THREE ESSAYS ON HIGHER EDUCATION OUTCOMES

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by

HEATHER HOFFMAN

Dr. Bradley Curs, Dissertation Supervisor

DECEMBER 2022

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

THREE ESSAYS ON HIGHER EDUCATION OUTCOMES

presented by Heather Hoffman,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Bradley Curs

Professor Peter Mueser

Professor Cory Koedel

Professor Matthew Foulkes

DEDICATION

In loving memory of Grandpa and Grandma Thomas.

Also, to Ruth Imogene.

If only your paths would have crossed.

ACKNOWLEDGEMENTS

Many people deserve thanks for making this dissertation and completing my doctoral studies possible. Above all, I owe a great deal of gratitude to my advisor, Dr. Brad Curs, who has spent countless hours meeting with me and providing me with feedback. The culmination of this work marks the end of nearly ten years of his advising me, first as his master's advisee and now as his doctoral advisee. Brad, thank you for seeing me through.

I would also like to thank my committee members, Drs. Peter Mueser, Cory Koedel, and Matt Foulkes, for their time and effort. I have so many thanks to give to Peter, who supported my independent study of labor economics and always answered my questions with great enthusiasm. I am also thankful for his great attention to detail and thorough feedback. Thanks to Cory for helping me finally "get" bias and his direct feedback, such as "sometimes less is more." Finally, thanks to Matt Foulkes for staying on the committee even though the dissertation shifted its focus away from rural populations back to my comfort zone of higher education policy and administration. I am also thankful to Matt for helping see my husband through his geography master's degree earlier this year.

I am also grateful for the opportunity to have worked with Drs. Oded Gurantz and Rajeev Darolia, who have both since left Mizzou. Thanks to Oded for taking me on as his graduate assistant for our project on honors students. As a full-time staff member and part-time student, there are few opportunities to collaborate with faculty, and I learned a great deal from working with Oded. I am thankful to Rajeev, who saw potential in me during my master's studies and encouraged me to apply for the doctorate. His mentorship in the early years of my program was essential for setting me up for success.

Several colleagues and friends at Mizzou also made this dissertation more manageable, especially during a pandemic and with a new baby. I am deeply appreciative of Dr. Jeni Hart, who is a fantastic supervisor, friend, and all-around person. Jeni, thank you for checking in with me about my dissertation and for never making me feel guilty for taking time off to work on it. I am so grateful for Jami Thompson, who covered my responsibilities when I was on parental leave and expected nothing in return. My Status of Women Committee crew (Dr. Nicole Campione-Barr, Dr. Mike Urban, and Michelle Heck) have been my cheerleaders through my master's degree and now with the doctorate. I am also grateful for my Truman School pals, Kristi Ressel, Dr. Trang Pham, and Dr. Chris Clause. Kristi, thanks so much for holding me accountable this summer and seeing me through the finish line—I hope to return the favor! And, special thanks to Ava Drayton for being the bestest work bestie.

Of course, nothing would be possible without the support of my family. I cannot thank my husband, John, enough for supporting me in my pursuit of graduate work and other endeavors. I am so grateful for Kathy Thomas for treating me like your own (not to mention, keeping me well-fed). Many thanks to my parents for always showing up and cheering me on. My sister, Heidi, has been my rock. My brother, Ryan, has been a constant reminder of all the fun I have been missing out on (a great motivator!). And I cannot leave out Kitty, who remained my furry and loyal lap companion throughout it all, up until his passing this last summer. Kitty, we did it!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THREE ESSAYS ON THE HIGHER EDUCATION OUTCOMES

Heather Hoffman

Dr. Bradley Curs, Dissertation Supervisor

ABSTRACT

This dissertation consists of three chapters examining issues relevant to higher education outcomes. The first essay examines the impact of state merit-aid program adoption on the stock of human capital across rural and urban populations. Using data from the 1990 and 2000 Decennial Censuses and the 2008-2012 American Community Survey, I utilize a staggered difference-in-difference methodology that exploits the exogenous variation in the timing of program adoption to produce causal estimates. Program adoption reduces bachelor's degree holders in rural counties by 1.2 percentage points. The second essay measures the impact of honors college participation on collegiate outcomes at a large, public-land grant midwestern university by leveraging strict eligibility criteria using a fuzzy regression discontinuity design. For compliers near the threshold, honors participation increases first- and senior-year cumulative grade-point averages but has no statistically significant effect on persistence or graduation rates. The third essay is a descriptive analysis that measures the extent to which education-job mismatch and its consequences vary by degree field for Ph.D. recipients from a large, public-land grant midwestern university from 2011 to 2020. The study relies on a novel dataset compiled from publicly accessible sources online. Relative to other degree fields, engineering Ph.D. graduates were the most likely to work in occupations for which they were over-educated or outside of their field. There is also some evidence that mismatch may be positively correlated with occupation-level earnings.

CHAPTER 1: INTRODUCTION

One of the ways in which higher education contributes to economic development is through the formation of human capital (Becker, 1964). By recruiting, training, and producing graduates who ultimately join the workforce, colleges and universities increase the stock of human capital that helps drive the economy. Moreover, college-educated workers contribute "spillover" effects that, in turn, enhance other workers' productivity (Moretti, 2004b, 2004a). While policymakers are vested in increasing college access for the masses, they also look to higher education to supply a labor pool of exceptionally talented or meritorious graduates that can help drive innovation and increase "spillover" effects on other less talented workers. This dissertation consists of three chapters examining issues relevant to higher education outcomes. Although the three essays vary in the topic of interest and approaches undertaken, each essay addresses a particular outcome of higher education.

In the first paper, I examine the impact of state merit-aid program adoption on the stock of human capital across rural and urban populations. State merit-based aid programs subsidize college for high school students that meet eligibility requirements, such as GPA or standardized test score minimums. Such programs are designed to increase residents' access to college and incentivize high-achieving students to stay in their home state (Bruce & Carruthers, 2014; Dynarski, 2004; Groen, 2011). However, past research has only focused on the effects of such programs at the state level (e.g., Fitzpatrick & Jones, 2012, 2016; Sjoquist & Winters, 2014) without considering the possible within-state effects, such as across rural and urban populations. Using data from the 1990 and 2000 Decennial Censuses and the 2008-2012 ACS, I utilize a staggered difference-in-difference methodology that exploits the exogenous variation in the timing of program adoption to produce causal estimates. I find that program adoption significantly reduces the share of bachelor's degree holders in rural counties, which increases the rural-urban college attainment gap by a small margin.

Honors programs and colleges seek to enhance the intellectual growth of academically talented students through targeted curricular and co-curricular activities (Astin, 1993). The second paper examines the impact of honors college participation on collegiate outcomes, such as cumulative grade-point average (GPA), first-to-second-year retention, and fourth-, fifth-, sixth-year graduation rates. While a few studies find that honors students have higher GPA, retention, graduate rates relative to non-honors students, this research has been largely limited to descriptive analyses (e.g., Bowman & Culver, 2018; Cosgrove, 2004; Furtwengler, 2015; Hartleroad, 2005; Rinn, 2007; Shushok, 2002); there is little evidence of the causal impacts of honors participation on collegiate outcomes. Using the Honors College at the University of Missouri (MU) as a case study, the analysis leverages honors eligibility criteria based on a strict set of cutoffs to produce casual estimates using regression discontinuity design. Honors participation has large positive effects on first year and senior year cumulative GPA, but no statistically significant effect on persistence or graduation rates.

The final paper examines the extent to which vertical mismatch and horizontal mismatch vary by degree field for Ph.D. recipients from the University of Missouri from 2011 to 2020. The study uses a novel dataset from Academic Analytics with employment information compiled using only publicly accessible sources online (e.g., LinkedIn, employer websites). In contrast to past research that uses survey data, this paper constructs more or less "objective" measures of mismatch using Standard Occupational Classification (SOC) codes provided by the US Bureau of Labor Statistics. While the study is purely descriptive, I find that Engineering graduates were most likely to work in occupations for which they were over-educated or in a job outside their degree field. An analysis of occupation-level earnings suggests that mismatch may have a positive effect on potential earnings.

References

Astin, A. W. (1993). *What matters in college? Four critical years revisited.* San-Francisco: Jossey-Bass.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press: Chicago and London.

Bruce, D. J., & Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, *81*, 30–44. https://doi.org/10.1016/j.jue.2014.01.006

Dynarski, S. M. (2004). The New Merit Aid. In *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It: Vol. I* (Issue September). http://www.nber.org/chapters/c10098.pdf

Fitzpatrick, M. D., & Jones, D. (2012). High education, merit-based scholarships and post-baccalaureate migration. In *National Bureau of Economic Research Working Paper Series*.

Fitzpatrick, M. D., & Jones, D. (2016). Post-baccalaureate migration and merit-based scholarships. *Economics of Education Review*, *54*, 155–172. https://doi.org/http://dx.doi.org/10.1016/j.econedurev.2016.07.002

Groen, J. A. (2011). Do Merit-Aid Programs Help States Build Skilled Workforces? *Change: The Magazine of Higher Learning*. https://doi.org/10.1080/00091383.2011.618080

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*(1–2), 175–212. https://doi.org/10.1016/j.jeconom.2003.10.015

Moretti, E. (2004b). Human capital externalities in cities. In *Handbook of regional and urban economics* (Vol. 4, pp. 2243–2291). Elsevier. https://learning.oreilly.com/library/view/handbook-of-regional/9780444509673/xhtml/S1574008004800087.htm

Sjoquist, D. L., & Winters, J. v. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, *80*, 39–40. https://doi.org/10.1016/j.jue.2013.10.003

CHAPTER TWO: THE HETEROGENEOUS EFFECTS OF STATE MERIT-AID PROGRAM ADOPTION: DO RURAL POPULATIONS BENEFIT LESS?

The shift from need-based aid to merit-based aid programs is one of the most dramatic changes to financial aid policy in the United States (US) over the last few decades (Groen, 2011). Since the early 1990s, nearly half of states have adopted a merit-aid program. In contrast to programs based on financial need, merit-based aid programs subsidize college for high school students who meet eligibility requirements, such as GPA or standardized test score minimums. Such programs are designed to increase residents' access to college and incentivize high-achieving students to stay in their home state (Bruce & Carruthers, 2014; Dynarski, 2004; Groen, 2011).

The extant literature on state merit-aid programs finds that, for the most part, they increase college enrollment (Bruce & Carruthers, 2014; Cornwell et al., 2006; Dynarski, 2000, 2004) and post-college retention within the student's state of residence (Fitzpatrick & Jones, 2016; Harrington et al., 2016; Sjoquist & Winters, 2014). There has been limited evidence to suggest that there may be heterogeneous effects by income and race (Dynarski, 2000). However, their potential within-state effects, such as across rural and urban populations, have not received any attention.

As a policy designed to increase the stock of college-educated workers in the state, merit-aid programs are interesting to analyze in the context of the widening college attainment gap between urban and rural areas. According to the U.S. Department of Agriculture, the share of urban residents ages 25 and over with bachelor's degrees grew from 26 to 33 from 2000 to 2015, while the share of rural degree holders grew from 15 to 19 percent (Marre, 2017).

It may be that adopting a merit-aid program increases net college attainment in the state, but most gains are found in urban areas. There is reason to believe that the effects of state merit-aid programs vary across rural and urban populations in ways that contribute to these growing disparities. For instance, rural high school graduates may be less likely to meet eligibility requirements, such as minimum standardized test scores. Rural school students comprise a historically disadvantaged population with regard to college access: relative to their urban counterparts, rural students are more likely to be of low-socioeconomic status (Adelman, 2002; Gibbs, 1998; Irvin et al., 2012; Lichter & Brown, 2014); first-generation (Irvin et al., 2012; Provasnik et al., 2007); and less academically prepared (Provasnik et al., 2007). As rural students are more likely than their urban counterparts to be disadvantaged, state merit-aid programs may do little to improve their college access.

Due to out-migration, state merit-aid programs may also impact rural and urban populations differently. Rural students who meet the eligibility requirements of their state programs will likely need to leave their hometowns to take advantage of merit aid at the state's qualifying colleges or universities. In this respect, state merit-aid programs may incentivize rural youth to leave their hometowns for college and eventually move to urban centers of production (Gibbs, 1990). While the extent of out-migration is difficult to track (Foulkes & Newbold, 2008), the "brain drain" of rural youth may partly be responsible for slowed population growth in rural areas (Cromartie, 2017). One estimate attributes 25 percent of the rural-urban college gap to rural student out-migration (Mykerezi & Jordan, 2014). From a state perspective, net college attainment may not be affected by the intrastate migration of college students; however, it may have implications for rural areas already struggling with slowed population growth or loss.

In the present study, I examine the effects of state merit-aid program adoption on the distribution of college-educated adults across rural and urban populations. In contrast to past research that focuses on the impact of merit aid on the individual student's decision to attend college or stay located in the state after graduation, I focus on the effect on the stock of human capital, which is operationalized as the share of college-educated adults at the county level. Specifically, I ask the research question: to what extent does merit-aid program adoption increase the stock of human capital of rural populations relative to urban populations?

Given the mechanisms described above, I hypothesize that the stock of human capital will grow more slowly in rural areas and ultimately serve to widen the urban-rural college gap. The analysis contributes to an understanding of how policies adopted by states to increase college-educated workers at the state level may have unintentional effects at the sub-state level.

The outline of the paper is as follows. The first section summarizes some of the research on the impact of merit-aid programs on college attendance and worker retention. I also include a brief discussion of the possible mechanisms that may be contribute to the differential impact of state merit-aid programs across rural and urban populations. The second section outlines my empirical strategy, which explains the choice of the staggered difference-in-difference framework, and describes the data used in detail. The final section presents the results of the analysis, including an event study analysis that accounts for heterogenous timing effects that may bias estimates found in the difference-in-differences analysis. As anticipated, program adoption negatively impacts bachelor's degree attainment and population growth in rural counties relative to urban counties. However, the magnitude of the effects is too small to be of practical significance.

Relevant Literature

The study relies on the quasi-experimental research on state merit-aid programs and literature on human capital theory. Most studies on state merit-aid programs focus on the effects on college enrollment or post-college retention. While studies on college enrollment tend to focus on individual states, more recent studies on post-college retention use pooled data to analyze the impact of merit-aid programs generally. As far as I know, there has not yet been an analysis of the effects of state merit-aid programs for sub-state geographies, such as the county level.

Effects of State Merit-Aid Programs on College Enrollment

Studies focused on individual broad-based state merit-aid programs have found generally positive effects on college enrollment within the state. The Georgia HOPE program, one of the oldest and most well-known programs, awards full tuition to high school students with a 3.0 GPA to attend one of the state's eligible institutions. Dynarski (2000), using a difference-in-differences estimation strategy with data from the 1989 to 1997 Current Population Survey (CPS) found that HOPE increased overall college attendance rates among all 18- to 19-year-olds by 7.0 to 7.9 percentage points in Georgia. After accounting for institution type and distinguishing scholarship from grant effects, Cornwell, Mustard, and Sridhar (2006) found a smaller (6%) but still positive effect of HOPE on college enrollment in Georgia, with the most significant gains at public (9%) and private (13%) four-year institutions.

Florida's Bright Futures program is another large-scale tuition-subsidy program that provides free tuition at four-year institutions for students meeting eligibility criteria using a tiered structure (e.g., students with a 3.0 GPA. and 20 ACT receive 75% of tuition and students with a 3.5 GPA. and 28 ACT receive 100% tuition). Zhang, Hu, and Sensenig (2013) analyze the effect of Bright Futures on college enrollment with data from the Integrated Postsecondary Education Data System (IPEDS) using a difference-in-differences framework. They find large and significant enrollment effects at Florida's public four-year and two-year institutions.

While studies generally find positive effects on enrollment, some findings suggest there may be heterogeneous effects for disadvantaged student populations (Cornwell et al., 2006; Dynarski, 2000; Ehrenberg et al., 2005; Heller & Marin, 2004). For example, Dynarski (2000) estimates the differential effects of Georgia HOPE by income and race and finds the program has widened the college enrollment gap between high- and low-income groups and between whites and Blacks. These findings underscore and lend to the criticism that merit-aid programs may exacerbate attendance gaps for disadvantaged students, such as between black and white students or between low- and high-income families (Dynarski, 2000). However, there has not yet been an analysis of heterogeneous effects on rural and urban populations.

Effects of State Merit-Aid Programs on Post-College Location

One of the more appealing features of state merit-aid programs is their propensity to prevent high school students from migrating to other states for college (Zhang, etc., 2013). Research shows that students who attend college in-state are 10–20% more likely to reside in-state than students who graduate out-of-state (Perry 2001; Groen 2004). The idea is that merit-aid programs may incentivize high-ability students to stay within the state to attend college and eventually enter the workforce.

The impact of state merit-aid programs on the post-college location has been an area of increasing interest. Using state administrative data, Harrington, Muñoz, Curs, and Ehlert (2016) analyzed the effect of Missouri's Bright Flight Program, which provides merit scholarships to the state's top ACT test-taker, on state labor force participation eight years later. Using a regression discontinuity approach with ACT as the running variable, they find a negative relationship between ACT scores and the probability of in-state employment overall, but the likelihood of in-state employment increased sharply at the threshold ACT score of 30.

Few studies, however, have access to administrative data. Given the difficulty of tracking migration flows, much of this literature measures the probability that college graduates reside in the state where they were born without accounting for where they attended college. Focusing on Florida's Bright Figures program, Hickman (2009) combines 2000 Census and 2001-2006 American Community Survey (ACS) data and uses a difference-and-differences approach to analyze the effect of the program on the probability that those exposed to the program in high school resided in the state when surveyed between the ages of 23 and 27. He estimated that Bright Futures increased the probability of locating in Florida for college-educated individuals by about 3 to 4 percentage points.

Finally, a few studies use pooled data to evaluate across multiple states to leverage the exogenous variation introduced by differences in the timing of program adoption. For instance, Sjoquist and Winters (2014) expanded the approach Hickman (2009) used to include 25 merit-aid programs adopted between 1991 and 2004 using microdata from the 2000 decennial Census long-form questionnaire and the 2001-2010 American Community Survey. As merit-aid programs vary in the stringency and amounts awarded, they identify nine programs as "strong" programs given their relatively greater participation rates and monetary awards in contrast to "weak" programs. Conditional on having a bachelor's degree, they find that merit-aid programs increase the probability that college-educated 24- to 30-year-olds reside in their state of birth after college by nearly 3.0 percentage points; however, the effect for weak merit-aid states was much smaller on average. Fitzpatrick and Jones (2016), using microdata from the 1990 and 2000

Census and 2001-2010 ACS, estimate the effect of merit-aid program adoption on the probability of both living in the same state and obtaining a bachelor's degree. They found that residents born in a merit-aid state exposed to the program at 18 were 0.9 percentage points more likely to live in the state between the ages of 24 and 32 than those born in ineligible cohorts. However, eligible cohorts were slightly less likely to have received a bachelor's degree, which they attribute to an increase in the likelihood of only finishing some college.

In summary, the extant literature on state merit-aid programs indicates that, for the most part, such programs increase college enrollment (Bruce & Carruthers, 2014b; Cornwell et al., 2006; Dynarski, 2000, 2004) and post-college retention in the state (Fitzpatrick & Jones, 2016; Harrington et al., 2016; Hickman, 2009; Sjoquist & Winters, 2014). These studies, however, do not measure the impact on the stock of human capital, nor do they address the possibility of within-state effects, such as at the county level.

Theoretical Assumptions

This paper is guided by several assumptions of human capital theory. According to human capital theory, individuals are utility maximizers and implicitly weigh the costs and benefits of an investment in human capital (Becker, 1993). Individuals balance the costs and benefits of a given investment in education and will continue to invest in human capital as long as the marginal benefit exceeds the marginal cost. In this respect, human capital theory predicts that financial aid, such as merit-aid awards, may increase the likelihood of one attending college by lowering the marginal cost of attendance, holding all else constant.

With regard to state merit-aid programs, I anticipate two different mechanisms at work that may result in differences in net college attainment across rural and urban areas. On the one hand, state merit-aid programs may not affect rural high school students whatsoever while increasing college attainment among urban students. Recall that state merit-aid programs have been shown to have little effect on relatively less advantaged students who are less likely to meet the eligibility requirements (Cornwell et al., 2006; Dynarski, 2000; Ehrenberg et al., 2005; Heller & Marin, 2004). By extension, such programs may do little to increase college access among rural students who are historically more likely than their urban counterparts to be of low-socioeconomic status (Adelman, 2002; R. Gibbs, 1998; Irvin et al., 2012; Lichter & Brown, 2014); first-generation (Irvin et al., 2012; Provasnik et al., 2007); and less academically prepared (Provasnik et al., 2007).

On the other hand, state merit-aid programs may have a positive effect on some rural individual college attainment, but it may not be observable among rural populations because these students ultimately migrate. As with the decision to invest in college, geographic migration decisions are also considered human capital investments in which individuals weigh the pecuniary and psychic costs of moving against the expected benefits (Borjas, 2015). For rural students, the decision to attend college is also a dual college-migration decision given the lack of geographically proximate institutions. For example, Gibbs (1998) estimates that only about 20% of all colleges are located in rural areas, and these are disproportionately two-year institutional types. If a rural student migrates to take advantage of merit aid at one of the state's qualifying institutions, then it would have no positive effect on the net college attainment of rural areas.

