Non-College Occupations and the Gender Gap in College Enrollment*

Amanda Chuan†

The Wharton School, University of Pennsylvania

November 1, 2017

Click here for latest version.

Abstract

Women used to lag behind but now exceed men in college enrollment. This paper shows that examining occupations which require only a high school degree ("non-college" occupations) can help resolve two puzzles related to this phenomenon. First, why do women attend college at greater rates than men today, when men work more and earn more than women? I document that non-college occupations for men are both more plentiful and higher paying than those for women. Next, I link the occupational inequality in the non-college labor market to the gap in college enrollment, by employing two empirical exercises to show that non-college jobs dramatically affect college-going decisions. Using employment changes in the oil and gas industry, I demonstrate that increases in men's non-college job opportunities lead male high school graduates to forego college enrollment. Using the automation of the office, I demonstrate that declines in the non-college employment opportunities of women lead female college enrollment to grow over time. Thus, the lower non-college job prospects of women contribute to women's higher college enrollment. This leads to the second puzzle: why did women historically attend college at lower rates than men, when women have always had worse non-college job prospects than men? I develop a theoretical model to demonstrate that both the importance and availability of non-college occupations for women contributed to women's initially low enrollment, as well as to the growth in female enrollment over time, such that women eventually overtook men in college-going.

Keywords: college enrollment, education and gender, human capital investment, occupational sorting
JEL Classifications: I23, I24, I26, J16, J22, J23, J24, N3

*I am deeply grateful to Iwan Barankay, Judd Kessler, Corinne Low, and Katherine Milkman for all their help, advice, and support. I thank Eduardo Azevedo, Emma Boswell, Kerwin Charles, Matthew Davis, Gilles Duranton, Ryan Fackler, Clayton Featherstone, Fernando Ferreira, Joseph Gyourko, Robert Jensen, Ben Keys, Ilyana Kuziemko, Olivia Mitchell, Karthik Nagarajan, Muriel Niederle, Alex Rees-Jones, Paul Sangrey, David Schindler, Katja Seim, Todd Sinai, Jeremy Tobacman, Petra Todd, Maisy Wong, Doug Webber, Karen Ye, Weilong Zhang, H2D2 Research workshop participants, WSAWBA conference participants, and Applied Economics student seminar participants at Wharton for excellent feedback. I am indebted to Daniel Keniston for generously providing the resource data on oil and gas production.

†The Wharton School, University of Pennsylvania, 3620 Locust Walk Suite 3000, Philadelphia, PA 19104. E-mail: achuan@wharton.upenn.edu

1 Introduction

Women are enrolling in college at greater rates than men, despite the fact that men have higher median earnings and higher labor force participation than women. This apparent contradiction has perplexed economists for decades.[1] I observe that for a high school graduate considering whether or not to attend college, the choice set appears dramatically different depending on gender. Men with only high school diplomas have viable, plentiful, and lucrative career prospects, especially given the plethora of blue-collar and trade occupations that pay highly based on physical strength, mechanical ability, or the willingness to face risky situations.[2] In 2015, jobs traditionally filled by men paid median incomes of $52,000 (truck driver), $53,000 (electrician), or $60,000 (police officer).[3] In contrast, the jobs typically held by women with only high school degrees are much lower paying. For example, jobs traditionally filled by women paid median incomes of $20,000 (cashier), $22,000 (housekeeper), and $24,000 (hairdresser).[4] If these occupational differences are broadly representative of the job prospects of men and women without a college degree, then it would be natural for women to enroll in college at greater rates than men.

The imbalance in occupations among workers with only high school degrees (hereafter, “non-college” workers) is an under-explored and overlooked reason for the greater college enrollment of women observed today. Moreover, this disparity contributed to the trends in the college gender gap over time, wherein women used to lag behind but now exceed men in college-going. This insight adds to the discussion on human capital investments by pointing out that: (1) women’s supposed “over-investment” in college is not an over-investment at all, given the few alternative options women have; and, similarly, (2) men’s comparative “under-investment” in college may not only arise from their greater barriers to human capital investment,[5] but also their more lucrative

[1] See Dougherty, 2005; Buchman and DiPrete, 2005; Goldin, Katz, and Kuziemko, 2006; Becker, Hubbard, and Murphy, 2010.

[2] Men have been shown to possess greater physical strength, mechanical ability, and tolerance for risk than women (see Miller et al., 1993; Blakemore et al., 2009; Croson and Gneezy, 2009).

[3] Bureau of Labor Statistics Occupational Employment and Wages, 2015.

[4] Ibid.

[5] Prior research has shown that men tend to be more impulsive, more myopic, and less risk averse than women, which may contribute to the lower high school graduation rates among men relative to women (see Bertrand and Pan, 2013; Becker, Hubbard, and Murphy 2010; Goldin, Kuziemko, and Katz 2006).

outside options when making the college-going decision.

This paper proceeds in two parts. The first part argues that women's bleak outside options make it natural for them to exceed men in college-going. In Section 2, I document evidence of the dramatic disadvantages faced by women with only high school diplomas. In Sections 3 and 4, I use two empirical exercises to show that these disadvantages directly contribute to the gender gap in college enrollment. In both exercises, I show that when the non-college occupations of one gender are disproportionately affected, there is a large gender difference in the college enrollment response, and therefore a significant change in the gap in college enrollment. Together, these results imply that the non-college labor market plays a large role in explaining why women attend college at greater rates than men today.

The second part of the paper addresses a related puzzle in the literature — why women first lagged behind and then exceeded men in college-going over time, when their non-college job prospects have always been worse than men's. In Section 4, I demonstrate that women's non-college employment opportunities dramatically declined over time, while men's non-college job options remained plentiful by comparison. In Section 5, I situate this dynamic in a theoretical model to explain how declining non-college job options for women and increasing female labor force participation complemented each other in contributing to the growth in female enrollment over time, such that female enrollment eventually surpassed male enrollment. Finally, I use the model to estimate the extent to which the change in college enrollment from 1970 to 2010 can be attributed to changes in non-college job options.

In particular, Section 2 documents stylized facts regarding the large disparity in non-college occupations facing male and female high school graduates. Using decennial census microdata, I document a "missing quadrant" of high paying non-college occupations for women. The majority of non-college occupations are male-dominated, while the few non-college occupations that employ women tend to exhibit low median earnings. I calculate the premium to college-going for men and women by constructing a weighted median of annual earnings, using the proportion of workers in each occupation as weights. I find that this premium is consistently higher for women than men by at least 30 log points from 1950 to 2010. Non-college women face even larger disadvantages when it comes to careers, as opposed to just jobs. Over the life cycle, non-college men earn roughly the same as college women by age, but non-college women make far less, experience little earnings growth over their careers, and are less likely to work in occupations that offer benefits. Overall, the job prospects of male high school graduates appear much more plentiful, higher paying, and more likely to be careers than the prospects of female high school graduates.

Does the imbalance in the non-college labor market translate to the gap in college enrollment? If so, do changes in non-college jobs shift enrollment rates for women, men, or both? Sections 3 and 4 address this question for men and women, respectively. Both sections show that shocks to specific occupations and industries change the non-college job prospects of women relative to men, and correspondingly change the college enrollment gap.

Section 3 uses employment changes in the oil and gas industry to demonstrate that increases in the non-college employment opportunities of men in this industry lead men to forego attending college. Jobs in the oil and gas industry (e.g., oil field worker or driller) are dominated by men, and employment changes in this industry have a larger impact on male employment than female employment. Using oil and gas production data from Allcott and Keniston (2016), I find that natural variation in oil and gas reserves predicts the capacity of different counties to increase or decrease employment for oil and gas workers. Exploiting this variation, I estimate that an additional 10% increase in oil and gas employment leads to an additional 1.4 percentage point reduction in college enrollment among male high school graduates. This estimate is economically and significantly greater than the estimated response of female college enrollment, which is effectively zero.

Section 4 demonstrates that automation led to dramatic declines in the non-college employment opportunities of young women, which led female college enrollment to grow over time. I build on the routine-biased technical change literature, which reports that automation displaced routine-intensive occupations and drastically changed the structure of the non-college labor market (see Autor, Levy, and Murnane, 2003). I present new evidence that routine-intensive occupations employed over 60% of the young non-college female work force in 1970, and demonstrate that women's non-college jobs were especially vulnerable to the displacing effect of automation. Guided by this finding, I then use a shift-share instrument that predicts exposure to automation to isolate the causal effect on college enrollment. I demonstrate that an additional percentage point decline in routine-intensive jobs led to a 0.7 percentage point increase in the female college enrollment rate, which was significantly greater than the effect on male enrollment. This empirical exercise illustrates that like men, women respond dramatically to their non-college employment opportunities, suggesting that the anemic non-college prospects women face today contribute to their greater college enrollment. Section 4 also highlights that the decline in routine-intensive jobs over the last 40 years has contributed to women's increase in college-going over that time.

Finally, Section 5 situates these findings in a theoretical model to simultaneously explain two puzzles: 1) why women attend college at greater rates than men now, even though men have earned more and worked more than women for most of history;[6] and 2) why women historically lagged behind men in college-going when their observed college premium was always higher than men's. The model illustrates that men's higher earnings and greater labor force attachment explain why male college enrollment shot up quickly and leveled off quickly. In contrast, the labor force participation of married women was initially low but grew substantially starting in the 1970s, making labor market outcomes more important for women *just as automation began to displace the bulk of their non-college job options*. The decline in women's non-college job prospects and the growth in female labor force participation were complementary in enabling women to realize their larger premium from schooling relative to men and propelling female college enrollment to surpass male college enrollment. A back-of-the-envelope calculation estimates that changes in non-college jobs account for 40% of the growth in female college enrollment and 28% of the change in male college enrollment between 1970 and 2010.

This paper addresses an old but open question, but takes a distinct approach from most of the related literature. Rather than discuss the ability of women to outperform men academi-

[6]A long literature on the gender gap in wages has shown that earnings for men tend to be higher than earnings for women, and that male labor force participation has been greater than female labor force participation (see Blau and Kahn, 2016 for a review).

cally (see Buchmann and Diprete, 2005; Jacob, 2002; Goldin, Katz, and Kuziemko, 2006; Becker, Hubbard, and Murphy 2010; Bertrand and Pan, 2013), or explore marriage market returns to attending college (see Chiappori, Iyigun, and Weiss, 2009; Chiappori, Salanie, and Weiss, 2015; Chiappori, Costas Dias, Meghir, 2015; Bronson, 2015; Zhang, 2016; Low, 2017), I focus on labor market returns. The majority of the literature on labor market returns and the college gender gap focuses on college jobs and uses structural models, Oaxaca decompositions, or panel data with the hope of isolating causal relationships (see Jacob, 2002; Dougherty, 2005; Charles and Luoh, 2003; Olivieri, 2015). In contrast, my paper uses exogenous variation in labor demand to test the hypothesis that the superior non-college job options of men leads to greater demand for college degrees among women. I show that women's enrollment rates increase when their non-college employment opportunities become scarce, and thus that their deteriorating non-college job options drove their college enrollment to grow and surpass that of men. I show that men's enrollment rates decline when their non-college labor market outcomes improve, and thus that their comparatively more plentiful non-college job options led a larger proportion of men than women to rationally forego attending college.

I make four contributions to the literature. To my knowledge, this paper is the first to connect gender differences in non-college job prospects to 1) the greater demand for a college degree observed today and 2) the trajectory of the college gender gap over time. The (few) other papers in this vein had other objectives, and therefore either do not show that changing non-college job opportunities lead to gender gaps in college enrollment (Charles, Hurst, and Notowidigdo, 2016), or do not show that individuals qualified to attend college actively forego college enrollment in the presence of more attractive outside options (Cascio and Narayan, 2015).[7]

Second, I leverage the task-based approach to measure occupational skill demands (Autor,

[7] Cascio and Narayan (2015) find that fracking increased high school drop-out rates among boys, but the focus of their paper is not to address gender differences in the *choice* to attend college. The mechanism for their findings may operate along dimensions other than choice. For example, if part-time jobs working in the oil and gas industry become more available for boys, boys may find it harder to balance a job with high school coursework, and therefore fail to complete high school even if they wished to graduate and attend college. In contrast, my paper finds that even among individuals qualified to attend college, men choose to forego college given an increase in oil and gas employment in their area.

Levy, and Murnane, 2003; Autor and Dorn, 2013), which provides more granular measures of the labor market returns to skill profiles than the conventional approach of examining wage gaps (see Goldin and Katz, 2008). Using this approach reveals that the returns to skills performed by non-college women declined relative to the skills performed by non-college men. I thus provide empirical facts that invite revisions of prior models, which overlook the role of declining non-college jobs for women in increasing the college premium for women (Welch, 2000; Rendall, 2010; Huang, 2014).

Third, I contribute to the literature on routine biased technical change. I demonstrate that automation propelled women to enter college at greater rates than before, by displacing their non-college employment opportunities. In contrast to the prior literature, which has mostly focused on how automation led to job market polarization (Autor, Levy, and Murnane 2003; Goos, Manning, and Salomans, 2009; Autor and Dorn, 2013; Goos, Manning and Salomans, 2014), I show that automation also affected the human capital investment decisions of women in irreversible ways.

Fourth, I present a simple model that resolves two contradictions. The first contradiction is that women attend college at greater rates than men, yet men have greater earnings and stronger labor force attachment (see Dougherty, 2005). The second contradiction is why men used to attend college at greater rates than women when outside options have always been worse for women.[8] My model demonstrates that men's greater earnings and labor force participation led them to attend college at higher rates than women at first, but that women's growing labor force participation allowed them to realize their greater labor market returns, which pushed women to eventually surpass men in college-going.

My results have direct implications for policy and future research. Policymakers have become increasingly concerned that men are lagging behind women in educational attainment (Rosin, 2015). Several countries have already implemented interventions intended to help men catch up, such as hiring more male teachers to serve as role models for boys, and tailoring class curricula to

[8]Becker, Hubbard, and Murphy (2010) identify a related mystery. For them, the true contradiction was why men surpassed women in college-going at first when a greater proportion of women were academically prepared to attend college relative to men.

appeal to boys (Rosin, 2015; The Economist, 2015).[9] The results of my paper suggest that these actions may be misguided. If men attend college at lower rates because they possess better outside options, then the gap in college enrollment may not be as inefficient as it seems, and interventions to minimize this gap may be ineffective at best and destructive at worst. Indeed, recent evidence from Carrell and Sacerdote (2013) indicates that interventions to encourage college-going have not shown much promise for men, and in their survey evidence, men cite larger expected earnings with only a high school degree as one key reason for choosing to not attend college. The aforementioned policy measures could even decrease welfare, for example if male teachers were hired at the expense of more qualified female teachers or if classroom curricula were changed to interest boys but ended up alienating girls. This paper suggests that before we devote public resources to eliminating educational differences between men and women, we should first re-examine why these differences exist in the first place, in order to determine the best role of policy in addressing the college gender gap.

2 Background and Stylized Facts

This section presents background information regarding the gender disparity in non-college jobs using raw data from the 2010 American Community Survey. Within the non-college labor market, men and women sort into different occupations, and the earnings of traditionally female occupations are far lower than the earnings of traditionally male occupations. Based on these differences alone, the observed college premium is dramatically greater for women than men. Women's disadvantages in the non-college labor market materialize not only in the form of lower annual earnings, but also worse career prospects in terms of lifetime earnings, earnings growth, and access to benefits. The combination of these stylized facts suggests that the stark imbalance in non-college job prospects makes it natural for women to enroll in college at higher rates than men.

Stylized Fact 1. *In the non-college labor market, there exists a "missing quadrant" of high paying jobs for women.*

[9]For example, Britain recently began a campaign to make reading more appealing to boys (Sommers, 2013).

Figure 2 depicts the median annual earnings percentile and worker gender composition for each non-college occupation in 2010.[10] Each data point is an occupation as defined by the 1990 Census Bureau occupational classification scheme. To capture labor market returns for individuals most likely to consider the college-going decision, I restrict the data to only 18-30 year olds.

The figure makes two important points. First, the majority of non-college occupations are male-dominated. In fact, almost 65% of all non-college occupations employ 20% or fewer women. Many of these occupations were trade or blue-collar occupations, which either paid highly for work that demanded physical strength or mechanical ability, or paid highly for work that was unpleasant. For example, table 1 shows that miners, machinists, and truck drivers reported median annual earnings that were between the 40th and the 80th percentile of median earnings for all occupations in 2010.[11] Second, male-dominated occupations pay more than occupations that employed female workers. The occupations that employ a non-trivial share of women have significantly lower median annual earnings than occupations that consisted of over 80% men. As shown in table 1, common jobs among non-college women include cashier, cosmetologist, or housekeeper, where annual median earnings fell below the 10th percentile of median earnings for all occupations.

These two points indicate that there exists a "missing quadrant" of high paying jobs for women. There exists both low and high paying jobs for men, but only low paying jobs for women. Figure 3 shows that there is a mirroring "missing quadrant" of low paying college jobs for men. This missing quadrant of low paying college jobs for men is to be expected, since no men would enter college to earn a low wage if there existed high paying alternatives which did not require a college degree. Together, figures 2 and 3 demonstrate that men typically sort into high paying occupations in the noncollege market, whereas women do not, which may then change the composition of men and women who elect to attend college.

The evidence indicates that gender differences in the allocation of workers to occupations lead non-college women to have much lower earnings than non-college men. This point, combined

[10] I define "non-college occupations" as occupations where over 50% of workers have never enrolled in college.

[11] Men could have had an easier time finding employment in these occupations, could have been more willing to work in these occupations given their high compensating wages, or some combination of both. In any case, the vast majority of non-college occupations employed very few women.

with the fact that occupational gender differences were much smaller in the college labor market, leads to a higher college premium among women than men. The next figure explores this in greater detail.

Stylized Fact 2. *Women have a higher observed college premium than men.*

The literature has typically used the difference in median log annual earnings between high school and college graduates as an approximate measure of the college premium in the labor market (see Goldin and Katz, 2008; Acemoglu and Autor, 2011). To determine how gender differences in occupations create different college premia between men and women, I weight median annual earnings by the share of workers in each occupation.[12] As a result, the only differences in college premia between men and women arise from differences in gender composition within occupations.

The results, shown in figure 4, are striking. The observed college premium in earnings for 18-30 year old workers is much higher for women than for men for all decades from 1950 to 2010. This difference was approximately 30 log points in 1950, rose to 50 log points by 1970 and 1980 right before women began to surpass men in college-going, and diminished to a little less than 40 log points by 2010 when the gap in college enrollment began to finally stop growing.[13] Appendix figure A.1 breaks down this difference by plotting the weighted median log wages by sex and education type, which reveals that the greater college premium of women is driven primarily by their much lower non-college median log wages. Within the college labor market, the gender difference in median log wages is much smaller, since gender-based segregation across occupations is not as pronounced as it is among non-college occupations. Thus, the occupational gender gap among non-college workers appears to create a large potential incentive for women to enroll in college at greater rates than men.

Stylized Fact 3. *Women face worse career prospects relative to men in the non-college labor market, in that non-college occupations that employ women tend to have lower earnings growth, are much less likely to offer retirement pensions, and are slightly less likely to offer health insurance.*

[12] The notes under figure A.1 describe how median annual earnings are computed in greater detail.

[13] This analysis is complementary to the results shown in Charles and Luoh (2003), which finds a higher college earnings premium for women than men using log median earnings instead of log median earnings weighted by occupation.

Attending college may change the workers ability to pursue a *career*, as opposed to just a job. Based on the raw data, the non-college market is more favorable to the career pursuits of men than women. I show that "traditionally male" non-college occupations exhibit characteristics that enable their workers to support a family and remain committed to the same occupation in the long-term. In particular, "traditionally male" non-college occupations exhibit earnings growth and the provision of benefits, such as retirement pensions and employer-sponsored health insurance. "Traditionally female" non-college occupations, on the other hand, have virtually no earnings growth, and are substantially less likely to provide retirement pensions. For women, having a well-paying career and a college degree go hand-in-hand, while men can excel in careers without a college degree.

To approximate how earnings grow with age over an individual's work life, I use the National Longitudinal Study of Youth 1979 cohort, and focus on the subsample of men and women with at least a high school diploma. I then split individuals by gender and college degree status to determine how calculated hourly wage rates differ between non-college and college workers.

Figure 5 presents the median log hourly wage rate by age for workers are between the ages of 25 and 55. Women's non-college career prospects appear much worse than men's, since the occupations non-college women tend to enter pay very little and exhibit low earnings growth. In particular, the figure demonstrates that 1) the median earnings of female non-college occupations are far lower at each age level than the median earnings for other occupations; and 2) the median earnings for male non-college occupations are approximately equivalent to the median earnings of female non-college occupations at each year of age. Figure 5 demonstrates that the large disadvantages non-college women experience in annual earnings relative to non-college men, shown in figure 4, become exacerbated when considering how these disadvantages multiply over the work life.

Careers tend to exhibit not only earnings growth but also the provision of benefits, such as retirement pensions and health insurance. The last two columns of table 1 summarize information about benefits taken from the 2010 Annual Social and Economic Supplement of the Current

Population Survey (CPS-ASEC).[14] Traditionally female non-college occupations are far less likely to have access to retirement pensions with their work, and are slightly less likely to be included in an employer group health plan. Among workers in traditionally male occupations like trucker driver, miner, and machinist, 46-62% reported that their job offered retirement pensions, compared to only 11-36% of workers in female-dominated occupations like cashier, cosmetologist, or housekeeper. Among these examples, over 95% of workers in traditionally male occupations reported being included in an employer group health plan, compared to less than 92% for workers in female-dominated occupations.

3 Male College Enrollment and the Oil and Gas Industry

Are non-college job options important in making the decision to attend college, or are other factors like innate ability, preparedness, and financial means sufficient to determine who goes to college and who does not? Theoretically, non-college jobs influence the returns to attending college and the opportunity cost of schooling. On the other hand, if empirically the earnings of college graduates were much higher than the earnings any high school graduate can expect to make, or if the social value of graduating from college was sufficiently high, then ability and means may be sufficient to explain most of the variation in college enrollment among high school graduates. Under these circumstances, any changes in non-college job opportunities may do little to actually shift college enrollment.

Therefore, the first question to address is whether changes in non-college jobs empirically lead to significant shifts in college enrollment, and whether men and women are differentially affected. I leverage the descriptive evidence in the previous section, which shows the large degree

[14] The American Community Survey (ACS) is large enough to estimate reliable summary statistics regarding work and earnings within occupations. Its measures regarding retirement income and health insurance, however, are too general for the purposes of this paper. Its retirement income questions ask about whether the respondent has income from retirement, survivorship, or disability benefits broadly. The questions in the ACS regarding insurance simply ask if the respondent is on employer-sponsored health insurance – the policyholder may be the respondent, the spouse, or another family member. In contrast, the CPS-ASEC Supplement asks individuals whether they have retirement income as a result of their employment, separate from survivorship payments, disability benefits, Social Security income, Veterans administration payments or other forms of income. The CPS-ASEC also asks individuals if they are the policyholder for their employer-sponsored health insurance.

of gender segregation in the non-college labor market. This section uses employment in the oil and gas industry, a male-dominated field, to isolate the causal effect of male non-college employment on college-going.

