

Earning More by Doing Less: Human Capital Specialization and the College Wage Premium *

Titan Alon

Northwestern University

November 2017

Abstract

This paper builds a model of human capital accumulation driven by increasing specialization of the workforce. Individuals increase the efficiency of time dedicated to human capital acquisition by focusing investments on narrower sets of skills. The evolution of secondary and post-secondary curricula in the United States from 1870-2000 confirms the presence of these changes in the *scope of specialization*. Quantitative exercises show that specialization can account for roughly 29% of the rise in the skill premium, and 25-30% of the rise in relative educational attainment from 1965-2005. The effect on the skill premium is largely due to a decline in specialization in high school, where vocational training was replaced with academic graduation requirements. The model's predictions are also consistent with international variation in the skill premium, attainment levels, and the organization of educational institutions. An important policy implication of the analysis is that making room for specialized occupational training in secondary schools could be an effective tool to tackle income inequality.

*I thank my adviser, Matthias Doepke, and my committee members, Joel Mokyr and David Berger, for their endless hours of guidance and support. A special thanks also to Matthew Notowidigdo for all his input at early stages of this project. Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208 (e-mail: titanalon@u.northwestern.edu).

1 Introduction

Few economists would refute that the *division of labor* is a primary driver of productivity and economic growth. Its centrality in economic thinking is apparent throughout the classical literature, and perhaps nowhere more famously than Adam Smith's (1776) discourse on the pin factory. Subsequent generations of economists including [Marshall \(1890\)](#); [Young \(1928\)](#); [Stigler \(1951\)](#); [Houthakker \(1956\)](#); [Arrow \(1979\)](#); [Rosen \(1978, 1983\)](#), and [Becker and Murphy \(1992\)](#) all discuss the gains from the division of labor prominently, often on equal footing with capital accumulation and mechanization¹.

In interpreting the classical literature on the gains from the division of labor, one can broadly group the underlying mechanisms into two categories: complementarity with mechanization and human capital accumulation. The first argues that increases in the division of labor split production into narrower sets of tasks for which specialized tools are more easily developed. Smith went so far as to suggest that the invention of new machinery itself was a consequence of the division of labor focusing workers on ever more specific tasks².

The second line of reasoning argues that the division of labor increases the skill of the workforce by allowing workers to focus human capital investments on only a subset of the total knowledge necessary for production. The claim follows from the fact that any skill can be improved by dedicating more time to it; hence, given any fixed time investment in education, an individual can achieve a greater level of mastery by focusing on a narrower set of skills. The human capital embodied in the workforce can then grow through a proliferation of specialists working together through a division of labor. The novelty of this *scope of specialization* channel is that human capital can grow even while the total time invested in skill acquisition remains unchanged. This contrasts with canonical human capital

¹For example, see Book IV, Chapter IX: *Division of Labour. The Influence of Machinery* in [Marshall \(1890\)](#).

²As Smith wrote in *The Wealth of Nations*, "It was the division of labour which probably gave occasion to the invention of the greater part of those machines, by which labour is so much facilitated and abridged. When the whole force of the mind is directed to one particular object, as in consequence of the division of labour it must be, the mind is more likely to discover the easiest methods of attaining that object than when its attention is dissipated among a great variety of things."

models in the literature (based on the work of Jacob Mincer) where growth occurs through increasing time investments in education (i.e. the *attainment channel*).

A simple example can help build the core intuition. Consider the employees of a translation company working across n different languages where all employees invest a fixed number of years, say s , to learn foreign languages. Suppose an individual's human capital in any language is strictly increasing in the time dedicated to studying it. If all translators work and are trained in all languages then on average they spend only s/n years studying each language and attain a shallow understanding. If instead individuals invest all of their s years in a single language, they would gain a greater level of mastery. Each worker could then handle translations in the language in which they specialized and the translation company, overall, would be providing a higher quality of service. The important thing to note here is that while the total time invested in human capital, s , is the same in both scenarios, the average human capital of a worker performing any task is higher with specialization.

This paper builds a model formalizing this *scope of specialization* channel for human capital accumulation. The model is one of *technical knowledge*, where production requires the execution of a differentiated set of tasks, each requiring different skills. Firms allow a division of labor whereby employees may work only on a subset of the tasks necessary for production. Skills training occurs in a competitive sector of schools which offer various *programs of study* differing in both the length of time and the *curriculum* offered. *Curricula* are partitions of the set of all skills necessary for production into non-overlapping subsets of skills called *fields of study*. Households therefore make two dimensional human capital decisions in choosing their programs of study, considering both the length of schooling and the breadth of skills acquired.

The costs and incentives for individual attainment choices follow the literature; conditional on choosing a specific set of skills, more time invested in education increases worker productivity, and hence earnings, but leads to less time in the labor market, resulting in forgone earnings. The incentive to specialize educational investments is similar; conditional on a time investment in education, individuals can narrow their focus on fewer skills, increasing their expertise, and then find

jobs which utilize those skills as intensively as possible. Since the initial time investment in education is independent of the subsequent cost of skill utilization, the former can be thought of as a fixed cost of human capital which gives rise to increasing returns to specialization. While not formally modeled, several researchers have identified precisely this feature of human capital markets as an incentive for specialization³, as Kenneth Arrow (1979) explains,

“Specialization in information can be considered. . . minimizing a set-up cost. When specializing, the individual. . . can spread the overhead of learning over a much longer run. . . So in our economy it is expedient for individuals to *specialize in some branch of knowledge* that is acquired partly through education.”

The cost of pursuing a more specialized education is pecuniary, stemming from the fact that more complex institutional arrangements must be adopted by schools in order to train and certify students in multiple differentiated courses of study⁴. Specifically, schools pay a fixed cost per field of study they offer in order to set up a corresponding department to administer education within the field. Therefore schools with more specialized curricular offerings must charge a higher tuition to recoup these institutional costs. The equilibrium level of educational specialization then balances the individual’s desire for greater earnings against the larger training costs incurred to support institutions that provide it. Along the model’s growth path, the real wage rises relative to these fixed costs, leading individuals to pursue more specialized training programs over time.

To investigate the strength of the *scope of specialization* mechanism as an organizing conceptual framework, I apply the model to study the long term dynamics of the college wage premium. Figure 1 summarizes the facts by displaying smoothed

³See also [Rosen \(1983\)](#).

⁴Another approach could have emphasized firm coordination costs, as in [Becker and Murphy \(1992\)](#). While the equilibrium formulations of the two approaches is nearly identical, the advantage of the approach taken here is that it is observable, as detailed microdata on curricular offerings and tuitions are publically available. Furthermore, the latter approach is more relevant for our empirical application which will consider changes in government policy which shift the cost structure of different programs of study. Alternatively, we could have emphasized the riskiness of human capital investments as a potential cost with individuals choosing not to specialize in order to protect themselves from skill obsolescence. I explore that approach and its implications in ongoing work.

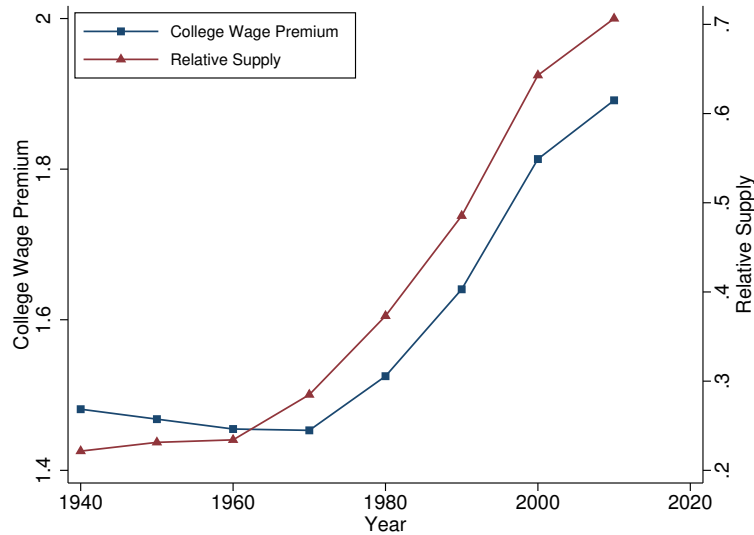


Figure 1: College Wage Premium and Relative Supply

trends from the decennial census in the college wage premium and the relative supply of college educated labor from 1940 onwards. The rise in relative earnings in the face of expanding supply from around 1980 onwards suggests a third factor pushing up the relative demand, or productivity, of the college educated relative to the high school educated. Identifying what factors are driving the trend in relative productivities is the focus of the enormous skill premium literature.

This paper argues that part of the concomitant rise in the skill premium and educational attainment is driven by the changing scope of specialization of secondary and tertiary education. The argument builds on a growing literature suggesting that the structure of institutions providing skill training in a society has important macroeconomic consequences for growth and inequality ([de la Croix, Doepke, and Mokyr, 2017](#)). Detailed historical curricular data and primary sources from the Department of Education spanning 1870-2000 provide direct evidence of changes in the degree of human capital specialization across the education sector. While historically both secondary and tertiary institutions provided specialized training, changes in federal policy beginning around 1970 all but eliminated this type of education in high schools. For college, the trend is most evident in the secular rise in fields of study over the last century. For high schools, where specialized

degrees do not exist, the focus is on the share of vocationally tracked students, which rose in the first half of the century to a peak of nearly 50% around 1970 before plummeting to single digits by the early 1990s.

To assess the quantitative importance of these changes, I nest the model into a generalized production framework allowing for skill-biased technological change and imperfect substitutability across education and age groups. The nested model allows us to study, net of general equilibrium effects, the independent contributions of skill-biased technological shocks alongside policy changes that shift the cost structure of schools offering different programs of study (i.e. specialization shocks). Given data on wages and attainments by cohort, the model separately over-identifies both series of shocks and indicates how they can be estimated non-parametrically using a log linear transformation of the data. The estimated specialization shock series displays a statistically significant trend break coinciding with the decline in secondary vocational education and rise in standardized, academically oriented curricula. Simulations which separately feed each set of shocks back through the model generate counter-factuals to assess the relative importance of each channel. The results suggest a sizeable role for the scope of specialization mechanism, which accounts for 29% of the rise in the college wage premium and 25-30% of the rise in the attainment ratio from 1965-2005.

The theory advocated here makes a number of contributions to the literature beyond offering an additional observable explanation for the long-term dynamics of the skill premium. In the sections that follow, I show that the model's predictions are consistent with several other important stylized facts in the skill premium literature, namely the presence of cohort effects identified by [Card and Lemieux \(2001\)](#) and cross-country variation in the skill premium documented by [Krueger, Perri, Pistaferri, and Violante \(2010\)](#). Even more importantly, the analysis suggests new educational policy tools to tackle income inequality. The economic model and accompanying empirical exercises suggest that paring back academic requirements to allow for specialized technical training at secondary schools could be an effective tool for tackling income inequality. The analysis also establishes that there is both domestic historical and cross-country precedent for the effectiveness of this approach. Such policies offer a meaningful alternative

to the prevailing wisdom that the main policy instrument should be promoting standardized academic curricula in secondary schools and subsidizing college matriculation.

The outline of the remainder of the paper is as follows; Section 2 situates this paper in the literature and provides additional motivating evidence for the approach taken. Section 3 reviews the history of curricula at secondary and tertiary education institutions from 1870-2000 to provide direct evidence of the paper’s mechanism. Section 4 builds the model and establishes several of its key properties. Section 5 develops the nested quantitative model and contains the results of the estimation and counter-factual exercises. Section 6 investigates the model’s consistency with international variation in the skill premium, attainments, and the structure of educational institutions. Section 7 discusses policy implications and concludes the paper.

2 Motivation and Literature

Several reasons motivate research into the long-term dynamics of the skill premium. First, the skill premium has tracked the U-shape trajectory of broader measures of income inequality over the last century (Piketty and Saez, 2003). This is in contrast to other widely studied margins of inequality, such as the male-female and black-white wage gaps, which have been declining for several decades. Second, the skill premium reflects the return to human capital—a major driver of modern economies—making it an important nexus of the literature studying the link between inequality, productivity, and economic growth. Finally, the question is policy relevant given the hundreds of billions of dollars expended annually on the education system and the popularity of education centric public policies to tackle income inequality and stimulate growth.

The existing literature may be broadly classified into three lines of reasoning: compositional, institutional, and technological. Compositional arguments have typically focused on the changing mix of workers and industries over the second half of the 20th century. This perspective alternatively takes changes in industrial composition due to trade or preferences, changes in workforce characteristics due to, among other things, increases in female labor force participation and

an aging population, or changes in the quality of graduates as the driving force behind the wage structure (Murphy and Welch, 1992; Card and Lemieux, 2001; Lemieux, 2006; Carneiro and Lee, 2011) Institutional arguments, in turn, focus on the erosion of unskilled workers' bargaining position brought on by the decline in unions, changes in compensation norms, and declines in the real minimum wage (Freeman, 1991; Lee, 1999; DiNardo, Fortin, and Lemieux, 1996; Card and DiNardo, 2002).

The final, and perhaps most popular, line of reasoning is the theory of skill-biased technical change (SBTC). Put tersely, this line of thinking rationalizes positive co-movements in observed relative supply and relative wages through outward shifts in the relative demand curve caused by technological change. Inspired in part by the SBTC theory's parsimony and empirical robustness, an enormous literature emerged trying to provide observable evidence of skill-biased technological change which had been originally justified indirectly by unaccountable movements in the residual of relative earnings. The most well received attempts alighted on computerization, or other advanced manufacturing techniques, co-opting the much older literature on capital-skill complementarities begun by Griliches in 1967 (Katz and Murphy, 1992; Berman, Bound, and Griliches, 1994; Doms, Dunne, and Troske, 1997; Autor, Katz, and Krueger, 1998; Autor, Levy, and Murnane, 2003; Dunne, Foster, Haltiwanger, and Troske, 2004). Several papers which extended the theory to full scale models and general equilibrium settings found similarly consistent results, both quantitatively and qualitatively (Heckman, Lochner, and Taber, 1998; Acemoglu, 1998; Krusell, Ohanian, Rios-Rull, and Violante, 2000). In a recent book, Goldin and Katz (2008) have shown how a broadened version of the SBTC hypothesis, one they term the Supply-Demand-Institutions (SDI) framework, is able to account for the dynamics of the skill premium over the near entirety of the 20th century.

Despite much promising evidence, doubts remain. The computerization theme could not reconcile implied slowdowns in skill-biased technical change with the rapid spread of computers in the nineties; longitudinal plant level data showed only weak correlation between advanced technologies and skill premiums, and the direction of causality was often unclear; and the very trend of skill biased

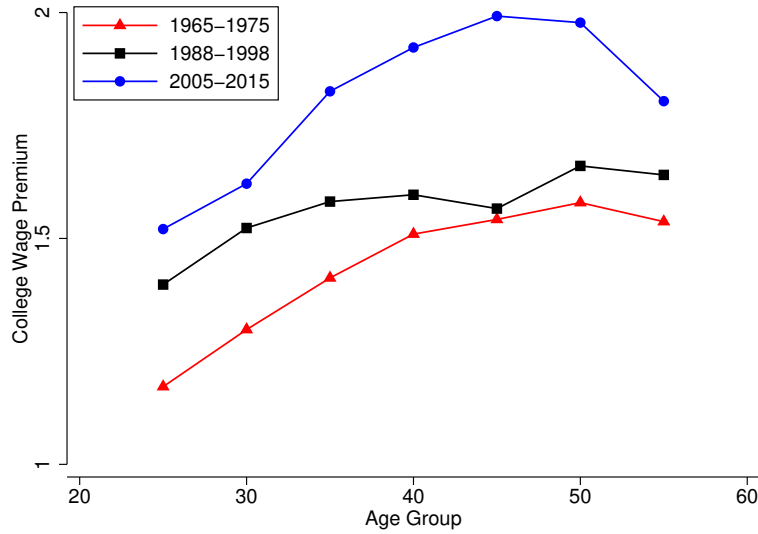


Figure 2: College Wage Premium by Age Group and Year

technical change was thought possibly ill at ease with measured slowdowns in productivity during earlier periods. Furthermore, [Goldin and Margo \(1992\)](#) uncovered a strong contraction in the wage premium during the first half of the twentieth century – a period of intense technological progress and diffusion⁵.

It is also not immediately clear how to reconcile the skill-biased technology hypothesis with other stylized facts on the skill premium documented in the literature. For example, [Card and Lemieux \(2001\)](#) find a substantial contribution of cohort effects to the growth of the college wage premium. Figure 2 illustrates the observation by plotting the college wage premium by age group and year for a sample period at the beginning, in the middle, and at the end of the Current Population Survey (CPS) data (for simplicity, we refer to these curves as *cohort profiles*). Comparing the early period, 1965-1975, to the middle period, 1988-1998, we see a "flattening" of the cohort profile as the skill premium of younger cohorts rises relative to the old. This was the primary observation made by [Card and Lemieux \(2001\)](#) showing that the premium's rise was being driven by cohorts entering the

⁵While these findings seem to have cast further doubt on the skill-biased theory as a long-term unifying framework, the authors rationalize these early results by characterizing progress as biased toward unskilled workers alongside a rapid increase in skilled labor.

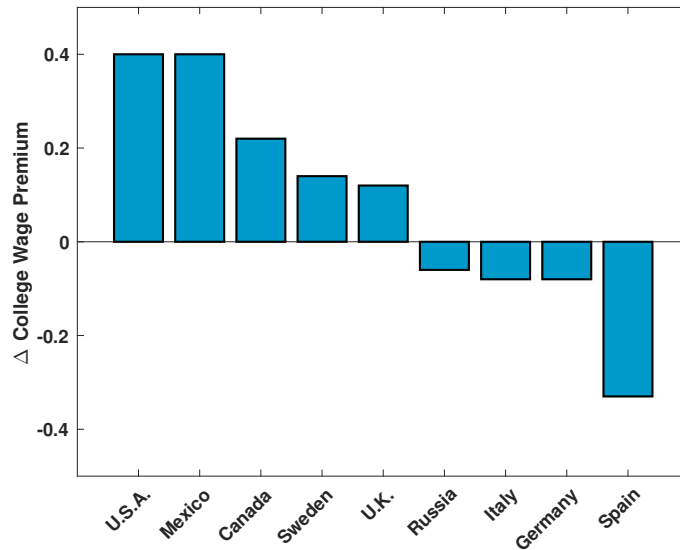


Figure 3: Cross-Country Changes in the College Wage Premium

labor market after the late 1970s⁶. Interestingly, if one extends their analysis using more recent data, we see that the hump shaped pattern of the cohort profile is eventually restored, with the curve for 2005-2015 looking much like the original 1965-1975 curve shifted upward. If the college wage premium was being driven primarily by capital deepening coupled with capital-skill complementarities, as in the technological hypothesis, then it is not clear why the rise would manifest itself only in young graduating cohorts after 1980, or why the effect would settle down after previous generations were replaced by these cohorts.