Empirical Strategy

In the current paper, I measure the extent to which state merit-aid program adoption impacts the stock of human capital across rural and urban populations. I use a treatment-comparison research design in which states that never adopted a merit program or adopted a program outside the window of analysis serve as the comparison group. As states adopted programs at various times, I use a staggered introduction difference-in-differences estimation strategy that exploits variation introduced by differential timing in program adoption. The difference-in-differences approach is a quasi-experimental research design that allows me to estimate the impact of an intervention by comparing within-group differences between treatment and control groups over time (Conley & Taber, 2011). The within-group change in the mean outcome before and after the timing of the intervention comprises the first difference, while the between-group mean difference in the outcome comprises the second difference (Angrist & Pischke, 2008).

Empirical Model

The central question of the analysis is to what extent does merit-aid program adoption increase the stock of human capital of rural populations relative to urban populations? I begin with the following baseline specification:

$$Y_{jt} = \alpha + \delta Merit_{jt} + \sigma_j + \tau_t + \varepsilon_{jt},$$

where Y_{jt} Is the proportion of residents aged 25 and above in county (*j*) and surveyed in the year (*t*) who have attained a particular college outcome. $Merit_{jt}$ is an indicator variable for whether a county *j* in survey year *t* includes residents exposed to the treatment at 18 years of age. Note that the analysis encompasses all 50 states, including 25 that adopted a merit-aid program from 1991 to 2004 and 25 that never adopted a program and served as pure controls. Table 1 presents the list of state merit-aid programs and the years in which each program was introduced. In addition, several states (Illinois, Maryland, Michigan, and Washington) repealed their programs in the mid-2000s (Sjoquist & Winters, 2015). I include these states in the analysis, but the treatment is “switched off” for later cohorts who would not have become 18 until after the states repealed the programs.

The baseline model also incorporates county (σ_j) fixed effects to control for time-invariant characteristics specific to each county and survey year (τ_t) fixed effects to control for time-varying characteristics. I do not include any demographic controls to avoid controlling for factors that the treatment variable may influence. For instance, state merit-aid programs may influence the age composition within rural areas if rural youth out-migrate to attend college, so controlling for age would be to control for the effect I am trying to measure.

I cluster standard errors at the state level as difference-in-differences approaches are vulnerable to issues of serial correlation when the same units are observed over long periods (Bertrand et al., 2004). Failure to cluster standard errors may produce standard errors that are too small, which contributes to Type 1 errors in which a true null hypothesis is rejected.

There is also a possibility that cohorts exposed to merit-aid programs occur systematically later than cohorts that have not been exposed. While cohort-specific fixed effects account for some systematic differences, I limit the analysis to 10 cohorts leading up until the last cohort is ineligible for merit aid and up to 10 cohorts lagging after the treatment to produce a 20-year event window, which reduces the extent to which estimates are biased to cohorts at the extremes.

Finally, the coefficient of interest is the difference-in-differences estimate (δ), which is interpreted as the average change in the share of residents in treated counties with a particular outcome relative to the change in untreated counties. For instance, for counties in the state of Missouri, which adopted a merit-aid program in 1997, I observe the change in college attainment for those aged 25 and above in survey years 2000 and 2010 relative to college attainment in the survey year 1990. I observe this change relative to counties in states that adopted a merit-aid program later, such as South Dakota, for which the survey year 2010 is the only treated survey year, and relative to counties in states that never adopted a merit-aid program.

To measure the extent to which state merit-aid program adoption impacts human capital levels across rural and urban populations, I interact rural status with the indicator variable for treatment status:

$$Y_{jt} = \alpha + \delta Merit_{jt} + \beta\left(Rural_{jt} * Merit_{jt}\right) + \sigma_j + \tau_t + \varepsilon_{jt,}$$

where δ captures the main effects of merit adoption in absence of being rural and β is the effect of adoption for rural counties specifically.

More recently, there has been increasing awareness around the potential for two-way fixed effects bias estimates in the presence of heterogeneous treatment effects over time (Goodman-Bacon, 2018). In particular, two-way fixed-effects models may more heavily weight observations in the middle than at the end. Second, using past treated units as controls for future treated units may flip the sign of the coefficient. Given the possibility of heterogeneity in timing, I present an event study analysis later in the paper.

Identifying Assumptions

The difference-in-differences approach is understood to produce causal estimates of the treatment effect if one can demonstrate that the control group serves as an appropriate counterfactual to the treatment group. The key identifying assumption is the parallel trends assumption, which holds that—in the absence of the treatment—differences between the treatment and control group will be constant over time (Angrist & Pischke, 2008). In other words, the expected between-group difference would be equal to zero had the treatment never been implemented. The parallel trends assumption is not directly testable, so common practice is to observe pretreatment trends to ensure that between-group changes were constant before the treatment. I present these tests in tandem with the heterogeneous analysis presented later in the paper.

Data

I extracted data from the National Historical Geographic Information System (NHGIS) database maintained by IPUMS (Manson et al., 2021). The NGHIS provides easy access to summary tables and time series of population data for all Decennial Censuses and the American Community Survey (ACS) for all US Census Bureau geography levels. I use data aggregated at the county level. Not only are counties the standard building block for researching regional population and economic trends, but they are relatively stable over time (US Department of Agriculture – Economic Research Service (USDA-ERS), 2021).

I used the 1990 and 2000 Decennial Censuses and the 2008-2012 ACS to cover periods before and after program adoption in all states. The Census is conducted every ten years by the US Census Bureau and captures basic population and housing data for all households, while the ACS replaced the Census long-form as of 2001 and uses monthly rolling samples throughout the decade. The ACS data include person-level weights to make the data nationally representative. I use the ACS multi-year datasets, which provide sufficient coverage and smaller margins of error for smaller populations, such as rural areas (Chand & Alexander, 2000).

Stock of Human Capital

I am interested in measuring the effect of state merit-aid program adoption on the stock of human capital, which is operationalized as the proportion of college-educated adults aged 25 and above at the county level. I relied primarily on the NGHIS time series table "Persons 25 Years and Over by Educational Attainment," which includes survey year, county and state FIPS codes, and total persons at each level of educational attainment across all survey years used in the sample. For the dependent variable, I calculated the total of persons at each level of educational attainment out of the total persons within each county for each survey year. I calculated the percentage of county residents with at least some college, an associate degree or above, a bachelor's degree or above, or a graduate degree or above. The primary outcome of interest is the percentage of county residents with at least a bachelor's degree.

Treatment Status

To preserve the confidentiality of respondents, the NGHIS time series tables, unfortunately, do not disaggregate educational attainment for each age at the county level. I determine treatment status based on whether a county within a particular survey year included *any* residents who would have been exposed to the treatment at 18. As the minimum age included in the time series table is 25 years old, I use the following calculation to determine treatment status:

$$(survey\ year - adoption\ year + 18) >= 25,$$

where the treatment status for each county-by-year cell is equal to "1" if the left-hand side of the equation is greater or equal to 25 and "0" otherwise. I refer to these county-by-year cells as "treated" and "untreated" cohorts.

Rural Status

One of the primary challenges in the current analysis was defining rural populations. The predominant definitions of rurality, however, are constructed by the Census Bureau and the Office of Management and Budget (OMB)—both of which define rurality in relation to what it is not (Cromartie & Gibbs, 2008). The U.S. Census Bureau's urban-rural definition fundamentally delineates geographical areas based on population density. The definition of "urban" encompasses "urbanized areas" of 50,000 or more and "urban clusters" between 2,500 and

50,000 people within proximity of an urbanized area. Typically, what is "rural" is defined relationally in terms of all population, housing, and territories outside of urban areas.

In contrast, OMB designates counties as "metro" or "nonmetro" based on economic and social integration. The OMB first defines "metropolitan statistical areas" (MSA) as core counties with one or more urbanized areas and outlying counties economically tied to the core counties as measured by work commuting. Typically, counties with an MSA are classified as "metro," while those without an MSA are "nonmetro." The economic basis of the OMB definition also makes it preferable to the more population density-based definition used by the Census; the Census definitions may be misleading if sparsely populated areas still have strong economic ties to adjacent urban areas. Of the two definitions, the OMB definitions are also the most used in empirical research (Cromartie & Gibbs, 2008).

I use 1993 OMB definitions in the present study to coincide with the earliest survey year (the 1990 Census). To determine the rural status of a county, I use the 1993 Rural-Urban Continuum Codes provided by the U.S. Department of Agricultural Economic Research Service. In the analysis, I define rural status as equal to "1" if the county is classified as nonmetro according to the continuum and "0" otherwise. The terms "rural" and "nonmetro" and "urban" and "metro" are used interchangeably in this paper.

Findings

Descriptive Summary

Table 2 presents descriptive statistics for the sample used in the primary specification. The analytical sample includes 9,416 observations across counties and survey years. Treated cohorts comprise about 53 percent of observations. The proportion of bachelor's degree holders among untreated cohorts is slightly higher (17.6%) than in treated cohorts (15.5%). The proportion of rural counties in untreated cohorts is also slightly higher (65.3%) than in treated cohorts (64.0%).

Effects of State Merit-Aid Adoption

College Attainment Levels

Table 3 reports the intent-to-treat effects of state merit-aid program adoption on college attainment at the county level. Panel A presents the results of the baseline specification described above. The first row of Panel A indicates that merit-aid program adoption had generally negative effects on college attainment for counties on average. I find that adoption decreased the share of county residents that were bachelor's degree holders by 0.4 percentage points on average. Program adoption also reduced the average share of county residents with at least some college, an associate degree, or a graduate degree, respectively.

In Panel B, I examine whether the effects of merit-aid program adoption vary by whether a county is urban or rural. The first row displays the main effects on urban counties. For urban counties, program adoption has a generally positive impact on all college attainment levels except for the share of residents with some college experience or more. In particular, the share of urban county residents who are bachelor's degree holders increased by an average of 0.7 percentage points and graduate degree holders by 0.4 percentage points, respectively. Both estimates are statistically significant. In contrast to the base model that indicates a negative effect for counties on

Next, I examine the interaction of the treatment with rural status. I find that the sign of the coefficient flips in the case of each college attainment variable. Specifically, relative to urban counties, I find that state merit-aid program adoption decreased share of bachelor's degree holders in rural counties by 1.9 percentage points at the 1 percent significance level. However, the overall reduction is only about 1.2 percentage points once the main effects are taken into account. I also observe slight negative effects on the share of rural county residents that hold at least an associate degree or a graduate degree. Together, the findings suggest that merit-aid adoption has a slight positive effect on the share of college residents that are credential holders in urban areas and a small negative effect in rural areas.

Population Growth

A potential complication when analyzing the effects of program adoption on college attainment across rural and urban counties is that migration may be mediating the effects on college attainment through population. To isolate migration effects, I would need to know whether an individual migrated after high school, but Census surveys typically only collect whether a respondent relocated within the past year or the past five years. Given these limitations, I instead focus on population growth, which can be affected by net migration.

To measure changes to population growth, I rely on county population estimates provided by the Surveillance, Epidemiology, and End Results Program (SEER), produced by the U.S. Census Bureau's Population Estimates Program. The SEER data provides intercensal annual population totals for each county by age from 1969-2019. An additional benefit of this dataset is that it allows me to determine the size of birth cohorts that would have been exposed to the treatment at the time of program adoption and in subsequent years.

Using SEER data, I estimate the following:

$$lnpop_{jtc} = \alpha + \beta Merit_{jtc} * \delta Rural_{jt} + \sigma_j + \tau_t * \gamma_c + lnpopat18_{jtc} + \varepsilon_{jt}$$

where $lnpop_{jtc}$ is the log-transformed population for county *j* in survey year *t* for birth cohort year c. In this model, treatment assignment is determined by whether program adoption would have occurred within the at c + 18 years of age in the state where they currently reside. Survey year and birth cohort year have been interacted to account for year-specific effects that may vary by birth cohort. To account for the possibility of migration effects, I control the size of each birth cohort at the age of 18 using the log-transformed variable $lnpopat18_{jtc}$.

As I focus on the effects of program adoption on post-college residence, the final sample is limited to cohorts aged 25 and above at the time of the survey. To control for the possibility of extreme cohorts over ten years before the introduction of such programs, I limit treated counties to cohorts that would have been 18 within ten years of program adoption. Counties in never treated states serve as pure controls, while counties in not yet treated states serve as controls for counties in treated states that later adopt.

Table 4 provides the effects of state merit-aid program adoption on population size by county and birth cohort. Panel A is the baseline model that shows the effects regardless of the county's rural status. Controlling for the cohort size at 18, I find that program adoption decreased the county population for each birth cohort by an average of 0.006, or 1.01 percent. These estimates are only slightly larger and are less precise than the estimates without the controls.

Next, I examine the heterogeneous effects of program adoption by rural status in Panel B. For urban counties, I find that adoption increased county population totals for each birth cohort by an average of 0.01 or 1.01 percent. Adoption decreased rural county population totals for birth cohorts relative to urban counties by an average of 0.022 or 2.2%. The effect size is small but statistically significant at the 1 percent level. The overall decrease is 0.012, which is slightly less than 1 percent. The findings suggest that program adoption negatively affects population growth in rural counties relative to urban counties; however, the effect sizes are practically zero.

Event Study Analysis

To address the issue of two-way fixed effects producing biased estimates (Goodman-Bacon, 2018), I implement an event study that allows me to observe the possible heterogeneity in treatment effects over time. I estimate the following:

$$lnpop_{jtc} = \alpha + \sum_{k=-10}^{10} \beta_k D_{jtc}^k * \delta Rural_{jt} + \sigma_j + \tau_t * \gamma_c + lnpopat18_{jtc} + \varepsilon_{jt},$$

where I include a set of dummy variables, D_{jtc}^k, which take a value of one if cohort *c* is *k* cohorts removed from the last cohort eligible for merit-aid in the state. As with the difference-in-differences model, I limit the dataset to cohorts aged 25 and above at the time of the survey. Counties in treated states again serve as pure controls, while counties in not yet treated states serve as controls for counties in states that later adopt them. Treated counties are limited to event years leading and lagging up to the adoption year.

Figure 1 plots the coefficient estimates and 95% confidence intervals from estimating the equation above. Panels A and B capture urban and rural counties, respectively. The coefficient at each point in time can be interpreted as the difference between the treated and untreated groups relative to *k* = -1, which is one year before program adoption. The never-treated units also contribute to the reference group.

First, I examine whether the evidence supports the absence of pretreatment trends for untreated cohorts. Although the plot shows a slightly upward pretreatment trajectory, neither plot shows strong evidence of pretreatment trends. Second, I examine the posttreatment coefficients, which are uniformly positive in the latter event years. Both plots, however, reveal large confidence intervals that grow larger with time. Furthermore, the confidence intervals straddle the y-axis at zero, so the effect remains practically zero.

Conclusion

Past research on state merit-aid programs has focused on the state-level effects, such as the probability of attending college or being located within the state after graduation. The current study contributes to the literature by analyzing the within-state effects of such programs, specifically whether they affect the distribution of human capital across rural and urban areas. In general, I find that adoption decreased the share of county residents that were bachelor's degree holders by an average of 0.4 percentage points. Once results are decomposed into rural and urban counties, I find that adoption significantly increases college attainment in urban counties by 0.7 percentage points and decreases college attainment in rural counties by 1.2 percentage points.

At the same time, the mechanisms in effect remain unclear. On the one hand, it may be that merit-aid programs are ineffectual among rural high school students, who as a disadvantaged population, are more likely to be on the margins of college access. On the other hand, it may be that rural areas have lower net college attainment because of out-migration of rural high school graduates leaving to attend college, which is difficult to track. To explore the possibility of out-migration effects, I examine the impact of merit-aid adoption on population growth across rural and urban areas. Relative to urban counties, I also find that merit-aid adoption decreased rural county population totals for birth cohorts by an average 2.2%.

Together, the findings suggest that state merit-aid programs may contribute slightly to widening the rural-urban college attainment gap or slowed population growth in rural areas. However, the extent to which the findings can be explained by differences in the effectiveness of merit-aid on rural students or the effects rural out-migration remains unclear. To address the former, future research might examine whether the impact of merit-aid adoption differs by broad-based and more targeted state merit-aid programs. The latter is more difficult to address given the challenges of tracking migratory flows; however, it may be possible with longitudinal datasets that are able to track rural high school students through college and beyond. Nevertheless, the growing rural-urban college gap speaks to the urgency of this research.

References

Adelman, C. (2002). The relationship between urbanicity and educational outcomes. In *Increasing access to college: Extending possibilities for all students* (pp. 15–34).

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Bruce, D. J., & Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, *81*, 30–44. https://doi.org/10.1016/j.jue.2014.01.006

Chand, N., & Alexander, C. H. (2000). Multi-year averages from a rolling sample survey. *Washington, DC, US Census Bureau*.

Conley, T. G., & Taber, C. R. (2011). Inference with "difference in differences" with a small number of policy changes. *Review of Economics and Statistics*, *93*(1), 113–125. https://doi.org/10.1162/REST_a_00049

Cornwell, C., Mustard, D. B. B., & Sridhar, D. J. J. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Program. *Journal of Labor Economics*, *24*(4), 761–786. https://doi.org/10.1086/506485

Cromartie, J. B. (2017). *Rural America at a Glance* (Issue November). https://www.ers.usda.gov/webdocs/publications/85740/eib-182.pdf?v=43054

Cromartie, J. B., & Gibbs, R. M. (2008). Rural Definitions in the U.S.: Concept and Practice. *Statistics on Rural Development and Agriculture Household Income*, 25. www.fao.org/es/ess/rural/wye_city_group/Rural_Definitions_in_the_US_Concept_and_Practice.pdf

Dynarski, S. M. (2000). Hope for whom? Financial aid for the middle class and its impact on college attendance. *National Tax Journal*, *3*(3), 1–52. https://doi.org/10.3386/w7756

Dynarski, S. M. (2004). The New Merit Aid. In *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It: Vol. I* (Issue September). http://www.nber.org/chapters/c10098.pdf

Ehrenberg, R. G., Zhang, L., & Levin, J. M. (2005). Crafting a class: The trade off between merit scholarship and enrolling lower-income students. *CHERI Working Paper #71*.

Fitzpatrick, M. D., & Jones, D. (2016). Post-baccalaureate migration and merit-based scholarships. *Economics of Education Review*, *54*, 155–172. https://doi.org/http://dx.doi.org/10.1016/j.econedurev.2016.07.002

Florida, R., Mellander, C., & Stolarick, K. (2008). Inside the black box of regional development—human capital, the creative class and tolerance. *Journal of Economic Geography*, *8*(5), 615–649.

Foulkes, M., & Newbold, K. B. (2008). Using alternative data sources to study rural migration: Examples from Illinois. *Population, Space and Place*, *14*(3), 177–188. https://doi.org/10.1002/psp.482

Gibbs, R. (1998). College completion and return migration among rural youth. In R. Gibbs, P. Swaim, & R. Teixeira (Eds.), *Rural education and training in the new economy: The myth of the rural skills gap* (1st ed., pp. 61–80).

Gibbs, R. M. (1990). *Going Away to College and Wider Urban Job Opportunities Take Highly Educated*. *10*(3), 35–44.

Gibbs, R. M. (1998). College completion and return migration among rural youth. In R. Gibbs, P. Swaim, & R. Teixeira (Eds.), *Rural education and training in the new economy: The myth of the* *rural skills gap* (1st ed., pp. 61–80). Iowa State University Press. https://eric.ed.gov/?id=ED422148

Goodman-Bacon, A. (2018). *NBER Working Paper Series Difference-in-Differences with Variation in Treatment Timing*. http://www.nber.org/papers/w25018

Groen, J. A. (2011). Do Merit-Aid Programs Help States Build Skilled Workforces? *Change: The Magazine of Higher Learning*. https://doi.org/10.1080/00091383.2011.618080

Harrington, J. R., Muñoz, J., Curs, B. R., & Ehlert, M. (2016). Examining the impact of a highly targeted state administered merit aid program on brain drain: Evidence from a regression discontinuity analysis of Missouri's Bright Flight Program. *Research in Higher Education*, *57*(4), 423–447. https://doi.org/10.1007/s11162-015-9392-9

Heller, D. E., & Marin, P. (2004). State merit scholarship programs and racial inequality. In *Civil Rights Project at Harvard University (The)*.

Hickman, D. C. (2009). The effects of higher education policy on the location decision of individuals: Evidence from Florida's Bright Futures Scholarship Program. *Regional Science and Urban Economics*, *39*(5), 553–562. https://doi.org/10.1016/j.regsciurbeco.2009.04.002

Irvin, M. J., Byun, S. Yong, Meece, J. L., Farmer, T. W., & Hutchins, B. C. (2012). Educational barriers of rural youth: Relation of individual and contextual difference variables. *Journal of Career Assessment*, *20*(1), 71–87. https://doi.org/10.1177/1069072711420105

Koricich, A., Chen, X., & Hughes, R. P. (2018). Understanding the effects of rurality and socioeconomic status on college attendance and institutional choice in the United States. *The Review of Higher Education*, *41*(2), 281–305. https://doi.org/10.1353/rhe.2018.0004

Lichter, D. T., & Brown, D. L. (2014). *The New Rural-Urban Interface: Lessons for Higher Education*. *29*(1).

Manson, S., Schroeder, J., van Riper, D., Kugler, T., & Ruggles, S. (2021). *IPUMS National Historical Geographic Information System: Version 16.0 [dataset].* IPUMS.