Oil and gas production has substantial effects on local labor markets (see Bartik et al., 2017; Feyrer, Mansur, and Sacerdote, 2016; Allcott and Keniston, 2015; Cascio and Narayan, 2015), particularly for non-college work. For example, in 2006, breakthroughs in hydraulic fracturing and horizontal drilling enabled unprecedented quantities of oil and gas production in North Dakota (NPR, 2011; Brown, 2013). Oil production catapulted from 40 million barrels to 150 million barrels within the span of five years from 2006 to 2011, which created sudden and enormous changes in the labor market returns to work in the oil and gas industry.[15] By some estimates, the oil boom created 35,000 new jobs in 2011, which is enormous for a state with a population of 670,000 (McChesney, 2011). Unemployment in North Dakota fell to 3.3% in 2012, the lowest in the entire United States. Wages for non-college work also saw drastic growth: average salaries for oilfield workers rose to $70,000-$100,000, and truckers routinely made $70,000-$80,000 a year (Gold, 2015).

The example of North Dakota illustrates the ramifications of oil and gas production on the labor market. To explore whether this influenced the demand for a college education across the entire United States over last few decades, I use fluctuations in oil and gas production from the contiguous United States from 1970 to 2010.

Upticks in oil and gas production increase the employment demand for not only workers directly involved in oil and gas production (e.g., oil-well drillers, miners, drillers of earth), but also other workers in related fields. Truck drivers, shippers, material handlers, material movers, and haulers are required to transport oil to refineries; welders, electricians, mechanics, installation technicians, and millwrights are required to build and maintain the equipment required to facilitate production; structural metal workers, construction workers, concrete pourers, and foremen are required to build residential and commercial properties. My analysis considers occupations directly

[15] This information is obtained from oil production data provided by Allcott and Keniston (2015), which is described in detail later.

employed in the oil and gas industry, as well as "related" industries where employment demand is positively correlated with oil and gas employment.

Work in the oil and gas industry is especially dangerous and requires intensive physical labor. The industry is considered one of the most dangerous in America, and the workplace death rate in North Dakota had grown to five times the national average since the oil boom began (Berzon, 2015). It is perhaps for these reasons that employment opportunities in the oil and gas industry have historically attracted overwhelmingly male, blue-collar workers (Eligon, 2013). Figure 6 graphs the composition of workers in the oil and gas industry by sex and education group. Among workers with at least a high school diploma, men comprise the majority of the workforce in the oil and gas industry, while college and non-college women each constituted less than 10% of the entire workforce. Male high school graduates comprised of 50-70% of the workforce in occupations with high labor shares in the oil and gas industry, such as truck, delivery, or tractor driver, laborers outside construction, or miners.

To identify a causal channel between non-college labor market outcomes and college enrollment, I exploit the fact that oil and gas production depends on the geology of the earth. There is a great deal of geographic heterogeneity in natural reserves, which influences the sites of active oil and gas production. When demand for oil and natural gas is high, areas rich in natural reserves are able to dramatically increase employment, as demonstrated by the example of the North Dakota boom. However, areas poor in natural reserves show little change in employment over time.

Figure 7 demonstrates that the geology of the earth determines oil and gas employment, by depicting employment in the oil and gas industry over time in "high-resource" (states with above median natural resources) and "low-resource" states.[16] There exists a great deal of fluctuation in male employment in high-resource states, but male employment in low-resource states remains relatively constant. Female employment in both high- and low-resource states also remains con-

[16] States with above median natural resource endowments are Alabama, Arkansas, California, Colorado, Florida, Illinois, Kansas, Kentucky, Louisiana, Michigan, Mississippi, Montana, North Dakota, New Mexico, Ohio, Oklahoma, Texas, Utah, West Virginia, and Wyoming. States with below median natural resource endowments are Arizona, Connecticut, Delaware, Georgia, Iowa, Idaho, Indiana, Massachusetts, Maryland, Maine, Minnesota, Missouri, North Carolina, Nebraska, New Hampshire, New Jersey, Nevada, New York, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Virginia, Vermont, Washington, and Wisconsin.

stant over time, since the labor share of employment in oil and gas was very low (less than 1%) for women. The figure illustrates that states with rich natural resources are able to expand or diminish employment in the oil and gas industry, while states with relatively poorer natural resources cannot. This influences the labor market outcomes of men far more than women, since jobs in the oil and gas industry comprise an extremely low share of female employment.

Using the geological variation in natural oil and gas reserves, I find that a 10% increase in mens non-college employment opportunities decreases male college enrollment by 1.4 percentage points. The male college enrollment response is significantly greater than the female college enrollment response, which is insignificant. This effect is strongest for individuals closest to the margin of college-going, and persists even after accounting for migration.

3.1 Data and Summary Statistics

Data on education, occupation, earnings, work, and demographic characteristics come from the Annual Social and Economic Supplement of the Current Population Surveys (CPS-ASEC), which are conducted every year jointly by the Bureau of Labor Statistics and the U.S. Census Bureau, and provided by the Integrated Public Use Microdata Series (Flood et al., 2015). The CPS-ASEC contains rich information regarding the occupations and industries in which each respondent worked, as well as detailed information regarding their earnings, hours and weeks worked, employment history, and schooling. Moreover, the CPS-ASEC contains rich data of migration patterns, which is especially useful when employing analysis that exploits spatial and time trends across labor markets.[17]

County-level data on natural reserves and oil and gas production were generously provided by Hunt Allcott and Daniel Keniston. Allcott and Keniston (2015) compile a unique data set of resource endowments at the county level in the contiguous United States from 1962-2012 using information from DrillingInfo (a market research company), the United States Energy Information Administration (EIA), and local reports and geological surveys. I only use data for the years 1970-

[17]For more detail regarding the samples used in the analysis, see Data Appendix C.

2010, since many of the earlier years contain missing data. Allcott and Keniston (2015) calculate the natural oil and gas reserve endowment per square area using the equation

$$r_c = \frac{\sum_{t=1960}^{T} \text{Production}_{ct} + \text{Proven Reserves}_{ct} + \text{Undiscovered Reserves}_{ct}}{\text{Area}_c}$$

Production_{ct} represents the production of oil or gas in year t in county c; $\text{Proven Reserves}_{ct}$ represent the reserves that oil and gas producers know to exist with relative certainty; $\text{Undiscovered Reserves}_{ct}$ are resources which oil and gas producers believe to exist due to the type of hydrocarbons found in the earth, but have not yet determined to exist with certainty. Since r_c is specific to the natural geographic composition of the earth, it should be exogenous to changes in labor demand in the oil and gas industry. The next subsection discusses the identification strategy in greater detail.

Appendix tables A.2-A.6 compare the observable characteristics in states with high resource endowments to states with low resource endowments for each year from 1970 to 2010.[18] Overall, resource endowments do not appear to determine significant differences across states: female college enrollment, male college enrollment, the proportion of women in a state, the proportion of blacks in a state, and the proportion by age bin do not differ systematically or significantly by the state's level of natural resources. In a few of the years, the proportion of individuals by different age bins are significantly different between high- and low-resource states - for example, the proportion of individuals between the ages of 18 to 25 differs significantly between high- and low-resource states in 2000, but for all other years, this relationship is insignificant.

This section examines the effect of employment in the oil and gas industry, as well as employment in the oil, gas, and "related" industries. I classify workers as being employed in the oil and gas industry if they worked in oil and gas extraction, petroleum refining, petroleum production, mining, trucking, or warehousing and storage. For "related" industries, I add workers who were not explicitly employed in the aforementioned categories but had skills that were transferable to the work commonly performed in the oil and gas industry, such as construction workers, material

[18] Only the years 1970, 1980, 1990, 2000, and 2010 are shown for brevity, but all other years are available upon request.

handlers, geologists, miners, excavation operators, drillers of earth, operators of machinery, and petroleum engineers. In the results, I separately display my regressions for employment in the oil and gas industry and in the oil, gas, and "related" industry. The results are comparable for both groups, although the effect sizes on college enrollment are understandably larger when workers in related industries and occupations are included.

Appendix table A.1 calculates the growth in employment share by gender and college status following national growth in oil and gas employment. I use two periods of time which experienced the most marked increase in national oil and gas employment: 1970-1980, and 2000-2010. The table demonstrates that the change in employment share is largest for non-college men by an order of magnitude for both employment in the oil and gas industry and employment in the oil, gas, and "related" industries. For college men, college women, and non-college women, the change in employment in the oil and gas industry or the oil, gas, and "related" industries as a proportion of total labor share is very small. Table A.1 provides further support for the evidence in figure 6 that changes in oil and gas employment affect the non-college labor prospects of men the most.

3.2 Instrumental Variable Strategy

Using the resource endowment measure r_c, Allcott and Keniston (2016) construct a modified shift-share instrumental variable that interacts county-level resource endowments with time-varying national employment in the oil and gas industry. I use their instrument in my estimation procedure, but my analysis aggregates resource endowments to the state level. I compute the instrument

$$\widehat{E}_{st} = r_s E_t \tag{3.1}$$

where E_t represents national employment in the oil and gas industry in year t. and $r_s = \sum_{c \in s} r_c$ represents the natural resource endowment in state s.

In the first stage, I regress natural log employment in the oil and gas industry in state s in year t on the instrument, a vector of state-level controls, and state fixed effects. The controls include percent female, percent black, and percent by ten-year age bin.

$$\ln(E_{st}) = \alpha_0 + \alpha_1 \ln(\widehat{E}_{st}) + \alpha_2 X_{st} + \theta_s + u_{st} \tag{3.2}$$

From the first stage, I obtain the linear prediction in employment demand for the oil and gas industry, $\widetilde{\ln E_{st}}$, which should be exogenous to any state-specific characteristics that would be correlated with oil and gas employment demand and college enrollment rates. Using the results from the first stage, I then estimate the effect of changes in employment demand on the male college enrollment rate, the female college enrollment rate, and the college gender gap (the male college enrollment rate - the female college enrollment rate). The second stage regression is

$$Y_{st} = \beta_0 + \beta_1 \widetilde{\ln(E_{st})} + \beta_3 X_{st} + \theta_s + \varepsilon_{st} \tag{3.3}$$

This instrument does a fairly accurate job of predicting actual state-level employment in the oil and gas industry. Table 2 presents the first-stage regression of the shift-share instrument on employment in the oil and gas industry (1) and employment in the oil, gas, and related industries (2). State-level extraction of oil, state-level extraction of gas, percent female, percent black, percent by age group, and state fixed effects are included as controls. Standard errors are clustered at the state level. All first-stage F-statistics exceed 10 and pass the Anderson-Rubin test of weak instruments. As expected, the instrument is stronger in predicting actual employment for workers in the oil and gas industry than workers in the oil, gas, and "related" industries, although the correlation between the instrument and employment is positive and significant in both cases. For each percentage point increase in the shift-share instrument, actual oil and gas employment increases by 0.736 percentage points and employment in the oil, gas, and "related" industries increases by roughly half that amount, at 0.380 percentage points ($p < 0.01$ in both cases).

Table 3 presents the results of a simple OLS of college enrollment on the instrument. The coefficients can be interpreted as the effect of oil and gas resource endowment on college enrollment. I control for oil production, gas production, proportion female, proportion black, proportion by age group, and state fixed effects to account for variation in college enrollment brought about by the demographic composition of the state. All regressions are clustered at the state level and robust against heteroskedasticity. An increase in the oil and gas resource endowment by 10% leads to a decline in male enrollment by 1.2 percentage points, while the effect on female enrollment is close to zero and insignificant. The net result on the college gender gap, defined as male enrollment less female enrollment, is a decline of 1.5 percentage points. The specification is the reduced form version of the 2SLS specification discussed next in the results subsection. The coefficient estimates reported in both the reduced form and the 2SLS specifications are very similar, since the effect of natural resources on oil and gas employment is close to 1, and the 2SLS coefficient estimate is the reduced form coefficient divided by the coefficient in the first stage.

The identification assumption in the 2SLS specification is that state-level oil and gas reserves are uncorrelated with the unobserved characteristics that influence both college enrollment and oil and gas employment. During upticks in oil and gas employment, states with rich natural resources increase their demand for oil and gas workers, and the effect of the geology of the earth on the demand increase in oil and gas workers should be uncorrelated with other factors that influence college enrollment. This assumption would be violated if increasing employment demand changed the composition of workers in a state by attracting migrants. I therefore run additional regressions using subsamples of only individuals who did not migrate across states in the last year or who did not move for work purposes.

3.3 Results

All regressions are conducted at the state-year level. All regressions control for demographic characteristics at the state level (proportion female, proportion black, proportion by age bin, oil production, gas production, base-year female enrollment, base-year male enrollment, and base-year employment in the oil and gas industry) and state fixed effects. Regressions are clustered at the state level and robust against heteroskedastic error terms.

My preferred specification is the instrumental variable regression. Table 4 uses the shift-share instrument to estimate how employment in the oil and gas industry influences the female college enrollment rate (column 1), the male college enrollment rate (column 2), and the gender gap in college enrollment (column 3). As predicted, changes in employment in the oil and gas industry do not affect female college enrollment: the coefficient estimate on female enrollment is positive, but very close to zero. In contrast, a 10% increase in oil and gas employment decreases male college enrollment by 1.2 percentage points and leads the college gender gap to decline by a corresponding 1.5 percentage points.

These results are robust even when accounting for migration. I replicate this analysis with the subsample of individuals who reported staying in the same state as the year before, and with the subsample of individuals who did not report moving for work-related reasons. These results are listed in tables 5 and 6, respectively.

Table 7 estimates how college enrollment rates respond to changes in employment opportunities in oil, gas, and related industries, where workers in industries positively affected by oil and gas booms (e.g., construction or trucking) are included. Overall, I find similar effects as in the previous regressions. The college enrollment rate for women is not significantly affected by fluctuations in employment demand. For men, however, the effect is even larger than before: a 10% increase in employment demand in oil, gas, and related industries leads to a 2.7 percentage point decline in the male college enrollment rate and a corresponding 3.1 percentage point decline in the college gender gap. The larger coefficients are unsurprising, given that oil and gas booms increase employment demand in not only the oil and gas industry but in many related industries which also happen to be male-dominated. If the independent variable incorporates variation in a larger share of the non-college male force, then one would expect the effect on male college enrollment rates to be at least as large as the effect shown in the prior regressions.

4 Female College Enrollment and Automation

Section 3 presents evidence that men's college-going decisions are extremely responsive to changes in non-college employment opportunities. Does this same effect hold for women? It may be the case that women attend college for different reasons than men (e.g., the greater likelihood of marrying a high earning spouse), so shifts in non-college employment opportunities may do little to shift college enrollment rates for women. This is important in determining how important non-college jobs are in contributing to the gender gap in college enrollment: do non-college jobs only affect the college gender gap by limiting the proportion of men who select into college-going? Or do non-college employment opportunities contribute to the growth in college enrollment for women as well?

I investigate this question by using the case of automation. Examining the correlation between college enrollment and non-college job opportunities alone is insufficient to isolate causal effects of non-college employment on college enrollment, since areas with high proportions of female college-goers will mechanically have lower shares of female workers in non-college jobs. Instead, I use an instrument for predicted automation exposure to show that a decline in non-college jobs for women, brought about by the automation of the office, leads to an increase in female college enrollment.

Automation led to dramatic changes in labor market (see Autor, Levy, and Murnane, 2003; Goos, Manning, and Salomans, 2009; Goos, Manning, and Salomans, 2014; Autor and Dorn, 2013; Autor and Acemoglu, 2011), particularly for women (Autor and Wasserman, 2013; Black and Spitz-Oener, 2010). I show that the continuous automation of the office decreased women's non-college job prospects and induced them to enter college at greater rates. This analysis complements the results from Section 3 by demonstrating that like male enrollment, female enrollment responds to changes in their non-college opportunities. Combining this result with the stylized facts in section 2 implies that the anemic options for women in today's non-college labor market are a key reason behind the greater proportion of women than men on college campuses today.

A secondary finding in this section is that automation led to historical growth in female college enrollment over time. Jobs that were displaced by automation, such as secretarial work, clerical work, telephone operators, and typists, employed the majority of non-college working women in 1970. From 1970 to 2010, the labor share of secretaries declined by 30%, while the labor share of typists declined by 86%. These large changes in key occupations for non-college female labor transformed the labor market, such that the labor market alternatives to college-going for women became increasingly scarce over time.

The adoption of automated systems by firms was an ongoing process throughout the 20th and 21st century. Automation significantly changed the content of jobs, by changing the marginal productivity of machines relative to that of human labor at certain tasks. To measure this change, the literature on automation focuses on the “routine”, “manual”, and “abstract” content of tasks performed in each occupation (see Autor and Dorn, 2013). “Routine” tasks are defined to be codifiable tasks that can be executed following an explicit set of rules. Technological development increasingly made it easier to write computer programs to execute these tasks, which had previously been performed by human labor. “Manual” tasks are defined as tasks required to be performed in person, such as physical tasks or service tasks. Finally, “abstract” tasks require more mental energy and involve more complex processes that could not be directly programmed, such as problem solving, management, and complex communication.[19] Prior work argues that automation directly substituted for routine jobs and complemented abstract and manual jobs.[20]

The other significant effect of automation, which has been overlooked and under-explored, is the disproportionate impact of automation on the occupations of women (Autor and Wasserman, 2013; Black and Spitz-Oener, 2010). Routine-intensive occupations were overwhelmingly dominated by female workers. In fact, I present new evidence that routine-intensive occupations employed over 60% of the high school graduate work force among women between the ages of 18 and 30. Moreover, “abstract”-intensive occupations tended to require a college degree, while occupations that were relatively “routine”- or “manual”-intensive did not. Thus, by displacing routine-intensive jobs but complementing abstract-intensive jobs, automation could have changed the labor market returns to attending college, and this change may have been stronger for women than for men. To my knowledge, this paper is the first to show that the decline in routine jobs sig-

[19]Research has shown that “abstract” tasks are becoming increasingly automated, but that this is a more recent phenomenon that began after the 1990s (Frey and Osborne, 2013; Hershbein and Kahn, 2016).

[20]Prior work has demonstrated that computers and routine tasks functioned as substitutes in production while computers and abstract tasks were complements (Brynjolfsson and Hitt, 2000; Bresnahan et al., 2002). Computers increased the marginal productivity of abstract tasks and labor demand for workers with abstract skills (Brynjolfsson and Hitt 2000; Bresnahan et al., 2002; Spitz-Oener 2008; Autor, Levy, and Murnane, 2002). Abstract tasks typically had larger educational requirements of workers, and the onset of computerization increased these educational requirements (Spitz-Oener, 2006; Brynjolfsson and Hitt, 2000; Autor, Levy, and Murnane, 2002).

nificantly increased women's college enrollment. Thus, automation directly changed the college gender gap over time, by helping drive women's enrollment to grow and eventually surpass men's enrollment.

This empirical exercise, combined with the results from Section 3, reveals that non-college employment opportunities have dramatic effects on the college enrollment rates of both men and women. Due to gender differences in the distribution of workers to occupations, shocks to certain occupations can change the gender disparity in the non-college labor market, and therefore the enrollment rate of women relative to men. Putting these findings together, it would then be natural for women's worse non-college job prospects to generate greater demand for a college degree among women relative to men.

4.1 Data

The analysis in this section utilizes the census microdata for all decades in 1950-2000 and American Community Survey (ACS) data for each year from 2001 to 2010. Both the census microdata and the ACS data are collected by the U.S. Census Bureau and provided by the Integrated Public Use Microdata Series (IPUMS; Ruggles et al., 2015). The census and ACS data are the largest publicly available data sets, making them some of the only data appropriate for occupation-level analyses of employment and wage trends at disaggregated levels of geography. The census data for 1950, 1960, and 1970 include 1% of the population. The census data for 1980, 1990, and 2000 include 5% of the population. The American Community Survey data include around 0.4% of the population for the years 2001-2004 and 1% of the population for the years 2005-2010. For the analysis in this section, I use either the sample of all men and women or the sample of 18-30 year old men and women. The data provide information on college enrollment, work characteristics, and demographic variables.

To measure how automation changed the demand for skill profiles over time, I use pre-existing occupational measures and the task-based approach for measuring the impact of automation that is typically used in the literature (Autor and Dorn, 2013; Autor, Dorn, and Hanson, 2014;

Goos, Manning, Salomans, 2014), following the suggestion in Autor (2013) that researchers re-use, recycle, and re-apply "off-the-shelf" measures of occupational skill requirements so that findings can be assessed under common metrics. In particular, I use the data set on work content compiled by Autor and Dorn (2013). Autor and Dorn (2013) uses the Dictionary of Occupational Titles (DOT) and the Occupational Information Network (O*NET) to construct measures of routine, manual, or abstract task content for each occupation.

The primary measure of occupational task content in my analysis is the composite measure of routine task intensity (RTI), which represents the relative routine-intensity of an occupation. It is constructed by Autor and Dorn (2013) using the routine-, manual-, and abstract-task measures for each occupation k:

$$\mathrm{RTI}_k = \ln(\mathrm{routine}_k) - \ln(\mathrm{manual}_k) - \ln(\mathrm{abstract}_k) \tag{4.1}$$

Occupations with high levels of the variable $\mathrm{routine}_k$ relative to the variables manual_k and $\mathrm{abstract}_k$ score high on RTI_k, while occupations with low levels of the variable $\mathrm{routine}_k$ relative to the other two task measure variables score low on RTI_k.

4.2 Descriptive Evidence

The descriptive evidence presented here serves two objectives. First, it demonstrates that young women's non-college employment opportunities were especially vulnerable to displacement by automation, relative to young women's college employment opportunities and young men's (college and non-college) employment opportunities. Second, it illustrates the variation that drives the identification of the instrumental variable approach.

I start by evaluating the result in the routine-biased technical change literature that the routine task intensity of occupations declined over time because automation displaced routine-intensive jobs (Autor and Dorn, 2013; Autor, Levy, and Murnane, 2003; Goos, Manning, and Salomans, 2014). Figure 8 graphs the average routine task intensity (RTI) in the labor force for 18-30 year old workers separately for men and women. A comparison between men and women reveals that the RTI content of women's jobs was much higher than the RTI content of men's jobs for all years in the data period, indicating that a greater proportion of the female labor force was employed in highly routine occupations relative to the male labor force. In 1950, average RTI was almost 0.8 standard deviations higher for women than for men.

Most importantly, figure 8 shows that the RTI of women's jobs plummeted from 1970 on, while the RTI of men's jobs stayed relatively steady at 0.2 standard deviations below the average. Thus, the documented decline in routine-intensity discussed in the prior literature appears to only exist for young women; for young men, RTI stayed relatively level. The evidence indicates that women's jobs drove the decline in routine task intensity among young workers. Increased automation, and the subsequent decline in the routine content of human labor, appears to have displaced women's job prospects by more than men's job prospects.