An additional important fact can be found in recent work by [Krueger et al. \(2010\)](#), which documents the substantial variation across developed countries in the dynamics of the skill premium⁷. Figure 3 plots the authors' estimates of changes in the college wage premium from roughly the early 1980s through 2000. The figure illustrates that the rise of the skill premium in the United States is not broadly representative; some countries, like Italy, Germany, and Spain, have seen substantial contractions in the skill premium over roughly the same period. While many potential explanations exist, it appears implausible that the large variation

⁶Those authors rationalize the facts with a partial equilibrium model of imperfect substitutability across age groups and exogenous variations in attainment levels across cohorts.

⁷The authors' work provides descriptive statistics on a variety of measures of inequality but does not advocate or test a specific hypothesis.

in experiences can be accounted for entirely by differences in technology across developed countries. Instead, the data suggests that institutional factors may be an important determinant of the trajectory of the skill premium.

A strength of the scope of specialization theory advocated here is that it can be reconciled with both of these additional stylized facts. The cohort trends are driven by the fact that policies constraining efficient human capital investments at the secondary level primarily affect generations entering high school after the policy shift. The result is rising inequality among the young relative to the old who were educated under the previous regime, causing the cohort profile to flatten out. Eventually all older generations are replaced by individuals graduating under the new regime and the cross-age dynamics disappear. These dynamics—a flattening of the profile along the transition to a new shifted up level—will in fact be a feature of the model developed in section 4 and will be a key source of variation in section 5.2 used to identify the policy shocks we study.

The international facts are also broadly consistent with the mechanism advocated here. One candidate explanation relevant for our purposes is the difference in educational institutions. While college curricula may differ slightly across the countries⁸, the extent of specialization in secondary schools varies immensely. Figure 4 depicts how these differences correlate with changes in the college wage premium, proxying for the degree of secondary specialization by the share of secondary school students in vocational programs⁹. Consistent with our theory, the figure documents a robust negative relationship¹⁰. In other words, more specialized secondary training programs are correlated with smaller changes in

⁸Several of the countries are signees of the Bologna Process which explicitly attempts to harmonize tertiary education curricula across countries.

⁹Except for the United States, all measures of vocational participation come from the OECD.Stat Labor Force Statistics database. For the United States, the estimate comes from the number of “vocational specialists” documented in the National Center for Education Statistics 1992 transcript study. This estimate is likely an upper bound for a truly harmonized measure.

¹⁰The one outlier is Spain. However, there are several reasons to believe that the data for Spain may be slightly unrepresentative of the underlying pattern we document. In the work of [Krueger et al. \(2010\)](#), the Spanish estimates are constructed for a shorter time period ending in the 1990s, use pre-tax data, and overlap with Spain’s ascension to the E.U. and the subsequent economic boom. The boom was also associated with a huge surge in college attainment in Spain, which today has one of the highest fractions of college educated workers in the OECD by a large margin. Nevertheless, more recent data from Spain, and other cross country patterns explored in Section 6, are also consistent with the results suggested here.

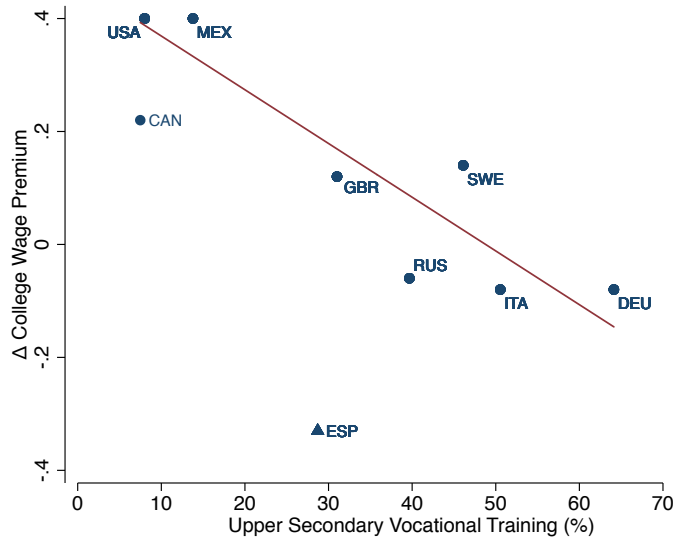


Figure 4: Cross-Country Changes in the College Wage Premium

the earnings of college to high school graduates. The theory ascribes this to the increased efficiency of educational investments at secondary institutions allowing for specialization. Further analysis of cross-country variation in the skill premium, attainment levels, and the arrangement of educational institutions is contained in Section 6.

This paper contributes to the literature by introducing a hitherto unexplored mechanism providing an observable foundation for the long-term dynamics of the skill premium which is also consistent with shifts in the cohort profile and variation in cross-country experiences. While the literature's focus thus far has been on the complementarity between capital and educational attainment, the argument here emphasizes the complementarity between human capital specialization and an increasing division of labor. This alternative approach creates a bridge between the very large literature on the skill premium and the very old literature on the consequences and causes of the division of labor (Smith, 1776; Young, 1928; Stigler, 1951; Houthakker, 1956; Rosenberg, 1976; Arrow, 1979; Rosen, 1978, 1983; Yang and Borland, 1991; Becker and Murphy, 1992; Jones, 2009) which has not yet been utilized to explain changes in income inequality¹¹. In doing so, this paper adds an

¹¹One line of research that comes close is the task-based framework advocated in Acemoglu

important new mechanism to a quickly growing literature seeking to understand the causes and consequences of rising income inequality ([Heathcote, Storesletten, and Violante, 2010](#); [Heathcote, Perri, and Violante, 2010](#)).

The quantitative application contributes insights to the education and labor literature beyond simply understanding the skill premium. Insofar as technical secondary education leads to middle skill jobs, its decline has also potentially contributed to the rise in labor market polarization documented in the literature ([Autor and Dorn, 2013](#)). More broadly, the focus on curricular change complements recent research on the costs and benefits of vocational versus academic education ([Arum and Shavit, 1995](#); [Acemoglu and Pischke, 1999](#); [Krueger and Kumar, 2004](#); [Fersterer, Pischke, and Winter-Ebmer, 2007](#); [Malamud and Pop-Eleches, 2010](#); [Lerman, 2012](#); [Hanushek, Schwerdt, Woessmann, and Zhang, 2017](#)) as well as offering a macroeconomic perspective on the importance of heterogeneity in fields of study on labor market outcomes ([Altonji, Blom, and Meghir, 2012](#); [Altonji, Kahn, and Speer, 2016](#); [Kirkeboen, Leuven, and Mogstad, 2016](#)). In identifying the catalytic role of federal policy, this paper also relates to research on optimal government policy in structuring educational institutions and their curriculums ([Kosters, 1999](#); [Hoxby, 2003](#); [Cunha and Heckman, 2007](#); [Dustmann, Fitzenberger, and Machin, 2008](#))

From a conceptual perspective, this paper introduces a growth model with the novel feature that human capital accumulation is driven by specialization, rather than educational attainment. By formalizing this *Smithian* channel of human capital, in contrast to the *Mincerian* tradition, the model constitutes a contribution to the literature studying the role of human capital in productivity growth and the extent to which variation in attainments alone is a useful empirical statistic ([Lucas, 1988](#); [Benhabib and Spiegel, 1994](#); [Mankiw, Weil, and Romer, 1992](#); [Klenow and Rodriguez-Clare, 1997](#); [Jones, 2014](#))¹². Furthermore, as the approach relies on a

and Autor (2011). However, that line of work is focused on exogenous models of specialization based on pre-existing abilities with optimal assignment and their interaction with automation. In contrast, the model developed below (and the classical literature cited above) is one of endogenous specialization based on individual education decisions and where the gains emerge from changes in human capital levels, rather than physical capital and automation. Nevertheless, as the introductory remarks make clear, the two approaches are inter-related and complementary to one another.

¹²See [Caselli \(2005\)](#) for a review.

differentiated task production framework, it connects to the literature stressing the importance of recognizing heterogeneity in skills and production tasks for understanding macroeconomic outcomes ([Jones, 2009](#); [Lazear, 2009](#); [Acemoglu and Autor, 2011](#); [Sanders and Taber, 2012](#)).

3 A Brief History of Schools, Curricula, and Skills

This section briefly reviews the history of curricular change at U.S. educational institutions since 1870 to provide direct evidence of changes in human capital specialization. A more detailed review, alongside reduced form evidence supporting the importance of specialization, is available in a companion paper [Alon \(2017\)](#).

Prior to 1870, most American secondary and collegiate education might be summarized by a single word: *classical*. There were no fields of study in the undergraduate curriculum; instead, all students followed the same prescribed course of study. A typical four year curriculum consisted primarily of logic, Latin, Greek, and Hebrew instruction with at least a year of metaphysics, ethics, and moral philosophy ([Kraus, 1961](#)). High schools were largely subservient to the colleges, most with preparatory curriculums similarly focused on the classical languages alongside a modicum of algebra, geography, history, and natural philosophy ([Alberty and Alberty, 1962](#)). This staunch traditionalism of American education would not last. As a new era driven by applied scientific knowledge took hold, the public would increasingly demand equal access to the skills demanded by the new economy and turned to the public education system to provide it.

Two events bear responsibility for bringing the classical college into contact with the modernizing currents of the age: the founding of the land-grant colleges and the arrival of the German research university model to the United States. The land-grant colleges were particularly revolutionary for the post-secondary curriculum. Alongside the classical curriculum, the new colleges introduced a range of engineering programs covering railways, mechanics, electrical, civil, irrigation, marine, and textiles; agricultural courses including animal husbandry, veterinary medicine, agronomy, horticulture, plant pathology, agricultural botany, agricultural chemistry, and farm management; as well as vocational studies in stenography, printing, telegraphy, mechanical drawing, machine design, ceramics,

journalism, architecture, business and many more (see [Report to the Commissioner of Education 1900, 1910](#)). Through the land-grant colleges, the scope of higher education in the United States expanded more in two decades than it had in the preceding two centuries. At the same time the new college accomplished what educational reformers had struggled and failed to do since colonial times—introduce secular, practical, and applied fields of study into institutions of higher education.

At the same time the land-grant colleges were being set up, a new model educational institutions arrived in the United States – the German research university. The university was different from existing American institutions in that it emphasized scientific research over teaching, focused on graduate and professional study, and advocated organizational structures with multiple specialized departments. Like the land-grant colleges, the university model legitimized the idea that post-secondary education could provide practical and occupationally oriented courses and gave the larger, established private colleges an avenue to respond to the growing popularity of the public colleges.

Through the land-grant colleges and university system the scope of higher education in the United States began to expand rapidly to encompass fields of study reflecting the various labor market demands of the day. While many colleges initially resisted the new university model on ideological grounds, it seems that the financial demands of the transformation were more often a determining factor. This would change by the end of World War II as government policies such as the Serviceman's Readjustment Act of 1944 and the Higher Education Act of 1965 massively increased federal subsidies to post-secondary institutions. The public's enthusiasm for practical, occupationally oriented education coupled with the surge in financial resources enabled the remaining classical institutions to transition to some form of the university model which, by the second half of the 20th century, would come to define American higher education.

Figure 5 tracks the effect these changes had on the expansion in fields of study at the tertiary level back to the mid 19th century. The top panel of the figure plots the number of fields with conferred degrees for a large midwestern research university and shows the emergence of differentiated fields of study in the late

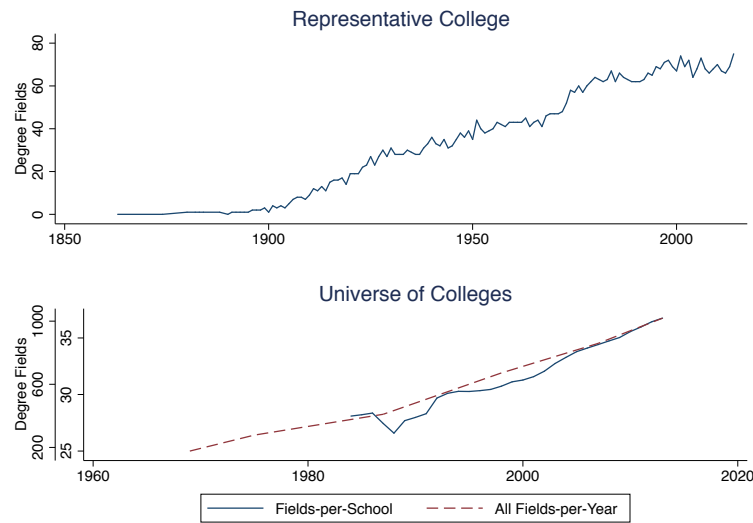


Figure 5: Number of fields of study at 4-year colleges

19th century. The bottom panel makes use of administrative data on fields of study for the *universe* of domestic 4-year colleges from the 1960s onwards, plotting both the total number of distinct degree fields recorded as well as the average across all institutions. The figures make clear the important role played by increasing specialization in driving the human capital accumulation of the college population.

By the beginning of the 20th century, while post-secondary education was expanding in scope and practicality, high schools were still bound to the classical curriculum. A decade earlier, the National Education Association's 1893 *Committee of Ten* report had successfully extended the classical curriculum of Northeastern academies to public high schools across the country. The system quickly came under pressure, however, as a surge in high school enrollments was reshaping the secondary school population and its life aims. By 1910, the eve of the high school movement, less than one-tenth of secondary students were preparing to enter college and were increasingly demanding access to skills training that would serve them in the modern labor market ([Report to the Commissioner of Education, 1910](#), Table 34). Responding to popular outcry as to the ill-suitedness of the preparatory curriculum for the majority of secondary students' life aims, the National Education Association established a second committee, the *Commission*

on the Reorganization of Secondary Education. The commission's report in 1918, the *Cardinal Principles of Secondary Education*, virtually overturned the classical approach to high school education. Among its most revolutionary aspects was the endorsement of vocational training as a fundamental principle of high school education, necessary to prepare all members of the public for working life.

The commission's publication coincided with action from Congress which passed the Smith-Hughes Act in 1917 providing the first direct federal support for nationwide secondary vocational education. In addition to providing federal dollars, the act established separate boards of education, curricula, and graduation requirements for vocational students. The curricular model followed the 50-25-25 model in which 50% of time was spent on hands-on work experience, 25% was dedicated to class study in closely allied subjects, and 25% on general education. Business groups and unions, like the National Association of Manufacturers and the American Federation of Labor, were closely involved in shaping the early programs and their curricula. Subsequent legislation over the next few decades, such as the George-Reed Act of 1929, the George-Deen Act of 1936, and the George-Barden Act of 1946, increased the scope and funding of federal support for vocational programs. Enrollments in vocational training programs surged in response to the federal initiatives. In 1917, just before the Smith-Hughes Act, there were approximately 200,000 vocational students in the United States with an annual budget of 3 million dollars; by 1946, there were 3.4 million vocational students with annual expenditures topping 176 million (Hayward and Benson, 1993). By 1950, vocational courses were the second most popular topic in high schools across the country, topped only by the Humanities.

The city of Chicago is perhaps the best paradigm of the transformations taking hold of secondary education in the early 20th century. According to Chicago's prospectus of courses, the city's high schools offered 11 distinct four-year courses and 10 two-year courses of study. The four-year courses of study included an (1) English course, (2) General course, (3) Foreign-language course, (4) Science course, (5) Normal preparatory course, (6) Business course, (7) Manual-training course, (8) Builders course, (9) Household Arts course, (10) Fine Arts course, and (11) Architectural course. The general courses were designed for non-specialized study

and were meant as preparatory courses for the colleges, normal, engineering, and scientific schools. In contrast, the 10 two-year courses of study were vocational in aim and designed for specialization. These programs included (1) Accounting, (2) Stenography, (3) Mechanical Drawing, (4) Design, (5) Advanced Carpentry, (6) Pattern-making, (7) Machine-shop work, (8) Electricity, (9) Household arts, and (10) Printing. According to the prospectus, each two-year course had an outlined curriculum that students were required to follow “in order to be well fitted for the occupation into which the major subject leads” ([Report to the Commissioner 1910](#), volume 1, chapter III).

By 1970, enrollments in secondary vocational education had peaked, but a growing number of challenges began to surface. Labor market demands were changing, but no system had been put in place to modernize the curriculum of vocational programs. Making matters worse, federal financial support had not kept up pace with student enrollments, and the share of federal financial support fell from roughly 50% in 1917 to 10% by 1970, eventually bottoming out at 5% by the end of the century ([Government Accountability Office, 1974](#); [Silverberg, Warner, Fong, and Goodwin, 2004](#)). At the same time, vocational programs were under pressure by civil rights activists demanding more equitable training access for disadvantaged minorities as a path to the middle class.

Congress attempted to act by passing amendments to the Vocational Education Act in 1968 and 1976 directly tying federal financial support to schools providing measurable evidence that they were modernizing vocational curriculums and expanding access. The legislation backfired, however, as schools lacked the resources and ability to monitor labor market outcomes of their vocational students. To meet federal demands, schools simply began introducing academic testing as a means to evaluate vocational programs. Moreover, the reallocation of funding lead to a surge of impoverished, special needs, english second language, and disabled populations into vocational programs. While well intentioned, federal legislation aimed at modernizing vocational programs ended up transforming them into remedial programs of last resort for society’s most difficult to educate students. A large part of the problem was that no additional financial resources or federal guidance was provided to help meet the increasing demands put on

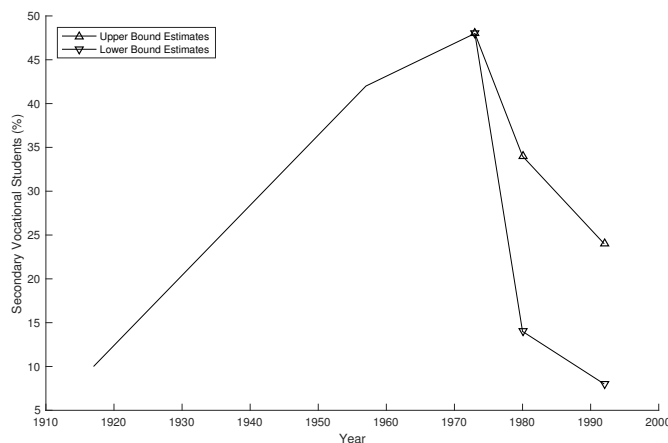


Figure 6: Share of vocational secondary school students

vocational programs ([Government Accountability Office, 1974](#)).

In this environment, education reformers arguing that the poor performance of high school graduates was the result of a lack of rigorous academic standards began to gain support. These efforts came to the fore in 1983 with the publication of a *Nation at Risk* which advocated for a nationally standardized, academic curriculum built around full sequences of English, science, mathematics, and humanities. In a historical irony, the new curricular recommendations were a modern take on those in the 1893 *Committee of Ten* report that had been overturned by the *Cardinal Principles* and introduction of vocational education nearly a century earlier. The efforts of academically minded reformers culminated in the *Common Core State Standards* by the National Governor's Association which extended academic graduation requirements across nearly the whole country. As the new academic requirements applied to all secondary students, the shift was a final death blow to the vocational curriculum. As a result, vocationally tracked enrollments plummeted from the 1970s through the early 1990s, when vocational tracks at secondary institutions had all but disappeared¹³.