Marre, A. W. (2017). *Rural Education at A Glance* (Issue April). https://doi.org/98

Moretti, E. (2004). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*(1–2), 175–212. https://doi.org/10.1016/j.jeconom.2003.10.015

Mykerezi, E., & Jordan, J. L. (2014). On rural-urban differences in human capital formation: Finding the ‘bottlenecks.’ *Journal of Rural Social Sciences*, *29*(1), 17–47.

Provasnik, S., KewalRamani, A., McLaughlin Coleman, M., Gilbertson, L., Herring, W., & Xie, Q. (2007). Status of Education in Rural America. In *National Center for Education Statistics*. https://doi.org/NCES 2007-040

Shapiro, J. M. (2006). Smart cities: Quality of life, productivity, and the growth effects of human capital. *Review of Economics and Statistics*, *88*(2), 324–335. https://doi.org/10.1162/rest.88.2.324

Sjoquist, D. L., & Winters, J. v. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, *80*, 39–40. https://doi.org/10.1016/j.jue.2013.10.003

Sjoquist, D. L., & Winters, J. v. (2015). State merit-based financial aid programs and college attainment. *Journal of Regional Science*, *55*(3), 364–390. https://doi.org/10.1111/jors.12161

Smith, M. H., Beaulieu, L. J., & Seraphine, A. (1995). Social capital, place of residence, and college attendance. *Rural Sociology*, *60*(3), 363–380. https://doi.org/10.1111/j.1549-0831.1995.tb00578.x

U.S. Department of Agriculture - Economic Research Service (USDA-ERS). (n.d.). *Rural Classifications*. Retrieved October 24, 2021, from https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/

U.S. Department of Agriculture - Economic Research Service (USDA-ERS). (n.d.). *Rural-Urban Continuum Codes, 1993.* Retrieved from https://wayback.archive-it.org/5923/20110914000642/http://www.ers.usda.gov/Briefing/Rurality/RuralUrbCon/priordescription.html

Winters, J. v. (2011). Why are smart cities growing? Who moves and who stays*. *Journal of Regional Science*, *51*(2), 253–270. https://doi.org/10.1111/j.1467-9787.2010.00693.x

Tables

Table 1

State Merit-Aid Programs Adopted from 1991-2004

State	Adoption year	Program name
Alaska	1999	Alaska Scholars
Arkansas	1991	Arkansas Academic Challenge Scholarship
California	2001	Competitive Cal Grant Program
Florida	1997	Florida Bright Futures Scholarship
Georgia	1993	Georgia HOPE Scholarship
Idaho	2001	Robert R. Lee Promise Scholarship
Illinois	1999-2004	Illinois Merit Recognition Scholarship
Kentucky	1999	Kentucky Educational Excellence Scholarship
Louisiana	1998	Louisiana TOPS Scholarship
Maryland	2002-2005	Maryland HOPE Scholarship
Michigan	2000-2008	Michigan Merit & Promise Scholarship
Mississippi	1996	Mississippi T.A.G. and E.S.G.
Missouri	1997	Missouri Bright Flight Scholarship
Nevada	2000	Nevada Millennium Scholarship
New Jersey	1997 (2004)	New Jersey OSRP (STARS)
New Mexico	1997	New Mexico Lottery Success Scholarship
New York	1997	N.Y. Scholarships for Academic Excellence
North Dakota	1994	North Dakota Scholars Program
Oklahoma	1996	Oklahoma PROMISE Scholarship
South Carolina	1998	South Carolina LIFE Scholarship
South Dakota	2004	South Dakota Opportunity Scholarship
Tennessee	2003	Tennessee HOPE Scholarship
Utah	1999	New Century Scholarship
Washington	1999-2006	Washington PROMISE Scholarship
West Virginia	2002	West Virginia PROMISE Scholarship

Note. Adapted from Sjoquist, D. L., & Winters, J. V. (2015). State merit-based financial aid programs and college attainment. Journal of Regional Science, 55(3), 364–390. https://doi.org/10.1111/jors.12161.

Table 2

Sample Means by County and Survey Year

	(1) All Cohorts	(2) Treated Cohorts	(3) Untreated Cohorts
Some College	0.424	0.406	0.444
Associate's Degree or Above	0.228	0.215	0.243
Bachelor's Degree or Above	0.165	0.155	0.176
Graduate Degree or Above	0.056	0.054	0.057
Rural Counties	0.653	0.64	0.668
N (Number of Counties x Survey Years)	9,416	5,027	4,389

Note. Statistics calculated using NGHIS time series table "Persons 25 Years and Over by Educational Attainment" for surveys 1990 and 2000 Decennial Censuses and 2008-2012 American Community Survey at the county level.

Table 3

Effects of Merit Aid on College Attainment Levels

	(1) Some college	(2) Associate degree	(3) Bachelor's Degree	(4) Graduate degree
Panel A. Base model				
Merit aid	-0.002	-0.005	-0.004*	-0.001
	(0.004)	(0.003)	(0.003)	(0.002)
Survey year FE	Yes	Yes	Yes	Yes
Observations	9,416	9,416	9,416	9,416
R-squared	0.877	0.789	0.688	0.494
Number of counties	3,142	3,142	3,142	3,142
Panel B. Rural county status				
Share of county	-0.003	0.004	0.007***	0.004**
(Main effects)	(0.006)	(0.004)	(0.003)	(0.002)
Share of county	0.002	-0.014***	-0.019***	-0.008***
X rural status	-0.005	-0.002	-0.002	-0.001
Survey year FE	Yes	Yes	Yes	Yes
Observations	9,416	9,416	9,416	9,416
R-squared	0.877	0.792	0.697	0.504
Number of counties	3,142	3,142	3,142	3,142

Note. Statistics calculated using 1990-2000 Census and 2008-2012 American Community Survey at the county level. Cell averages are weighted by the number of observations. Standard errors are clustered at the state level. ***, **, and * indicate statistically significant coefficients at the one, five, and ten percent levels, respectively.

Table 4

Effects of Merit Aid on Population

	(1)	(2)
Panel A. Base Model		
Merit Aid	-0.008	-0.006
	-0.008	-0.005
Population at 18 years old		0.348***
		-0.034
Observations	9,416	9,416
R-squared	0.877	0.789
Observations	1,401,276	1,401,276
Number of Counties	1,676	1,676
Survey year F.E.	Yes	Yes
Birth cohort F.E.	Yes	Yes
Population at 18 Control	No	Yes
Panel B. Nonmetro status		
Merit Aid	0.030**	0.01
(Main effects)	-0.012	-0.007
Merit Aid	-0.051***	-0.022***
X nonmetro status	-0.012	-0.006
Population at 18 years old		0.340***
		-0.033
Observations	1,400,415	1,400,415
R-squared	0.065	0.088
Number of counties	1,675	1,675
Survey year F.E.	Yes	Yes
Birth cohort F.E.	Yes	Yes
Population at 18 Control	No	Yes

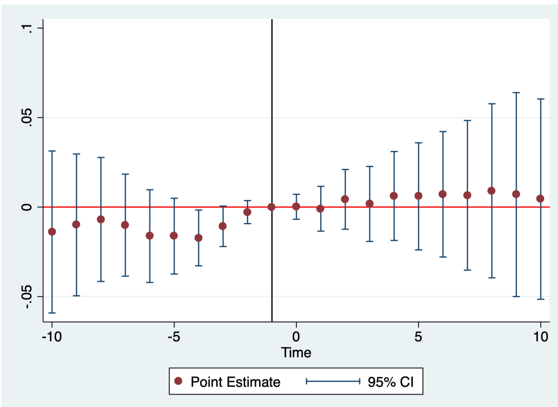
Note. Statistics calculated using 1990-2000 Census and 2008-2012 American Community Survey at the county level. Cell averages are weighted by the number of observations. Standard errors are clustered at the state level. ***, **, and * indicate statistically significant coefficients at the one, five, and ten percent levels, respectively.

Figures

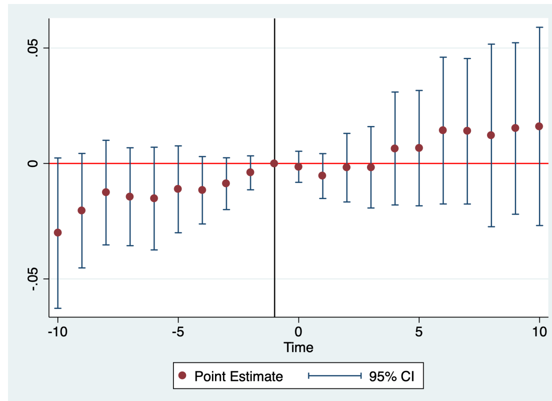
Figure 1

Event Study Figures for College Attainment by Urban and Rural County Status

A. Urban Counties



B. Rural Counties



CHAPTER THREE: THE IMPACT OF HONORS PARTICIPATION ON COLLEGIATE OUTCOMES

Collegiate honors programs are increasingly one way in which U.S. colleges and universities strive to recruit and enhance the experience of academically talented students. It is estimated that over 1,500 honors programs and colleges exist nationwide (Scott et al., 2017). Such programs typically offer targeted curricular and co-curricular opportunities, including academically rigorous versions of general education courses, smaller classes, increased faculty interaction and mentoring, and networking opportunities (Astin, 1993; Boazman et al., 2012). To be eligible to participate, students typically must demonstrate prior academic ability and achievement, usually based on high school grade point average (GPA) and standardized test scores (e.g., ACT, SAT) (Achterberg, 2005; Digby, 2005).

In the past two decades, there has been an increase in the number of honors colleges established, with more than 200 now in existence (Long, 2002). Honors colleges differ from honors programs in so far as they are infrastructurally more complex and require more resources (Scott et al., 2017). In particular, they are more likely to have separate residential housing options for honors students and additional scholarship opportunities (Long, 2002). Honors colleges are also concentrated at four-year public institutions (Scott et al., 2017). One possible reason is that public institutions look to honors colleges as a recruiting strategy that promises to deliver high-ability students with a less costly alternative to more selective liberal arts colleges (Long, 2002). In this respect, the investment in honors colleges coincides with broader trends at both the state and institutional levels to attract academically talented students through merit-based aid (Singell & Tang, 2012).

Popular sentiment in U.S. higher education and beyond suggests that honors programs are beneficial for students (Digby, 2005). However, research on undergraduate honors students and the programs that serve them is somewhat limited (Cosgrove, 2004; Rinn & Plucker, 2019). A few studies suggest that honors programming has a positive effect on college outcomes, such as GPA (Hartleroad, 2005; Rinn, 2007; Furtwengler, 2015; Cosgrove 2004; Spisak & Squires, 2006), first year to the second-year retention (Keller and Lacy, 2013), and graduation rates (Keller and Lacy, 2013; Slavin, Coladarci, and Pratt, 2008). However, these studies are only able to produce simple correlational effects by comparing honors and non-honors students. They do not account for unobserved heterogeneity that may bias estimates of the impact of honors participation.

The current study estimates the impact on honors participation on collegiate outcomes, such as cumulative GPA, retention, and graduation. The setting is the University of Missouri (MU), which is a public land-grant and research-intensive university with one of the nation's oldest honors colleges. Until Fall 2019, first-time college students were automatically admitted to the honors college if they met a combination of GPA, high school class rank, and ACT requirements. Using entering freshman cohorts from fall 2008 to 2016, the analysis takes advantage of eligibility criteria based on strict set of cutoffs. These cutoffs allow for a regression discontinuity design in the causal effects of honors college participation can be estimated by comparing honors and non-honors participants directly above and below the given threshold.

As honors programs and colleges require substantial financial and human resources, the findings provide insight on whether such programs contribute to student success (Bowman & Culver, 2018).

Relevant Literature

Despite the popularity of honors programs and colleges, there is limited evidence on whether honors participation benefits students or merely reflects the positive selection of high-ability students who may be already motivated for success (Rinn & Plucker, 2019). The literature review summarizes the extant research on the relationship between honors participation and each of the collegiate outcomes analyzed in the current paper. The limitations of the existing research, in so far as the studies only control for observables, is then discussed. In contrast, the current study is the first to use a regression discontinuity framework to estimate the effect of honors participation on collegeiate outcomes.

Studies find a generally positive relationship between honors participation and GPA (Bowman & Culver, 2018; Cosgrove, 2004; Furtwengler, 2015; Hartleroad, 2005; Rinn, 2007; Shushok, 2002). For instance, Rinn (2007) compared GPAs of honors and non-honors students with a 1,300 SAT scores and above (the cutoff for honors eligibility). Honors students across class levels had significantly higher GPAs on average (3.74) than non-honors students (3.26), after controlling for SAT scores. In some cases, positive estimates found in the first year disappear by the fourth year. For instance, Shushok (2002, 2006) found that while honors students had significantly higher cumulative GPAs by the end of their first year (honors 3.41, non-honors 3.18), the differences were no longer statistically significant by the end of the fourth year of study (honors 3.46, non-honors 3.40).

There is also some evidence the positive relationship between honors participation and retention and graduation rates in various single-institution studies (Bowman & Culver, 2018; Cosgrove, 2004; Keller & Lacy, 2013; Slavin et al., 2008). Shushok (2002, 2006) found that honors students were 7 percentage points more likely to return their second year relative to non-honors at the, while Slavin et al. (2008), at the University of Main, estimated that the odds of an honors student returning the fall semester were roughly three times greater than the odds of a non-honors student returning, controlling for SAT and high school rank. Finally, Keller & Lacy (2013) found that honors students were five percentage points more likely to return for their second year of schooling, controlling for previous academic achievement and other background characteristics. They also estimate that honors participation is associated with an increased probability of graduating four, five, and six years by 8.4 percentage points, 12.3 percentage points, and 14 percentage points, respectively.

While the research reports generally positive effects of honors participation, one of the inherent problems with measuring the effects of honors participation is the selection problem inherent in naïve comparisons of honors and non-honors (Bowman & Culver, 2018). In the existing literature, studies often attempt to overcome the selection issue by limiting their analysis to honors and high-ability non-honors students and controlling on observables, such as high school GPA or test scores (e.g., Cosgrove, 2004; Hartleroad, 2005; Rinn, 2007). For instance, Rinn (2007) compared GPAs of honors and non-honors students with a 1,300 SAT scores and above (the cutoff for honors eligibility).

Cosgrove (2004) also limits their sample honors and high-ability non-honors students (1150 or better SAT) in their multisite study of four-year colleges in Pennsylvania.

There are also a few studies that use propensity score analysis to attempt to reduce bias by matching similar honors students and non-honors students (Bowman & Culver, 2018; Furtwengler, 2015; Shushok, 2002, 2006). However, a limitation of propensity score analysis, however, is that they still rely on observables. There may still be unobserved heterogeneity between honors participants and non-participants that may serve to bias the estimates if treatment receipt depends on unobserved variables that are correlated with the outcome. For example, students who choose to participate in honors programming (after qualification) may already be high achievers, have positive academic self-perceptions, and are motivated for success (Rinn & Plucker, 2019).

The Setting

The setting is the University of Missouri (MU), a public land-grant and research-intensive university with over 30,000 students—roughly 22,500 of whom are undergraduate students. Founded in 1958, the Honors College is one nation's first honors colleges and has been ranked among the top public honors colleges (Willingham, 2014). The Honors College strives to enhance the experience of academically talented students through specialized curricular and co-curricular activities and events designed to both challenge and engage. The curriculum includes over 200 unique courses per year, whether offered directly by the college itself as "general honors" or as field-specific honors courses through the academic departments. Students who complete 20 credit hours graduate with an honors certificate. The Honors College also offers a variety of co-curricular and extra-curricular activities and events designed to enhance the "Honors

Experience," such as private seminars with distinguished speakers, site tours, study abroad programs, and guest visits to graduate seminars. First-year students also have an option to live in a residential hall community with their honors peers.

Until Fall 2019, first-time college students were automatically admitted to the Honors College if they met a combination of eligibility criteria that included ACT score, core GPA and high school percentile ranking (HSPR). The admissions criteria followed a tiered structure: 29 ACT score, 3.91 core GPA, and 95th HSPR; 30 ACT score, 3.74 core GPA, and 90th HSPR; and 31 ACT score; 3.51 core GPA, 85th HSPR. The core GPA was calculated by the University using high school grades from all English, science, social studies, and foreign language courses. In the case that a student's high school did not use a ranking system, the Honors College waived the percentile rank requirement. In addition, students who did not meet the above criteria could also apply for admission.

Data

To examine the impact of honors participation on college student outcomes, we use individual-level administrative data. The University of Missouri provided data on students' background characteristics, high school performance, and standardized test scores for all applicants within the top 75th percentile of their graduating high school class with a minimum core GPA of 3.0. honors eligibility criteria are well above these cutoffs, thus limiting the sample in this way raised no concerns.

We also received enrollment and transcript data for all applicants that were admitted to MU. We limited the sample to first-time college enrollees who were admitted in the fall semesters of 2008 through 2016, with outcomes tracked through Spring 2022. We excluded entering cohorts before 2008 because of changes to the administrative record-keeping database that took place at that time. Entering cohorts after 2016 were excluded to allow a minimum of six years to transpire after enrollment to calculate graduation rates. The final analytical sample included approximately 18,944 students.

Honors Eligibility

We received ACT composite scores and subtest scores for each applicant that submitted scores to MU. Table 1 presents the total students that met each eligibility threshold by highest composite ACT score. For example, 4,511 students who met the 29 ACT minimum threshold also met the 3.91 core GPA and 95th percentile criteria, while 10,107 students also met the 3.51 GPA and 85th percentile criteria.

Empirical Strategy

One of the limitations to estimate the effect of honors programming is that honors and non-honors students may differ in ways systematically that are correlated with the outcomes. For instance, students who meet the eligibility requirements of honors programs may differ from students who are not with regard to academic achievement, coursework, motivation, and experiences (Bowman & Culver, 2018). Further, honors eligible students who choose to participate may differ from eligible students who choose not to participate. In addition to their achievement orientation and motivation, they may also have different academic self-perceptions (Rinn & Plucker, 2019). Therefore, naïve estimates of honors programming, not only capture program effects, but the influence pretreatment characteristics as well.

The design of honors eligibility criteria allows us to use a regression discontinuity design (RDD) to control for unobserved heterogeneity to obtain unbiased estimates of the impact of honors college eligibility on a variety of outcomes (Angrist & Pischke, 2008).

RDD approaches can be "sharp" in which the running variable is a deterministic function of treatment receipt (i.e., compliance is perfect) or "fuzzy" in which the running variable is highly correlated with treatment receipt but is not entirely deterministic (Lesik, 2006; van der Kaauw, 2002). In the current paper, the fuzzy RDD approach is preferable as the running variable is not entirely deterministic of treatment receipt. For instance, not every student who is honors eligible completes an honors course and, in other cases, students who are not honors eligible may enroll in an honors course. We determine treatment receipt based on whether a student completed at least one honors course.

The fuzzy RDD removes noncompliance bias through a two-stage procedure similar to the instrumental variables empirical strategy (Angrist & Pischke, 2008; Cunningham, 2021). The first stage of the fuzzy RDD estimates the probability of receiving the treatment by regressing the actual treatment status on the assignment variable and indicator variables based on the decision rule:

$$T_i = \alpha_1 + \gamma_0 D_i + f_1(Z_i) + \varepsilon_i, \tag{1}$$

where T_i is equal to 1 if student *i* received the treatment (i.e., took at least one honors course), and 0 otherwise; D_i is equal to 1 if individual *i* is assigned to the treatment based on the cutoff rule, and 0 otherwise; Z_i is the running variable recentered at 0 for individual *i*; $f_1(Z_i)$ is the relationship between cutoff and treatment assignment; and ε_i is the random error of the first-stage regression, which is assumed to be identically and independently distributed.

To determine the causal impact of the treatment, the first stage equation is estimated jointly with the second stage equation:

$$Y_i = \alpha + \beta_0 \hat{T}_i + f_2(Z_i) + \mu_i, \tag{2}$$

where Y_i is the outcome (GPA, persistence, graduation rate) for student i; $\hat{T}_i$ is the predicted likelihood of honors participation based on the running variable from the first-stage equation; $f_2(Z_i)$ is the relationship between the running variable and treatment assignment for individual i; and μ_i is the random error of the first-stage regression, which is assumed to be identically and independently distributed.

The coefficient of interest is β_0, which can be interpreted as the local average treatment effect (LATE) of receiving the treatment on compliers near the threshold (Angrist et al., 1996; Imbens & Angrist, 1994). Following Lee and Card (2008), we cluster standard errors by the running variable to account for the grouping of students by their ability level.

Rather than implement multiple cutoff scores for the running variable, we focus on first ACT score as the primary threshold. However, using whole ACT points creates a coarse running variable (e.g., 28, 29, 30), so we averaged the four subtest scores (English, Math, Reading, and Scientific Reasoning) to arrive at new composite scores with four potential points per score (e.g., 28.0, 28.25, 28.5, 28.75). Therefore, we use the ACT score of 28.5 (which rounds up to 29) as the primary threshold for determining honors eligibility. We further limit the sample both above and below the ACT threshold to students who also met the minimum core GPA (3.51) and percentile rank (85th) requirements.