Appendix figure A.2 decomposes the change in RTI into its three component parts: the routine intensity measure, the manual intensity measure, and the abstract intensity measure. The raw data show that the decline in RTI for women is driven entirely by the decline in the routine intensity measure.[21] These trends are consistent with the evidence in appendix table A.7, which summarizes correlations between routine and abstract work for college and non-college female workers from 1950 to 2000. For women, the strong positive correlation between not attending college and working in routine-intensive jobs dissipates over the decades, while the positive correlation between attending college and working in abstract-intensive jobs becomes stronger and larger.

Figure 8 suggest that women's jobs experienced declines in routine-intensive task content while men's jobs did not. Did the automation of the office displace some jobs more than others in way that affected women more than men? The left panel of figure 9 shows that this appears to be the case. I separately plot the labor share of high- and low-RTI occupations and find that labor share for women in high-RTI occupations declined, whereas the labor share for men in high-

[21] Abstract intensity increased by the same extent for both men's and women's jobs, while manual intensity remained relatively constant during this time.

RTI occupations did not.[22] Women in high-RTI occupations peaked at a little over 25 percent of the labor force in 1970, before declining precipitously to about 20 percent of the labor force in 2000. In contrast, women in low-RTI occupations, and men in high- and low-RTI occupations did not experience declines in labor share. In fact, their labor share actually rose slightly during this period.

The natural next question is: did automation affect the college-going margin for women? The right panel of figure 9 breaks down the change in labor share for college and non-college women by high- and low-RTI occupations. The employment share of high-RTI non-college women peaked at 14 percentage points before dropping almost 60% by 2000. For non-college women in low-RTI occupations, labor share remained steady, and for college women, the labor share increased during this period. These trends suggest that the displacement of jobs by automation documented by the prior literature disproportionately impacted the non-college job prospects of women. Simultaneously, the labor share of college women in both high- and low-RTI occupations increased. The results point to an asymmetrical effect of automation on labor market prospects, where the occupations that employed a large share of the non-college female workforce declined in labor share but the occupations that employed college women did not.

Finally, automation fundamentally changed the labor structure of non-college occupations. Figure 10 graphs the density of occupations by proportion of female workers in each occupation. The left panel displays the density for "non-college" occupations, in which the majority of workers had only high school degrees, while the right panel displays the density for "college" occupations, in which the majority of workers had college degrees. There were striking changes in gender composition among college and non-college occupations from 1970 to 2010. In 1970, the majority of non-college occupations were male-dominated (less than 30% women), some non-college occupations were female-dominated (at least 70% women), and very few occupations were "gender-equitable" (30-70% women). The female-dominated occupations that form the mass at

[22]To accord with measures commonly used in the literature, I define high-RTI occupations as occupations in the top third of RTI in 1980, and low-RTI occupations as occupations in the bottom third of RTI in 1980 (see Autor and Dorn, 2013). Because that the graph only depicts the labor share for occupations at the top and bottom third of RTI, the labor shares do not sum to one in any year.

the right of the 1970 non-college occupation distribution were all highly routine-intensive occupations: stenography, typist, secretary, telephone operator, etc. In contrast, college occupations were overwhelmingly male-dominated.

Over time, as automation displaced routine-intensive jobs, the mass at the right of the non-college occupation distribution declined and eventually disappeared. By 2010, almost all non-college occupations were male-dominated, and the non-college labor market became a relatively inhospitable place for women. In contrast, the number of female-dominated or gender-equitable college occupations rose. The descriptive evidence suggests that in the 1970s, women had job options in the non-college market but relatively few job options in the college market. Over time, their non-college job prospects declined while college occupations became more accessible. By 2010, the reverse is true. Very few non-college jobs were accessible to women, and the occupations that employed women had significantly lower wages, as shown by figure 2. On the other hand, women's access to the college labor market dramatically expanded, since college occupations that used to be traditionally male are now gender-equitable or even female-dominated.

4.3 Identification

To instrument for the decline of routine-intensive employment opportunities, I follow the approach of Autor and Dorn (2013) and construct a modified shift-share instrument that predicts the employment share of routine-intensive occupations in a local labor market. The logic behind this instrument is that local labor markets with higher 1950 shares of routine-intensive employment ("routine employment share") experienced greater automation than local labor markets with low 1950 shares of routine-intensive employment. The instrument is constructed as follows:

$$\widehat{\mathrm{RSH}}_c = \sum_{i=1}^{I} E_{i,c,1950} \mathrm{RSH}_{i,-c,1950} \tag{4.2}$$

where a local labor market is a commuting zone, indexed by c. $E_{i,c,1950}$ represents the employment share of industry i in commuting zone c in 1950. $RSH_{i,-c,1950}$ represents the share of routine occupations in industry i in all states except the state with commuting zone c.

In the first stage regression, I interact the shift-share instrument with a matrix of year dummies to nonparametrically predict the effect of the instrument on the actual labor share of routine-intensive employment in future years. The idea behind the identification strategy is that local labor markets with high baseline shares of industries that experienced a large amount of automation later on will experience larger displacement of women's non-college labor market opportunities later on. The instrument relies on the assumption that high 1950 shares of industries that automated later on should influence employment opportunities in future years, but not directly influence college enrollment rates in future years.

The first stage regression is

$$\mathrm{RSH}_{ct} = \alpha_0 + \alpha_1 \widehat{\mathrm{RSH}}_{c,1950} \times \mathbf{1}(\text{year} = t) + \alpha_2 X_{ct} + \phi_t + \theta_c + e_{ct} \tag{4.3}$$

where c indexes local labor market, t indexes the year, X_{ct} is a vector of controls for local labor market c at year t, ϕ_t is a vector of year dummies, and θ_c is a vector of fixed effects for local labor markets. I use the estimates from the first stage regression to predict variation in actual routine share employment for each year t, denoted by $\widetilde{RSH}_{ct}$.

The second stage regression is

$$Y_{ct} = \beta_0 + \beta_1 \widetilde{RSH}_{ct} + \beta_2 X_{ct} + \phi_t + \theta_c + \varepsilon_{ct} \tag{4.4}$$

The IV regression estimates the effect of declining employment opportunities in routine-intensive industries on the female college enrollment rate, the male enrollment rate, and the college gender gap (defined as the male college enrollment rate less the female college enrollment rate).

The first stage regression obtains the variation in actual routine employment share due to the industry composition of a local labor market in the base year of 1950, weighted by the national routine employment share of each industry in 1950. The measure is compiled from industry characteristics in 1950, which pre-date the changes in automation that occurred starting in the 1970s. The instrument should therefore net out any post-1950 correlations between employment opportunities and college enrollment, as long as these relationships are independent of 1950 industry composition. Moreover, since the instrument takes the average routine share of employment per industry for all states except the one that contains the commuting zone of interest, it nets out local labor market shocks that influence educational outcomes along dimensions other than changes in the employment share of routine-intensive occupations.

The exclusion restriction specifies that industry composition in a base year influences college enrollment decisions in a future year only through changing non-college occupations in that future year. This instrument leverages the argument in Autor and Dorn (2013) that labor markets with large baseline shares of industries high in routine-intensive work were the ones with greater demand for automation. Since automation displaced routine-intensive work, the instrument should predict future declines in job market opportunities for workers in routine-intensive occupations. My first-stage results, presented in table 8, are consistent with this argument. The correlation between the actual employment share of routine-intensive occupations and the instrumental variable is negative for all years starting in 1970, when routine-intensive employment shares first began declining. The correlation grows strictly more negative with each successive decade, which is also consistent with the story that the growth of automated processes in the workplace lead to persistent contractions in employment demand among routine-intensive occupations.[23]

There are a number of alternative explanations that lead to a violation of this exclusion restriction. First, one might argue that increased automation in different labor markets could have made it easier to attend school through decreasing the costs of finishing high school or expanding the resources of post-secondary institutions. For this alternative hypothesis to explain my findings, automation would have had to affect men and women differentially, since I find a significantly larger increase in female enrollment relative to the insignificant effect on male enrollment. This appears unlikely.

[23]First stage regressions were also conducted using median wages in routine-intensive occupations (not shown). Autor and Dorn (2013) do not find that automation uniformly decreased wages in the way that it did with employment. My findings are similar. When the instrument is interacted with year dummies, the resulting coefficients do not appear to be significantly negative and decreasing. The first stage regressions show that wages do not experience monotonic declines in labor markets with high predicted 1950 routine share employment, which violates a necessary condition of the LATE theorem for IV estimation (Angrist and Pischke, 2009). Because the instrument is a better predictor of employment opportunities in routine-intensive occupations, the analysis focuses on the relationship between employment opportunities in routine-intensive occupations and college enrollment.

Another alternative hypothesis is that high 1950 levels of routine-intensive employment are correlated with omitted characteristics that influence both non-college employment and schooling choices. For example, local labor markets with social norms that were conducive to women working may have had higher 1950 routine employment shares. These social norms could then have encouraged more women to attend college twenty years later. Here, it is important to note that my identifying variation draws from *predicted* routine employment shares, not actual routine employment shares. The variation in my specification arises from the industry composition in a local labor market in 1950. In other words, local labor markets with high 1950 shares of the industries that happened to automate faster later on were the markets that had high college enrollment among women (but not men) later on. By constructing predicted routine employment share using a shift-share approach, the instrumental variable strategy nets out the confounding effects of actual initial market conditions, as well as unobservable characteristics correlated with actual initial market conditions.

4.4 Results

I find that declining routine-intensive occupations, which employed the majority of the non-college female workforce among young workers, increased the college enrollment rate significantly more for women than for men. The main instrumental variable regression results are reported in table 9 for the sample of 18-25 year olds. Table 10 reproduces the regression for the larger sample of 18-30 year olds. For both tables, the first two columns report the results for women, the second two columns report the results for men, and the last two columns report the results with the gender gap (male enrollment minus female enrollment) as the dependent variable. All regressions include fixed effects for commuting zone, year, and region. The even-numbered columns also include commuting-zone level controls for total population, proportion of women, proportion of blacks, proportion of Hispanics, proportion by ten-year age bin.

Table 9 reports the main regression estimates, where the outcome variable is the proportion of 18-25 year olds who have ever enrolled in college. Decreasing the share of routine-intensive occupations by an additional percentage point leads the female enrollment rate among 18-25 year olds to increase by 0.50 percentage points ($p < 0.05$), as shown in column (1). In contrast, the effect on male enrollment, shown in column (3), is very close to zero (point estimate of -0.04) and statistically insignificant. The coefficient estimates are both economically and statistically significantly greater for women than for men. Column (5) shows that the net impact on the college gender gap (male enrollment less female enrollment) is a decline of 0.46 percentage points ($p < 0.01$). Columns (2), (4), and (6) add demographic controls at the commuting zone level, which allow for variation in enrollment rates due to the demographic composition of individuals within the commuting zone. I find that in all cases, the point estimates do not significantly change after the inclusion of demographic controls. Column (2) shows that the effect size increases directionally, such that an additional percentage point decline in the labor share of routine-intensive occupations increases female enrollment by 0.74 percentage points ($p < 0.01$). Column (4) show that the effect on male enrollment remains insignificant. Again, the estimated effect on female enrollment is economically and statistically significantly greater for women than for men. Finally, the net impact on the college gender gap, shown in column (6), is a decline of 0.54 percentage points.

Table 10 expands the sample to the proportion of 18-30 year olds who have ever enrolled in college. Since the sample now includes individuals who are further from the margin of college-going, the point estimates noticeably decline. As shown by column (1), decreasing the share of routine-intensive occupations by an additional percentage point leads to an increase in female enrollment by 0.35 percentage points. In contrast, column (3) shows that the effect of routine occupations on male enrollment is much smaller and insignificant. Column (5), which presents the results on the college gender gap (male enrollment less female enrollment), shows that the corresponding effect is a 0.18 decline in the college gender gap. As with the sample of 18-25 year olds, I find that the point estimates do not significantly change after including demographic controls. Column (2) shows that adding demographic controls directionally magnifies the effect of routine-intensive occupations, such that the estimated effect of an additional percentage point decline in routine-intensive labor share on female enrollment is now 0.48 percentage points. The effect on male enrollment, shown in column (4), remains insignificant. The net effect on the college gender gap is a 0.25 percentage point decline, shown in column (6).

5 Explaining Time Trends in the Reverse College Gender Gap: Theoretical Model

So far, the paper makes the case that women's worse non-college job prospects contribute in major ways to their greater college enrollment rate. But women's non-college prospects have always been worse than men's, so why have women not always exceeded men in college-going? The literature on the college gender gap has identified two symmetric puzzles: first, why did women attend college at greater rates than men after 1980, when men have always worked more and earned more than women?[24] Second, why did men attend college at greater rates than women before 1980, when women have always had a higher observed college premium than men?[25] This section presents a theoretical model that reconciles both of these contradictions.

The theoretical framework demonstrates that non-college jobs played a growing role in women's college-going decisions, and that this contributed to the growth and eventual dominance of women in college classrooms. Since this paper focuses on the role of labor market returns, the model purposefully abstracts from other factors that have already been shown to contribute to the college gender gap, such as abilities (see Becker, Hubbard, and Murphy, 2010; Jacob, 2002; Bertrand and Pan, 2013; Goldin, Kuziemko, and Katz, 2006), marriage market outcomes (see Chiappori, Iyigun, and Weiss, 2009; Chiappori, Costa Dias, and Meghir, 2015; Chiappori, Salanie, and Weiss, 2015; Bronson, 2015; Zhang, 2016; Low, 2017), or gender differences in preferences (see Niederle and Vesterlund, 2010), by treating these factors as equal between men and women. The model assumes three key differences between men and women: expected wage rates, time available for labor, and exposure to fertility risk. Within the model, these three differences are sufficient to explain why men exceeded women in college-going at first while women exceeded

[24]For literature that identifies this question, see DiPrete and Buchmann 2008; Jacob 2002; Zafar 2013; Becker, Hubbard, and Murphy 2010; Goldin, Katz, and Kuziemko 2006

[25]Becker, Hubbard, and Murphy (2010) and Goldin, Katz, and Kuziemko (2006) have identified men's initially greater enrollment rate to be the major puzzle in the literature.

men in college-going later on. In addition, one natural implication of this model is that the greater enrollment rate of women leads to the lower college wage rates for women compared to men. In other words, the gender wage gap among college workers is a direct result of the gender gap in college enrollment.

5.1 Model Setup

Individuals live for two periods. In each period, they have quasilinear utility over consumption c_t and leisure ℓ_t, as well as a fixed amount of housework that must be completed. Individuals maximize their utility by choosing how to allocate their remaining time net of housework.

In period 0, individuals must decide whether to attend college. The decision to attend college is denoted $s \in \{0,1\}$, where $s=0$ represents the choice to not attend college and $s=1$ represents the choice to attend college. Individuals make the decision to attend college based on their decisions regarding expected utility in periods 1 and 2.

In period 1, all individuals are single. They must allocate their time net of housework between college s, labor h_1, and leisure ℓ_1. If they work in period 1, they will receive expected wage rate $\underline{w}$. College enrollees must pay the costs of attending college, which consist of monetary costs d and idiosyncratic non-monetary costs ε, where ε is drawn from the distribution $G(\varepsilon)$.[26] In addition, attending college requires z units of time, where $z=1$ is sufficient to obtain a college degree. As I will discuss in detail later, with probability q individuals face an unplanned pregnancy and expect to complete only $z<1$ of their college requirements.

In period 2, all individuals marry, have a child if the wife did not have one in period 1, and pool their income with their spouse. Individuals allocate their time net of housework between labor h_2 and leisure ℓ_2. Importantly, couples can pool their time to complete the household production required by the family. Their expected wage rate in period 2 is determined by whether a college degree was earned at the end of period 1, as denoted by sz, where $w(sz) \in \{\underline{w}, \overline{w}\}$. Individuals who

[26] Following the formulation of Becker, Hubbard, and Murphy (2010), ε can be considered an ability cost. High-ability individuals have low non-monetary costs of college, while low-ability individuals have high non-monetary costs of college.

do not earn a college degree by the end of period 1 ($sz < 1$) receive wage rate $\underline{w}$, and individuals who receive a college degree by the end of period 1 ($sz = 1$) receive wage rate $\overline{w}$, with $\overline{w} > \underline{w}$. Individuals also expect to receive k in spousal earnings.

In period 2, the maximization problem is given by[27]

$$\begin{aligned} V_2(s,z) &= \max_{c_2,\ell_2} c_2 + \ln(\ell_2) \\ \text{subject to } &\alpha\big[w(sz)\underbrace{[T_2 - \ell_2]}_{h_2} + k\big] = c_2 \end{aligned} \tag{5.1}$$

In period 1, the maximization problem is given by

$$\begin{aligned} V_1(s,z) &= \max_{c_1,\ell_1} c_1 + \ln(\ell_1) - \varepsilon s + \beta V_2(s,z) \\ \text{subject to } &\underline{w}[\underbrace{T_1 - \ell_1}_{h_1} - sz] = c_1 + dsz \end{aligned} \tag{5.2}$$

In period 0, the utility maximization problem is given by

$$\max_s \mathbb{E}V_1(s,z) = \max_s \ (1-q)V_1(s,z=1) + qV_1(s,z<1) \tag{5.3}$$

5.2 Gender

Denote women by the subscript f and men by the subscript m. Men and women differ in three key ways. First, I model the fact that men sort into higher paying occupations relative to women as a higher expected wage rate for men than women. Here, expected wage rates can be considered the sum of earnings in each occupation weighted by the probability of filling an occupation. A decline in non-college employment opportunities would be represented as a decline in the expected wage rate. In this formulation, men have higher wage rates than women within education groups. Following the data, I assume $\overline{w}_m > \overline{w}_f$ and $\underline{w}_m > \underline{w}_f$. The gender disparity among workers without a college degree is larger than the gender disparity among college graduates: $\overline{w}_f - \underline{w}_f > \overline{w}_m - \underline{w}_m$.

[27] The main results of the model are generalizable to the case where utility with respect to leisure follows a function v, where v is quasiconcave, twice differentiable, and homogeneous of a degree between 0 and 1.

Second, men and women have potentially different time net of housework to allocate to labor h_t, leisure ℓ_t, and schooling s. In period 1, single men and women without children have the same amount of housework they must complete. They will have an equal amount of time net of housework $T_1 = T$ to allocate to labor, leisure, and schooling. In period 2, men and women marry, have a child if the woman did not have a child in period 1, and pool their time to complete the housework needed for the family. The time devoted to housework will differ between men and women, because married couples can specialize. Assume households are efficient and that one member can complete all the housework needed by the household. The higher wage rate of men implies lower opportunity costs for women to engage in housework, assuming that both are equally efficient at it and that the marginal productivity of time in housework is constant. The comparative advantage of men in market work leads women to spend time completing all the housework needed by the family, following Becker's theory of household specialization (Becker 1981, 1985).[28] Let T_2 represent the time net of housework in period 2, with $T_2 = T_f$ for wives, $T_2 = T_m$ for husbands, and $T_f < T < T_m$. To remain consistent with observed trends, the model assumes married men will always work. In other words, men's time net of housework T_m is high enough that it is always optimal for married men (both with and without college degrees) to work.

Lastly, women face fertility risk but men do not ($q_m = 0$). With probability $q_f > 0$, women will have an unplanned pregnancy in period 1 while single. Having an unplanned pregnancy introduces the expectation that female college enrollees will leave school without fulfilling the time requirements necessary to earn a college degree. In the state where women do not have an unplanned pregnancy in period 1, $z = 1$. In the state where women have an unplanned pregnancy in period 1, $z < 1$. They would then receive expected wages $\underline{w}_f$ instead of $\overline{w}_f$ in period 2. All women know the probability of an unplanned pregnancy q_f prior to making their decision to attend college in period 0. They know whether or not they have an unplanned pregnancy once period 1 starts, *after* their college-going decision is made but *before* their labor or leisure decisions in each period

[28] An alternative formulation which achieves the same result is to assume a comparative advantage in housework for women, which leads men to specialize in market work and explains their higher expected wage rates (Becker, 1985; Galor and Weil, 1996).

are made.

Other parameters that might differ between men and women are expected earnings from the spouse k, share of household income α, and discount factor for period 2 utility β. However, the crucial differences explored in the model are the three key differences described above. For ease of exposition I will assume α, β, and k are the same between men and women.

5.3 When is enrollment higher for men than women? When is enrollment higher for women than men?

Based on equations (5.1)-(5.3) and the three key differences between men and women, the schooling decisions for men and women can be derived. Individuals choose to attend college if and only if they receive draws of ε below a threshold college-going value, which corresponds to γ_m for men and γ_f for women. The appendix lists the derivations, details, and analysis. The key points are summarized below.

First, men's higher earnings and higher labor force participation make their college-going decisions strictly more responsive to labor market returns than the college-going decisions of women. Since labor market returns for men were high, male enrollment shot up quickly and leveled off quickly. In contrast, female enrollment grew gradually but steadily over time. Initially, womens college-going decisions were not very responsive to their labor market returns, for many potential reasons. The model focuses on two reasons that have been identified by the literature: housework responsibilities kept most married women out of the labor force, and the risk that an unanticipated pregnancy would prevent women from finishing educational investments created uncertainty in whether women could capitalize on their labor market returns. Increased access to contraceptive technologies and advances in household production that decreased the amount of housework performed by women increased the responsiveness of women's college-going decisions to labor market returns over time.

Second, the model delivers closed form solutions regarding whether and how much women will work, based on their expected wage rates and the time they have available after housework. Expected non-college wage rates directly figure into women's college-going decisions even when married women do not work, since non-college wage rates represent the opportunity cost of college attendance when women are single. In contrast, expected college wage rates only play a role in women's college-going decisions when it is optimal for female college graduates to attend college. Consequently, women derive labor market benefits from earning a college degree only when female college graduates become efficient enough at housework to have time for market work.

Third, fertility risk mediates the slope of the growth in female enrollment, in that declines in fertility risk increase the responsiveness of women's college enrollment decisions to their labor market returns. When fertility risk q is above $1-\frac{\overline{w}_m-\underline{w}_m}{\overline{w}_f-\underline{w}_f}$, female enrollment will always be below male enrollment. When fertility risk declines below $1-\frac{\overline{w}_m-\underline{w}_m}{\overline{w}_f-\underline{w}_f}$, it becomes possible for female enrollment to surpass male enrollment.

Figure 11 delivers the final result of the model. The left panel summarizes the role of increasing housework efficiency and declining fertility risk on how wage rates affect women's college-going. The x-axis is T_i, time net of housework for gender i. The figure depicts how the threshold college-going value for gender i, γ_i, changes as T_i increases.