The role played by shifting government policy is apparent not only in the nar-

¹³There are currently 900 full time vocational high schools in the United States, a tiny fraction of the 37,000 private and public high schools that exist. Imperfect records exist on vocational tracks at comprehensive high schools, though what data does exist suggest that by and large they no longer offer unified sequences geared toward specific occupations.

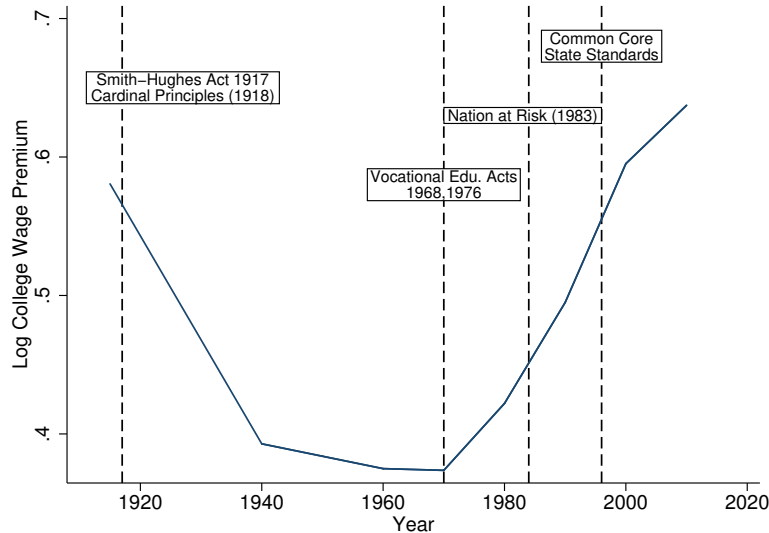


Figure 7: Education Policies and the College Wage Premium

rative of historical legislation, but also in the distribution of federal education financing. The total federal budget for vocational education at the secondary *and* post-secondary levels has remained around one billion dollars from the 1980s until today ([Government Accountability Office, 1974](#); [Gordon, 2008](#)). In contrast, the sum total of grants, subsidized loans, tax credits, and work-study funding-programs rooted in the Higher Education Act of 1965—reached over 200 billion dollars by 2013 ([PEW, 2015](#)). The funding priorities seem even more lopsided when one considers that nearly 2/3 of individuals do not matriculate at college and rely on secondary schools for their labor market training.

Figure 6 tracks the rise and fall of vocationally tracked students over the last century. Measuring the trajectory of these programs is not straightforward due to the lack of a national system of degree and certifications or even a consistent definition of who should be considered a vocational student. In an effort to create a harmonized series, the figure combines data from several sources to track vocational enrollments consistent with the curricular description in the 1909 Chicago prospectus as including students who have completed 2 carnegie units in a single occupational area. Due to the ambiguity of historical sources, it is hard to do so precisely, but nearly all attempts confirm the hump shaped trajectory.

The history makes clear that the structure of educational institutions has remained anything but constant over the last 150 years. The secular rise in college majors and vocationalism in the first half of the century highlights the important role of specialization in driving human capital accumulation and the labor market preparedness of students at both educational levels. After 1970, however, the efficiency of human capital investments in secondary schools was constrained by the decline in vocational education and rise of standardized academic curricula. As a result, secondary schools today face many of the same challenges as in the early 20th century; their curricula are focused nearly exclusively on college prep while nearly two-thirds of their student population will not proceed to college.

Figure 7 illustrates how these changes coincide with the long-term dynamics of the skill premium by extending the Census data back to 1915 using estimates from [Goldin and Katz \(2008\)](#) and plotting them alongside the important turning points in educational policy identified above. Consistent with our economic framework, the introduction and growth of secondary vocational education after 1917 coincided with a long term decline in the premium driven by large gains by the high school population. In contrast, the decline in specialized secondary education after 1970 accompanied a rise in the college wage premium and surge in matriculation. The next section formalizes these ideas by introducing the model and showing how government policies which shift the cost structure of different programs of study can effect the efficiency of human capital investments, attainment decisions, and the earnings of households. Combining the model with microdata from 1965 onwards confirms that these effects have been a substantial and statistically significant determinant of wages, attainments, and productivities.

4 A Model of Human Capital Specialization

This section introduces the benchmark model and highlights how specialization can drive human capital accumulation. Overlapping generation of finitely-lived, heterogenous households make two dimensional human capital investments, choosing both the length and breadth of their studies. After the equilibrium is characterized, several propositions establish the model's response to shocks affecting the schools' costs of providing certain programs of study. These exercises

are meant to reflect the type of policy shocks identified in the history section above and will motivate the empirical exercises in the next section.

4.1 Production

Production of the economy's single final good requires the execution of K distinct tasks. A competitive sector of firms possessing a constant return to scale technology hires labor to perform each of the $\kappa \in K$ tasks necessary for production. The productivity of a worker i assigned to some task κ is determined by his competency, or human capital, in the task denoted by $h_{i\kappa}$. A worker with skill level $h_{i\kappa}$ who dedicates $\ell_{i\kappa}$ physical hours to performing task κ will supply $h_{i\kappa}\ell_{i\kappa}$ *effective hours* of productive services at task κ . Letting $L_{\kappa t}$ be the total effective hours dedicated to production services in task κ we may express production of the final good by (suppressing time subscripts on all variables)

$$Y = AF(L_1, \dots, L_K, \dots, L_K) \quad (1)$$

where A is a neutral technological parameter and $F : \mathbb{R}^K \rightarrow \mathbb{R}$ exhibits constant returns to scale and is continuous, increasing, and concave in each of its inputs. The final goods technology is operated by a competitive industry, so that effective hours supplied to any task will be paid a task-wage, w_κ , equal to their marginal product.

4.2 Schooling and Education

All schools possess a common education technology which transforms study time dedicated to any skill κ , s_κ , into area-specific human capital, h_κ . Importantly, the technology is strictly increasing in study time so that more time dedicated to studying some skill always results in a higher level of human capital. There is heterogeneity in ability in the population, denoted by $a \sim F$, so that individuals of different ability achieve different levels of human capital conditional on the same time investments. The inclusion of heterogeneity in ability is not critical for the results developed below. Nevertheless, given that heterogeneity in ability features prominently in the returns to schooling literature, I include it in the baseline model

to illustrate that all the results are consistent with such a framework. I assume the following functional form:

$$h(a_i, s_\kappa) = a s_\kappa^\theta \quad 0 < \theta \leq 1$$

The educational technology is operated by a competitive sector of schools. Schools choose how many students to admit (n) and set tuition levels (p) as well as determining the *program of study* which is defined by the number of years of study, s , and the *curriculum*, m . A *curriculum* m is a partition of the set of K skills used in production into m non-overlapping¹⁴ equally sized subsets of skills called *fields of study*. As a matter of notation, let m_j denote the j – *th* field of study in a curriculum of type m . Students attending a school with program of study (s, m) choose a field of study m_j and allocate their studying time s uniformly across all skills contained within their field¹⁵.

The main difference in curriculums is the breadth of training they offer. A curriculum of type m offers fields of study containing K/m skills so that an individual pursuing a program of study (s, m) will dedicate sm/K time to each skill within his field, emerging with human capital level $h(a, sm/K)$ in each skill within his specialization. Hence, conditional on any time investment in human capital, s , individuals face a tradeoff in the breadth of skills: more specialized curriculums (higher m) yield higher levels of human capital in a narrower range of skills. Since in what follows we'll be focusing only on interior equilibrium and K is simply a scaling constant unidentified relative to the scale of technology, we normalize $K \equiv 1$ to economize on notation¹⁶.

Offering more diverse curricula is costly and requires schools to adopt more

¹⁴In principle schools could offer overlapping *fields of study*. However, in the equilibrium of the model being developed, no school will choose to do so and so we impose this condition in the definition for simplicity.

¹⁵The assumption that time is spent uniformly is for simplicity, as in [Becker and Murphy \(1992\)](#). One could easily extend this to consider education policies that mix specialized and generalized education which require some fraction β of time can be dedicated to general education and only $1 - \beta$ into specialization. All results below will be the same up to a scaling provided that $\beta \leq 1/2$ so that individuals emerge with more human capital within their field of specialization.

¹⁶Alternatively one could think of K as a very large index so that the summation may be represented as an integral over the unit interval.

complex institutional arrangements. Specifically, each school must pay a fixed cost c_m for each field of study it offers in order to operate a *department*, whose purpose it is to organize, monitor, and certify education within a field. These institutional costs are what ultimately constrain the incentive to specialize as more specialized programs of study demand more complex institutional arrangements which require charging higher tuition levels to support. Specifically, each school solves:

$$\pi \equiv \underset{p,m,n,s}{\text{maximize}} \quad psn - c_n n^{v_n} - c_m m^{v_m}$$

where $c_n, c_m > 0$ and $v_n > 1$ and $v_m \geq 1$ parameterize the schools' cost function¹⁷. The purpose of including c_n is twofold: first, it ensures that schools have a finite size and so avoids aggregate scale effects in the model's equilibrium and, second, it shows how the microeconomic source of variation in program costs can emerge from policies that directly affect the cost of maintaining programs or those which alter the schools' cost of scale (see below). Competition and free-entry in the education sector then implies that in equilibrium all profits are driven to zero for each type of program of study. As a result, in equilibrium, annual tuition costs for a program of type (s, m) is given by:

$$p(s, m) = \frac{f m^\nu}{s} \quad (2)$$

where $f \equiv \left[v_n / (v_n - 1) \right]^{\frac{v_n-1}{v_n}} c_n^{\frac{1}{v_n}} c_m^{\frac{v_n-1}{v_n}}$ and $\nu \equiv \frac{v_n-1}{v_n} v_m$ are compound parameters. From equation 2 it is clear that the tuition cost of providing any program of study (s, m) is increasing in f , which itself increasing in both the costs of scale c_n and the cost of establishing specialized departments, c_m . In what follows, we refer to the compound parameter f as the *cost of specialization*.

¹⁷The assumption that $v_n > 1$ ensures that schools have finite size so that there are no scale effects in the equilibrium we study. The assumption that $v_m \geq 1$ is a generalization of the fixed costs of establishing new departments. Modelling institutional costs of new departments purely as a fixed cost would correspond to the subcase $v_m = 1$.

4.3 Households

The economy is populated by finitely lived overlapping generations of heterogeneous households each of unit mass. Households are heterogeneous along two dimensions: ability a_i and the opportunity cost of school ϕ_i , denoted in utility terms. The heterogeneity in costs, ϕ_i , is meant to parsimoniously capture the large number of factors studied in the literature which cause individuals to have different opportunity costs of schooling, such as financial frictions or family background. Modelling the heterogeneity in this way allows the equilibrium to support a distribution of attainment levels while remaining agnostic on the underlying determinant, which is beyond the scope of this paper and already extensively studied in the literature¹⁸. Throughout, we assume these two margins of heterogeneity, opportunity costs and ability, are independently distributed in the population subject to the distribution functions G and F , respectively. In Appendix E, I explore the implications for selection into higher education of more general joint distribution functions defined over ability and opportunity costs.

Each household is endowed with one unit of time in each period of life and has access to perfect capital markets at fixed flow interest rate r . Each household makes human capital, consumption, and labor supply decisions. The first stage of life is spent in full time schooling and the human capital decisions are represented by a choice of a program of study and field, (s, m_j) . In the second stage of life, individuals enter the labor market and choose how to allocate work hours across production tasks $\ell_{\kappa t}$. We refer to an individual i 's distribution of work hours across tasks $\{\ell_{\kappa t}\}_{\kappa}$ as their *job*. The time budget constraint requires the sum of all work hours sum to unity in each period. When workers choose jobs that involve assigning zero hours to some tasks κ' , we say that production features a *division of labor*. Finally, while on the job, human capital grows at exogenous rate g_e as experience accumulates through learning-by-doing. The growth rate in human capital on the job allows us to match the upward sloping lifecycle of earnings

¹⁸Note that both distributions are permitted to be degenerate mass points. Specifically, we allow that $\phi_i = 0$ for all individuals. Without the heterogeneous utility costs of schooling an equilibrium would still exist but all individuals would choose the same level of schooling $s = \frac{\theta}{(r+v-g_a)}$. While this is no problem from the perspective of illuminating the specialization mechanism, it will make our second goal of comparing earnings across education groups impossible and hence this is why we introduce the heterogeneous costs.

evident in the data, which will be necessary to control for when comparing cohorts.

The full household problem can therefore be written:

$$\begin{aligned}
& \underset{c_t, s, m_j, \ell_{kt}}{\text{maximize}} && \int e^{-(\rho+\delta)t} \log(c_t) - \phi_i s \\
& \text{subject to} && \int e^{-(r+\delta)t} c_t = \int_s e^{-(r+\delta)t} I_t(a, s, m, \ell) - p(s, m) s \\
& && \sum_k \ell_{kt} = 1
\end{aligned}$$

where δ is the flow rate of death ensuring finite lifespans, r is the fixed interest rate, and ρ represents time preference. $I_t(a, s, m_j, \ell)$ denotes the period t income of an individual born in period 0 with ability a , job ℓ , and who pursued program of study (s, m_j) in school so that $\forall t \geq s$ we have:

$$I_t(a, s, m_j, \ell) \equiv \sum_{\kappa \in m_j} w_{\kappa t} h_{\kappa t}(a, sm) \ell_{\kappa, t}$$

along with

$$h_{\kappa, s} = h(a, sm) \quad \text{and} \quad \dot{h}_{\kappa, t} = g_e$$

The expressions on the right-hand side of the budget constraint highlight the tradeoffs faced by households in making human capital investments. Increasing either the time dedicated to education, s , or the specialization of human capital investments, m , will increase household earnings in the labor market. The cost of spending more time in school is foregone earnings, which is captured by the fact that earnings (and experience) only begin to accrue after period s . These costs are augmented by the heterogeneous utility cost of schooling, ϕ_i , which generates a distribution of attainment levels when G is non-degenerate. In contrast, the cost of pursuing a more specialized curriculum is the higher tuition costs that must be paid to attend a school with more specialized departments (see the expression for equilibrium tuitions in 2). While in principle the costs of specialization are likely broader than simply the associated education costs—for example, the uncertainty in demand that might exist for specific types of skills¹⁹—here we choose to focus on

¹⁹I explore the *riskiness* of specialization as a determining equilibrium cost in an upcoming

the institutional costs since, (i) they are measureable in the data, and (ii) historical evidence suggests they are the most likely channel through which the changes in government policy have induced changes in human capital investments.

4.4 Equilibrium and Results

The following proposition summarizes the model's equilibrium.

Definition 1. *The model's equilibrium is characterized by:*

- *A set of task wages $\{w_\kappa\}$.*
- *A distribution of admissions $\{n\}$, programs of study $\{s, m\}$, and tuitions $p(s, m)$.*
- *A distribution of consumption $\{c\}$, labor supply $\{\ell\}$, and human capital investments $\{s, m\}$.*

such that:

- *Each household type (a, ϕ) is maximizing lifetime utility.*
- *Firms and schools are maximizing profits and both sectors support free entry.*
- *Factor markets clear: supply and demand of effective hours for each task κ are equated, $L_\kappa^D = L_\kappa^S$.*
- *Education markets clear: supply and demand for programs of study (s, m) clear.*

As neutral technology grows at a constant rate, the model supports a balanced growth path, the key properties of which are given by the follow proposition:

Proposition 1 (Balanced Growth Path). *As neutral technology A_t grows at constant rate g_a the model supports a balanced growth path where:*

- *Consumption, output, and incomes grow at a constant rate.*
- *The distribution of educational attainments, s , is constant.*

companion paper.

- *Human capital grows at a constant rate, driven by specialization.*
- *The college wage premium and the cohort profile are constant.*

Proof of Proposition 1. See Appendix A.

The proposition highlights the model's main feature that human capital accumulation can occur even while the distribution (and hence average level) of educational attainment remains constant. Note that the model supports a distribution of program types (s, m) so that in equilibrium the education sector is populated by many schools offering programs of varying lengths of study and curricular specialization to accommodate students of different ability and opportunity cost of schooling. Along the growth path, however, the distribution of program lengths remains unchanged but the number of fields of study is growing.

The intuition for the result is as follows: the growth in (neutral) technology shifts out the labor demand curves for all tasks, raising all task-wages. The shift has no effect on the distribution of attainment levels since the increase in returns to schooling are exactly offset by higher foregone earnings for those who study longer. Hence, instead of studying longer, individuals choose to economize on their schooling investments through increased specialization. This is because the institutional costs associated with increased specialization, denoted in monetary units, falls relative to the wage rate. As a result, successive generations adopt more complex and specialized training programs. Hence, even while the average level of attainment remains unchanged, human capital accumulation occurs and contributes to the growth rate in output, consumption, and incomes.

We can now consider policies, consistent with history reviewed above, that alter the cost for individuals to specialize their human capital acquisition. From a modelling perspective, these policies can be implemented in several ways. For example, we might explicitly “standardize” the curriculum by forcing the same distribution of human capital investments on all workers, or we can restrict the set of skills $\kappa \in K$ individuals can pursue, or yet still we can model policies as changing the cost of specialization, f , of different schools. While all three options will yield the same qualitative predictions, in what follows I pursue the third approach since it will prove to be the most empirically tractable, allows changes

to occur endogenously rather than by fiat, will allow for gradual changes in specialization over time, and will allow for some level of specialization to persist even after the policy change—all of which appear to be realistic features. The microeconomic sources of these changes in f are highlighted by the equilibrium tuition levels in equation 2. These include increases in the costs of scale c_n or to the cost of maintaining specialized programs c_m . Such changes could emerge from decreases in government support, curricular standardization, or changes in graduation requirements. From the historical discussion above, it is plausible to believe that all three factors have played a meaningful role.

In the results that follow, and the quantitative exercises below, we consider changes to f directly and its effect on the economy's equilibrium and the model's transition dynamics. To begin the analysis, first consider the case of an unanticipated change in the institutional costs of specialization.

Proposition 2 (Specialization Shock). *Consider an unanticipated rise in the parameter f . Then the new equilibrium will have lower levels of human capital, income, output, and consumption but educational attainment levels will not change. Along the transition path, the income of the young will fall relative to the old.*

Proof of Proposition 2. See Appendix A.

