of the most severe labor market contractions since the Great Depression. While the contemporary harm to workers unable to find jobs is well-studied in both the popular and academic press, the enduring effects on new entrants are less well understood. A small but growing economics literature has explored the impact that labor market conditions at the time of entry have on the long-term career profiles of men, with most studies finding a negative impact on wages and/or job prestige that is persistent, lasting at least 6 to 8 years, and often longer. The subject has made for an interesting empirical question because different economic theories of labor supply offer different predictions.[1] Given the richness of theories about female labor supply, however, it is somewhat puzzling that little work has examined the effects of graduating during a recession on women. This paper aims to extend the literature by examining how initial negative labor market shocks affect a not-well-studied demographic group—high school graduate women—and whether these impacts are different from those for men. Unlike much of the earlier literature, particular emphasis is placed on channels besides wages, such as the extent and timing of employment and enrollment in higher education.

There are several reasons why high school graduate women are both an economically interesting and important group to study. First, because they arguably have a greater set of alternative uses for their time, they are likely to respond to initial adverse labor market conditions differently than men. For example, housework and child care are well-known to be more plausible choices for women than men.[2] Indeed, given the relatively higher labor supply elasticity found for women relative to

[1]For example, competitive spot labor markets suggest short-lived effects of entry conditions, while job-matching and search models allow for longer-lasting impacts. Kahn (2010) offers a longer discussion of the predictions of these models.

[2]According to the March *Current Population Survey*, for the period 1975 through 1999, an average of 24 percent of women aged 18 to 40 reported not being in the labor force at any point during the year. Of these 24 percent, 70

men (Blau and Kahn, 2007), women first entering the labor force in a period of depressed real wages should be less likely than men to be working, implying that margins other than wages are worth investigating. If home production represents a viable alternative to market work for these women,[3] whether due to less stigma or greater productivity, do women at the margin select into it when the effective market wage offer falls, as it does in a recession? On the other hand, beginning in the early 1980s, women began to outnumber men in college-going. As Goldin (2006) has noted, the generation of women who graduated high school in the late 1970s and early 1980s had greater career ambitions than did their parents; rather than trying to find a job during a recession right after high school or starting a family, perhaps obtaining more education represented another appealing option. Whether more young women are on the work-home production margin or work-education margin–and how these compare with men—is an open empirical question.

Second, certain stylized facts suggest that the long-lasting effects, or scarring, found among more educated men might not occur among less educated women. Studies to date have generally shown that negative effects tend to be larger, and more durable, for workers with steeper earnings profiles. For example, the economics Ph.D. and MBA graduates studied by Oyer (2006, 2007) suffer larger and longer-lasting earnings and placement penalties than do the college graduates studied by Kahn (2010) and Oreopoulos *et al.* (2012), and the college graduates in these two studies in turn fare worse than the prime-age male workers in Beaudry and DiNardo (1991), Genda, Kondo, and Ohta (2010), and Brunner and Kuhn (2009). While different methodologies stress caution in making inferences, the additional finding from Brunner and Kuhn (2009) that effects are more severe for white-collar than blue-collar workers is strongly suggestive that the shape of the earnings profile is related to the impacts of labor market entry during a recession. If the earnings profiles of high school graduate women are relatively flat, these women may recover much better than the previously studied men. High school graduate men, whose earnings profiles fall between these groups, might similarly be expected to fare worse than the women but better than the highly educated men.

Third, studying these women allows a more complete understanding of the short and long-run welfare implications of business cycles for workers, and how they vary by fundamental economic parameters, such as the elasticity of labor supply with respect to wages or the return to labor market experience. Intertemporal behavioral shifts may play a sizable role in determining the impact of an initial adverse labor demand shock on certain worker groups, and analyses that focus predominantly on wages may miss this part of the story. Moreover, given the severity of the 2008-2009 recession,

percent reported their primary reason for not working as "taking care of home/family."

[3] Lechner and Wiehler (2007) present evidence consistent with this hypothesis for less-skilled women in Austria.

understanding the behavioral responses of high school graduates and the consequences for their long-term well-being is of tremendous concern to policymakers and the public.

In order to understand how and why choices might vary across education groups and sex, in the next section I present an informal discrete-choice discussion that focuses the analysis and nests the possibilities for wage impacts seen in previous studies. Section 2.3 discusses the data set, the National Longitudinal Survey of Youth 1979, and the estimation strategy used for identification. The fourth section presents the main results of how an initial labor demand shock affects the probability of being employed, showing that high school graduate women—but not men—substitute away from market work in the first few years following graduation. This section also provides checks for robustness and explores some channels through which the observed effects might be operating. Section 2.5 offers a brief conclusion and thoughts for future research.

2.2 Labor Demand Shocks and Time Use Decisions

Suppose at time t individuals can choose among working in the market, pursuing education, or engaging in home production. Each of these choices grants some flow of utility to the individual that period based on her characteristics, and the choices also have the potential to affect utility flows in future periods. For example, working grants the individual a wage, and the experience gained while working can raise the wage in subsequent periods. Likewise, enrolling in postsecondary school is costly in terms of both money and effort, but additional education can also raise future wages. Home production may possibly affect future productivity in home production, but it does not affect future flows from either working or education.[4]

An individual's optimal choice at time t is the one that grants the highest expected utility, and this depends on an individual's characteristics at that point. For instance, those who have worked more in the past face higher wage offers, on average, than those who have worked less, making working again a more compelling choice. Similarly, if education becomes more costly to pursue the more educated one is (Heckman, Lochner, and Todd, 2006), further education is less appealing unless the return rises commensurately. More generally, greater levels of schooling or work experience make home production a less appealing option, because the opportunity cost (in foregone wages) is greater.

For new high school graduates, in particular, the decision process is somewhat cleaner, as they all have the same education, minimal work experience,[5] and are of nearly the same age. Choices

[4]For evidence that home production is highly substitutable with market work, at least in the short run, see Burda and Hamermesh (2009).

[5]The literature has not reached consensus on whether work experience in high school affects later labor market outcomes. Ruhm (1997) finds positive effects on work participation and earnings around age 30 for those who worked in

will be determined chiefly by exogenous characteristics, such as cognitive ability—which affects the psychic and financial costs of education and thus its net return—the return to working—both the initial wage offer and the return to experience—and the value attached to home production.

When a negative labor demand shock occurs, as during a recession, and wage offers are reduced, how are choices affected?[6] It depends on the size of the shock relative to the value of the alternative options, and while the latter depend on levels of schooling and experience, there is also reason to believe they vary by sex even when schooling and experience are held constant. Women, for instance, likely find childrearing and housework (i.e., home production) far more feasible than men, who often suffer greater stigma from these choices (Goldin, 2006). Additionally, because women on average have more career interruptions than men, often family related, joblessness is less likely to convey a negative signal of worker quality to employers for women than it is for men. Indeed, Light and Ureta (1995) show that women's wage penalties associated with re-entering market work after a nonworking spell are small relative to those for men. Furthermore, women's greater amount of time spent outside market work may serve to lower the lifetime returns to education. Each of these factors suggests that more women than men are on the margin between work and home production, and thus a negative labor demand shock would be expected to induce women from the former to the latter more than men.

As simple as this framework is, it can put into context why different effects should be expected than the results found by Kahn (2010), Oreopoulos *et al.* (2012), and Oyer (2006, 2007). Because these papers study those with bachelor's or graduate degrees, the wage offer is much higher than it would be for the less educated.[7] Additionally, the value of additional schooling is likely to be less, both because schooling becomes more costly and because the return decreases. Consequently, even with a labor demand shock that reduces the wage offer, continuing to work is likely to remain the best option for most of this population, and the initial impact on these workers will almost fully show up in observed wages.

Furthermore, the framework can also shed light on how long the effects from an early demand shock can be expected to last. It is well known that wage experience profiles tend to be steeper for those more highly attached to the workforce, particularly the more educated. In the context here, the return to experience is likely to be larger for these groups than it is for new high school

high school, but Hotz *et al.* (2002), using the same data, argue and present evidence that once unobserved endogeneity is taken into account, the positive effects become statistically indistinguishable from zero.

[6]For evidence that the unemployment rate serves as a decent proxy for labor demand, and the procyclicality of wage offers, see Hall (2005) and Elsby, Michaels, and Solon (2009).

[7]The single other paper that has investigated women specifically, to my knowledge, is Kondo (2008). Because her focus is on wages, and her sample includes women with different initial education levels, she does not examine the behavioral substitution predicted here.

graduates, especially new high school graduate women.[8] Using the NLSY79, I find that a simple Mincer equation estimate for female high school graduates implies a return to actual experience of about 3 percent among new workers. This is considerably less than the implied estimates of 10 to 13 percent for new male workers found by Heckman, Lochner, and Todd (2006). Over time, not only will the negative impact on wages persist for the labor friction-related reasons of implicit contracting (Beaudry and DiNardo, 1991) or reduced accumulation of industry or occupation specific capital (Kahn, 2010), these effects compound as wage growth occurs from a smaller initial base. With a flatter wage profile, not only is the initial penalty from foregoing work smaller, so is the compounding effect. This implies weaker long-term wage effects among individuals who do work, and, because the implicit wage offers are penalized less, contemporaneous labor demand shocks are likely to matter more than the initial labor demand shock for the less educated, and for women more than men. Consequently, negative employment effects for less educated women are likely temporary.

2.3 Data and Empirical Strategy

2.3.1 Discussion of Data

To investigate the presence and magnitude of behavioral shifts, I employ the *National Longitudinal Survey of Youth, 1979* cohort. The *NLSY79* is a detailed panel data set that first interviewed 12,686 individuals aged 14 to 22 in 1979 and conducted follow-up interviews annually thereafter until 1994, when subsequent interviews became biennial. Not only does it contains extensive information on education and work history, its respondents graduate high school during the mid 1970s through early 80s, a period that experienced a recession, a recovery, and another recession.[9]

The panel data in the NLSY79 provide several advantages over repeated cross-sections from the *Current Population Survey* (*CPS*). By using the restricted geocoded version of the data set, I can precisely identify the exact year, month, and geography of high school graduation.[10] This allows me to restrict the sample to the population of interest: "on-time" high school graduates who leave high school the year they turn 17, 18 or 19. Additionally, as the *CPS* is a household sample, individuals attending college and living in a dormitory may not be counted properly, which is particularly

[8] This would be true, for example, if there are complementarities between education and work experience.

[9] For more information on the *NLSY79*, see the Bureau of Labor Statistics' web site: `http://www.bls.gov/nls/nlsy79.htm`.

[10] Measurement error in the *CPS* on the timing and location of receipt of the high school diploma may be substantial. First, the date of diploma receipt is not consistently available, and back-of-the-envelope calculations from the NLSY79 suggest considerable dispersion in the age at diploma receipt, making imputation based on age unwise. Second, high school graduation is generally commingled with GED receipt. Heckman and LaFontaine (2007) review why GEDs are not the same as actual diplomas and document that 10 to 15 percent of high school completers in the *CPS* are GED recipients.

problematic as college enrollment is an outcome of interest. Perhaps most importantly, however, is that the data provide an ability measure, the Armed Forces Qualifying Test (AFQT).[11] Although no measure of ability is perfect, the AFQT is likely to be a good proxy for the ability variable discussed in Section 2.2. Since respondents ranged in age from 15 to 23 at the time of administration, it is necessary to adjust the scores to make them comparable across cohorts. I create age-specific z-scores by regressing the raw scores on year of birth dummies and then dividing the residuals from this regression by their sample cohort standard deviations.

Table 2.1: Summary Statistics of Selected Variables

Panel A: Time Invariant Variables	*Men*		*Women*	
	Mean	Std Dev.	Mean	Std Dev.
Age-standardized AFQT	0.275	0.847	0.202	0.803
National unemployment rate at HS grad	7.44	1.24	7.40	1.21
Panel B: Time Varying Variables	*Men*		*Women*	
	Mean	Std Dev.	Mean	Std Dev.
Years since HS graduation	13.21	7.32	13.36	7.38
Currently in labor force	0.855	0.352	0.775	0.418
Currently employed	0.796	0.403	0.716	0.451
Annual weeks employed	43.3	15.9	38.1	19.7
Annual hours worked	1935	960	1431	932
Currently enrolled	0.131	0.338	0.130	0.337
Log hourly wages ($1977)	1.632	0.585	1.404	0.527

Notes: See data appendix for variable definitions and construction.

The *NLSY79* data are then linked to external labor market indicators. For the principal independent variable of interest, the labor market demand shock at the time of high school graduation, I match the national annual average unemployment rate of all workers using year of graduation. I use the national unemployment rate rather than state unemployment rates for two reasons. First, it is measured much more precisely than the state unemployment rates. Second, it is available for all cohorts in the data whereas the state unemployment rate series begin only in 1976, and thus cannot be matched to the 1975 high school graduation cohort.[12] Summary statistics for all the variables mentioned are shown separately by sex in Table 2.1.

[11] More accurately, the respondents took a 10-component test known as the Armed Services Vocational Aptitude Battery (ASVAB). The AFQT is a weighted sum of four of the components: arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations.

[12] This comes at the expense of potentially greater identifying variation in the state unemployment rates. While I do not use these latter rates in the current paper, analysis using them in earlier drafts showed similar, albeit slightly smaller and noisier, results than those using national unemployment rates.

2.3.2 Empirical Strategy

Organizing the data into a panel format, I estimate the following reduced-form equation motivated by the framework in the preceding section:

$$
\begin{aligned}
y_{it} = \beta_0 + \sum_{m=1} \beta_{1m}[UR_{i0} * 1(t_i = m)] + \beta_2 AFQT_i + \beta_3 t_i \\
+\beta_4 t_i^2 + \sum_r \beta_{5r}[1(year_{it} = r)] + \beta_{\mathbf{6}}\mathbf{X_i} + \varepsilon_{it},
\end{aligned} \tag{2.1}
$$

where i indexes the individual and t indexes the number of years elapsed since high school graduation. The dependent variable y_{it} is a measure of work attachment, college enrollment, or wages. $AFQT_i$ is the age-standardized z-score, t_i is the number of years since graduation (i.e., potential experience), $\sum_r \beta_{5r}[1(year_{it} = r)]$ is a set of calendar year-of-observation dummies to control for contemporaneous labor demand factors,[13] and $\mathbf{X_i}$ are time-invariant family background controls.[14] The terms $[UR_{i0} * 1(t_i = m)]$ represent a set of variables for each possible year elapsed since graduation, where each variable takes on the value of the national unemployment rate at the time of high school graduation for $t = m$ and 0 otherwise. The corresponding set of coefficients, $\{\beta_{1m}\}$, gives the impact of the high school graduation unemployment rate on y_{it} *at each year after graduation*, with identification coming from temporal variation in the unemployment rate.[15] When y_{it} is continuous, as for wages or number of weeks worked, equation (2.1) is estimated by OLS. When y_{it} is a binary variable, as for work status or college enrollment, the probability that y_{it} equals 1 is modeled as the probit analogue to equation (2.1) and estimated by (quasi-)maximum likelihood.[16] For each dependent variable, the estimation is done separately by sex, although tests of equality across sexes are based on a fully interacted model. Finally, because the survey waves are fielded at different points in the year over time, models for dependent variables that refer to a point-in-time measure (work status, labor force participation, wages) also include a set of calendar month dummies in order to control for seasonal effects.

[13] Results do not appreciably change if a linear term of the contemporaneous unemployment rate interacted with time since high school graduation is added as an additional control.

[14] These background controls include a set of dummies for mother's education, father's education, race, whether one's mother worked when one was 14, whether one was born in the South, and whether one grew up speaking a language other than English. I have also estimated equation (2.1) excluding the family background controls. Consistent with these controls being essentially uncorrelated with the unemployment rate at high school graduation, the point estimates are nearly identical, although standard errors are slightly larger.

[15] This semi-parametric specification is generally to be preferred over a low-order polynomial interaction with potential experience and the initial unemployment rate, despite it being more demanding of the data. If the treatment effect is not smooth over time, or fades quickly relative to the estimation horizon, a polynomial specification may not capture effects properly. Appendix Figures 1 and 2 contrast the semi-parametric specification with one using a quartic in experience for different outcomes.

[16] Because ε_{it} is likely correlated across time for a given individual, the observations are not strictly independent, and thus the estimator is only "quasi-ML."

2.3.3 Validity of Empirical Strategy

Identification of the causal effect of the high school graduation unemployment rate on later outcomes rests on the assumption that the unemployment rate is uncorrelated with the timing of high school graduation. If there is correlation between the two, then high school graduation is itself endogenous and the estimates will be biased and inconsistent. To test this assumption, Table 2.2 shows the results of regressing whether an individual graduated high school on time (at age 17, 18, or 19, and with a regular diploma) on the unemployment rate in the year in which the individual turned 17 or 18. The first two columns look at women while the last two look at men. In none of the cases is the point estimate statistically significant at even the 10 percent level, and the magnitudes are generally quite small. Consequently, the assumption that the unemployment rate is uncorrelated with the timing of high school graduation cannot be rejected.

Table 2.2: Tests of Correlation Between High School Graduation and Unemployment Rate

	(1)	(2)	(3)	(4)
	Women	Women	Men	Men
Mean dep variable	0.758	0.758	0.724	0.724
HS Grad UR, age 17	0.003		0.009	
	(0.010)		(0.007)	
HS Grad UR, age 18		0.009		0.003
		(0.006)		(0.004)
(Pseudo) R-squared	0.194	0.194	0.210	0.209
Observations	2938	2938	2810	2810

* significant at 10%, ** significant at 5%, *** significant at 1%

Notes: The estimates report the results of marginal effects (averaged over sample individuals) from probit regressions of on-time high school graduation on the unemployment rate during the year when the respondent turns age 17 or 18. The samples in this table are restricted to individuals who have valid AFQT scores; they *include* nongraduates and those who graduate after age 19. The unit of observation is an individual. All regressions include as covariates the age-standardized AFQT score and family background controls. Standard errors (in parentheses) are clustered at the year-of-birth level; using standard Huber-White "robust" variance estimation produces nearly identical standard errors.

Another issue is the potential for omitted variable bias, as use of the national unemployment rate does not allow for heterogeneity or trends across cohorts in the outcome variables. I investigate this possibility by making the assumption that ability, as measured by the AFQT, is highly correlated with the unobserved cohort heterogeneity parameter. Appendix Table 2.1 presents sample means and standard errors of the normalized AFQT measure by high school graduation cohort. Two things bear mentioning. First, most of the scores are significantly positive, at around 0.2 standard deviations above zero. This is commensurate with the fact that the sample is conditioned on those who graduated from high school (full diploma) and on time (age 17, 18 or 19); this group is expected to be slightly above average. Second, most of the means are not significantly different from one another; the two exceptions are 1976 and 1983. The difference between 1976 and most of the other years is small, about 0.1 standard deviations. The 1983 mean, on the other hand, is a clear outlier with a mean of 0.5 standard deviations below zero. As the youngest cohort in the NLSY79 turned 15 during the calendar year 1979, the 1983 high school graduation cohort consists exclusively of students who turned 19 that year. This group may disproportionately comprise students who were held back a year, individuals who are presumably of lower than average ability. Because this finding calls into question the comparability of the 1983 cohort with the others, the subsequent regressions were also run omitting this cohort, as well as the 1976 graduation cohort; results were not appreciably affected. Thus, cohort differences do not appear to be an important concern for estimation.[17]

2.3.4 Regarding Inference

For purposes of inference, it would be desirable to allow for arbitrary correlation of the errors within a high school graduation year cohort, of which there are nine (see Moulton, 1986, and Bertrand, Duflo, and Mullainathan, 2004). However, when the number of clusters is small, both the standard normal distribution and the t-distribution with few degrees of freedom may prove poor approximations of the distribution of the test statistic, making standard asymptotic inference problematic. Two alternative approaches for variance estimation when there are few clusters have been suggested (Cameron, Gelbach, and Miller, 2008): a nonparametric cluster bootstrap and a wild cluster bootstrap of the t-statistic. The first of these methods relies on re-sampling the data by clusters, running the estimation, and then deriving the distribution of the point estimates across the re-samples. The second is a more complicated technique that relies on simulating data under the null hypothesis. Unfortunately, there are difficulties using either technique to calculate standard errors in the current setting. The nonparametric cluster bootstrap will produce some re-samples with very little variation in the key regressor—the unemployment rate at high school graduation—resulting in point estimates that behave erratically. Indeed, Cameron *et al.* present simulations that show that this method does not perform well when the number of clusters is fewer than 20.[18] The wild cluster bootstrap, on the other hand, is designed for continuous (not binary) outcomes and to yield rejection

[17] The data appendix discusses further approaches used to check the validity of the empirical strategy against variation across cohorts.

[18] In particular, the distribution of the point estimates is highly kurtotic, with data points in the tails occurring when variation in the high school unemployment rate is small.

regions rather than standard errors, per se. Furthermore, recent work suggests the approach may not be robust to model mis-specification (Kline and Santos, 2011).

Because of the shortcomings of these variance estimation approaches, I have chosen to present standard errors that cluster the error structure on individuals. These standard errors are generally larger than those from clustering on high school graduation year, so this approach will lead to more conservative inference. (Wald tests of joint significance are based on the "robust" estimator of the covariance matrix that accounts for individual-level clustering.) However, because it is not feasible to reliably allow for intra-graduation cohort correlation, some caution should be used in inferring statistical significance.

2.4 Results

2.4.1 The Work Decision

The results from estimation of the model for whether the respondent is currently working appear in column 1 of Table 2.3 (women) and Table 2.4 (men). These estimates show how changes in the unemployment rate at high school graduation affect the likelihood of working for each of the 15 years following graduation.[19] For women, the effects are negative, large in magnitude, and statistically significant for the first few years after graduation. In the first year, the average woman is 2.5 percentage points (about 4 percent) less likely to be employed per percentage point rise in the unemployment rate. Thus, a woman who graduated in a severe recession like that of 1982, in which the unemployment rate rose 3 percentage points above its long-term average, would be 7.5 percentage points (or about 12 percent) less likely to be working one year after graduation.[20] Following this sharp drop in the first year, the net effect begins to diminish, with the magnitude of the estimate falling by almost half, to 1.4 percentage points, in the second year, and another

[19] Although the underlying regressions use all available experience years, to save space, Tables 2.3 and 2.4 report effects only up to 15 years after graduation; significant effects past this horizon will be discussed as they arise.

[20] For comparison, this 7.5 percentage point reduction is of approximately the same magnitude as Angrist and Evans (1998) find for the motherhood penalty using data on twin births, and it is about two-thirds of the difference in employment rates between females who are exactly high school graduates and those with some college (Bureau of Labor Statistics).

Table 2.3: Working Status by Experience Year, Women

	Mean work	(1) Work (=1)	Mean weeks worked	(2) Weeks worked last year	Mean hours worked	(3) Hours worked last year
HS Grad UR:						
1 year after	0.583	-0.0246***	32.2	-0.895**	1029	-56.94***
		[0.0087]		[0.354]		[15.53]
2 years after	0.652	-0.0136**	34.0	-0.621**	1129	-42.87***
		[0.0069]		[0.304]		[13.33]
3 years after	0.674	-0.0097*	35.0	-0.470*	1189	-34.23***
		[0.0057]		[0.265]		[11.63]
4 years after	0.686	-0.0082	36.1	-0.342	1295	-20.42*
		[0.0052]		[0.244]		[10.76]
5 years after	0.753	0.0010	38.1	-0.095	1455	-0.56
		[0.0047]		[0.233]		[10.39]
6 years after	0.768	0.0025	39.5	0.085	1530	7.91
		[0.0047]		[0.230]		[10.49]
7 years after	0.781	0.0032	39.3	0.033	1526	4.56
		[0.0049]		[0.235]		[10.87]
8 years after	0.781	0.0024	39.4	0.023	1538	3.83
		[0.0051]		[0.243]		[11.38]
9 years after	0.767	-0.0009	39.0	-0.072	1537	1.46
		[0.0056]		[0.252]		[11.94]
10 years after	0.764	-0.0020	39.2	-0.063	1543	0.26
		[0.0058]		[0.262]		[12.51]
11 years after	0.765	-0.0027	29.1	-0.102	1536	-2.17
		[0.0059]		[0.271]		[13.04]
12 years after	0.749	-0.0059	39.1	-0.129	1530	-5.30
		[0.0061]		[0.277]		[13.40]
13 years after	0.747	-0.0077	39.0	-0.187	1513	-10.25
		[0.0062]		[0.283]		[13.78]
14 years after	0.756	-0.0082	39.3	-0.194	1524	-11.63
		[0.0061]		[0.286]		[14.03]
15 years after	0.765	-0.0086	39.5	-0.207	1518	-14.49
		[0.0064]		[0.288]		[14.21]
Wald test 1-4 years zero effect (p-value)		0.003		0.039		< 0.001
(Pseudo) R-squared		0.031		0.037		0.045
Observations		34,660		50,657		50,157
Unique women		2,211		2,209		2,209

* significant at 10% ** significant at 5% *** significant at 1%

Notes: The estimates report the marginal effect of a one percentage point increase in the unemployment rate in the year of high school graduation on the likelihood of being employed x years after high school graduation. Column 1 shows average marginal effects(averaged across sample observations) from a probit; columns 2 and 3 show OLS coefficients. All regressions also include AFQT z-score, a quadratic in years since high school graduation, dummies for years of observation, and family background controls (mother's education, father's education, race dummies, a dummy for whether the mother was working when respondent was 14, a dummy for whether the respondent was born int he South, and a dummy for whether the respondent grew up speaking a foreign language). Column 1 also includes month of interview dummies. Observations are person-years, and the number of unique individuals refers to respondents who have at least one observation in the sample. Standard errors (in brackets) are clustered at the individual level. See text for discussion on inference.