Identifying Assumptions

The RDD is a quasi-experimental method used to estimate the causal treatment effect of policies or interventions in situations in which assignment to the treatment is based on whether the value of a continuous rating score falls above or below a given cutoff (Imbens & Lemieux, 2008). In a sharp RD, all subjects with a score above the threshold will receive the treatment while those below do not. The difference in outcomes between the treatment and the control groups can be considered the causal effect of the treatment for subjects around the threshold. The discontinuity in outcomes between the control and treatment groups at the threshold is otherwise known as the local average treatment effect (LATE).

The identification strategy for the RDD rests on the key assumption that the relationship between the running variable and outcome is appropriately specified so that those right above and below the cutoff are equal in expectation (Imbens & Lemieux, 2008). First, the running variable and the cutoff must be independently determined. Second, the outcome-running variable relationship must be continuous. As these conditions are met, the RDD provides unbiased causal estimates of the treatment impact for those with rating score values at or near the threshold. These assumptions are tested in the following sections.

Independence of the Running Variable

Research shows that public knowledge of treatment assignment rules and cutoffs may lead to manipulation of the assignment variable to receive the treatment (Imbens & Lemieux, 2008; Lee & Lemieux, 2010; McCrary, 2008). In the case of the Honors College, since students are aware of the honors eligibility criteria before applying to the University, it is plausible that students may retake the ACT exam if they did not receive a sufficient score on their first instance. For instance, Harrington et al. (2016) found that students near the cutoff for Missouri's Bright Flight program, a composite score of 30, were more likely to retake the ACT. Given the propensity of Missouri students to retake, we examined the discontinuity for both the first ACT score and maximum ACT score.

We tested the independence assumption by visually inspecting the density of the running variable near the cutoff score using a histogram (McCrary, 2008). A smooth density curve suggests there has not been any manipulation of the running variable near the cutoff, while a discontinuity near the cutoff suggests otherwise. A discontinuity in the frequency of observed values of the running variable near the cutoff indicates possible manipulation of the running variable, which creates concerns for the validity of the RDD design.

Figure 1 presents a histogram of the first ACT scores for all applicants. As expected, the density curve of scores is smooth among first-time test takers. By comparison, Figure 2 presents a histogram of the maximum ACT scores for all applicants. We observe some sort of discontinuity in the distribution of ACT scores at 27 and 31, which suggests that students may have retaken the test to reach some sort of threshold, but distance from the honors cutoff of 29 suggests that students may be responding to an event other than honors eligibility. For instance, we confirmed that MU has automatic scholarship level requirements which coincide with both scores, while the state of Missouri also has the Bright Flight program, which provides in-state scholarships to students in the top 3 percent of all Missouri test takers.

Continuity of the Outcome-Running Variable

The RDD also requires evidence that there is a smooth relationship between the outcome and the running variable at the cutoff score. While we would expect to see a discontinuity in the outcomes at the threshold, it is also necessary to provide evidence that the discontinuity is to be attributed to the treatment and not a concurrent event that also affects the outcomes (Imbens & Lemieux, 2008). We tested this assumption by examining the pretreatment covariates as dependent variables of the running variable (i.e., the estimated effect of HC eligibility on covariates). A lack of discontinuity in pretreatment covariates at the threshold provides empirical evidence for the continuity of the expected potential outcomes. Ultimately, this exercise tested for random assignment around the discontinuity point (Imbens & Lemieux, 2008).

Table 2 presents the covariate smoothness at the HC eligibility threshold based on initial ACT scores at various bandwidths. The baseline column includes control group means and standard deviations for students one point below the threshold (i.e., ACT scores 27.5 through 28.25) for each covariate. As a baseline for the control group, approximately 8.1% of students directly below the threshold have taken the SAT with an average score of about 1,227.9. About 17.4% are first-generation, 69.6% are in-state residents, and 58.5% are female. In terms of race/ethnicity, most baseline students are White (94.3%), followed by Black (2.5%), and Hispanic (2.4%).

The second column of Table 2 presents the relationship between HC eligibility (i.e., ACT score of 28.5 or greater) and each of the covariates for students within 1.5 ACT points of the threshold (i.e., ACT scores of 27.0 through 30.0). At the threshold, we find that HC eligibility increases the probability of taking the SAT by 2.3 percentage points and increases SAT scores by about 7.0 points on average. HC eligibility has a slightly negative effect on the probability of being first-generation, an in-state resident, or Black, while the probability of being female, Asian, or Hispanic increases slightly. Most covariates are not statistically significant at any bandwidth, except the probability of being Hispanic, which is statistically significant across most bandwidths, and the probability of being Black, which is statistically significant at the bandwidths of 3.0 and 3.5.

In contrast, Table 3 presents the covariate smoothness at the threshold based on maximum ACT at various bandwidths. We find that honors eligibility has a significant positive effect on SAT score, being female, being Asian, and being Hispanic at various bandwidths, and a significant negative effect on being first-generation, in-state residency status, and being Black. Although we do not have strong suspicions that students are retaking the ACT to be admitted to the Honors College specifically—but rather to increase their odds of acceptance to MU or some other college, earn the Bright Flight scholarship, or meet some other benchmark—this retaking may be contributing to some imbalance at the honors eligibility threshold based on maximum ACT.

Given the slight imbalances caused by using the maximum ACT score, our analysis primarily focuses on the first ACT score to be conservative with our estimates and reduce potential biases due to retaking the ACT. Based on a visual inspection of the data, we choose to present most of our estimates at the bandwidth of 1.5 (i.e., ACT scores of 27.0 through 30.0), where we find the density of observations to be smoothest around the cutoff. This also helps us to avoid any confounding effects that may be caused by scholarship cutoffs at the same thresholds.

Findings

First-Stage Estimates: Likelihood of Honors Participation

As described above, assignment to the treatment based on honors eligibility does not perfectly predict actual participation in honors courses. The fuzzy regression discontinuity estimation requires strong first-stage estimates to ensure the precision and validity of the second-stage results. Table 4 presents the reduced form results of the first-stage regression using the first ACT score as the running variable for each outcome. Note that the sample is limited to students entering in the fall semester and we were provided semester-by-semester data, so our indicator variables for GPA and persistence capture the total number of fall semesters since enrollment (e.g., first fall, second fall, third fall, and so on). The first year corresponds to what we commonly refer to as the freshman year, while the fourth year corresponds to what we know as the senior year. We omit the persistence indicator for the first fall semester as it includes everyone in the sample, whereas we are more interested in whether students return the following year (e.g., their sophomore year) and so on. However, our indicators for graduation are coded slightly differently, telling us whether a student graduated by the end of their fourth, fifth, or sixth years, respectively.

Table 4 presents the results of the first-stage regression. The first-stage results are similar among several of the outcomes within the same bandwidth because there is no change in both dependent and independent variables for outcomes with similar sample sizes. At the bandwidth of 1.5, all coefficients are positive and statistically significant at the 1% level, although the magnitude of the coefficients is small, which may reduce the statistical power of the model and increase standard errors in the second stage. All F-statistics are also below 10, which indicates a somewhat weak instrument (Stock et al., 2002). However, Figure 3 plots the relationship between first ACT composite source and the proportion of students that took at least one honors course. The discontinuity observed near the threshold confirms that there is a relationship between being honors eligible and honors participation.

Second-Stage Estimates: Effect of Honors Participation on Outcome

Table 5 presents the second-stage regressions for each outcome. For comparison purposes, we include a baseline group of students 1.0 ACT point below the threshold (ACT scores 27.25 through 28.25) and their standard deviations. Among the baseline group, we find that the average cumulative GPA by the end of the first fall semester is approximately 3.5-grade points, which increases to an average of 3.6-grade points by the end of senior year. At the bandwidth of 1.5, we find positive effects of honors participation on each GPA variable, with statistically significant estimates on first year and senior year GPAs. For instance, honors participation increases the average cumulative GPA by the end of the first year by 0.2-grade points and the end of senior year by 0.3-grade points for compliers near the threshold. The magnitude of these estimates is large, so they are of both practical and statistical significance.

Table 5 also presents estimates related to student persistence and graduation. Among the baseline control group, we find that approximately 95.4% return the fall semester of their sophomore year. At the bandwidth of 1.5, we find that honors participation increases the probability of persisting to sophomore year by 11.9 percentage points for compliers near the threshold, but the finding is not statistically significant. In terms of graduation rates, roughly 68.3% of the baseline control group graduated within four years, 86.8% in five years, and 88.8% within six years, respectively. At the bandwidth of 1.5, we see honors participation increases four-year graduation rates by 18.2 percentage points, five-year rates by 15.4 percentage points, and six-year rates by

12.7 percentage points, respectively. While the estimates are large, none are statistically significant. The sign for four-year graduation rate also switched to negative across bandwidths.

Alternative Functional Forms

The model above assumes the functional form between the outcome and ACT score is linear. However, misspecification of the functional form of the decision variable may serve to bias the estimates of the treatment effect (Angrist & Pischke, 2008). To test the robustness of our estimates, we present various functional forms for selected outcomes in Table 6. The first column is the local linear regression discussed above while the second column presents the quadratic fit. We do not explore the cubic functional form due to limited degrees of freedom given only 6 observations above and below the cutoff at the bandwidth of 1.5. In Column 2, we find that the quadratic term is not statistically significant for any of the outcomes. We also present scatterplots for the linear (Figure 3) and quadratic (Figure 4) model to visually inspect the functional forms. The scatterplots confirm that the linear fit is the best specification as there is not strong evidence of a quadratic term for each of the outcomes.

Conclusion

While past research suggests that honors participation has a positive effect on collegiate outcomes, these studies have been largely limited to descriptive analysis (Furtwengler, 2015; Hartleroad, 2005; Keller & Lacy, 2013; Rinn, 2007; Slavin et al., 2008). This study presents contributes to the growing literature of the impact of honors participation by using a regression discontinuity design (RDD) to present casual estimates. In particular, we find that honors participation increases average first-year

GPA by 0.2-grade points and senior-year GPA by 0.4-grade points for compliers near the threshold. However, these findings may reflect that honors courses have different grade distributions than non-honors courses. In addition, while we find generally positive effects on persistence and graduation rates, the point estimates are noisy, and none are statistically significant.

As honors programs require considerable financial and human resources, the findings of the present study shed light on whether honors programs contribute to student success (Bowman & Culver, 2018). Generally, more quantitative research of high-achieving students within the context of post-secondary honors programs is necessary to increase validity and generalizability of the results and conclusions regarding the effects of participating in honors (Furtwengler, 2015). The study also has implications in the context of the increasing popularity of accountability systems, such as performance-based funding whereby states use metrics, such a persistence, retention, and time to graduation, to justify the allocation of their funds to institutions of higher education (Burke & Modaressi, 2001).

The analysis is limited in a few ways, however. As with RDDs more generally, there is limited external validity as the local average treatment effect occurs within a relatively narrow bandwidth. If the findings were to apply to any other settings, it would most likely be other public research universities with similar honors eligibility criteria. Regarding internal validity, students' propensity to retake ACT contributed to imbalances in several of the covariates and reduced the integrity of the maximum ACT score as a running variable. The reliance on initial ACT scores to determine eligibility ultimately produced less precise estimates. Finally, the study is limited to measuring honors participation via course-taking, which is only one aspect of the experience. Future research can benefit from examining the relationship between more specific forms of honors participation, such as residential housing or mentoring, and various outcomes.

References

Achterberg, C. (2005). What is an honors student? *Journal of the National Collegiate Honors Council*, *170*, 75–84. https://digitalcommons.unl.edu/nchcjournal

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455. https://doi.org/10.1080/01621459.1996.10476902

Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Boazman, J. K., Sayler, M. F., & Easton-Brooks, D. (2012). Mediating factors of personal wellbeing in gifted college students: Early-college entrants and honors college students. *Journal of Social Research & Policy*, *3*(2), 111–131.

Bowman, N. A., & Culver, K. C. (2018). When do honors programs make the grade? Conditional effects on college satisfaction, achievement, retention, and graduation. *Research in Higher Education*, *59*(3), 249–272.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2017). *Mobility report cards: The role of colleges in intergenerational mobility* (NBER Working Paper No. 23618). National Bureau of Economic Research. https://www.nber.org/papers/23618

Cosgrove, J. (2004). The impact of honors programs on undergraduate academic performance, retention, and graduation. *Journal of the National Collegiate Honors, 137*. https://digitalcommons.unl.edu/nchcjournal/137

Cunningham, S. (2021). Causal Inference: The Mixtape. In *Causal Inference: The Mixtape*. Yale University Press.

Digby, J. (2005). *Peterson's honors programs: The official guide of the National Collegiate Honors Council* (4th ed).

Furtwengler, S. R. (2015). Effects of participation in a post-secondary honors program with covariate adjustment using propensity score. *Journal of Advanced Academics*, *26*(4), 274–293. https://doi.org/10.1177/1932202X15603365

Hahn, J., Todd, P., & van der Klaauw, W. (1999). *Evaluating the effect of an antidiscrimination law using a regression-discontinuity design* (NBER Working Paper No. 7131). National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w7131/w7131.pdf

Harrington, J. R., Muñoz, J., Curs, B. R., & Ehlert, M. (2016). Examining the impact of a highly targeted state administered merit aid program on brain drain: Evidence from a regression discontinuity analysis of Missouri's Bright Flight Program. *Research in Higher Education*, *57*(4), 423–447. https://doi.org/10.1007/s11162-015-9392-9

Hartleroad, G. E. (2005). Comparison of the academic achievement of first-year female honors program and non-honors program engineering students. *Journal of the National Collegiate Honors Council*, *Fall/Winter*, 109–120.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467. https://doi.org/10.2307/2951620

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Keller, R. R., & Lacy, M. G. (2013). Propensity score analysis of an honors program's contribution to students' retention and graduation outcomes. *Journal of the National Collegiate Honors Council, 14*(2). https://digitalcommons.unl.edu/nchcjournal/397/

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.

Long, B. T. (2002). *Attracting the Best: The Use of Honors Programs to Compete for Students*.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714.

Rinn, A. N. (2007). Effects of programmatic selectivity on the academic Achievement, academic self-Concepts, and aspirations of gifted college students. *Gifted Child Quarterly, 51*(3), 232–245. https://doi.org/10.1177/0016986207302718

Rinn, A. N., & Plucker, J. A. (2019). High-Ability College Students and Undergraduate Honors Programs: A Systematic Review. *Journal for the Education of the Gifted*. https://doi.org/10.1177/0162353219855678

Scott, R., Smith, P., & Cognard-Black, A. (2017). Demography of Honors: The Census of U.S. Honors Programs and Colleges. *Journal of the National Collegiate Honors Council*, *18*(1), 189–224. http://digitalcommons.unl.edu/nchcjournal/548

Shushok, F. (2002). *Educating the best and the brightest: Collegiate honors programs and the intellectual, social and psychological development of students*.

Shushok, F. (2006). Student Outcomes and Honors Programs: A Longitudinal Study of 172 Honors Students 2000-2004. *Journal of the National Collegiate Honors Council*, *7*(2), 85–96. http://digitalcommons.unl.edu/nchcjournal/

Singell, L. D., & Tang, H.-H. (2012). The Pursuit of Excellence: An Analysis of the Honors College Application and Enrollment Decision for a Large Public University. *Research in Higher Education*, *53*(7), 717–737. https://doi.org/10.1007/s11162-012-9255-6

Slavin, C., Coladarci, T., & Pratt, P. A. (2008). Is student participation in an honors program related to retention and graduation rates? *Journal of the National Collegiate Honors Council--Online Archive*, 66.

Stock, J., Yogo, M., & Wright, J. (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business and Economic Statistics*, *20*.

Willingham, J. (2014). *A review of fifty public university honors programs with additional reviews of five regional public university honors programs.*

Tables

Table 1

Summary of HC Eligibility Thresholds, 2008-2016

	Partner Criteria		
Primary Criteria	3.91 GPA & 95th Percentile	3.74 GPA & 90th Percentile	3.51 GPA & 85th Percentile
29 ACT	1,953	3,395	4,604
30 ACT	1,493	2,495	3,297
31 ACT	1,065	1,727	2,206
Total	4,511	7,617	10,107

Note. Up until Fall 2019, the University of Missouri used the above criteria to automatically admit students to the Honors College.

Table 2

Covariate Smoothness at the Threshold, First ACT Composite Score

		Bandwidth					
	Baseline	1.5	2.0	2.5	3.0	3.5	4.0
Took SAT	0.081	0.023	0.007	0.005	0.002	-0.005	-0.006
	(0.273)	(0.017)	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)
SAT Score	1227.870	7.034	30.769	18.459	20.720	23.193+	17.373
	(74.874)	(19.551)	(23.002)	(15.032)	(13.527)	(13.051)	(12.281)
First-generation	0.174	-0.023	-0.005	-0.021	-0.019	-0.015	-0.014
	(0.380)	(0.025)	(0.021)	(0.019)	(0.017)	(0.016)	(0.015)
In-state	0.696	-0.044	-0.029	-0.019	-0.008	0.008	0.011
	(0.460)	(0.028)	(0.024)	(0.022)	(0.020)	(0.018)	(0.017)
Female	0.585	0.008	0.005	0.019	0.027	0.006	0.015
	(0.493)	(0.031)	(0.027)	(0.024)	(0.022)	(0.020)	(0.019)
Asian	0.026	0.012	0.007	0.002	-0.002	0.003	0.003
	(0.158)	(0.011)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
Black	0.025	-0.004	-0.007	-0.012	-0.013*	-0.012*	-0.008
	(0.155)	(0.010)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)
Hispanic	0.024	0.017*	0.022*:	0.017*	0.012+	0.013*	0.011+
	(0.153)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)
White	0.943	-0.001	-0.008	0.003	0.006	0.002	-0.000
	(0.232)	(0.015)	(0.013)	(0.011)	(0.010)	(0.010)	(0.009)
Minority (Black/Hi:	0.073	0.020	0.019	0.008	-0.002	0.003	0.005
	(0.261)	(0.017)	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)
N	2099	4160	5421	6615	7675	8669	9651

Note. + p<0.1, * p<0.05, ** p<0.01. Baseline estimates based on control group means and standard deviations for students one ACT point below threshold. Coefficients are treatment effects at various bandwidths (e.g., 1.5 ACT points within the threshold).

Table 3

Covariate Smoothness at the Threshold, Maximum ACT Composite Score

		Bandwidth					
	Baseline	1.5	2.0	2.5	3.0	3.5	4.0
Took SAT	0.064	0.015	0.010	0.014	0.014	0.005	0.002
	(0.244)	(0.015)	(0.012)	(0.011)	(0.010)	(0.010)	(0.009)
SAT score	1193.110	18.996	33.851+	29.476*	16.047	23.236+	23.404+
	(77.765)	(18.141)	(19.182)	(14.296)	(13.126)	(12.607)	(11.973)
First-generation	0.191	-0.012	0.002	0.002	-0.009	-0.016	-0.016
	(0.393)	(0.023)	(0.020)	(0.018)	(0.017)	(0.016)	(0.015)
In-state	0.729	-0.069**	-0.077**	-0.070**	-0.066**	-0.041*	-0.033*
	(0.444)	(0.025)	(0.021)	(0.019)	(0.018)	(0.017)	(0.016)
Female	0.601	0.072*	0.071**	0.075**	0.075**	0.061**	0.066*
	(0.490)	(0.028)	(0.024)	(0.022)	(0.020)	(0.019)	(0.017)
Asian	0.025	0.028**	0.014	0.008	0.004	0.006	0.008
	(0.155)	(0.010)	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)
Black	0.026	-0.003	0.003	-0.001	-0.000	-0.001	0.003
	(0.161)	(0.009)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)
Hispanic	0.023	0.014+	0.019**	0.017**	0.014*	0.013*	0.008
	(0.150)	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)
White	0.942	-0.016	-0.022+	-0.009	-0.008	-0.008	-0.011
	(0.233)	(0.014)	(0.012)	(0.011)	(0.010)	(0.009)	(0.009)
Minority	0.072	0.035*	0.035**	0.025*	0.019+	0.020+	0.021*
	(0.259)	(0.015)	(0.013)	(0.012)	(0.011)	(0.010)	(0.010)
N	2605	4831	6473	7887	9034	10073	11045

Note. + p<0.1, * p<0.05, ** p<0.01. Baseline estimates based on control group means and standard deviations for students one ACT point below threshold. Coefficients are treatment effects at various bandwidths (e.g., 1.5 ACT points within the threshold).