The result again makes clear the importance of considering the specialization channel. In the original BGP, attainment levels were constant while human capital was growing. In the proposition above, we see also that shocks to f may decrease the level of human capital by raising the cost structure of more specialized curricula, even while the levels of attainment remain constant. The main point is that attainment levels alone are an insufficient statistic for the level of human capital. Building on this result, we can now consider more directly a case relevant for our purposes by considering a specialization shock to f for all educational attainment levels below some cutoff level of schooling. This exercise can be thought of as a first pass at trying to understand the economics of a change in policy which constrains the ability to specialize in high school relative to college. The following proposition establishes the result:

Proposition 3 (Relative Specialization Shock). *In response to an unanticipated rise in the cost of specialization f for all programs of study of length $s < s_{\text{college}}$,*

- *The college wage premium and the relative supply of college labor will increase.*
- *The college wage premium among young cohorts rises relative to older cohorts along the transition path.*
- *In the new equilibrium, the cohort profile will be shifted up and parallel to the original curve.*

Proof of Proposition 3. *See Appendix A.*

The second proposition links the model, the shocks identified in the history section, and the stylized facts in the skill premium literature. Specifically, the model shows how shocks which raise the relative cost of providing specialized training at the secondary level²⁰ can generate model dynamics consistent with the rise in the skill premium and attainment levels (as in [Katz and Murphy \(1992\)](#)) in addition to generating twisting in the cohort profile (as in [Card and Lemieux \(2001\)](#)). The rise in the college wage premium and the twisting of the cohort profile are intuitive to see from the model. The wage premium rises because the specialization channel becomes costlier for non-college graduates, leading to lower levels of human capital from this group which persist throughout their life. The dynamic response to the shock generates cohort effects because it operates through educational institutions which are primarily attended by the young. The fact that attainments rise is also intuitive but less apparent from the equilibrium of the model since the expression for the optimal program length of study in the BGP (see equation 8) is not a direct function of f . What drives the result is that the asymmetric shock in proposition 3 introduces a kink in the value function around s_{col} , leading some mass of students whose optimal length of study is near (but below) a college degree to pursue a longer program which also allows a greater degree of specialization.

²⁰Or, alternatively, reduce the cost of providing specialized education in colleges relative to high schools.

The results above suggest that, from a theoretical perspective, the scope-of-specialization mechanism is a feasible explanation for the stylized facts on the skill premium. The next section builds on these results by showing that the mechanism is also significant and substantial in an empirical sense. To demonstrate the theory's value to the literature, we embed the model in a production framework which also nests the [Katz and Murphy \(1992\)](#) and [Card and Lemieux \(2001\)](#) frameworks to show that the scope of specialization channel provides additional explanatory power above and beyond these two widely accepted mechanisms in the literature.

5 A Quantitative Assessment

This section aims to quantify the effects of the policies studied above. To do so, we allow the cost parameters to vary over time and between colleges and high schools in order to assess the extent to which movements in the relative cost of specialization, f_h/f_c , has contributed to the stylized facts in the introduction. To establish the mechanism is meaningful relative to theories in the literature, we embed the one sector model above into a more general production framework which allows for skill biased technical change and imperfect substitutability between education levels and age groups. The aim is to estimate a series of *specialization shocks* f_h/f_c alongside changes in skill biased technology A_c/A_h net of general equilibrium and supply effects. We then interpret the estimated shocks in light of the historical evolution of educational policy before feeding them back through the model to construct counter-factual simulations of their quantitative significance.

5.1 A Nested Model

Following the literature, we assume that high school and college labor are imperfectly substitutable in the production of the economy's final good so that the aggregate production function can be written

$$Y = [(A_c Y_c)^\gamma + (A_h Y_h)^\gamma]^{\frac{1}{\gamma}}$$

where γ determines the elasticity of substitution between education groups, Y_e is the value of inputs provided by workers of education level $e \in \{\text{college (c), high}$

school (h)), and A_e are sector specific technologies. Furthermore, within each sector, labor services are generated by overlapping generations of imperfectly substitutable workers of the same education level, so that

$$Y_e = \left[\sum_j Y_{ej}^\eta \right]^{\frac{1}{\eta}}$$

where η governs the elasticity of substitution between age groups, and Y_{ej} is the output of individuals of age j and education level e . The production technology for each group of workers in both sectors, Y_{ej} , is the same as the one-sector model above (see equation 1).

Note that the framework above nests three models into a single production framework. For example, if $\eta = 1$ and $\theta = 0$, we have the canonical Katz-Murphy framework. If only $\theta = 0$, we have the Card-Lemieux set up. When $\gamma = 1$ and $\eta = 1$ we have the one sector scope of specialization model from the last section. The problem faced by households is the same as above, except now the educational attainment choice is binary. The school problem is also the same except now only two programs length can be offered and the cost parameters determining f are allowed to vary by sector and over time (and will be estimated below).

Finally, to characterize transition paths off equilibrium it will be necessary to parameterize the distribution function G of the opportunity costs of schooling ϕ . The choice will determine how elastic the household attainment decisions are to changes in relative earnings. There are many potential parametric choices for G and the literature provides no guidance on which ones might be preferred. In the empirical work that follows, we will use the functional form

$$G(x|d, \alpha) \equiv \frac{d \exp(x)^\alpha}{1 + d \exp(x)^\alpha}$$

where $d, \alpha > 0$ are parameters. This functional form yields several advantages over alternatives. First, it will yield a constant elasticity of relative supply, mirroring the constant elasticity of relative demand that will emerge from the production side of the model. This congruence makes characterizing the off-equilibrium behavior of the model very tractable in addition to avoiding situations in which the response

in relative attainment levels may change drastically and non-monotonically in response to a change in relative earnings. Second, the congruent structures in relative supply and demand for skilled workers will also allow us to express the model's observables as a log-linear function of the underlying shocks (see proposition 4 below), greatly simplifying the implementation and transparency of the estimation procedure²¹. Finally, the parameterization of G is straightforward to match to the data as, in equilibrium, d and α will govern the scale and curvature, respectively, of the relative supply function²².

5.2 Identifying Shocks

Our primary goal is to use the structure of the model and data on attainments and wages to estimate a sequence of shocks to the cost of specializations f_h/f_c and relative technologies A_c/A_h that best fits the data. It is not possible to separately identify the two shock series of the model relying solely on the BGP. This is because along the BGP the two relative shock terms enter symmetrically, up to proportion, in determining relative wages and attainment levels and hence are observationally equivalent and unidentified. This should not be surprising as it speaks to the main conceptual point being advocated here that the long-run trajectory of attainments and the skill premium can be equivalently explained by skill-biased technology or changes in the relative scope of specialization.

To achieve identification, we exploit the model's differential response off the BGP to each type of shock. Identification comes from the fact that off the BGP the wage structure responds differently to specialization shocks than to technology shocks. This approach has already been alluded to in proposition 3 above. Heuristically, we can think of the strategy as comparing the pivoting in the cohort wage profile²³ to its vertical shifts. The intuition is that shifts in relative specialization costs will affect each cohort in the labor market differently depending on when they were enrolled in school, causing the structure of wages across age groups to twist. In

²¹The shocks would of course still be identifiable with other parametric assumptions but these would require simulations or a Maximum Likelihood or Nonlinear Least Squares approach.

²²As an alternative robustness check we conduct simulations assuming a lognormal distribution for G and find our main conclusions little changed.

²³Recall the cohort profiles is the curve tracing the college wage premium by age group in a given time period as in figure 2

contrast, changes in sectoral technologies will symmetrically effect all workers in a given sector the same way, causing the cohort profile to shift up or down. Observing the wage structure over time then and comparing the pivots to the vertical shifts will allow us to separately identify the two shock series.

Of course, the identification strategy above is confounded by general equilibrium forces. Changes in relative technologies, for example, will alter the net payoff for incoming generations of young workers deciding between college and high school education, inducing cohort effects which twist the wage structure. Similarly, changes in the cost of specialization will change the size and composition of cohorts entering each sector, causing sectoral wages to shift and hence affect relative earnings across all ages. These general equilibrium interdependencies generate endogeneities which confound our identification strategy. To address these challenges, we exploit the structure of the model to quantify the extent to which each type of shock spills-over through relative prices and remove this source of variation from the data in order to directly identify the underlying shocks. Most of the parameters which govern the strength of these forces have already been widely estimated in the literature, and so we are able to use them directly in the estimation procedure (see section 5.4 for details). Given a suitable parameterization, both shock series are independently identified provided sufficient micro data exists on wages and attainments by cohort. The following proposition formalizes the identification condition.

Proposition 4 (Identification). *Define \mathbf{E} to be a vector of length $N \times T$ of the log relative earnings of each age group in N for each observed time period in T and \mathbf{N} be a length $N \times T$ vector of the corresponding log attainment levels. Let $\tilde{\mathbf{W}}$ be a $N \times T$ dimensional vector whose entries are the time t log ratio of total sector labor earnings in the college sector over the high school sector. Then*

$$\mathbf{D} \equiv \mathbf{A} + \frac{\theta}{\nu} \mathbf{f}$$

where \mathbf{A} and \mathbf{f} are $N \times T$ vectors of the period t relative technologies $\log(A_{ct}/A_{ht})$ and cohort $t - n$ relative specialization shocks $\log(f_{h,t-n}/f_{c,t-n})$ and

$$\mathbf{D} = \frac{1}{\eta} \left[\mathbf{E} - \mathbf{B}_0 - \mathbf{B}_1 \tilde{\mathbf{W}} - \mathbf{B}_2 \mathbf{N} \right]$$

where $\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2$ are known matrices of constants composed of the model's parameters. Then provided that $T \geq 2$ the model is identified and the shocks may be estimated by

$$\mathbf{A}, \mathbf{f} \equiv \operatorname{argmin} \left(\mathbf{D} - \mathbf{A} - \frac{\theta}{\nu} \mathbf{f} \right)' \Omega \left(\mathbf{D} - \mathbf{A} - \frac{\theta}{\nu} \mathbf{f} \right)$$

for any $NT \times NT$ weighting matrix Ω .

The intuition for the result is straightforward. The matrix \mathbf{D} begins with the relative earnings profiles in each period and then nets out all the direct effects, such as accumulated experience and variation in age-specific relative supplies, and indirect general equilibrium effects, such as movements in relative wages, until the only remaining terms are the relative technology and cohort specialization shocks. The first of these shocks is specific to a time period, the second is specific to a cohort, and so projecting the residual earnings terms onto a set of time and cohort fixed effects allows us to estimate the shock series without the confounding endogeneity concerns. Given observations for N age groups over T periods we have $N \times T$ observations to estimate $T - 1$ relative technology shocks and $N + T - 2$ relative specialization shocks²⁴ which is possible provided that $NT > 2T + N - 2$, or more simply $T \geq 2$. With our data sources, the shock series are both over identified as we observe multiple age groups within each time period and observe each cohort multiple periods over time. Exploiting the log linearity of the set-up then we can estimate the shocks using weighted least squares where the Ω matrix can be an arbitrary weighting matrix. The actual shock series can then be retrieved by rescaling and exponentiating the estimated log series.

5.3 Data

The primary data source is the Census Current Population Survey (CPS). From the CPS, we derive variables on individual income, educational attainment, labor force status, hours worked, and demographic characteristics for the U.S. economy from 1965-2015. The baseline sample consists of the entire working age population 18-65 whom are full-time, full-year workers not in the public sector, military, or attending school and whom have earned either a high school or bachelor's

²⁴Allowing the base year technology and specialization shocks to be absorbed by the constant.

degree²⁵. Earnings are defined as average weekly total pre-tax personal income to account for variation in hours and weeks worked across individuals in the sample. The high school population is defined as those with either a high school diploma, GED, or have completed 12th grade. The college population includes those with a Bachelor's degree or have completed 4 or 5 years of post-secondary education and do not hold any advanced degrees. All observations with either non-positive, top-coded, or missing values for any of the variables defining the sample above are dropped.

In the estimation procedure below, we bin observations into five-year age groups before calculating earnings and attainment levels. The choice to bin the observations, a procedure common in the literature, serves three purposes here: to increase precision, to smooth our estimates, and to keep the dimensionality of the simulations manageable²⁶. As identifying variation across cohorts will be important we similarly bin year observations into five-year groupings so that our identified cohorts represent the same group of people over time. Furthermore, the baseline sample used to estimate the shock series in section 5.2 will span 1965-2005 and only include age groups over 25. The time period is defined to omit any time bins that contain years from the great recession and its aftermath. The focus on age groups over 25 is to mitigate the fact that the age of completion and entrance to college has crept upward over the sample period we study. While increasing the transparency and identification of our empirical exercise, as a robustness check we verify that our main findings are not sensitive to either restriction.

Other raw data sources used in our analysis include the Federal Reserve Economic Database (FRED) and the National Center for Educational Statistics' (NCES) Integrated Postsecondary Education Data System (IPEDS). Details on the variables derived from these sources and their usage is contained in the next section.

²⁵Full-time is defined as working at least 35 hours a week. Full year is defined as working at least 40 weeks a year.

²⁶In the simulations of the transition paths we will need to calculate a trajectory of wages and earnings for each age group. Binning reduces the number of variables to keep track of from 48 to 8 while still allowing us to characterize lifecycle earnings and the cohort profiles.

5.4 Parameterization

The parameters used in the empirical exercises are summarized in Table 1. The values in the table are either calibrated to match widely accepted values in the empirical literature or, when these are unavailable, estimated to match relevant moments derived from the model's equilibrium conditions. The first group includes the values for γ , η , θ , g_c , and g_h ²⁷. With the exception of θ , each of these parameters is directly estimated in the corresponding literature. As the value of θ governs the marginal productivity of time spent studying, it is set to match the average returns to years of schooling holding constant the level of specialization, which has been heavily studied and is reviewed in Card (2001)²⁸. The value for the interest rate, r , is set to its average over the sample period studied and the discount rate, β , is calibrated to the inverse of the interest rate for simplicity and to avoid degeneracies in light of fixed interest rate borrowing. The growth rate of TFP, g_a , is set to its long-term utilization-adjusted trend in the post-war period, as estimated by Fernald (2014)²⁹.

The values for the parameters d and α , which govern the elasticity of *supply* of college goers, and the compound parameter ν , which governs the elasticity of education costs with respect to curricular specialization, do not appear elsewhere in the empirical literature and so must be estimated here. From the model's equilibrium conditions, we have that the values of d and α determine the level and curvature, respectively, of the relative supply of college goers as a function of relative lifetime incomes. These two parameters are set by matching the levels of attainments and the college wage premium in 1965 and 2005. The estimated value of d implies that if the return to college was identical to that of high school,

²⁷The estimates of the returns to experience by education level, g_c and g_h , are taken as the average estimates across men and women in table 5 of the Connolly and Gottschalk (2006) paper, which itself draws on microdata from the Survey of Income and Program Participation.

²⁸Specifically, we model high school as a four year time investment compared to college which is a nine year investment, to be consistent with manner that the data is binned. Hence, ceteris paribus, the return to college in the model can be write $(9/4)^\theta$ and so taking the literature's average estimate of a return to a year of schooling at 8.5% we choose θ so that the above expression equals $(1.085)^5$.

²⁹The rate is estimated as the average annualized percent change in the post-war period up through 2007. Estimates are very similar if one instead uses the non-farm multifactor productivity estimates available through FRED.

Parameter	Value	Source
γ	0.50	Katz and Murphy (1992) , Acemoglu and Autor (2012)
η	0.80	Card and Lemieux (2001)
θ	0.50	Returns to Schooling, Card (2001)
ν	1.55	Estimated, NCES Tuition and Curricular Data
d	0.12	Estimated, Attainment and Earnings Trends (CPS)
α	1.52	Estimated, Attainment and Earnings Trends (CPS)
β	0.94	Inverse of Interest Rate r
r	6.5%	Average Effective FF Rate (FRED)
g_c	2.8%	Connolly and Gottschalk (2006)
g_h	1.9%	Connolly and Gottschalk (2006)
g_a	1.3%	Long-Term TFP Growth Rate, Fernald (2014)

Table 1: Model Parameters

roughly 10% of individuals would still pursue a college degree. The value of α implies that if the *lifetime* earnings of a college graduate relative to a high school graduate increased by 1% the relative supply *ratio* of college graduates to high school graduates would increase by roughly 5%³⁰.

The compound parameter value ν governs the elasticity of tuition costs with respect to curricular complexity. Taking the log of the equilibrium tuition levels from equation (2) yields

$$\log(p) = \log(f/s) + \nu \log(m)$$

which suggests that ν may be estimated by regressing log tuition levels on log curricular offerings. Using a balanced panel of tuitions and the number of degree fields offered for all domestic accredited four-year public and not-for-profit colleges who awarded more than one-degree a year between 2001-2014³¹ in the

³⁰Note that the increase is in the ratio of college to high school graduates, the actual percentage increase in the number of college graduates would be smaller.

³¹Estimation is restricted to this period since consistent measures of tuitions and curricula are

NCES data we estimate ν by

$$\log(p_{it}) = \beta_0 + \nu \log(m_{it}) + a_t + \epsilon_{it}$$

where the subscript i indexes colleges and t indexes time periods. The time fixed effects a_t are included to capture nominal effects in tuition levels and any other time trends in the cost of college. To ensure our estimates are representative, we weight all school observations by the total number of degrees they award in a given year. The estimation results as well as a number of alternative robustness checks are included in appendix Table 3.

5.5 Empirical Results

The two estimated shock series are plotted in Figure 8, along with their 95% confidence sets. The panel on the left plots the relative specialization shocks $f_{h\tau}/f_{c\tau}$ for each cohort τ against the cohort's high school graduation year. The panel on the right plots the ratio of the sectoral technology series A_{ct}/A_{ht} for each year.

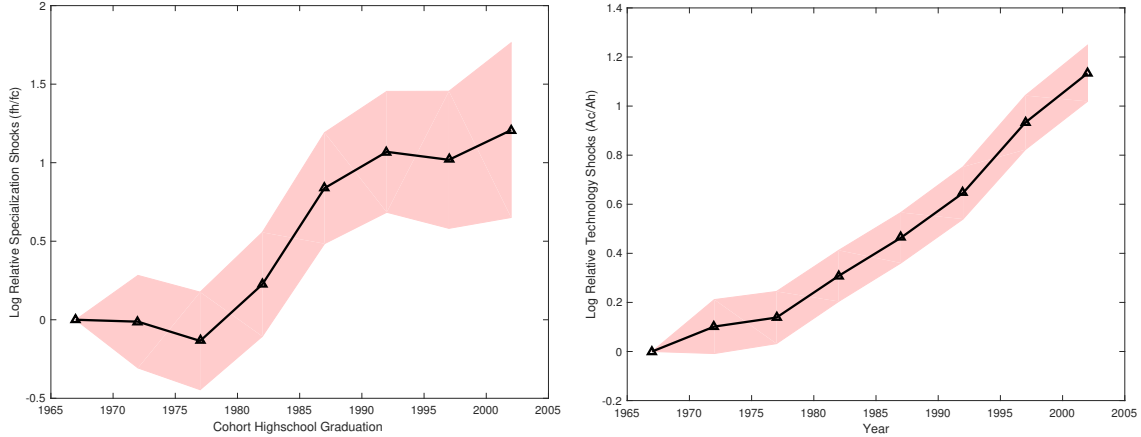


Figure 8: Estimated Model Shocks

The first thing to note is the strong secular trend in relative sectoral technologies, consistent with the literature's findings that models with linear trends in

available for the universe of colleges. Including observations from the NCES from earlier periods would require harmonization of the tuition data.

skill-biased technological change fit the data very well. Here, we did not impose any assumption of linearity in our estimation procedure and yet the results confirm, non-parametrically, the presence of a nearly log-linear trend in relative technologies. *Ceteris paribus*, the rate of increase in the ratio of sectoral technologies corresponds roughly to a 1.5% annual trend in the aggregate production function³², which is roughly one-half to three-quarters of the reduced-form, partial-equilibrium estimates in the literature.