The effect of increasing household efficiency on how female college-going responds to wage rates is represented by increasing time net of housework for women, $T_i = T_f$. Female college-going threshold γ_f grows as T_f increases, represented by right-ward movement along the x-axis. This growth stems entirely from the result that increasing T_f increases the strength of the college-going response to wage rates. This growth is discontinuous, depending on the relationship between time net of housework T_f and wage rates $(\underline{w}_f, \overline{w}_f)$. The figure also graphs male enrollment, γ_m which grows as time net of housework for men $T_i = T_m$ increases (represented by a right-ward shift along the same x-axis).

The effect of declining fertility is represented by the shift from $\gamma_f(\widetilde{q}_f)$ to $\gamma_f(\widehat{q}_f)$, where $\widetilde{q}_f > 1-\frac{\overline{w}_m-\underline{w}_m}{\overline{w}_f-\underline{w}_f} > \widehat{q}_f$. Again, $1-\frac{\overline{w}_m-\underline{w}_m}{\overline{w}_f-\underline{w}_f}$ is the threshold below which it is possible for female enrollment to surpass male enrollment. For this reason $\gamma_f(\widehat{q}_f)$ crosses γ_m, but $\gamma_f(\widetilde{q}_f)$ never crosses γ_m.

Proposition 1 summarizes the conditions which create gender differences in college enrollment.

Proposition (Proposition 1a)**.** *If $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$, there exists a $T_i = T_{mf}$ where $\gamma_f(q_f, T_{mf}) = \gamma_m(T_{mf})$.*

Men will exceed women in college enrollment if $q_f > 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$ or if $T_f < T_{mf}$ and $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$.

Proposition (Proposition 1b)**.** *If $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$, there exists a $T_i = T_{mf}$ where $\gamma_f(q_f, T_{mf}) = \gamma_m(T_{mf})$. For any arbitrary $\widehat{T}_m > T_{mf}$, there exists $\widehat{T}_f < \widehat{T}_m$ where $\gamma_f(q_f, \widehat{T}_f) = \gamma_m(\widehat{T}_m)$. Then, $\forall T_f \in (\widehat{T}_f, \widehat{T}_m)$, $\gamma_f(q_f, T_f) > \gamma_m(\widehat{T}_m)$.*

Women will exceed men in college enrollment if women experience fertility risk q_f, time net of housework T_f, and men experience time net of housework T_m, where $q_f < 1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f}$, $T_m = \widehat{T}_m > T_{mf}$, and $T_f \in (\widehat{T}_f, \widehat{T}_m)$.

Proof. See theory appendix B.

Proposition 1b demonstrates that necessary conditions for women to exceed men in college enrollment are that fertility risk q_f must fall below $1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f}$ and that housework must fall to a point where it is optimal for college women to work (in other words, time net of housework T_f must exceed $\frac{1}{\alpha \overline{w}_f}$). Once these two conditions are met, it is possible for women to take advantage of their higher labor market returns. Because $\overline{w}_f - \underline{w}_f > \overline{w}_m - \underline{w}_m$, the slope of female enrollment γ_f exceeds the slope of male enrollment γ_m. As long as housework time for women is sufficiently low, female college-going will be higher than male college-going even if women have less time for work and lower wage rates than men.

The right panel of figure 11 represents the change in female college-going threshold γ_f given a decline in the female non-college wage rate $\underline{w}_f$. Consider a decline in $\underline{w}_f$ to $\underline{\underline{w}}_f$, which pushes the y-intercept up and shifts the vertical axis $\frac{1}{\alpha \underline{w}_f}$ further to the right, increasing the slope of γ_f. This change is represented by the shift from $\gamma_f(\underline{w}_f)$ to $\gamma_f(\underline{\underline{w}}_f)$. Since γ_f both shifted up and now has a steeper slope, there is a lower value of T_{mf} for which $\gamma_f(\underline{\underline{w}}_f, T_{mf}) = \gamma_m(T_{mf})$.[29] For any $\widehat{T}_m$, define $\widehat{T}_f$ and $\underline{\underline{\widehat{T}}}_f$ to be such that $\gamma_m(\widehat{T}_m) = \gamma_f(\underline{w}_f, \widehat{T}_f) = \gamma_f(\underline{\underline{w}}_f, \underline{\underline{\widehat{T}}}_f)$. It can be shown that $\underline{\underline{\widehat{T}}}_f < \widehat{T}_f$.[30]

[29] See appendix B for a formal proof.
[30] See appendix B for a formal proof.

This result is significant because it shows that declines in non-college wage rates for women complement increasing housework efficiency and decreasing fertility risk in enabling female enrollment to grow and overtake male enrollment. A decline in non-college wage rates enable female college enrollment to exceed male college enrollment at lower levels of household efficiency and higher levels of fertility risk. Declining employment opportunities in the non-college market therefore help explain not only why women overtook men in college enrollment, but also why the overtaking occurred as early as the 1980s, when female labor force participation was still quite low at 50%.

5.4 Estimating the impact of non-college jobs on aggregate trends

To show that non-college occupations played an important role in the evolution of the college gender gap, I conduct a back-of-the-envelope calculation to determine how much of the aggregate change in college enrollment can be explained by changes in non-college employment for men and women. Table 11 shows the change in non-college employment for men and women. Using the point estimates from Sections 3 and 4, I find that the decline in non-college jobs for women can explain 82.3% of the change in female enrollment over time and that the changes in non-college jobs for men can explain 28.2% of the changes in male enrollment over time.

However, changes in non-college employment arise from both supply and demand effects. Supply-driven declines in non-college employment may arise, for example, from workers obtaining college degrees at higher rates for reasons unrelated to employment changes, non-college workers choosing to leave the labor force, or influxes of non-college workers from foreign countries. Demand-driven changes in non-college employment, on the other hand, stem from changes in employer demand for non-college workers. To more precisely estimate changes in college enrollment that stem from demand-driven changes in non-college jobs, I perform a simple variance decomposition which utilizes the relationship between the ordinary least squares estimator and the two-stage least squares estimator to back out the proportion of an aggregate change that can be attributable to demand changes (Autor, Dorn, and Hanson, 2013). My back-of-the-envelope calculations show that non-college jobs explain about 40.3% of the change in female enrollment and 14% of the change in male enrollment from 1970 to 2010.

Next, I use the model results to estimate the counterfactual college gender gap based on changes in non-college jobs alone by holding all other factors that influence college enrollment fixed at 2010 levels. I use the derivative of the schooling rule in the model to obtain a closed form expression of how college enrollment responds to non-college employment:

$$\left[\frac{\partial(\text{college enrollment})}{\partial(\text{non-college employment})}\right]_t = \beta(\text{time spent at work}_t) - \text{time spent at school} \tag{5.4}$$

The equation produces a measure of the responsiveness of female college-going to non-college employment. This responsiveness depends on the amount of time worked in the labor market. Historically, female enrollment was low – in 1970, only 30% of married women participated in the labor market at all – making the responsiveness of female college-going to non-college employment low. If married women had always worked as much as they did in 2010, they would have been far more responsive to their non-college labor market conditions. How would the trajectory of female college enrollment have changed in this case?

I perform a back-of-the-envelope counterfactual calibration exercise where I multiply the first term in equation (5.4) with the ratio of the time spent at work in 2010 over the time spent at work in a prior year t. This provides a rough approximation of how responsive female college-going would have been if women always worked as much as they did in 2010. I then take information from the American Time Use Survey, the American's Use of Time Survey, and the Time Use in Economic and Social Accounts Study to obtain measures of time spent in school and time spent at work for women (Aguiar and Hurst, 2007; Sayer, 2014).

Figure 12 presents the counterfactual estimation of college enrollment for men and women, where changes in college enrollment arise solely from changes in non-college jobs. If women had always worked as much as they did in 2010, they would have been much more responsive to their non-college job prospects. Based on the counterfactual estimation, their relatively anemic options in the non-college labor market would have pushed their college enrollment to exceed the college enrollment of men for all years in the estimation exercise. In other words, if women had always worked in the labor market as much as they have in recent years, they would have never lagged behind men in college enrollment. This back-of-the-envelope exercise provides suggestive evidence that the low hours women used to work were a key reason behind why women did not attend college at higher rates than men before 1980.

6 Conclusion

The greater college enrollment of women over men has been a long-standing open question. While most of the literature has focused on how college-going decisions are driven by preparedness, marriage market concerns, social concerns, or labor market outcomes for college graduates, this paper provides new evidence that the labor market for high school graduates plays a key role in explaining this gender gap. I document the large gender disparity in non-college job options and demonstrate that these disparities create unequal demand in a college degree between men and women. I then construct a theoretical model to explain how the gender imbalance in non-college job options can rationalize the greater enrollment of men before the 1980s and the greater enrollment of women after the 1980s, despite the fact that women's observed college premium has been consistently higher than men's during this time.

This paper speaks to the importance of outside options in contributing to the large difference in human capital investments between men and women. My findings demonstrate that men may not be "under-investing" in education as much as it may at first seem. Much of the public debate on the college gender gap has focused on how myopia, poor behavior in school, and lack of interest in learning present barriers to men from optimally investing in their education (Economist, 2015). In contrast, I show that one key reason behind why men enroll in college at lower rates than women is because they have more attractive alternatives to attending college. Thus, even if everyone behaved rationally, men would still be expected to enroll in college at a lower rate than women.

In addition to explaining present conditions, this paper rationalizes trends in the college gender gap over time, which have puzzled social scientists for decades. I demonstrate that the increasing rate of automation disproportionately displaced the non-college job options of young women just as female labor force participation began to grow substantially, which in turn led female college enrollment to increase at rates higher than male college enrollment. At the same time, men's non-college job opportunities remained plentiful by comparison, leading a greater proportion of men than women to rationally forego attending college. The combination of these factors contributed to both the greater college enrollment of men prior to 1980 and the greater college enrollment of women after 1980.

The results presented here raise further questions that merit exploration. Since women have access to fewer lucrative options with only a high school degree, higher earnings are required to induce the marginal man to enter college relative to the marginal woman. Average wages for male college graduates will therefore be higher than average wages for female college graduates. The gender gap in college enrollment thus creates a persistent gender gap in earnings among college workers. The large steady gap in college enrollment between men and women may explain why the gender gap in wages has failed to close, despite efforts from governments and firms alike. To my knowledge, this paper is the first to reveal a tension between the gender gap in college enrollment and the gender gap in wages, wherein interventions to narrow the gender gap in wages will widen the gender gap in college-going, and interventions to narrow the gender gap in college enrollment will widen the gender gap in wages.

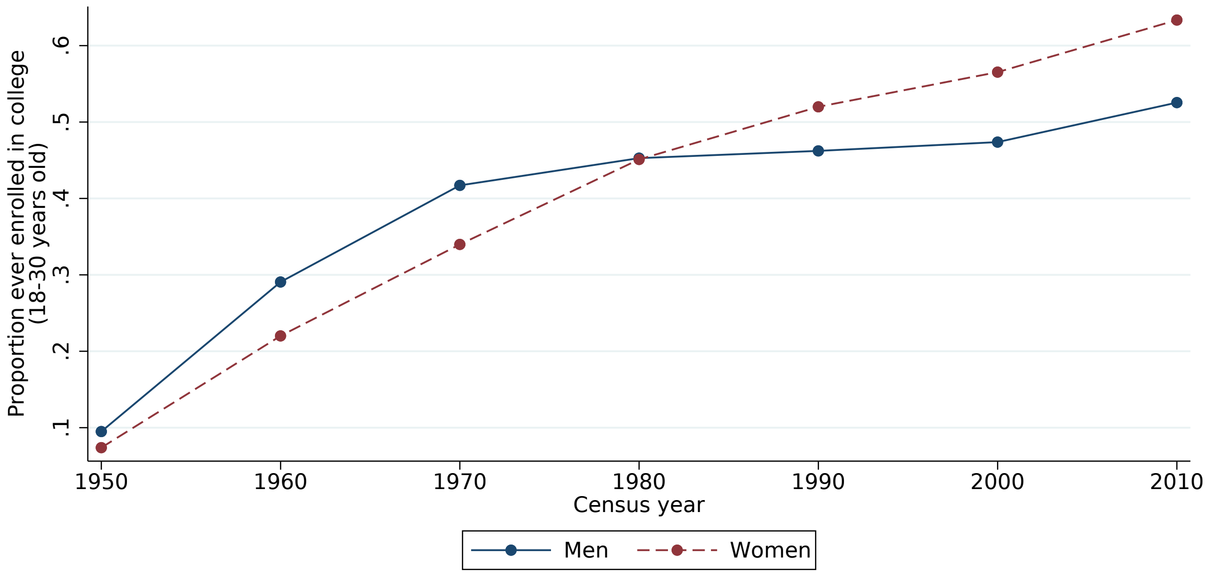
A second, related implication of this paper is that women are more likely than men to choose non-STEM fields among college enrollees. Since the opportunity cost of attending college is lower for women than men, women have greater freedom to major in a less lucrative field and still make the investment in attending college worthwhile. Recent work by Card and Payne (2017) support this prediction. They show that men are 13 percentage points more likely than women to major in a STEM field, and that 9 of these 13 percentage points can be attributed to the higher college enrollment rate of women.

Overall, highlighting the role of the non-college labor market in the college gender gap yields the insight that different outside options lead men and women to self-select into attending college at differential rates. The marginal college-going woman will differ from the marginal college-going man, and this creates persistent differences in the fields that men and women choose, the average wages of men and women across the population of college workers, and a variety of other employment outcomes.

A third implication of this paper is that greater study should be devoted to non-college jobs in order to determine the optimal role of policy in individuals' private education decisions. If men choose to forego valuable college investments due to high paying non-college job prospects, future research should focus on what these jobs are. Do they pay enough to support a family over a lifetime? Are they viable career paths? How do people who forego formal human capital investment to work in these jobs weather adverse labor market shocks in the future? In a companion paper (Chuan, 2017), I take one step in this direction by estimating a structural Roy model to show that some men can indeed maximize lifetime earnings by foregoing a college degree. However, more work to investigate the non-college labor market is needed in order to determine the welfare consequences of foregoing a college degree.

Figures

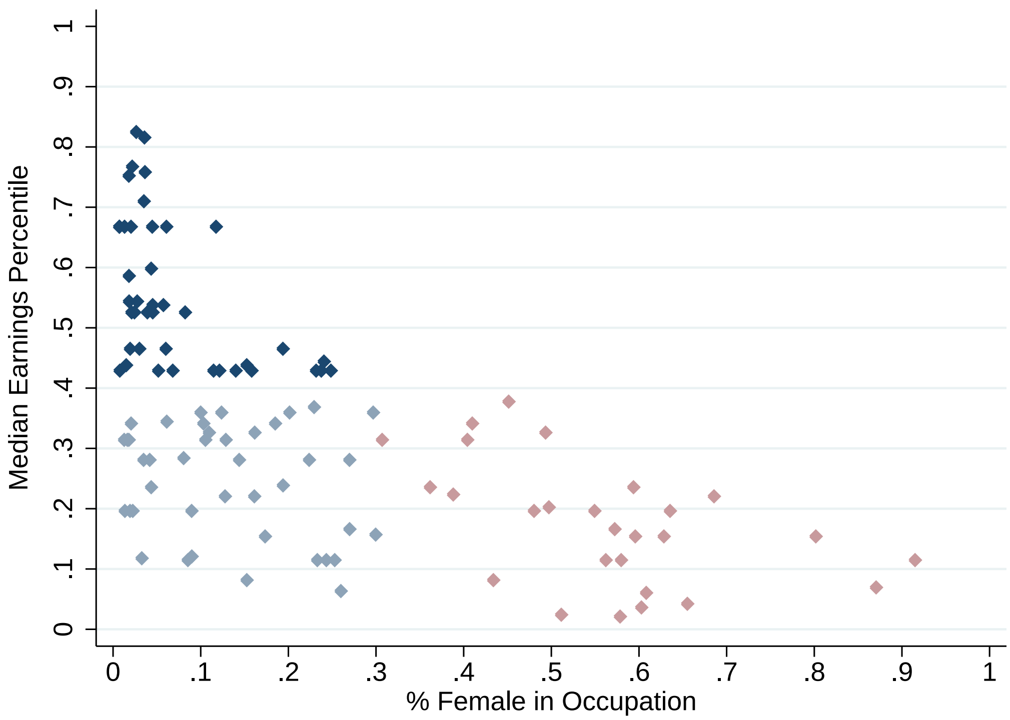
FIGURE 1
COLLEGE ENROLLMENT BY GENDER



Notes: Figure 1 shows the proportion of men and women between the ages of 18 and 30 who have ever enrolled in college. Before the 1980s, the proportion of men ever enrolled in college was greater than that of women. The gender gap in college enrollment closed when women's college enrollment rate converged to that of men. After the 1980s, the gender gap in college enrollment reversed when the college enrollment rate of women surpassed that of men. The male college enrollment rate has leveled off since the 1980s while the female college enrollment rate continued to increase from 1980 to 2010. The figure uses census microdata for each decade in 1950-2000 and American Community Survey (ACS) data for each year in 2001-2010.

FIGURE 2

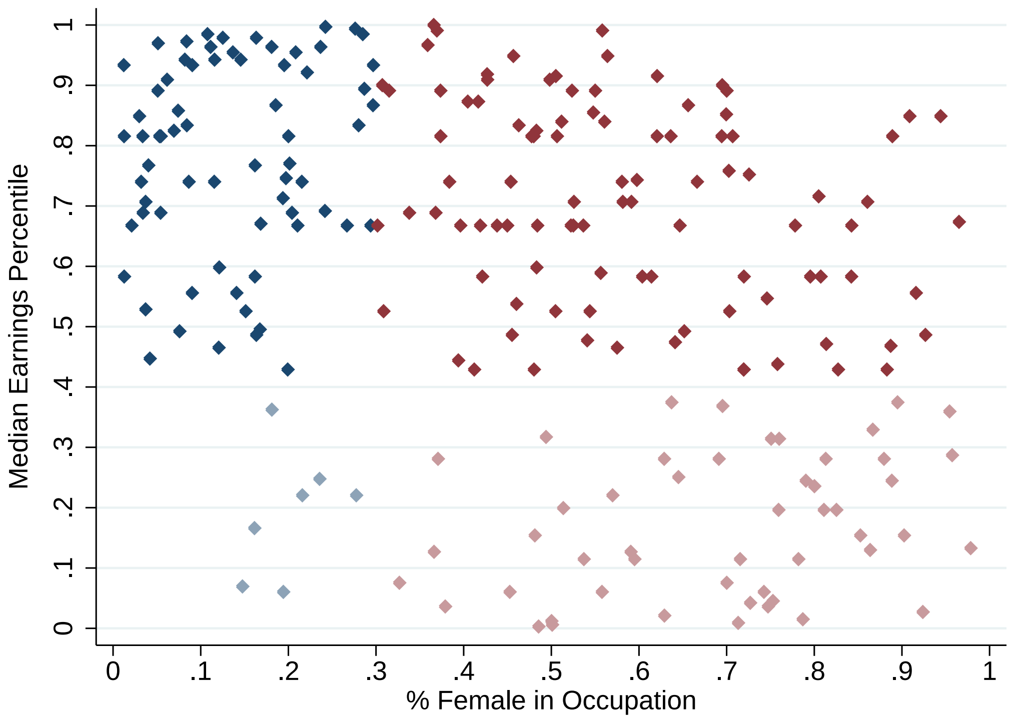
NON-COLLEGE OCCUPATIONS BY GENDER COMPOSITION AND PERCENTILE MEDIAN EARNINGS



Notes: The figure depicts a scatter plot of all non-college (at least 50% workers with only a high school degree) occupations by proportion of female workers and median earnings percentile. First, the majority of occupations (over 60%) are male-dominated, with 20% or fewer female workers. Second, occupations which employ a non-trivial fraction of women pay significantly lower median earnings than male-dominated occupations. The figure uses 2010 American Community Survey (ACS) data and the definition of occupation based on the 1990 Census Bureau occupational classification scheme. To focus on the non-college labor structure for young workers, only 18-30 year olds are included in the calculation of worker composition and median earnings percentile.

FIGURE 3

COLLEGE OCCUPATIONS BY GENDER COMPOSITION AND PERCENTILE MEDIAN EARNINGS



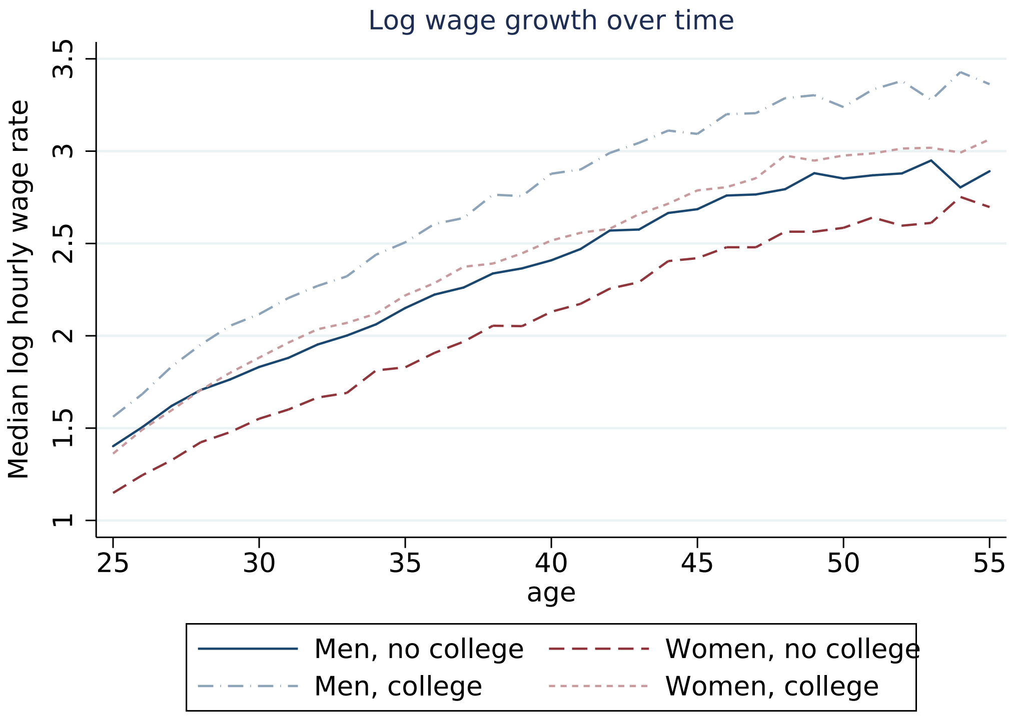
Notes: The figure depicts a scatter plot of all college (at least 50% workers who were college enrollees) occupations by proportion of female workers and median earnings percentile. The figure uses 2010 American Community Survey (ACS) data and the definition of occupation based on the 1990 Census Bureau occupational classification scheme. To focus on the non-college labor structure for young workers, only 18-30 year olds are included in the calculation of worker composition and median earnings percentile.