Table 2.4: Working Status by Experience Year, Men

	Mean work	(1) Work (=1)	Mean weeks worked	(2) Weeks worked last year	Mean hours worked	(3) Hours worked last year
HS Grad UR:						
1 year after	0.594	0.0102	32.9	0.123	1199	-17.70
		[0.0092]		[0.340]		[17.17]
2 years after	0.640	0.0058	34.6	0.029	1318	-21.61
		[0.0075]		[0.303]		[15.46]
3 years after	0.667	-0.0005	36.0	-0.105	1402	-28.72**
		[0.0065]		[0.268]		[13.96]
4 years after	0.695	-0.0053	37.6	-0.159	1528	-28.83**
		[0.0060]		[0.244]		[12.93]
5 years after	0.767	-0.0024	40.4	-0.051	1730	-18.39
		[0.0055]		[0.230]		[12.42]
6 years after	0.816	-0.0016	42.3	-0.044	1846	-19.07
		[0.0052]		[0.219]		[12.15]
7 years after	0.846	-0.0023	44.1	-0.043	1962	-18.49
		[0.0049]		[0.2111]		[12.09]
8 years after	0.865	-0.0035	45.0	-0.132	2020	-24.16**
		[0.0048]		[0.209]		[12.21]
9 years after	0.881	-0.0037	45.8	-0.217	2080	-27.85**
		[0.0046]		[0.207]		[12.31]
10 years after	0.892	-0.0039	46.0	-0.334	2118	-31.64**
		[0.0045]		[0.207]		[12.51]
11 years after	0.893	-0.0049	46.3	-0.425**	2155	-34.10***
		[0.0046]		[0.206]		[12.68]
12 years after	0.905	-0.0042	46.6	-0.509**	2190	-36.41***
		[0.0044]		[0.207]		[12.87]
13 years after	0.904	-0.0051	46.7	-0.574***	2182	-42.41***
		[0.0047]		[0.209]		[13.09]
14 years after	0.931	-0.0006	47.3	-0.562***	2224	-40.33***
		[0.0037]		[0.208]		[13.16]
15 years after	0.913	-0.0030	47.3	-0.620***	2233	-41.86***
		[0.0047]		[0.207]		[13.16]
Wald test 1-4 years zero effect (p-value)		0.100		0.171		0.012
Wald test 1-4 years, men same as women		0.010		0.017		0.003
(Pseudo) R-squared		0.098		0.108		0.142
Observations		30,673		44,381		43,905
Unique women		2,006		2,004		2,004

* significant at 10%   ** significant at 5%   *** significant at 1%

Notes: See notes to Table 2.3.

third, to 1 percentage point, by the third year. One can easily reject the null hypothesis of no effect of the initial unemployment rate on the likelihood of working in the first four years (p = 0.003).

However, full recovery is reached five years out, and there appears to be no significant effect of the high school graduation unemployment rate on women's work status, either negative or positive, after this point.

Men, on the other hand, suffer no short-run employment penalty to graduating in a recession (Table 2.4). In fact, the point estimates are actually positive for the first two years, although none of the estimates is individually statistically significant. Although the joint test of no effect over the first four years is marginally significant (p=0.100), the change in sign of the point estimates rules out a negative effect. Moreover, the initial employment effects between women and men are statistically different at the 1 percent level. This difference can perhaps best be seen in Figure 2.1, which plots out the estimates for both sexes over the 15-year time horizon.[21]

Furthermore, the negative work effects for women are *not* driven by women entering unemployment. Figure 2.2 presents estimates based on equation (2.1) using labor force participation as the dependent variable. The pattern of estimates for women is nearly identical to the work estimates shown in Figure 2.1. The pattern for men is remarkably similar as well, except for the first two years out, where labor force participation effects are slightly more positive than work estimates (although they are not statistically significantly different). These results are consistent with the average high school graduate woman having a more valuable alternative option to working than the average high school graduate man.[22]

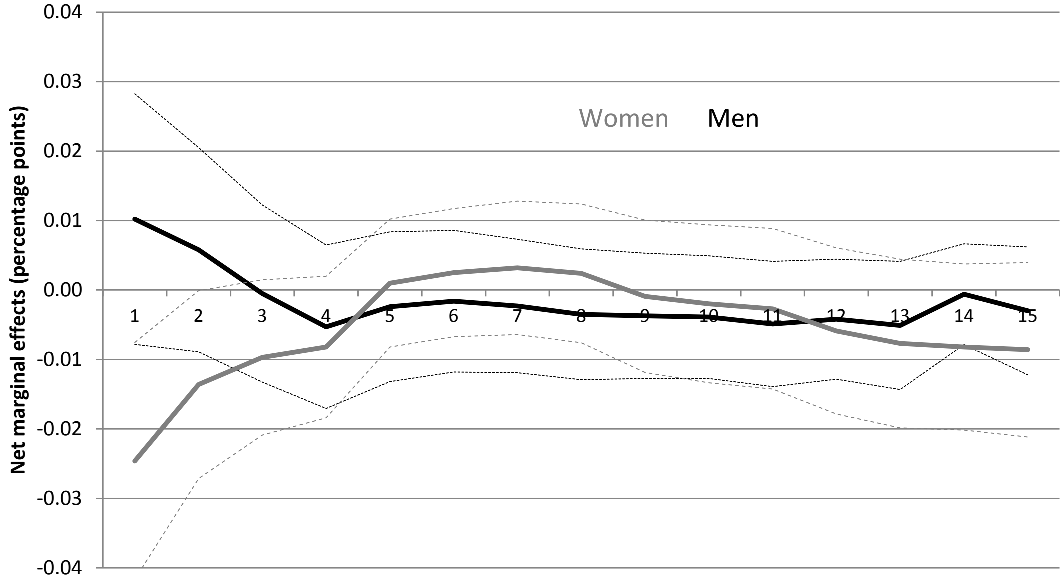
The remaining columns of Tables 2.3 and 2.4 show that these work effects are robust to alternative measures of work intensity. Column 2 reports the effects of the unemployment rate on the number of weeks worked over the calendar year, and column 3 similarly shows effects for the number of hours worked over the calendar year.[23] For women, the pattern of the coefficients is quite similar to the binary measure of work status: a strong negative estimate at one year after graduation that fades away over the next two to three years. The magnitudes are also comparable. The initial 0.9 week (57 hour) reduction in work at one year amounts to a 2.8 (5.5) percent decrease, close to the 4 percent decline for work status. Tests on the joint significance of effects over the first four years also give similar inference: rejections of no effect at p=0.039 for weeks worked and $p<0.001$ for

[21] Tests of joint significance from 5 years until the end of the data horizon (including past year 15) indicate no significant effect for women (p=0.122), a marginally significant impact for men (p=0.051), and no statistical difference between the two (p=0.212).

[22] The negative work results for women appear to be driven by those who don't go to college: restricting the sample to those who don't enroll in the first four years after high school produces effects approximately 50 percent larger in the first two years.

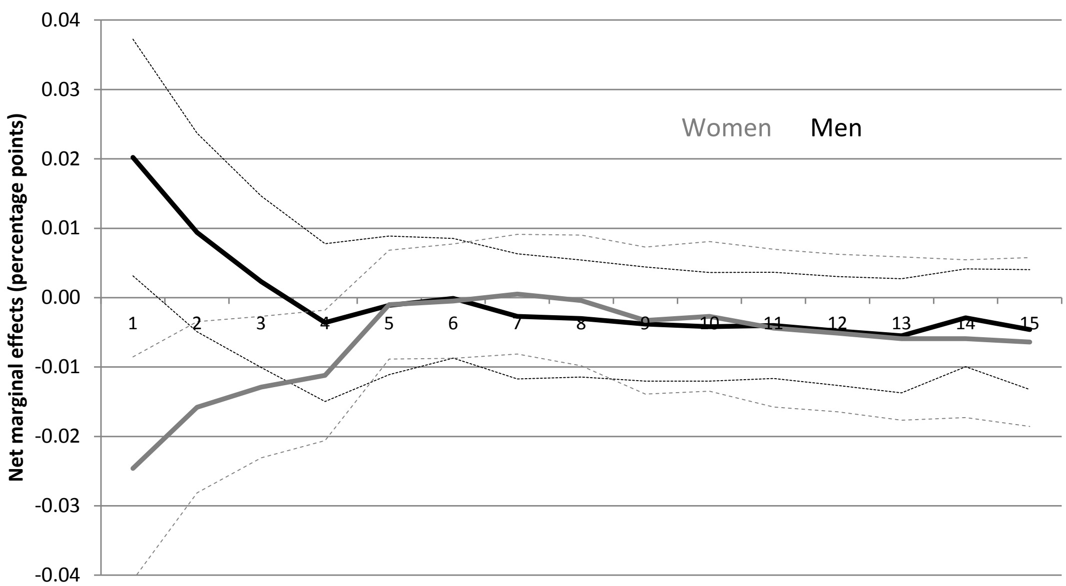
[23] As these measures are based on work history schedules, they are available for every year, including after the survey switched to a biennial format. In contrast, the work status variable in column 1 is based on the employment status recode variable, which is available only through 1998. Both these factors serve to increase the number of observations in columns 2 and 3.

Figure 2.1: Net Marginal Effects of HS Graduation Unemployment Rate on the Probability of Working, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of working per percentage point increase in the national unemployment rate at the time of high school graduation. Dashed lines represent point-wise 95-percent confidence intervals.

Figure 2.2: Net Marginal Effects of HS Graduation Unemployment Rate on the Probability of Being in the Labor Force, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of being in the labor force per percentage point increase in the national unemployment rate at the time of high school graduation. Dashed lines represent point-wise 95-percent confidence intervals.

hours worked. (Although joint tests over the remaining time horizon also lead to rejection of no effect—p=0.013 for weeks worked, p=0.003 for hours worked—these are harder to interpret given the estimates' change in sign, and the magnitudes are only one-fourth the size as in the first few years.)

The weeks worked by men soon after graduation, like work status, is not affected by the high school unemployment rate (p=0.171 for the first four years), but there does appear to be a slight negative effect on hours worked. More specifically, men cut their annual hours by about 1.5 to 2 percent per percentage point increase in the unemployment rate over their late teens and early 20s (p=0.012). Since there is no reduction in weeks worked, though, this effect stems from a shorter work week. This decline is still significantly smaller than that for women: the null of equal effects between the sexes can be rejected at the 1 percent level for hours worked (and the 5 percent level for weeks worked). Also different from the women is that the effect of the initial unemployment rate has enduring effects on the men. This negative impact starts showing at about 10 years after high school and persists for another 10 years (not fully shown in the table).[24] Over this horizon, men work about 0.5 fewer weeks and 35 to 40 fewer hours per year for each percentage point increase in the unemployment rate. These estimates imply that 10 to 20 years after graduating high school during a severe recession, men's weeks (hours) worked would be about 4 (7) percent lower than if they had graduated during a typical labor market. The persistence in male negative employment effects at the intensive margin suggests that the labor frictions described by Kahn (2010), in which the initial labor demand shock affects accumulation of occupation or industry-specific human capital, may manifest in ways other than wages.

2.4.2 College Enrollment

Because the negative work (and labor supply) effects shown in Table 2.3 last around four years, the typical length of study for a baccalaureate degree, the initial labor demand shock possibly induces women from working into education.[25] The first column of Table 2.5 demonstrates this is not the case. Examining the first four rows of coefficients, there is no effect at all of the unemployment rate on college enrollment rates of women. Men, on the other hand, *are* induced into attending college shortly after graduation, and the magnitudes of these effects are considerable:

[24] The null hypothesis that there is no effect for men from year 5 through the end of the horizon is soundly rejected ($p<0.001$ for weeks worked, p=0.005 for hours worked). Despite the differences in magnitude of the point estimates, the joint effects for men over this interval are not statistically distinguishable from those for women.

[25] Betts and McFarland (1995) and Turner (2003) find positive effects of the unemployment rate on college enrollment; Betts and McFarland show that the effect is concentrated among two-year colleges, and Turner that the effect is much stronger among people who have been out of high school for several years (i.e., who are in their mid and late 20s).

Table 2.5: College Enrollment and Log Wages by Experience Year

	College Enrollment				Log wages	
	Mean enroll, women	(1) *Women*	*Mean enroll, men*	(2) *Men*	(3) *Women*	(4) *Men*
HS Grad UR:						
1 year after	0.472	-0.0048	0.453	0.0173*	-0.0131*	-0.0240***
		[0.0091]		[0.0091]	[0.0074]	[0.0091]
2 years after	0.413	-0.0025	0.421	0.0188**	-0.0153**	-0.0184**
		[0.0075]		[0.0076]	[0.0065]	[0.0081]
3 years after	0.358	-0.0004	0.384	0.0196***	-0.0136**	-0.0164**
		[0.0062]		[0.0065]	[0.0056]	[0.0072]
4 years after	0.315	0.0018	0.366	0.0219***	-0.0115**	-0.0157**
		[0.0052]		[0.0058]	[0.0052]	[0.0067]
5 years after	0.176	-0.0078**	0.233	0.0075	-0.0070	-0.0098
		[0.0038]		[0.0048]	[0.0051]	[0.0065]
6 years after	0.125	-0.0075**	0.161	0.0004	-0.0027	-0.0053
		[0.0031]		[0.0039]	[0.0053]	[0.0064]
7 years after	0.110	-0.0049	0.132	-0.0004	-0.0022	-0.0026
		[0.0030]		[0.0037]	[0.0056]	[0.0066]
8 years after	0.105	-0.0020	0.111	-0.0007	-0.0026	-0.0036
		[0.0030]		[0.0035]	[0.0060]	[0.0069]
9 years after	0.111	0.0027	0.097	-0.0004	-0.0047	-0.0020
		[0.0034]		[0.0035]	[0.0065]	[0.0071]
10 years after	0.086	0.0018	0.087	-0.0001	-0.0042	-0.0008
		[0.0030]		[0.0034]	[0.0069]	[0.0075]
11 years after	0.080	0.0034	0.069	-0.0013	-0.0055	-0.0014
		[0.0030]		[0.0029]	[0.0073]	[0.0078]
12 years after	0.080	0.0055*	0.065	-0.0010	-0.0047	-0.0020
		[0.0031]		[0.0029]	[0.0077]	[0.0082]
13 years after	0.076	0.0072**	0.055	-0.0008	-0.0068	-0.0019
		[0.0031]		[0.0028]	[0.0080]	[0.0084]
14 years after	0.078	0.0091***	0.052	-0.0002	-0.0059	-0.0032
		[0.0033]		[0.0028]	[0.0083]	[0.0086]
15 years after	0.076	0.0112***	0.050	0.0003	-0.0096	-0.0047
		[0.0034]		[0.0028]	[0.0085]	[0.0089]
Wald test 1-4 years zero effect (p-value)		0.775		0.001	0.047	0.040
Wald test 1-4 years, men same as women			0.133		0.075	
(Pseudo) R-squared		0.196		0.280	0.272	0.300
Observations		40,135		35,225	29,862	29,661
Unique individuals		2,211		2,007	2,194	1,988

* significant at 10%   ** significant at 5%   *** significant at 1%

Notes: The estimates in columns 1 and 2 report the marginal effect (averaged across relevant observations) from a probit of a one percentage point increase in the unemployment rate in the year of high school graduation on the likelihood of being enrolled in college x years after high school graduation. Columns 3 and 4 have real log hourly wage as the dependent variable and report estimates from an OLS regression. All regressions include the same controls as those in Table 2.3. Observations are person-years, and the number of unique individuals refers to respondents who have at least one observation in the sample. Standard errors (in brackets) are clustered at the individual level. See text for discussion on inference.

a one percentage point rise in the unemployment rate increases the enrollment rate by about 2 percentage points, or between 4 and 6 percent.[26] Although one just fails to reject that the effects on enrollment are the same for men and women over the first four years as a whole (p = 0.133), the differences for each of these years individually are statistically significant at the 5 percent level (10 percent for the first year).

The effects for men are somewhat higher than what Betts and McFarland (1995) or Turner (2003) find for recent high school graduates, but these studies did not look at enrollment separately by sex. If the non-response of women is averaged in with the positive response of men, the estimates are in line with both studies. As women on the margin of attending college are likely to work considerably less over their lifetime than similar men, and would thus be expected to have a lower return to attending college, their lack of immediate enrollment response is consistent with the theory described in Section 2.2. This effect would be enhanced if credit constraints are more binding during times of weaker labor demand (Christian 2007).[27]

Interestingly, there is a modest but generally statistically significant net negative effect on enrollment rates for women five and six years after graduation. Perhaps initial college-going for the average woman in the sample is relatively inelastic to labor market conditions at high school graduation, but continuing in college (or going to graduate school) is not.[28]

This might be the case, for example, if an individual who has amassed debt to pay for college may be reluctant or unable to accumulate more by continuing one's studies, and instead might prefer to work or at least put off further schooling. Some support for this hypothesis can be found by looking at the rows for the effects 12 or more years after graduation: net effects for women are now significant *and positive* in column 1. (Although Table 2.5 truncates experience years beyond 15, these later interaction effects are also positive through year 20, and are driven by women who have not yet completed college.) In contrast, there are negligible enrollment effects for men.[29] Going back to school may be optimal for women if returns to schooling are larger than returns to work experience (Light and Ureta, 1995), and the value of home production falls relatively quickly after

[26] These enrollment effects are concentrated among the men who continue to work, suggesting that education may be acting as a substitute for on the job training if Kahn's (2010) match-based training job friction story is correct. It is also worth noting that these enrollment effects do translate into higher educational attainment: men obtain about 0.08 more years of schooling per percentage point increase in the high school unemployment rate.

[27] Indeed, stratifying the enrollment analysis by whether the respondent had a father who was a manager or professional, a proxy for household wealth, showed weak positive enrollment effects for women with white-collar fathers and weak negative effects for women with blue-collar fathers, although these results were not statistically significant.

[28] Further analysis showed that the negative effect five to six years out was driven equally by those without a BA and those in graduate school.

[29] The positive effects for women at 12 to 15 years just fail to be significantly different from those of the men (p = 0.118), but they are individually statistically significant for years 13, 14, and 15 (and beyond, although these are not shown).

prime child-bearing years.

2.4.3 Wages

Finally, columns 3 and 4 of Table 2.5 present results for log real hourly wages among working individuals. If women, on average, view home production as a more feasible alternative to working than men, more women at the margin (near their reservation market wage) will choose not to work in the market, as shown in Tables 2.3 and 2.4. Although wage offers are lowered for everyone, the substitution out of working by women at the margin, who have lower average wages, creates positive selection among the women who remain working. As a result, the effect of the initial unemployment rate on observed wages should be smaller for women than for men.

The wage estimates in Table 2.5 show that both men and women experience a temporary negative wage impact after graduating in a weak labor market. A severe recession that raises the unemployment rate by 3 to 4 percentage points reduces women's wages by about 5 percent over the next four years, and men's wages by about 7 percent. The effects for both sexes are statistically significant at the 5-percent level for each of the first four years after graduation as well as jointly over this horizon. Although it is only marginally significant ($p = 0.075$), one can also reject that men and women suffer the same wage impact.

Furthermore, the wage estimates here, even for the men, are a little over half of what Oreopoulos *et al.* (2012) find for similar cohorts of Canadian college graduate men and approximately one-third of what Kahn (2010) finds in her sample of NLSY79 college graduate white men. Additionally, the wage effects fade away more rapidly here than in those studies; by six years out, the wages of both women and men who graduated in a recession are barely distinguishable from those of their more fortunate peers.[30] However, as there are lingering negative effects on the work intensity of men (Table 2.4), there is some persistence of a slight loss in their annual earnings. In context, the results here and in the other papers suggest that if there are wage frictions in the form of implicit contracting or match-based training, their effect on earnings depends integrally on the underlying wage profiles of affected groups. The flatter the wage profile, the less persistent and severe labor-demand-induced wage frictions appear to be.

[30] Joint tests over the horizon from 5 years out show statistical significance ($p=0.044$ for women and $p=0.015$ for men; $p=0.801$ for their difference), but the point estimates are generally less than one-third the size of those over the four years after graduation.

2.4.4 Other Results

While it would be desirable to test directly for the impact on home production using the same methodology, limitations of the data make this infeasible. However, the timing of certain specific channels of home production, notably first marriage and childbearing, can be investigated. One should note that if the unemployment rate affects the marriage market by lowering the wages of potential spouses, a positive own substitution effect toward family formation can be offset by a negative income effect from lower earnings of mates. Although most of the literature on the economic determinants of fertility has found evidence for procyclical fertility since the 1970s (see Butz and Ward, 1979, and Andersson, 2000, for a review), these papers have focused on married women or on all women of childbearing age, with little specific attention toward the young and unmarried. An exception is Kondo (2011), who finds support for opposing own-substitution and mate-income effects: young women who experience higher unemployment rates are likely to marry and have children sooner, but these effects are countered by the unemployment rates that young men face.

In Figures 2.3 and 2.4, I present results from a modification of equation (2.1) that show the impact of the initial unemployment rate on the probability of having been married or having borne a child.[31] While the women evince no change in the likelihood of either event shortly after graduation, the men do shy away from marriage in their early 20s, findings broadly consistent with Kondo (2011), despite her different methodology.

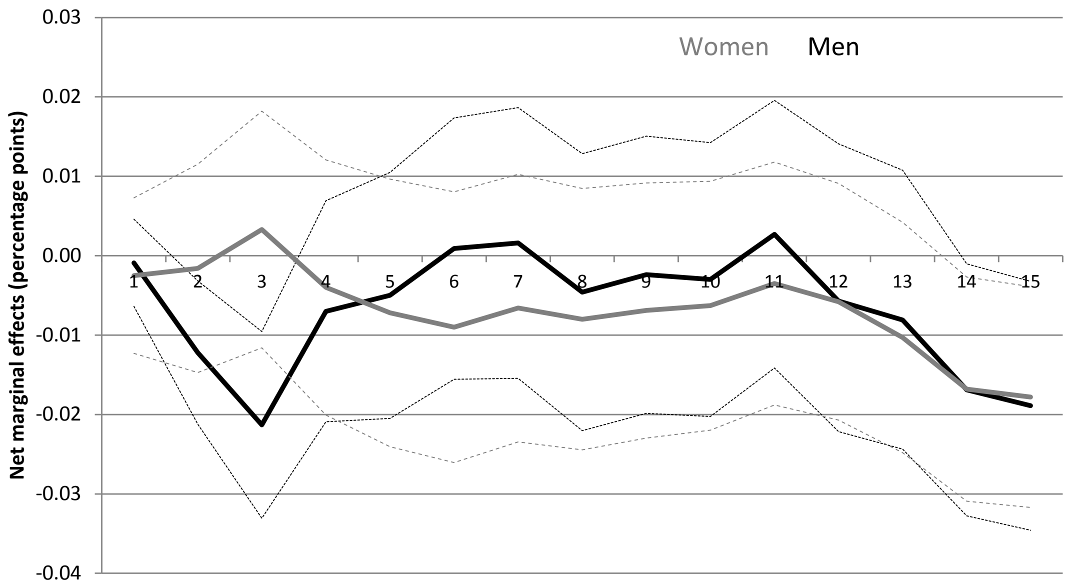
There is some suggestive evidence, however, that women graduating high school in a recession do in fact partially shift toward housework. Using a single cross-section of time-use data in the *NLSY79* from 1981 with state unemployment rates, and parameterizing the impact of the unemployment rate linearly with time elapsed, women (but not men) increased their time spent in housework by about 10 percent per percentage point increase in the unemployment rate, with this effect fading over the next four years. Of course, as the identification strategy is different, this finding should be treated cautiously. Some additional evidence comes from the *American Time Use Survey* (*ATUS*). Using these data, which run from 2003 through 2010, Aguiar, Hurst, and Karabarbounis (2011) demonstrate that in the 2008-2009 recession, both women and men reduced their time spent on market work by about 6 percent, with about one-third of the difference being spent on home production. However, they looked at men and women of all ages and education levels. With the same data, I examined how the national unemployment rate at age 18 affected the daily home-production time

[31] Because these outcomes are one-time events, equation (2.1) is estimated without the time-varying controls. A version that also included a linear trend in year of birth to capture trends unrelated to the unemployment rate produced similar results.

of high school graduates aged 18 through 24.[32]

Women increased this time between 9 and 14 minutes (about 9 percent) per percentage point increase in the unemployment rate between the ages of 18 and 22; men, on the other hand, saw insignificant increases of only 1 to 5 minutes (about 5 percent). While none of these time use results is definitive, and better data would certainly be helpful, the patterns are consistent with women being more likely than men to substitute into home production following graduating high school in a recession.

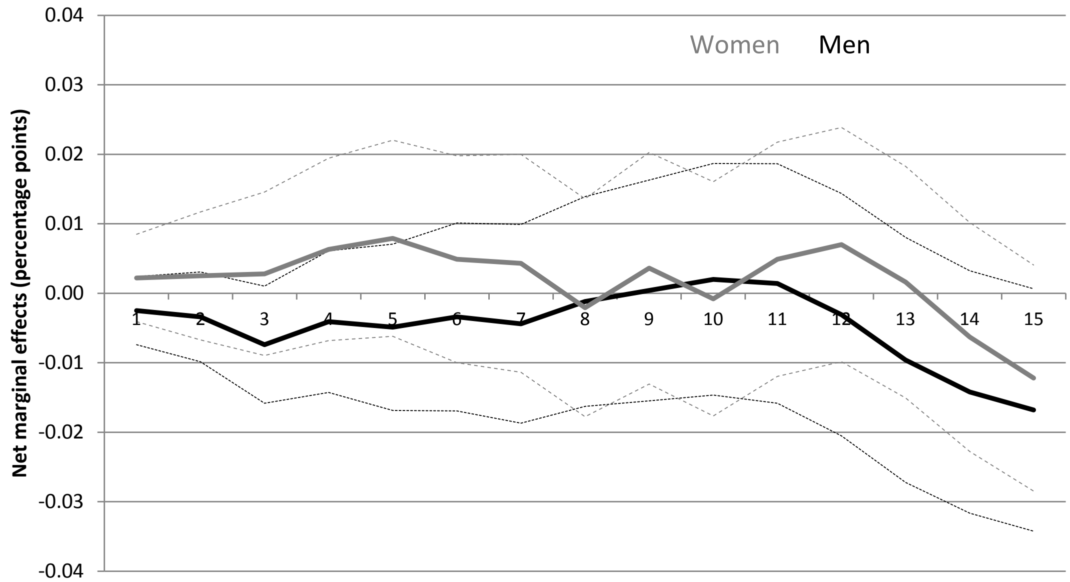
Figure 2.3: Net Marginal Effects of HS Graduation Unemployment Rate on Marriage Probability, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of first marriage per percentage point increase in the national unemployment rate at the time of high school graduation. Dashed lines represent point-wise 95-percent confidence intervals.

[32] More specifically, I combined the time use variables for "household activities" and "caring for and helping household members" and regressed this measure on a variant of (2.1) that excludes family background variables other than race (as they aren't available) but includes day-of-the-week controls. Detailed results are available upon request from the author.