In addition to the trend in relative technologies, the estimates indicate a statistically significant trend break in the costs of specialization beginning for cohorts entering high school in the mid 1970s. Note that this is above and beyond what could be explained by variations in cohort-specific attainment levels. From the confidence bands, it is clear to see that one can statistically test and reject that the levels are the same before and after the level shift. Consistent with our theory, the break coincides closely with the peak and subsequent demise of mainstream vocational training at the secondary level in the United States and continues through the early push toward standardization and college-preparatory oriented curricula in the mid 1980s. The fact that the timing of the break in the specialization shock series coincides with the historical decline in specialized secondary education and rise in standardized, academic curricula strongly suggests it as a candidate explanation for the transition.

Before decomposing the contributions of each set of shocks, we can evaluate the model's ability to fit the aggregate trends. Figure 9 plots the time-series data on the college wage premium and relative supply of college graduates from the census data alongside the model's predictions of these series, derived by feeding the shocks back through the model and calculating the off-equilibrium transition paths. All simulations are constructed subject to rational expectations, where subsequent generations of households correctly anticipate the sequence of shocks occurring during their lifetime. The figures illustrate that the model does very well at capturing the dynamics of both the skill premium and relative supply, especially considering that the shock estimation procedure did not explicitly target either

³²Recall the aggregate production technology $Y_t = [(A_{ct}Y_{ct})^\gamma + (A_{ht}Y_{ht})^\gamma]$ so that the trend in the relative technology series within the aggregate production function corresponds to the estimate shocks above raised to γ .

of these two series and nearly all the parameters were set using independent estimates from the literature. The model does well even with the last two data points for the bins around 2007 and 2012, both of which are out-of-sample for our estimation procedure. The model and data deviate most in the late 1970s where the model under-predicts the attainment levels and, partially in consequence, over predicts the college wage premium. One reason for the deviation is that in the data the share of college workers in a given cohort changes over time, due to chance, differential labor market attachments, and late matriculation, while in the model these forces are absent, so the level of attainment within a cohort stays constant once in the labor market. Nevertheless, the key aspects of the data which we seek to decompose are captured by the model, those are the timing and magnitude of the run-up in the skill-premium and long run trends in attainment rates.

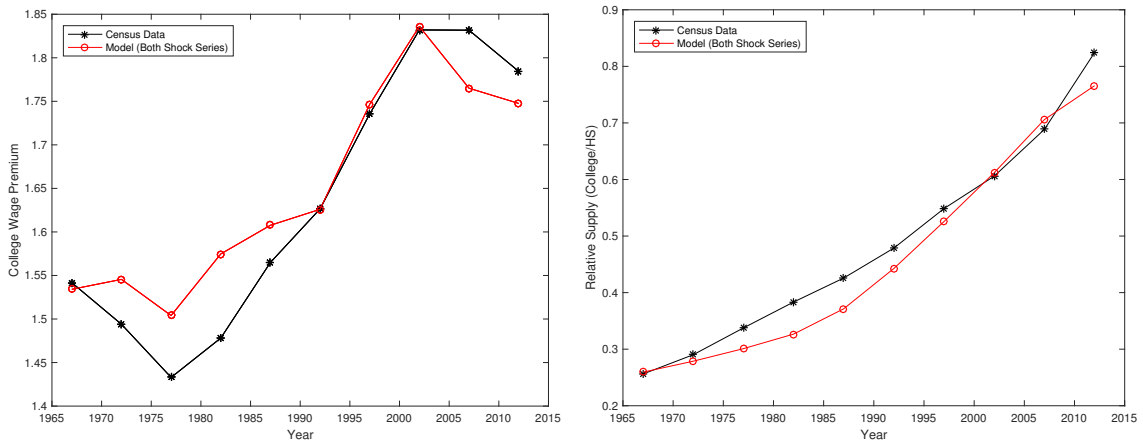


Figure 9: Model Fit: Wage Premium and Educational Attainment

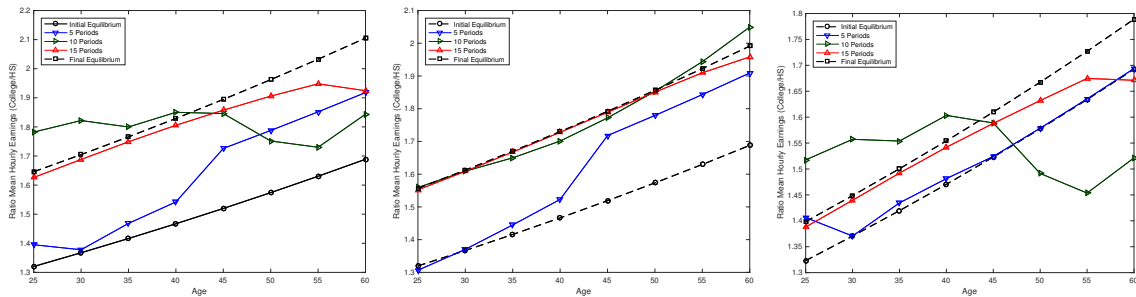


Figure 10: Simulated Cohort Profile Transitions

Figure 10 illustrates how our identification scheme worked in backing out the

estimated shocks. The figure plots the cohort profiles—the skill premium by age group in each year—at the beginning, in the middle, and at the end of the transition dynamics. The first panel plots the result when both shocks are fed through the model, the middle panel does so for only the technology shocks, and the last panel deals with only the specialization shocks. Note that the cohort profiles are linear here, rather than hump shaped, due to our assumption that learning on the job occurs at a constant rate and that labor market attachment for both education groups is constant over the lifecycle³³. As in the stylized facts in figure 2, the first panel shows how the cohort profiles generated by the model pivot and flatten out as the economy transitions to the new equilibrium where the profile is parallel and shifted up from its original location. Furthermore, in keeping with our identification strategy, the next two panels show that this pivoting is due primarily to the specialization shocks. The figures show that, given our parameterization, the general equilibrium spillovers from changes in relative technologies do not contribute much to pivots in the cohort profile. Taken together, the results suggest that the cohort effects in the college wage premium documented in the literature may be wholly the result of changes in the education curricula documented above.

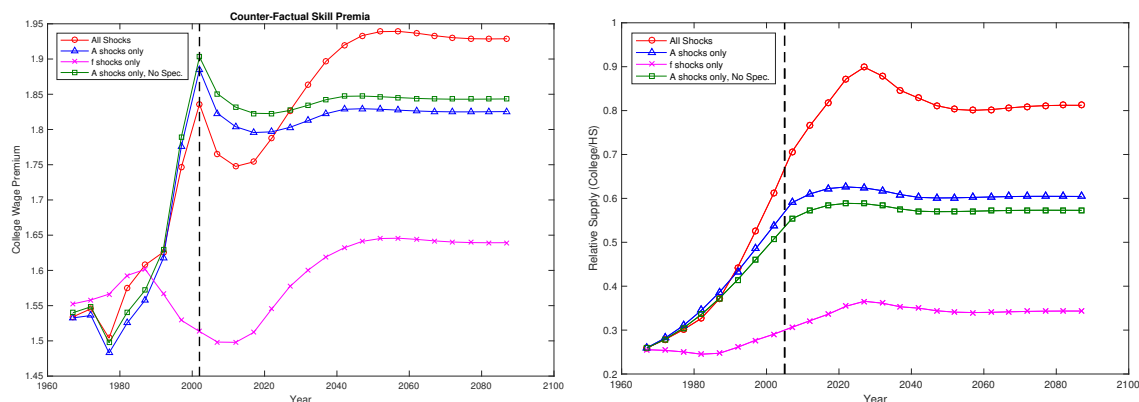


Figure 11: Simulated Long-Run Equilibria

Finally, we can turn to quantifying the relative contribution of the two shocks to the skill premium and attainment trends. One challenge in establishing an accurate decomposition is that off the equilibrium path the alternative shock series generate

³³Using more general returns to experience on job to get hump shaped cohort profiles does not change the main results below but substantially increases the number of parameters in the model. Furthermore, no readily available estimates of such parameters are available in the literature.

substantially different transition dynamics. As a result, the relative contribution of each shock will depend on the particular point along the transition chosen to conduct the evaluation. To avoid these issues, I conduct the decomposition by focusing on the long-run equilibrium which the model converges to as a result of each of the shock processes. The conceptual exercise is to consider the long run equilibria that would emerge if after the shock series we estimated no other changes occurred and the economy settled down into a new steady state. In the long-run steady state that emerges the contributions of the shock series would be constant and could be exactly decomposed into additive contributions. Figure 11 illustrates the long run convergence of the several shock scenarios studied and how they can admit additive decompositions. The dashed vertical line signifies the end of the simulation period constructed by feeding the estimated shocks through the model and the period following it is determined by the implied dynamics of the model adjusting to the series of shocks. By comparing the long run level differences of each of the scenarios we can assign relative contributions to each shock series in an effort to quantify its importance.

The estimated contributions are summarized in Table 2. The top panel reports percent changes and the bottom panel reports the corresponding levels. In-sample entries correspond to the period from 1965-2005 used to estimate the model shocks. Column 3 reports the model's prediction of the long-run equilibrium the economy converges to as a result of both the estimated shock series. Columns 4-6 report the long-run results of each shock series separately. Note that columns 4 and 5 provide a decomposition in that they always add up to the model's long-run predictions in column 3. Column 6 reports the results of simulating only the sector specific technology shocks while preventing the endogenous response in relative specializations it induces. The purpose of the final column is to get a complete sense of the role played by human capital specialization in the model.

The results in table 2 suggest a quantitatively important role for human capital specialization. Comparing the counter-factual estimates in columns 4 and 5 to their total in column 3 suggests that the demise in specialized training in secondary schools can account for nearly one-third of the rise in the skill premium and one-quarter of the rise in the college attainment ratio. If one takes a broader view

	Data	Model	Counterfactuals			
	<i>in-sample</i>	<i>in-sample</i>	<i>long-run</i>	<i>f shocks</i>	<i>A shocks</i>	<i>A shocks, no spec</i>
<i>Changes</i>						
<i>Wage Premium</i>	17%	18%	24%	7%	17%	18%
<i>Relative Supply</i>	86%	86%	113%	28%	85%	80%
<i>Levels</i>						
<i>Wage Premium</i>	1.83	1.84	1.93	1.64	1.83	1.85
<i>Relative Supply</i>	0.61	0.61	0.80	0.34	0.60	0.57

Table 2: Shocks Contributions by Type

of the total contribution of human capital specialization in the model, accounted for in column 6, its effect on the wage premium is largely unchanged but its contribution to college attainment rises to roughly thirty percent. The contrast is intuitive and results from the fact that specialization in human capital is a source of labor productivity. As skill biased technological change raises the demand for college graduates it also incentivizes them to invest more in specialized training to take advantage of the higher wages in their sector. The result is a knock on effect which raises the return to a college education above and beyond the direct effect of technological change and hence leads more individuals to choose college matriculation. If, as in column 6, the level of specialization does not adjust, then the net return to college is lower and so less individuals find it worthwhile to pursue a college degree.

6 Cross-Country Evidence

In the preceding section, we reviewed the evidence in favor of the scope of specialization mechanism in the United States time series using a structural model. In this section, we review the international evidence. While the rise in the skill premium and accompanying rise in the educational attainment has been relentless in the United States over the last half century, the experience across other developed countries has been varied. Recent work by [Krueger et al. \(2010\)](#) highlights the differential experiences across countries using detailed microdata spanning

roughly the two decades from the 1980s to the early 2000s. At one end, there is the United States where the college wage premium grew by 0.4 accompanied by large increases in educational attainment. At the other end, there is Germany where the college wage premium *fell* by 0.08 over roughly the same period and with a college attainment level amongst the lowest in the OECD³⁴.

In attempting to make sense of the cross-country variation in the skill premium the potential explanations proliferate. Nevertheless, it is clear that the technological or supply driven hypotheses widely embraced in explaining the experience of the United States would struggle with the international observations. There is scant evidence to suggest that the level of technology in general, or computerization in particular, differs sufficiently across developed countries to explain the magnitude of differences in the college wage premium. Furthermore, supply driven explanations would struggle with the fact that considerable variation in the wage premium persists even after supply effects are accounted for. The cross-country variation therefore suggests that an institutional framework may be necessary to reconcile the facts, one with a broader perspective on the role of education and the acquisition of skills. Crafting such a hypothesis would require there to exist significant variation in the organization and content of educational systems across developed countries. While it would be hard to argue for such a distinction at the tertiary level, especially for signees of the Bologna accord, there is an unequivocal basis to argue that such variation exists at the secondary level.

As an illustration, consider again comparing the United States and Germany, whose long-term dynamics in the skill premium and systems of secondary education span the gamut of those observed in the cross-section. Details of the secondary education system in the United States were reviewed in section 3 and highlight its focus on standardized and academically oriented college preparatory curriculum. Germany, in contrast, maintains a multi-track, multi-tier secondary school systems where students can earn specialized degrees and certifications in

³⁴KPPV results actually suggests that the experience of Spain was even more extreme, with decline in the college wage premium of 0.33. However, there is reason to believe that Spain's experience is an outlier given that the period studied coincides exactly with Spain's ascension to the European Union and an accompanying boom in economic activity and sharp rises in educational attainment. The data on Spain also corresponds to a shorter time period ending in the mid-1990s in contrast to the USA data which spans 1980-2006 and Germany which spans 1983-2003.

roughly 350 distinct nationally recognized occupations³⁵. Around 2/3 of German secondary students select into one of these vocationally tracked programs.

Broadly speaking, non-tertiary schools in Germany separate students into vocational or academic tracks by ages 10-12. By 5th grade, students can choose one of three schools to pursue: *Hauptschule*, *Realschule*, or *Gymnasium*³⁶. The *Gymnasium* spans grade 5 through 12 and is the most academic of the three tracks, leading to an *Abitur* degree which qualifies students to attend university upon graduation. In terms of content, the curriculum of the *Gymnasium* is most comparable to that of mainstream U.S. high schools.

The *Realschule* curriculum is less academic and more practically oriented than the *Gymnasium*, focusing on preparing students for the labor market. The program spans grades 5 through 10 and graduates of this track often spend their final 2-3 years of school either in an apprenticeship program or specialized full-time vocational schools awarding certifications for certain occupations. Graduates of the *Realschule* also have the option of spending their final two years in the *Gymnasium* to earn the *Abitur* or to enter the specialized technical *Fachoberschulen* schools in preparation for a polytechnic college.

The final track, the *Hauptschule*, is the least academic of the three and focuses on basic education and vocational skills. As with the *Realschule*, this track terminates in 10th grade with most graduates entering apprenticeship programs in various industries. Alternatively, students also have the option to transfer to the *Realschule* in the final years to earn that qualification.

Students who choose a specialized occupational track spend roughly 3 days a week at an apprenticeship, being trained and getting hands on experience, and 2 days attending a *Berufsschule*, specialized vocational schools linked to their particular occupation. The allocation of time in the vocational track is similar to the 50-25-25 rule that characterized early American vocational education. The *Berufsschule* are divided by trade and industry and are intended to provide a generalized context to the learning that goes on at work. Alternatively, students

³⁵The precise number of occupational degrees varies slightly between years

³⁶Alternatively, there is also the *Mittelschule* which combines the first two tracks, as well as the *Gesamtschule* which combines all three

may also choose to pursue full time specialized vocational training without an apprenticeship at a *Berufsfachschulen*, though this track is less common than the apprenticeship scheme. A number of other remedial vocational training schools are also available for those who have not earned intermediate secondary degrees, were not accepted to apprenticeships, or have special learning needs.

The content and graduation requirements of the vocational tracks are tightly regulated by a federal umbrella organization known as the *Bundesinstitut fuer Berufsbildung* (BiBB). While many stakeholders contribute to the process, it is German enterprises whose labor demands the program is meant to serve—through the affiliated Chamber of Commerce and Industry—which have the largest hand in shaping the content, implementation, and graduation requirements of each program. The system ensures that the vocational training received does indeed fit the labor market demands of German companies. Surveys of participating companies conducted by the BiBB in 2004 found that 94% of companies that offer apprenticeships say they do so because trainees fit the company's needs best and 93% said the programs allow the acquisition of specialists not available in the labor market. The stated motivations of participating businesses highlight again the tight link between the value of human capital specialization and a production process that makes use of the division of labor. In addition to shaping the curriculum, German enterprises also shoulder the lion's share of the financing for vocational training programs and apprenticeships, contributing roughly 71% of the \$39 billion annually spent to maintain the programs on top of what they pay in apprenticeship wages ([Hippach-Schneider, Krause, and Woll, 2007](#)).

It is interesting to compare the roots of the modern German system to the historical experience of the United States. Germany, like the United States, faced a crisis of confidence in its secondary vocational program in the late 1960s. Again, as in the United States, there were concerns that many of the programs were out of step with the changing demands of the labor market, that course taking had become diffuse, and that funding and planning wasn't adequate. However, while the response in the United States was to pursue policies which subsidized and supported college matriculation, Germany chose to increase investment and planning in its vocational programs. Somewhat ironically, the 1969 law which established the

BiBB and modernized the secondary vocational programs in Germany was called the Vocational Training Act – the same name borne by the legislation in the United States which would eventually lead to the demise of secondary vocational training in this country. The key reforms in Germany that were missing in the United States seem to have been the creation of a federal authority which could recognize or remove occupations from a national system of specialized secondary degrees and the close involvement of businesses in funding and shaping the vocational programs. The close involvement of enterprises and labor groups was, in fact, also a key feature of the early decades of vocational education in the United States.

Despite the historical causes, the contrast between secondary education in Germany and the United States today could not be clearer. Germany offers students a choice between an academic track or a variety of highly specialized, occupationally oriented tracks while the United States predominantly offers an undifferentiated, academic college-preparatory curriculum. Most other countries fall somewhere in between these poles, for example Canada and Mexico are more like the United States while Switzerland, Austria, and Denmark are more like Germany.

The discussion above makes clear that the differences between countries at the secondary school level offer substantial scope to test the implications of our economic framework. Figure 12 takes a first step toward evaluating the cross-country evidence in terms of our model by plotting cross-country growth rates in the college wage premium calculated by Krueger et al. (2010) against the share of upper-secondary school students receiving specialized vocational education³⁷. While focusing on changes in the college wage premium restricts us to a small sample of countries for which estimates are available, the exercise has the benefit of avoiding the pitfalls of comparing levels of wages across countries. Consistent

³⁷Data on vocational training is taken from the OECD.Stat database for all countries except the United States, for which no estimate exists. For the OECD.Stat data, the shares correspond to the 1998-2000 averages, the earliest data available which overlaps with our study period. For Canada, the earliest data available is for 2012. The OECD derived results are not sensitive to using longer average estimates or even averages derived from the full sample 1998-2012. The number for the United States corresponds to the share of "vocational specialists" found in the NCES High School Transcript study. The definition of vocational specialists are those students with at least 4 credits in a single occupational area with at least two of them above the introductory level. This estimate is likely an upper bound on true level of vocational specialization in the United States relative to foreign benchmarks.