Table 4

Fuzzy Regression Discontinuity Estimates at Various Bandwidths (First Stage)

	Bandwidth					
Outcome	1.5	2.0	2.5	3.0	3.5	4.0
GPA 1	0.198***	0.123***	0.132***	0.130***	0.127***	0.138***
	(0.014)	(0.026)	(0.023)	(0.021)	(0.019)	(0.018)
N	11,924	5,394	6,582	7,636	8,629	9,609
F-Test	326.3	22.46	43.42	76.31	128.5	205.8
Prob > F	0	2.20e-06	0	0	0	0
GPA 2	0.162***	0.132***	0.139***	0.134***	0.133***	0.144***
	(0.032)	(0.027)	(0.024)	(0.022)	(0.020)	(0.018)
N	3,901	5,089	6,208	7,200	8,118	9,026
F-Test	3.061	17.99	39.44	73.15	117.4	189.4
Prob > F	0.0803	2.26e-05	3.62e-10	0	0	0
GPA 3	0.159***	0.131***	0.136***	0.129***	0.131***	0.141***
	(0.033)	(0.028)	(0.025)	(0.022)	(0.021)	(0.019)
N	3,669	4,802	5,857	6,782	7,648	8,506
F-Test	3.310	16.94	37.50	72.09	107	175.7
Prob > F	0.0689	3.92e-05	9.75e-10	0	0	0
GPA 4	0.204***	0.175***	0.171***	0.163***	0.162***	0.177***
	(0.037)	(0.032)	(0.028)	(0.025)	(0.023)	(0.022)
N	2,769	3,622	4,427	5,125	5,764	6,411
F-Test	0.923	6.885	24.12	47.70	79.76	117.5
Prob > F	0.337	0.00873	9.37e-07	0	0	0
Persist 2	0.152***	0.122***	0.131***	0.130***	0.126***	0.138***
	(0.031)	(0.026)	(0.023)	(0.021)	(0.019)	(0.018)
N	4,160	5,421	6,615	7,675	8,669	9,651
F-Test	4.530	22.84	44.32	77.61	129.4	207.1
Prob > F	0.0334	1.80e-06	0	0	0	0
Graduated in 4 years or less	0.152***	0.122***	0.131***	0.130***	0.126***	0.138***
	(0.031)	(0.026)	(0.023)	(0.021)	(0.019)	(0.018)
N	4,160	5,421	6,615	7,675	8,669	9,651
F-Test	4.530	22.84	44.32	77.61	129.4	207.1
Prob > F	0.0334	1.80e-06	0	0	0	0
Graduated in 5 years or less	0.152***	0.122***	0.131***	0.130***	0.126***	0.138***
	(0.031)	(0.026)	(0.023)	(0.021)	(0.019)	(0.018)
N	4,160	5,421	6,615	7,675	8,669	9,651
F-Test	4.530	22.84	44.32	77.61	129.4	207.1
Prob > F	0.0334	1.80e-06	0	0	0	0
Graduated in 6 years or less	0.152***	0.122***	0.131***	0.130***	0.126***	0.138***
	(0.031)	(0.026)	(0.023)	(0.021)	(0.019)	(0.018)
N	4,160	5,421	6,615	7,675	8,669	9,651
F-Test	4.530	22.84	44.32	77.61	129.4	207.1
Prob > F	0.0334	1.80e-06	0	0	0	0

Note. + p<0.1, * p<0.05, ** p<0.01. Coefficients are treatment effects at various bandwidths (e.g., 1.5 ACT points within the threshold).

Table 5

Fuzzy Regression Discontinuity at Various Bandwidths (Second Stage)

		Bandwidth					
Outcome	Baseline	1.5	2.0	2.5	3.0	3.5	4.0
GPA 1	3.504	0.172**	0.250	0.176	0.270	0.294*	0.259*
	(0.509)	(0.079)	(0.216)	(0.179)	(0.168)	(0.161)	(0.137)
GPA 2	3.499	0.240	0.188	-0.003	0.092	0.063	0.020
	(0.508)	(0.205)	(0.213)	(0.182)	(0.172)	(0.163)	(0.142)
GPA 3	3.526	0.002	0.065	0.192	0.239	0.204	0.158
	(0.506)	(0.206)	(0.211)	(0.182)	(0.177)	(0.162)	(0.141)
GPA 4	3.552	0.321*	0.309*	0.213	0.230	0.102	0.087
	(0.498)	(0.182)	(0.178)	(0.163)	(0.158)	(0.148)	(0.128)
Persist 2	0.954	0.119	0.110	0.075	0.049	0.055	0.037
	(0.210)	(0.087)	(0.092)	(0.077)	(0.071)	(0.068)	(0.058)
Graduated in 4 years or less	0.683	0.182	0.136	-0.059	-0.059	-0.011	-0.046
	(0.466)	(0.191)	(0.203)	(0.172)	(0.159)	(0.151)	(0.130)
Graduated in 5 years or less	0.868	0.154	0.192	0.080	0.088	0.131	0.094
	(0.339)	(0.140)	(0.150)	(0.124)	(0.114)	(0.110)	(0.094)
Graduated in 6 years or less	0.888	0.127	0.154	0.086	0.090	0.109	0.070
	(0.316)	(0.131)	(0.140)	(0.116)	(0.108)	(0.103)	(0.089)

Note. + $p<0.1$, * $p<0.05$, ** $p<0.01$. Baseline estimates based on control group means and standard deviations for students one ACT point below threshold. Coefficients are treatment effects at various bandwidths (e.g., 1.5 ACT points within the threshold).

Table 6

Alternative Functional Forms

	(1) Linear	(2) Quadratic
GPA 1	0.241	0.055
	(0.202)	(0.249)
ACT	-0.001	0.067
	(0.035)	(0.101)
ACT*Greater than 28.5	0.032	0.020
	(0.035)	(0.129)
ACT squared		0.032
		(0.052)
ACT squared*Greater than 28.5		-0.066
		(0.067)
Persist 2	0.114	-0.062
	(0.084)	(0.099)
ACT	-0.013	0.049
	(0.015)	(0.041)
ACT*Greater than 28.5	-0.001	-0.006
	(0.014)	(0.053)
ACT squared		0.029
		(0.021)
ACT squared*Greater than 28.5		-0.063**
		(0.027)
Grad 4	0.232	0.034
	(0.188)	(0.226)
ACT	-0.031	0.089
	(0.033)	(0.092)
ACT*Greater than 28.5	0.004	-0.143
	(0.032)	(0.120)
ACT squared		0.061
		(0.048)
ACT squared*Greater than 28.5		-0.038
		(0.062)

Note. + $p<0.1$, * $p<0.05$, ** $p<0.01$. Coefficients presented at the bandwidth of 1.5.

Figures

Figure 1

Histogram of First ACT Composite Score

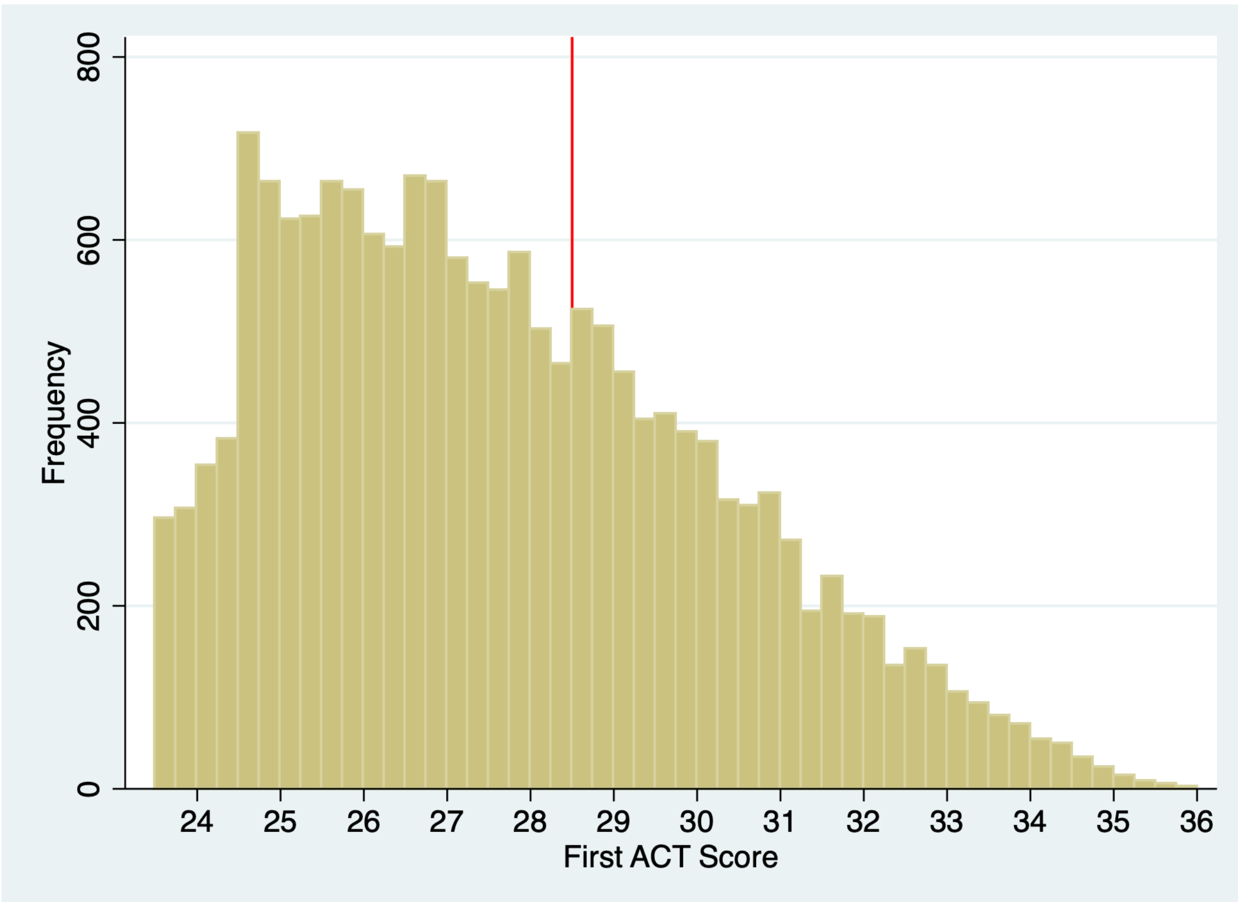


Figure 2

Histogram of Maximum ACT Composite Score

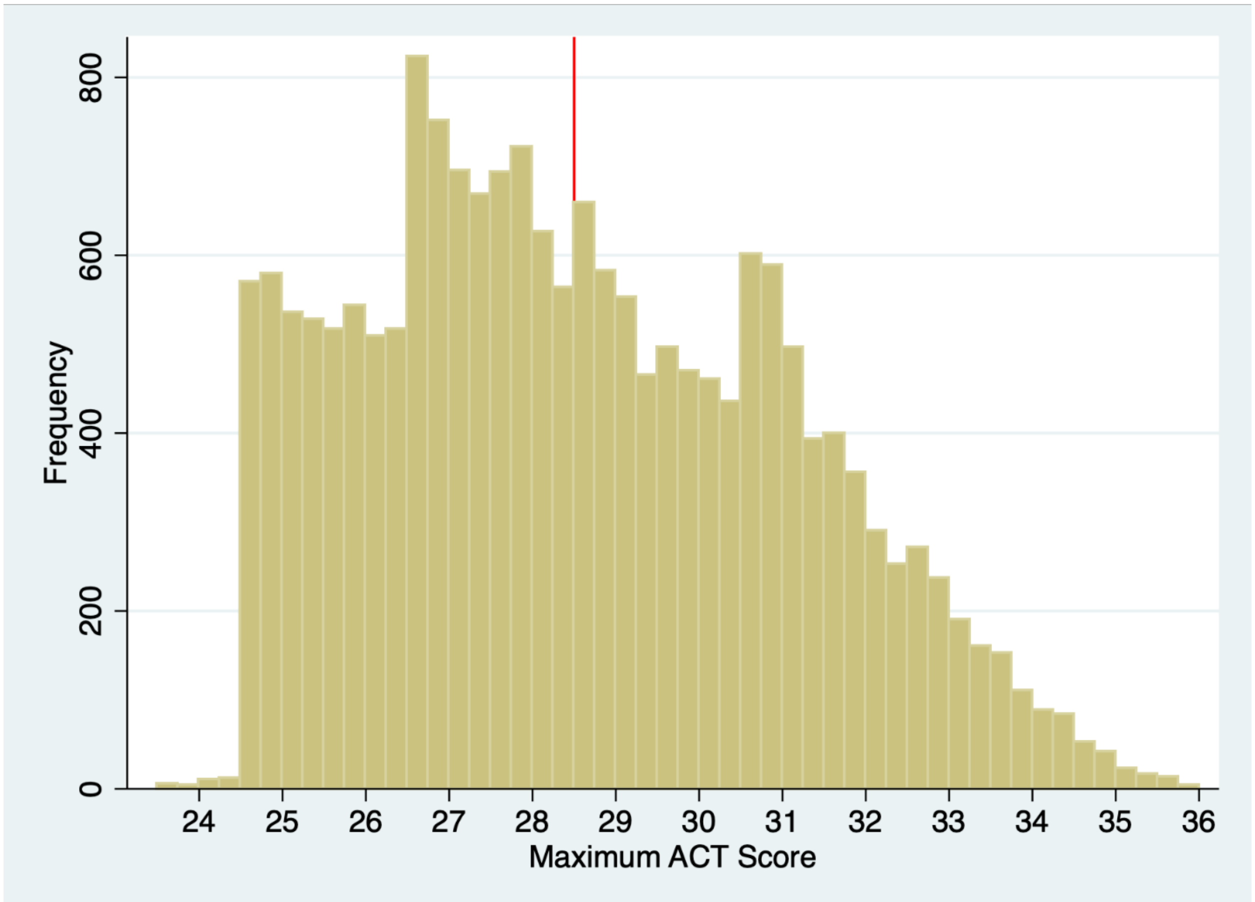


Figure 3

The Proportion of Honors Participants by First ACT Composite Score (First Stage)

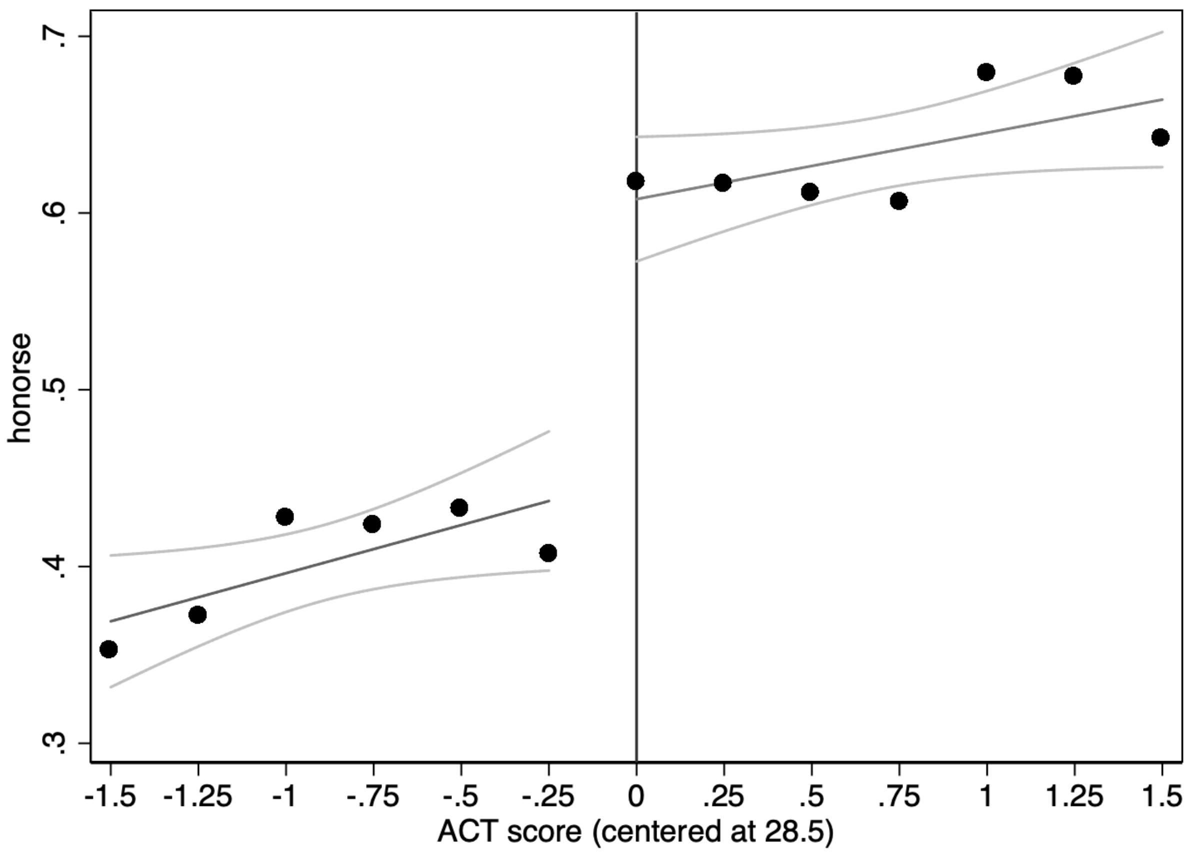


Figure 3

Fuzzy Regression Discontinuity Estimates (Linear)

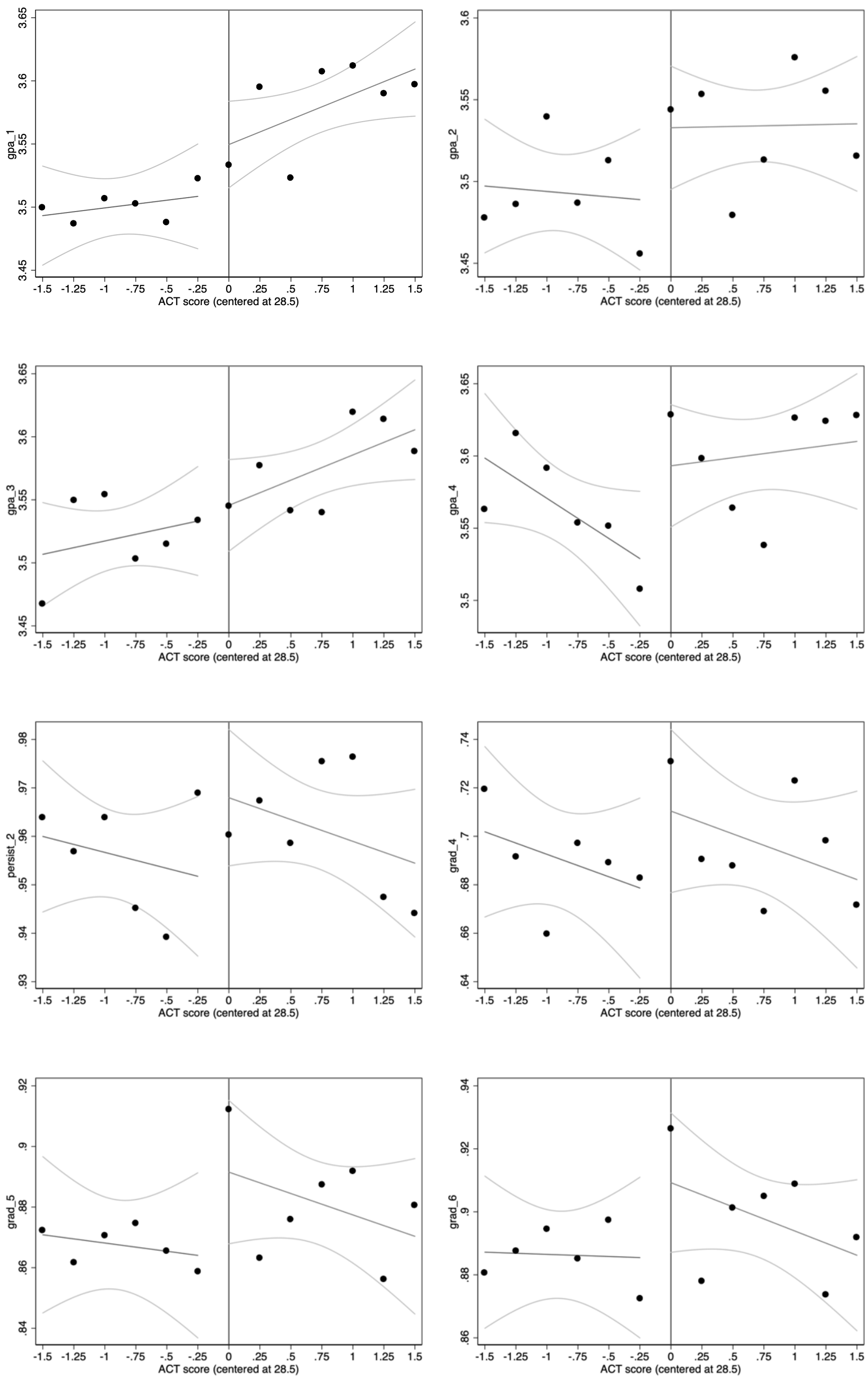
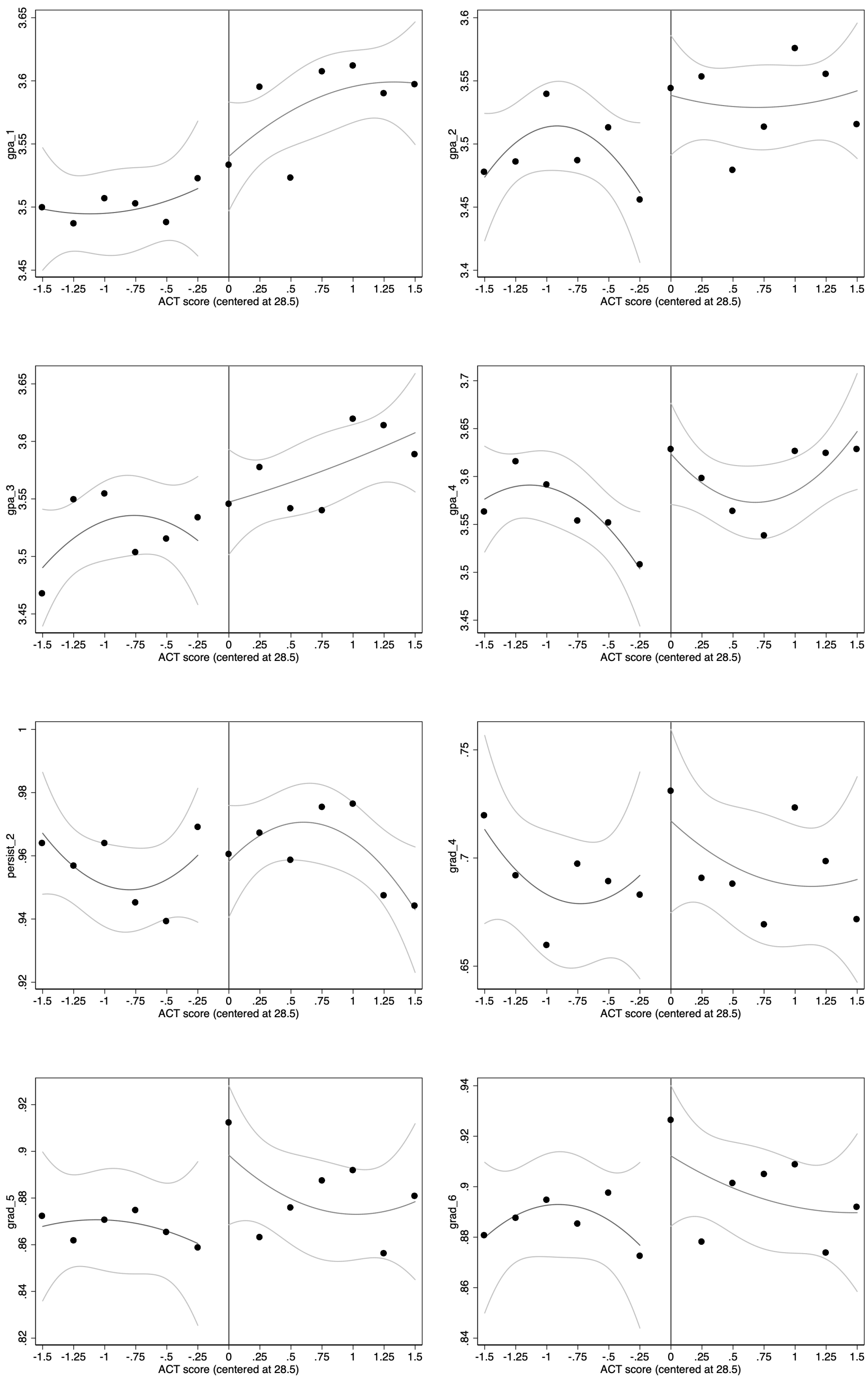


Figure 4

Fuzzy Regression Discontinuity Estimates (Quadratic)



CHAPTER FOUR: EDUCATION-JOB MISMATCH AMONG PH.D. HOLDERS: AN INSTITUTIONAL STUDY

Today's Ph.D. graduates are much less likely to enter the professoriate. According to the 2020 Survey on Earned Doctorates (SED), only 40% of recent Ph.D. graduates reported employment commitments in the academic sector[1] (National Science Foundation, 2021). In contrast, approximately 40% of graduates reported commitments in industry or business, 7.7% in government, 6.0% in nonprofit, and 6.6% in other sectors (National Science Foundation, 2021).