Figure 2.4: Net Marginal Effects of HS Graduation Unemployment Rate on Childbirth Probability, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of first childbirth per percentage point increase in the national unemployment rate at the time of high school graduation. Dashed lines represent point-wise 95-percent confidence intervals.

2.5 Discussion and Conclusion

This paper set out to provide and test the implications of a framework for how men and women (and the more educated and less educated) may vary in their behavioral responses to beginning labor market entry during a recession. In contrast with previous related literature that found a substantial negative wage impact but no effect on labor supply in samples of highly educated men, the empirical evidence presented here suggests a different outcome for the less well educated. Women, but not men, who graduate high school in an adverse labor market are less likely to be in the workforce for the next four years, but longer-term effects are minimal. Further, while men increase their enrollment as a short-run response to weak labor demand, women do not; instead, they appear temporarily to substitute into home production, most likely in the form of housework. Additionally, the wages of high school graduate women who do continue to work—in both the short-run and the long-run—are less affected by the unemployment rate at the time of graduation than are the wages of similarly educated men, and both groups are less affected than the college graduates previously studied.

This analysis extends our understanding of the short and long-run implications of business cycles for workers by focusing on the half of the population—women—that has been understudied in this literature. More directly, it illustrates how certain fundamental economic parameters, such as the return to labor market experience, are integral in determining both how and how long individuals are affected by poor labor markets. Studies that look to identify welfare costs predominantly through lost wages are likely to be inadequate for the less educated, particularly less educated women.

Of course, the absence of long-term labor supply and wage effects does not mean that women who graduate high school in a recession are not affected negatively over the long run. A fertile area for future research exists in exploring how temporary reductions in the likelihood of working found here translate more thoroughly into lifetime income,[33] health, the selection into and duration of marriage, the quantity and quality of investments in children,[34] and many other possible dimensions of social interest.

With the 2008-2009 recession having brought unemployment rates into the double digits for the first time in a generation, some caution should be exercised in generalizing the results found here to the women coming of age today. Blau and Kahn (2007) and others have shown that women's labor supply elasticity has fallen and their return to experience has grown in the past quarter century, although they have not yet approached the levels of men. These trends serve to lower the relative value of home production as an alternative and may concentrate the impact of graduating in a recession more toward wages and less in temporary reductions in working. As a consequence, the longer term ramifications for women graduating high school today may be more substantial than in the past.

[33] See Jacobsen and Levin (1995).
[34] See Hotz and Miller (1988).

2.6 Appendices

2.6.1 Data Appendix

Sample and Variable Construction

The 12,686 individuals initially surveyed in the *NLSY79* comprise different three sample groups: a nationally representative sample of youths aged 14 to 21 in 1979, an over-sample of the poor and racial minorities, and a military sample. Because the focus of the analysis in this paper is on the choices of high school graduates, I use the first of these samples in the analysis. (The over-sample is not used because not all of its respondents are followed for the entire time horizon.) I restrict the estimation sample to individuals who graduated high school on time (at age 17, 18, or 19) and have valid AFQT scores. Appendix Table 2.2 provides details about how conditioning the sample affects the sample size.

Details on the construction of the key variables used in the analysis are presented below.

Unemployment Rate: The national unemployment rate is the non-seasonally-adjusted annual average for all labor force participants age 16 and older as calculated by the Bureau of Labor Statistics (BLS) from the Current Population Survey (`www.bls.gov/cps`).

AFQT: The Armed Forces Qualifying Test is derived from four of the ten components of the Armed Services Vocational Aptitude Battery (ASVAB), with the score on each component being a variable in the NLSY79. Specifically, the raw AFQT score is given by: *arithmetic reasoning score* + *word knowledge score* + *paragraph comprehension score* + 0.5**numerical operations*. I regress this raw AFQT score on year of birth dummies using the 1979 probability sampling weights. The residuals from this regression are converted into z-scores by subtracting the mean and dividing by the standard deviation within each year of birth.

Labor Force Participation: This binary variable is created from the employment status recode variable for each survey year through 1998, the last year in which the employment status recode variable exists. It is coded 1 if the respondent has a current job or is looking for one and coded 0 otherwise.

Working: This binary variable is also created from the employment status recode variable for each survey year through 1998. It is coded 1 if the respondent has a current job and 0 otherwise.

Annual Weeks Worked and Annual Hours Worked: The *NLSY79* has a complete work history section beginning in January, 1978. Each calendar week describes the labor force status of the individual and the number of hours worked. These variables are summed across the weeks in a calendar year to create annual measures.

College Enrollment: This binary variable is created within the *NLSY79* based on start and stop dates at post-secondary institutions and indicates whether the individual was enrolled as of May 1 of the calendar year of the survey.

Log wages: The *NLSY79* asks for the hourly rate of pay for up to five jobs in every survey wave. I construct hourly wages using the rate of pay variable from the first job (which is the current or most recent job) among respondents who are employed at the the time of the survey. These wages are converted to 1977 dollars using the CPS-U-RS and then are transformed by natural logarithm. Outliers below $1 or above $200 (in 1977 dollars) are excluded from analysis.

Cohort Size Effects

The *NLSY79* cohorts are born from 1957 through 1964 and consist of the younger Baby Boomers, an unusually large cohort. The fertility rate, or the number of births per 1000 women aged 15 to 44, peaked in 1957 at around 120 and then began a steady decline that lasted until the mid 1970s, reaching a low of around 65. Nonetheless, the Baby Boom is generally dated as lasting through 1964 because the fertility rate, though falling, remained historically high—it was about 105 in 1964, approximately the level in 1949, three years after the start of the Baby Boom.[35] The sheer size of the cohort may have important implications on schooling and labor market decisions. Falaris and Peters (1992) show that the size of both past and future cohorts (that is, the timing of birth relative to whether the birth rate is rising or falling) affects both the amount of education an individual receives and the age at which one completes formal schooling. Specifically, cohorts born during the upswing of the cycle tend both to get more education and take longer per additional year to get it than do cohorts born during the cycle downswing; cohorts born at peaks or troughs fall in between. Thus, in the *NLSY79*, we might expect to see slightly less education and earlier labor market entry for the younger cohorts. However, Falaris and Peters find that the cyclical effects for women, while statistically significant, are quite small relative to those for men. The authors hypothesize that the gender difference may be due to women's smaller total labor supply and thus weaker incentives to obtain more school in order to mitigate the negative wage effects of excess supply. This explanation can nest with business cycle effects on women's labor supply, but it suggests caution in disentangling the demographic cycle from the business cycle.

A potential shortcoming of the national rate regressions is the inability to control for cohort size. By restricting the sample to women who graduated high school at more or less the same age, any indicator for cohort size would be almost perfectly collinear with the national unemployment rate in the equation. If Falaris and Peters are right, however, omitted variable bias from missing cohort effects in the national rate analysis should be trivial. Furthermore, since cohort size is falling with time in the *NLSY79* sample, the results of Welch (1979) suggest that the younger cohorts should be faced with higher wage offers (and, hence, incentives to participate in the labor market), *ceteris paribus*, than the older cohorts. But it is the younger cohorts in the sample who experienced the highest unemployment rates upon graduation: 9.7 for the 1982 grads and 9.6 for the 1983 grads. Thus, to the extent that cohort effects are present, we would expect the bias to go *against* finding negative unemployment rate effects.[36]

Sample Attrition

Approximately one-third of the viable sample has attrited by survey year 2004. However, most of the attrition occurs relatively late in the sample (mid 1990s and afterward). Thus, if much of

[35] These numbers are from U.S. Vital Statistics: `http://www.cdc.gov/nchs/products/pubs/pubd/vsus/vsus.htm`.

[36] A crude approach to control for cohort effects in the national unemployment rate regressions is to include a linear time trend in year of birth or year of high school graduation. Doing so, however, has no notable effect on any of the unemployment rate coefficient estimates.

the effect of the initial labor demand shock is concentrated relatively soon after graduation, as the framework suggests, estimation should not be significantly plagued by sample attrition. Nevertheless, to address the potential problem, each estimation equation was run on two samples, one using all available person-years and the other using only individuals who were interviewed every survey year. The resulting sets of estimates were not appreciably different; sample attrition does not seem to be a major problem.

Appendix Table 2.1: Sample Means of Standardized AFQT by High School Graduation Cohort

High School Graduation Year	*Mean*	*Std error*	*Significantly different from:*
1975	0.296	0.051	1976**, 1983***
1976	0.115	0.048	1977*, 1980**, 1982*, 1983***
1977	0.231	0.047	1983***
1978	0.208	0.048	1983***
1979	0.183	0.046	1983***
1980	0.278	0.043	1983***
1981	0.183	0.049	1983***
1982	0.236	0.052	1983***
1983	-0.543	0.129	–

Notes: Significance levels are from t-tests of differences in means for the specified years. (* significant at 10%, ** significant at 5%, *** significant at 1%.)

Appendix Table 2.2: Sample Sizes and Attrition in the NLSY79 Cross-Section Samples

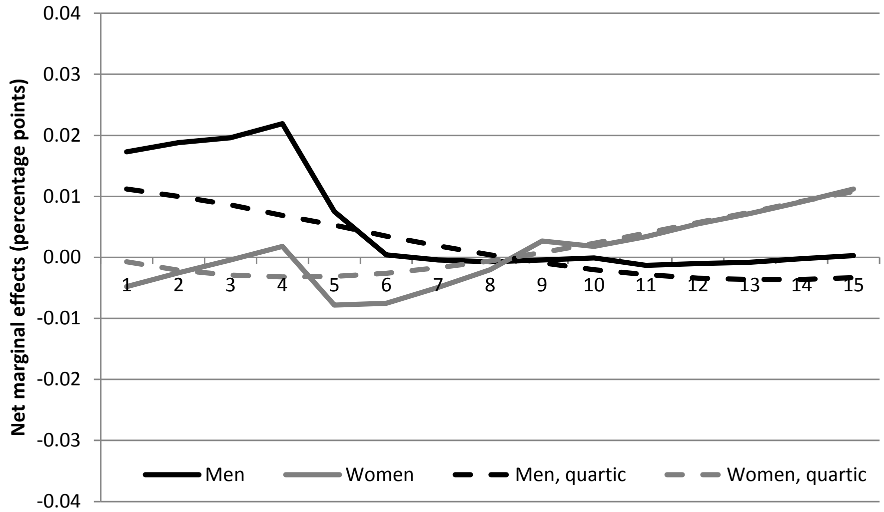
Women

	Number	Percent	Percent of HS on-time grads
1) Female respondents in cross-section sample	3108	100.0%	–
2) + also graduate high school	2409	77.5%	–
3) + also graduate at age 17-19 (on time)	2337	75.2%	100.0%
4) + also have valid AFQT score	2211	71.1%	94.6%

Men

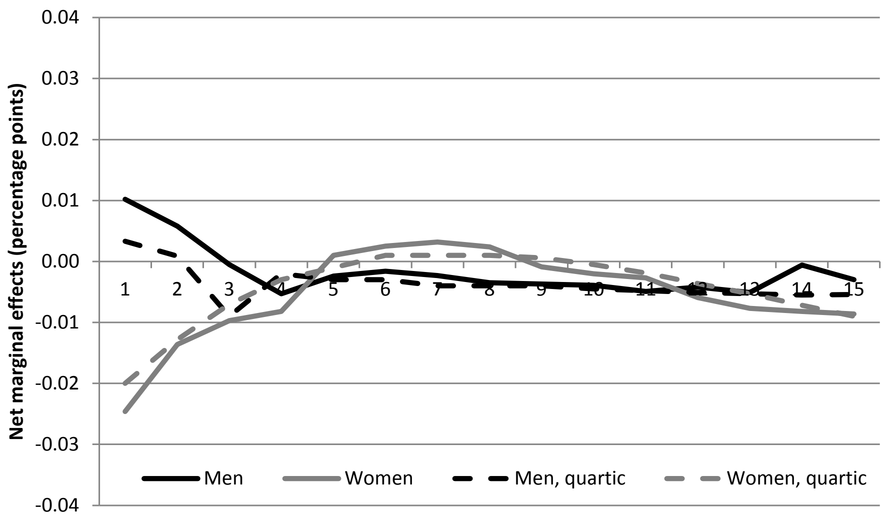
	Number	Percent	Percent of HS on-time grads
1) Male respondents in cross-section sample	3003	100.0%	–
2) + also graduate high school	2208	73.5%	–
3) + also graduate at age 17-19 (on time)	2149	71.6%	100.0%
4) + also have valid AFQT score	2006	66.8%	93.3%

Appendix Figure 2.2: Net Marginal Effects of HS Graduation Unemployment Rate on the Probability of College Attendance, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of being enrolled in college per percentage point increase in the national unemployment rate at the time of high school graduation. Solid lines show effects estimated by year; dashed lines show effects from a quartic in actual time since graduation.

Appendix Figure 2.1: Net Marginal Effects of HS Graduation Unemployment Rate on the Probability of Working, by Years Since Graduation



Notes: Data represent the net marginal effect (in percentage points) on the likelihood of working per percentage point increase in the national unemployment rate at the time of high school graduation. Solid lines show effects estimated by year; dashed lines show effects from a quartic in actual time since graduation.

2.7 References

Andersson, Gunnar. 2000. "The Impact of Labour-Force Participation on Childbearing Behaviour: Pro-Cyclical Fertility in Sweden during the 1980s and the 1990s." *European Journal of Population* 16(4): 293-333.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. 2011. "Time Use During Recessions." NBER Working Paper No. 17259.

Beaudry, Paul, and John DiNardo. 1991. "The Effect of Implicit Contracts on the Movement of Wages Over the Business Cycle: Evidence from Microdata." *Journal of Political Economy* 99(4): 665-688.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1): 249-275.

Betts, Julian, and Laurel McFarland. 1995. "Safe Port in a Storm: The Impact of Labor-Market Conditions on Community College Enrollments." *Journal of Human Resources* 30(4): 741-765.

Blau, Francine D., and Lawrence M. Kahn. 2007. "Changes in the Labor Supply Behavior of Married Women: 1980-2000." *Journal of Labor Economics* 25(3): 393-438.

Brunner, Beatrice, and Andreas Kuhn. 2009. "To Shape the Future: How Labor Market Entry Conditions Affect Individuals' Long-Run Wage Profiles." IZA Discussion Paper No. 4601.

Burda, Michael C., and Daniel S. Hamermesh. 2010. "Unemployment, Market Work and Household Production." *Economics Letters* 107(2): 131-133.

Butz, William, and Michael Ward. 1979. "The Emergence of Countercyclical U.S. Fertility." *American Economic Review* 69(3): 318-328.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3): 414-427.

Christian, Michael S. 2007. "Liquidity Constraints and the Cyclicality of College Enrollment in the United States." *Oxford Economic Papers* 59(1): 141-169.

Elsby, Michael W. L., Ryan Michaels, and Gary Solon. 2009. "The Ins and Outs of Cyclical Unemployment." *American Economic Journal: Macroeconomics* 1(1): 84-110.

Falaris, Evangelos M., and H. Elizabeth Peters. 1992. "Schooling Choices and Demographic Cycles." *Journal of Human Resources* 27(4): 551-574.

Genda, Yuji, Ayako Kondo, and Souichi Ohta. 2010. "Long-Term Effects of a Recession at Labor Market Entry in Japan and the United States." *Journal of Human Resources* 45(1):157-196.

Goldin, Claudia. 2006. "The Quiet Revolution that Transformed Women's Employment, Education, and Family." Richard T. Ely Lecture. *American Economic Review* 96(2): 1-21.

Hall, Robert E. 2005. "Employment Efficiency and Sticky Wages: Evidence from Flows in the Labor Market." *Review of Economics and Statistics* 87(3): 397-407.

Heckman, James J., and Paul A. LaFontaine. 2007. "The American High School Graduation Rate: Trends and Levels." NBER Working Paper No. 13670.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." *Handbook of the Economics of Education*, Volume 1. E. Hanushek and F. Welch, eds. Amsterdam: Elsevier Science Publishers. 307-458.

Hotz, V. Joseph, and Robert A. Miller. 1988. "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply." *Econometrica* 561(1): 91-118.

Hotz, V. Joseph, Lixin Colin Xu, Marta Tienda, and Avner Ahituv. 2002. "Are There Returns to the Wages of Young Men from Working While in School?" *Review of Economics and Statistics* 84(2): 221-236.

Jacobsen, Joyce P., and Lawrence M. Levin. 1995. "Effects of Intermittent Labor Force Attachment on Women's Earnings." *Monthly Labor Review* 118(9): 14-19.

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

Kline, Patrick, and Andres Santos. 2011. "Higher Order Properties of the Wild Bootstrap Under Misspecification." NBER Working Paper No. 16793.

Kondo, Ayako. 2008. "Differential Effects of Graduating During A Recession Across Race and Gender." Columbia University mimeo.

Kondo, Ayako. 2011. "Gender Specific Labor Market Conditions and Family Formation." Osaka University mimeo.

Lechner, Michael, and Stephan Wiehler. 2007. "Kids or Courses? Gender Differences in the Effects of Active Labor Market Policies." IZA Discussion Paper No. 2740.

Light, Audrey, and Manuelita Ureta. 1995. "Early-Career Work Experience and Gender Wage Differentials." *Journal of Labor Economics* 13(1): 121-154.

Moulton, Brent R. 1986. "Random Group Effects and the Precision of Regression Estimates." *Journal of Econometrics* 32(3): 385-397.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession: Hysteresis and Hetereogeneity in the Market for College Graduates." *American Economic Journal: Applied Economics* 4(1): 1-29.

Oyer, Paul. 2006. "Initial Labor Market Conditions and Long-Term Outcomes for Economists." *Journal of Economic Perspectives* 20(3): 143-160.

Oyer, Paul. 2007. "The Making of an Investment Banker: Stock Market Shocks, Career Choice, and Lifetime Income." Stanford GSB mimeo.

Ruhm, Christopher. 1997. "Is High School Employment Consumption or Investment?" *Journal of Labor Economics* 15(4): 735-776.

Turner, Sarah E. 2003. "Pell Grants as Fiscal Stabilizers." University of Virginia mimeo.

Welch, Finis. 1979. "The Effect of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy* 87(5) Part 2: Education and Income Distribution: S65-S97.

CHAPTER III

The Opt-In Revolution? Contraception and the Gender Gap in Wages

3.1 Introduction

During the 1980s, the long-standing U.S. gender gap in pay narrowed rapidly. The median annual wage and salary earnings of women working full-time, full-year rose from roughly 60 percent of men's earnings in 1979 to 69 percent a decade later. Not only was this a striking departure from the stability of women's relative pay during the 1970s, but the speed of women's convergence in the 1980s was also faster than during the 1990s and the 2000s.

The correlates of the narrowing of the gender gap in the 1980s are well documented: the decade witnessed a convergence in measured labor market skills between men and women. Expecting to remain in the labor-force longer, women born in the 1950s (who came of age in the 1970s) narrowed the gender gap in college going and completion, attaining professional degrees, and working in non-traditionally female occupations (Goldin 2004, 2006). Increases in demand for skills that benefited women relative to men increased the returns to women's investments in market skills (Blau and Kahn 1997, Welch 2000). Widening wage inequality among women may have also encouraged women to invest in market skills and led more able women to select into full-time employment (Mulligan and Rubinstein 2008). Each of these factors may have contributed to and resulted from the growth in women's work experience (O'Neill and Polachek 1993, Wellington 1993).

The root causes of these tremendous changes are less clear. Two important but elusive candidates include the resurgence of the women's movement in the late 1960s and early 1970s and the new legal protections afforded to women under the 1964 Civil Rights Act (and later federal enforcement) that reduced overtly-discriminatory hiring and compensation practices-both of which should have changed attitudes and norms about women's employment. Recent literature suggests oral contraception, often called "the Pill," as another important candidate. Its diffusion to younger, unmarried women improved their ability to time births, altered their expectations about future childbearing, and reduced the cost of altering career investments to reflect their changed expectations. The timing of its diffusion during the 1960s and 1970s also fits well with the slow growth in women's wages during the 1970s (as younger women invested more in their human capital) and the rapid convergence in the gender gap during the 1980s (when these women enjoyed the returns on their human capital investments and accumulated labor-market experience). To quantify the importance of the Pill, Goldin and Katz (2002) use state-by-birth-cohort changes in the age of consent from 21 to 18 for medical care and, thereby, prescription birth control. Based upon extensions of this empirical strategy, the recent literature links "early access to the Pill" to delays in marriage (among college goers) and motherhood, changes in selection into motherhood, increased educational attainment, labor-force participation, and occupational upgrading among college graduates (Goldin and Katz 2002, Bailey 2006, Guldi 2008, Hock 2008, Ananat and Hungerman 2012). Although these studies imply that the Pill benefitted individual women's careers, its effect on aggregate wages need not be large or even positive due to changes in the composition of working women and increased labor supply. No study, however, has considered the impact of these many changes on the gender gap in compensation.

This chapter examines the role of the Pill in altering women's life-cycle wages and its ultimate implications for convergence in the gender gap during the 1980s and 1990s. Following earlier work, our empirical strategy leverages state-by-birth-cohort changes in laws reducing the age of consent for medical care and access to prescription birth control for unmarried women under age 21. We extend the literature by providing two new tests of this empirical strategy's identifying assumptions. Using the 1970 *National Fertility Study*, we show that early access laws doubled Pill use among women between the ages of 18 and 20-precisely the ages affected by access laws-but not beyond age 21, when the laws did not bind. In addition, we test the excludability of Pill access laws (i.e., the assumption that early legal access to the Pill was conditionally, randomly assigned) using the *National Longitudinal Survey of Young Women* (*NLS-YW*). Among 18 family background characteristics that should not have been affected by these legal changes, early access to the Pill is correlated with only one at the 10 percent level-no more than would be expected by chance.

Using longitudinal wage information from the *NLS-YW*, our main results show that early access to the Pill *lowered* women's wages in their early twenties (corresponding to the 1970s) but raised their wages in their thirties and forties (corresponding to the 1980s and 1990s). By their late forties, women with early access to the Pill earned a statistically-significant hourly premium of 8 percent-enough to account for between a third and half of the total hourly wage gains for these cohorts over their peers born a decade earlier. Consistent with the well-known relationship of women's wage growth to cumulative labor-force experience, our decomposition indicates that almost two thirds of the Pill-induced wage premium at the mean is explained through its effect on women's labor-force experience. Another third of the premium is due to changes in educational attainment and occupational choice.

The *NLS-YW* also sheds light on the mechanisms for these effects. Stratifying our sample by measures of high school "IQ score" reveals that the flexibility conferred by the Pill had no measurable impact on the education or experience of lower IQ women. Both middle and higher IQ women, however, raised their educational attainment in their twenties and, in their thirties, acquired more labor-market experience and increased their representation in non-traditionally female occupations. Interestingly, the Pill's largest effects on work experience accrued to women in the middle of the IQ distribution with some college, not to the high-achieving women who have been the focus of earlier studies. In keeping with this finding, early access to the Pill had the largest impact on the lifecycle wages of women in the middle of the IQ distribution. Thus, the rapid narrowing of the gender gap during the 1980s reflected, in part, a Pill-induced revolution in middle-ability women planning for and opting into paid work.

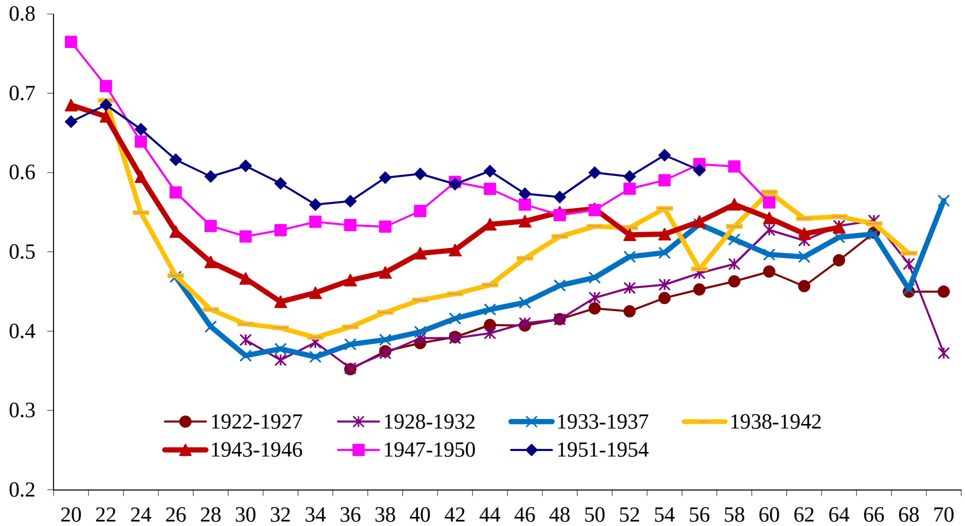
3.2 The Revolution in Women's Work

Aggregate statistics documenting women's wages from the 1950s and 1960s only hint at the tremendous changes in women's earning capacity. Goldin (1990: Table 3.1) shows that women's real wages fell relative to men's from the 1950s to the 1960s; from the 1960s through the mid-1970s, the gap in pay remained constant at roughly 60 percent (Blau, Ferber and Winkler 2010: Figure 51). Beginning in the 1980s, the gender gap in wages narrowed substantially. Although this narrowing has continued to the present, its pace has slowed since the mid-1990s. To provide context for our cohort-age based investigation, this section uses the 1964 to 2009 March *Current Population Surveys* (*CPS*) to describe by age and cohort the changes in women's wages and labor-force outcomes, what Goldin (2006) dubbed the "quiet revolution."[1] We also present statistics *relative to men* to underscore the convergence in outcomes.

Figure 3.1 shows the evolution of mean annual wage and salary earnings in 2000 dollars (PCE deflator) for seven different birth cohorts of women relative to men-a measure of the age-specific gender gap for the following cohorts: those born from 1922 to 1927 (called mid-1920s), 1928 to 1932 (early 1930s), 1933 to 1937 (mid-1930s), 1938 to 1942 (early 1940s), 1943 to 1946 (mid-1940s), 1947

[1] We use CPS rather than the NLS, because the CPS contain information on older cohorts and their larger sample sizes make our series less noisy. Data from the NLS-YW augment this discussion when informative.