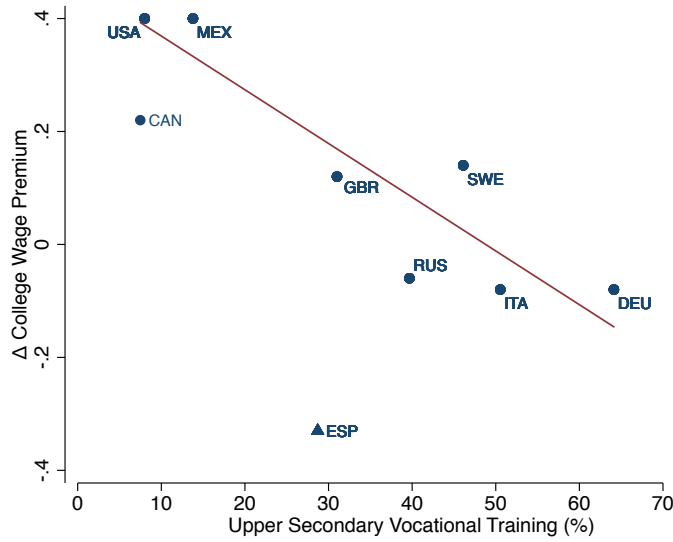


Figure 12: Long-Term Change in College Wage Premium and Secondary School Specialization

with our model, there appears to be a robust link between the two variables with countries that offer more vocational training experiencing less of an increase in the college wage premium³⁸.

If we focus on purely cross-sectional variation, we may extend our analysis to all countries in the OECD database. Figures 13 through 18 do precisely this. Figure 13 plots the implied ratio of labor productivity between college and high school educated workers against the intensity of secondary vocational training³⁹. While the productivity ratio is typically associated with skill-biased technology, the figure here documents a significant negative correlation between the measured ratio and the extent of secondary educational specialization. The result is consistent with our theory which argues that more specialized secondary education increases the productivity of the secondary school population by allowing more efficient use of time invested in human capital. Similarly, figure 14 documents the significant

³⁸As already discussed above, however, Spain is an outlier given that its sample period is short and overlaps with its ascension to the European Union and subsequent booms in the economy and educational attainment.

³⁹The level of skill bias is calculated using the canonical model in the skill premium literature assuming that the elasticity of substitution is common across countries and equal to 2. Specifically, the skill bias is calculated as $A_c/A_h = \log(W_c/W_h) + 1/2\log(L_c/L_h)$.

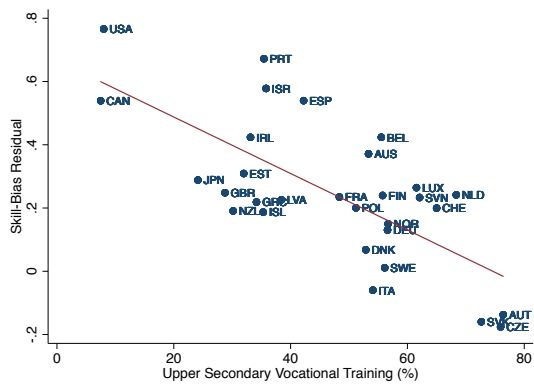


Figure 13: Skill-Bias

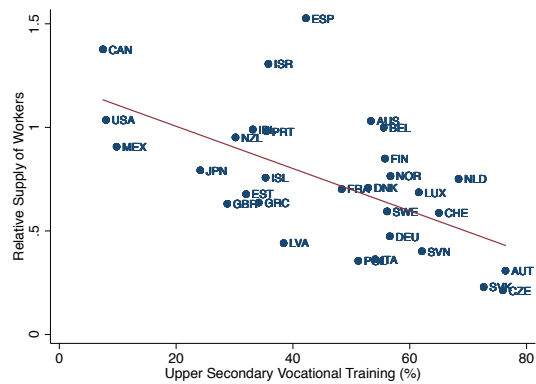


Figure 14: Relative Supply

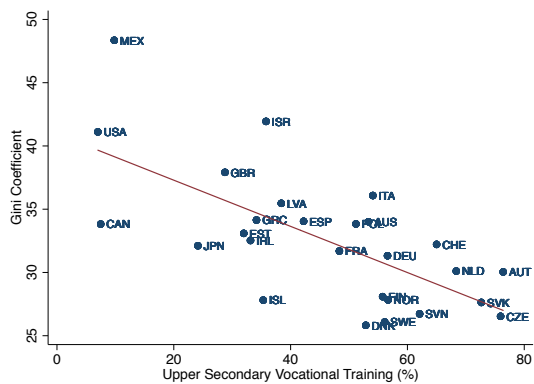


Figure 15: Gini coefficient

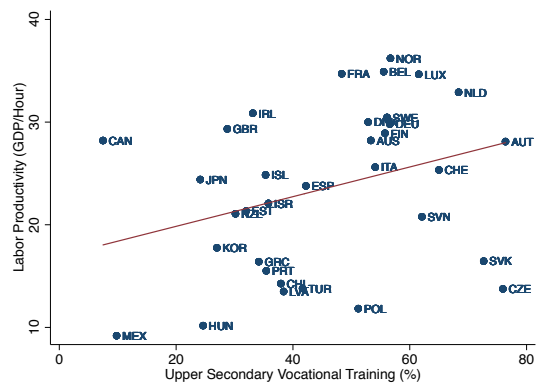


Figure 16: Labor Productivity

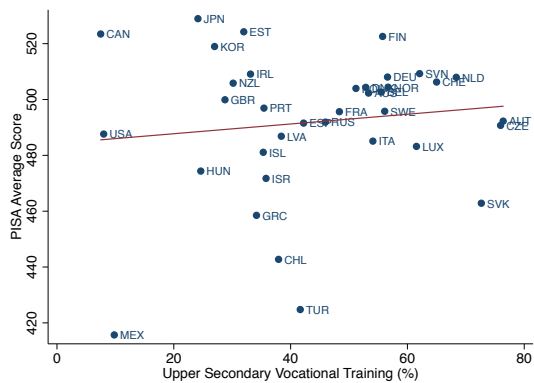


Figure 17: Average PISA Scores

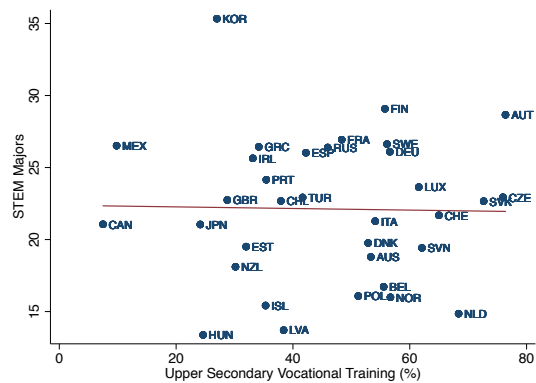


Figure 18: Tertiary STEM Degrees

negative relationship between the college attainment ratio⁴⁰ and the intensity of secondary vocational training, indicating that countries with more specialized secondary school training have a smaller fraction of workers who attend college. This too is predicted by our framework, as the relative return to college versus high school is smaller when workers can specialize in occupationally oriented human capital at the secondary level.

Figures 15 and 16 go beyond the immediate implications of the model to document the relationships between specialization at the secondary level and broader measures of human capital and inequality. Figure 15 plots the country wide Gini coefficient and documents that countries with more specialized secondary education have lower levels of income inequality overall. The result is consistent with the spirit of our model that specialized training at the secondary level allows a larger fraction of the population to efficiently invest their time in human capital acquisition. It is also consistent with the idea that specialized secondary technical education is a path to middle-skill jobs, precisely those that are disappearing from the United States and driving part of the rise in inequality (Autor and Dorn, 2013). Figure 16 plots the overall level of labor productivity against the intensity of vocational course taking. Though here the relationship is more tentative, the positive correlation again suggests that allowing specialization in secondary education can increase the efficiency of human capital investments in the non-tertiary education sector and hence raise the labor productivity of the workforce overall if the secondary school population is sufficiently large.

Finally, figures 17 and 18 assess whether the results are driven by other unobserved differences in the quality of education systems which are correlated, but unrelated to, the intensity of secondary vocational training. Figure 17 plots average PISA scores against the intensity of secondary vocational training and shows no notable correlation. This rules out that our results are being driven by the possibility that countries with high vocational training may also possess an overall better quality of secondary schools which raises the wages of high school graduates independently of specialization. Similarly, figure 18 looks at the share of tertiary degrees awarded in STEM fields. The concern here is that countries with high

⁴⁰College educated workers are defined by ISCED 5-6 categories, high school educated workers correspond to ISCED 3-4.

vocational training may produce less STEM graduates and therefore have lower college wage premiums, independently of the secondary training regimes. The lack of a correlation in the figure between STEM tertiary degrees and share of secondary vocational training seems to rule out this possibility as well. Taken together, these figures suggest that it is indeed the differences in secondary curricula specialization that is driving the results.

7 Conclusion

This paper introduces a new channel for human capital accumulation driven by increasing specialization and the division of labor. The mechanism, while novel in the modern literature, is based on a long tradition in economic thought dating back at least to [Smith \(1776\)](#). The usefulness of the theory as an organizing conceptual framework is demonstrated by its ability to simultaneously reconcile several prominent stylized facts in the literature on the skill premium. Taking the model to the data confirms its quantitative relevance, as the mechanism can account for nearly one-third of the rise in the college wage premium and attainment ratio in recent decades.

The framework above also suggests several promising avenues for future research. For example, in the empirical literature it is common practice to measure the contribution of human capital by average educational attainments. The arguments developed above, however, suggest that this is an insufficient and perhaps even misleading statistic; as the model illustrates, allowing specialization at the secondary schools increased the efficiency of human capital investment at that level while decreasing the incentive for college matriculation. Gaining a better understanding of the precise interaction of these two channels for human capital accumulation then seems to be an important and empirically relevant path for future research.

Even more subtly, the framework above provides interesting implications for the optimal provision of human capital. Recently, many prominent researchers have suggested that the fast pace of technological progress and automation demand general skills training for workers to be resilient in an ever changing labor market. While this observation is undoubtedly true, it neglects the fact that technological

progress has been accompanied by substantial increases in the division of labor, with many jobs now demanding a narrower set of specific technical skills. In this environment, the breadth provided by general skills training can fail to provide the level of narrow expertise demanded in a specialized production process. Put another way, while general skills may be complementary to increased mechanization, it is also true that specialized skills are complementary to increases in the division of labor. More work is needed to better understand the tension between these two forces on the optimal provision of human capital in a modern economy.

Finally, and perhaps most importantly, the arguments advocated here have important policy implications. As already stated, there has been much discussion in economics, policy, and press circles about the growing need to address the yawning wage gap between different segments of society. Partly in homage to many of the existing theories in the literature reviewed above, the most commonly proposed remedies often include changes to compensation laws, trade policy, and increasing access to higher education. A different perspective emerges, however, in reviewing again the history of education and the skill premium in the United States. The early 20th century contraction in the wage gap coincided with the transformation of secondary schools from simple preparatory institutions into major loci of occupational training. There was often a tight articulation between the secondary school curriculum and the fast growing occupations of the day. However, in recent decades, this focus on specialized occupational preparation has disappeared from high schools. Where it does still exist, it is often focused on the same occupational and industrial sectors that it was a century ago. This need not be the case. Occupational education at the secondary level can be revitalized to reflect today's fast growing labor demands for certain *technical skills*, such as software programming, technician work, and certain healthcare jobs. Some community colleges and private un-accredited institutions—such as the quickly proliferating "hacker schools"⁴¹—are already seeing and effectively responding to the demand for this type of specialized technical training around the country.

⁴¹See for example "Web-Era Trade Schools, Feeding a Need for Code", *The New York Times*, October 2014; "As Tech Booms, Workers Turn to Coding for Career Change", *New York Times*, July 2015; and the many related articles in the series.

Much more can be done to integrate their approach into the secondary school system.

While the university will always have an important role to play in teaching certain advanced skills and training professionals, as well as its more general function as an institution of cultural enrichment and research, it need not also be the sole locus of job training in the modern economy. This is especially unfair when one considers that a large part of the financial cost associated with college attendance and admission reflects the fact that college is also a consumption and leisure good. This conflation, while giving college its unique and lasting character, also imposes barriers to job training that are inessential and prevent many, particularly the less advantaged, from gaining access. With this perspective, the relationship between college attendance and income inequality may be turned on its head. Tackling inequality need not require increasing attendance at costly and exclusive colleges, but rather removing the need to do so in seeking the meaningful, practical, and useful training necessary for a modern worker.

A Appendix: Proofs

Proof Proposition 1. Using the fact that the production function is CRS we can rewrite output

$$Y_t = A_t F(s_{1t}, \dots, s_{Kt}) L_t \quad (3)$$

where L is total effective hours at time t and $s_{\kappa t}$ is the share of effective hours dedicated to task κ , where clearly the shares must sum to unity. Since all labor demand curves slope downward and the cost of acquiring any skill is identical, in equilibrium effective hours will be distributed so as to equate marginal products and hence also task wages, $w_{\kappa t}$. Equating the first order conditions then for shares s_i, s_j we can simplify so that $F_i = F_j$ for all i, j and so the equilibrium distribution of effective hours across tasks does not depend on A .

Let s^* denote the K dimensional vector of equilibrium shares of effective hours across tasks. The first order condition with respect to total effective hours pins down the common wage rate per effective hour as

$$w_t = AF(s^*)$$

which grows at the rate of technology, g_a , along the balanced growth path. To economize on notation, let us normalize $F(s^*)$ to unity⁴². Plugging the wage into the period t earnings for an individual of age j who chose to study for s years yields

$$\begin{aligned} I_t(a, s, m, \ell) &= w_{t+j} e^{g_e(j-s)} a(sm)^\theta \\ &= w_t e^{g_a j + g_e(j-s)} a(sm)^\theta \end{aligned}$$

Plugging into the expression for lifetime income net of schooling costs, $W(s, m)$, and substituting in the equilibrium expression for tuition levels from equation 2 yields⁴³:

$$W(s, m) \equiv \int_s^\infty e^{-(r+\delta-g_a-g_e)j-g_e s} w_t a_i s^\theta m^\theta \partial j - f m^\nu$$

⁴²This is without loss of generality since it is a constant not separately identified from the level of technology A .

⁴³Assuming $r + \delta - g_a - g_e > 0$ so discounted values are finite

and taking first order conditions with respect to m yields

$$[m] \quad \frac{e^{-(r+\delta-g_a)s}}{r+\delta-g_a-g_e} \theta w_t a_i s^\theta m^{\theta-1} = \nu f m^{\nu-1}$$

which can be solved for explicitly for m

$$m = \left[\frac{e^{-(r+\delta-g_a)s}}{r+\delta-g_a-g_e} \frac{\theta}{\nu f} w_t a_i s^\theta \right]^{\frac{1}{\nu-\theta}} \quad (4)$$

and then plugging back into the the expression for net lifetime income yields

$$W(s) \equiv \left(1 - \frac{\theta}{\nu}\right) \left(\frac{\theta}{\nu f}\right)^{\frac{\theta}{\nu-\theta}} (r+\delta-g_a-g_e)^{-\frac{\nu}{\nu-\theta}} e^{-\frac{\nu(r+\delta-g_a)s}{\nu-\theta}} w_t^{\frac{\nu}{\nu-\theta}} a_i^{\frac{\nu}{\nu-\theta}} s^{\frac{\nu\theta}{\nu-\theta}} \quad (5)$$

whose only choice variable is now s . Next, we take the first order condition with respect to consumption and differentiate with respect to time yielding the standard continuous time Euler equation

$$[c] \quad \frac{\dot{c}_t}{c_t} = r - \rho \quad (6)$$

and hence for simplicity we assume $r = \delta$ so that we have a flat consumption profile⁴⁴ which, combined with the budget constraint, allows us to solve for equilibrium consumption level in each period as

$$\bar{c} = (r + \delta)W(s) \quad (7)$$

where we use the expression for lifetime income in terms of s only, $W(s)$, in equation 5. Plugging \bar{c} into the utility function and taking first order conditions with respect to s allows us to solve for optimal schooling:

$$[s] \quad \frac{1}{\rho + \delta} \left[-\frac{\nu(r + \delta - g_a)}{\nu - \theta} + \frac{\nu\theta}{\nu - \theta} \frac{1}{s} \right] - \phi_i = 0$$

rearranging yields an analytical expression for the optimal years of schooling for

⁴⁴Note that the assumption $r = \delta$ is not necessary but made for convience to simplify the notation

individual i as

$$s(\phi_i) = \frac{\nu\theta}{(\nu - \theta)(\rho + \delta)\phi_i + \nu(r + \delta - g_a)} \quad (8)$$

which makes clear that individuals with higher opportunity cost of schooling, ϕ_i , will dedicate less time to schooling. Plugging back into the expression for the optimal level of specialization we have

$$m_t(a_i, \phi_i) = \psi(\phi_i) f^{-\frac{1}{\nu-\theta}} w_t^{\frac{1}{\nu-\theta}} a_i^{\frac{1}{\nu-\theta}} \quad (9)$$

where

$$\psi(\phi_i) \equiv \left[\frac{e^{\frac{-\nu\theta(r+\delta-g_a)}{(\nu-\theta)(\rho+\delta)\phi_i + \nu(r+\delta-g_a)}}}{r + \delta - g_a - g_e} \frac{\theta}{\nu} \left(\frac{\nu\theta}{(\nu - \theta)(\rho + \delta)\phi_i + \nu(r + \delta - g_a)} \right)^\theta \right]^{\frac{1}{\nu-\theta}}$$

is a function only of parameters and so (net) lifetime income for individuals born in time t can be expressed

$$W_t(a_i, \phi_i, w_t) = \left(\frac{\nu - \theta}{\theta} \right) \psi(\phi_i)^\nu f^{-\frac{\theta}{\nu-\theta}} w_t^{\frac{\nu}{\nu-\theta}} a_i^{\frac{\nu}{\nu-\theta}} \quad (10)$$

We can now characterize the balanced growth path. First, note that the total effective hours supplied by individuals of age j in the labor market at time t can be written

$$L_{t,j} = \int_\phi \int_a e^{g_e(j-s(\phi))} a s(\phi)^\theta m_{t-j}(a, \phi)^\theta dF(a) dG(\phi) \quad (11)$$

Plugging in the expression for m and taking time derivatives yields

$$\frac{\dot{L}_{t,j}}{L_{t,j}} = \frac{\theta}{\nu - \theta} \frac{\dot{A}_t}{A_t} \equiv \frac{\theta}{\nu - \theta} g_a$$

for all age groups j . As a result, total effective hours supplied at time t grows at the same rate since

$$L_t = \int_j L_{t,j} dj$$

And since the final goods production function is constant returns to scale we have

that aggregate output grows at rate

$$\frac{\dot{Y}_t}{Y_t} = \frac{\dot{A}_t}{A_t} + \frac{\dot{L}_t}{L_t} = \frac{\nu}{\nu - \theta} g_a$$

Similarly, we see from equation 10 that (net and gross⁴⁵) lifetime earnings are growing at rate

$$\frac{\dot{W}_t}{W_t} = \frac{\nu}{\nu - \theta} g_a$$

for all individual types. From equation 7 then we have that consumption for each individual type is also growing at the same constant rate and so to then is aggregate consumption.