Over the past two decades, there has been mounting pressure for universities to adapt doctoral education to the changing demands of the labor market (Cardoso et al., 2022). Organizations such as the Council of Graduate Schools and the American Association of Universities have launched initiatives to improve data collection around career outcomes to better understand the diverse career pathways of Ph.D. students (Allum et al., 2014). However, not much is known about career outcomes beyond academia. Most research on doctoral career outcomes has focused on academic employment (e.g., Austin, 2002; Curtin et al., 2016). Many studies are also discipline-specific (e.g., Gibbs et al., 2014; Perley, 2019; Siegfried & Stock, 1999) or limited to STEM fields (e.g., Bender & Heywood, 2009; Conti & Visentin, 2015; Shauman, 2017).

From a policy standpoint, one area of particular interest is whether Ph.D. holders are well-matched for their chosen occupations. An educational-job mismatch is defined broadly as the extent to which workers' skills or education levels are above, below, or not related to their current job requirements (Mahuteau et al., 2014). It is well-established that not only do job

[1] Excludes postdoctoral appointments.

mismatches lead to lower wages, but they also reduce worker satisfaction and productivity (see Hartog, 2000 for literature review). Given the substantial time and resources invested in a Ph.D., including public funding, any labor market inefficiencies generated by mismatches are much higher for Ph.D. recipients than for other educational levels (Di Paolo & Mañé, 2016).

Drawing on cross-sectional data for Ph.D. graduates from the University of Missouri (MU) from 2011 to 2020, this paper analyzes the extent to which education-job mismatches vary by field of study and the potential consequences of mismatch on earnings. The analysis is motivated by the following research questions. First, to what extent does each type of mismatch (i.e., vertical, horizontal, vertical and horizontal) vary by degree field? Second, to what extent is each type of mismatch associated with a wage penalty, and how does this vary by degree field?

While education-job mismatches can take on various forms, I focus on whether Ph.D. graduates work in occupations for which they are over-educated (i.e., vertical mismatch) or outside of their field of study (i.e., horizontal mismatch). The study uses a novel dataset from Academic Analytics, a business intelligence firm that uses standardized processes to comprehensively collect alumni outcomes data by accessing publicly available data online. Where previous studies on education-job mismatch have relied on vertical and horizontal measures based on worker-self-assessments, the present analysis uses objective measures based on Standard Occupational Classification (SOC) codes. To the best of my knowledge, the current study is the first to address the consequences of vertical and horizontal mismatch among doctoral degree holders within the U.S. The study also contributes to a better understanding of the diverse career pathways of Ph.D. graduates and provides an approach that can be replicated at other universities.

Relevant Literature

While education-job mismatch among university graduates has been extensively researched, the studies on mismatch among Ph.D. holders are quite limited. In the following literature review, I describe some common approaches used to measure vertical and horizontal mismatch and some findings from studies that estimate the wage penalties of mismatch. Next, I discuss the empirical research concerning Ph.D. holders, including some of the limitations. To the best of my knowledge, this is the first study to analyze both types of mismatches across disciplines and estimate the potential wage effects of mismatch. Finally, I outline some theoretical assumptions that motivate and guide the current study.

Approaches to Measuring Education-Job Mismatch

Vertical Mismatch

Vertical mismatch refers to a situation in which a worker possesses a level of education above or below that required for their job and is often measured in terms of over-education, under-education, over-skilling, and under-skilling (McGuinness et al., 2018). The literature identifies three approaches typically used to measure vertical mismatch: the subjective method based on worker self-assessments, the empirical method based on the mean level of education within a given occupation, and the job evaluation method based on assessments by professional job analysts (Leuven & Oosterbeek, 2011).

With respect to the existing literature, much of the empirical research on education-job mismatch is devoted to over-education and measuring its effect on wages (McGuinness et al., 2018). Typically, these studies construct an indicator variable for mismatch and incorporate it into the right-hand sign of the Mincerian earnings function to estimate the effect of over-education on mismatched workers relative to their adequately matched counterparts. In their comprehensive review of literature, Hartog (2000) estimates the returns to overeducation are about half to two-thirds of the returns to required education. Vertical mismatch has also been linked to other non-fiduciary consequences, such as worker dissatisfaction (Tsang et al., 1991), higher turnover (Mcgoldrick & Robst, 1996), and reduced worker productivity (Tsang, 1987).

Horizontal Mismatch

In contrast, horizontal mismatch refers to a situation in which the worker is employed in a job unrelated to their field of study (McGuinness, 2006; Salas-Velasco, 2021; Verhaest et al., 2017). Horizontal mismatch is often measured using subjective questions asking the respondent to assess the degree to which their current job is related to the study field of their highest qualification (Robst, 2007; Verhaest et al., 2017) but is sometimes measured independently by comparing a field of study variable with occupation codes (Sellami et al., 2017).

More recently, horizontal mismatch has received increasing attention (see Somers et al., 2019 for a literature review). As with the literature on vertical mismatch, studies that estimate the impact of horizontal mismatch on earnings typically incorporate an indicator for mismatch (in some cases, an ordinal variable for perfectly matched, weakly matched or mismatched) into the right-hand side of the Mincerian earnings equation (Hadavand et al., 2019). Several studies have found evidence of a wage penalty for horizontally mismatched individuals (Bender & Roche, 2013; Nordin et al., 2010; Robst, 2007a; Robst & Vangilder, 2016).

There is also some evidence that horizontal mismatch is determined by the extent to which workers possess general skills as opposed to occupation-specific skills. For instance, Robst (2007) found that university graduates with degrees that emphasized general skills (e.g., English, social sciences, liberal studies) had a higher likelihood of horizontal mismatch in contrast to degrees that emphasized occupation-specific skills (e.g., computer sciences, health professions, engineering). However, the wage penalties associated with mismatch were greatest for degree fields emphasizing occupation specifically. In their study conducted in Sweden, Nordin et al. (2010) also found that horizontal mismatch was associated with larger wage penalties where degree fields emphasize more specialized skills.

Education-Job Mismatch for Ph.D. Graduates

A few studies examine education-job mismatch among Ph.D. holders specifically (Di Paolo & Mañé, 2016; Gaeta, 2015a; Gaeta et al., 2022; Park et al., 2018; Shauman, 2017). Several studies find that Ph.D. holders that are over-educated earn significantly less than their adequately matched counterparts. For instance, Park et al. (2018) examined over-education among doctorate holders using the Korean survey of Careers and Mobility of Doctorate Holders. Park et al. estimate that about 44% of graduates reported being over-educated in their current position, with a significant wage penalty of 6.5% relative to their well-matched counterparts.

Other studies find significant wage penalties only when including a measure of over-skilling, which is defined in terms of whether the worker has skills they possess that are not necessary to their current position. Using a 2011 survey on the early labor market experiences of Ph.D. holders in Spain, Di Paolo and Mañé (2016) found no statistically significant wage penalty associated with being over-educated or over-skilled independently. However, Ph.D. holders who were over-educated and over-skilled faced a significant wage penalty of 12% compared to their adequately matched counterparts.

Bender and Heywood (2009, 2011) are the only studies that address horizontal mismatch among doctoral holders. Using data from the 1997 and 1997 Survey of Doctoral Recipients (SDR), a nationally representative longitudinal survey of Ph.D. students in the hard and social sciences, they find that the prevalence and wage penalties of mismatch differed significantly by sector. For instance, while Ph.D. holders in academia were less likely to report any mismatch, they experienced a more significant wage penalty (13.8%) due to mismatch than that experienced by nonacademic workers (9.8%). Bender & Heywood (2011) find worse effects for those with Ph.D. graduates in the hard sciences, followed by the social sciences.

For the most part, the extant literature on education-job mismatch for Ph.D. holders finds significant wage penalties associated with vertical and horizontal mismatch. There is also some evidence that prevalence and wage penalties may vary by field (Bender & Heywood, 2009, 2011). However, these studies have been limited to independently analyzing vertical or horizontal mismatches. It may be that the effect of over-education interacts with the impact of field-of-study mismatch. Furthermore, the studies within the context of the US have been limited to STEM fields due to their reliance on the Survey of Doctoral Recipients.

Note that Shauman (2017) addresses both vertical and horizontal mismatch for Ph.D. holders within the scope of one study. They also use the same horizontal mismatch measure used by Bender & Heywood (2009, 2011) from the Survey of Doctoral Recipients. Still, their analysis is limited to the early careers of STEM Ph.D. graduates.

Another potential limitation of the literature is the reliance on purely subjective measures of vertical and horizontal mismatch, which may be biased and introduce measurement error (Leuven and Oosterbeek 2011). Studies rely on surveys that ask respondents to assess the level of education required for the job (e.g., Park et al., 2018) or whether the job is related to their degree (e.g., Bender & Heywood, 2009, 2011). However, subjective approaches may be biased in the case of over-educated workers who may be too apathetic to respond to surveys or reluctant to admit to mismatch (McGuinness, 2006).

Theoretical Assumptions

A few assumptions guide this paper that are consistent with a human capital framework. I assume that individuals choose a particular level of education and a field of study with an expectation of working in a profession related to the skills acquired (Nordin et al., 2010; Robst, 2007a) and that the decision includes some knowledge of potential earnings associated with that profession (Betts, 1996). I also assume that individuals acquire different skills during their doctoral studies—some of which are general and others that are occupation-specific.

Education-job mismatches arise, in part, due to supply and demand dynamics. Human capital theory explains this as a temporary mismatch between workers' skills and the firm's technology (Becker, 1964). For example, an oversupply of skilled workers may force jobseekers to temporarily accept jobs below their level of education and/or outside their field of study (Wolbers, 2003; Verhaest et al., 2017). Employers may also become more discriminating and hire workers that are over-educated or within the correct field of study (Verhaest et al., 2017). If individuals do not obtain a position that matches their level of education or the content of their field of study, then it is considered an undesirable outcome for both the individual and society as educational resources are wasted. This represents an inefficient allocation of human capital.

Data

The setting of the current study is the University of Missouri (MU), which is a public land-grant university that is one of 65 member institutions of the American Association of Universities (AAU) and ranks in the top 50 institutions in terms of the number of doctorate degrees awarded each year (National Center for Science and Engineering Statistics, 2019). MU has an enrollment of 30,000+ undergraduate and graduate students, with approximately 300 Ph.D. degrees awarded annually across 60 doctoral programs.

The study uses a cross-sectional snapshot of Ph.D. recipients who graduated from MU over a 10-year window from 2011 to 2020. The dataset is constructed from multiple data sources. First, employment data was compiled from Academic Analytics, a company previously known for faculty analytics that recently expanded its services into tracking alumni data. For a fee, AcA uses a team of researchers to data mine publicly accessible internet sources (e.g., LinkedIn, employer websites) and compile employment information for each graduate, such as employer, position, location, and estimated median earnings based on occupation and industry.

One of the advantages of this dataset is that each job record found is assigned a Standard Occupational Classification (SOC) code. According to the U.S. Bureau of Labor Statistics, SOC codes are used by federal agencies to classify workers into occupational categories to collect, calculate, or disseminate data. Occupations are grouped in each SOC based on job duties and, in some cases, based on skills, education, and training.

The 2018 SOC taxonomy includes 23 major group codes ending with 0000 (e.g., 25-0000 Educational Instruction and Library Occupations), 98 minor group codes (e.g., 25-1000 Postsecondary Teachers), 459 broad group codes ending with 0 (e.g., 25-1060 Social Sciences Teachers, Postsecondary), and 867 detailed group codes ending in a number other than 0 (e.g., 25-1063 Economics Teachers, Postsecondary). Academic Analytics provides SOC codes at the major group level and the broad group level. I use the broad group-level SOC code, which is the more detailed of the two.

The study also relies on administrative data provided by MU for information on academic background characteristics and demographics. Academic background variables include degree field (including 2020 Classification of Instructional Programs (CIP) codes), graduation year, and ability controls, such as cumulative GPA and time-to-degree (TTD). Socio-demographic variables include indicator variables for sex, race/ethnicity, and citizenship, as well as age and a quadratic term for age. Since the analysis focuses on current employment outcomes, I drop anyone above retirement age (above 65 years old). Note that for comparison purposes, I have grouped sixty (60) degree fields into eight (8) broader degree fields used by the 2020 Survey of Earned Doctorates (see Table 1 for a complete account of which University degrees fall under each general category).

Occupation-Level Measures of Mismatch

Vertical Mismatch Measure

Vertical mismatch measures the extent to which workers possess a level of education above or below that required for their job (McGuinness et al., 2018). For the present analysis, I am only concerned with over-education. I use data from the Occupational Information Network (O*NET) 18.0 database maintained by the US Department of Labor/Employment and Training Administration to determine education level. The O*NET database reports the required level of education for each SOC code using incumbent surveys and analyses by occupational experts. I determined the percentage of O*NET respondents who specified that a doctoral or post-doctoral degree is required for their job performance. The vertical mismatch variable was constructed based on the percentage of respondents for whom a doctoral degree is not required for their current occupation. An individual was coded as a vertical mismatch (VERTICAL=1) if they were in an occupation with an unrelated CIP code. O*NET uses SOC codes at the detailed level (e.g., 25-1063.00 Economics Teachers, Postsecondary), so I aggregate the results at the broad group level (e.g., 25-1060 Social Sciences Teachers, Postsecondary).

Horizontal Mismatch Measure

Horizontal mismatch measures the extent to which a worker's field of study is related to the content of their job (McGuinness, 2006; Salas-Velasco, 2021; Verhaest et al., 2017). To determine which occupations are within each degree field, I utilize the 2020 CIP-to-SOC crosswalk provided by the National Center of Educational Statistics (NCES), created in partnership with the Bureau of Labor Statistics (BLS). The NCES CIP-to-SOC crosswalk matches 6-digit 2020 CIP codes with 6-digit SOC codes based on their descriptions. The idea is that the academic program associated with each CIP code needs to provide the skills and knowledge required to perform the occupation related to each SOC. I consider the occupation to be within the degree field if the CIP code provided by NCES matches the CIP code of the field-of-study provided by the University. Thus, an individual is coded as a horizontal mismatch (HORIZONTAL=1) if they are in an occupation with an unrelated CIP code. Note that the NCES CIP-SOC crosswalk uses SOC codes at the detailed group level (e.g., 25-1063 Economics Teachers, Postsecondary), which I recode to their corresponding broad level SOC code (e.g., 25-1060 Social Sciences Teachers, Postsecondary).

To facilitate the interpretation of the results, Figure 1 provides some examples of an education-job mismatch (i.e., vertical, horizontal, and both) for the reference category. For Mathematics and Computer Sciences, an Assistant Professor of Mathematics is an example of an occupation that is both a vertical and horizontal match. In contrast, a Software Engineer is an example of a vertical mismatch, and Assistant Professor in a different field is a horizontal mismatch, and a Chief Executive is simultaneously a vertical and horizontal mismatch.

Median Earnings by SOC Code and Industry

AcA provides median annual earnings for each record based on the broad-level SOC code and the North American Industry Classification System (NAICS) code. NAICS is the standard system used by Federal statistical agencies to classify business establishments to collect, analyze, and publish statistical data related to the U.S. business economy. If the SOC code is this accepted classification for occupation, then the NAICS code might be considered the equivalent classification for the business sector or industry. The wage data for each combination of SOC and NAICS codes was sourced by Academic Analytics from the U.S. Bureau of Labor Statistics. In the current analysis, I use the natural log of wages to reduce positively skewed data.

Other Adjustments

A few adjustments were made manually to the dataset given the reliance on broad-level SOC codes and SOC codes in general, as well as the use of the O*NET and NCES CIP-to-SOC crosswalks.

First, the broad-level SOC code 11-9030 for Education and Childcare Administrators includes administrators in higher education (e.g., Dean, Chair) and administrators in other educational sectors (e.g., daycare, elementary school) with very different educational requirements. In the case of administrators employed in higher education, I use the information on their secondary appointments (e.g., professor, instructor) to manually code vertical and horizontal mismatch variables for each record individually. Out of 156 observations, I identified 24 whose secondary position would not have required a doctorate (e.g., lecturer, part-time adjunct faculty) and 2 whose secondary position was in a different field compared to their field of study.

Second, there is not currently a SOC code encompassing post-doctoral fellow (postdoc) positions, as these are sometimes viewed as continuing education opportunities rather than employment. I manually coded all postdocs as vertical matches since a doctoral degree is a known job requirement. However, I visually inspected the department for each postdoc to determine if their current appointment could be considered within the degree field. Out of 209 postdocs, I determined all to be horizontally matched or within the degree field. The lack of SOC code also meant that I did not have access to salary data as I did with other SOC codes, so I relied on median salary by broad field from the 2020 Survey of Earned Doctorates. For instance, Life Sciences graduates employed in academia as a postdoc in 2020 earned a median salary of $50,000. Since all postdocs were vertically and horizontally matched, this approach created no issues in measuring salary disparities between matched and mismatched postdocs within field.

Third, the O*NET database may under-report the level of education required for some postsecondary teaching positions. Out of 11 broad group SOC codes pertaining to postsecondary teachers, I found that 3 (Physical Sciences, Math and Computer Science, and Health) fell just below the 50% threshold used to determine whether an occupation is doctoral level. As it is commonly accepted that a Ph.D. is a requirement for teaching at the postsecondary level and each of these SOC codes falls within the margin of error (i.e., within 10 points), I recoded any records falling into these specific SOC codes as doctoral level. These changes impacted 226 observations.

Finally, there were a few limitations to using the CIP-to-SOC crosswalk provided by NCES, which I used to determine whether a graduate's occupation was within the same field of as their field of study. First, the crosswalk did not fully account for postsecondary teaching options within field for some CIP codes (for example, postsecondary teaching in education for education graduates and postsecondary teaching in health for exercise physiology graduates). I adjusted seven CIP-to-SOC code combinations, resulting in 99 records becoming horizontal matches. Second, the crosswalk only provided one occupation (SOC 21-1012 for Guidance Counselors), so I expanded the options to include psychologists (SOC 19-3030) and counselors (SOC 21-2010), which resulted in 23 records being changed to horizontal matches.