Figure 3.1: The Evolution of the Real Annual Wage Earnings of Women Relative to Men by Age and Birth Cohort



Notes: Annual labor earnings include income from all jobs, including self-employment. The series is adjusted for inflation to year 2000 dollars using the personal consumption expenditures deflator (BEA 2009). Data are weighted using CPS sample weights and collapsed into two-year age groups. Source: 1964-2009 March CPS.

to 1950 (late 1940s), and 1951 to 1954 (early 1950s).[2] For cohorts born before the 1940s, the relative wage series have similar age profiles. Beginning with cohorts born in the early 1940s, the gender gap increases less rapidly (i.e., the pay of women relative to men falls less rapidly) in women's twenties and rebounds more quickly after age 30. For 34 year-olds, annual incomes increased from 39 percent of similarly aged men for the 1938 to 1942 cohort to 55 percent for cohorts born less than a decade later.

Large changes in relative wage and salary earnings followed dramatic relative increases in women's *pre-market* and *post-entry* career investments. Goldin, Katz, and Kuziemko (2006) show that the share of women (relative to men) attending and completing college accelerated for cohorts born after the mid-1930s. Labor-force participation during the childbearing years grew rapidly as well. At the extensive margin, participation of 30-year-old women born in the mid-1940s increased by 16 percentage points (from a base of 39 percent) over cohorts born a decade earlier. For women born in the early 1950s, this statistic increased another 14 percentage points.[3] Because the labor-force participation of men was stable over this period, these increases imply a narrowing in the cohort-based gender gap in participation, shown as a flattening of the relative labor-force participation

[2]This divides the cohorts of the National Longitudinal Surveys of Mature and Young Women into roughly equal-sized groups. Wage and salary earnings in Figure 3.1 exclude farm, business or self-employment income. Our sample excludes those who report zero earnings, but Figure 3.1 makes no further sample restrictions.

[3]Statistics for women alone are computed using the March *CPS*, but *only* statistics relative to men are presented for brevity.

series plotted in Figure 3.2, panel A). Women's greater labor-force participation also translated into considerably more work experience (cf. O'Neill and Polachek 1993, Wellington 1993). In the NLS-YW, we calculate that women born in the early 1950s worked 3000 more hours between ages 24 and 40 than did women born in the mid-1940s—an increase of 1.5 full-time, 50-week years.[4]

Changes in the nature of women's work for pay—along with their experience—also coincide with the narrowing of the cohort-based gender gap. The fraction of women working in professional or managerial jobs in their mid-thirties was roughly twice as high for cohorts born in the mid-1940s as for cohorts born a decade earlier. Figure 3.2, panel B, shows that, after accounting for the increase in the share of men working in professional and managerial jobs, women's representation in these fields at age 30 increased by 25 percentage points between the cohorts born in the early and late 1940s and another 24 percentage points for cohorts born in the early 1950s.

Although the remarkable, late-twentieth-century transformation in women's careers is well known, its catalysts are less well understood. Women may have been pulled into the labor force by changes in demand reflecting increasing enforcement of anti-discrimination legislation or skill- (and gender-) biased technological change (Welch 2000, Black and Juhn 2000, Weinberg 2000, Black and Spitz-Oener 2010). At the same time, rapidly changing ideas about women's work and roles in the workplace (Fernandez, Fogli, and Olivetti 2004, Fernandez and Fogli 2009, and Fortin 2009), shifts in divorce rates (Stevenson and Wolfers 2007), and the availability of better colleges and better education at the same colleges (Goldin and Katz 2010) may have increased the supply of women's skills to the market. The next sections describe the potential importance of the Pill for young women's decisions and wages and outline our empirical strategy for quantifying its role within the broader social and economic changes of the last 40 years.

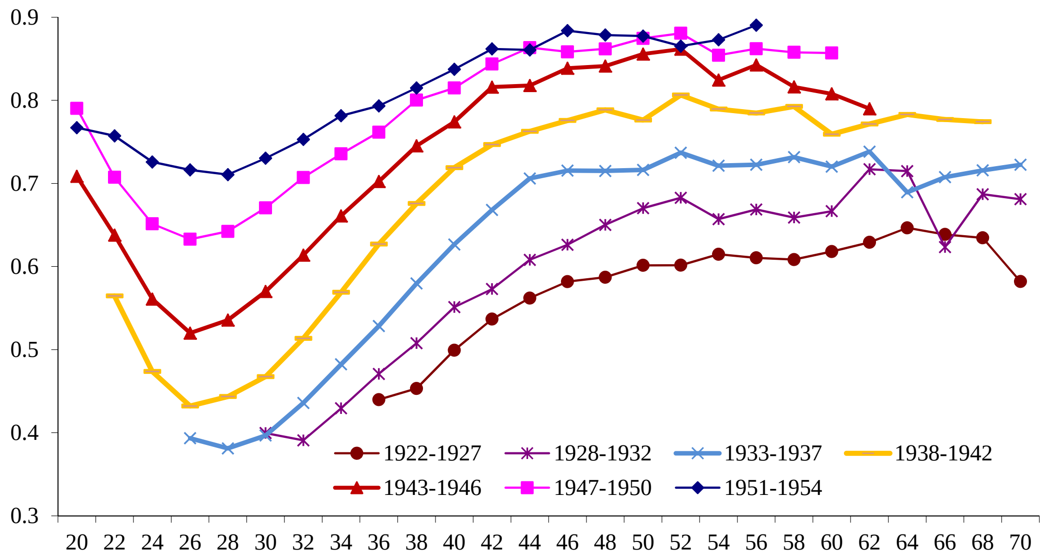
3.3 Was This an Opt-In Revolution? The Expected Effects of Changes in Pill Access on Women's Lifecycle Wages

The diffusion of oral contraception, first released for the regulation of menses in 1957 and approved by the U.S. Food and Drug Administration as a contraceptive in 1960, had an important impact on younger women's ability to time births and plan future childbearing. Women born in the early 1940s (who would be young adults in the early 1960s) would have been the first with access to the Pill in late adolescence when they made decisions about family formation, childbearing, and

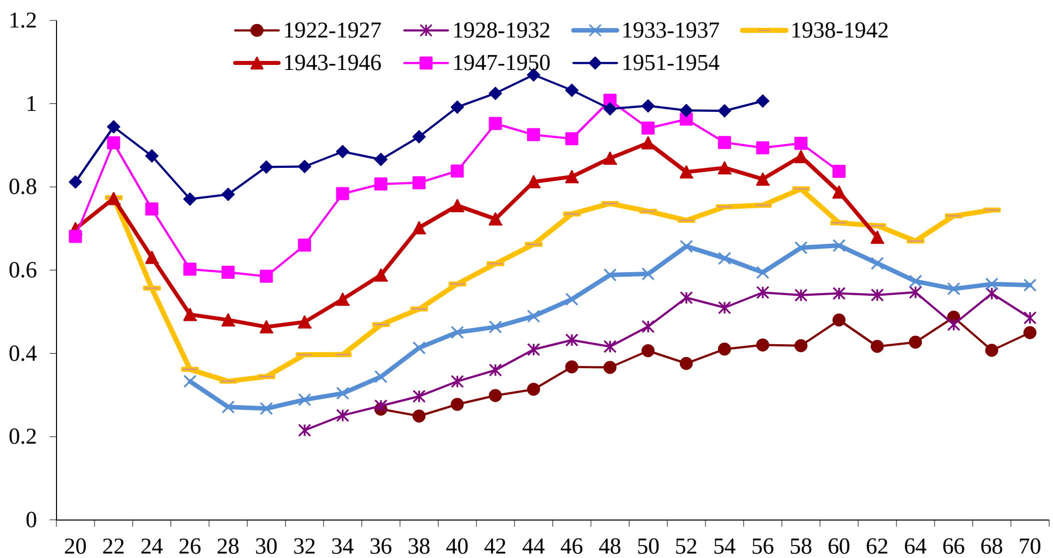
[4] We cannot compare these estimates with cohorts born earlier than the mid-1940s, as the Mature Women were first interviewed when they were between the ages of 30 and 45. Therefore, we are missing information on these older cohorts' labor-force participation at younger ages. For construction of these experience measures, see Appendix 3.A.

Figure 3.2: The Evolution of Human Capital Investments by Age and Birth Cohort

A. Share of Women Participating in the Labor Force Relative to Men



B. Share of Women Working in Professional and Managerial Jobs Relative to Men



Notes: Share participating in the labor force is constructed from a binary variable indicating whether the respondent was employed or looking for a job at the time of the survey. Job groups are coded using the 3-digit Census occupational codes in the CPS. Women are counted in a job category only if they are employed at the time of the survey. Data are weighted using CPS sample weights and collapsed into two-year age groups. Source: 1964-2009 March CPS.

career investments. They would have also been the first to gain autonomy in deciding to use contraception (rather than sharing it with their partners), the first to be able to make decisions about contraception at a time separate from intercourse, and the first to benefit from the reliability and *expectation* of birth predictability the Pill conferred over the entirety of their childbearing years and early careers. Changes in *expectations* are key. Even women who would not have married or had a child before age 22 without the Pill may have altered their career investments as their expectations about future childbearing changed.

The difficulty of parsing the Pill's effect on women's wages relates to the timing of its appearance. By cause or coincidence, the Pill's diffusion coincided with important changes in norms and ideas about women's work and the end of the baby boom. Following Goldin and Katz (2002) and Bailey (2006), our empirical strategy makes use of state-level variation *within* birth cohorts in "early legal access to the Pill" (*ELA*), which allowed younger women to consent for medical care. As described in Bailey (2006), most legal changes were due either to judicial expansions in the rights of legal minors or to legislative changes that lowered the age of majority to 18. The timing of changes in *ELA* differed considerably across states (the earliest change was in 1960 and the latest in 1976), but the common feature of these laws is that they gave physicians latitude to prescribe oral contraception to unmarried women under 21 without consulting parents (Paul, Pilpel, and Wechsler 1974, 1976). State-by-birth cohort variation in *ELA*, therefore, facilitates comparisons of labor-force outcomes for women who gained legal access to the Pill earlier (typically at their 18th birthdays) to those who gained access at 21.

This three-year difference in access to the Pill during a formative life stage potentially affected a host of decisions. Having access to the Pill at age 18, for instance, directly reduced the cost of delaying childbearing and marriage to enter or stay in college.[5] Even among those who did not attend college, better fertility control reduced the cost of remaining at a job long enough to obtain a promotion or additional training. In addition to decreasing the costs of investing, access to the Pill at 18 may have altered the *expected* returns to early human capital investments. All else equal, the same early human capital investment would yield larger *expected* lifetime returns if women *anticipated* being in the labor force more or being more successful in achieving their career aspirations. In short,

[5] A lower risk of childbearing at ages 18 to 19 may have also affected when and whom women married, which could have an independent effect on their careers (Chiappori and Oreffice 2008). Staying in college longer could allow marriage to a more educated man and, therefore, increase a woman's nonwage income and reduce her labor-supply (Ge 2008). On the other hand, staying in college longer should increase a woman's own earnings and, therefore, increase her options outside of marriage. If this leads to greater divorce, women would have lower nonwage incomes and, therefore, tend to work more at older ages (and younger ages, to the extent that women are risk averse and forward looking). For both reasons, marriage delay may improve women's career outcomes independently of fertility delay (Loughran and Zissimopolous 2009, Miller 2011).

earlier access to the Pill should have both reduced the costs of and increased the expected returns to early career investments—predictions consistent with the empirical literature: Hock (2008) and Ananat and Hungerman (forthcoming) show Pill access affected college enrollment and education; Bailey (2006) shows that it increased women's labor-force attachment; and Goldin and Katz (2002) find that it increased college women's representation in non-traditionally female professions.

This theoretical framework suggests three (potentially reinforcing) mechanisms linking *ELA* to steeper wage and salary earnings profiles. First, *ELA* may have increased labor-force participation, which enabled women to accumulate more labor-market experience and job- or firm-specific capital. Thus, women with *ELA* would experience more rapid wage growth. We call this mechanism the "experience mechanism." Second, women with *ELA* may have shared the costs of gaining on-the-job human capital by accepting lower initial wages but then enjoyed larger wage growth with tenure. We call this channel the "on-the-job-investment mechanism." Third, *ELA* may have increased school enrollment and participation in training programs, which should lower wage earnings at younger ages, and increase them following school exit. We call this channel the "formal human-capital investment mechanism."

Our empirical estimates of the effect of *ELA* on wages should be interpreted cautiously for two reasons. The first relates to the off-setting effects of labor supply. Because *ELA* could increase labor-force participation for large numbers of women thus reducing the capital-to-labor-ratio, its effect on any one woman may be larger than its effect on an entire birth cohort, which our analysis recovers. The magnitude of these supply-side effects depends (among other things) on the degree of substitutability of male and female labor in production. The closer substitutes men and women are in production, the smaller the labor-supply effect and the more likely the overall effect of *ELA* on wages will be positive (due to its effect on human capital accumulation). Our analysis recovers estimates that include this labor-supply effect, so our estimates will tend to understate the effect of the Pill on an individual woman's wages, especially in the shorter-run (at younger ages) before firms adjust their capital stock.

The second reason relates to selection. Because wages are only observed for labor market participants, the observed impact of *ELA* on women's wage growth will be larger than the effect on the average woman if the Pill differentially affects human capital investments and labor supply of higher ability women. If, for instance, early access to the Pill causes higher ability woman to continue in their education and makes them less likely to work in their early twenties, then the *ELA*-induced growth in wages will reflect both the returns to these greater investments and changes in the composition of working women to favor those of higher ability. Our analysis explores these compositional effects explicitly by breaking our sample into three IQ tertiles (based upon a composite developed from high school aptitude tests) and examining the effects of *ELA* for women within each of these tertiles.

3.4 Data and Empirical Strategy for Identifying the Impact of the Pill on Wages

Our analysis uses the rich, longitudinal data of the *National Longitudinal Survey of Young Women* (*NLS-YW*), which contains interviews beginning in 1968 for 5,159 women, ages 14 to 24, with 21 subsequent interviews. Crucial is that the *NLS-YW* sampled women born from 1943 to 1954, cohorts that varied in their early legal access to the Pill. Although this data set is smaller than those used in earlier studies, the restricted version contains information on the legal state of residence for the respondents at age 21. We use residence at age 21 (which should be reported as parents' residence for unmarried, college women) to infer treatment status with considerably less error than previous studies.[6]

The *NLS-YW* confers several additional advantages. It contains a rich set of pre-treatment outcomes for testing the validity of our empirical strategy and also facilitates an analysis of heterogeneity in the impact of the Pill by socio-economic status and high school IQ of the respondent, which allows us to understand the ways in which the Pill influenced the selection of women into paid work.[7] Finally, the *NLS-YW* provides information on women's wage earnings in every survey year as well as their career investments including educational attainment, job training and certification, and labor-force participation (weeks and hours). Repeated reports of women's labor-force participation allows us to construct measures of their cumulative labor-force experience and link the Pill to this important correlate of women's wage gains.

3.4.1 Empirical Specification

Our empirical strategy follows the previous literature with several modifications. We estimate the following linear regression models for continuous dependent variables,

$$Y_{iacs} = \sum_g \beta_g ELA_{cs} D_{g(a)} + \sum_g \lambda_g D_{g(a)} + \sum_s \lambda_s D_s + \sum_c \lambda_c D_c + \eta_{iacs}, \tag{3.1}$$

[6] Restricting the sample to those with valid date of birth (cohort) and state of residence information reduces the sample to 4354. Both Goldin and Katz (2002) and Bailey (2006) use repeated cross-sections that contain no information on an individual's state of residence at ages 18 to 21. As a result, Goldin and Katz (2002) and Bailey (2006) infer *ELA* based upon the reported birth state or state of residence *at the time of the survey* respectively.

[7] Appendix 3.A describes the survey questions and coding of each variable.

where Y is the outcome of interest for individual i, at age a, who was born in year c = 1943, 1944, ..., 1953 (also referred to as "birth cohort"), and residing in state s = 1, 2, ..., 51 at age 21. Fixed effects for state of residence, $\sum_{s=2}^{51} \lambda_s D_s$ where $D_s = 1$ if i resided in state s at age 21, and single year-of-birth cohorts, $\sum_{c=1944}^{1953} \lambda_c D_c$ where $D_c = 1$ if i was born in year c, are included in all specifications. The dummy variables $D_{g(a)}$ are set to 1 if the respondent's age fell into the five-year age group, g (14-19, 20-24, ..., or 45-49). Standard errors for all models are robust to heteroskedasticity and clustered at the state level.[8]

Early legal access to the pill, ELA_{cs}, is equal to one if a woman born in year c would have had access to oral contraception before age 21 in her state of residence at age 21, and interactions of ELA with the age-group dummy variables allow its effect to vary across the lifecycle. Therefore, the key parameters of interest, the β_g terms, measure differences in the outcome of interest in age group g between women with and without early legal access to the Pill. It is worth noting that β_g will understate the impact of early Pill access for three reasons: local compliance and enforcement were imperfect; many young women could not have afforded the Pill even when it was legal; and young women may have driven across state lines to obtain it.

The main modification to Bailey (2006) is that we rely upon a revised legal coding (see Appendix 3.B). This updated legal coding reduces measurement error in ELA and allows the estimation of more precise effects over the lifecycle. Because these laws are not used elsewhere in the literature, the following section establishes their relationship with Pill use and subjects them to validity checks using detailed information on pre-treatment characteristics.

3.4.2 Validity of Using *ELA* to Identify the Impact of the Pill

One important assumption required to obtain consistent estimates of β_g is that ELA is uncorrelated with the error term after conditioning on state, age-group and birth-cohort fixed effects, or $cov(ELA, \eta \mid \mathbf{Z}) = 0$, where $\mathbf{Z}$ captures the fixed effects in equation (3.1).

One reason that $cov(ELA, \eta \mid \mathbf{Z})$ may not be zero is that ELA may not be conditionally, randomly assigned at baseline. That is, a systematic correlation between omitted characteristics and ELA could drive the relationship between ELA and outcomes. Because the *NLS-YW* contain rich information on respondents' backgrounds at age 14 *before treatment with ELA*, we test this possibility using the

[8] For dichotomous dependent variables, we estimate probits and report average partial effects (APEs). The standard errors are calculated using a non-parametric bootstrap method with states as clusters (1,000 repetitions).

following specification,

$$X_{ics} = \gamma ELA_{cs} + \sum_{s} \lambda_s D_s + \sum_{c} \lambda_c D_c + \varepsilon_{ics}, \tag{3.2}$$

where X is a pre-treatment characteristic and other notation remains as previously described. Thus, γ measures the residual correlation between ELA and pre-treatment characteristics that could indicate correlations with other, unobserved characteristics. (This approach is akin to testing for balance in observable characteristics in a controlled experiment.) Failure to reject $\gamma = 0$ is consistent with conditional random assignment of early legal access to the Pill. Although the power of this test is limited by our small sample sizes, it provides a strong validity test of the empirical strategy.

Table 3.1 reports the results of this exercise for 18 pre-treatment characteristics including a binary variable for whether the respondent's father was born in the U.S.; a binary variable for whether the respondent's father/mother worked for pay or held a professional job when she was 14 (four separate outcomes); an occupational prestige index for the father, conditional on working; a socio-economic status index for the respondent's parents in 1968; a binary variable for whether the respondent resided on a farm or in a rural area at age 14; a binary variable for whether the respondent had access to magazines, newspapers or a library card at age 14 (three separate outcomes); a binary variable for whether the respondent lived in a household with two parents at age 14; the number of siblings a respondent had; the highest grade completed by father/mother by 1968 (two separate outcomes); the number of years of schooling parents wanted the respondent to obtain when she was age 14; the atypicality of the respondent's mother's job (conditional upon mother working; negative numbers represent more atypical outcomes); and the respondent's IQ score in high school (see Appendix 3.A for details). Each column represents a separate, least-squares regression estimate of γ.[9] Consistent with treating ELA as conditionally, randomly assigned, only one of the 18 estimates is statistically significant at the ten percent level—no more than expected by chance. It is also reassuring that the pattern of correlations suggests no consistent relationship

[9]Linear probability models are used for binary outcomes to circumvent potential problems with disclosure. The results are robust to using negative binomials and probits where appropriate.

Table 3.1: Relationship of *ELA* to Pre-Treatment Respondent Characteristics

	Father worked for pay	Father held professional job	Mother worked for pay	Mother held professional job	Duncan index of occupation of head	Family socio-economic status in 1968
ELA	-0.020	0.023	0.003	0.046	0.692	-0.288
	(0.012)	(0.029)	(0.029)	(0.029)	(1.617)	(1.664)
Observations	4352	3930	3754	1426	3930	4100
R-squared	0.01	0.04	0.03	0.05	0.07	0.14
Mean of D.V.	0.929	0.195	0.387	0.126	31.625	99.917
	Magazines available	Newspapers available	Respondent held library card	Lived in two-parent household	Number of siblings in 1968	Father born in U.S
ELA	-0.017	-0.019	-0.012	-0.016	-0.138	-0.017
	(0.029)	(0.022)	(0.033)	(0.025)	(0.194)	(0.012)
Observations	4341	4345	4346	4354	4323	4353
R-squared	0.07	0.09	0.13	0.03	0.07	0.05
Mean of D.V.	0.637	0.833	0.695	0.816	3.586	0.959
	Highest grade completed by father in 1968	Highest grade completed by mother in 1968	Parents' desired education for respondent	Index of atypicality of mother's job	Respondent's IQ score in 1968 (age-adjusted)	Rural residence
ELA	0.065	0.101	-0.105	0.033	1.189	0.027
	(0.241)	(0.210)	(0.179)	(2.490)	(1.430)	(0.030)
Observations	3228	3893	3907	1786	2879	4348
R-squared	0.12	0.09	0.02	0.05	0.08	0.09
Mean of D.V.	10.044	10.313	13.337	29.909	102.091	0.256

Notes: See Appendix 3.A for more information on survey questions and variable coding. Characteristics are measured at age 14, unless otherwise indicated. Each of the separate regressions also includes a set of state of residence and birth cohort fixed effects. Heteroskedasticity-robust standard errors are corrected for clustering at the state level and are presented in parentheses below each estimate.

between ELA and the pre-treatment characteristics. For instance, ELA is negatively associated with father's employment and with family socio-economic status, but is positively associated with mother's education and professional employment.

Even if ELA is conditionally, randomly assigned, another reason that $cov(ELA, \eta \mid \mathbf{Z})$ may not be zero is that ELA is packaged with other policy changes. Although the history of these legal changes makes this unlikely, one concern is that cohorts with ELA were differentially treated with abortion access by chance—a treatment that could have a similar effect. Although data limitations mean that abortion access cannot be measured directly, our analysis accounts for this possibility by augmenting our equation (3.1) with a rich set of abortion controls:

$$\begin{aligned} Y_{iacs} = &\sum_g \beta_g ELA_{cs} D_{g(a)} + \sum_g \gamma_g EAA_{cs} C50_c D_{g(a)} + \sum_g \theta_g ELA_{cs} EAA_{cs} C50_c D_{g(a)} \\ &+ \delta LnDist_s C50_c + \sum_g \lambda_g D_{g(a)} + \sum_s \lambda_s D_s + \sum_c \lambda_c D_c + \eta_{iacs} \qquad (3.1'), \end{aligned}$$

where EAA represents "early access to abortion" and is equal to 1 if an individual resided (at age 21) in Alaska, California, the District of Columbia, Hawaii, New York or Washington, states that legalized abortion in 1970. *C50* is equal to 1 for birth cohorts born in 1950 or later, because the early legalization of abortion in 1970 could not have affected Pill use or fertility timing among 18 to 20 year-olds *before* 1970 (cohorts born before 1950). It is also important to note that any cohort-invariant, state-level differences in access to abortion will be captured in the state effects. The interaction of *EAA* and *C50* with age-group dummies allows the differential evolution of outcomes for state-birth-cohort groups exposed to legal abortion in their state of residence before their 21st birthdays. Separate interactions of *EAA* and *C50* with *ELA* and age-group dummies allow early abortion access and early access to the Pill to be complements or substitutes. Finally, cross-state travel to obtain abortion is accounted for by inclusion of log distance to the nearest large city providing legal abortions to out-of-state residents (Buffalo, New York City, Los Angeles, San Francisco, or the District of Columbia), $LnDist_s$, for cohorts born in 1950 or later (cf. Joyce, Tan, and Zhang 2010). Therefore, the key parameters of interest, β_g, measure differences in outcomes in age group g between women with and without *ELA* for cohorts that did not have early access to abortion in their home state after adjusting for cohort-level changes in cross-state travel for abortion.[10]

[10] Disclosure limitations from the Research Data Center prevent us from reporting the estimates on EAA and the ELA-EAA interactions, although we can summarize these findings generally. We find that early abortion access does have independent effects on many (but not all) of the outcomes we examine, of a comparable magnitude to ELA. The coefficients on the interactions are consistent with the Pill and abortion acting as substitutes, which agrees with Ananat and Hungerman (2012), although the estimates are seldom statistically significant. The inclusion of these abortion controls has a negligible effect on the ELA point estimates, as can be seen by comparing estimates here to

Finally, we test the sensitivity of our results in four alternative specifications of (3.1'): one with linear, state-specific time trends; another with controls for Vietnam casualties[11]; another using only a balanced sample of individuals (those missing information in any year or attriting are omitted); and another using state where the respondent attended high school to match to *EL*A rather than state of residence at 21.[12]

3.4.3 The Relevance of Early Legal Access for Pill Use

Testing the relevance of *ELA* for women's use of the Pill is more difficult, because the *NLS-YW* contains no information on young women's contraceptive decisions. Goldin and Katz (2002) examined this question with a single cross-sectional data set (*1971 National Study of Young Women*, *NSYW71*) and found that *ELA* increased Pill use among 17 to 19 year-olds by 4 percentage points (40 percent), but it is unclear how this evidence bears upon this analysis for two reasons. One reason is that Goldin and Katz (2002) used a different legal coding, which means their estimates may not generalize to the coding used in this paper. A second and more important reason is that the single cross-section of data in the *NLSY71* cannot be used to estimate the implicit first stage of this analysis, because state and cohort fixed effects cannot be included. Key for our investigation is that *ELA* increased Pill use at ages 18 to 20 *after conditioning on year of birth and state fixed effects*.

The *1970 National Fertility Survey* (*NFS*), which asked ever-married women to recall Pill use over the decade of the 1960s, allows us to examine this question directly for the subset of women who were ages 18 to 21 before 1970 and women who were married by 1970. We re-estimate equation (3.2) where X is a binary dependent variable equal to 1 if a respondent first used the birth control pill before age a, where $a = 18, 19, \ldots, 22$. If *ELA* mattered for Pill use at ages 18 to 20, we would expect γ to be positive.