Furthermore, along the BGP the distribution of educational attainments (length of study s) is constant but human capital is still growing at a constant rate. The constancy of attainment levels follows directly from the equilibrium expression for s in equation 8, which only depends on fixed parameters of the model and the time-invariant distribution of costs ϕ . Despite the constancy of attainments, aggregate human capital is still growing at a constant rate $\frac{\theta}{\nu - \theta} g_a$. The source of human capital accumulation here is the constant growth rate in specialization which can be seen directly from equation 9.

To see this more directly, recall that initial level of human capital for a worker of type (a, ϕ) is $h_0(a, \phi) = as(\phi)^\theta m(a, \phi)^\theta$, which is growing at rate $\frac{\theta}{\nu - \theta} g_a$ through specialization, independently of type. Since human capital at any stage of life is just equal to the initial level of human capital augmented by a constant rate of return to experience g_e , this means that subsequent generations human capital is growing at a constant rate at each point in life for all workers and therefore so it aggregate human capital. This growth in human capital through specialization is actually what is driving the growth in effective hours supplied by all workers and is therefore responsible for a fixed fraction $\frac{\theta}{\nu}$ of the growth in output and incomes.

Finally, we establish the ratio of wages across education groups and age groups is constant along the BGP. We show this by establishing that earnings for all types of workers at each point in life are growing at the same constant rate. Specifically,

⁴⁵Note that gross lifetime earnings is just net lifetime earnings times $\frac{\theta}{\nu - \theta}$

the income of individual i at age j in time t can be expressed

$$I_{ijt} = w_t e^{(j-s(\phi_i))g_e} a_i s(\phi_i)^\theta m_{t-j}(a_i, \phi_i)^\theta \quad (12)$$

which grows at constant rate

$$\frac{\dot{I}_{ijt}}{I_{ijt}} = g_a + \frac{\theta}{\nu - \theta} g_a = \frac{\nu}{\nu - \theta} g_a$$

Furthermore, since the distribution across attainment levels is constant along the BGP, the ratio of *average earnings* (whereas the above is total earnings) also remains constant as average values grow at the same rate since population and cohort sizes do not change.

□

Proof Proposition 2. The proof follows immediately from the characterization of the BGP in proposition 1. From equation 8 it is clear that the level of educational attainment is not effected by f . From equation 9 its clear that specialization falls (rises) when f increases (decreases). The effect this has on incomes and consumption then follows from equations 7, 10, and 12. The effect on output then follows immediately from the fall in effective hours supplied, seen in equation 3.

To see that effect on cross-age incomes we can take the ratio of two individuals of the same type of ages j and k where $j < k$ using equation 12 so that the ratio of the young to the old is

$$\frac{I_{jt}}{I_{kt}} = e^{(k-j)(\frac{\theta}{\nu-\theta}g_a - g_e)} \left(\frac{f_{t-k}}{f_{t-j}} \right)^{\frac{\theta}{\nu-\theta}}$$

where f_τ is the compound specialization cost parameter for individuals born in period τ . Hence, if the specialization cost parameter rises after cohort $t - k$ enters the market but before $t - j$ does, the income of the young will fall relative to the old, as in the proposition. \square

Proof of Proposition 3. Consider a shock that raises the cost of specialization to $\hat{f} > f$ for individuals who choose $s < s_{col}$. The first claim in the proposition is that the earnings of those with attainment level s_{col} rises relative to those with $s < s_{col}$. The relative average earnings of individuals in age groups τ and τ' can be written

$$\begin{aligned} \frac{I(s_{col})}{I(s)} &= \frac{E(w_t e^{g_e(\tau-s_{col})} a_i s_{col}^\theta m_{t-\tau}^\theta | s_{col})}{E(w_t e^{g_e(\tau'-s)} a_i s^\theta m_{t-\tau'}^\theta | s)} \\ &= \frac{e^{g_e(\tau-s_{col})} w_{t-\tau}^{\frac{\theta}{\nu-\theta}} s_{col}^{\frac{\theta\nu}{\nu-\theta}} f^{-\frac{\theta}{\nu-\theta}} E\left(a_i^{\frac{\theta}{\nu-\theta}} | s_{col}\right)}{e^{g_e(\tau'-s)} w_{t-\tau'}^{\frac{\theta}{\nu-\theta}} \hat{f}^{-\frac{\theta}{\nu-\theta}} E\left(a_i^{\frac{\theta}{\nu-\theta}} | s\right)} \propto \left(\frac{\hat{f}}{f}\right)^{\frac{\theta}{\nu-\theta}} \end{aligned} \quad (13)$$

when $s < s_{col}$ and where we used the fact that $a_i \perp \phi_i$. The expression makes clear that when the cost of specializing for attainment levels $s < s_{col}$ rises from f to $\hat{f} > f$ the relative earnings between any two individuals, one with a college degree and one with less than a college degree, will increase in equilibrium proportionally to the change in relative costs. This is the first part of the proposition.

The second part of the proposition claims that the relative number of individuals who choose attainment level s_{col} relative to those who choose $s \leq s_{col}$ will also weakly increase. The result is not immediately apparent since the optimal expression for attainment choices before the policy change does not directly depend on f . However, after the policy change, there will be a kink in the value functions as a function of attainment level at s_{col} that must be considered separately. To establish the result, I show that there exists a $\phi^* > \phi_{col}$ (where ϕ_{col} is the utility cost cutoff of those choosing college before the change) such that after the policy all $\phi_i \in [\phi_{col}, \phi^*]$ will choose s_{col} whereas before the policy they chose $s < s_{col}$.

Let $s(\phi_i)$ define the attainment levels of those with heterogenous utility cost ϕ_i before the policy change and define $\phi_{col} = s^{-1}(s_{col})$ to be the heterogenous utility cost of those who decide to attend college before the policy shift. Its clear to see that all those with $\phi < \phi_{col}$ chose attainment levels above s_{col} before the policy and will have no incentive to change the choice after the policy shift. The question then is if there are some individuals with $\phi > \phi_{col}$ that are induced to attend college after the policy. We show below that there are a set of individuals near the threshold of college for which this is the case.

Now consider an individual with ϕ such that $s(\phi) < s_{col}$ before the policy change (i.e. $\phi > \phi_{col}$, since $s(\phi)$ is decreasing in ϕ). Let $V(\phi, s(\phi), \hat{f})$ denote the discount lifetime utility for this individual if they stay at their originally optimal choice of s at the new \hat{f} and let $V(\phi, s_{col}, f)$ be their discounted lifetime utility if they switch to attainment level s_{col} . The cost of switching from attainment level $s(\phi)$ to s_{col} can be expressed $\phi(s_{col} - s(\phi))$ and so the individual switches if

$$V(\phi, s_{col}, f) - V(\phi, s(\phi), \hat{f}) \geq \phi(s_{col} - s(\phi))$$

Using equations 5 and 7 we can re-express the condition as

$$(\rho + \delta)^{-1} \left[\frac{\theta}{\nu - \theta} \log \left(\frac{\hat{f}}{f} \right) - \frac{\nu(r + \delta - g_w)}{\nu - \theta} (s_{col} - s(\phi)) + \frac{\nu\theta}{\nu - \theta} \log \left(\frac{s_{col}}{s(\phi)} \right) \right] \geq \phi(s_{col} - s(\phi))$$

Rearranging gives

$$\begin{aligned} \frac{1}{\nu} \log \left(\frac{\hat{f}}{f} \right) + \log \left(\frac{s_{col}}{s(\phi)} \right) &\geq \frac{(\nu - \theta)(\rho + \delta)\phi + \nu(r + \delta - g_w)}{\nu\theta} (s_{col} - s(\phi)) \\ &= \frac{s_{col}}{s(\phi)} - 1 \end{aligned}$$

where we substituted the expression for $s(\phi)$ from equation 8. Now we define the function g such that

$$g(x) = \frac{1}{\nu} \log \left(\frac{\hat{f}}{f} \right) + \log \left(\frac{s_{col}}{x} \right) - \frac{s_{col}}{x} + 1$$

for $x \leq s_{col}$. Taking the first order condition with respect to x :

$$g'(x) = -\frac{1}{x} + \frac{s_{col}}{x^2} > 0$$

which is clearly increasing in x for the domain $x < s_{col}$. Furthermore, define x^* such that

$$g(x^*) = 0$$

Using the fact that g is strictly increasing in the domain and that

$$g(s_{col}) = \frac{1}{\nu} \log \left(\frac{\hat{f}}{f} \right) > 0$$

we have that $x^* < s_{col}$ and furthermore (again by the fact that $g' > 0$)

$$g(x) \geq 0 \quad \forall x \in [x^*, s_{col}]$$

Now define $\phi^* \equiv s^{-1}(x^*)$ and note that since s is decreasing in ϕ we have that $\phi^* > \phi_{col}$. Hence, we've shown that after the policy change all $\phi \in (\phi_{col}, \phi^*]$ who had originally selected attainment levels less than college will now be induced to go to college after the policy change. This proves the second part of the proposition.

The final part of the proposition is that along the transition the cohort profiles for those with attainment levels under s_{col} and those college graduates will flatten out. This result follows immediately from equation 13. Consider the relative earnings of individuals of age $\tau, \tau' > s_{col}$ ⁴⁶ where $\tau' > \tau$ and where those of age τ' graduated before the policy change and those of age τ graduated after the policy change. Let $\tilde{I}(\tau)$ be the *relative* earnings of those with a college degree and those with an attainment level less than college in cohort τ . Using the expressions in equation 13 then we have

$$\frac{\tilde{I}(\tau')}{\tilde{I}(\tau)} = \left(\frac{\hat{f}}{f} \right)^{-\frac{\theta}{\nu-\theta}}$$

Hence after the policy change the gradient of relative earnings between the two groups will decrease since $\hat{f} > f$. This implies that if the gradient slopes upward⁴⁷ the cohort profile will flatten during the transition before reaching its new equilibrium which has the same slope but is shifted upward proportionally to the change in relative costs. This completes the proof of the proposition. \square

⁴⁶The restriction is so that both groups actually have positive earnings since before this age the college age group is still in school.

⁴⁷The gradient always slopes up in the data and will in the model provided that the return to experience is sufficiently high compared to the growth rate of neutral technology A.

Lemma 1 (Ability and Specialization). *Let $m_{e,j}(a)$ be the optimal level of specialization chose by individual in cohort j in sector e with ability level a . We may write*

$$m_{e,j}(a) = a^{\frac{1}{\nu-\theta}} m_{e,j}(1) \equiv a^{\frac{1}{\nu-\theta}} m_{e,j}$$

so that the influence of ability on the optimal level of specialization is multiplicatively seperable so that we may write the optimal choice of specialization of any ability type a within a sector e and cohort j as the product of the individual's ability raised to $\frac{1}{\nu-\theta}$ and the optimal specialization of an individual in the same education and cohort group.

Proof of Lemma 1. The result follow immediately from the FONC for specialization in equation 9. \square

Proof Proposition 4. Relative *total earnings* of individuals age k in time t (cohort $t - k$) of different education levels can be expressed

$$\begin{aligned} \frac{E_{ckt}}{E_{hkt}} &= \frac{w_{cjt} L_{cjt}}{w_{hjt} L_{hjt}} \\ &= \left(\frac{A_{ct}}{A_{ht}} \right)^\gamma \left(\frac{Y_{ct}}{Y_{ht}} \right)^{\gamma-\eta} \left(\frac{L_{ckt}}{L_{hkt}} \right)^\eta \end{aligned} \quad (14)$$

where L_{ejt} is the total effective hours supplied by workers of education level e of age j in period t . Similarly, the relative *total earnings* of cohorts $t - j$ and $t - k$ in time t within the same education sector e is

$$\frac{E_{ejt}}{E_{ekt}} = \left(\frac{L_{ejt}}{L_{ekt}} \right)^\eta$$

solving for L_{ejt} and plugging into the sectoral production function yields

$$\begin{aligned}
Y_{et} &= \left[\sum_j L_{ejt}^\eta \right]^{\frac{1}{\eta}} \\
&= \left[\sum_j \left(\left(\frac{E_{ejt}}{E_{ekt}} \right)^{\frac{1}{\eta}} L_{ekt} \right)^\eta \right]^{\frac{1}{\eta}} \\
&= \left[\sum_j E_{ejt} \right]^{\frac{1}{\eta}} E_{ekt}^{-\frac{1}{\eta}} L_{ekt} \equiv \tilde{W}_{et}^{\frac{1}{\eta}} E_{ekt}^{-\frac{1}{\eta}} L_{ekt}
\end{aligned}$$

where $\tilde{W}_{et} \equiv \sum_j E_{ejt}$ is the total time t income of all workers in sector e . Substituting the expression back into equation 14 and simplifying yields

$$\begin{aligned}
\frac{E_{ckt}}{E_{hkt}} &= \left(\frac{A_{ct}}{A_{ht}} \right)^\gamma \left(\frac{\tilde{W}_{ct}^{\frac{1}{\eta}} E_{ckt}^{-\frac{1}{\eta}} L_{ckt}}{\tilde{W}_{ht}^{\frac{1}{\eta}} E_{hkt}^{-\frac{1}{\eta}} L_{hkt}} \right)^{\gamma-\eta} \left(\frac{L_{ckt}}{L_{hkt}} \right)^\eta \\
&= \left(\frac{A_{ct}}{A_{ht}} \right)^\gamma \left(\frac{\tilde{W}_{ct}}{\tilde{W}_{ht}} \right)^{\frac{\gamma-\eta}{\eta}} \left(\frac{L_{ckt}}{L_{hkt}} \right)^\gamma \left(\frac{E_{ckt}}{E_{hkt}} \right)^{-\frac{\gamma-\eta}{\eta}}
\end{aligned}$$

Rearranging to solve for the ratio of earnings yields

$$\frac{E_{ckt}}{E_{hkt}} = \left(\frac{A_{ct}}{A_{ht}} \right)^\eta \left(\frac{\tilde{W}_{ct}}{\tilde{W}_{ht}} \right)^{\frac{\gamma-\eta}{\gamma}} \left(\frac{L_{ckt}}{L_{hkt}} \right)^\eta$$

Now recall that the total effective hours supplied by individuals of education level e of age k in time t can be expressed as

$$\begin{aligned}
L_{ekt} &= \int_{i \in ekt} (1 + g_e)^{k-s_e} s_e^\theta a_i m_{e,t-k} (a_i)^\theta \\
&= (1 + g_e)^{k-e_E} s_e^\theta m_{e,t-k}^\theta \int_{i \in ekt} a_i^{\frac{\nu}{\nu-\theta}} \\
&= (1 + g_e)^{k-e_E} s_e^\theta m_{e,t-k}^\theta \mathbb{E} \left(a_i^{\frac{\nu}{\nu-\theta}} | i \in ekt \right) \\
&= (1 + g_e)^{k-e_E} s_e^\theta m_{e,t-k}^\theta \mathbb{E} \left(a_i^{\frac{\nu}{\nu-\theta}} \right)
\end{aligned}$$

where going from the first to second line we make use of Lemma 1 and use the fact that the distribution of abilities is stationary and independent of the distribution of costs ϕ_i in going from line three to four. Plugging back into the expression for the relative supply of total effective hours and using the FONC for the optimal level of specialization yields

$$\frac{L_{ckt}}{L_{hkt}} = \left(\frac{s_c}{s_h} \right)^\theta \frac{(1+g_c)^{k-s_c}}{(1+g_h)^{k-s_h}} \left(\frac{f_{h,t-k}}{f_{c,t-k}} \right)^{\frac{\theta}{\nu}} \left(\frac{W_{c,t-k}}{W_{h,t-k}} \right)^{\frac{\theta}{\nu}}$$

where we can further substitute into the expression using the inverse of the equilibrium educational attainment condition (derived from G)

$$\frac{W_{c,t-k}}{W_{h,t-k}} = \left(\frac{1}{d} \right)^{\frac{1}{b\rho}} \left(\frac{N_{c,t-k}}{N_{h,t-k}} \right)^{\frac{1}{b\rho}}$$

plugging the above expressions back into relative total earnings and dividing by the number of workers in each group yields relative *mean* earnings (denoted \tilde{E})

$$\begin{aligned} \frac{\tilde{E}_{ckt}}{\tilde{E}_{hkt}} &= \frac{E_{ckt}}{E_{hkt}} \left(\frac{N_{c,t-k}}{N_{h,t-k}} \right)^{-1} \\ &= \left(\frac{A_{ct}}{A_{ht}} \right)^\eta \left(\frac{\tilde{W}_{ct}}{\tilde{W}_{ht}} \right)^{\frac{\gamma-\eta}{\gamma}} \left(\frac{s_c}{s_h} \right)^{\theta\eta} \left(\frac{(1+g_c)^{k-s_c}}{(1+g_h)^{k-s_h}} \right)^\eta \left(\frac{f_{h,t-k}}{f_{c,t-k}} \right)^{\frac{\theta}{\nu}\eta} \left(\frac{1}{d} \right)^{\frac{\theta}{\nu b\rho}\eta} \left(\frac{N_{c,t-k}}{N_{h,t-k}} \right)^{\frac{\theta}{\nu b\rho}\eta-1} \end{aligned}$$

and so defining

$$\begin{aligned} D_{kt} \equiv \log \left(\frac{\tilde{E}_{ckt}}{\tilde{E}_{hkt}} \right) &- \frac{\gamma-\eta}{\gamma} \log \left(\frac{\tilde{W}_{ct}}{\tilde{W}_{ht}} \right) - \left(\frac{\theta\eta - \nu b\rho}{\nu b\rho} \right) \log \left(\frac{N_{c,t-k}}{N_{h,t-k}} \right) \\ &- \log \left[\left(\frac{s_c}{s_h} \right)^{\theta\eta} \left(\frac{(1+g_c)^{k-s_c}}{(1+g_h)^{k-s_h}} \right)^\eta \left(\frac{1}{d} \right)^{\frac{\theta}{\nu b\rho}\eta} \right] \end{aligned}$$

allows us to write

$$\frac{1}{\eta} D_{kt} = \log \left(\frac{A_{ct}}{A_{ht}} \right) + \frac{\theta}{\nu} \log \left(\frac{f_{h,t-k}}{f_{c,t-k}} \right)$$

for all age groups k in time t . Stacking all NT observations into one vector gives the matrix expression in the proposition and completes the proof. \square

B Appendix: Figures

C Appendix: Tables

	Full Sample			Balanced Panel		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Curriculum</i>	1.042*** (0.009)	1.559*** (0.000)	1.958*** (0.000)	1.492*** (0.019)	1.708*** (0.021)	1.961*** (0.016)
Time Effects	yes	yes	yes	yes	yes	yes
Weighted	no	yes	yes	no	yes	yes
Trunc. Sample	no	no	yes	no	no	yes
N	28,459	28,459	12,738	22,846	22,846	12,615
R^2	0.96	0.98	0.99	0.97	0.98	0.99

Table 3: Robustness Estimates of ν

In the model simulations we use specification (2) since this is our preferred specification and close to the midpoint of other specifications. The dependent variable is log published out-of-state tuition and required fees for full-time undergraduates, in order to avoid subsidies effects for local students. The independent variable measure complexity of curriculuar offerings as the log number of fields with more than one conferral in the specific year for each school . The sample includes all public and not-for-profit 4 year degree granting instutions accredited in the united states. Since we want schools that have "full offerings" we also consider subsamples which include only school that have at least the mean number of course offerings—that is the Mcutoff. The time period used is 2001-2014, where consistent measures of tuitions and curricular classications are available without making adjustments.