FIGURE 4

LOG WAGE GAP (WEIGHTED BY OCCUPATION) BETWEEN COLLEGE AND HIGH SCHOOL GRADUATES

Notes: The figure depicts the observed college premium for men and women based on occupational differences alone. Median earnings are calculated by summing over the median earnings of each occupation, weighted by the share of each worker type employed in that occupation (where type is indexed by college enrollment status and sex). Figure 4 demonstrates that among 18-30 year olds, the difference in median wages between college graduates and high school graduates is consistently and substantially larger for women than for men.

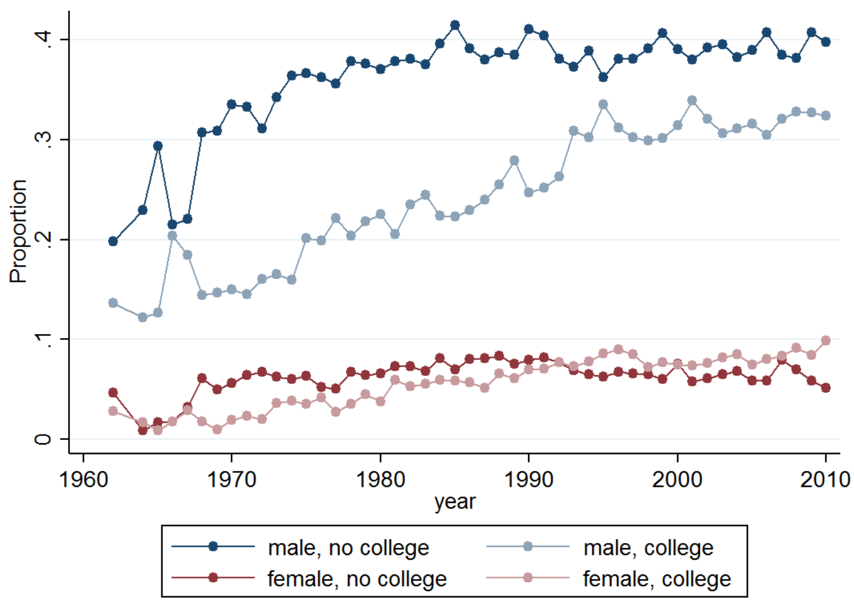
FIGURE 5
MEDIAN ANNUAL EARNINGS BY AGE



Notes: Figure 5 shows the median annual earnings among four occupation categories. All groups exhibit some growth in annual earnings over time, with college male occupations exhibiting the highest earnings at all ages, non-college male occupations making about as much as college female occupations, and non-college female occupations exhibiting the lowest earnings at all ages. The gender gap in annual wages *and* lifetime earnings is smaller among college occupations than non-college occupations. Women appear to face a larger earnings disadvantage in the non-college market than the college market. In addition, in contrast to the other groups, non-college female occupations exhibit almost no earnings growth over the working lives of their workers. This is consistent with the notion that non-college women tend to fill occupations that are not "careers", which typically exhibit some earnings growth with tenure.

FIGURE 6

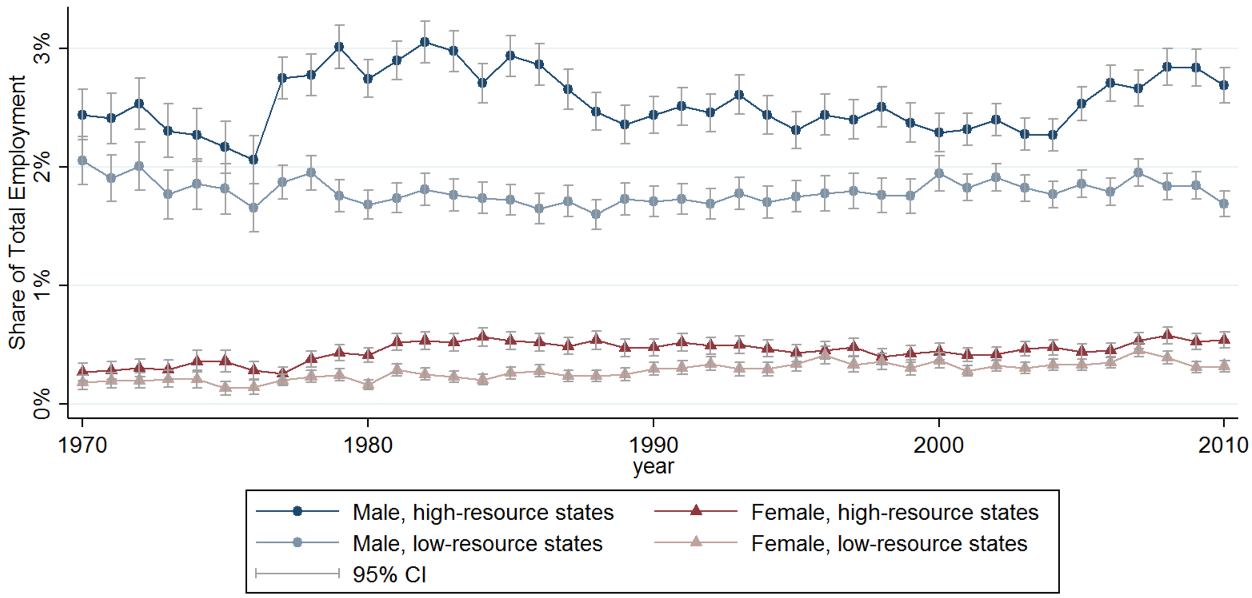
OIL AND GAS EMPLOYMENT, WORKER COMPOSITION



Notes: Figure 6 depicts the composition of workers by gender and education in the oil and gas industry. Male non-college workers comprise most of the workforce in the data period. College and non-college women make up less than 10% of the workforce each. The evidence suggests that male workers would be most affected by changes in the employment demand of the oil and gas industry, since they make up the overwhelming majority of workers in oil, gas, and related industries. This figure uses data from the CPS-ASEC for the years 1970-2010.

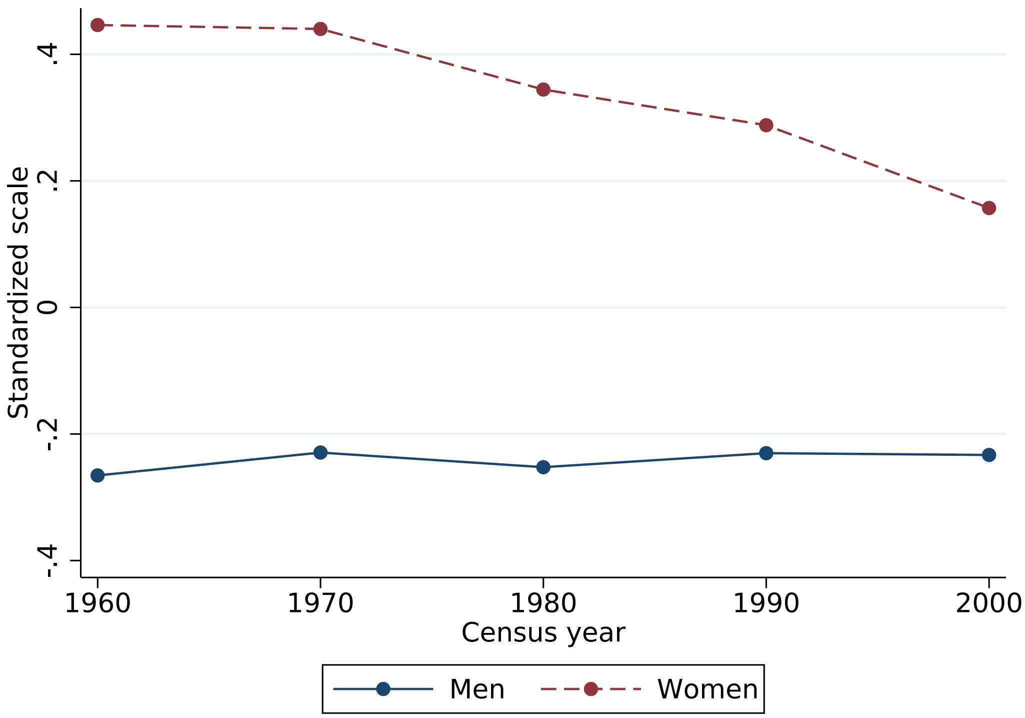
FIGURE 7

SHARE OF EMPLOYMENT IN OIL AND GAS INDUSTRY



Notes: Figure 7 graphs the share of employment in the oil and gas industry by gender and whether the state is a high- or low-resource state. In both high- and low-resource states, employment of men in the oil and gas industry far exceed employment of women. Substantial employment fluctuations are only found among male employment in high-resource states. Employment of men in low-resource states, women in high-resource states, and women in low-resource states remains relatively constant despite booms and busts in the oil and gas industry during this period. The figure provides evidence that natural resources matter in determining employment in the oil and gas industry, and that these natural resources substantially determine the employment rates of men but have little effect on the employment rates of women.

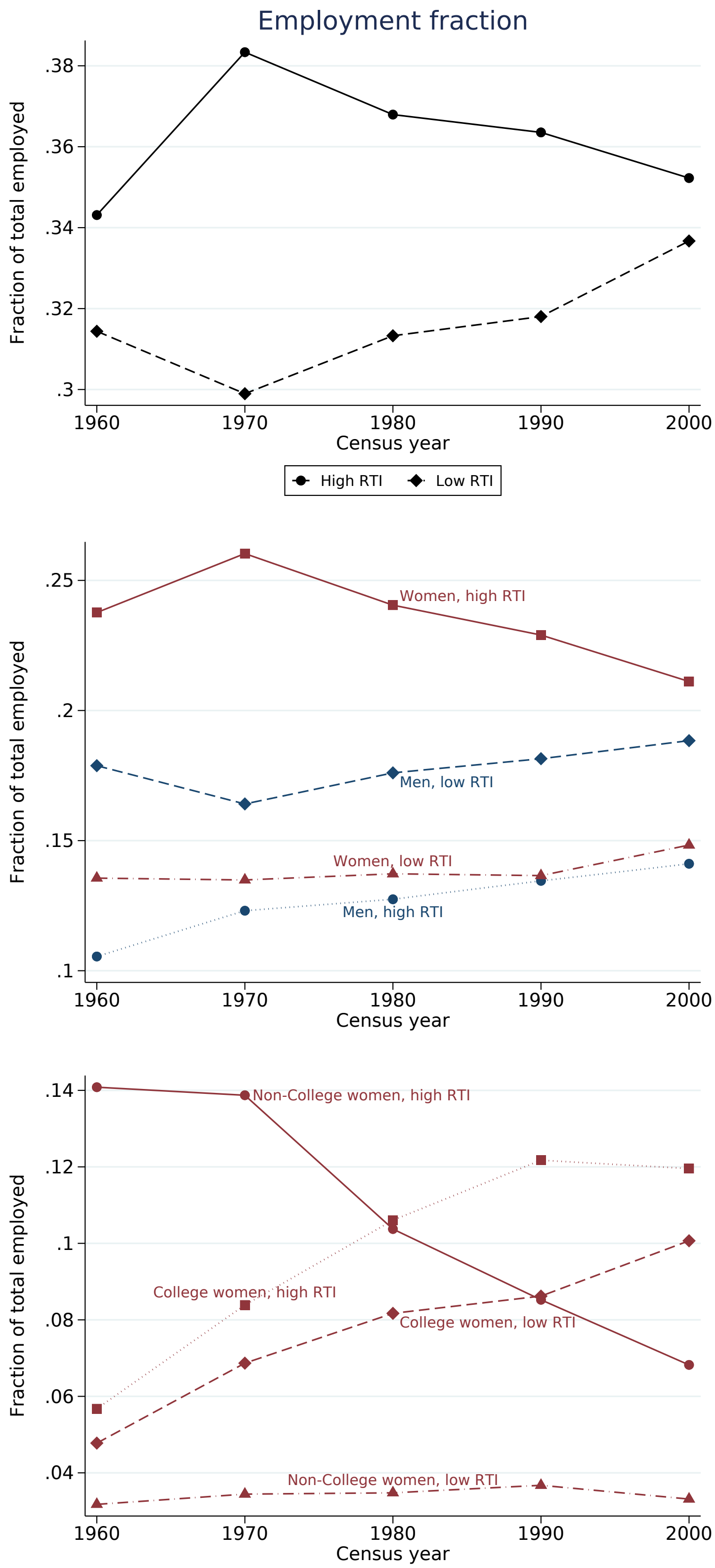
FIGURE 8
ROUTINE TASK INTENSITY (RTI) IN LABOR MARKET OVER TIME



Notes: Figure 8 depicts the measure for routine task intensity (RTI) in the labor force for young men (blue) and young women (red). This figure shows that the displacement of high-RTI jobs by automation fell on women but not men among young workers. The evidence suggests that the employment opportunities of women were most affected by the erosion of routine-intensive jobs. Data obtained from census microdata, ACS data, and the job characteristic measures constructed by Autor and Dorn (2013). Only individuals between the ages of 18 and 30 are included.

FIGURE 9

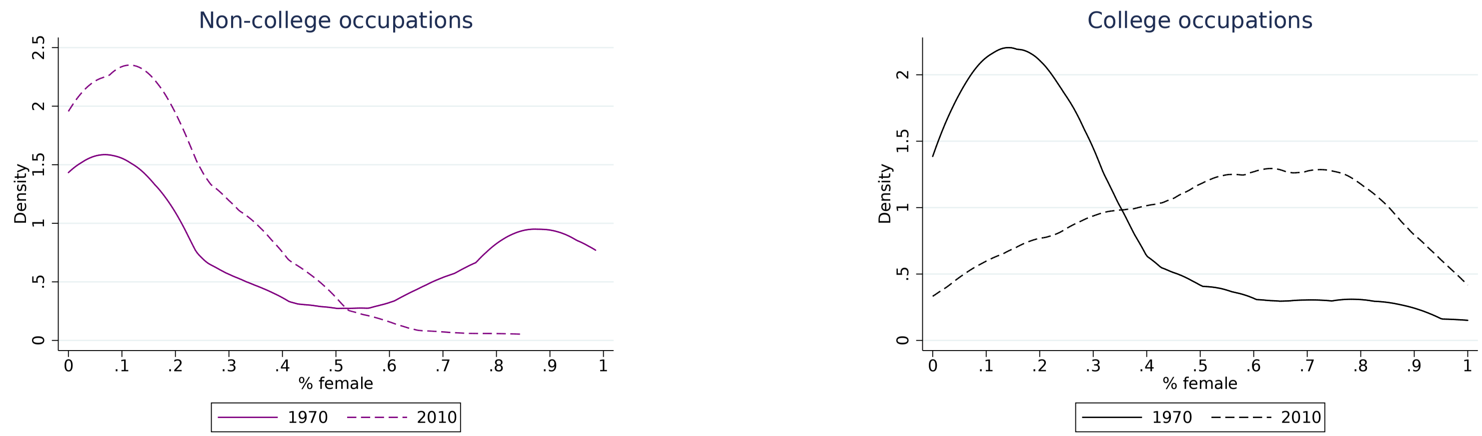
EMPLOYMENT CHANGES OVER TIME, BY ROUTINE-INTENSITY OF JOB TASKS



Notes: Figure 9 graphs the proportion of employed workers by high (top third in 1950) and low (bottom third in 1950) routine task intensity (RTI), the measure of how routine the tasks in an occupation are. The top panel shows that beginning in 1970, the share of high-RTI occupations declined while the proportion of low-RTI occupations increased, consistent with the literature on the decline in routine-intensive tasks over time. The middle panel splits the relationship by gender. Importantly, the decline in highly routine occupations corresponds to a decline in the employment share for *women but not men*. For men, the employment share of high-RTI labor actually increased steadily during this period. Low-RTI labor share increased for both men and women. The bottom panel breaks down this relationship even further by education. The decline in routine-intensive employment is entirely driven by *non-college women in high-RTI occupations*. For all other groups, employment share did not decline. Female college workers gained employment share in both high- and low-RTI jobs. The fraction of non-college women in low-RTI occupations remained relatively unchanged during this period.

FIGURE 10

OCCUPATIONAL DISPERSION FOR COLLEGE AND NON-COLLEGE WORKERS

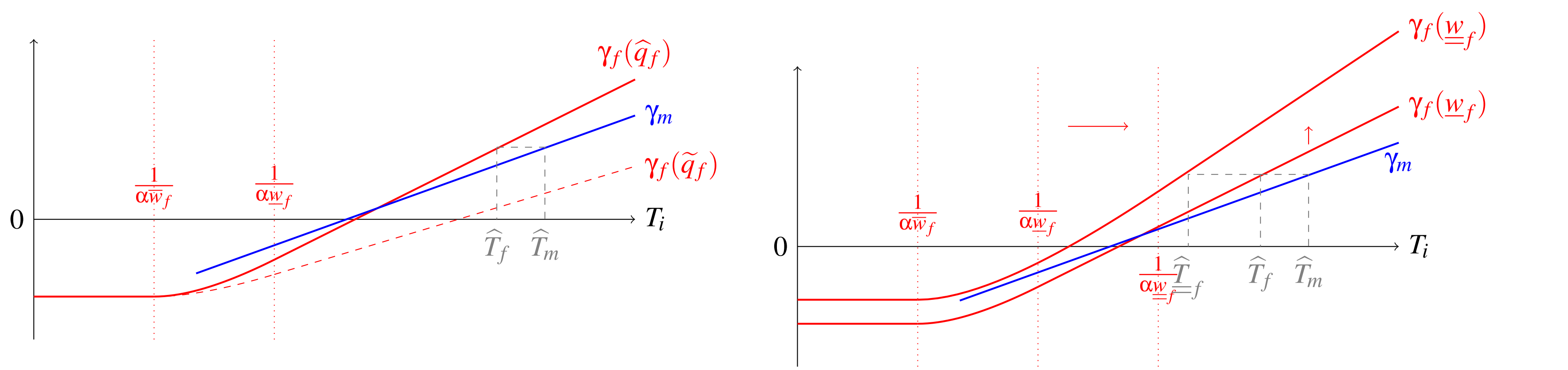


Notes: Figure 10 graphs the distribution of occupations by percent female for 1970 and 2010, for "non-college" occupations (left) and "college" occupations (right). "College" occupations are those with over 50% college workers; "non-college" occupations are those with over 50% non-college workers. 1950 is excluded since few occupations were comprised of over 50% college workers. Only individuals between the ages of 18 and 30 are included.

In 1970, over 60% of occupations employed at least 80% men among both college and non-college occupations, while a much smaller proportion of occupations were female-dominated, and very few occupations employed relatively equal amounts of men and women. The female-dominated occupations all concentrated among non-college occupations. Over time, non-college occupations continued to be segregated by gender, with few occupations that employed relatively equivalent amounts of men and women. However, the female-dominated non-college jobs declined in number, while college jobs became more gender-equitable. In other words, the jobs that used to employ non-college women declined, while the jobs that used to employ a majority of college men became accessible to college women.

FIGURE 11

THRESHOLD COLLEGE-GOING VALUES FOR MEN AND WOMEN



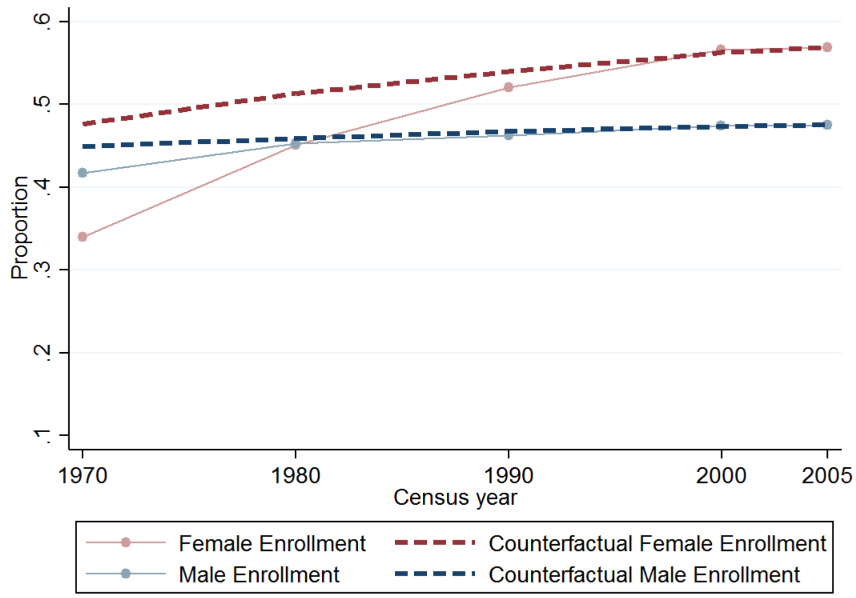
Notes: The left panel of figure 11 illustrates the impact of household production efficiency and the role of declining fertility risk on the enrollment of women relative to men. γ_m represents the growth in male enrollment as men's time net of housework increases, while γ_f represents the growth in female enrollment as women's time net of housework increases. As household production efficiency increases, the time needed to complete housework decreases and individuals have more time to work. The labor market returns to attending college therefore increase, which leads college enrollment to increase. Since women complete the housework required by the family, increasing household efficiency should increase women's enrollment by more than men's enrollment.

Secondly, the decline in fertility risk for women q_f is represented by a shift from $\gamma_f(\widetilde{q}_f$ to $\gamma_f(\widehat{q}_f)$. The figure demonstrates that for high fertility risk $\widetilde{q}_f$, women can never surpass men in college enrollment, while for low fertility risk $\widehat{q}_f$, women can surpass men in college enrollment as long as household production is sufficiently efficient.

The right panel of figure A.3 demonstrates that declines in non-college wages complement household production efficiency and contraceptive technology in accelerating growth in female enrollment. A decline in non-college wage rates from $\underline{w}_f$ to $\underline{\underline{w}}_f$ shifts the γ_f function up and increases the slope of growth. The decline in non-college wage rates therefore makes it possible for women to surpass men in college-going at a higher level of fertility risk and lower level of household production efficiency than before.

FIGURE 12

REAL AND COUNTERFACTUAL COLLEGE ENROLLMENT RATES



Notes: The bold, dashed lines represent the counterfactual college enrollment rates, where changes in college enrollment arise only from changes in the share of non-college jobs over time. The pale solid lines represent the true college enrollment rate. The graph demonstrates that if women had always worked as much as they did in 2010, women would have never lagged behind men in college enrollment.

Tables

TABLE 1
EXAMPLES OF NON-COLLEGE OCCUPATIONS, 2010

Occupation, 1990 basis	Percent of female workers	Earnings percentile	Work has pension plan	Employer-sponsored health insurance
Cashiers	71%	3%	36%	92%
Housekeepers and cleaners	82%	4%	24%	89%
Hairdressers and cosmetologists	88%	9%	11%	60%
Miners	3%	81%	62%	99%
Machinists	4%	60%	62%	97%
Truck, delivery, and tractor drivers	5%	41%	46%	95%

Notes: Table 1 lists examples of non-college occupations. The first three occupations (cashiers, housekeepers and cleaners, hairdressers and cosmetologists) are “traditionally female” occupations. The last three occupations (miners, machinists, and truck drivers) are “traditionally male” occupations. The “female” and “male” occupations differ greatly in terms of annual earnings and whether or not they provide benefits such as retirement pensions and health insurance. “Male” occupations pay between the 40th to the 80th percentile of median annual earnings among all occupations in 2010, while “female” occupations pay below the 10th percentile of median annual earnings. “Male” occupations are significantly more likely to offer a pension retirement plan for workers, and slightly more likely to offer employer-sponsored health insurance. The provision of these benefits indicate that “male” occupations are more likely to have the trappings of careers than “female” occupations.