Before presenting the results, several limitations of the data should be noted. First, the sample is restricted to ever-married women. Because women treated with early access to the Pill tended to delay marriage (cf. Goldin and Katz 2002, Appendix Table 3.1), unmarried young women not in the 1970 *NFS* may have been among those with the strongest response to *ELA*. This would lead our estimates to *understate* the impact of *ELA* on Pill use. Second, the 1970 *NFS* provides information

those without abortion controls in Figures 3.3 and 3.4.

[11] Using data from the National Archives on the Vietnam Conflict, the specification in equation (3.1') is augmented with controls for state-level casualties. These controls include state-specific annual death rates lagged one, two, and three years; and cohort-specific, state-level death rates within two years of a woman's date of birth.

[12] Due to disclosure requirements on implicit sample sizes, we cannot include all of these controls and restrictions in one specification. More details on each specification can be found in Appendix 3.A.

on a smaller set of cohorts and identifying variation than does the *NLS-YW* analysis. In order to estimate γ using a balanced panel, the analysis restricts the sample to the birth cohorts of 1942 (age 18 in 1960) to 1948 (age 22 in 1970), which results in 1,985 observations. Implicitly, this limits the states transitioning to *ELA* to Georgia, Kentucky, Mississippi, Ohio, and Washington. Finally, stigma-induced underreporting of Pill use among young, unmarried women with *ELA* who started systematically earlier would also lead to an understatement of the impact of *ELA* on Pill use.

Despite these limitations, these data provide strong evidence that *ELA* increased Pill use at the appropriate ages. Panel A of Table 3.2 presents separate regressions of equation (3.2) for first Pill use before a given age. By chance, it appears that women in the five states that transitioned to *ELA* before 1968 were significantly *less* likely to use the Pill before age 18—a bias that works against our finding effects.[13] However, Pill use by age 18 (before age 19) was 17 percentage points higher—an increase of roughly 140 percent over the national mean use at that age. Pill use by age 20 was 16 percentage points higher, an increase of 43 percent over the national mean. These striking differences fall sharply to a statistically-insignificant 5 percentage points at age 21, when women without *ELA* could obtain the Pill legally.[14]

Panel B of Table 3.2 explores heterogeneity in this effect by the community size of the primary sampling unit. We implement this by augmenting equation (3.2) with a dummy variable for non-metropolitan area as well as the interaction of this variable with *ELA*. Not surprisingly the strongest responses to *ELA* occurred in metropolitan areas. Consistent with changes in *ELA* increasing access to the Pill at age 18, use of the Pill in metropolitan areas with *ELA* was 30.4 percentage points higher—2.5 times the national mean in metro areas. This difference was 13.7 percentage points in less populated areas. Use of the Pill before age 21 was 26.9 percentage points, or 77 percent, higher among women with *ELA* in metro areas and 12.7 percentage points, or 31 percent higher, in non-metro areas, and these estimates are virtually unchanged with the inclusion of state linear time trends (see Appendix Table 3.2). For metro and non-metro areas, the difference in Pill use for women with *ELA* fell to 10 percentage points and 3 percentage points, respectively, by age 22, when early access laws ceased to bind.

[13] The relatively small standard error for the estimate of Pill use before age 18 appears to be an artifact of heteroskedasticity. When calculated under the assumption of homoskedasticity, the standard error is 0.042, rendering the estimate statistically insignificant. We also experimented with variance estimation by clustering at the state-by-year-of-birth level (instead of at the state level) and by using standard Huber-White methods. These alternative approaches did not weaken our inferences for any of the other estimates.

[14] Although omitted here for brevity, we also find that these differences in use translated into meaningful differences in marriage timing (cf. Goldin and Katz 2002) and age at first birth (cf. Bailey 2006, 2009): women with *ELA* delayed marriage by an average of 0.42 years and motherhood by 0.25 years.

Table 3.2: The Impact of *ELA* on Pill Use among Ever Married Women

	(1)	(2)	(3)	(4)	(5)
	1=Used Pill before age 18	1= Used Pill before age 19	1= Used Pill before age 20	1= Used Pill before age 21	1= Used Pill before age 22
Mean of DV	*0.034*	*0.119*	*0.226*	*0.369*	*0.506*
Panel A: Pill Use					
ELA	-0.056	0.171	0.188	0.158	0.050
	(0.017)	(0.204)	(0.142)	(0.084)	(0.040)
R-squared	0.048	0.105	0.124	0.136	0.127
Panel B. Pill Use Heterogeneity					
ELA	-0.052	0.304	0.254	0.269	0.105
	(0.020)	(0.168)	(0.117)	(0.088)	(0.061)
ELA x Non-metro area	-0.004	-0.167	-0.084	-0.142	-0.073
	(0.014)	(0.060)	(0.057)	(0.067)	(0.055)
R-squared	0.049	0.108	0.125	0.137	0.128
Observations	1985	1985	1985	1985	1985
Fixed effects	S, Y	S, Y	S, Y	S, Y	S, Y

Notes: Panel A presents the estimates of equation (3.2), while Panel B presents estimates from equation (3.2) augmented with a dummy for non-metropolitan area and the interaction of this dummy with *ELA*. Both panels use are estimated with a linear probability model on the 1942 to 1948 birth cohorts from the 1970 National Fertility Survey, which sampled ever married women. These cohorts are chosen so that the youngest women (born in 1948) were at least 22 in 1970 and that the oldest women (born in 1940) would have varied in their legal access to the Pill by age 21. All regressions include state fixed effects (S) and cohort fixed effects (Y). Heteroskedasticity-robust standard errors are corrected for clustering at the state level and are presented in parentheses below each estimate.

Stronger results in metropolitan areas are consistent with the difficulty of getting contraceptives anonymously in small towns or rural areas (even when legal).[15]

Although these results provide the best evidence in the literature of the relevance of *ELA*, we caution against using them as a denominator to approximate average treatment effects for Pill *use* on the treated (ATT) for several reasons. First, the sample of married women and stigma about reporting premarital Pill use may lead this analysis to understate the true effect of Pill access on Pill use, which would inflate estimates of the treatment effect on the treated. Second, the external validity of these results is difficult to establish. Not only was the 1970 *NFS* not designed to be representative at the state level, but the estimates for the handful of states that transition to *ELA* (cohorts of 1942 to 1948) during our sample period may not represent the effects for the full set of cohorts (1943 to 1953) considered in the analysis. Finally, even if the effect of *ELA* on Pill use lies in our estimated range of 16 to 19 percentage points, dividing other *ELA* effects by this amount yields the ATT only if *ELA* has zero effect on women who did not use the Pill. That would not be the case if the *option* to use the Pill affects human capital investment or if there are general equilibrium effects or demand-side responses to Pill diffusion. For instance, as more women enter the workplace with *ELA*, women in these markets who did not use the Pill may benefit from reductions in employers' statistical discrimination. Our intention-to-treat estimates in the following section include these general equilibrium effects, but our estimates of Pill use in the *NFS* do not.

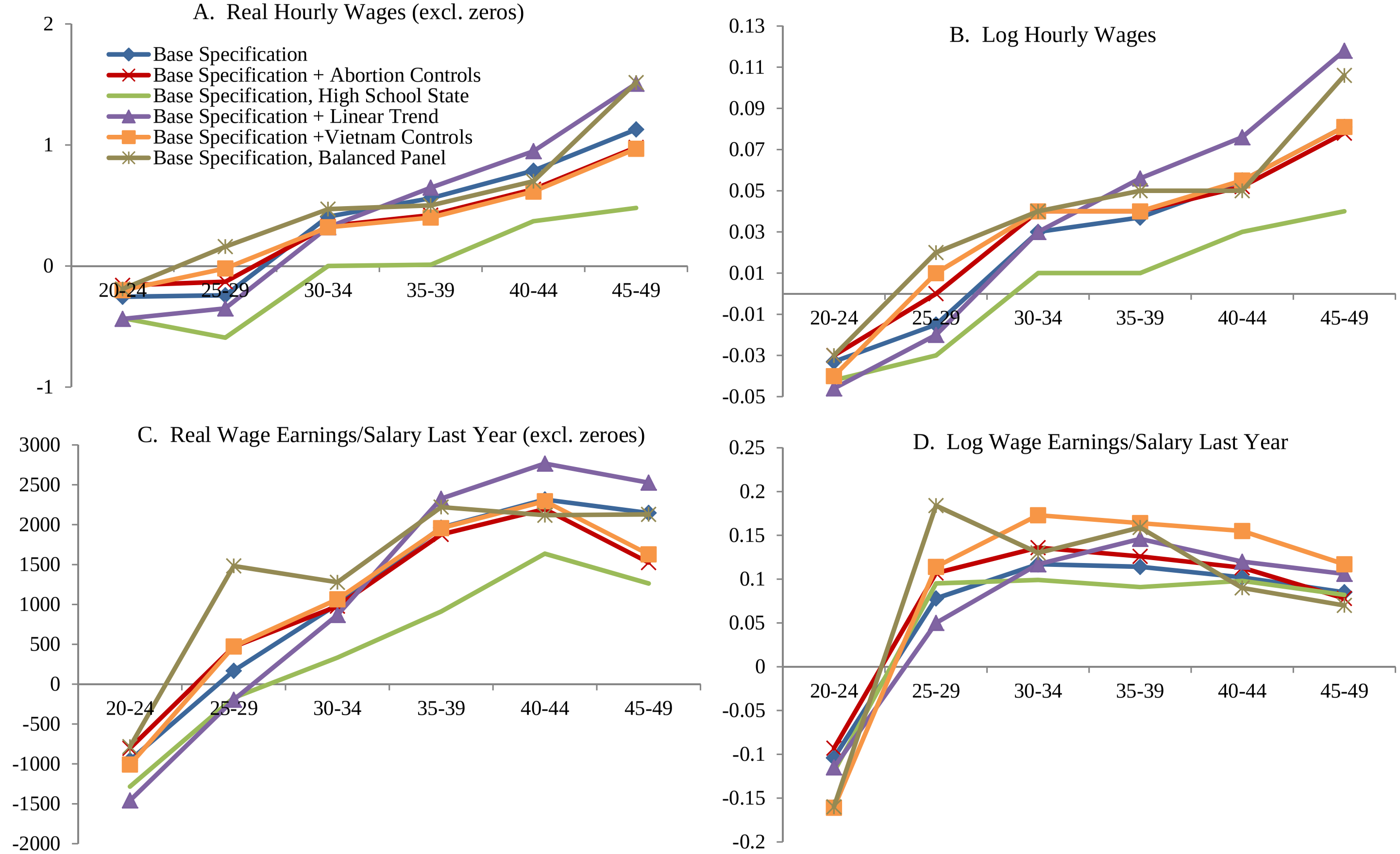
3.5 Results: How the Pill Affected Women's Lifecycle Wages

3.5.1 The Effect of the Pill on Women's Wages

Figure 3.3 plots the effect of *ELA* on women's life cycle wage earnings for four dependent variables in each of four panels. The figure includes our baseline specification (using equation (3.1)), a specification with abortion controls (using equation (3.1')), and the four alternative specifications described above. Throughout the results section, our discussion focuses on the magnitudes of our estimates with abortion controls (3.1'), but it is important to note that the estimates from each of the other five specifications are generally not statistically different from those in (3.1').

[15] Knowing the town doctor—or knowing that your parents did—or potentially being observed by your neighbor entering the local Planned Parenthood may have deterred many young women from seeking a prescription for the Pill—even if it was legal. Moreover, small town physicians may have been less willing to prescribe the Pill to unmarried women even when legal.

Figure 3.3: The Effects of Early Access to the Pill on Lifecycle Wage Earnings



Notes: Wage earnings are in 2000 dollars using the personal consumption expenditures deflator (BEA 2009). Each panel plots β_g from six different regressions: baseline specification (equation 3.1), baseline + abortion controls (equation 3.1') which corresponds to our tables, and four variants of equation 3.1': one with linear, state-specific time trends; another including controls for Vietnam casualties; another using only a balanced sample of respondents (those missing information in any year or attriting are omitted); and another using state where the respondent attended high school to match to *ELA* (see footnotes 20 and 24 regarding selection problems with this sample). Source: NLS-YW.

Across the six specifications, samples (including and excluding nonworking women), and definitions of the dependent variable, Figure 3.3 shows a consistent pattern. Women with *ELA* earned less in terms of hourly and annual wages in their early twenties, but their wage and salary earnings grew more rapidly than their counterparts as they aged.[16] At ages 20 to 24, working women with *ELA* earned 3 percent less in hourly terms (Table 3.3 columns 1 and 2) and 9 percent less on an annual basis (Table 3.3 columns 3 and 4). By their early forties, women with *ELA* earned a statistically significant premium of 5 percent hourly and 11 percent annually. This implies they earned 63 cents more per hour and roughly 2,200 dollars more per year. Notice that the annual amount is substantially larger than the 1,300 dollars implied by the hourly increase for a full-time, full-year worker, which is consistent with *ELA* also affecting labor-force participation.[17] Column 5 confirms this. Including women who did not work increases the *ELA* annual wage premium to 2,700 dollars per year.

Although previous work links the diffusion of the Pill among younger, unmarried women to increased educational attainment (Hock 2008), women's lifecycle labor-force participation (Bailey 2006), and marital outcomes and occupational upgrading among college graduates (Goldin and Katz 2002), none of these studies explores the implications of these changes for women's wages, which is this paper's objective. The following sections extend the literature by reexamining these mechanisms and explicitly linking them to wages. For thoroughness, we replicate previous findings in the literature for a sample of all women and compare our findings, which are based on different cohorts and measures of *ELA*, to previous estimates. In addition, we add to the literature on the Pill's labor-market effects by examining novel outcomes such as on-the-job training and cumulative labor-market experience (Section 3.5.2) and by considering how the Pill changed selection into human capital investments and paid work across ages (Section 3.5.3).

3.5.2 Mechanisms for the Pill's Effect on Wages

Our theoretical framework provides three potentially reinforcing explanations for *ELA*'s effects on wage profiles. The experience mechanism suggests that the initial *increase in women's labor-force participation* could have depressed wages at younger ages but increased wages later as these

[16] Although the estimates are not statistically different, it is noteworthy that using high school state rather than state at age 21 reduces the effect of *ELA* on wages. This is the case because we are less likely to have information on high school state for women who left the state for college. (Note that our estimates of college enrollment in Table 3.4 are also much smaller for this sample.) Because women attending out-of-state colleges may have been the most able or ambitious, it makes sense that our wages estimates are slightly smaller when we omit them.

[17] The annualized value of the hourly premium may also differ from the annual wages because the compensation information represents different pay periods. Hourly wages are from the most recent job, whereas annual wage and salary earnings reflect earnings in the previous calendar year from 1968 to 1993 and in the previous 12 months after 1994.

Table 3.3: The Impact of Early Access to the Pill on Wages and Annual Incomes

		(1)	(2)		(3)	(4)		(5)
	Mean real hourly wages excl. zeros	Real hourly wage (excl. zeros)	Log real hourly wage	*Mean real wages/salary last year excl. zeros*	Wage or salary last year (excl. zeros)	Log real annual wage	*Mean real wages/salary last year incl. zeros*	Wage or salary last year (incl. zeros)
ELA * Ages 20-24	*7.88*	-0.160	-0.030	*9943*	-804	-0.093*	*7661*	-1,187*
		(0.315)	(0.025)		(681)	(0.053)		(625)
ELA * Ages 25-29	*9.60*	-0.130	0.000	*15610*	469	0.107**	*10911*	202
		(0.347)	(0.028)		(741)	(0.046)		(721)
ELA * Ages 30-34	*10.62*	0.330	0.040	*18116*	978	0.136**	*12452*	803
		(0.332)	(0.028)		(731)	(0.059)		(683)
ELA * Ages 35-39	*11.74*	0.420	0.040	*21173*	1,878**	0.126**	*15442*	1,449*
		(0.333)	(0.027)		(749)	(0.050)		(744)
ELA * Ages 40-44	*12.84*	0.634*	0.052**	*24493*	2,196**	0.113**	*19184*	2,674***
		(0.334)	(0.024)		(919)	(0.045)		(892)
ELA * Ages 45-49	*14.29*	0.980**	0.078**	*28148*	1,526*	0.078*	*25238*	3,376***
		(0.448)	(0.031)		(781)	(0.047)		(919)
Fixed effects		Y, S, A	Y, S, A		Y, S, A	Y, S, A		Y, S, A
Observations		46388	46388		51277	51277		68169
Unique women		4210	4210		4245	4245		4351
R-squared		0.22	0.27		0.01	0.10		0.01

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Wages are adjusted to 2000 dollars using the PCE deflator (BEA 2009). All regressions include state fixed effects (S); cohort fixed effects (Y); age group fixed effects (A); controls for abortion access; and abortion access controls interacted with ELA as described in equation (3.1'). Heteroskedasticity-robust standard errors are clustered at the state level and presented in parentheses below each estimate.

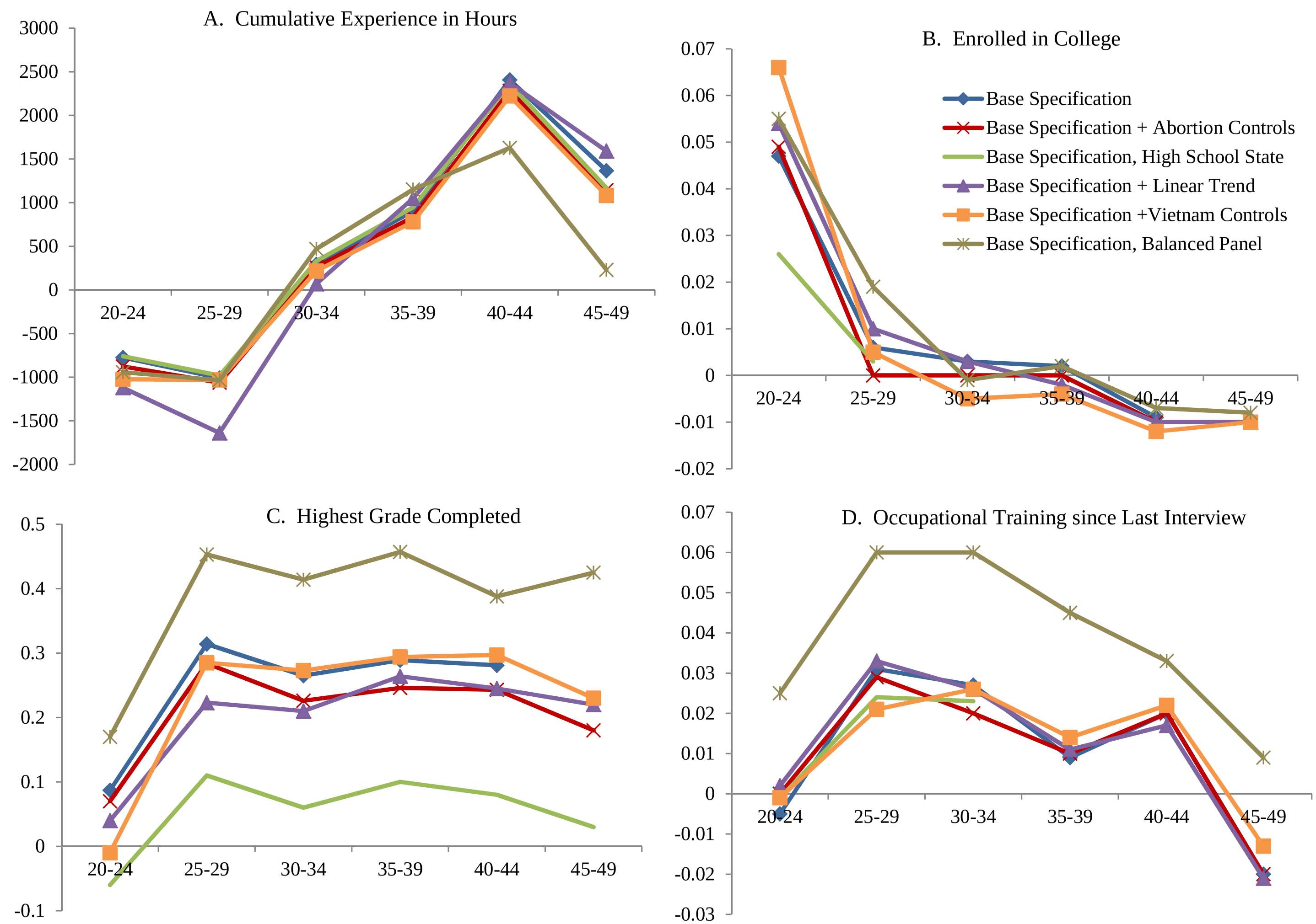
women accumulated labor-market experience and/or job/firm-specific capital. The on-the-job training mechanism requires *no initial or longer-run differences in labor-force participation*, but suggests that workers with *ELA* increased their on-the-job human capital investments, which would also result in steeper wage earnings profiles. The formal human capital investment mechanism is consistent with women *reducing their initial labor-force participation* as they invested in their education or training and then reaping the returns to these early investments when they returned to the labor market, which would also result in steeper wage earnings profiles. Each of these explanations likely operated to some degree in practice, so our exploration of the Pill's labor-force participation effects here aims to shed light on the predominant mechanism for its observed wage effects. Importantly, each of these explanations postulates *different* labor-force participation and human capital investment patterns.

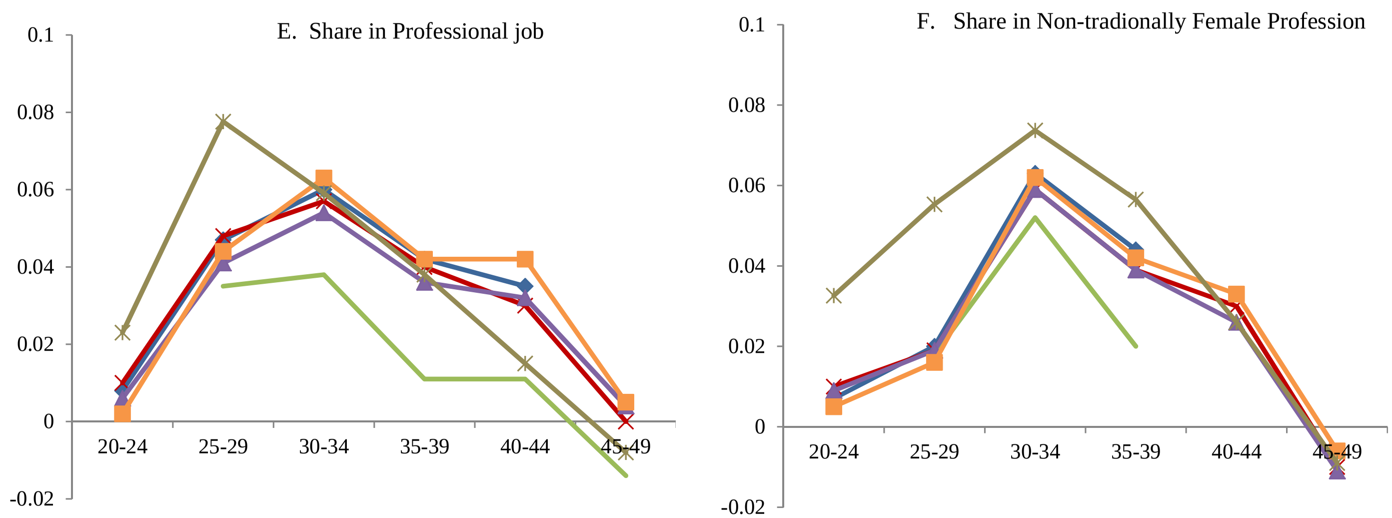
As a starting point, we examine the effect of *ELA* on women's labor market participation at the extensive (1 = in the labor force) and intensive margins (using "usual weekly hours" for working women) and find that women with *ELA* participated *less* in their early twenties and more in their late twenties and thirties.[18] These differences in labor-force participation resulted in different cumulative experience profiles as shown in Figure 3.4A and column 1 of Table 3.4, which define women's cumulative work experience as weeks worked multiplied by usual weekly hours summed across survey waves (see Appendix 3.A for more details). The results show that women with *ELA* had worked 18 percent fewer hours by their late twenties but erased this deficit during their thirties. By their early forties, women with *ELA* had amassed the equivalent of 1.15 years more of full-time, full-year work (2,300 more hours)—an increase of over 10 percent relative to their same-aged peers without *ELA*, and about 30 percent larger than the increase found by O'Neill and Polachek (1993) between cohorts born in the mid-1930s and those born a decade later.[19]

[18] These findings are consistent with Bailey's (2006) results using repeated cross-sections from the March *CPS*, but the magnitudes in the *NLS-YW* are larger than in the *CPS* but less precisely estimated owing to significantly smaller sample sizes. These differences in magnitude are expected because Bailey's (2006) use of current state of residence (rather than residence at age 21) should attenuate her results. For brevity, we omit estimates for labor-force participation from this paper and compare our *NLS-YW* estimates to Bailey (2006) in this footnote. At ages 25 to 34, women with *ELA* were roughly 3.8 percentage points, or 6 percent, more likely to work for pay in the *NLS-YW*; Bailey reports an almost identical estimate (3.9 percentage points for women ages 26 to 30) but her estimate is smaller at 1.6 percentage points for women ages 31 to 35. The *NLS-YW* also shows a larger effect in the late thirties than the *CPS*, although the *NLS-YW* estimate is statistically insignificant. The effect of *ELA* on hours worked (excluding zeros) in the *NLS-YW* is not as comparable, because it asks usual hours worked whereas the CPS asks the number of hours worked in the CPS reference week. The effects at older ages are larger for usual hours worked in the *NLS-YW*, where women 30 to 34 years-old worked one additional hour per week on average, 2.5 percent more than their counterparts without *ELA*; 35 to 44 year-olds worked 1.3 to 1.7 additional hours, or 3.5 to 4.8 percent more. Full results are available upon request.

[19] The comparison with O'Neill and Polachek is approximate, both because they analyze slightly different groups of women and because their measure of labor market experience is different. In particular, they count years in which at least 26 weeks were worked as a full year of experience; changes at the extensive margin or changes on the intensive margin that do not cross the 26-week threshold are thus missed by their measure.

