ESSAYS ON EDUCATION AND DISCRIMINATION

by

YUANTING WU

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Jennifer Hunt

And approved by

New Brunswick, New Jersey

October, 2020

ABSTRACT OF THE DISSERTATION

Essays on Education and Discrimination

by YUANTING WU

Dissertation Director:

Jennifer Hunt

In Chapter 1, I investigate the migration response of college students to tuition differences between states, using variation introduced by tuition regional reciprocity agreements. Out-of-state students generally pay higher tuition than in-state students, but reciprocity agreements reduce the premium paid by students from other states in the agreement (sometimes to zero). I examine migration between directed pairs of states, with the tuition difference faced by a potential migrant as the covariate of interest. By instrumenting the tuition difference with a binary variable indicating the pair of states' membership in a common regional reciprocity agreement, I find that a one percent decrease in the nonresident tuition of the destination state due to the regional reciprocity agreements would increase nonresident students' inflow to the destination state by 0.4-0.5%. The reduced form shows that having a regional reciprocity agreement between states increases college migration between states by 29%.

In Chapter 2, I provide a new method to decompose discrimination by Chinese employers into customer and coworker discrimination. Using data from an online job board, I relate employer advertisements for beautiful and tall applicants to occupational job requirements as measured by the American O*NET data. I find that employers hiring in occupations with more contact with customers are more likely to require beautiful applicants in their job ads and employers hiring in occupations with more contact with coworkers are more likely to require tall applicants in their job ads. Customer discrimination plays a more important role in terms of contributions to the R-squared for both beauty and height requirements than coworker discrimination.

The determinants of requiring tall applicants are similar for ads requesting males and females. For beauty, on the other hand, the effect of customer contact is driven by jobs requesting females, while the effect of coworker contact is driven by jobs requesting males.

In Chapter 3, I compare how the gender wage gap evolves with age for occupations with different levels of contact with the public, coworkers, and customers. I use the O*NET data describing occupational job requirements to create indices of contact by occupation. I merge these indices with worker data from the Current Population Survey. I find suggestive evidence that the gender wage gap grows faster with age in occupations with greater overall contact; the evidence is stronger and statistically significant for occupations with high direct customer contact and with high public contact. This link is stronger for non-college graduates than more educated workers. I hypothesize that perceptions of how male and female beauty change with age could explain the results.

Acknowledgements

I am so grateful that I have this opportunity to thank all the people I love. First, I would like to express my deepest gratitude to my advisor and mentor Professor Jennifer Hunt. I first met Professor Hunt at her class, not knowing know that a special and wonderful journey had just started. Her labor class introduced labor economics to me and awoke my interests in education and discrimination. Her empirical microeconometrics class helped me build a foundation in empirical analysis. I liked all the materials in her class, so it was straightforward to make the choice of becoming her student. It has really been an honor to be her student and I think that is the best choice I have made in my life. Professor Hunt helped me make the hardest transition from a student to a researcher. She is the person who helped me find my first research idea, who checked my first piece of code, who read my first paragraph and who gave me the first comment. I still remember all those days and nights she worked with me back and forth editing on my papers. Every time I got my paper back from her, I would always get overwhelmed by the first glance at her pages full of comments and edits and then impressed by her magic edits that put words together and her insightful comments that helped me to clarify the full picture and inspired me to think more. Step by step, she guided me to develop my own research, which I would not dare to think of before. This thesis would not be possible without her. Professor Hunt's influence on me is beyond academia and research. Based on her excellent records on teaching and research, my trust in her extends to almost everything. She is always so nice and patient, sharing her life experiences and wisdom with me and never gives me vague answers to the sometimes strange questions I ask. I know that life will not be easy and the research journey is even harder. I know that I still have a lot of growing, learning and work to do. But, having her as my example, I am confident to continue that journey and not afraid of anything.

I am heavily indebted to my other committee members for their help in reviewing my thesis and for their tremendous support, guidance, and inspiration for my research and career. My gratitude goes to Professor Amanda Agan. Every time I was about to give up, her passion and enthusiasm about research encouraged me and cheered me up. Every time I was out of ideas, her intelligence inspired me and helped me find new solutions. Under her influence, I have started to enjoy doing research. My gratitude also goes to Professor Bingxiao Wu. She has always been willing to help and she has guided my research from a different perspective. She emphasizes the value of practice and trains me with good data process habits and skills for empirical analysis.