D Appendix: Data

E Appendix: Selection and Heterogeneity in Ability

In this section we consider alternative assumptions about the joint distribution of abilities a_i and costs ϕ_i and how they effect our results. Analyzing this feature is important as our empirical strategy exploits variation across cohorts and, as attainment levels rise, selection into college can change the relative ability of each group of students in a way that drives this variation in the data. Specifically, in analyzing the college wage premium, the important statistic is how the average ability of those in college compares with those in high school:

$$\frac{\mathbb{E}[a_i | i \in college]}{\mathbb{E}[a_i | i \in highschool]} \quad (15)$$

Specifically, it is important how the statistic above changes as attainment levels change. In the baseline model presented in the paper, we assumed costs and abilities were independently distributed. Combined with log utility, this assumption meant that selection into college was based on heterogeneity in costs alone and so ability played no role since:

$$\frac{\mathbb{E}[a_i | i \in college]}{\mathbb{E}[a_i | i \in highschool]} = \frac{\mathbb{E}[a_i | \phi_i > c]}{\mathbb{E}[a_i | \phi_i < c]} = \frac{\mathbb{E}[a_i]}{\mathbb{E}[a_i]} = 1$$

where c is a constant representing the model's equilibrium cutoff for college matriculation. Notice that as c falls (i.e. attainment increases), the ratio will remain unchanged. With more general utility functions, selection into college would depend on both costs ϕ_i and ability a_i . In particular, with CRRA preferences it would depend on the ratio of the two: individuals are more likely to go to college if they have low opportunity cost of schooling or high ability.

It should be clear then that in the more general context what is important is the joint distribution over costs and abilities $H(a, \phi)$. For example, if differences in costs were large compared to differences in ability, and the two were positively correlated, we could get perfect negative selection on ability. Similarly, if they were perfectly negatively correlated, we could get perfect positive selection on ability. The point is that, with no additional information, almost anything is possible if one constructs H appropriately.

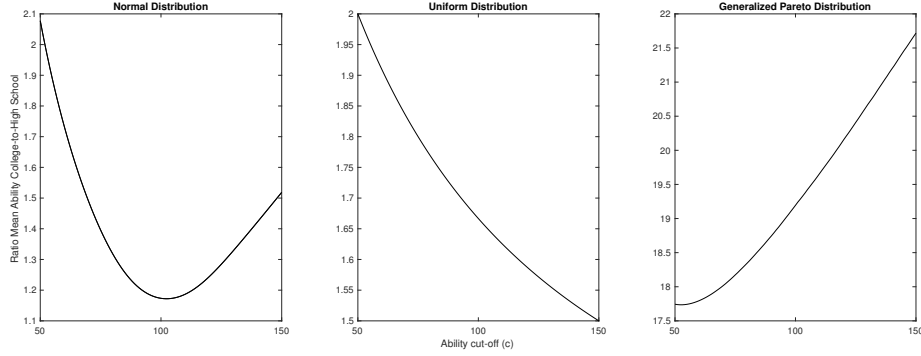


Figure 19: Relative Ability Under Perfect Selection

Of course, not every possibility is equally plausible a priori. Most people likely believe that opportunity cost and ability are negatively correlated so that, coupled with selective college admissions, it is the highest ability individuals which end up in college. The interesting thing for our purposes is that even in this context, where selection on ability is perfect, how the ratio in equation 15 changes as attainments increase is not a foregone conclusion. To see this, consider that with perfect selection we may rewrite equation 15 as:

$$\frac{\mathbb{E}[a_i | i \in college]}{\mathbb{E}[a_i | i \in highschool]} = \frac{\mathbb{E}[a_i | a_i > c]}{\mathbb{E}[a_i | a_i < c]} \quad (16)$$

for some cutoff ability level c . When college attainment levels increase the cutoff c drops. Hence, while it is true that when attainments rise and the best ability high school students go to college the average ability of remaining high school students (the denominator) falls, it is also true that the marginal college goer has gotten worse and so the average ability of college students (the numerator) also falls. The effect on the ratio is therefore ambiguous and depends on the shape of the distribution of abilities in the population.

To illustrate this point, figure 19 plots how the ratio in equation 16 evolves as the cutoff level changes for three different distributions: the normal distribution, the uniform distribution, and the generalized pareto distribution⁴⁸. Parameters are

⁴⁸The specific parameters are: $\mu = 100, \sigma = 10$ for the normal distribution; bounds of 50 to 150 for the uniform distribution; and for the Pareto distribution I use a location parameter of 0, a scale parameter of 25, and shape parameter equal to $3/4$.

chosen so that all three distributions have mean of 100 and the plot considers changes in the neighborhood of the mean. Note that the cutoff falls when attainments rise, so one can interpret rising attainments by reading the axis from right to left. While all three cases seem reasonable, a priori, the implications of each for how selection effects relative abilities are completely different and may even switch directions as attainments change.

It is even possible to construct distributions of ability such that under perfect selection the ratio stays constant for all attainment level. Consider, for example, the distribution with finite positive support $x \in [a_L, a_H]$ and density function:

$$f(x) \equiv \frac{\lambda}{2} x^{-\frac{3}{2}} \quad \text{where} \quad \lambda \equiv \frac{a_H^{\frac{1}{2}} a_L^{\frac{1}{2}}}{a_H^{\frac{1}{2}} - a_L^{\frac{1}{2}}} \quad (17)$$

A little bit of algebra confirms that this distribution has a constant ratio such that

$$\frac{\mathbb{E}[a_i | a_i > c]}{\mathbb{E}[a_i | a_i < c]} = \left(\frac{a_H}{a_L} \right)^{\frac{1}{2}} \quad \forall c \in [a_L, a_H]$$

The main point is that whether or not ability heterogeneity plays a role in driving the college wage premium depends completely on functional form assumptions which are, for the most part, unobservable and irrefutable. This is a very unappealing property from the perspective of applied work and motivated the independence assumption used in the baseline model. By the same token, its impossible to determine the extent that our empirical results are biased by heterogeneity in ability and selection. Given that attainment levels in the US are below 50% (i.e. the cutoff is to the right of the mean in the figure above), the normal and pareto distribution would suggest that our results are a lower bound as changes in ability have dampened our estimates of rising specialization shocks; the uniform assumption suggests precisely the opposite; and the custom distribution in equation 17 suggests our estimates are exactly correct and unbiased.

Given the prominence with which heterogeneity in ability appears in the education and labor literatures, the results above suggest that more clever empirical work

is needed to determine the best modeling choices for integrating heterogeneous ability into models with endogenous education and labor decisions.

References

- Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality*. *The Quarterly Journal of Economics* 113(4), 1055–1089.
- Acemoglu, D. and D. Autor (2011). Chapter 12 - skills, tasks and technologies: Implications for employment and earnings. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, pp. 1043 – 1171. Elsevier.
- Acemoglu, D. and D. Autor (2012). What Does Human Capital Do? A Review of Goldin and Katz's *The Race between Education and Technology*. NBER Working Papers 17820, National Bureau of Economic Research.
- Acemoglu, D. and J.-S. Pischke (1999). The structure of wages and investment in general training. *Journal of Political Economy* 107(3), 539–572.
- Alberty, H. and E. Alberty (1962). *Reorganizing the High-School Curriculum*, 3rd ed. Macmillan Company.
- Alon, T. (2017). Schools, skills, and curricula in the 20th century. *Working Paper*.
- Altonji, J. G., E. Blom, and C. Meghir (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics* 4(1), 185–223.
- Altonji, J. G., L. B. Kahn, and J. D. Speer (2016). Cashier or consultant? entry labor market conditions, field of study, and career success. *Journal of Labor Economics* 34(S1), S361–S401.
- Arrow, K. J. (1979). The division of labor in the economy, the polity, and society. In G. O'Driscoll (Ed.), *Adam Smith and Modern Political Economy*. Iowa State University.
- Arum, R. and Y. Shavit (1995). Secondary vocational education and the transition from school to work. *Sociology of Education* 68(3), 187–204.

- Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review* 103(5), 1553–97.
- Autor, D. H., L. F. Katz, and A. B. Krueger (1998). Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics* 113(4), 1169–1213.
- Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration*. *The Quarterly Journal of Economics* 118(4), 1279–1333.
- Becker, G. and K. M. Murphy (1992). The division of labor, coordination costs, and knowledge. *The Quarterly Journal of Economics* 107(4), 1137–1160.
- Benhabib, J. and M. M. Spiegel (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics* 34(2), 143 – 173.
- Berman, E., J. Bound, and Z. Griliches (1994). Changes in the demand for skilled labor within u.s. manufacturing: Evidence from the annual survey of manufacturers. *The Quarterly Journal of Economics* 109(2), 367–397.
- Campbell, P. B., M. N. Orth, and P. Seitz (1981). Patterns of participation in secondary vocational education: A report based on transcript and interview data of the 1979 and 1980 nlsy cohort. Technical report, Office of Vocational and Adult Education.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica* 69(5), 1127–1160.
- Card, D. and J. E. DiNardo (2002). Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles. *Journal of Labor Economics* 20(4), 733–783.
- Card, D. and T. Lemieux (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The Quarterly Journal of Economics* 116(2), 705–746.

- Carneiro, P. and S. Lee (2011). Trends in quality-adjusted skill premia in the united states, 1960-2000. *American Economic Review* 101(6), 2309–49.
- Caselli, F. (2005). Chapter 9 accounting for cross-country income differences. *Handbook of Economic Growth* 1, 679 – 741.
- Cockrill, A. and P. Scott (1997). Vocational education and training in germany: trends and issues. *Journal of Vocational Education and Training* 49(3), 337–350.
- Connolly, H. and P. Gottschalk (2006). Differences in wage growth by education level: do less-educated workers gain less from work experience? urn:nbn:de:101:1-20090825291.
- Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- Davis, S. J. and J. Haltiwanger (1991). Wage dispersion between and within u.s. manufacturing plants, 1963-1986. Working Paper 3722, National Bureau of Economic Research.
- de la Croix, D., M. Doepke, and J. Mokyr (2017). Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *The Quarterly Journal of Economics*. *Forthcoming*.
- DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica* 64(5), 1001–1044.
- Doms, M., T. Dunne, and K. Troske (1997). Workers, wages, and technology. *The Quarterly Journal of Economics* 112(1), 253–290.
- Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske (2004). Wage and productivity dispersion in united states manufacturing: The role of computer investment. *Journal of Labor Economics* 22(2), 397–429.
- Dustmann, C., B. Fitzenberger, and S. Machin (Eds.) (2008). *The Economics of Education and Training*. Studies in Empirical Economics. Physica.

- Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. *Federal Reserve Bank of San Francisco Working Paper*.
- Fersterer, J., J.-S. Pischke, and R. Winter-Ebmer (2007). Returns to Apprenticeship Training in Austria: Evidence from Failed Firms. IZA Discussion Papers 3104, Institute for the Study of Labor (IZA).
- Freeman, R. B. (1991). How Much Has De-Unionisation Contributed to the Rise in Male Earnings Inequality? NBER Working Papers 3826, National Bureau of Economic Research, Inc.
- Goldin, C. and L. Katz (2008). *The Race Between Education and Technology*. Belknap Press for Harvard University Press.
- Goldin, C. and R. A. Margo (1992). The great compression: The wage structure in the united states at mid- century. *The Quarterly Journal of Economics* 107(1), 1–34.
- Gordon, N. (2008). The changing federal role in education finance and governance. In H. F. Ladd and E. B. Fiske (Eds.), *Handbook of Research in Education Finance and Policy*, Chapter 17, pp. 295–313. New York: Routledge.
- Government Accountability Office (1974). *What is the role of Federal assistance for vocational education?* Washington D.C.: U.S. General Accounting Office.
- Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017). General Education, Vocational Education, and Labor-Market Outcomes over the Lifecycle. *Journal of Human Resources* 52(1), 48–87.
- Hanushek, E. A. and L. Woessmann (2010). Education and economic growth. In D. J. Brewer and P. J. McEwan (Eds.), *Economics of Education*. Elsevier.
- Hayward, G. C. and C. S. Benson (1993). Vocational-technical education: Major reforms and debates 1917-present. Technical Report ED/OVAE-93-7, Office of Vocational and Adult Education, Washington D.C.
- Heathcote, J., F. Perri, and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the united states, 1967 - 2006. *Review of*

- Economic Dynamics* 13(1), 15 – 51. Special issue: Cross-Sectional Facts for Macroeconomists.
- Heathcote, J., K. Storesletten, and G. L. Violante (2010). The macroeconomic implications of rising wage inequality in the united states. *Journal of Political Economy* 118(4), 681–722.
- Heckman, J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics* 1(1), 1–58.
- Hippach-Schneider, U., M. Krause, and C. Woll (2007). Vocational education and training in germany. Technical report, European Centre for the Development of Vocational Training (Cedefop).
- Houthakker, H. S. (1956). Economics and biology: Specialization and speciation*. *Kyklos* 9(2), 181–189.
- Hoxby, C. (Ed.) (2003). *The Economics of School Choice*. The University of Chicago Press.
- Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies* 76(1), 283–317.
- Jones, B. F. (2014). The human capital stock: A generalized approach. *American Economic Review* 104(11), 3752–77.
- Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics* 107(1), 35–78.
- Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection*. *The Quarterly Journal of Economics* 131(3), 1057–1111.
- Klenow, P. and A. Rodriguez-Clare (1997). The neoclassical revival in growth economics: Has it gone too far? In *NBER Macroeconomics Annual 1997, Volume 12*, pp. 73–114. National Bureau of Economic Research, Inc.

- Kosters, M. H. (Ed.) (1999). *Financing College tuition: Government Policies and Educational Priorities*. The AEI Press.
- Kraus, J. W. (1961). The development of a curriculum in the early american colleges. *History of Education Quarterly* 1(2), 64–76.
- Krueger, D. and K. B. Kumar (2004). Skill-specific rather than general education: A reason for us-europe growth differences? *Journal of Economic Growth* 9(2), 167–207.
- Krueger, D., F. Perri, L. Pistaferri, and G. L. Violante (2010). Cross-sectional facts for macroeconomists. *Review of Economic Dynamics* 13(1), 1 – 14. Special issue: Cross-Sectional Facts for Macroeconomists.
- Krusell, P., L. E. Ohanian, J.-V. Rios-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica* 68(5), 1029–1053.
- Lazear, E. (2009). Firm specific human capital: A skill weights approach. *Journal of Political Economy* 117(5), 914 – 940.
- Lee, D. S. (1999). Wage inequality in the united states during the 1980s: Rising dispersion or falling minimum wage? *The Quarterly Journal of Economics* 114(3), 977–1023.
- Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review* 96(3), 461–498.
- Lerman, R. I. (2012). Can the united states expand apprenticeship? lessons from experience. IZA Policy Paper 46, Institute for the Study of Labor (IZA), Bonn.
- Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics* 22(1), 3–42.
- Malamud, O. and C. Pop-Eleches (2010). General education versus vocational training: Evidence from an economy in transition. *The Review of Economics and Statistics* 92(1), 43–60.

- Mankiw, G. N., D. Weil, and D. Romer (1992). A contribution to economic growth. *The Quarterly Journal of Economics* 107, 407–37.
- Marshall, A. (1890). *Principles of Economics*. London: Macmillan and Co., Ltd.
- Murphy, K. M. and F. Welch (1992). The structure of wages. *The Quarterly Journal of Economics* 107(1), 285–326.
- National Commission on Excellence in Education (1983). *A Nation at Risk: The imperative for educational reform*. Washington D.C.: Government Printing Office.
- National Education Association (1893). *Report of the Committee on Secondary School Studies*. Washington D.C.: Government Printing Office.
- North, D. (1981). *Structure and Change in Economic History*. Norton.
- Office of Education (1900). *Report to the Commissioner of Education*. Washington D.C.: Government Printing Office.
- Office of Education (1910). *Report to the Commissioner of Education*. Washington D.C.: Government Printing Office.
- Office of Education (1915). *Report to the Commissioner of Education*. Washington D.C.: Government Printing Office.
- Office of Education (1928). *Report to the Commissioner of Education*. Washington D.C.: Government Printing Office.
- Office of Education (1948). *Report to the Commissioner of Education*. Washington D.C.: Government Printing Office.
- PEW (2015). Federal and state funding of higher education. Technical report, The Pew Charitable Trusts.
- Piketty, T. and E. Saez (2003). Income inequality in the united states, 1913–1998. *The Quarterly journal of economics* 118(1), 1–41.
- Rosen, S. (1978). Substitution and division of labour. *Economica* 45(179), 235–250.

- Rosen, S. (1983). Specialization and human capital. *Journal of Labor Economics* 1(1), 43–49.
- Rosenberg, N. (1976). Another advantage of the division of labor. *Journal of Political Economy* 84(4), 861–868.
- Rudolph, F. (1962). *The American College and University*. Athens: University of Georgia Press.
- Sanders, C. and C. Taber (2012). Life-cycle wage growth and heterogeneous human capital. *Annual Review of Economics* 4(1), 399–425.
- Silverberg, M., E. Warner, M. Fong, and D. Goodwin (2004). National assessment of vocational education: Final report to congress. Technical report, Department of Education.
- Smith, A. (1776). *The Wealth of Nations*. London: W. Strahan and T. Cadell.
- Snyder, T. D. (1993). Higher education. In T. D. Snyder (Ed.), *120 Years of American Education: A Statistical Portrait*. National Center for Education Statistics.
- Sonnenberg, W. (1993). Elementary and secondary education. In T. D. Snyder (Ed.), *120 Years of American Education: A Statistical Portrait*. National Center for Education Statistics.
- Stasz, C. and S. Bodilly (2004). Efforts to improve the quality of vocational education in secondary schools: Impact of federal and state policies. Technical Report ED-00-C0-0002, U.S. Department of Education, Washington D.C.
- Stigler, G. J. (1951). The division of labor is limited by the extent of the market. *Journal of Political Economy* 59(3), 185–193.
- Yang, X. and J. Borland (1991). A microeconomic mechanism for economic growth. *Journal of Political Economy* 99(3), 460–482.
- Young, A. (1928). Increasing returns and economic progress. *History of Economic Thought Articles* 38, 527–542.