Figure 3.4: The Effects of Early Access to the Pill on Lifecycle Human Capital Investments





Notes: See notes to Figure 3.3.

This pattern of reduced labor-force participation is the reverse of the labor-supply shift needed to decrease wages at younger ages. Similarly, the on-the-job training channel is also inconsistent with early career dips in labor supply: if fewer women are working for pay, more cannot be accumulating on-the-job training at these ages. The Pill-induced accumulation of experience is most consistent with the formal human-capital investment channel, which postulates that *ELA* women used the Pill to make more investments in formal schooling and training early in their careers and enjoyed the returns on these investments in terms of steeper wage profiles, which also encouraged greater labor-force attachment, as they aged.

Panels B through F of Figure 3.4 examine *ELA*'s effect on these more formal human capital investments including women's college enrollment, years of education, occupational training, and professional occupations for the six specifications; Table 3.4 presents estimates in tabular form. The results provide a rich picture of Pill-induced changes in women's career investments. College enrollment was 4.9 percentage points, or 20 percent, higher for women with *ELA* in their early twenties but not at later ages (Table 3.4 column 2; Figure 3.4B).[20] Their advantage in grades completed (Table 3.4 column 3; Figure 3.4C) peaks in their late twenties, at a little more than one quarter of a year and erodes a bit as women without *ELA* returned to school in their thirties. A difference of one quarter of a year of schooling, however, persists through the early forties. In addition to completing more formal education in their early twenties, women with *ELA* were 15 percent more likely to report occupational training (Table 3.4 column 4, Figure 3.4D) in their late twenties. Although reports of occupational training remain modestly elevated for *ELA* women at older ages, the estimates are not statistically different from zero.

Women's greater human capital investments also appear in their occupational choices, which capture both observed (more formal education) as well as unobserved career investments (such as more career commitment or effort) (see Appendix 3.A for more information on occupational coding). With *ELA*, women were 17 to 30 percent (4 to 6 percentage points) more likely to be working in a professional or managerial job during their late twenties and thirties, respectively (Table 3.4 column 5, Figure 3.4E). Half of this increase in the late twenties, and all of it during the thirties, was due to entry into non-traditionally female professional occupations—professions other than nursing or teaching (Table 3.4 column 6, Figure 3.4F).

[20] Estimates are 30 percent larger than our baseline estimate (0.066 for a 27 percent increase) when we include controls for Vietnam mobilization. Estimates are 50 percent smaller (0.026 for an 11 percent increase) when we use high school state. Using high school state reduces our estimates because we are less likely to have information on high school state for women who went out of state to college. Thus, our sample of women for whom we have high school state disproportionately drops out-of-state college enrollees. These estimates are larger than reported in Hock's (2008) working paper. Using the October *CPS*, he finds—using a different measure of *ELA*—that college enrollment was roughly 2.5 percentage points higher among 21 and 22 year olds with ELA.

Table 3.4: The Impact of Early Access to the Pill on Human Capital Accumulation and Occupational Upgrading

	(1)	(2)	(3)	(4)	(5)	(6)
	Cumulative Experience in Hours	1= Enrolled in College	Highest Grade Completed	1=Occupational training since last interview	1= in Professional Job	1=in Non-traditional Job
ELA * Age 20-24	-876**	0.049**	0.070	0.000	0.010	0.010
	(369)	(0.022)	(0.136)	(0.013)	(0.013)	(0.008)
ELA * Age 25-29	-1,062**	0.000	0.284**	0.029***	0.048***	0.019*
	(443)	(0.008)	(0.131)	(0.011)	(0.019)	(0.011)
ELA * Age 30-34	263	0.000	0.226*	0.020	0.057***	0.059***
	(405)	(0.013)	(0.132)	(0.016)	(0.021)	(0.016)
ELA * Age 35-39	836	0.000	0.246*	0.010	0.040	0.039**
	(550)	(0.010)	(0.133)	(0.018)	(0.023)	(0.019)
ELA * Age 40-44	2,282***	-0.010	0.243*	0.020	0.030	0.030
	(784)	(0.010)	(0.129)	(0.022)	(0.027)	(0.020)
ELA * Age 45-49	1,143	-0.010	0.180	-0.020	0.000	-0.010
	(988)	(0.007)	(0.145)	(0.020)	(0.021)	(0.018)
Fixed Effects	Y,S,A	Y,S,A	Y,S,A	Y,S,A	Y,S,A	Y,S,A
Observations	61736	57373	78809	63013	73737	73737
Unique women	4329	3702	4354	4323	4354	4354
(Pseudo) R-squared	0.62	0.15	0.15	0.03	0.07	0.09
Mean of DV for 20-24	2723	0.241	12.09	0.203	0.086	0.044
Mean of DV for 25-29	5929	0.077	12.52	0.188	0.163	0.080
Mean of DV for 30-34	10758	0.072	12.85	0.245	0.199	0.137
Mean of DV for 35-39	16098	0.065	12.99	0.285	0.242	0.202
Mean of DV for 40-44	22609	0.049	13.13	0.310	0.249	0.225
Mean of DV for 45-49	30010	0.029	13.28	0.324	0.242	0.218

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: Columns (2) and (4)-(6) report average marginal effects from probit specifications; columns (1) and (3) report coefficients from OLS regressions. All regressions include state fixed effects (S); cohort fixed effects (Y); age group fixed effects (A); controls for abortion access; and abortion access controls interacted with ELA as described in equation (3.1'). Heteroskedasticity-robust standard errors are clustered at the state level and presented in parentheses below each estimate.

It is also interesting that differences in professional work erode with age, as female professionals with *ELA* retire.[21]

Together, more investments in formal human capital and greater labor-market attachment contributed to women's steeper age-earnings profiles. But given *ELA*'s reduction in labor-supply during women's early twenties, the *decrease* in working women's wages at those ages remains an open question. It is also unclear to what extent changes in the composition of women investing in their human capital and working for pay drive the increase in women's wages at older ages. We address both questions in the next section.

3.5.3 Heterogeneous Effects of the Pill and the Role of Workforce Composition in Wage Growth

In addition to shifting women's investments in their human capital, early access to the Pill may have shifted *which women* pursued an education, went to graduate or professional school, and got promoted. If higher ability women disproportionately used the Pill to make career investments with the expectation of working longer, and thus were initially more likely to be out of the labor force, then women working during their early twenties may have been negatively selected. As higher-ability women entered the work force in their later twenties after having made their career investments, their greater skills (unobserved and observed) would lead their earnings profiles to be steeper than those of less skilled women. Moreover, less skilled women may have seen their earnings fall as their more skilled counterparts began working. In short, access to the Pill may have altered selection into the labor market at younger ages, which could help explain the effect of the Pill on age-earnings profiles shown in Figure 3.3.

To examine the importance of selection, we use a composite of respondents' performances on aptitude tests from their high school transcripts, which was reported to the *NLS-YW* in 1968 and called an "IQ score" in the documentation. IQ is available for only two-thirds of the sample, so we divide respondents into IQ tertiles (low, middle, and high) to maintain samples sizes large enough for disclosure.[22] Equation (3.1') is then estimated for each of the IQ tertiles separately. We also examine heterogeneous effects of *ELA* by educational attainment (any versus no college) and, for education outcomes, family background (socio-economic status tertiles of families when the respondent was

[21] Our estimates are larger than those found in Goldin and Katz (2002, Table 5), who use a sample of U.S. born college graduate women ages 30 to 49 and find that the Pill increased the share in professional occupations, excluding teachers and nurses, by 0.4 percentage point (3 percent). One reason for the difference may be that their estimate includes women in their forties, where we find smaller effects.

[22] Griliches, Hall and Hausman (1978) point out that these IQ composite scores are missing "almost at random" in the National Longitudinal Survey of Young Men, which is also the case in the *NLS-YW*. See Appendix 3.A for details on the composite score.

14). Whereas IQ tertile measured in high school is not affected by *ELA* directly (cf. Table 1), educational attainment is (Table 3.4). The latter breakdown should be viewed as a description to help us explore how different groups of women differentially benefitted from early access to the Pill.

Table 3.5: Heterogeneity in the Impact of Early Access to the Pill on Real Hourly Wages

	(1)	(2)	(3)	(4)	(5)
Sample	Lower third of IQ distribution	Middle third of IQ distribution	Upper third of IQ distribution	No College	Some College
ELA * Age 20-24	-0.670	0.580	-0.390	-0.260	-0.730
	(0.634)	(0.623)	(0.444)	(0.294)	(0.529)
ELA * Age 25-29	-0.190	0.980	0.460	-0.110	0.050
	(0.580)	(0.724)	(0.477)	(0.293)	(0.518)
ELA * Age 30-34	-0.956	1.873**	0.720	0.060	0.760
	(0.519)*	(0.759)	(0.669)	(0.306)	(0.583)
ELA * Age 35-39	-0.120	1.888**	0.540	-0.190	1.346**
	(0.654)	(0.794)	(0.577)	(0.410)	(0.662)
ELA * Age 40-44	-0.420	2.216**	0.790	0.550	1.347**
	(0.958)	(0.944)	(0.632)	(0.479)	(0.611)
ELA * Age 45-49	0.720	2.302**	3.046***	0.797*	2.677***
	(1.043)	(0.939)	(1.010)	(0.470)	(0.907)
Observations	10468	14165	16788	40229	21785
Unique women	793	975	1112	2895	1456
R-squared	0.18	0.21	0.23	0.17	0.26
Mean of DV for 20-24	5.59	6.49	7.18	5.49	7.21
Mean of DV for 25-29	5.89	6.79	8.69	5.52	9.51
Mean of DV for 30-34	6.59	7.19	8.94	6.18	9.74
Mean of DV for 35-39	7.44	8.40	10.79	7.16	11.42
Mean of DV for 40-44	8.34	9.89	12.79	8.34	13.63
Mean of DV for 45-49	10.02	12.59	16.04	10.33	16.76

Notes: This table uses a specification similar to column (1) of Table 3.3. Each column presents estimates from a separate regression. Unlike Table 3.3, this table includes zero wages in the left-hand-side variable. We cannot report results excluding the zeros among the separate groups for disclosure reasons, but they follow a pattern similar to that shown above. Columns (1) to (3) break women into thirds of the IQ distribution, and columns (4) and (5) divide women into no college and some college. All other notes are as in Table 3.3.

Table 3.5 begins this analysis by examining the effect of *ELA* on women's hourly wages by IQ tertile and college attainment.[23] Whereas *ELA* reduces or has no significant effect on earnings for the lowest IQ tertile (column 1), it increases them in the middle and upper third of the IQ distribution (columns 2 and 3) for women aged 30 to 49. Almost all of the wage gains accrued to women in the

[23] We note that the results in Table 3.5 are from samples that included observations with zero earnings, unlike Table 3.3, which included only observations with positive earnings. This change was unfortunately necessary for disclosure reasons but does not affect the patterns we observe.

middle of the IQ distribution, where the effects are largest both absolutely and relatively. For this group, women with *ELA* enjoyed greater hourly wages throughout their twenties and the premium grew to a statistically-significant 20 percent at ages 30 to 49.

It is worth noting that the estimates in this table are from a more flexible version of the regression model that allows the state, cohort and age group fixed effects to vary by IQ group. The fact that *ELA* had an effect within the middle IQ group suggests that the labor market gains described previously are not the sole result of shifts in the composition of the workforce. Furthermore, if the wage effects of *ELA* were driven by changing selection into the labor market by women with different ability levels, we would expect the overall wage effects from models without IQ controls to be substantially larger than those from Table 5's models that stratify by IQ tertile. Instead, Table 3.3 and Table 3.5 imply similar average estimates (compare the *ELA* estimates averaged across the three IQ tertiles in Table 3.5 to the overall population estimates in Table 3.3).[24]

The fact that the wage effects are strongest for women attending some college suggests that one mechanism for these middle-IQ women was college enrollment. Although *ELA* conferred little if any wage premium for women without college (column 4), women with some college (column 5) experienced lower wages in their early twenties (perhaps as they worked at temporary jobs) but a 12-percent wage premium in their late thirties.[25] The effects for the highest IQ group are considerably smaller and not statistically significant at any age below 44, which suggests these women may have already been taking advantage of their educational and career opportunities without *ELA*. In contrast to these positive effects, the lowest IQ women with *ELA* suffered a statistically significant wage reduction of roughly 15 percent in their early thirties. Although this negative effect is consistent with the Pill increasing crowding in jobs where lower IQ women were working or decreasing the relative skills of lower IQ women, the estimate is not robust to the inclusion of state linear time trends (Appendix Table 3.3). The lack of wage benefits for lower IQ women may be related to the limited returns to human capital investments in low-skilled jobs or the absence altogether of these women's investments in their human capital, which we examine next.

[24] There are two other reasons why the averages of the estimates in Table 3.5 might differ from those in Table 3.3: the smaller sample in Table 3.5 (excluding women with missing IQ information) and the different outcome variable (including women with zero earnings). We further confirmed that the averages of the ability-group specific *ELA* estimates are also similar to the overall estimates when the samples both include women with zero earnings: the former tend to be smaller at younger ages but larger for women in their forties.

[25] It is worth noting that the estimated effects of *ELA* by college attainment in Tables 3.5 (for wages) and 3.7 (for experience) may be downward biased because of compositional effects. If the marginal women who attended college because of *ELA* were on average higher ability than the women with *ELA* and no college, but of lower average ability than the women who attended college even without *ELA*, the estimated effects of *ELA* on average wages and experience for each group will appear lower than the actual impact on individuals in either group.

Table 3.6: Heterogeneity in the Impact of Early Access to the Pill on Highest Grade Completed

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Lower third of IQ distribution	Middle third of IQ distribution	Upper third of IQ distribution	Lower third SES distribution	Middle third SES distribution	Upper third SES distribution
ELA * Age 20-24	-0.507**	0.240	0.170	0.220	-0.140	0.200
	(0.205)	(0.198)	(0.185)	(0.141)	(0.218)	(0.316)
ELA * Age 25-29	-0.409*	0.360	0.420**	0.480***	0.020	0.340
	(0.207)	(0.228)	(0.191)	(0.147)	(0.242)	(0.274)
ELA * Age 30-34	-0.431**	0.386*	0.426**	0.410***	0.000	0.280
	(0.206)	(0.224)	(0.197)	(0.161)	(0.246)	(0.288)
ELA * Age 35-39	-0.401**	0.437*	0.505**	0.434***	0.080	0.270
	(0.197)	(0.220)	(0.202)	(0.161)	(0.253)	(0.309)
ELA * Age 40-44	-0.494**	0.455*	0.449**	0.427**	0.080	0.270
	(0.215)	(0.243)	(0.191)	(0.175)	(0.254)	(0.274)
ELA * Age 45-49	-0.380	0.330	0.584***	0.425**	0.030	0.200
	(0.239)	(0.243)	(0.207)	(0.190)	(0.267)	(0.296)
Observations	13538	17550	20982	25101	24538	24798
Unique women	793	975	1112	1392	1366	1342
R-squared	0.19	0.19	0.23	0.12	0.19	0.26
Mean of DV for 20-24	11.87	12.40	13.30	10.98	12.26	13.22
Mean of DV for 25-29	12.05	12.74	14.08	11.21	12.66	14.01
Mean of DV for 30-34	12.28	13.02	14.39	11.53	12.94	14.35
Mean of DV for 35-39	12.35	13.16	14.58	11.63	13.07	14.52
Mean of DV for 40-44	12.45	13.27	14.72	11.72	13.26	14.64
Mean of DV for 45-49	12.55	13.45	14.87	11.86	13.39	14.77

Notes: This table uses the specification in column (3) of Table 3.4. Each column presents estimates from a separate regression. Columns (1) to (3) break women into thirds of the IQ distribution, and columns (4) to (6) divide the sample into thirds of the distribution of family background characteristics. SES is available for more women than IQ score, so the sample sizes in columns (4)-(6) are larger. All other notes are as in Table 3.4.

The next set of tables explores how the Pill affected human capital investments and paid work by IQ and childhood SES. The estimates in Table 3.6, which uses highest grade completed as a dependent variable, are roughly consistent with the pattern of *ELA*'s effects on wages. *ELA*'s effects on education are large and positive in the middle of the IQ distribution and negative for the lowest IQ group. (These negative effects may reflect higher IQ women crowding out lower IQ women in colleges.) Unlike the wage estimates, however, *ELA*'s effects on education are also large and statistically significant for the highest IQ tertile. By age forty, *ELA*'s effects for the middle and upper IQ groups translate into a 0.4 to 0.5 year schooling advantage. The right side of the table shows that *ELA*'s effects are largest for women from the *lowest* SES households (columns 4 through 6). Women with *ELA* from the most disadvantaged backgrounds attained roughly half of a year more education than their peers (column 4). This is a large effect, amounting to roughly one third of the difference in grades completed between women in the low and middle SES groups.[26] Although our data do not reveal whether these effects arise at the stage of high school completion, college admission, or class standing and persistence, it is clear that higher IQ women with access to the Pill—especially those from disadvantaged households—were more likely to continue their educations. Thus, *ELA* shifted women's educational attainment into more of a meritocracy.

Is the heterogeneity in the Pill's effects by IQ apparent for labor-force attachment as well? Table 3.7 uses cumulative labor-force experience to examine this question. As with education, the effect of *ELA* on labor-force experience is largest for women in the middle third of the IQ distribution and with some college. Middle IQ women (column 2) with *ELA* had accumulated 2,200 to 4,800 additional hours of work experience by their early thirties to late forties. Women in the highest IQ group (column 3) with *ELA* also participated more, but these effects on experience are smaller and less precise. Echoing the wage results, the effects of *ELA* on labor-force experience are largest for women with some college (column 5).[27]

In summary, the data provide strong support that the Pill influenced *which women* invested in their careers and shifted into paid work. Given the lack of labor-supply or schooling gains for low IQ women, the Pill appears to have induced positive selection into higher education as well as the labor market. This analysis also shows different responses to early access to the Pill across IQ tertiles. While lower IQ women with *ELA* did not gain ground in terms of education or experience, both middle and higher IQ women raised their educational attainment and those with some college became more likely to work for pay. Interestingly, the Pill's largest effects on work experience accrued to women in the middle of the IQ distribution, not to the high achievers who have been the focus of other studies. Thus, our findings highlight the different ways in which women across the IQ distribution used the flexibility conferred by early access to the Pill to opt into paid work.[28]

[26] The effect of *ELA* on college enrollment among 20 to 24 year-olds for the lowest IQ group was 0.9 percentage points (s.e. 3.6, mean 12 percent); it was 3.9 (s.e. 3.5, mean 19 percent) and 5.9 percentage points (s.e. 2.7, mean 37 percent) for the middle and upper IQ groups, respectively. The effect of *ELA* on college enrollment among 20 to 24 year-olds for the lowest SES group was 11.3 percentage points (s.e. 3.8 percentage points), an implied increase of 108 percent (of the mean of 10.5 percent). It was 3.9 percentage points (s.e. 4.1, mean 21 percent) and 2.1 percentage points (s.e. 3.0, mean 36 percent), respectively, for the middle and upper SES groups.

[27] We also directly estimated the effect of *ELA* by IQ tertile and college attendance on (binary) labor force participation. The heterogeneity in effects is similar: women in the middle IQ tertile in their late twenties and early thirties show the largest increases in participation. Higher IQ women also show increased participation at these ages, but the estimates are smaller and less precise. Women with some college show significant participation responses to *ELA* as well, with significantly lower rates in their early twenties, followed by significantly higher rates over the next decade.

[28] Another potential mechanism for the Pill's wage effects is its interaction with the marriage market and the size of spousal earnings. To investigate this "marriage-market channel," Appendix Table 3.4 examines the relationship of *ELA* with both the likelihood of never having married (panel A) and the likelihood of having divorced (panel B) by IQ group and college attendance. In almost all cases, we cannot reject that the likelihood of having married is unrelated to *ELA*. In contrast, divorce rates were significantly higher for women with *ELA* in the lower IQ groups and among women without any college. Women in the lowest third of the IQ distribution with *ELA* were almost twice as likely to divorce (9.7 percentage points) by their late twenties (panel B, column 1). Similarly, *ELA* women with no college

Table 3.7: Heterogeneity in the Impact of Early Access to the Pill on Cumulative Experience

	(1)	(2)	(3)	(4)	(5)
Sample	Lower third of IQ distribution	Middle third of IQ distribution	Upper third of IQ distribution	No College	Some College
ELA * Age 20-24	-1,083	409	-397	-871*	-1,056*
	(1,299)	(964)	(720)	(499)	(593)
ELA * Age 25-29	-1,253	278	-389	-928*	-920
	(1,295)	(1,043)	(695)	(552)	(615)
ELA * Age 30-34	-688	2,214*	654	45	862
	(1,145)	(1,150)	(802)	(450)	(722)
ELA * Age 35-39	-153	3,015**	1,377	346	2,045**
	(1,371)	(1,313)	(872)	(693)	(871)
ELA * Age 40-44	40	4,778***	1,853*	2,095**	3,001***
	(1,761)	(1,701)	(983)	(861)	(1,026)
ELA * Age 45-49	-600	3,701*	1,379	1,492	2,344*
	(2,251)	(2,242)	(1,228)	(1,075)	(1,331)
Observations	12469	16531	20181	47925	26150
Unique women	790	975	1112	2898	1456
R-squared	0.610	0.637	0.679	0.582	0.703
Mean of DV for 20-24	2533	3152	2793	2833	2432
Mean of DV for 25-29	5160	6103	6340	5382	6516
Mean of DV for 30-34	9558	10755	11432	9755	12104
Mean of DV for 35-39	14822	15936	17151	14662	18106
Mean of DV for 40-44	20975	21570	23838	20752	25111
Mean of DV for 45-49	27775	29652	31933	27964	33133

Notes: This table uses the specification in column (1) of Table 3.4. Each column presents estimates from a separate regression. Columns (1) to (3) break women into thirds of the IQ distribution, and columns (4) and (5) divide women into no college and some college. All other notes are as in Table 3.4.

3.6 Decomposing Pill-Induced Wage Gains

To quantify the contribution of each of these different human capital investments to the estimated Pill premium in wages, we decompose women's *ELA*-induced log hourly wage premium in their late forties into five components: formal education, on-the-job training, cumulative experience, occupational choice, and changes in marital status (that affect wages through the income of a spouse). We present results using the standard Blinder-Oaxaca decomposition at the mean (Blinder 1973; Oaxaca 1973) and the recentered influence function procedure (RIF) proposed in Firpo, Fortin, and Lemieux (2009), which generalizes Blinder-Oaxaca to other quantiles. This approach has the

were almost 34 percent (4.4 percentage points) more likely to divorce. However, these effects are for the wrong groups of women to be driving the wage effects. Although they are strong for women in the middle of the IQ distribution, they appear for those without any college—not the middle IQ women who pursued college. In short, little evidence points to divorce and the absence of a second earner as the explanation for the wage effects.

advantage of not being sensitive to the decomposition order and permits a richer characterization of the importance of Pill-induced changes in productive characteristics at different points in the skill distribution. To implement both procedures, we restrict the estimation sample to the last available wage observation for each woman in the 45 to 49 age group and use women without *ELA* as the reference group.

Table 3.8 quantifies how much of the difference in the log hourly wage premium of women with *ELA* at various points along the wage distribution can be explained (in an accounting sense) by each of the characteristics. Panel A reports the Blinder-Oaxaca decompositions at the mean and shows that cumulative experience accounts for just under two-thirds of the Pill premium. Education and occupation each account for another sixth of the gap, with both job training and marriage having negligible effects. Together, these five factors explain over 90 percent of the *ELA* wage premium at the mean.

What do our estimates imply about the returns to education and experience for women? Women with *ELA* obtained 0.18 years more schooling by their late forties (Table 3.4, column 3), which increased their wages by 0.015 log-points (Table 3.8, panel A), for an implied return of 0.083 (=0.015/0.18). If we also attribute the entire 0.014 log-point increase in wages (Table 3.8, panel A) from occupational upgrading to schooling, the total return to women's schooling would be 0.161 (=0.029/0.18). These estimates are both within a plausible range of Heckman, Lochner and Todd's (2006) 0.128 estimate of the returns to education for white men in 1990 (p. 326). For the same group, Heckman, Lochner and Todd estimate coefficients on experience and experience squared of 0.1301 and -0.0023, respectively (Ibid). Applying these returns to experience to our estimates indicates that, from an initial experience level of 15 years, that 0.57 years more experience (Table 3.4, column 1) would increase women's log-wages by 0.034 ($0.1301 * 0.57 - 0.0023 * (15.57^2 - 15^2)$). Our decomposition attributes more than that, 0.056 log points, to the 0.57 years of additional experience, which is also reasonable if the returns to women's experience are higher than the returns for men or level off less quickly (cf. Weinberger and Kuhn 2010).

The results of the RIF procedure, shown in panel B, are consistent with the Oaxaca-Blinder decompositions, with experience accounting for the largest share of the premium, followed by education and occupation.[29] The relative roles of experience and education-occupation, however, vary at different points in the wage distribution.

[29] The decomposition results are also similar if we use the semi-parametric approach of DiNardo, Fortin, and Lemieux (1996) to re-weight the characteristics of women without *ELA* to resemble those of women with *ELA* at different points in the distribution.

Table 3.8: Decomposition of the Impact of Early Access to the Pill on Log Hourly Wages

		Effect of					
Statistic	Total Difference	Education	Job Training	Experience	Occupation	Marriage	Unexplained Difference
Panel A: Oaxaca-Blinder Decomposition							
Mean	0.088	0.015	-0.003	0.056	0.014	0.000	0.006
		(17.0)	(-3.4)	(63.6)	(15.9)	(0.0)	(6.8)
Panel B: Recentered Influence Function Decomposition							
10^{th} percentile	0.077	0.003	-0.001	0.053	-0.003	-0.001	0.026
		(3.9)	(-1.3)	(68.8)	(-3.9)	(-1.3)	(35.1)
25th percentile	0.077	0.005	-0.004	0.066	0.007	0.001	0.003
		(6.5)	(-5.2)	(85.7)	(9.1)	(1.3)	(2.6)
50^{th} percentile	0.106	0.014	-0.005	0.072	0.013	-0.003	0.015
		(13.2)	(-4.7)	(67.9)	(12.3)	(-2.8)	(14.2)
75^{th} percentile	0.073	0.017	-0.004	0.074	0.028	0.000	-0.042
		(23.3)	(-5.5)	(101.4)	(38.4)	(0.0)	(-57.5)
90^{th} percentile	0.104	0.023	0.000	0.040	0.012	-0.001	0.028
		(22.1)	(0.0)	(38.5)	(11.5)	(-1.0)	(26.9)

Notes: The numbers represent the difference in log hourly wages at different points in the distribution between women (aged 45 to 49) with and without *ELA* after adjusting for the specified factors using the indicated decomposition (reference group is those without *ELA*). Share of total difference are presented in parentheses. The unexplained difference is the residual not accounted for by the five factors. The total difference at the mean (0.088) differs slightly from the estimate reported in Table 3.3 (0.078) because the numbers here are based on a single observation per woman.