I also benefited a lot from interaction with other distinguished professors and scholars at Rutgers University. My gratitude goes to Professor Roger Klein, Professor Richard Mclean, Professor Todd Keister, Professor Colin Campbell, Professor Roberto Chang, Professor John Landon-lane, and Professor Michael Bordo for building my foundation in economics. My gratitude also goes to Professor Hilary Sigman, Professor Carolyn Moehling, Professor Anne Morrison Piehl, Professor Ira Gang, and Professor Jacob Bastian for the comments and suggestions for my papers. My gratitude also goes to Professor Oriol Carbonell-Nicolau, Professor Xiye Yang, Professor Yuan Liao, Professor Zhifeng Cai, Professor Thomas Prusa, and Professor Barry Sopher for their help and guidance in the job market.

I would like to thank all the friendly and kind staff in the economics department: Linda Zullinger, Janet Budge, Donna Ghilino, Debra Holman, Ashley Pavlis.

I would like to thank my friends and peers. My gratitude goes to the labor group members: Hyunjoo Jang, Suwei Luo, Monica Mogollon Plazas, Yilin Wu, Yutian Yang, Kihwan Bae, and Mike Cassidy for the discussions and comments. My gratitude goes to the Hot-pot Lover Group: Xuan Zou, Mark Avery, Wanjun Li, Humberto Martinez, Hang Miao, Lu Wang, Fanle Wei, Yan Xue, Weijia Peng, Weinan Yan for the friendship and companionship.

Finally, my gratitude goes to my parents back in China. Although they might not

have had a clear idea of what I have been doing, this has not stopped them from supporting me. They are on my side like they always are. Dad and Mom, I made it!

Dedication

To my mom Yunqiu and dad Linhua

Contents

List of Figures

List of Tables

Introduction

A crucial ingredient for an efficient labor market is high-quality matching between workers and jobs: having workers in the jobs which fit them best. On the one hand, this means that people can work in firms where they can have the best opportunity and training to maximize their productivity given their human capital and work in the jobs that are most suitable for their skills without the constraints of gender, race, age, and appearance. On the other hand, it also means that people acquire the level and type of human capital best suited to their inherent abilities and talents.

In my dissertation, I examine issues related to the matching of both people to skills and workers with skills to jobs. In my first chapter, I investigate whether an interstate tuition policy helps students move to the places where the level and type of education offered by the colleges are a better match with the students' talents and skills. In my second and third chapters, I seek evidence for the presence of beauty discrimination, which might obstruct workers perceived as less attractive from working in the jobs that best match with their skills. I examine this in both Chinese job ads and American wages and I explore the sources of this discrimination.

I explore the matching of people to skills in Chapter 1. High-quality matching of students to the appropriate college (and major) maximizes national efficiency in the production of human capital. College tuition is an important determinant of where students attend college, but out-of-state students face higher tuition than in-state students in public universities. The high out-of-state tuition prices for nonresident students reduces the efficiency of human capital production if students who match best to an out-of-state school and would have attended it had there been no distinction between students for tuition instead attend an in-state school due to financial concerns.

In Chapter 1, I find that a decrease in out-of-state tuition, which implies a decrease in the gap between out-of-state and in-state tuition, would increase nonresident students' inflow into that state. This indicates that there exist students who would have attended out-of-state universities in the absence of a tuition gap. I conclude that the interstate tuition gap leads to inefficiency.

I investigate the matching of workers with skills to jobs in Chapter 2 and Chapter 3. The high-quality matching of workers with skills to jobs means that workers work in jobs that are suitable for their skills. Customer discrimination might lead employers to hire less from a discriminated group than in the absence of discrimination in order to cater to customers. It is profitable because employers can attract more customers by avoiding hiring from the discriminated group. However, this might cause welfare loss because of the loss in productivity due to insufficient hiring of the discriminated group and the losses in earnings suffered by the discriminated group might outweigh the extra happiness the dominant group gains from avoiding the discriminated group.

In Chapter 2 and Chapter 3, I find evidence suggesting the presence of customer discrimination in both job ads and wages, particularly as it relates to perceived beauty, in the labor market. In China, I find that employers hiring in occupations with more contact with customers are more likely to require beautiful applicants in the job ads. In the United States, I find that the gender wage gap grows faster with age in occupations with greater overall contact and the link is stronger and statistically significant for occupations with high direct customer contact. I hypothesize that any interpersonal contact effects are likely to reflect perceived beauty-based discrimination. I conclude that the quality of the matching of workers with skills to appropriate jobs still has room for improvement.

Though the matching of both people to skills and workers with skills to jobs are imperfect, I argue that the job-matching issue I study is a greater concern because it affects more people. The interstate college tuition gap affects only those who obtain some college education as well as a few who might be deterred from going to college at all, while discrimination in jobs can affect women, Blacks, Hispanics, and many other vulnerable groups.