Methods

The paper aims to answer several questions related to education-job mismatch for Ph.D. Specifically, to what extent does each type of mismatch vary by degree field? Second, to what extent is each type of mismatch associated with a wage penalty, and how does this vary by degree field? The choice of degree field is likely to be correlated with unobserved factors, such as ability, that are also correlated with education-job alignment (Leuven and Oosterbeek, 2011) and or earnings (Carenvale, et. al., 2011). Therefore, the goal of the current paper is to produce a rigorous descriptive analysis that documents some central tendencies by degree field with respect to the skills and earnings potential of the jobs Ph.D. graduates ultimately obtain. According to Loeb et al. (2017), despite the popularity of causal methods, descriptive analysis can stand on its own as research when it identifies phenomena or patterns in data that have not previously been recognized using low-inference, low-assumption methods that use no or minimal statistical adjustments (Loeb et al., 2017).

Probability of Education-Job Mismatch by Degree Field

First, I measure the extent to which the probability of mismatch varies by degree field using a binomial probit model. I estimate the following:

$$\Pr\,(Mismatch)_{ijc} = f(X_{ijc}\beta + Z_j\alpha + \varepsilon_{ijc}), \qquad (1)$$

where mismatch is equal to one if an individual is "mismatched" according to the various mismatch measures for individual (i) in degree field (j) for graduation cohort (c); X_{ijc} is a vector of socio-demographic variables (age and its quadratic, sex, race/ethnicity, and citizenship) and academic background variables (cumulative GPA and time-to-degree); and Z_j denotes degree field. In addition, standard errors are clustered by degree field. Finally, the coefficient of interest is α. I report the effect sizes of coefficients as average marginal effects, which indicate the change in the probability of being mismatched associated with a one-unit change in the respective independent variable, holding all other independent variables constant.

Wage Effects of Education-Job Mismatch by Degree Field

Next, I examine how wage effects from being mismatched vary across degree fields. As such, I interact an indicator variable for mismatch with each degree field. While I do not have individual salary information, Academic Analytics provides the median wages of each graduate's occupation based on the SOC and NAICS codes. I utilize a degree field fixed effects model to compare "mismatched" graduates with "matched" graduates within the same degree field. I estimate the following, which is a modified version of the Mincerian equation:

$$\text{LnW}_{ijc} = X_{ijc}\beta + Z_j\alpha + \delta_j(Z_j * Mismatch_{ij}) + \varepsilon_{ijc}, \qquad (2)$$

where W_{ijc} is the natural logarithm of occupation-level wages for individual (i) in degree field (j) for graduation cohort (c), X_{ijc} is a vector of socio-demographic and academic background characteristics, Z_j captures degree field fixed effects, and $Mismatch_{ij}$ is an indicator variable equal to 1 if the graduate is mismatched and 0 otherwise. Standard errors are clustered by degree field.

The coefficient of interest is δ_j, which captures the extent to which the effect of mismatch on the log of occupational-level wages varies by degree field. Since the dependent variable is the natural log of occupation-level earnings, I present exponentially transformed coefficients to simplify the interpretation of results; the coefficients provide the percent increase (or decrease) in the response for every one-unit increase in the independent variable.

Results

Descriptive Summary

Out of the 3,153 Ph.D. degrees awarded from 2011 to 2020, Academic Analytics found current employment information for approximately 77% of the population, excluding individuals of retirement age and records that were not assigned a SOC code (n=119). Table 2 presents a descriptive summary of academic and demographic variables for Ph.D. holders for whom an outcome was found compared with the total population of Ph.D. graduates. In both the sample and the population, graduates of life sciences comprise the largest proportion, with almost a quarter of observations, followed by education at about 15%. Engineering graduates are slightly underrepresented in the analytics sample at 11.2% relative to the population at 13.3%.

The academic background and socio-demographic characteristics of the sample and population are also similar. The average cumulative GPA for both is 3.8. The time-to-degree for the sample is slightly faster for the analytical sample (5.4 years) relative to the population (5.5 years). With regard to demographics, the sample is 46.7% female, 52.1% white, and 33.8% non-residential/international students. White individuals are slightly overrepresented in the analytical sample, while non-residential/international students are slightly underrepresented. This is not unexpected given that career outcomes may be more difficult to identify online for non-resident/international students that return to their home countries after graduation. Finally, the average age of the sample and the population is about 39 years old.

Probability of Education-Job Mismatch by Degree Field

The first research question addressed is whether the probability of being mismatched for each type of mismatch varies by degree field. Figure 2 shows the percentage of graduates within each degree field that fall into each mismatch category. Overall, the vertical mismatch rate was 43.9%, the horizontal mismatch rate was 43.3%, and rate of both simultaneously was 33.4%. Across degree fields, graduates of engineering had the highest likelihood for each mismatch measure, with 65.3% of graduates being vertically mismatched, 61.9% being horizontally mismatched, and 52.4% being both. Further, education and physical sciences and earth sciences had vertical mismatch rates above 50%. For the reference group of mathematics and computer sciences, graduates were much less likely to be horizontally mismatched (29.2%) than vertically mismatched (49.4%)

Table 4 presents the preferred specification of the binomial probit model, which estimates the likelihood of mismatch by degree field controlling for academic background and sociodemographic characteristics. Marginal effects to assist with interpreting the probit coefficients. The marginal effects coefficients can be interpreted as the change in the probability of mismatch associated with a one-unit change in the broad degree (i.e., from 0 to 1) relative to the reference group, holding all other independent variables constant. In the current study, the field of mathematics and computer sciences serve as the reference group.

Probability of Vertical Mismatch by Degree Field

The first two columns present the probit coefficients and average marginal effects of vertical mismatch, respectively. Relative to mathematics and computer sciences, education, engineering, and physical sciences and earth sciences graduates are more likely to be vertically mismatched after controlling for academic and socio-demographic characteristics. However, differences are only statistically significant for engineering graduates. Relative to math and computer science graduates, engineering graduates were 18.4 percentage points more likely to be vertically mismatched, holding all other variables constant.

Graduates of humanities and arts, life sciences, psychology and social sciences, and other (non-STEM) fields were all significantly less likely than the reference group to be vertically mismatched. Among these, humanities and arts graduates had the least probability of vertical mismatch. Relative to math and computer science graduates, humanities and arts graduates were 18.5 percentage points less likely to be vertically mismatched, holding all other variables constant.

Probability of Horizontal Mismatch by Degree Field

The third and fourth columns present the probit coefficients and average marginal effects of horizontal mismatch. Relative to mathematics and computer sciences, graduates of almost all fields were more likely to be horizontally mismatched, holding all other variables constant. Most fields were also significantly more likely to be mismatched (except for other fields). Among these, engineering graduates had the greatest probability of horizontal mismatch. Relative to math and computer science graduates, engineering graduates were 32.5 percentage points more likely to be horizontally mismatched, holding the other variables constant. In contrast, humanities and arts and other fields had the lowest relative likelihood of horizontal mismatch.

Probability of Vertical and Horizontal Mismatch by Degree Field

The fifth and final columns present the probit coefficients and average marginal effects of field on the likelihood of both vertical and horizontal mismatch. Most fields have a relatively higher probability of both being both vertically and horizontally mismatched than mathematics and computer sciences graduates. Again, graduates of engineering are the most likely to be vertically and horizontally mismatched. Relative to computer and math science graduates, engineering graduates are 26.0 percentage points more likely to be in a job in which they are over-educated and outside of their field of study. Only graduates of other (non-STEM) fields were relatively less likely to be dually mismatched.

Wage Effects of Education-Job Mismatch by Degree Field

Next, I examine the relationship between mismatch and occupation-level earnings by degree field. Table 5 presents the results of the second equation outlined above, which includes degree fixed effects.

Wage Effects of Vertical Mismatch by Degree Field

The first column presents the consequences of vertical mismatch on occupation-level earnings. Overall, I find that vertically mismatched graduates entered occupations for which the average median earnings were higher than their well-matched counterparts within the same field. For mathematical and computer sciences graduates, vertically mismatched workers entered occupations in which the average median wages were approximately 23.4% higher than those who entered occupations for which they were well-matched, holding other variables constant. Engineering graduates who were over-educated experienced the greatest wage premiums at 24.8%, while graduates of physical Sciences and earth sciences had the smallest wage premiums at about 5%.

Wage Effects of Horizontal Mismatch by Degree Field

The second column presents the consequences of horizontal mismatch on occupation-level earnings. Again, I find that horizontally mismatched graduates entered occupations for which the average median earnings were higher than those of their well-matched counterparts within the same field. For mathematical and computer sciences graduates, horizontally mismatched workers enter occupations in which the average median wages were approximately

18.3% higher than those who entered occupations for which they were well-matched, holding other variables constant. In contrast to vertical mismatch, life sciences graduates experienced the greatest premiums to working outside their field at 34.7%, followed by Psychology and Social Sciences at 33.3%. Physical sciences and earth sciences graduates had the lowest wage premiums at 10.7%.

Wage Effects of Vertical and Horizontal Mismatch by Degree Field

The first column presents the consequences of vertical and horizontal mismatch on occupation-level earnings. As with vertical and horizontal mismatch individually, I find that graduates who were both vertically and horizontally mismatched entered occupations for which the average median earnings were higher than those of their well-matched counterparts within the same field. However, the magnitude of the effect size is smaller than with vertical and horizontal mismatch individuals. For mathematical and computer sciences graduates, dually mismatched workers enter occupations in which the average median wages were approximately 15.2% higher than those who entered occupations for which they were well-matched, holding other variables constant. As with horizontal mismatch, Life sciences graduates experienced the greatest premiums to working outside of their field at 27.1%, followed by psychology and social sciences at 26.9%. The lowest wage premiums associated with being dually mismatched were found among the Physical Sciences and Earth Sciences at 11.5%.

Heckman Correction

One possible concern of data gathered through web scraping is the potential for sample selection bias. In particular, the employment outcomes were only observable for those Ph.D. holders who had a publicly accessible record of their employment online that Academic Analytics could identify. This raises concerns about potential bias as there may be unobservables correlated with being over-educated or outside of one's field *and* the propensity to have a LinkedIn profile. For instance, some degree fields may promote the use of LinkedIn, such as business, but not other degree fields, such as education.

To address the problem, I estimate a bivariate probit model incorporating a sample selection estimation procedure to obtain unbiased estimates (Heckman, 1979). Gaeta (2015) utilized a similar procedure in their analysis of vertical mismatch among Ph.D. holders in Italy to address endogeneity concerns related to being only able to observe vertical mismatch for those who were employed at the time of the survey; in particular, there may be unobservable individual factors that correlate with the probability of overeducation/overskilling and the propensity to get a job. Similarly, I am concerned that there may be unobservable individual factors that correctly with the probability mismatch and having a professional online presence. I use the following selection equation to account for the selection effect:

$$Y_i^{Select} = (Z_i\lambda + \varepsilon_{2i} > 0) \qquad (3)$$

where Y_i^{Select} takes the value of 1 if *i*-th individual was found and 0 otherwise, *Z* is a vector of covariates, λ is a vector of coefficients to be estimated, and ε_2 is the normally distributed error term. In addition, vector *Z* includes the same controls as used in the primary specification described above.

The Heckman selection estimation procedure requires the *Z* vector in the above equation to include covariates that can be legitimately excluded from the *X* in equation (1). This requires finding a variable that does not affect labor market mismatch but can affect the likelihood that an employment outcome is found. In this case, I instrument the popularity of surnames using the 2010 Census names database tabulating the most frequently occurring surnames. In this case, I assigned each surname their rank order according to the database if it was within the top 1,000 and 0 otherwise. The idea is that more popular surnames may have a decreased likelihood of being found because it may be difficult to distinguish between records of the same name. Surnames less common in the United States, such as those not of English language origin, may be less likely to be ranked by the Census or appear on LinkedIn if the graduate is a non-resident and returns to their home country.

Table 5 presents the probit coefficients for the main specification for vertical mismatch results only, with and without standard errors, followed by the model with the Heckman selection estimation. The results appear to be robust to alternative specifications. Moving from left to right, there is no change in the direction of the coefficients. Except for Life Sciences, there is also no change in the significance. The Heckman correction shows slight changes in magnitude, but all results remain statistically significant. These results have been plotted in Figure 3 with 95% confidence intervals. The figures are almost identical, which speaks further to the robustness of the results across various specifications.

Conclusion

In the U.S., the research on education-job mismatch among Ph.D. holders has been limited to STEM fields and relies on purely subjective measures of vertical and horizontal mismatch. The current study took a different approach by using SOC codes to construct occupation-level measures of mismatch. First, I find that the probability of each type of mismatch varies significantly by degree field. Across all fields, graduates of Engineering have the highest likelihood of each type of mismatch, holding all other variables constant. In other words, graduates of Engineering are the most likely to currently have a job in which they are over-educated and is unrelated to their field of study. I also find that mismatch is associated with a wage premium in all fields, ranging from about 5% to 25% for vertical mismatch and 11% to

34% for horizontal mismatch. While these estimates capture occupation-level earnings only (not individual wages), the findings suggest that mismatched Ph.D. graduates are entering occupations for which the earnings potential on average is higher relative to their adequately matched counterparts.

The findings are somewhat unexpected given results on university graduates that find that the probability of mismatch is lower for fields that empathize occupation-specific skills as opposed to general skills (e.g., Robst, 2007b). For instance, Robst, (2007b) found that university graduates of Engineering had the lowest incidence of horizontal mismatch and the highest wage premiums. In contrast, it may be that doctoral-level education includes a combination of more general skills (e.g., critical thinking, communication) that allows graduates to enter occupations outside their field of study without much consequence. Further, there may be a high demand for advanced research skills in nonacademic settings. These findings have implications for those looking to improve doctoral education through better understanding of diverse career pathways.

Some of the results might be interpreted with caution, however. While I use more or less objective approaches to construct mismatch measures, there is likely to be measurement error associated with using occupation-level measures. While subjective measures may be prone to bias, occupation-level mismatch measures may also have measurement error, if systematic, could produce bias. Future research might benefit from using more detailed SOC codes or occupational crosswalks that better account for the unique requirements of academic jobs.

References

Allum, J. R., Kent, J. D., & McCarthy, M. T. (2014). *Understanding PhD Career Pathways for Program Improvement: A CGS Report*. Washington, DC: Council of Graduate Schools.

Bender, K. A., & Heywood, J. S. (2009). Educational Mismatch among Ph.D.s. In R. Freeman & D. Goroff (Eds.), *Science and Engineering Careers in the United States* (pp. 229–256). University of Chicago Press. https://doi.org/10.7208/chicago/9780226261904.003.0008

Bender, K. A., & Heywood, J. S. (2011). Educational mismatch and the careers of scientists. *Education Economics*, *19*(3), 253–274.

Bender, K. A., & Roche, K. (2013). Educational mismatch and self-employment. *Economics of Education Review*, *34*, 85–95. https://doi.org/10.1016/J.ECONEDUREV.2013.01.010

Conti, A., & Visentin, F. (2015). Science and engineering Ph.D. students' career outcomes, by gender. *PLOS ONE*, *10*(8), e0133177. https://doi.org/10.1371/JOURNAL.PONE.0133177

Di Paolo, A., & Mañé, F. (2016). Misusing our talent? Overeducation, overskilling and skill underutilisation among Spanish PhD graduates. *The Economic and Labour Relations Review*, *27*(4), 432–452. https://doi.org/10.1177/1035304616657479

Gaeta, G. L. (2015). Was it worth it? An empirical analysis of over-education among PhD recipients in Italy. *International Journal of Social Economics*, *42*(3), 222–238. https://doi.org/10.1108/IJSE-08-2013-0186/FULL/XML

Gaeta, G. L., Lavadera, G. L., & Pastore, F. (2022). The effect of job–education vertical mismatch on wages among recent PhD Graduates: Evidence from an instrumental variable analysis. *Italian Economic Journal*, *8*(1), 197–225. https://doi.org/10.1007/s40797-021-00167-y

Gibbs, K. D., McGready, J., Bennett, J. C., & Griffin, K. (2014). Biomedical science Ph.D. career interest patterns by race/ethnicity and gender. *PLOS ONE*, *9*(12), e114736. https://doi.org/10.1371/JOURNAL.PONE.0114736

Hartog, J. (2000). Over-education and earnings: where are we, where should we go? *Economics of Education Review*, *19*(2), 131–147. https://doi.org/10.1016/S0272-7757(99)00050-3

Heckman, J. J. (1979). *Sample Selection Bias as a Specification Error*. *47*(1), 153–161.

Mahuteau, S., Mavromaras, K., Sloane, P., & Wei, Z. (2014). Horizontal and vertical educational Mismatch and Wages. *25th Australian Labour Market Research*.

Mcgoldrick, K., & Robst, J. (1996). Gender differences in overeducation: A Test of the Theory of differential overqualification. *The American Economic Review*, *86*(2), 280–284.

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, *20*(3), 387–418. https://doi.org/10.1111/J.0950-0804.2006.00284.X

McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, *32*(4), 985–1015. https://doi.org/10.1111/joes.12254

National Science Foundation. (2021). *Doctorate Recipients from U.S. Universities: 2020: Vol. NSF 16-300* (Issue December). https://ncses.nsf.gov/pubs/nsf22300/report

Nordin, M., Persson, I., & Rooth, D. O. (2010). Education-occupation mismatch: Is there an income penalty? *Economics of Education Review*, *29*(6), 1047–1059. https://doi.org/10.1016/j.econedurev.2010.05.005

Park, K., Jang, D., & Shahiri, H. I. (2018). Over-education among doctorate holders in the Korean labor market. *Asia-Pacific Social Science Review*, *18*(1), 32–45.

Perley, D. A. (2019). Gender and the career outcomes of Ph.D. astronomers in the United States. *Publications of the Astronomical Society of the Pacific*, *131*(1005), 114502. https://doi.org/10.1088/1538-3873/AB0CC4

Robst, J. (2007). Education and job match: The relatedness of college major and work. *Economics of Education Review*, *26*(4), 397–407. https://doi.org/10.1016/j.econedurev.2006.08.003

Robst, J., & Vangilder, J. (2016). Salary and job satisfaction among economics and business graduates: The effect of match between degree field and job. *International Review of Economics Education*, *21*, 30–40. https://doi.org/10.1016/J.IREE.2015.11.001

Salas-Velasco, M. (2021). Mapping the (mis)match of university degrees in the graduate labor market. *J Labour Market Res*, *55*, 14. https://doi.org/10.1186/s12651-021-00297-x

Sellami, S., Verhaest, D., Nonneman, W., & van Trier, W. (2017). The impact of educational mismatches on wages: The influence of measurement error and unobserved heterogeneity. *B.E. Journal of Economic Analysis and Policy*, *17*(1), 1–38. https://doi.org/10.1515/bejeap-2016-0055

Shauman, K. (2017). Gender differences in the early employment outcomes of STEM doctorates. *Social Sciences*, *6*(1), 24. https://doi.org/10.3390/socsci6010024

Siegfried, J. J., & Stock, W. A. (1999). The labor market for new Ph.D. economists. *Journal of Economic Perspectives*, *13*(3), 115–134. https://doi.org/10.1257/JEP.13.3.115

Somers, M. A., Cabus, S. J., Groot, W., & van den Brink, H. M. (2019). Horizontal mismatch between employment and field of education: Evidence from a systematic literature review. *Journal of Economic Surveys*, *33*(2), 567–603. https://doi.org/10.1111/JOES.12271

Tsang, M. C. (1987). The impact of underutilization of education on productivity: A case study of the U.S. Bell companies. *Economics of Education Review*, *6*(3), 239–254. https://doi.org/10.1016/0272-7757(87)90003-3

Verhaest, D., Sellami, S., & van der Velden, R. (2017). Differences in horizontal and vertical mismatches across countries and fields of study. *International Labour Review*, *156*(1). https://doi.org/10.1111/j.1564-913X.2015.00031.x

Tables

Table 1

Classification of Degree Fields by Broader Field

Broad Field	Field	Broad Field	Field
Education	Agricultural Education & Leadership	Mathematics and	Computer Science
	Curriculum & Instruction	Computer Sciences	Mathematics
	Educational Leadership & Policy Analysis		Statistics
	Educational School & Counseling Psychology	Other (Non-STEM)	Accountancy
	Information Science & Learning Technologies		Business Administration
	Learning Teaching & Curriculum		Communication
	Music Education		Food & Hospitality Systems
	Special Education		Human Environmental Sciences
Engineering	Biological Engineering		Journalism
	Chemical Engineering		Social Work
	Civil & Environmental Engineering	Physical Sciences and	Chemistry
	Electrical & Computer Engineering	Earth Sciences	Geological Sciences
	Industrial Engineering		Physics
	Mechanical & Aerospace Engineering	Psychology and Social	Anthropology
	Nuclear Engineering	Sciences	Economics
Humanities and Arts	Art History & Archaeology		Political Science
	English		Psychological Sciences
	History		Psychology
	Philosophy		Public Affairs
	Romance Languages & Literature		Rural Sociology
	Theatre		Sociology
Life Sciences	Agricultural & Applied Economics		
	Agronomy		
	Animal Sciences		
	Biochemistry		
	Biological Sciences		
	Biomedical Sciences		
	Nutrition & Exercise Physiology		
	Genetics Area Program		
	Informatics		
	Medical Pharmacology & Physiology		
	Molecular Microbiology & Immunology		
	Natural Resources		
	Neuroscience		
	Nursing		
	Pathobiology Area Program		
	Plant Insect Microbial Sciences		
	Soil, Environmental & Atmospheric Sciences		
	Speech, Language & Hearing Sciences		

Note. Degree fields are used by the University of Missouri and classified into broad degree fields used by the 2020 Survey of Earned Doctorates.