Consistent with Table 3.5's result that the largest wage effects occur for women in the middle of the IQ distribution, panel B shows that the total log-wage differential associated with *ELA* varies non-monotonically across the distribution and is largest (0.106) at the median. Furthermore, education and occupation explain relatively more of the wage gap (and cumulative experience relatively less) higher in the wage distribution, which accords with the results from Table 3.7 showing stronger cumulative experience in the middle rather than highest IQ group. At the 25th percentile the five components explain nearly all of the wage gap while at the median they explain about 85 percent of the gap; at the 75th percentile, they actually over-explain the gap, suggesting they may be offset by other factors near the top of the wage distribution.

3.7 The "Opt-In" Revolution

In 2003, Lisa Belkin's *New York Times Magazine* article, "The Opt-Out Revolution," reopened the debate about the reasons for persistent differences in women's and men's labor market outcomes. In particular, she argued that the women who might have been the professional equals of men *chose not to be*—these women "opted out" to raise their children. Shang and Weinberg (2009) find some evidence that college graduate women have begun to have more children, but these changes seem small relative to the Opt-In Revolution that began 50 years ago.

This paper quantifies the role of the Pill in catalyzing this revolution. As the Pill provided younger women the *expectation* of greater control over childbearing, women invested more in their human capital and careers. Most affected were women in the middle of the IQ distribution and with some college, who experienced remarkable wage gains over their lifetimes. To put our results into perspective, the Pill-induced effects on wages amount to roughly one-third of the total wage gains for women in their forties born from the mid-1940s to early 1950s.[30] Our decomposition shows that almost two thirds of these Pill induced gains (at the mean) can be attributed to increasing labor-market experience and another third is due to greater educational attainment and occupational upgrading.

What do our estimates imply about the importance of the Pill in narrowing the gender gap from 1980 to 2000? To answer this, we simulate a counterfactual hourly wage distribution from the 1980,

[30] This estimate is obtained by comparing the coefficients for *ELA**40-44 and *ELA**45-49 in Table 3.3 to the total change in wage rates for women in their 40s between the 1943-46 and the 1951-1954 cohorts in the *NLS-YW*. Weinberger and Kuhn (2010) distinguish between changing "levels," the starting wage at labor-force entry, and "slopes," the growth in wages after entry, and argue that changes in "slopes" can account for one third of the narrowing in the gender gap over the last 40 years—a number they argue provides a reasonable upper bound for the importance of all post-schooling investments. Our measures of career investment combine both pre-market investments (e.g., college and occupational choice, which should shift levels) and post-market investments (e.g., labor market experience and on-the-job training, which should shift slopes).

1990, and 2000 population censuses by removing age-specific estimates of early legal access to the Pill from the earnings of cohorts born after 1940 (Table 3.3, column 2) and compute the actual hourly wage distribution for men and women in 1980, 1990 and 2000.[31] From 1980 to 1990, the actual gender gap in real hourly wages for 25 to 49 year olds closed by 0.126 log points, and the simulated gender gap closed by 0.113 log points. From 1990 to 2000, the actual gender gap in real hourly wages closed by 0.074 log points, and the simulated gender gap closed by 0.051 log points. Our main estimates, therefore, imply that 10 percent of the narrowing in the gender gap during the 1980s and 31 percent during the 1990s can be attributed to early access to the Pill. While improvements in contraception play an important role in increasing women's earnings, our results also implicitly highlight the importance of other factors. The unexplained component of cross-cohort changes due to, for example, shifts in the demand for women's labor (e.g., anti-discrimination legislation and enforcement or changes in preferences) as well as shifts in the quality of women's education remain substantial.

Did the Pill unleash the Opt-In Revolution? Our results provide no conclusive answer. They may understate the Pill's broader influence because our empirical strategy does not allow us to explore the effect of changes in access to the Pill beyond age 20 and fails to capture the potentially large social multiplier effects. For instance, the Pill's availability likely altered norms and expectations about marriage and childbearing and firms' decisions to hire and promote women-even among cohorts without legal access to the Pill. Thus, the effects of the Pill may be larger than we find, though it is not clear how much larger. Even these conservative estimates, however, suggest that the Pill's power to transform childbearing from probabilistic to planned shifted women's career decisions and compensation for decades to come.

[31] Real hourly wage is total wage and salary earnings of last year divided by the product of weeks worked last year and usual hours worked per week and divided by the PCE deflator to get year 2000 dollars. The estimates use IPUMS person weights and exclude real hourly wage outliers of less than \$2 or more than \$200. The sample contains native-born women ages 25 to 49 whose wages were not imputed and who were not self-employed. The simulated log hourly earnings values are adjusted by subtracting the estimates in column 2 of Table 3.3 for women who were born in or after 1940 and born in a state where they would have had early access to the Pill.

3.8 Appendices

3.8.1 Appendix 3.A: Data and Specifications

This appendix summarizes the creation of the variables used in the analysis as well as the construction of the alternative specifications used for Figures 3.3 and 3.4. The independent variables, including the key *ELA* measure, are described first, followed by the sequence of dependent or outcome variables. (The dependent variables are available in every wave of the survey unless otherwise stated.) Finally, each alternative specification is discussed.

Age and year of birth: Determining the age of the respondents at each survey is crucial, both in identifying early legal access, which is age dependent, and because the effects of early legal access are likely to vary over the lifecycle. Both age at time of interview and date of birth (month and year) are asked in various waves of the survey; however, they are not always consistent. Date of birth was asked in 1968, 1977, 1978, 1982, 1988 and 1991 and confirmed or corrected in 1995, 1997, 1999, 2001, and 2003. Of the 5,159 women in the sample, 94 (1.8 percent) had conflicting birth date reports, and another 818 (15.9 percent) had only a single report. For the conflicting cases, all available data were used to check birth reports, but, in most cases, the modal reported year and month of birth was used.[32] From the date of birth information, age at the end of each survey year (not at the time of interview) was constructed for consistency between early and later waves.[33]

State of residence: The geocoded version of the *NLS-YW*, available at Census Research Data Centers, contains the state of residence of each respondent for each wave of the survey. Using respondents' age information and variables pertaining to mover status in the public-use data, one can construct variables for the state of residence at key ages (such as 18, 19, 20, and 21) for most but not all respondents. In some cases, women exit the sample before they reach the key ages; in others, women in the older cohorts who move frequently during the key ages are not observed until they are older. Nonetheless, for each of the key ages (18 through 21), between 80 and 90 percent of the respondents were successfully matched to a state of residence.

Early Legal Access to the Pill (*ELA*): By researching state laws, the authors compiled a list of the years in which each state legally allowed unmarried women (of age 20 or younger) to have access to the birth control pill (see Appendix 3.B: Legal Variables). Using the restricted version of the *NLS-YW*, state of residence at each survey is observed and the respondents' state of residence at age 21 is used to generate the *ELA* variable. A respondent's *ELA* status was coded 1 if her year of birth plus 20 was greater than or equal to the year in which her residence state at age 21 first allowed legal access. State of residence at age 21 rather than age 20 was used because it was identifiable for more women (4,419 versus 4,398) and the correlation between the two was high (r = 0.94).

Early Abortion Access (EAA): Five states (Alaska, California, Hawaii, Washington, and New York) and the District of Columbia legalized abortion in 1970, three years before *Roe v. Wade*. We code a respondent as having *EAA* if she lived in one of the above areas at age 21 and was born in 1950 or later; these are the cohorts of women who had legal abortion access in their states of

[32] The exact code is available from the authors upon request.

[33] The early waves sampled respondents in the early months of the year but later waves sampled respondents in later months.

residence before the age of 21. To address the possibility that women crossed state lines to obtain an abortion, we also constructed a measure of the distance in miles between each state's population centroid in 1970 and the closest major location providing abortions in the pre-*Roe* period (District of Columbia, Los Angeles, San Francisco, Buffalo, and New York City. This distance was then transformed into its natural logarithm.

Age at first marriage: Although age at first marriage is directly asked in 1968, this is useful only for women who had been married prior to the first interview. To determine marital ages for the rest of the sample, three additional sources are used: (a) marital histories, (b) changes in current marital status, and (c) timing of changes in marital status. Marital history questions are asked in 1978, 1983, 1997, 1999, 2001, and 2003. In 1978 and 1983, the questions ask about up to the three most recent marriages (including the current one); in the latter years, only the date of the most recent marriage is asked. Current marital status is asked in every survey year. Changes in marital status are reported in 1969 and 1970 and every survey year from 1985 onwards. We observe no first marriage date for 809 women. This outcome is only used in Appendix Table 3.1.

Wages and salary earnings: Hourly rates of pay for the current or most recent job (measured in cents) and annual wage and salary earnings from the previous calendar year are available for years 1968 through 1993. For 1995 through 2003, the hourly rate of pay variable is for the first (main) job, and annual wage and salary earnings are for the previous 12 months rather than the previous calendar year. Information on wages and salary earnings excludes farm, business, or self-employment income. Each of the wage, earnings, and income variables is converted from nominal to 2000 dollars using the PCE deflator and then converted to natural logarithms. Although there is no effective top code to hourly wages, annual earnings are subject to censoring from above, with the top code varying across years. (Generally, fewer than 2 percent of women have top-coded earnings in any year.) In the analysis, hourly wage outliers (less than 2 or more than 100 real dollars) are excluded.

Cumulative experience: We measure cumulative work hours at the start of each calendar year as the sum of hours of work reported since 1967. We approximate hours of work with the product of usual weekly hours and our best estimate for the number of weeks worked each year.

We rely on three sets of questions to compute number of weeks worked. In 1968, 1969, 1975, 1977, 1980, 1982, 1985 and 1987, respondents were asked to report the number of weeks they worked in the previous calendar year. In 1970, 1971, 1972, 1973, 1978, 1983, 1988, 1991 and 1993, the survey asked the number of weeks worked since the last eligible interview, regardless of whether or not that interview took place. In 1970, 1971, 1972, 1973, 1995, 1997, 1999, 2001 and 2003, they survey asked weeks worked since the last actual interview. We combine these measures as available, being careful to avoid double-counting. (This procedure is complicated and idiosyncratic to each survey wave; the code used is available upon request.)

Despite our best efforts, we note that it is not possible to create a truly comprehensive measure of weeks worked for several reasons. First, there are some gaps in coverage for which no weeks worked questions were asked: The initial shift from calendar year to survey period leads to a small time period (generally under 6 weeks) for which we have no measure of weeks worked. The size of this coverage gap increases over time. For example, we miss nine to eleven months between the 1973 interview and January 1, 1974, and the entire calendar year of 1975. Second, item non-response for a question regarding weeks worked poses a significant problem because cumulative experience is dependent on all past responses. It is only possible to recover cumulative experience for women who miss an interview and are subsequently re-interviewed *if* the later interview asks about weeks worked since the last actual interview.

Our main measures address these concerns with additional sample restrictions or assumptions. We address the coverage issue by rescaling the experience measure to a base of full coverage. We effectively assume that the fraction of weeks *observed* working is the same as the fraction of weeks *elapsed* spent working; that is, we scale the cumulative weeks worked measure by the ratio of total weeks elapsed to total weeks for which there is coverage. For the second problem, we exclude women once they have an episode of an item non-response for the weeks worked question. For the third problem, we restrict estimation to women who have a valid weeks report in every survey wave (no missed interviews and no item non-response). None of these alternate measures, whether used individually or all together, changes the qualitative pattern of results we find of *ELA* on cumulative experience. The numbers and estimates reported in Table 3.4 apply the first and second measures but exclude the third in the interest of maintaining a larger sample size.

College enrollment: Using questions that asked about current enrollment in an academic program of study, as well as the highest grade completed, a respondent was coded as enrolled in college (a binary variable) if she was enrolled and the highest grade completed was at least 12. As a result, "college enrollment" includes all forms of academic post-secondary education but excludes vocational/occupational training. Note that women who did not graduate from high school are excluded (coded as missing).

Highest grade completed: The basis of these variables is the set of revised highest grade completed questions. Although the "revised" set has supposedly been cleaned and corrected of errors found in the original highest grade completed questions, an inspection revealed that several problems remained, and these were often some form of non-monotonic progression. Five hundred thirteen women (10.0 percent) had at least one discrepancy, but in most cases these were minor, such as a jump up or down of one grade in a single survey wave before returning to trend. The "revised" variables were cleaned further of likely misreports using responses from previous and later years. Specifically, "jump" deviations that last only a single wave (in some cases, two waves) are smoothed by replacing these values with those that occur both before and after the deviation. For example, a woman whose highest reported grade is 12 in 1975 and 1977, 10 in 1978, and 12 in 1980 and 1982, would have the 1978 value recoded to 12. This procedure leaves 205 women (4.0 percent) with a non-correctable discrepancy, such as multiple, non-monotonic jumps; these respondents are flagged and excluded from the analysis. Including these women alters the results very little.

Labor-force participation: Labor-force participation (LFP) is based on the employment status recode (1968 through 1993) or monthly labor recode (1995 through 2003) variables. The LFP dummy variable takes the value of 1 if the respondent is employed at the time of the survey (whether at work or not) or unemployed, and 0 otherwise. Note that choice of specific activities in the survey for non-labor-force participants changed between 1993 and 1995, when the *NLS-YW* adopted the new CPS definitions. Results using this measure are reported in footnote 21.

Usual weekly hours: These variables are based on a question asking about the usual hours worked per week at the respondent's job. For most years, the job is defined to be either the one currently held or the job most recently held since the last interview; however, in 1970, 1971, 1972, 1973, 1978, and 1983, the question pertains to the current job only. In these cases, another question specifically referring to the usual hours worked at the most recent job is used to supplement the current job question to maintain comparability: Respondents with missing values for the current job only question are replaced with the usual hours worked from the most recent job question. Finally, because responses in some years are top-coded at 99 hours while some are not, values above 99 are recoded to exactly 99. This affects no more than 1 to 3 women in any year and has a negligible impact on the estimates.

Occupational training: Although the *NLS-YW* asks several questions throughout the survey waves about occupational training, the questions are not completely consistent across waves. In 1968 and again from 1980 through 2003, the survey asked whether respondents had undergone (a) any on-the-job training since the last interview, and (b) any other occupational or vocational training. From 1969 to 1978, however, these two different types of training were co-mingled in a single training question. For consistency, both training types are combined into a single (binary) indicator that captures whether the respondent underwent any form of vocational or occupational training, on-the-job or otherwise, since the last interview. The estimation sample for training includes only respondents who were not currently attending an academic program, because training questions were asked only of respondents not enrolled in an academic program until 1975.

Occupation: For each wave of the survey, there is a variable containing the 3-digit Census code of the respondent's current or most recent job. Through 1993 the variable is for current or most recent job; for 1995 through 2003, when the new (circa 1994) CPS definitions were used, the variable for job 1 (the main job) is used. Unfortunately, a consistent coding is not available in the data. The coding at the beginning of the survey is based on the 1960 scheme, and it is available through 1993. Coding based on the 1980 scheme begins in 1980 and runs through 1999; the 1990 scheme runs from 1993 through 2001; and the 2000 scheme runs from 1995 through 2003. Thus, there is significant overlap for several years. In the interest of creating a longer series, we aggregate the different coding schemes by collapsing the 3-digit job codes into four groups that can be made consistent over the entire time period. We use a coding scheme as soon as it becomes available, so we use the 1960 scheme for data years 1968 through 1978, the 1980 scheme for years 1980 through 1991, the 1990 scheme in 1993, and the 2000 scheme for years 1995 through 2003. The four groups are: all professional and managerial jobs, non-traditionally female professional and managerial jobs, clerical and sales jobs, and all other jobs. "All professional and managerial jobs" generally includes any 3-digit code that falls under the "professional, technical and kindred workers" or "managers, officials, and proprietors except farm" categories (or their equivalent) from any of the coding schemes. "Non-traditionally female professional and managerial jobs" is a subset of the first category that excludes the traditionally female occupations of nurses and elementary, secondary, and not elsewhere classified (n.e.c.) teachers. "Clerical and sales jobs" includes 3-digit codes listed under the clerical or sales categories, and "all other jobs" includes all 3-digit codes not in one the previous groups, including craftspeople, operatives, agricultural workers, and service jobs. The complete list of 3-digit Census job codes to our four groups by coding scheme is available by request. For the analysis in Table 3.4, a woman must be currently employed to be counted in one of the four job groups; if she reported a 3-digit code in the survey but also reports not being currently employed, we code her as a zero in all four job categories.

IQ and Childhood Family Socioeconomic Status: The 1968 wave of the *NLS-YW* included a questionnaire for the high schools of the respondents, which in addition to asking about school characteristics also asked for the most recent intelligence or aptitude test of the respondent. Scores were reported for 3,530 of the respondents (though almost none for respondents born in 1953). (Griliches, Hall and Hausman (1978) provide an assessment of whether scores are missing at random in the *National Longitudinal Survey of Young Men*, the nearly identical survey for men, and conclude that they very nearly are.) The agency that processed the *NLS-YW*, the Center for Human Resource Research (CHRR), converted these scores from various tests composites to a unified "IQ score" based on a normally-distributed national population with mean 100 and standard deviation 15. (More information on this procedure can be found at http://jenni.uchicago.edu/evo-earn/IQ.pdf.) Based on this distribution and the unified score, a respondent was also classified into an IQ quantile and stanine. Using information from the initial survey wave on father's occupation and education, mother's education, eldest sibling's education, and availability of reading material at home, CHRR also constructed a summary family socioeconomic status variable to follow a normal distribution with mean 100 and standard deviation 30. Our analysis breaks these measures into tertiles.

Attrition: In most cases, the empirical analysis has made no attempt to restrict the sample to non-attriters. The decision to exploit every person-year observation was made in order to maximize sample size. One of our sensitivity checks, reported in Figures 3.3 and 3.4, shows that findings based upon a balanced panel of individuals are very similar to those reported in the paper. In addition, regressions, available upon request, show no correlation between each year's interview status and *ELA*.

Variables Used in Table 3.1 Balancing Tests

1. **Father worked for pay**: binary variable equal to one if a respondent's father worked for pay when respondent was 14. About 93 percent of the sample had a father working for pay at age 14. (Note: This is *not* conditional on having a father in the HH).

2. **Father held professional job**: binary variable equal to one if a respondent's father had a "professional" job when respondent was 14. "Professional" has the same coding as in the main results, based on 1960 occupational definitions. About 20 percent of the sample had a father working in a professional job. (Note: This is conditional on having had a father working at age 14).

3. **Mother worked for pay**: binary variable equal to one if a respondent's mother worked for pay when respondent was 14. This was *not* asked of respondents who lived with their mother as the sole parent. About 39 percent of the effective sample had a mother working for pay at age 14. (Note: This *is* conditional on having a father (or other male adult) in the HH).

4. **Mother held professional job**: binary variable equal to one if a respondent's mother had a "professional" job when respondent was 14. "Professional" has the same coding as in the main results, based on 1960 occupational definitions. About 13 percent of the sample had a mother working in a professional job. (Note: This *is* conditional on having had a mother working at age 14).

5. **Duncan index of household head**: Duncan index socioeconomic job score of head of household when respondent was age 14, as created by CHRR in the data. Values are conditional on the head (not necessarily father) working when respondent was 14. (The scale runs from 3 to 97).

6. , **Socio-economic status**: socioeconomic index of respondent's parents in 1968, as provided in the data. Based on father's occupation and education, mother's education, eldest sibling's education, and availability of reading material at home. By construction, SES is distributed N(100,900).

7. **Magazines in home**: binary variable equal to one if a respondent had magazines available at home when she was age 14. About 64 percent of the sample did.

8. **Newspapers in home**: binary variable equal to one if a respondent had newspapers available at home when she was age 14. About 83 percent of the sample did.

9. **Respondent held library card**: binary variable equal to one if a respondent had a library card when she was age 14. About 70 percent of the sample did.

10. **Two-parent household**: binary variable equal to one if a respondent lived in a household with two parents (including step-parents) at age 14. About 80 percent of the sample lived with two parents at age 14.

11. **Number of siblings**: number of siblings of respondent in 1968 (not necessarily in the household); we can't reliably determine whether this includes step- and half-siblings.

12. **Father born in U.S.**: binary variable equal to one if a respondent's father was born in U.S./Canada. About 96 percent of sample had the father born in U.S./Canada.

13. **Highest grade completed by father**: highest grade completed by father, in 1968. Conditional on having a father in household. Item non-response is relatively high; *ELA*, however, is uncorrelated with whether father's HGC is observed.

14. **Highest grade completed by mother**: highest grade completed by mother, in 1968. Conditional on having a mother in household. Item non-response is relatively high; *ELA*, however, is uncorrelated with whether mother's HGC is observed.

15. **Parents' education goals for respondent**: number of years of schooling respondent's parents want respondent to obtain, when respondent was 14.

16. **Atypicality index of mother's job**: atypicality index of respondent's mother's job when respondent was 14, conditional on respondent's mother working then. Atypicality index is the female percentage of an occupation minus the percent of the experienced civilian labor force that was female in 1970; negative numbers indicate more atypical occupations.

17. **Respondent's IQ score**: continuous IQ score of respondent. Reference distribution is independent national norm, not empirical sample. Only two-thirds of entire sample had an IQ or achievement test administered; while these two-thirds were slightly above national norms, the presence of an IQ score is uncorrelated with *ELA*.

18. **Rural residence**: binary variable equal to one if a respondent resided on a farm/ranch or in another rural area at age 14. About 26 percent of the sample lived in a rural area at age 14.

Alternative Specifications

Figures 3.3 and 3.4 include six specifications: one following equation (3.1) called our baseline specification, one following equation (3.1') that augments our baseline specification with abortion controls, and four alternative specifications of (3.1') described below. Tabular presentation of estimates from equation (3.1') are presented as the main tables of the paper.

Linear state-specific time trends: The specification in equation (3.1') is augmented with the interactions of each state of residence dummy with the year of observation.

Vietnam casualties: Using data from the National Archives on the Vietnam Conflict (`http://www.archives.gov/research/military/vietnam-war/electronic-records.html`), the specification in equation (3.1') is augmented with controls for state-level casualties. These controls include state-specific annual death rates lagged one, two, and three years; and cohort-specific, state-level death rates within two years of a woman's date of birth.

Balanced panel: The specification in equation (3.1') is estimated on a sample that is restricted to women who are interviewed in every survey wave from 1968 through 2003 and successfully answer all relevant questions (no item non-response).

High school state: This specification uses state of residence during high school (rather than at age 21) for all state-based variables. Like state of residence at age 21, this variable is created using each wave's state of residence, move histories, and tenure at current residence. Because older cohorts are father removed from high school age, they are less likely to be successfully matched, particularly if they moved frequently. (While this problem exists for state of residence at age 21, it is more pronounced for high school state.)

3.8.2 Appendix 3.B: Legal Coding

The coding used in this paper relies upon the updated coding of Bailey and Guldi (2009) and differs from the coding used in Bailey (2006) for 15 states. These differences in coding reflect two main changes: (1) Non-specific female age of majority statutes are not treated as emancipation for the purpose of consenting for medical care unless this is specifically noted in the statute. As a result, the coding changes in 4 states. (2) Statutes were interpreted incorrectly, enforcement was ambiguous, or earlier statutes, policy changes or attorney general decisions were found. These changes affected coding in 11 states; in six of these cases, the date of legal change shifts by only one or two years. These legal changes are summarized in the table below, and then the explanation of each of the changes is discussed in detail, including legal citations by state.

Dates of Legal Change Granting Early Access to the Pill

State	Bailey (2006)	Bailey and Guldi (2009)	Different?	Reason for recoding
Alabama	1971	1971		
Alaska	1960	1960		
Arizona	1972	1972		
Arkansas	1960	1973	X	FAOM → AOM
California	1972	1972		
Colorado	1971	1971		
Connecticut	1972	1972		
Delaware	1972	1972		
District of Columbia	1971	1971		
Florida	1974	1974		
Georgia	1968	1968		
Hawaii	1970	1972	X	TFP → AOM
Idaho	1963	1972	X	FAOM → AOM
Illinois	1973*	1969		
Indiana	1973	1973		
Iowa	1973	1972	X	Earlier AOM
Kansas	1970	1970		
Kentucky	1968	1965/68?	X	Ambiguous
Louisiana	1972	1972		
Maine	1971	1969	X	Earlier AOM
Maryland	1967	1971	X	TFP → MM
Massachusetts	1974	1974		
Michigan	1972	1972		
Minnesota	1973	1972	X	Earlier AGD
Mississippi	1966	1966		
Missouri	1976	1973	X	Earlier AGD
Montana	1971	1971		
Nebraska	1972	1972		

Nevada	1969	1973	X	FAOM → AOM
New Hampshire	1971	1971		
New Jersey	1973	1973		
New Mexico	1971	1971		
New York	1971	1971		
North Carolina	1971	1971		
North Dakota	1972	1972		
Ohio	1965	1960	X	MM
Oklahoma	1966	1972	X	FP → AOM
Oregon	1971	1971		
Pennsylvania	1971	1970	X	Earlier MM
Rhode Island	1972	1972		
South Carolina	1972	1972		
South Dakota	1972	1972		
Tennessee	1971	1971		
Texas	1974	1974		
Utah	1962	1975	X	FAOM → AOM
Vermont	1972	1972		
Virginia	1971	1971		
Washington	1971	1968	X	AOM → FP
West Virginia	1972	1972		
Wisconsin	1973	1972	X	Earlier AOM
Wyoming	1969	1969		
Differences in coding			15	

Notes: Legal change is coded as the earliest date, at which an unmarried, childless women under age 21 could legally consent for medical treatment without parental or spousal consent. A full legal appendix and scans of statutes are available from Bailey and Guldi (2009). FAOM → AOM: lower female age of majority changed to the legal majority for men and women for all purposes. FP → AOM: family planning law changed to age of majority law; AOM → FP indicates the reverse. TFP → AOM/MM: erroneously coded treatment for pregnancy statute changed to be the date for the change in legal age of majority/mature minor doctrine. Earlier AGD/AOM/MM indicates that an earlier attorney general decision/age of majority/mature minor doctrine was located. *Illinois is a typo in the published version of Bailey (2006) that the author did not catch before publication. The correct coding and the coding used in her analysis is 1969. See below for more details.