The proportion of female workers and earnings percentile data were obtained from the 2010 American Community Survey (ACS). The benefits data were obtained from the 2010 Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC). The American Community Survey provides reliable summary statistics regarding work and earnings within occupations. Its measures regarding retirement income and health insurance, however, are too general for the purposes of this paper. In contrast, the CPS-ASEC Supplement asks individuals whether they have retirement income as a result of their employment, separate from survivorship payments, disability benefits, Social Security income, Veterans administration payments or other forms of income. The CPS-ASEC also asks individuals if they are the policyholder for their employer-sponsored health insurance.

TABLE 2

FIRST STAGE REGRESSION: BARTIK INSTRUMENT ON EMPLOYMENT IN OIL & GAS INDUSTRY

	(1) Oil/gas employment	(2) Oil/gas employment, expanded
ln oil/gas boom	0.736***	0.380***
	(0.160)	(0.0842)
Constant	1.734	0.0477
	(1.338)	(0.702)
Observations	1633	1642
R^2	0.117	0.341
F	19.05	74.68

Notes: Regressions at the state-year level. First stage regression of actual oil and gas employment on shift-share prediction of oil and gas employment. Column (2) uses employment in oil, gas, and related industries as dependent variable. State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 3

OLS: RESOURCE ENDOWMENT ON COLLEGE ENROLLMENT

	(1) Female college enrollment	(2) Male college enrollment	(3) College gender gap
ln oil/gas boom	0.0342	-0.115***	-0.149***
	(0.0285)	(0.0262)	(0.0276)
Constant	0.659***	-0.419*	-1.079***
	(0.238)	(0.219)	(0.230)
Observations	1642	1642	1642
R^2	0.639	0.469	0.219

Notes: Regression of shift-share instrument on college enrollment rates. Regressions at the state-year level. Column (1) examines effect on female enrollment rates, column (2) examines effect on male enrollment rates, and column (3) examines effect on the the difference (male enrollment rates less female enrollment rates). State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 4

2SLS REGRESSION: OIL/GAS EMPLOYMENT ON COLLEGE ENROLLMENT

	(1) Female college enrollment	(2) Male college enrollment	(3) College gender gap
ln oil/gas employment	0.0303	-0.136***	-0.166***
	(0.0350)	(0.0413)	(0.0462)
Constant	0.502***	-0.0805	-0.582***
	(0.165)	(0.194)	(0.218)
Observations	1633	1633	1633

Notes: Regression of oil and gas employment on college enrollment rates, using a shift-share instrument for oil and gas employment. Regressions at the state-year level. Column (1) examines effect on female enrollment rates, column (2) examines effect on male enrollment rates, and column (3) examines effect on the difference (male enrollment rates less female enrollment rates). State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 5

2SLS REGRESSION: OIL/GAS EMPLOYMENT ON COLLEGE ENROLLMENT, SUBSAMPLE WITHOUT MIGRATION

	(1) Female college enrollment	(2) Male college enrollment	(3) College gender gap
ln oil/gas employment	0.0334	-0.142***	-0.175***
	(0.0375)	(0.0450)	(0.0507)
Constant	0.502***	-0.0805	-0.582***
	(0.165)	(0.194)	(0.218)
Observations	1633	1633	1633

Notes: Regression of oil and gas employment on college enrollment rates, using a shift-share instrument. Regressions at the state-year level. Column (1) examines effect on female enrollment rates, column (2) examines effect on male enrollment rates, and column (3) examines effect on the difference (male enrollment rates less female enrollment rates). State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Subsample of workers who did not move. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 6

2SLS REGRESSION: OIL/GAS EMPLOYMENT ON COLLEGE ENROLLMENT, SUBSAMPLE WITHOUT MIGRATION FOR WORK

	(1) Female college enrollment	(2) Male college enrollment	(3) College gender gap
ln oil/gas employment	0.0293	-0.138***	-0.167***
	(0.0352)	(0.0421)	(0.0469)
Constant	0.502***	-0.0805	-0.582***
	(0.165)	(0.194)	(0.218)
Observations	1633	1633	1633

Notes: Regression of oil and gas employment on college enrollment rates, using a shift-share instrument. Regressions at the state-year level. Column (1) examines effect on female enrollment rates, column (2) examines effect on male enrollment rates, and column (3) examines effect on the difference (male enrollment rates less female enrollment rates). State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Subsample of workers who did not migrate for work purposes. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 7

2SLS REGRESSION: OIL, GAS, AND RELATED EMPLOYMENT ON COLLEGE ENROLLMENT

	(1) Female college enrollment	(2) Male college enrollment	(3) College gender gap
ln oil/gas employment, expanded	0.0458	-0.269***	-0.314***
	(0.0682)	(0.0819)	(0.0903)
Constant	0.520**	-0.311	-0.831***
	(0.218)	(0.261)	(0.288)
Observations	1642	1642	1642

Notes: Regression of employment in oil, gas, and related industries on college enrollment rates, using a shift-share instrument. Regressions at the state-year level. Column (1) examines effect on female enrollment rates, column (2) examines effect on male enrollment rates, and column (3) examines effect on the difference (male enrollment rates less female enrollment rates). State-level controls include oil production, gas production, percent female, percent black, and percent by 10-year age bin. Fixed effects for state included. Standard errors (clustered at state level) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 8

FIRST STAGE REGRESSION: BARTIK INSTRUMENT ON ROUTINE EMPLOYMENT SHARE

	(1) Routine employment share	(2) Routine employment share
predicted routine employment share * 1980	-0.123***	-0.132***
	(0.0232)	(0.0265)
predicted routine employment share * 1990	-0.336***	-0.356***
	(0.0272)	(0.0302)
predicted routine employment share * 2000	-0.493***	-0.460***
	(0.0267)	(0.0304)
Observations	2888	2888
F	193.8	127.1
Demo. ctrls		Yes
Commuting zone dummies	Yes	Yes
Year dummies	Yes	Yes

Note: Regressions at the commuting zone-year level. First stage regression of actual routine share employment on Bartik prediction of routine share employment (based on 1950 industry shares). Column (2) adds demographic controls for total population, proportion female, proportion black, proportion Hispanic, and proportion by age bin. Fixed effects for state, year dummies, and region dummies included. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 9

2SLS REGRESSION: EMPLOYMENT OUTCOMES ON COLLEGE ENROLLMENT (18-25)

	(1) % female enrollment	(2) % female enrollment	(3) % male enrollment	(4) % male enrollment	(5) college gender gap	(6) college gender gap
Routine share employment	-0.502**	-0.742***	-0.0423	-0.206	0.460***	0.536***
	(0.200)	(0.191)	(0.182)	(0.175)	(0.118)	(0.148)
Observations	2888	2888	2888	2888	2888	2888
First Stage F-stat	193.8	127.1	193.8	127.1	193.8	127.1
RMSE	0.0554	0.0442	0.0572	0.0463	0.0348	0.0345
Demo. ctrls		Yes		Yes		Yes
Commuting zone dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Effect of declining routine-intensive employment on college enrollment among 18-25 year olds, using Bartik instrument for employment in high-routine occupations. Regressions at the commuting zone-year level. Columns (2), (4), and (6) include demographic controls for total population, proportion female, proportion black, proportion Hispanic, and proportion by age bin. Fixed effects for state, year dummies, and region dummies included. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 10

2SLS REGRESSION: EMPLOYMENT OUTCOMES ON COLLEGE ENROLLMENT (18-30)

	(1) % female enrollment	(2) % female enrollment	(3) % male enrollment	(4) % male enrollment	(5) college gender gap	(6) college gender gap
Routine share employment	-0.351**	-0.476***	-0.172	-0.222	0.179**	0.254**
	(0.164)	(0.158)	(0.173)	(0.168)	(0.0824)	(0.105)
Observations	2888	2888	2888	2888	2888	2888
First Stage F-stat	193.8	127.1	193.8	127.1	193.8	127.1
RMSE	0.0451	0.0355	0.0503	0.0400	0.0275	0.0270
Demo. ctrls		Yes		Yes		Yes
Commuting zone dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Effect of declining routine-intensive employment on college enrollment among 18-30 year olds, using Bartik instrument for employment in high-routine occupations. Regressions at the commuting zone-year level. Columns (2), (4), and (6) include demographic controls for total population, proportion female, proportion black, proportion Hispanic, and proportion by age bin. Fixed effects for state, year dummies, and region dummies included. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 11

AGGREGATE CHANGES IN NON-COLLEGE EMPLOYMENT

Year	Change in non-college jobs for women	Change in non-college jobs for men
1970-1980	-0.079	-0.085
1980-1990	-0.069	-0.072
1990-2000	-0.058	-0.048
2000-2005	-0.017	-0.017

Notes: The table presents the change in the labor share of non-college jobs for women and the change in the log labor share of non-college jobs for men.

References

Aguiar, M., Hurst, E., 2007. Measuring Trends in Leisure: The Allocation of Time over Five Decades. The Quarterly Journal of Economics 122(3): 969-1006.

Allcott, H., Keniston, D., 2015. Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America.

Angrist, J., Pischke, J., 2009. Mostly Harmless Econometrics: An Empiricst's Companion. Princeton, New Jersey: Princeton University Press.

ASVAB Fact Sheet. http://official-asvab.com/docs/asvab_fact_sheet.pdf

Autor, D., 2013. The "Task Approach" to Labor Markets: An Overview. NBER Working Paper 18711.

Autor, D., Dorn, D., 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. American Economic Review 103(5): 1553-1597.

Autor, D., Levy, F., Murnane, R., 2002. Upstairs, Downstairs: Computers and Skills on Two Floors of a Large Bank. Industrial and Labor Relations Review 55(2), 432-447.

Autor, D., Levy, F., Murnane, R., 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. Quarterly Journal of Economics 118(4), 1279-1333.

Autor, D., Wasserman, M., 2013. Wayward Sons: The Emerging Gender Gap in Labor Markets and Education. Third Way, March 2013.

Bailey, M., 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.

Beaudry, P., Lewis, E., 2014. Do Male-Female Wage Differentials Reflect Differences in the Return to Skill? Cross-City Evidence from 1980-2000. American Economic Journal: Applied Economics 6(2), 178-194.

Becker, G., Hubbard, W., Murphy, K., 2010. Explaining the Worldwide Boom in the Higher Education of Women. Journal of Human Capital 4(3), 203-241.

Bertrand, M., Pan, J., 2013. The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. American Economic Journal: Applied Economics 5(1), 32-64.

Berzon, A., 2015. Oil Deaths Rise as Bakken Boom Fades. The Wall Street Journal. March 12, 2015.

Black, S., Spitz-Oener, A., 2010. Explaining Womens Success: Technological Change and the Skill Content of Womens Work. Review of Economics and Statistics 92(1), 187-194.

Blakemore, J., Berenbaum, S., Liben, L., 2009. Gender development. New York: Psychology Press; 2009.

Bresnahan, T., Brynjolfsson, E., Hitt, L., 2002. Information Technology, Workplace Organization, and the demand for Skilled Labor: Firm-Level Evidence. Quarterly Journal of Economics 117(1), 339-376.

Brown, C., 2013. North Dakota Went Boom. The New York Times Magazine. Jan. 31, 2013.

Brynjolfsson, E., Hitt, L., 2000. Beyond Computation: Information Technology, Organizational Transformation and Business Performance. Journal of Economic Perspectives 14(4), 23-48.

Buchmann, C., DiPrete, T., 2006. The Growing Female Advantage in College Completion: The Role of Family Background and Academic Achievement. American Sociological Review 71(4), 515-541.

Bureau of Labor Statistics, U.S. Department of Labor. National Longitudinal Survey of Youth 1979 cohort, 1979-2010 (rounds 1-24). Produced and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH: 2012.

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, 2014-15 Edition.

Cascio, E., Narayan, A., 2015. Who Needs a Fracking Education? The Educational Response to Low-Skill Biased Technological Change. NBER Working Paper 21359.

Charles, K., Luoh, M., 2003. Gender Differences in Completed Schooling. The Review of Economics and Statistics 85(3), 559-577.

Cortes, G., Jaimovich, N., Nekarda, C., Siu, H., 2014. The Micro and Macro of Disappearing Routine Jobs: A Flows Approach. NBER Working Paper 20307.

Cortes, G., Jaimovich, N., Siu, H., 2016. Disappearing Routine Jobs: Who, How, and Why? NBER Working Paper 22918.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic Literature 47(2), 1-27.

Dougherty, C., 2005. Why Are the Returns to Schooling Higher for Women than for Men? The Journal of Human Resources 40(4), 969-988.

Eckel, C., Grossman, P., 2002. Sex and risk: experimental evidence. In: Plott, C., Smith, V., (Eds.), Handbook of Experimental Economics Results, vol. 1 Elsevier, New York.

Economist, 2015. Manhood. May 23, 2015.

Ellis, B., 2011. Double your salary in the middle of nowhere, North Dakota. CNNMoney.

Eligon, J., 2013. An Oil Town Where Men Are Many, and Women Are Hounded. The New York Times. Jan. 15, 2013.

Gebrekidan, S., 2012. Shale boom turns North Dakota into No. 3 Oil Producer. Reuters. Mar. 8, 2012.

Gold, R., 2015. Crude-Oil Price Collapse Takes Toll on Williston. March 12, 2015.

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

Goldin, C., Katz, L., 2010. The Race between Education and Technology. Harvard University Press: Cambridge, Massachusetts.

Goldin, C., Katz, L., Kuziemko, I., 2006. The Homecoming of American College Women: The Reversal of the College Gender Gap. NBER Working Paper Series 12130/

Goos, M., Manning, A., Salomans, A., 2014. Explaining Job Polarization: Routine-Biased Technical Change and Offshoring. American Economic Review 104(8), 2509-2526.

Heckman, J., Stixrud, J., Urzua, S., 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. Journal of Labor Economics 24(3), 411-482.

Hershbein, B., Kahn, L., 2016. Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings.

Jacob, B., 2002. Where the boys aren't: non-cognitive skills, returns to school, and the gender gap in higher education. Economics of Education Review 21, 589-598.

Jaimovich, N., Siu, H., 2012. The Trend is the Cycle: Job Polarization and Jobless Recoveries. NBER Working Paper 18334.

Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. Integrated Public Use Microdata Series, Current Population Survey: Version 3.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

Leyk, D., Gorges, W., Ridder, D., Wunderlich, M., Ruther, T., Sievert, A., Essfeld, D., 2007. Hand-grip strength of young men, women and highly trained female athletes. European Journal of Applied Physiology 99, 415-421.

Manning, A., 2006. CentrePiece. Centre for Economic Performance. 11(1): 1316.

McChesney, J., 2011. Oil Boom Puts Strain on North Dakota Towns. NPR. December 2, 2011.

Miller, A., MacDougall, J., Tarnopolsky, M., Sale, D., 1993. Gender differences in strength and muscle fiber characteristics. European Journal of Applied Physiology and Occupational Physiology 66(3), 254-262.

Neal, D., Johnson, W., 1996. The Role of Premarket Factors in Black-White Wage Differences. Journal of Political Economy 104(5), 869-895.

National Public Radio, 2011. New Boom Reshapes Oil World, Rocks North Dakota. Sept. 25, 2011.

O*Net, 2015. http://www.onetonline.org. Accessed July 7, 2015.

Prada, M., Urzua, S., 2014. One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Wages. NBER Working Paper 20752.

Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

Rosen, S., 1986. The Theory of Equalizing Differences, in *Handbook of Labor Economics,* ed. by Orley Ashenfelter and Richard Layard. Amsterdam: North Holland, pp. 641-692.

Rosin, H., 2010. The End of Men. The Atlantic, July/August 2010 Issue.

Rosen, S., Willis, R., 1979. Education and Self-Selection. Journal of Political Economy 87(5), S7-S36.

Sayer, L., 2014. "Trends in Women's and Men's Time Use, 1965-2012: Back to the Future?" in Gender and Couple Relationships, edited by Susan M. McHale, Valerie King, Jennifer Van Hook, Alan Booth. Pennsylvania State University National Symposium on Family Issues (NSFI) book series, Springer.

Smith, A., 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* London: W. Strahan and T. Cadell.

Spitz-Oener, A., 2006. Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. Journal of Labor Economics 24(2), 235-269.

Spitz-Jener, A., 2008. The Returns to Pencil Use Revisited. Industrial and Labor Relations Review 61(4), 502-517.

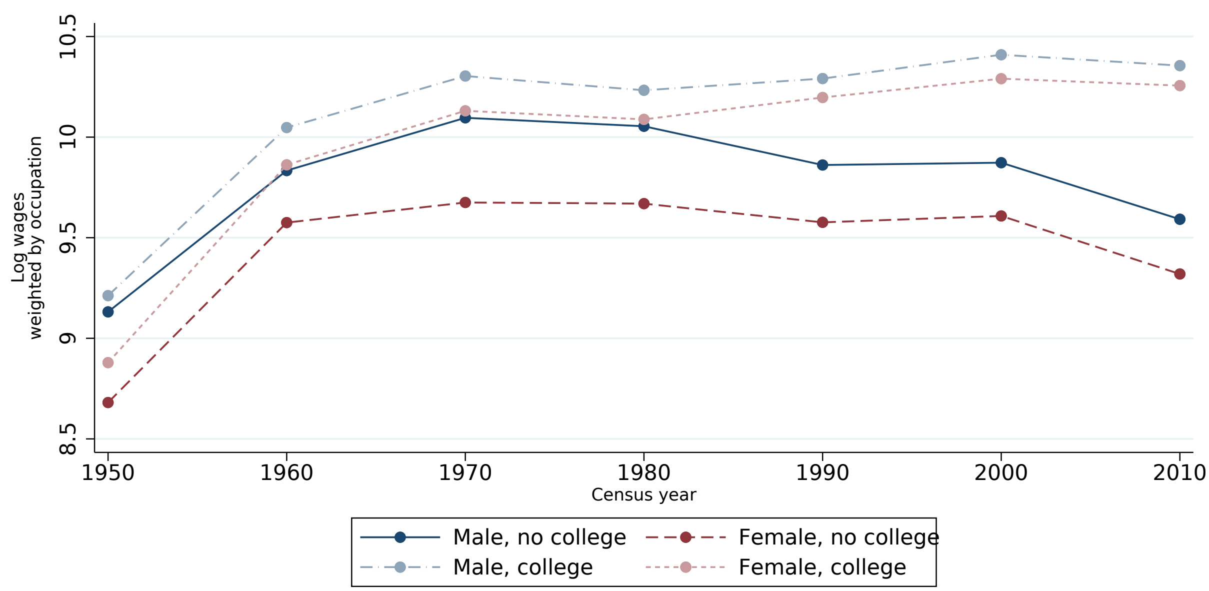
Welch, F., 2000. Growth in Women's Relative Wages and in Inequality among Men: One Phenomenon or Two? American Economic Review 90(2), 444-449.

A Appendix: Tables and Figures

Figures

FIGURE A.1

LOG WAGE GAP (WEIGHTED BY OCCUPATION) BETWEEN COLLEGE AND HIGH SCHOOL GRADUATES



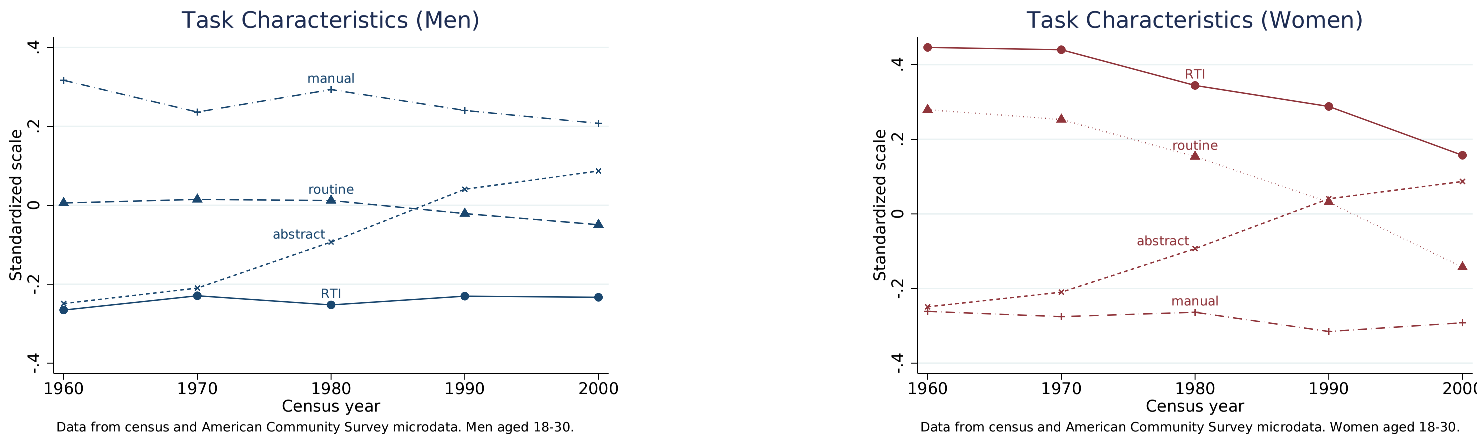
Notes: Figure A.1 depicts median earnings weighted by the labor share of each worker type in an occupation, where worker type is indexed by college status and sex. The explicit formula used to calculate weighted median earnings is

$$\ln \text{median earnings}_{\text{gender,college}} = \ln\left(\sum_{\text{occ}}\left[\text{median}(\text{earnings}_{\text{occ}})\frac{\text{total workers in occ}_{\text{gender,college}}}{\text{total workers}_{\text{gender,college}}}\right]\right)$$

Any differences in median earnings between worker types stems entirely from distributional differences of worker types across occupations. The data come from the decennial census microdata from 1950 to 2000 and from the 2010 American Community Survey (ACS) data. To focus on the discrepancy among young workers, only 18-30 year olds are included.