Table 2

Descriptive Summary of Independent and Control Variables

	Population				Analytical Sample			
	min	max	mean	sd	min	max	mean	sd
Broad Degree Field								
Education	0.000	1.000	0.153	0.291	0.000	1.000	0.148	0.355
Engineering	0.000	1.000	0.133	0.295	0.000	1.000	0.112	0.315
Humanities and Arts	0.000	1.000	0.091	0.307	0.000	1.000	0.093	0.291
Life Sciences	0.000	1.000	0.240	0.313	0.000	1.000	0.242	0.429
Mathematics and Computer Scien	0.000	1.000	0.066	0.309	0.000	1.000	0.069	0.254
Physical Sciences and Earth Scier	0.000	1.000	0.078	0.302	0.000	1.000	0.075	0.264
Psychology and Social Sciences	0.000	1.000	0.119	0.282	0.000	1.000	0.124	0.329
Other (non-S&E)	0.000	1.000	0.121	0.310	0.000	1.000	0.136	0.343
Cohort								
2011	0.000	1.000	0.096	0.295	0.000	1.000	0.102	0.302
2012	0.000	1.000	0.097	0.296	0.000	1.000	0.103	0.304
2013	0.000	1.000	0.094	0.291	0.000	1.000	0.095	0.293
2014	0.000	1.000	0.096	0.295	0.000	1.000	0.100	0.299
2015	0.000	1.000	0.105	0.307	0.000	1.000	0.109	0.311
2016	0.000	1.000	0.110	0.313	0.000	1.000	0.115	0.319
2017	0.000	1.000	0.107	0.309	0.000	1.000	0.104	0.306
2018	0.000	1.000	0.108	0.310	0.000	1.000	0.109	0.311
2019	0.000	1.000	0.101	0.302	0.000	1.000	0.097	0.296
2020	0.000	1.000	0.087	0.282	0.000	1.000	0.067	0.251
Academic Background Characteristics								
Cumulative GPA	3.029	4.000	3.822	0.189	3.029	4.000	3.826	0.187
Time to degree	0.000	23.880	5.515	1.974	0.000	23.270	5.426	1.866
Socio-Demographic Characteristics								
Female	0.000	1.000	0.472	0.499	0.000	1.000	0.467	0.499
White	0.000	1.000	0.498	0.500	0.000	1.000	0.521	0.500
Black	0.000	1.000	0.033	0.179	0.000	1.000	0.033	0.178
Asian or Pacific Islander	0.000	1.000	0.020	0.139	0.000	1.000	0.017	0.129
Hispanic	0.000	1.000	0.023	0.148	0.000	1.000	0.022	0.146
Multiracial/Other/Not Specified	0.000	1.000	0.069	0.254	0.000	1.000	0.069	0.254
Non-Resident/International	0.000	1.000	0.357	0.479	0.000	1.000	0.338	0.473
Age	27.000	78.000	39.660	7.594	27.000	65.000	39.313	6.897

Note. Descriptive summary tabulated from administrative data provided by the University of Missouri. Population (N=3,119) includes all Ph.D. recipients that graduated from the University of Missouri during calendar years 2011-2020. The analytical sample (n=2,422) includes records for which Academic Analytics was able to identify an employment record excluding individuals over the age of 65 and records with a "N/C" SOC code. The category of Multiracial/Other/Not Specified includes non-residential/international students. Due to changes in record-keeping, some observations have zero time to degree.

Table 3

Percentage of Graduates by Mismatch Type and Degree Field

Degree Field	*n*	Vertical mismatch (%)	Horizontal mismatch (%)	Vertical and horizontal mismatch (%)
All	2422	43.9%	43.3%	33.5%
Education	358	50.3%	47.8%	34.9%
Engineering	271	65.3%	62.0%	52.4%
Humanities and Arts	226	30.5%	29.2%	27.4%
Life Sciences	587	44.5%	45.0%	36.5%
Mathematics and Computer Sciences	168	49.4%	29.2%	26.2%
Physical Sciences and Earth Sciences	182	53.8%	47.3%	40.1%
Psychology and Social Sciences	300	36.7%	46.0%	31.0%
Other (Non-STEM)	330	26.1%	32.4%	17.9%

Note. Limited to the analytical sample.

Table 4

Probability of Mismatch by Degree Field

	Verticial mismatch		Horizontal mismatch		Verticial and horizontal mismatch	
Variable	Probit coe.	Dy/Dx	Prob coe.	Dy/Dx	Prob coe.	Dy/Dx
Broad Degree Field (Reference=Mathematics and Computer Sciences)						
Education	0.059	0.023	0.482***	0.180***	0.229***	0.080***
	(0.068)	(0.027)	(0.062)	(0.023)	(0.064)	(0.022)
Engineering	0.485***	0.184***	0.853***	0.325***	0.698***	0.260***
	(0.021)	(0.007)	(0.014)	(0.007)	(0.017)	(0.008)
Humanities and Arts	-0.495***	-0.185***	0.018	0.006	0.026	0.009
	(0.063)	(0.023)	(0.056)	(0.019)	(0.059)	(0.019)
Life Sciences	-0.111***	-0.043***	0.404***	0.150***	0.249***	0.087***
	(0.041)	(0.016)	(0.022)	(0.007)	(0.032)	(0.011)
Physical Sciences and Earth Sciences	0.085	0.033	0.419***	0.156***	0.323***	0.115***
	(0.063)	(0.025)	(0.043)	(0.017)	(0.048)	(0.018)
Psychology and Social Sciences	-0.333***	-0.127***	0.454***	0.169***	0.119***	0.040***
	(0.040)	(0.015)	(0.033)	(0.011)	(0.032)	(0.011)
Other (Non-STEM)	-0.557***	-0.205***	0.095***	0.033***	-0.304***	-0.090***
	(0.042)	(0.016)	(0.034)	(0.012)	(0.038)	(0.012)
Academic Background Characteristics						
Cumulative GPA	-0.189		-0.316**		-0.376**	
	(0.190)		(0.147)		(0.185)	
Time to degree	0.072***		0.021*		0.030**	
	(0.021)		(0.013)		(0.014)	
Socio-Demographic Characteristics						
Age	0.037		0.041		0.016	
	(0.087)		(0.085)		(0.072)	
Age squared	-0.091*		-0.046		-0.053	
	(0.048)		(0.049)		(0.046)	
Sex (Male = Reference)						
Female	0.001*		0.001		0.001	
	(0.001)		(0.001)		(0.000)	
Race/ethnicity (White = Reference)						
Black	0.053		0.071*		0.078*	
	(0.043)		(0.040)		(0.042)	
Asian or Pacific Islander	0.110		-0.139		-0.079	
	(0.181)		(0.133)		(0.147)	
Hispanic	0.139		-0.005		0.089	
	(0.247)		(0.266)		(0.194)	
Multiracial/Other/Not Specified	-0.383***		-0.351		-0.451**	
	(0.126)		(0.272)		(0.185)	
Citizenship (U.S. Citizen = Reference)						
Non-Resident/International	-0.094		-0.145		-0.170	
	(0.076)		(0.096)		(0.114)	
Cohort FE	YES	YES	YES	YES	YES	YES
N	2,374	2,374	2,374	2,374	2,374	2,374

Note. + p<0.1, * p<0.05, ** p<0.01.

Table 5

Occupation-Level Returns to Schooling by Type of Mismatch and Degree Field

VARIABLES	Vertical mismatch	Horizontal mismatch	Vertical and horizontal mismatch
Mismatch	23.395***	18.304***	15.180***
	(0.112)	(0.081)	(0.078)
Education	0.272	8.476***	8.847***
	(0.002)	(0.111)	(0.062)
Engineering	1.408**	5.324***	3.791***
	(0.007)	(0.029)	(0.026)
Humanities and Arts	-8.329***	-1.804	1.234
	(-0.052)	(-0.017)	(0.010)
Life Sciences	-0.370	16.406***	11.909***
	(-0.002)	(0.102)	(0.097)
Other (Non-STEM)	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Physical Sciences and Earth Sciences	-18.397***	-7.646***	-3.635***
	(-0.279)	(-0.042)	(-0.032)
Psychology and Social Sciences	1.500	14.988***	11.689***
	(0.014)	(0.107)	(0.112)
Cohort Fixed Effects	YES	YES	YES
Degree Fixed Effects	YES	YES	YES
Observations	1,844	1,844	1,844
R-squared	0.161	0.209	0.209

Note. + p<0.1, * p<0.05, ** p<0.01.

Table 6

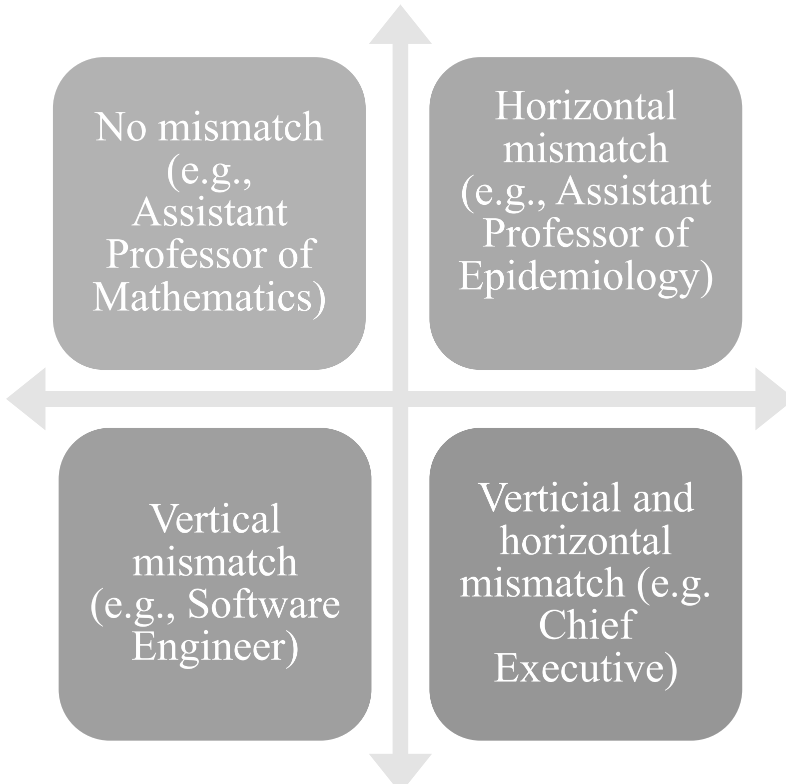
Alternative Specifications of Vertical Mismatch

Variable	Vertical mismatch		Vertical mismatch with standard errors		Heckman correction	
	Probit coe.	Dy/Dx	Prob coe.	Dy/Dx	Prob coe.	Dy/Dx
Broad Degree Field (Reference=Mathematics and Computer Sciences)						
Education	0.059	0.023	0.059	0.023	0.060	0.023
	(0.128)	(0.050)	(0.068)	(0.027)	(0.068)	(0.026)
Engineering	0.485***	0.184***	0.485***	0.184***	0.484***	0.184***
	(0.128)	(0.048)	(0.021)	(0.007)	(0.021)	(0.007)
Humantities and Arts	-0.495***	-0.185***	-0.495***	-0.185***	-0.495***	-0.184***
	(0.139)	(0.052)	(0.063)	(0.023)	(0.062)	(0.021)
Life Sciences	-0.111	-0.043	-0.111***	-0.043***	-0.111***	-0.043***
	(0.117)	(0.046)	(0.041)	(0.016)	(0.040)	(0.016)
Physical Sciences and	0.085	0.033	0.085	0.033	0.085	0.033
Earth Sciences	(0.142)	(0.056)	(0.063)	(0.025)	(0.063)	(0.025)
Psychology and	-0.333***	-0.127***	-0.333***	-0.127***	-0.332***	-0.127***
Social Sciences	(0.127)	(0.049)	(0.040)	(0.015)	(0.038)	(0.013)
Other (Non-STEM)	-0.557***	-0.205***	-0.557***	-0.205***	-0.557***	-0.205***
	(0.130)	(0.048)	(0.042)	(0.016)	(0.042)	(0.017)
Controls	YES	YES	YES	YES	YES	YES
Clustered SE's			YES	YES	YES	YES
N	2,483	2,483	2,374	2,374	3,080	3,080

Note. + $p<0.1$, * $p<0.05$, ** $p<0.01$.

Figure 1

Education-Job Mismatch for Mathematics and Computer Sciences



Note. Diagram adapted from Manuel Salas-Valesco (2021).

Figure 2

Percentage of Graduates by Mismatch Type and Degree Field

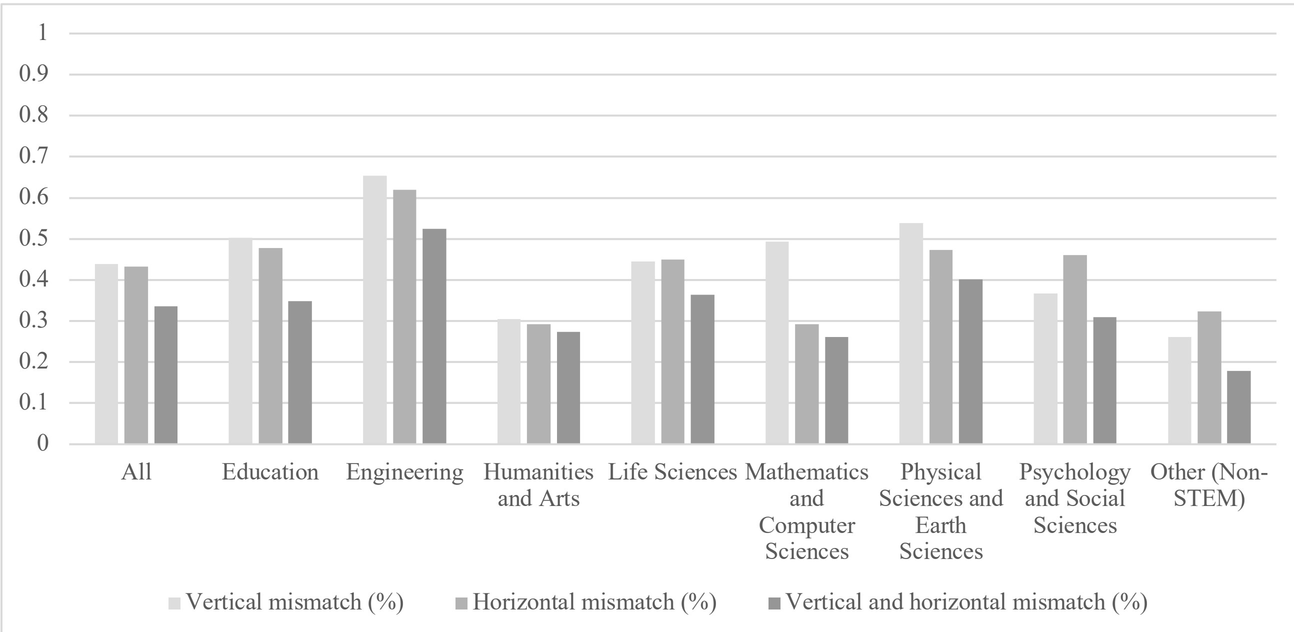
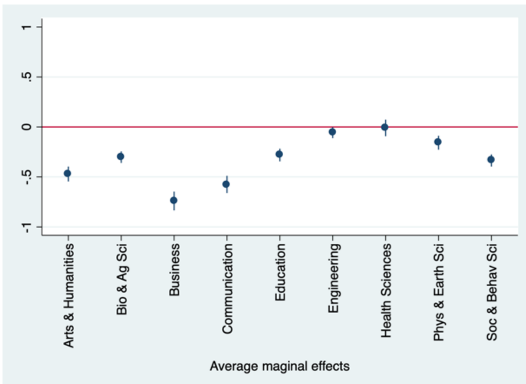


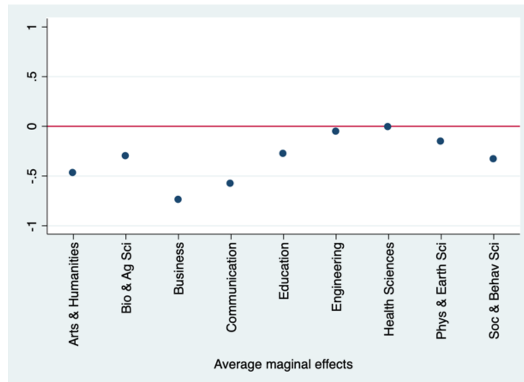
Figure 3

Alternative Specifications of Vertical Mismatch

A. Probit Model

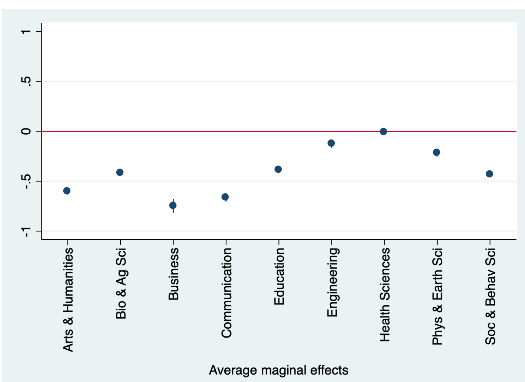


B. Probit with Clustered Standard Errors



C. Probit Model with Clustered Standard Errors

and Heckman Estimation Procedure



CHAPTER FIVE: CONCLUSIONS

Each of the essays presented above addressed a different outcome of higher education. As the findings and implications vary substantially, I summarize each individually.

In the first paper, I examine the impact of state merit-aid program adoption on the stock of human capital across rural and urban populations. Using data from the 1990 and 2000 Decennial Censuses and the 2008-2012 ACS, I utilize a staggered difference-in-difference methodology that exploits the exogenous variation in the timing of program adoption to produce causal estimates. I find that program adoption significantly reduces the share of bachelor's degree holders in rural counties, which increases the rural-urban college attainment gap by a small margin (1.2 percentage points). However, the extent to which the observed effect can be attributed to the lack of efficacy of such policies on rural students or outmigration effects remains unclear. Future research might benefit from examining whether the impact of merit-aid adoption differs by broad-based and more targeted state merit-aid programs or using longitudinal datasets that are able to track rural high school students through college and beyond.

The second paper examines the impact of honors college participation on collegiate outcomes. For compliers near the threshold, honors participation increases average first-year GPA by 0.2-grade points and senior-year GPA by 0.4-grade points. However, these findings may reflect that honors courses have different grade distributions than non-honors courses. In addition, we find generally positive effects on persistence and graduation rates at a narrow bandwidth, but none of the estimates are statistically significant. Given the substantial resources invested in honors programming, the findings may have implications in terms of whether such programs contribute to student success (Bowman & Culver, 2018).

The final paper examines the extent to which vertical mismatch and horizontal mismatch vary by degree field for Ph.D. recipients from the University of Missouri from 2011 to 2020. While the study is purely descriptive, I find that Engineering graduates were most likely to work in occupations for which they were over-educated or in a job outside their degree field. An analysis of occupation-level earnings suggests that mismatch may have a positive effect on potential earnings. The findings are somewhat unexpected, given previous research suggesting that education-job mismatch results in greater wage penalties for degree fields emphasizing occupation-specific skills (Robst, 2007). At the same time, the study was somewhat limited, given the use of broad-level Standard Occupational Codes (SOC). Future research might benefit from using more detailed SOC codes to reduce measurement error.

References

Astin, A. W. (1993). *What matters in college? Four critical years revisited.* San-Francisco: Jossey-Bass.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press: Chicago and London.

Bruce, D. J., & Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, *81*, 30–44. https://doi.org/10.1016/j.jue.2014.01.006

Dynarski, S. M. (2004). The New Merit Aid. In *College Choices: The Economics of Where to Go, When to Go, and How to Pay For It: Vol. I* (Issue September). http://www.nber.org/chapters/c10098.pdf

Fitzpatrick, M. D., & Jones, D. (2012). High education, merit-based scholarships and post-baccalaureate migration. In *National Bureau of Economic Research Working Paper Series*.

Fitzpatrick, M. D., & Jones, D. (2016). Post-baccalaureate migration and merit-based scholarships. *Economics of Education Review*, *54*, 155–172. https://doi.org/http://dx.doi.org/10.1016/j.econedurev.2016.07.002

Groen, J. A. (2011). Do Merit-Aid Programs Help States Build Skilled Workforces? *Change: The Magazine of Higher Learning*. https://doi.org/10.1080/00091383.2011.618080

Moretti, E. (2004a). Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, *121*(1–2), 175–212. https://doi.org/10.1016/j.jeconom.2003.10.015

Moretti, E. (2004b). Human capital externalities in cities. In *Handbook of regional and urban economics* (Vol. 4, pp. 2243–2291). Elsevier. https://learning.oreilly.com/library/view/handbook-of-regional/9780444509673/xhtml/S1574008004800087.htm

Sjoquist, D. L., & Winters, J. v. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, *80*, 39–40. https://doi.org/10.1016/j.jue.2013.10.003

VITA

Heather Hoffman is Director of Analytics and Assessment for the Graduate School at the University of Missouri (MU). Her research interests include economics of education, labor economics, and higher education administration. She holds a master's degree in Educational Leadership & Policy Analysis, a bachelor's degree in Sociology, and a bachelor's degree in Women's & Gender Studies. Previously, she worked as a senior analyst with the MU Office of Institutional Research & Quality Improvement and as an analyst with MU Advancement.