Arkansas: Bailey (2006) coded the 1948 Arkansas statute that stipulated that females over 18 were of the age of majority [AR Code §9-25-101 (1987), AR Stat. Ann. §57-103 (1947)], but it is unclear that this law treated women as legal adults except for marriage. Effective July, 1973, Arkansas passed a law allowing pregnant minors of any age to consent to medical care other than abortion (Merz et al. 1995: footnote 150; Acts 1973, No. 32, §1, p.1028). The law provided that *any* female could consent to medical treatment or procedures "for herself when in given [sic.] connection with pregnancy or childbirth, except the unnatural interruption of a pregnancy" [AR R.S. §82-363 (1976)]. The statute goes on to grant the power of consent to "any unemancipated minor of sufficient intelligence to understand and appreciate the consequences of the proposed surgical or medical treatment or procedures" [ibid.]. Bailey and Guldi (2009), therefore, code a mature minor doctrine as of 1973.

Hawaii: Bailey (2006) erroneously codes a "treatment for pregnancy" statute as a mature minor doctrine: "The consent to the provision of medical care and services by public and private hospitals or public and private clinics, or the performance of medical care and services by a physician licensed to practice medicine, when executed by a female minor who is or professes to be pregnant" [HI Rev. Stat. §577A-2 (1999), L. 1968, c. 58]. Under this law, *only* minors professing to be pregnant or having a venereal disease could consent to "medical care," defined as "the diagnosis, examination and administration of medication in the treatment of venereal diseases and pregnancy" [L. 1968, c. 58, §4]. This law did not permit non-pregnant teens to be treated or prescribed contraception legally. Bailey and Guldi (2009) code the legal change in the age of majority, effective March 28, 1972, which lowered the age of majority to 18.

Idaho: Bailey (2006) codes a female age of majority statute [ID Code Ann. §31-101 (1932)], but it is unclear whether consent to contraception would have been covered under this statute. Bailey and Guldi (2009) found a 1972 amendment that equalized the ages of majority for males and females at 18 and extended this majority for *all* purposes [ID Code §32-101 (1983); am. 1972, ch. 117, S1, p. 233].

Iowa: Bailey (2006) codes the change in the legal age of majority to 18 in 1973. Bailey and Guldi (2009) located and code an earlier change in the legal age of majority from 21 to 19 in 1972 [IA Code Ann. §599.1 (1954), Acts 1972 (64 G.A.) ch. 1027, §49; Acts 1973 (65 G.A.) ch. 140, §49].

Kentucky: Bailey and Guldi (2009) codes a law, effective January 1, 1965, that lowered the legal age of majority "for all purposes" in Kentucky to 18 [KY R.S. §2.015 (1967), enacted Acts 1964, ch. 21, §1]. Because this Council of State Governments publication in 1973 noted that this 1965 had law prompted "a good deal of confusion [about the exact privileges granted to those 18 and older] and four years later [a] clarifying statute was passed" [1972: pp.12-3], Bailey (2006) codes the 1968 amendment to the age of majority statute that included the clause "all other statutes to the contrary notwithstanding" [KY Acts ch. 100, §1, approved March 25, 1968] that clarified the interpretation of the statute.

Maine: Bailey (2006) codes a change in the legal age of majority passed in 1971 which lowered the legal age of majority to 18 [1 M.R.S.A. §73 (1979); 1969, c. 433 §8; 1971 c. 598, §8]. Bailey and Guldi (2009) located an earlier statutory change in the age of majority, effective October 1, 1969, which lowered the legal age of majority in Maine from 21 to 20.

Maryland: Bailey (2006) erroneously codes a "treatment for pregnancy" statute based upon Merz *et al.* (1995: footnote 388), which notes that minors could consent to medical treatment for "alcohol and drug abuse, venereal diseases, pregnancy, contraception other than sterilization, and in cases of rape or sexual abuse" since June 1, 1967. However, the specific language relating to contraception was not added until 1971. The original statute, effective June 1, 1967, restricted the law to "apply . . . to minors who profess to be in need of hospital or clinical care or services or medical or surgical care or services to be provided by a physician licensed to practice medicine, whether because of suspected pregnancy or venereal disease, regardless of whether such professed suspicions of pregnancy or venereal disease are, or are not subsequently substantiated on a medical basis" [MD Laws 1967 ch. 468]. Therefore, Bailey and Guldi (2009) code the 1971 revision to the 1967 statute that eliminated the restriction to pregnant minors or minors suspected to be pregnant.

Minnesota: Bailey (2006) codes the change in the age of majority to 18 effective June 1, 1973 [Minn. Stat. §518.54(2) (1990)]. One year prior to the change in the age of majority, on May 27, 1971, a series of statutes concerning the consent to medical care of minors became effective. One section provides for an extension of the rights of emancipated minors [MN Stat. Ann. §144.341 (1989); see also CA Civil Code §34.6 (1982)]. Although ambiguous in their applicability to consent for birth control, a 1972 Attorney General decision interpreted these statutes as "not making it a crime for physicians to furnish birth control devices to minors" [From LexisNexis Academic: Minn. Stat. §144.341-144.347, 617.251 (1971), No. 494-b-39, 1972 Minn. AG LEXIS 35]. The interpretation of these statutes remained in dispute for some time; they were again challenged in *Maley v. Planned Parenthood of Minnesota, Inc.* Cir. Case No. 37769 (Minn. Dist. Ct., Third Jud. Dist., Jan. 5, 1976). In this case, six couples filed a class action lawsuit, seeking to prevent Planned Parenthood from providing contraceptive services to unemancipated minors without parental consent (Paul, Pilpel and Wechsler, 1974; http://www.popline.org/docs/730457). However, the Minnesota District Court upheld the constitutionality of sections 144.343 and 144.344, writing that "under these sections Planned Parenthood could provide minors with contraceptive information and services without parental consent, unless a parent specifically notifies Planned Parenthood that he/she does not wish his/her child to receive such services" (DHEW 1978, p.244).[34] This decision, therefore, reinforced the attorney general's broad interpretation of the statute. Legally, Planned Parenthood could provide contraceptives to unmarried minors as long as they had not been explicitly informed by parents. Bailey and Guldi (2009), therefore, revise the coding to reflect the 1972 attorney general decision.

Missouri: Bailey (2006) coded the P*lanned Parenthood of Central Missouri v. Danforth* decision [428 U.S. 52 (1976)], in which the Supreme Court ruled that the state could not prohibit minors from obtaining abortions and, by extension, contraception. Bailey and Guldi (2009) located an earlier Attorney General decision issued in March of 1973 stating that "no law prohibits physicians from prescribing contraceptives to minors who do not have parental consent or who have not been emancipated by marriage or other means" [DHEW 1978, p. 253, citing Op. Atty. Gen. 3/9/1973].

Nevada: Bailey (2006) codes a 1969 lower female age of majority statute, but this statute was in effect since at least 1930 and applied only to women's ability to enter into contracts [NV C.L. §300 (1930); NV R.S. §129.010 (1963); see also DHEW 1974, p. 236]. Bailey and Guldi (2009) code a 1973 amendment to the age of majority statute which equalized the ages of majority for males and females at 18 [N.R.S. §129.010 (2003); 1973, p. 1578].

[34] Though the final *Maley* ruling was not issued until 1976, according to Paul, Pilpel and Wechsler (1974), the district court came to the same conclusion during a preliminary stage of the case in 1973.

Ohio: Ohio courts adopted a mature minor doctrine as early as 1956. The *Lacey v. Laird* [166 Ohio St. 12, 139 N.E. 2d 25 (1956)] opinion states:

> A charge that this 18-year-old plaintiff [who had nose surgery when she was 18 without her parents' consent] could not consent to what the jury could have found was only a simple operation, would seem inconsistent with the conclusion of our General Assembly, that any female child of 16 can prevent the taking of liberties with her person from being raped merely by consenting thereto at the time such liberties are taken.My conclusion is that performance of a surgical operation upon an 18-year-old girl with her consent will ordinarily not amount to an assault and battery for which damages may be recoverable even though the consent of such girl's parents or guardian has not been secured [139 N.E. 2d at 34].

Legal interpretations held that minors could consent to minor surgery and general medical care under this decision (DHEW 1974: 265), but Ohio also had an anti-obscenity statute. Ohio's statute originally passed in 1885 and banned the dissemination of information and supplies relating to contraception. The words "for the prevention of conception" were removed from Ohio's statute in 1965, so Bailey (2006) coded 1965 as the earliest date that an unmarried minor could obtain the Pill legally. However, Ohio's statute went on to note that "nothing in this section [about contraception and obscenity] or the next two sections shall be construed to affect teaching in regularly chartered medical colleges, or the publication of standard medical books, or the practice of regular practitioners of medicine, or druggists in their legitimate business" [OH R.S. 7027 (1896)] [April 30, 1885: 82 v. 184]. It is not clear how to interpret this physician and pharmacist exceptions, which makes it unclear whether to code Ohio as 1960, when the Pill was introduced (this assumes that the obscenity statute was not binding for physicians), or 1965, when the law was amended to omit language about contraception (this assumes the obscenity statute was binding for physicians).

Oklahoma: Bailey (2006) coded a family planning statute [OK Stat. Ann. Tit. 63 Ch. 32, §2071-5 (1984)]. Although no explicit eligibility requirements are stated in the statutes, the Department of Health Education and Welfare (DHEW) contacted the state about their policy and reported that, "[a]ll categories of adults apparently are eligible for family planning services; no exclusions were noted in the CFPPD survey and none appear in the written policies. According to the Division of Maternal and Child Health's *Guidelines for Family Planning Programs*, 'minors may be accepted for services if: 1) ever married or ever pregnant; 2) bearing acceptable proof of impending marriage; 3) accompanied by parent or guardian requesting services; 4) referred by a recognized agency, a doctor, a nurse, or a clergyman'... [However,] contraceptive advice may be given in *all* cases where the 'health needs of the patient make it advisable..."' (1974, p.271). Because these policies only allow legal minors who are pregnant to obtain contraceptive *advice*, Bailey and Guldi (2009) code the change in the legal age of majority which was amended and effective in August 1, 1972, which equalized the ages of majority for men and women at 18 [OK Stat. Ann. Tit. 15 §13 (1972); L. 1972, c. 221, §1].

Pennsylvania: Bailey (2006) coded a mature minor doctrine effective in 1971, but Bailey and Guldi (2009) located an earlier mature minor statute, enacted on February 13, 1970 and effective in April 1970, that allowed any minor 18 or over to consent to medical care: "Any minor who is eighteen years of age or older... may give effective consent to medical, dental and health services for himself or herself, and the consent of no other person shall be necessary" [PA Stat. tit. 35, §10101 (1977)].

Utah: Bailey (2006) coded the lower age of female majority, but this statute's application was unclear with respect to medical care. Policy documents indicate there was considerable ambiguity regarding whether physicians could prescribe birth control to unmarried women under age 21. On July 21, 1971, the Attorney General advised "not to provide family planning information or services to minors without parental consent 'until such time as the state legislature may adopt appropriate legislation.' ... In support of this view the Attorney General cites the common law requirement of parental consent in the absence of an emergency, plus the expression of legislative intent inferred from the statute dealing with prophylactics..." (DHEW 1974: 300 citing Op. Atty. Gen. No. 71-017, July 21 1971). Bailey and Guldi (2009), therefore, code the amendment to this statute in 1975 to make both men and women legal adults at the age of 18 for all purposes [L. 1975, ch. 39, §1, approved March 24, 1975].

Washington: Bailey (2006) codes the legal age of majority "for all purposes" which changed from 21 to 18 in 1971. Bailey and Guldi (2009) located an earlier policy change and code 1968, because a Washington Board of Health Policy directed that all persons were eligible for family planning without parental consent, including never-pregnant, never-married minors [WAC248-128-001 for Board of Health policy adopted August 3, 1967, codified July 1, 1968].

Wisconsin: Bailey (2006) erroneously coded the date of 1973 as the year the legal change in age of majority to 18 became effective [WI Laws 1971, ch. 213; see also DHEW (1978: 363)]. In fact, this statute became effective in March 23, 1972. Bailey and Guldi (2009), therefore, code 1972.

Appendix Table 3.1: The Impact of *ELA* on the Timing of First Marriage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Age at first marriage	1=Married before 19	1=Married before 20	1=Married before 21	1=Married before 22	1=Married before 23	1=Married before 24
Mean of DV	*21.2*	*0.270*	*0.396*	*0.505*	*0.597*	*0.671*	*0.721*
ELA	0.427	-0.064	-0.059	-0.020	-0.018	-0.004	-0.007
	(0.270)	(0.022)	(0.023)	(0.024)	(0.029)	(0.031)	(0.033)
Observations	3786	4210	4204	4200	4200	4200	4200
(Pseudo) R-squared	0.04	0.04	0.03	0.03	0.03	0.03	0.02
Fixed effects	S, Y	S, Y	S, Y	S, Y	S, Y	S, Y	S, Y

Notes: This table presents estimates of *ELA* on age of first marriage among those ever married (column 1) and binary indicators for whether the respondent was married before age a, for $a = 19, \ldots, 24$. The table uses the 1943 to 1953 birth cohorts from the *NLS-YW*. The sample in columns (2) through (7) includes women who never get married and the estimates represent average partial effects from a probit. Changes in sample size across columns (2) through (7) are due to dropping of observations that do not contribute to the likelihood. The R-squareds for columns (2) through (7) are pseudo (McFaddens) R-squareds. All regressions include state fixed effects (S) and cohort fixed effects (Y). Heteroskedasticity-robust standard errors are corrected for clustering at the state level and are presented in parentheses below each estimate.

Appendix Table 3.2: The Impact of *ELA* on Pill Use among Ever Married Women, with State Linear Time Trends

	(1)	(2)	(3)	(4)	(5)
	1=Used Pill before age 18	1= Used Pill before age 19	1= Used Pill before age 20	1= Used Pill before age 21	1= Used Pill before age 22
Mean of DV	*0.034*	*0.119*	*0.226*	*0.369*	*0.506*
Panel A: Pill Use					
ELA	-0.072	0.381	0.204	0.210	0.133
	(0.030)	(0.167)	(0.209)	(0.106)	(0.046)
R-squared	0.070	0.138	0.141	0.156	0.142
Panel B. Pill Use Heterogeneity					
ELA	-0.067	0.443	0.246	0.292	0.169
	(0.030)	(0.142)	(0.185)	(0.114)	(0.065)
ELA x Non-metro area	-0.006	-0.102	-0.072	-0.141	-0.070
	(0.014)	(0.055)	(0.067)	(0.071)	(0.059)
R-squared	0.070	0.139	0.142	0.157	0.143
Observations	1985	1985	1985	1985	1985
Fixed effects	S, Y	S, Y	S, Y	S, Y	S, Y
State linear time trends	Yes	Yes	Yes	Yes	Yes

Notes: See notes to Table 3.2

Appendix Table 3.3: Heterogeneity in Real Hourly Wage Growth: State Linear Time Trends

	(1)	(2)	(3)	(4)	(5)
Sample	Lower third of IQ distribution	Middle third of IQ distribution	Upper third of IQ distribution	No College	Some College
ELA * Age 20-24	-1.283**	0.380	-0.610	-0.621**	-0.907*
	(0.631)	(0.590)	(0.491)	(0.266)	(0.493)
ELA * Age 25-29	-0.530	0.780	0.240	-0.350	-0.190
	(0.558)	(0.716)	(0.491)	(0.257)	(0.510)
ELA * Age 30-34	-0.800	1.868**	0.700	0.100	0.760
	(0.558)	(0.761)	(0.687)	(0.303)	(0.625)
ELA * Age 35-39	0.450	2.040**	0.740	0.120	1.610**
	(0.745)	(0.810)	(0.623)	(0.409)	(0.735)
ELA * Age 40-44	0.520	2.477**	1.080	1.042**	1.691**
	(1.108)	(0.935)	(0.704)	(0.470)	(0.710)
ELA * Age 45-49	2.121*	2.625**	3.507***	1.524***	3.184***
	(1.204)	(0.980)	(1.067)	(0.445)	(0.891)
Observations	10468	14165	16788	40229	21785
Unique women	793	975	1112	2895	1456
R-squared	0.20	0.22	0.25	0.19	0.28
Mean of DV for 20-24	5.59	6.49	7.18	5.49	7.21
Mean of DV for 25-29	5.89	6.79	8.69	5.52	9.51
Mean of DV for 30-34	6.59	7.19	8.94	6.18	9.74
Mean of DV for 35-39	7.44	8.40	10.79	7.16	11.42
Mean of DV for 40-44	8.34	9.89	12.79	8.34	13.63
Mean of DV for 45-49	10.02	12.59	16.04	10.33	16.76

Notes: See notes to Table 3.5. The estimates here include state linear time trends.

Appendix Table 3.4: Heterogeneity in the Impact of Early Access to the Pill on Marriage and Divorce Propensities

	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
	A. Never Been Married					B. Ever Been Divorced				
	Lower third of IC dist.	Middle third of IC dist.	Upper third of IC dist.	No College	Some College	Lower third of I(dist.	Middle third of IC dist.	Upper third of I(dist.	No College	Some College
ELA * Age 20-24	-0.120	0.020	0.060	-0.020	0.010	0.010	0.010	ND	0.000	-0.010
	(0.079)	(0.066)	(0.079)	(0.029)	(0.052)	(0.027)	(0.015)		(0.008)	(0.005)
ELA * Age 25-29	-0.030	0.020	0.030	-0.010	0.030	0.097*	0.085**	0.020	0.044*	0.020
	(0.060)	(0.045)	(0.059)	(0.021)	(0.0501	(0.057)	(0.035)	(0.031)	(0.024)	(0.018)
ELA * Age 30-34	-0.040	0.030	0.030	-0.010	0.030	0.070	0.070	0.020	0.020	0.020
	(0.050)	(0.033)	(0.047)	(0.020)	(0.035)	(0.066)	(0.047)	(0.039)	(0.030)	(0.025)
ELA * Age 35-39	-0.030	0.030	0.020	-0.020	0.020	0.030	0.070	0.020	0.000	0.010
	(0.044)	(0.029)	(0.044)	(0.017)	(0.031)	(0.074)	(0.051)	(0.019)	(0.030)	(0.032)
ELA * Age 40-44	-0.050	0.020	0.040	-0.023*	0.010	0.010	0.020	0.020	-0.030	0.010
	(0.044)	(0.026)	(0.045)	(0.015)	(0.028)	(0.078)	(0.055)	(0.047)	(0.032)	(0.034)
ELA * Age 45-49	-0.050	0.010	0.030	-0.030	0.010	0.010	0.030	0.000	-0.040	0.010
	(0.039)	(0.025)	(0.047)	(0.017)	(0.031)	(0.079)	(0.032)	(0.047)	(0.033)	(0.037)
Observations	12605	16698	20330	48548	26371	13540	18284	21575	54006	26439
Unique women	788	972	1112	2898	1456	776	966	1109	2895	1450
Pseudo R2	0.23	0.33	0.32	0.24	0.35	0.22	0.21	0.19	0.19	0.18
Mean of DV for 20-24	0.459	0.415	0.510	0.347	0.665	0.029	0.030	0.027	0.039	0.013
Mean of DV for 25-29	0.223	0.145	0.187	0.159	0.276	0.106	0.121	0.092	0.127	0.065
Mean of DV for 30-34	0.156	0.080	0.114	0.119	0.165	0.205	0.209	0.180	0.224	0.148
Mean of DV for 35-39	0.131	0.064	0.087	0.104	0.116	0.303	0.287	0.256	0.301	0.226
Mean of DV for 40-44	0.129	0.062	0.083	0.098	0.110	0.381	0.358	0.319	0.373	0.288
Mean of DV for 45-49	0.120	0.057	0.086	0.091	0.107	0.466	0.422	0.368	0.441	0.345

Notes: This tables presents mean marginal effects of equation (3.1') from a probit. Each column presents estimates from a separate regression on the indicated groups. "ND" indicates that disclosure requirements were not met for this estimate. All other notes are as in 3.4

3.9 References

Acemoglu, Daron, David H. Autor, and David Lyle. 2004. "Women, War and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid-Century." *Journal of Political Economy* 112(3): 497-551.

Ananat, Elizabeth Oltmans, and Daniel M. Hungerman. 2012. "The Power of the Pill for the Next Generation: Oral Contraception's Effects on Fertility, Abortion, and Maternal and Child Characteristics." *Review of Economics and Statistics* 94(1): 37-51.

Bailey, Martha J. 2006. "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply." *Quarterly Journal of Economics* 121(1): 289-320.

Belkin, Lisa. 2003. "The Opt-Out Revolution." *New York Times Magazine*. Sunday, October 26.

Black, Sandra E., and Chinhui Juhn. 2000. "The Rise of Female Professionals: Are Women Responding to Skill Demand?" *American Economic Review* 90(2): 450-455.

Black, Sandra E., and Alexandra Spitz-Oener. 2010. "Explaining Women's Success: Technological Change and the Skill Content of Women's Work." *Review of Economics and Statistics* 92(1): 187-194.

Blau, Francine D., Marianne A. Ferber, and Anne E. Winkler. 2010. *The Economics of Women, Men and Work*. New York: Prentice Hall.

Blau, Francine D., and Lawrence M. Kahn. 1997. "Swimming Upstream: Trends in the Gender Wage Differential in 1980s." *Journal of Labor Economics* 15(1): 1-42.

Blau, Francine D., and Lawrence M. Kahn. 2004. "The U.S. Gender Pay Gap in the 1990s: Slowing Convergence." NBER Working Paper 10853.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Variables." *Journal of Human Resources* 8(4): 436-455.

Bureau of Economic Analysis. 2009. National Income Product Accounts, Table 1.1.4: Personal Consumption Expenditures Price Index. Available at: http://www.bea.gov/national/nipaweb/SelectTable.asp?Selected= Accessed 8/17/2009.

Chiappori, Pierre-Andre, and Sonia Oreffice. 2008. "Birth Control and Female Empowerment: An Equilibrium Analysis." *Journal of Political Economy* 116(1): 113-140.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-1044.

Duncan, Otis. 1961. "A Socioeconomic Index for All Occupations." In ed. J. Reiss, Jr., *Occupations and Social Status*. New York: Free Press of Glencoe. 109138.

Fernandez, Raquel, and Alessandra Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1(1): 146-177.

Fernandez, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." *Quarterly Journal of Economics* 119(4): 1249-1299.

Firpo, Segio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77(3): 953-973.

Fortin, Nicole. 2009. "Gender Role Attitudes and Women's Labor Market Participation: Opting-Out, AIDS, and the Persistent Appeal of Housewifery." Mimeo.

Goldin, Claudia. 1990. *Understanding the Gender Gap: An Economic History of American Women.* New York: Oxford University Press.

Goldin, Claudia. 2004. "The Long Road to the Fast Track: Career and Family." *Annals of the American Academy of Political and Social Science* 596(1): 20-35.

Goldin, Claudia. 2006. "The Quiet Revolution that Transformed Women's Employment, Education, and Family." *American Economic Review* 96(2): 1-21.

Goldin, Claudia, and Lawrence Katz. 2002. "The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions." *Journal of Political Economy* 110(4): 730-770.

Goldin, Claudia, and Lawrence Katz. 2010. "Putting the Co in Education: Timing, Reasons, and Consequences of College Coeducation from 1835 to the Present." NBER Working Paper 16281.

Goldin, Claudia, Lawrence Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20(4): 133-156.

Griliches, Zvi, Bronwyn H. Hall, and Jerry A. Hausman. 1978. "Missing Data and Self-Selection in Large Panels." *Annales de L'Insee* 30/31: 137-176.

Heckman, James J., Lance J. Lochner, and Petra Todd. 2006. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." In eds. Eric A. Hanushek and Finis Welsh, *Handbook of the Economics of Education*, vol. 1. Amsterdam: Elsevier. 307-358.

Loughran, David S., and Julie M. Zissimopoulous. 2009. "Why Wait? The Effect of Marriage and Childbearing on the Wages of Men and Women." *Journal of Human Resources* 44(2): 326-49.

Miller, Amalia. 2011. "The Effects of Motherhood Delay on Career Path." *Journal of Population Economics* 24(3): 1071-1100.

Mulligan, Casey, and Yona Rubinstein. 2008. "Selection, Investment and Women's Relative Wages over Time." *Quarterly Journal of Economics* 123(3): 1061-1110.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14(3): 693-709.

O'Neill, June, and Solomon Polachek. 1993. "Why the Gender Gap in Wages Narrowed in the 1980s." *Journal of Labor Economics* 11(1): 205-228.

Paul, Eve, Harriet Pilpel, and Nancy Wechsler. 1974. "Pregnancy, Teenagers and the Law, 1974." *Family Planning Perspectives* 6(3): 142-147.

Paul, Eve, Harriet Pilpel, and Nancy Wechsler. 1976. "Pregnancy, Teenagers and the Law, 1976." *Family Planning Perspectives* 8(1): 16-21.

Shang, Qingyan, and Bruce Weinberg. 2009. "Opting for Families: Recent Trends in the Fertility of Highly Educated Women." Mimeo. November.

Stevenson, Betsey, and Justin Wolfers. 2007. “Marriage and Divorce: Changes and their Driving Forces.” *Journal of Economic Perspectives* 21(2): 27-52.

Weinberg, Bruce. 2000. “Computer Use and the Demand for Female Workers.” *Industrial and Labor Relations Review* 53(2): 290-308.

Weinberger, Catherine, and Peter Kuhn. 2010. “Changing Levels or Changing Slopes? The Narrowing of the U.S. Gender Earnings Gap, 1959-1999.” *Industrial and Labor Relations Review* 63(3): 384-406.

Welch, Finis. 2000. “Growth in Women’s Relative Wages and in Inequality among Men: One Phenomenon or Two?” *American Economic Review Papers and Proceedings* 90(2): 444-449.

Wellington, Allison. 1993. “Changes in the Male/Female Wage Gap, 1976-85.” *Journal of Human Resources* 28(2): 383-411.