FIGURE A.2

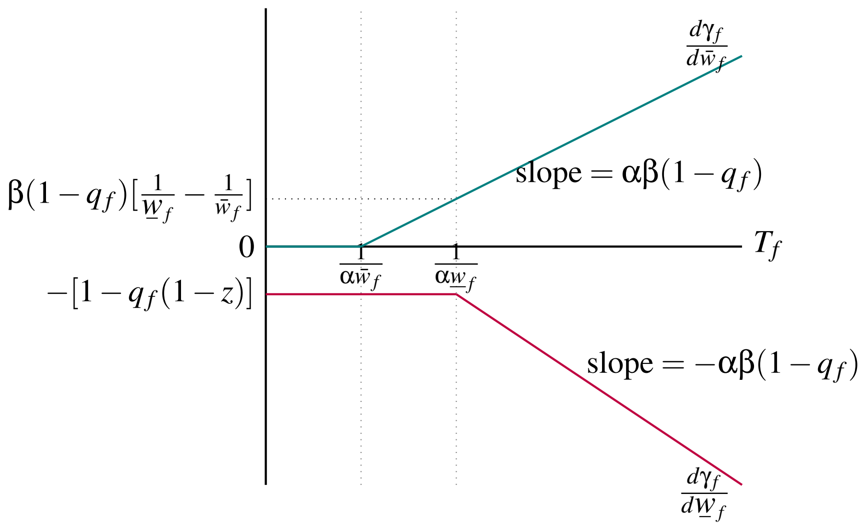
JOB CHARACTERISTICS OVER TIME



Notes: Figure A.2 depicts average routine-intensity, manual-intensity, abstract-intensity, and RTI (the composite measure) in the labor force for young men (left) and young women (right). This figure shows that the displacement of high-RTI jobs by automation fell on women but not men among young workers. The evidence suggests that the employment opportunities of women were most affected by the erosion of routine-intensive jobs. Data obtained from census microdata, ACS data, and the job characteristic measures constructed by Autor and Dorn (2013). Only individuals between the ages of 18 and 30 are included.

FIGURE A.3

EFFECT OF T_f ON RELATIONSHIP BETWEEN WAGE RATES AND COLLEGE-GOING



Notes: The figure depicts the effect of increasing time net of housework, T_f, on the responsiveness of women's college-going to their wage rates. The effect of college wage rates on women's college-going, $\frac{d\gamma_f}{d\bar{w}_f}$, is weakly positive, while the effect of non-college wage rates on women's college-going, $\frac{d\gamma_f}{d\underline{w}_f}$, is strictly negative. When $T_f < \frac{1}{\alpha\bar{w}_f}$, married women do not work and only non-college wage rates influence the college-going decision, since they represent the opportunity costs to attending college in period 1 while single. When $T_f \in [\frac{1}{\alpha\bar{w}_f}, \frac{1}{\alpha\underline{w}_f}]$, only married women with college degrees work, and $\frac{d\gamma_f}{d\bar{w}_f}$ increases with T_f. Finally, when $T_f > \frac{1}{\alpha\underline{w}_f}$, married women without college degrees also join the work force and $\frac{d\gamma_f}{d\underline{w}_f}$ becomes more negative with increasing T_f.

Tables

TABLE A.1
EMPLOYMENT CHANGES IN OIL AND GAS INDUSTRY

	Oil and Gas Industry		Oil, Gas, and Related Industries	
	1970-1980	2000-2010	1970-1980	2000-2010
Non-college men	1.18%	0.42%	4.18%	1.65%
College men	0.73%	0.09%	0.81%	-0.63%
Non-college women	0.02%	-0.07%	0.29%	0.07%
College women	0.24%	0.13%	0.55%	-0.06%

Notes: Table A.1 shows the change in employment in the oil and gas industry or in the oil, gas, and related industries as a proportion of total employment for men and women based on college status. The table examines two decades in which national growth in oil and gas employment was large: 1970-1980 and 2000-2010. In both of of these periods, non-college men experienced much larger changes in employment share in the relevant industries than college men, college women, or non-college women. This provides further evidence that employment changes in these industries impact the non-college labor market returns of men.

TABLE A.2
SUMMARY STATISTICS BY STATE RESOURCE LEVEL, 1970

	Low-Resource	High-Resource	Difference
Female college enrollment	0.347	0.374	-0.0272
	(0.0173)	(0.0187)	(0.0256)
Male college enrollment	0.423	0.464	-0.0414
	(0.0277)	(0.0127)	(0.0331)
% female	0.518	0.513	0.00586
	(0.00205)	(0.00397)	(0.00421)
% black	0.111	0.122	-0.0111
	(0.0260)	(0.0275)	(0.0381)
% 18-25 years old	0.119	0.126	-0.00731
	(0.00295)	(0.00298)	(0.00425)
% 26-35 years old	0.125	0.118	0.00688
	(0.00343)	(0.00446)	(0.00553)
% 36-45 years old	0.120	0.110	0.00930**
	(0.00274)	(0.00297)	(0.00406)
% 46-55 years old	0.111	0.113	-0.00260
	(0.00395)	(0.00342)	(0.00538)
% 56-65 years old	0.0908	0.0925	-0.00175
	(0.00446)	(0.00574)	(0.00715)
% older than 65 years old	0.0864	0.0885	-0.00209
	(0.00508)	(0.00643)	(0.00807)
Observations	10	8	

Table A.2 reports the results of a t-test comparing the characteristics of high-resource states with the characteristics of low-resource states. Overall, there are few significant differences in observed composition between high- and low-resource states. The proportion of 36-45 year olds is significantly different ($p < 0.05$). These differences are insignificant for almost all other years. Standard errors in parentheses. Stars denote significant differences between high- and low-resource states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A.3
SUMMARY STATISTICS BY STATE RESOURCE LEVEL, 1980

	Low-Resource	High-Resource	Difference
Female college enrollment	0.396	0.400	-0.00456
	(0.00999)	(0.0172)	(0.0187)
Male college enrollment	0.396	0.408	-0.0126
	(0.00878)	(0.0129)	(0.0151)
% female	0.514	0.513	0.000504
	(0.00181)	(0.00229)	(0.00289)
% black	0.0845	0.106	-0.0215
	(0.0156)	(0.0227)	(0.0266)
% 18-25 years old	0.150	0.148	0.00234
	(0.00210)	(0.00315)	(0.00364)
% 26-35 years old	0.157	0.157	0.0000911
	(0.00224)	(0.00386)	(0.00420)
% 36-45 years old	0.117	0.110	0.00706**
	(0.00202)	(0.00207)	(0.00296)
% 46-55 years old	0.103	0.0978	0.00543*
	(0.00207)	(0.00231)	(0.00314)
% 56-65 years old	0.0936	0.0907	0.00285
	(0.00166)	(0.00262)	(0.00296)
% older than 65 years old	0.0964	0.101	-0.00490
	(0.00255)	(0.00454)	(0.00487)
Observations	28	20	

Table A.3 reports the results of a t-test comparing the characteristics of high-resource states with the characteristics of low-resource states. Overall, there are few significant differences in observed composition between high- and low-resource states. The proportion of 36-45 year olds is significantly different ($p < 0.05$), and the proportion of 46-55 year olds is marginally significantly different ($p < 0.10$). Standard errors in parentheses. Stars denote significant differences between high- and low-resource states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A.4
SUMMARY STATISTICS BY STATE RESOURCE LEVEL, 1990

	Low-Resource	High-Resource	Difference
Female college enrollment	0.479	0.479	-0.000114
	(0.0111)	(0.0195)	(0.0211)
Male college enrollment	0.456	0.428	0.0282
	(0.0111)	(0.0201)	(0.0214)
% female	0.520	0.528	-0.00748
	(0.00353)	(0.00528)	(0.00611)
% black	0.0846	0.102	-0.0174
	(0.0176)	(0.0258)	(0.0301)
% 18-25 years old	0.150	0.156	-0.00586
	(0.00370)	(0.00521)	(0.00621)
% 26-35 years old	0.223	0.227	-0.00413
	(0.00616)	(0.00812)	(0.0100)
% 36-45 years old	0.187	0.185	0.00238
	(0.00530)	(0.00496)	(0.00755)
% 46-55 years old	0.130	0.132	-0.00139
	(0.00387)	(0.00398)	(0.00569)
% 56-65 years old	0.113	0.108	0.00568
	(0.00418)	(0.00545)	(0.00675)
% older than 65 years old	0.145	0.137	0.00807
	(0.00570)	(0.00575)	(0.00832)
Observations	28	20	

Table A.4 reports the results of a t-test comparing the characteristics of high-resource states with the characteristics of low-resource states. Overall, there are no significant differences in observed demographic characteristics between high- and low-resource states. Standard errors in parentheses. Stars denote significant differences between high- and low-resource states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A.5

SUMMARY STATISTICS BY STATE RESOURCE LEVEL, 2000

	Low-Resource	High-Resource	Difference
Female college enrollment	0.629	0.611	0.0181
	(0.0170)	(0.0207)	(0.0267)
Male college enrollment	0.569	0.576	-0.00701
	(0.0160)	(0.0175)	(0.0241)
% female	0.516	0.518	-0.00150
	(0.00388)	(0.00389)	(0.00565)
% black	0.0967	0.0991	-0.00236
	(0.0186)	(0.0203)	(0.0279)
% 18-25 years old	0.137	0.153	-0.0159**
	(0.00486)	(0.00492)	(0.00710)
% 26-35 years old	0.172	0.169	0.00290
	(0.00583)	(0.00803)	(0.00967)
% 36-45 years old	0.217	0.207	0.0101
	(0.00501)	(0.00760)	(0.00874)
% 46-55 years old	0.172	0.168	0.00388
	(0.00386)	(0.00436)	(0.00587)
% 56-65 years old	0.115	0.111	0.00375
	(0.00375)	(0.00571)	(0.00655)
% older than 65 years old	0.149	0.151	-0.00229
	(0.00552)	(0.00803)	(0.00941)
Observations	28	20	

Table A.5 reports the results of a t-test comparing the characteristics of high-resource states with the characteristics of low-resource states. Overall, there are few significant differences in observed demographic characteristics between high- and low-resource states. In 2000, the proportion of 18-25 year olds differs significantly between low- and high-resource states ($p < 0.05$). However, this difference is insignificant for all other years. Standard errors in parentheses. Stars denote significant differences between high- and low-resource states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A.6
SUMMARY STATISTICS BY STATE RESOURCE LEVEL, 2010

	Low-Resource	High-Resource	Difference
Female college enrollment	0.659	0.641	0.0178
	(0.0118)	(0.0165)	(0.0197)
Male college enrollment	0.566	0.565	0.00101
	(0.0137)	(0.0169)	(0.0217)
% female	0.517	0.514	0.00285
	(0.00141)	(0.00202)	(0.00239)
% black	0.102	0.102	-0.000169
	(0.0175)	(0.0222)	(0.0279)
% 18-25 years old	0.141	0.143	-0.00209
	(0.00389)	(0.00497)	(0.00623)
% 26-35 years old	0.166	0.165	0.000610
	(0.00484)	(0.00655)	(0.00797)
% 36-45 years old	0.166	0.174	-0.00788
	(0.00325)	(0.00402)	(0.00514)
% 46-55 years old	0.188	0.178	0.0105
	(0.00428)	(0.00492)	(0.00655)
% 56-65 years old	0.152	0.148	0.00403
	(0.00442)	(0.00600)	(0.00728)
% older than 65 years old	0.152	0.156	-0.00404
	(0.00429)	(0.00778)	(0.00830)
Observations	28	20	

Table A.6 reports the results of a t-test comparing the characteristics of high-resource states with the characteristics of low-resource states. Overall, there are no significant differences in observed demographic characteristics between high- and low-resource states. Standard errors in parentheses. Stars denote significant differences between high- and low-resource states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A.7
CORRELATIONS BETWEEN FEMALE AND WORK TYPE

Year	Corr(female no college, routine)	Corr(female college, abstract)
1950	0.142***	0.0433***
1970	0.133***	0.107***
1980	0.0775***	0.117***
1990	0.0341***	0.147***
2000	0.00420***	0.156***

Notes: The table presents pairwise correlations of worker type and task-intensity. Non-college women are significantly more likely to work in routine-intensive occupations, but this likelihood declines over time. College women are significantly more likely to work in abstract-intensive occupations, and this likelihood increases over time. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Appendix: Theory

B.1 College-going decision for men

The assumption $T_m > \frac{1}{\alpha \underline{w}_m}$ guarantees that married men will always work in period 2, and the assumption $T > \frac{1}{\underline{w}_m} + 1$ guarantees that single men will have time for both leisure and schooling in period 1. The marginal returns to leisure are strictly decreasing while the marginal costs are constant, guaranteeing a unique solution to the utility maximization problem defined by equations (5.1), (5.2), and (5.3) for men.

Proposition (Proposition B.1)**.** *The schooling rule for men can be defined by a threshold value γ_m, such that men will attend college if and only if they draw non-monetary costs ε where $\varepsilon < \gamma_m$.*

$$s_m = \mathbf{1}\left[\gamma_m > \varepsilon\right] \tag{B.1}$$

where $\gamma_m = \beta\left[\alpha T_m[\bar{w}_m - \underline{w}_m] + ln\left(\frac{\underline{w}_m}{\bar{w}_m}\right)\right] - \underline{w}_m - d$

oof For men, the indirect utility from choosing to attend college is given by

$$V_1(s=1,z=1) = \beta\left[\alpha\bar{w}_m[T_m - \frac{1}{\alpha\bar{w}_m}] + \ln\left(\frac{1}{\alpha\bar{w}_m}\right)\right] + \underline{w}_m\left[T - \frac{1}{\underline{w}_m} - 1\right] - d - \varepsilon + \ln\left(\frac{1}{\underline{w}_m}\right) \tag{B.2}$$

The indirect utility from choosing to not attend college is given by

$$V_1(s=0,z=1) = \beta\left[\alpha\underline{w}_m\left[T_m - \frac{1}{\alpha\underline{w}_m}\right] + \ln\left(\frac{1}{\alpha\underline{w}_m}\right)\right] + \underline{w}_m\left[T - \frac{1}{\underline{w}_m}\right] + \ln\left(\frac{1}{\underline{w}_m}\right) \tag{B.3}$$

Men will attend college if and only if $V_1(s=1,z=1) - V_1(s=0,z=1) > 0$. Taking the difference between (B.2) and (B.3), we have

$$V_1(s=1,z=1) - V_1(s=0,z=1) = \beta\left[\alpha T_m[\bar{w}_m - \underline{w}_m] + \ln\left(\frac{\underline{w}_m}{\bar{w}_m}\right)\right] - \underline{w}_m - d - \varepsilon \tag{B.4}$$

The schooling rule is therefore defined as

$$s_m = \mathbf{1}\left[\beta\left[\alpha T_m[\bar{w}_m - \underline{w}_m] + \ln\left(\frac{\underline{w}_m}{\bar{w}_m}\right)\right] - \underline{w}_m - d > \varepsilon\right] \tag{B.5}$$

Equation (B.1) demonstrates that men choose to attend college if and only if their future discounted additional earnings in period 2 exceeds the future discounted loss in utility from less leisure in period 2, their foregone earnings from attending college in period 1, and the total (monetary and non-monetary) college costs in period 1.

B.2 College-going decision for women

The focus of this model is the change in labor force participation of married women over time, so for ease of exposition I assume that single women with and without children have time for both leisure and schooling in period 1 ($T > \frac{1}{\underline{w}_f} + 1$, $T'_f \geq \frac{1}{\underline{w}_f} + 1$). Married women may or may not work in period 2. If they do work, female college graduates will work strictly more than women without college degrees.[31]

Proposition (Proposition B.2)**.** *The college enrollment rule for women is given by the threshold value γ_f, such that women will attend college if and only if they draw non-monetary costs ε where $\varepsilon < \gamma_f$. Threshold value γ_f takes on different values depending on whether or not it is optimal for married women to work.*

$$s_f = \mathbf{1}\left[\gamma_f > \varepsilon\right] \tag{B.6}$$

where

$$\gamma_f = \begin{cases} -[1-q_f+q_f z)](\underline{w}_f + d) \\ \textit{if wives do not work } (T_f \leq \frac{1}{\alpha \overline{w}_f}) \\ \\ \beta(1-q_f)\left[\alpha \overline{w}_f T_f - 1 - ln(\alpha \overline{w}_f T_f)\right] - (1-q_f+q_f z)(\underline{w}_f + d) \\ \textit{if only female college graduates work } (\frac{1}{\alpha \overline{w}_f} < T_f \leq \frac{1}{\alpha \underline{w}_f}) \\ \\ \beta(1-q_f)\left[\alpha T_f(\overline{w}_f - \underline{w}_f) + ln\left(\underline{w}_f / \overline{w}_f\right)\right] - (1-q_f+q_f z)(\underline{w}_f + d) \\ \textit{if all wives work } (\frac{1}{\alpha \underline{w}_f} < T_f) \end{cases}$$

Proof. For women, the indirect utility from attending college depends on whether or not it is optimal for them to work when married in period 2.

Case 1: It is not optimal for married women to work in period 2 ($T_f \leq \frac{1}{\alpha \overline{w}_f}$). The indirect utility from choosing to attend college is

$$\begin{aligned} \mathbb{E}V(1,z) = (1-q_f)\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - 1) - d - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha k + \ln\big(T_f\big)\big]\Big] + \\ q_f\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - z) - dz - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha k + \ln\big(T_f\big)\big]\Big] \end{aligned} \tag{B.7}$$

The indirect utility from choosing to not attend college is

$$\mathbb{E}V(0,z) = \underline{w}_f(T - \frac{1}{\underline{w}_f}) + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha k + \ln\big(T_f\big)\big] \tag{B.8}$$

[31]This stems from the assumption that utility is quasilinear in consumption and that women with college degrees earn higher wage rates than those without: $\overline{w}_f > \underline{w}_f$.

Subtracting B.8 from B.7, provides the schooling rule:

$$s_f = \mathbf{1}\left[\mathbb{E}V(1,z) - \mathbb{E}V(0,z) = -[1-q_f(1-z)][\underline{w}_f + d] - \varepsilon > 0\right] \tag{B.9}$$

Case 2: It is only optimal for married women with college degrees to work in period 2 ($\frac{1}{\alpha\bar{w}_f} \leq T_f < \frac{1}{\alpha\underline{w}_f}$). The indirect utility from choosing to attend college is

$$\mathbb{E}V(1,z) = (1-q_f)\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - 1) - d - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha[\bar{w}_f(T_f - \frac{1}{\alpha\bar{w}_f}) + k] + \ln\big(\frac{1}{\alpha\bar{w}_f}\big)\big]\Big] +$$
$$q_f\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - z) - dz - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha k + \ln\big(T_f\big)\big]\Big] \tag{B.10}$$

The indirect utility from choosing to not attend college is

$$\mathbb{E}V(0,z) = \underline{w}_f(T - \frac{1}{\underline{w}_f}) + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha k + \ln\big(T_f\big)\big] \tag{B.11}$$

The difference in indirect utilities obtained from subtracting B.11 from B.10 is given by

$$s_f = \mathbf{1}\Big[\mathbb{E}V(1,z) - \mathbb{E}V(0,z) =$$
$$\beta(1-q_f)\big[\alpha\bar{w}_f(T_f - \frac{1}{\alpha\bar{w}_f}) - \ln(\alpha\bar{w}_f T_f)\big] - [1-q_f(1-z)][\underline{w}_f + d] - \varepsilon > 0\Big] \tag{B.12}$$

Case 3: It is optimal for married women to work in period 2 ($T_f > \frac{1}{\alpha\underline{w}_f}$). The indirect utility from choosing to attend college is given by

$$\mathbb{E}V(1,z) = (1-q_f)\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - 1) - d - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha[\bar{w}_f(T_f - \frac{1}{\alpha\bar{w}_f}) + k] + \ln\big(\frac{1}{\alpha\bar{w}_f}\big)\big]\Big] +$$
$$q_f\Big[\underline{w}_f(T - \frac{1}{\underline{w}_f} - z) - dz - \varepsilon + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha[\underline{w}_f(T_f - \frac{1}{\alpha\underline{w}_f}) + k] + \ln\big(\frac{1}{\alpha\underline{w}_f}\big)\big]\Big] \tag{B.13}$$

The indirect utility from choosing to not attend college is given by

$$\mathbb{E}V(0,z) = \underline{w}_f(T - \frac{1}{\underline{w}_f}) + \ln\big(\frac{1}{\underline{w}_f}\big) + \beta\big[\alpha[\underline{w}_f(T_f - \frac{1}{\alpha\underline{w}_f}) + k] + \ln\big(\frac{1}{\alpha\underline{w}_f}\big)\big] \tag{B.14}$$

The difference in indirect utilities obtained from subtracting B.14 from B.13 yields

$$s_f = \mathbf{1}\left[\mathbb{E}V(1,z) - \mathbb{E}V(0,z) = \beta(1-q_f)\left[\alpha T_f(\bar{w}_f - \underline{w}_f) + \ln\left(\frac{\underline{w}_f}{\bar{w}_f}\right)\right] - [1-q_f(1-z)][\underline{w}_f + d] - \varepsilon > 0\right] \tag{B.15}$$

$\square$

Based on Proposition B.2, the schooling rule for women will always depend on women's non-college wage rates, even if they will not work in period 2. However, women's schooling rule does not depend on college wage rates if it is not optimal for female college graduates to work.

B.3 Effect of wage rates on college-going

Men choose to attend college if and only if they draw non-monetary costs $\varepsilon \sim G(\varepsilon)$ below γ_m, and women choose to attend college if and only if they draw $\varepsilon \sim G(\varepsilon)$ below γ_f. Proposition B.3 summarizes how these threshold values change given changes in non-college wage rates.

Proposition (Proposition B.3)**.** *The effect of non-college wage rates in decreasing college-going is strictly larger in magnitude for men than women.*

$$0 > \frac{d\gamma_f}{d\underline{w}_f} > \frac{d\gamma_m}{d\underline{w}_m} \tag{B.16}$$

where $\frac{d\gamma_m}{d\underline{w}_m} = -\alpha\beta\big[\underbrace{T_m - \frac{1}{\alpha\underline{w}_m}}_{h^*_{2m}(sz<1)}\big] - 1$, *and* $\frac{d\gamma_f}{d\underline{w}_f} = -\alpha\beta(1-q_f)max\big[\underbrace{T_f - \frac{1}{\alpha\underline{w}_f}}_{h^*_{2f}(sz<1)}, 0\big] - [1-q_f(1-z)]$

Proof. Taking the derivative of the threshold college-going value for men, γ_m, with respect to men's non-college wage rates, $\underline{w}_m$, we have

$$\frac{d\gamma_m}{d\underline{w}_m} = -\alpha\beta\big[\underbrace{T_m - \frac{1}{\alpha\underline{w}_m}}_{h^*_{2m}(sz<1)}\big] - 1 \tag{B.17}$$

Taking the derivative of the threshold college-going value for women γ_f with respect to women's non-college wage rates, $\underline{w}_f$, we have

$$\frac{d\gamma_f}{d\underline{w}_f} = \begin{cases} -[1-q_f(1-z)] & \text{if } T_f \leq \frac{1}{\alpha\underline{w}_f} \\ -\alpha\beta(1-q_f)\big[\underbrace{T_f - \frac{1}{\alpha\underline{w}_f}}_{h^*_{2f}(sz<1)}\big] - [1-q_f(1-z)] & \text{if } T_f > \frac{1}{\alpha\underline{w}_f} \end{cases} \tag{B.18}$$

Optimal time spent at work h^*_2 is increasing in wage rates $\underline{w}$, so $\underline{w}_f < \underline{w}_m$ implies that $h^*_{2f} < h^*_{2m}$. Moreover, $(1-q_f) < 1$ and $1-q_f(1-z) < 1$. Therefore, each term in equation (B.18) is smaller in magnitude than each term in equation (B.17) and

$$\frac{d\gamma_f}{d\underline{w}_f} < \frac{d\gamma_m}{d\underline{w}_m} < 0 \tag{B.19}$$

$\square$

The effect of non-college wages in decreasing college-going rates is stronger for men than women, for three reasons. First, women's non-college wage rates are lower than men's, which decreases the amount of labor time women optimally *choose* to supply. Second, married women's time net of housework is less than married men's, which leaves them with less time they *can* convert to labor. Third, the probability of an unplanned pregnancy decreases the likelihood that women who attend college will earn a college degree, which decreases the period 2 expected earnings gain from attending college. The effect of any increase in non-college wage rates on college-going is smaller given this lower expected earnings gain. On an additional interesting side note, the risk of leaving school having only finished partway *decreases* the period 1 cost of college for women, by decreasing the earnings women expect to forego while in school. The combination of all these factors make women's college-going decisions less responsive to non-college wage rates than men's college-going.

On the other hand, the effect of *college* wage rates on college-going is positive, strictly for men and weakly for women. Much of the analysis for college wage rates mirror that for non-college wage rates, so explicit expressions for the effects of college wage rates are relegated to the appendix. There are two results to highlight. First, men's college-going decision is more responsive to their college wage rates than women's. Second, women's college wage rates have no effect on women's college-going if female college graduates do not work in period 2, but women's non-college wage rates always have a negative effect on women's college-going (even if they do not work in period 2). These results are formalized in proposition B.4.

Proposition (Proposition B.4)**.** *The effect of college wage rates in increasing college-going is strictly larger in magnitude for men than women.*

$$\frac{d\gamma_m}{d\bar{w}_m} > \frac{d\gamma_f}{d\bar{w}_f} \geq 0 \tag{B.20}$$

where the last relationship holds with equality if married female college graduates do not work ($T_f \leq \frac{1}{\alpha\bar{w}_f}$).

Proof. Taking the derivative of the threshold college-going value for men, γ_m, with respect to men's college wage rates, $\bar{w}_m$, we have

$$\frac{d\gamma_m}{d\bar{w}_m} = \alpha\beta[\underbrace{T_m - \frac{1}{\alpha\bar{w}_m}}_{h^*_{2m}(sz=1)}] > 0 \tag{B.21}$$

Taking the derivative of the threshold college-going value for women, γ_f, with respect to women's college wage rates, $\bar{w}_f$, we have

$$\frac{d\gamma_f}{d\bar{w}_f} = \begin{cases} 0 & \text{if } T_f \leq \frac{1}{\alpha\bar{w}_f} \\ \alpha\beta(1-q_f)\underbrace{\left[T_f - \frac{1}{\alpha\bar{w}_f}\right]}_{h^*_{2f}(sz=1)} & \text{if } T_f > \frac{1}{\alpha\bar{w}_f} \end{cases} \tag{B.22}$$

Since it is assumed that $T_m > \frac{1}{\alpha\bar{w}_m}$, the amount of time men spend at work is always positive. Therefore, the effect of college wage rates on the college-going threshold γ_m for men is strictly positive. If it is optimal for women with college degrees to work $(T_f > \frac{1}{\alpha\bar{w}_f})$, the effect of their college wage rates on their college-going threshold γ_f is also positive. However, because 1) $\bar{w}_f < \bar{w}_m$ and optimal time spent at work is increasing in wage rates, and because 2) women have less time net of household production than men do $T_f < T_m$, women will spend strictly less time at work than men: $h^*_{2f}(sz=1) = T_f - \frac{1}{\alpha\bar{w}_f} = T_m - \frac{1}{\alpha\bar{w}_m} = h^*_{2m}(sz=1)$. If it is not optimal for married women with college degrees to work $(T_f \leq \frac{1}{\alpha\bar{w}_f})$, then the effect of college wage rates $\bar{w}_f$ on γ_f is 0. □

B.4 Roles of increasing household production efficiency and declining fertility risk

B.4.1 Increasing household production efficiency

Increases in the efficiency of household production decrease the time needed for housework. Since wives complete all the housework needed by the family, this is modeled as an increase in T_f, the time net of household production that wives can allocate to leisure and market labor in period 2. In this section I will explore how increases in T_f influence the effect of wages on college-going for women.

Figure A.3 illustrates how the effect of wage rates on college-going changes as T_f increases. For values of T_f below $\frac{1}{\alpha\bar{w}_f}$, wives do not work in period 2. Only non-college wage rates influence enrollment, since women must still forego earnings in order to attend school in period 1 even if they did not expect to work in period 2. As T_f increases past $\frac{1}{\alpha\bar{w}_f}$, it becomes optimal for only wives with college degrees to work. The effect of college wage rates on female enrollment grows discontinuously from 0 to $\alpha\beta(1-q_f)[T_f - \frac{1}{\alpha\bar{w}_f}]$, since women now experience an expected earnings gain in period 2 from obtaining a college degree. As T_f increases past $\frac{1}{\alpha\underline{w}_f}$, it becomes optimal for wives without college degrees to join wives with college degrees in the workforce. The negative effect of non-college wages on female enrollment jumps from $-[1-q_f(1-z)]$ to $-\alpha\beta(1-q_f)[T_f - \frac{1}{\alpha\underline{w}_f}] - [1-q_f(1-z)]$. If women expect to work even without a college degree, pursuing a college degree now entails sacrificing their non-college earnings, so the effect of non-college wage rates is stronger than the case where women don't expect to earn anything if they did not attend college.

B.4.2 Declining fertility risk

I will next examine how changes in the probability of an unplanned pregnancy, q_f, influences the response of female enrollment to changes in college and non-college wage rates. Taking the derivatives of $\frac{d\gamma_f}{d\underline{w}_f}$ and $\frac{d\gamma_f}{d\bar{w}_f}$ with respect to q_f, we have

$$\frac{d^2\gamma_f}{d\underline{w}_f dq_f} = \alpha\beta \max[T_f - \frac{1}{\alpha\underline{w}_f}, 0] + 1 - z \text{ where } z < 1 \tag{B.23}$$

$$\frac{d^2\gamma_f}{d\bar{w}_f dq_f} = -\alpha\beta \max[T_f - \frac{1}{\alpha\bar{w}_f}, 0] \tag{B.24}$$

As access to contraceptive technologies expanded, the probability of an unplanned pregnancy, q_f, declined. Equations (B.23) and (B.24) show that this increased the role of expected wage rates in the college-going decisions of women. First, declining fertility risk increased the probability that women who enrolled in college would actually obtain a college degree. Consequently, the expected earnings gain to attending college rose. Any changes in wage rates would have a larger effect on the expected earnings gain to attending college, as shown by both equations (B.23) and (B.24). Second, women expected to spend more time in college since they had a lower probability of dropping out. Any changes in non-college wage rates would then have a larger effect in changing the total foregone earnings of college enrollment, as shown by the second term in equation (B.23).

As contraceptive technologies decreased the probability of unplanned pregnancies, college and non-college wage rates became increasingly important in the college-going decision of women. First, declining fertility risk increased the expected earnings gain from attending college, because women had higher expectations that they would complete college and receive wage rate $\bar{w}_f$ in period 2. A decline in non-college wage rates or an increase in college wage rates would therefore have a larger impact in increasing the expected period 2 earnings gains of attending college, as shown by equations (B.23) and (B.24). Second, the lower likelihood of leaving school partway means women expected to spend more time in college. Any declines in non-college wage rates had a larger impact in decreasing the earnings women must forego to attend college, as shown by equation (B.23). [32]

B.5 When is enrollment higher for men than women? When is enrollment higher for women than men?

Figure 11 delivers the final result of the model. The left panel summarizes the role of increasing housework efficiency and declining fertility risk on how wage rates affect women's college-going. The x-axis is T_i, time net of housework for gender i. The figure depicts how the threshold college-going value for gender i, γ_i, changes as T_i increases.

The effect of increasing household efficiency on how female college-going responds to

[32] In addition, equation (B.23) demonstrates that fertility risk influences the effect of non-college wage rates whether or not women work ($\frac{d^2\gamma_f}{d\underline{w}_f dq_f} > 0$), while equation (B.24) demonstrates that fertility risk only influences the effect of college wage rates when women with college degrees work ($\frac{d^2\gamma_f}{d\bar{w}_f dq_f} < 0$ if and only if $T_f > \frac{1}{\alpha\bar{w}_f}$).

wage rates is represented by increasing time net of housework for women, $T_i = T_f$. Female college-going threshold γ_f grows as T_f increases, represented by right-ward movement along the x-axis. This growth stems entirely from the result that increasing T_f increases the strength of the college-going response to wage rates. This growth is discontinuous, depending on the relationship between time net of housework T_f and wage rates $(\underline{w}_f, \overline{w}_f)$. The figure also graphs male enrollment, γ_m which grows as time net of housework for men $T_i = T_m$ increases (represented by a right-ward shift along the same x-axis).

The effect of declining fertility is represented by the shift from $\gamma_f(\widetilde{q}_f)$ to $\gamma_f(\widehat{q}_f)$, where $\widetilde{q}_f > 1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f} > \widehat{q}_f$. Again, $1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f}$ is the threshold below which it is possible for female enrollment to surpass male enrollment. For this reason $\gamma_f(\widehat{q}_f)$ crosses γ_m, but $\gamma_f(\widetilde{q}_f)$ never crosses γ_m.

Proposition B.5, which is the same as Proposition 1 in the main text, summarizes the conditions which create gender differences in college enrollment.

Proposition (Proposition B.5a)**.** *If $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$, there exists a $T_i = T_{mf}$ where $\gamma_f(q_f, T_{mf}) = \gamma_m(T_{mf})$.*

Men will exceed women in college enrollment if $q_f > 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$ or if $T_f < T_{mf}$ and $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$.

Proposition (Proposition B.5b)**.** *If $q_f < 1 - \frac{\overline{w}_m - \underline{w}_f}{\overline{w}_f - \underline{w}_f}$, there exists a $T_i = T_{mf}$ where $\gamma_f(q_f, T_{mf}) = \gamma_m(T_{mf})$. For any arbitrary $\widehat{T}_m > T_{mf}$, there exists $\widehat{T}_f < \widehat{T}_m$ where $\gamma_f(q_f, \widehat{T}_f) = \gamma_m(\widehat{T}_m)$. Then, $\forall T_f \in (\widehat{T}_f, \widehat{T}_m)$, $\gamma_f(q_f, T_f) > \gamma_m(\widehat{T}_m)$.*

Women will exceed men in college enrollment if women experience fertility risk q_f, time net of housework T_f, and men experience time net of housework T_m, where $q_f < 1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f}$, $T_m = \widehat{T}_m > T_{mf}$, and $T_f \in (\widehat{T}_f, \widehat{T}_m)$.

Proof. The slope of γ_m is $\alpha\beta[\bar{w}_m - \underline{w}_m]$. For $T_f > \frac{1}{\alpha \underline{w}_f}$, the slope of γ_f is $\alpha\beta(1 - q_f)[\bar{w}_f - \underline{w}_f]$. For ease of exposition, we have assumed that α and β are the same between men and women. However, because $\bar{w}_f - \underline{w}_f > \bar{w}_m - \underline{w}_m$, there exists a $q_f \in (0,1)$ such that the slope of γ_f, $\alpha\beta(1 - q_f)[\bar{w}_f - \underline{w}_f]$, exceeds that of γ_m, $\alpha\beta[\bar{w}_m - \underline{w}_m]$. For a sufficiently large value of T_i, $\gamma_f(T_i) \geq \gamma_m(T_i)$. For any arbitrary $\widehat{T}_m$ above this level, $\gamma_f(\widehat{q}_f, \widehat{T}_m) > \gamma_m(\widehat{T}_m)$ (as depicted by figure 11). Since both γ_m and γ_f are continuous with constant slope, there exists a $\widehat{T}_f < \widehat{T}_m$ such that $\gamma_f(\widehat{q}, \widehat{T}_f) = \gamma_m(\widehat{T}_m)$. For all $T_f \in (\widehat{T}_f, \widehat{T}_m)$, $\gamma_f(\widehat{q}, T_f) > \gamma_m(\widehat{T}_m)$. In other words, there exists an interval $(\widehat{T}_f, \widehat{T}_m)$ such that for all values of time net of housework for women T_f within this interval, the college-going threshold for women γ_f will exceed the college-going threshold for men γ_m at $\widehat{T}_m$. $\square$

Proposition 5b demonstrates that necessary conditions for women to exceed men in college enrollment are that fertility risk q_f must fall below $1 - \frac{\overline{w}_m - \underline{w}_m}{\overline{w}_f - \underline{w}_f}$ and that housework must fall to a point where it is optimal for college women to work (in other words, time net of housework T_f must exceed $\frac{1}{\alpha \overline{w}_f}$). Once these two conditions are met, it is possible for women to take advantage of their higher labor market returns. Because $\overline{w}_f - \underline{w}_f > \overline{w}_m - \underline{w}_m$, the slope of female enrollment γ_f exceeds the slope of male enrollment γ_m. As long as housework time for women is sufficiently low, female college-going will be higher than male college-going even if women have less time for work and lower wage rates than men.

The right panel of figure 11 represents the change in female college-going threshold γ_f given a decline in the female non-college wage rate $\underline{w}_f$. Consider a decline in $\underline{w}_f$ to $\underline{\underline{w}}_f$, which pushes the y-intercept up and shifts the vertical axis $\frac{1}{\alpha \underline{\underline{w}}_f}$ further to the right, increasing the slope of γ_f. This change is represented by the shift from $\gamma_f(\underline{w}_f)$ to $\gamma_f(\underline{\underline{w}}_f)$. Since γ_f both shifted up and now has a steeper slope, there is a lower value of T_{mf} for which $\gamma_f(\underline{\underline{w}}_f, T_{mf}) = \gamma_m(T_{mf})$, as shown by Claim 1 below. For any $\widehat{T}_m$, define $\widehat{T}_f$ and $\underline{\underline{\widehat{T}}}_f$ to be such that $\gamma_m(\widehat{T}_m) = \gamma_f(\underline{w}_f, \widehat{T}_f) = \gamma_f(\underline{\underline{w}}_f, \underline{\underline{\widehat{T}}}_f)$. Claim 2 shows that $\underline{\underline{\widehat{T}}}_f < \widehat{T}_f$.

Claim 1. *A decrease in non-college wage rates $\underline{w}_f$ to $\underline{\underline{w}}_f$ leads the female enrollment threshold, γ_f to intersect the male enrollment threshold, γ_m, at a lower value of T_i.*

Proof. Denote $f_1(T_i) = \gamma_f(\underline{w}_f, T_i) - \gamma_m(T_i)$ and denote $f_2(T_i) = \gamma_f(\underline{\underline{w}}_f, T_i) - \gamma_m(T_i)$. Since $\gamma_f(\underline{\underline{w}}_f, T_i) > \gamma_f(\underline{w}_f, T_i)$, then $f_2(T_i) > f_1(T_i) \forall T_i$. Both $f_1(T_i)$ and $f_2(T_i)$ are strictly increasing, but initially negative. Since $f_2(T_i) > f_1(T_i) \forall T_i$, it intersects the y-axis first at a lower level of T_i.

Let $f_2(T_2) = 0$ and let $f_1(T_1) = 0$. Then $T_2 < T_1$. □

Claim 2. $\underline{\underline{\widehat{T}}}_f < \widehat{T}_f$.

Proof. Choose any $\widehat{T}_m > T_{mf}$. Denote $g_1(T_i) = \gamma_f(\underline{w}_f, T_i) - \gamma_m(\widehat{T}_m)$ and denote $g_2(T_i) = \gamma_f(\underline{\underline{w}}_f, T_i) - \gamma_m(\widehat{T}_m)$. Since $\gamma_f(\underline{\underline{w}}_f, T_i) > \gamma_f(\underline{w}_f, T_i)\ \forall T_i$, $g_2(T_i) > g_1(T_i)\ \forall T_i$. Therefore, $g_2(T_i)$ intersects the y-axis at a lower level of T_i than $g_1(T_i)$.

Let $\widehat{T}_f, \underline{\underline{\widehat{T}}}_f$ be such that $g_1(\widehat{T}_f) = 0$ and $g_2(\underline{\underline{\widehat{T}}}_f) = 0$. Then $\underline{\underline{\widehat{T}}}_f < \widehat{T}_f$. □

This result is significant because it shows that declines in non-college wage rates for women complement increasing housework efficiency and decreasing fertility risk in enabling female enrollment to grow and overtake male enrollment. A decline in non-college wage rates enable female college enrollment to exceed male college enrollment at lower levels of household efficiency and higher levels of fertility risk. Declining employment opportunities in the non-college market therefore help explain not only why women overtook men in college enrollment, but also why the overtaking occurred as early as the 1980s, when household efficiency and declining fertility risk had shown significant advancements but had not yet reached current levels.

B.6 Extension: Gender Wage Gap

Figure 11 reveals that women are willing to accept lower college wage rates than men to enter the college market, due to the large imbalance in non-college wage rates between men and women. The difference between college and non-college wage rates is greater for women ($\overline{w}_f - \underline{w}_f > \overline{w}_m - \underline{w}_m$) and this is sufficient to entice a greater proportion of women to choose to attend college than men, even though college wage rates are higher for men than women ($\overline{w}_m > \overline{w}_f$). The empirical disparity in non-college job prospects, documented in Section 2, may therefore explain why the gender wage gap in college occupations has been so persistent. Since women have access to fewer lucrative options in the non-college labor market, they are willing to enter the college market for lower pay relative to men, and therefore men continue to enjoy greater college earnings than women on average.

One implication of this argument is that external measures to narrow the gender gap in wages will widen the gender gap in college enrollment. Policy interventions to make women's college earnings equal to that of men will lead female college enrollment, γ_f, to increase relative to male college enrollment, γ_m. For symmetric reasons, external measures to narrow the gender gap in college enrollment will widen the gap in wages. The model suggests that gender differences in the non-college labor market link the gender gap in college enrollment with the gender gap in college earnings. Current policies aimed at leveling one inequality may exacerbate the other, since prior research has overlooked the role of the non-college market in contributing to both.

C Appendix: Data

C.1 Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) Data

In Sections 2 and 3, I use data from the Annual Social and Economic Supplement of the Current Population Surveys (CPS-ASEC), which is jointly conducted by the U.S. Census Bureau and the Bureau of Labor Studies and provided by the Integrated Public Use Microdata Series (IPUMS; Flood et al., 2015). I use the sample of 16-64 year olds from the years 1970-2010, although for many measures I restrict the sample to just 18-25 year olds or 18-30 year olds. In most of the analysis, I restrict the sample to full-time (at least 35 hours worked per week), full-year (at least 40 weeks worked per year) workers using the definitions employed by Acemoglu and Autor (2011).

I code college enrollment using the harmonized EDUC variable. Individuals are coded as having enrolled in college if they reported at least 1 year of college or some college but no degree. Individuals are coded as having never enrolled in college if they report having no more than a high school diploma or equivalent. Individuals who did not report an education level are excluded from the analysis.

Workers are coded potentially affected by the oil and gas industry if they belong to the following industries: petroleum and coal production, mining (including oil and gas extraction), trucking services, warehousing and storage. Workers are coded as potentially affected by the oil, gas, and related industries if they belonged to any of the aforementioned industries or were employed in the following occupations: construction inspectors; inspectors and compliance officers; metallurgical and materials engineers; petroleum, mining, and geological engineers; chemical engineers; electrical engineers; industrial engineers; mechanical engineers; geologists; drillers of earth, construction trades (not elsewhere classified); extractive occupations (drillers of oil wells, explosives workers, miners, other mining occupations); supervisors of motor vehicle transportation; truck, delivery, and tractor drivers; transport equipment operatives; material moving equipment operators; helpers, constructive, and extractive occupations.

Annual earnings data is obtained from the income from INCWAGE, the pre-tax income from wages and salary variable. Earnings are top-coded at the 95th percentile of reported earnings and bottom-coded at the 1st percentile of reported earnings. All annual earnings are inflated to 2008 dollars. Only workers who reported being in the labor force, whether employed or unemployed, are included.

All regressions using CPS-ASEC data are conducted at the state-year level. Microdata are aggregated up to the state level using person-level weights for CPS supplement data. For the subsample analysis of non-migrants in Section 3 (table 5), I only include workers who reported living in the same house, moving within the county, or moving to a different county in the same state. I exclude workers who reported moving between states or moving abroad. I also exclude workers who did not respond to the migration questions. Workers are classified as moving for work (and therefore excluded from the sample in table 6) if they report that they moved: for a new job or transfer, to look for work or lost a job, for an easier commute, or for other job-related reasons.

C.2 Census and American Community Survey (ACS) Microdata

In Sections 2 and 4, I use the decennial census microdata from 1950 to 2000 and the annual American Community Survey (ACS) microdata from 2001-2010, which are both conducted by the U.S. Census Bureau and provided by the Integrated Public Use Microdata Series (IPUMS, Ruggles et al., 2015). I use the sample of 16-64 year olds, but for some analyses I restrict the sample to just 18-25 year olds or 18-30 year olds. In most of the analysis, I restrict the sample to full-time (at least 35 hours worked per week), full-year (at least 40 weeks worked per year) workers using the definitions employed by Acemoglu and Autor (2011).

The college enrollment variable is constructed using the harmonized EDUCD variable. Individuals are coded as having ever enrolled in college if they report having at least some college education. Individuals are coded as having never enrolled in college if their highest reported level of educational attainment was a high school diploma or equivalent. Individuals who did not report an education level were excluded from the analysis.

Annual earnings data is obtained from the variable INCWAGE, the pre-tax individual income from wages and salary. Annual earnings are only computed for workers who report working for wages or salary. Individuals who report being self-employed or an unpaid family worker, and individuals who report working no weeks in the previous year, are excluded. Annual earnings are topcoded at the pre-determined Census topcode levels, which vary from year to year. Annual earnings are bottom coded as the 1st percentile of reported earnings for each year. All earnings are inflated to 2008 dollars.

The census and ACS data are merged to the occupational task intensity data compiled by Autor and Dorn (2013) using the OCC1990 variable, which is harmonized across all years. The Routine Task Intensity (RTI) measure is the primary measure I use to determine how "routine-intensive" an occupation is. An occupation is classified as highly routine-intensive occupation if its RTI measure scores in the top third of all RTI. Out of 330 total occupations, 113 occupations fit this criterion.

All regressions are conducted at the commuting zone-year level. The census and ACS data are merged to corresponding commuting zones using the crosswalks provided by Autor and Dorn (2013). Demographic characteristics, occupations, education, earnings, and work variables are aggregated up to the commuting zone level using person-level weights.