Another reason why job-matching may be a greater concern is that while discrimination and constraints in college choice both reduce wages, discrimination is more likely to prevent a person from being employed. Loss of employment means a person's basic needs are less likely to be met, and according to the Maslow's hierarchy of needs, the basic needs should take priority and need to be satisfied first. In addition, low income and high unemployment due to discrimination might bring other social concerns such as crime.

I argue that the matching problem in jobs is worse in China than in U.S. In Chapter 2, I find that beauty, age, and gender profiling are still quite common in job ads, while these practices have been prohibited in U.S. by law since 1964. With the absence of the regulations and laws against discrimination in China, the discrimination in job ads found in Chapter 2 might just be the tip of the iceberg. In addition, the discrimination found in job ads in China in Chapter 2 is different from the discrimination found in wages in the U.S. in Chapter 3. Chinese employers are excluding an entire group of people from being considered, which indicates a conscious and strong prejudice towards that group of people, while discrimination reflected in wages might be the result of subconscious behaviors, which is less strong.

The results of my three chapters suggest several solutions that might help to improve the matching. In Chapter 1, I show that regional tuition reciprocity agreements can reduce the interstate tuition gap and encourage interstate migration for college students. These existing regional agreements set good examples for policy makers, who might be induced to institute a national wide tuition reciprocity policy. In Chapter 2 and Chapter 3, I find that the source of discrimination related to perceived beauty mainly comes from customers. Therefore, many current discrimination related policies targeting employers might be inappropriately targeted, while policies regulating the behaviors of customers are more crucial. For example, it might be helpful to launch more public service advertisements to encourage interactions with the discriminated groups.

Chapter 1

The Impact of Interstate Tuition Differences on College Student Migration: Evidence from Regional Reciprocity Agreements

1.1 Introduction

Nonresident students pay as much as two or three times the in-state tuition in public postsecondary institutions in the United States. In 2017, the average ratio of nonresident to resident tuition in public postsecondary institutions was around 2.6, though it varied considerably across states: for example, Massachusetts had the highest ratio with 7.3 and South Dakota had the lowest ratio with 1.2.[1] Tuition is clearly an important concerns when students are deciding where and which college to attend.[2] Many students have to restrict their choices to in-state schools due to financial concerns, limiting their ability to attend the college which is the best fit, and hence leading to inefficiencies (Knight and Schiff 2019). In Fall 2016, around 81% of the first-time bachelor's degree seeking residents in the United States studied in their home states with the largest being 91% in Utah (Snyder, Brey and Dillow 2019). It is natural to ask the degree to which this high share represents a distortion caused by barriers to out-of-state enrollment and whether there are policies that can improve efficiency.

[1] Data from 2017 Integrated Postsecondary Education Data System (IPEDS).

[2] In 2012, 67% of high school seniors in the High School Longitudinal Study of 2009 reported that cost of attendance was very important in influencing their college choice, and 29% said it was somewhat important (LaFave, Kelly and Ford 2018).

The formation of regional reciprocity agreements is one such intervention that reduces the tuition public institutions charge nonresidents from other member states. In 1957-1958, the New England Board of Higher Education (NEBHE) established the first-ever regional reciprocity agreement, the Regional Student Program (RSP), to share higher education resources and expand educational opportunities for residents in the New England area. With this agreement, nonresidents from the member states in New England can study in member states at discounted tuition ranging from the same rate as residents to only 150% of resident tuition depending on the state and program they attend. Since then, following the RSP, other regions in the United States have launched similar regional reciprocity agreements. The Southern Regional Education Board (SREB) launched the Academic Common Market (ACM) program in 1979; the Western Interstate Commission for Higher Education (WICHE) established the Western Undergraduate Exchange (WUE) program in 1988; and the Midwestern Higher Education Compact (MHEC) set up the Midwest Student Exchange Program (MSEP) in 1994. As of 2018, these regional reciprocity agreements cover 45 states in the United States in total.[3] According to the annual report of the RSP of NEBHE, in the academic year 2018-2019, a full-time nonresident undergraduate saved an average of $7,900 by taking advantage of the program. These programs have drawn limited attention from economists.

In this paper, I ask what impact tuition gaps between in- and out-of-state tuition have on college student migration. The challenge for identifying the causal effect of the tuition gap on college migration is that tuition is endogenous, meaning it could be correlated with migration determinants that are unable to be controlled for, including the quality of education and a sudden migration shock. It is hard to rule out bias from these omitted variables without a credible empirical identification analysis

I exploit the regional reciprocity agreements to estimate the effect of tuition on migration of college students, using regional reciprocity as an instrument for the tuition gap between states. I use data on membership of the four major regional reciprocity agreements in the United States since 1958. I also analyze the effect of

3 There are 6 states in the RSP; 14 states in the ACM; 15 states in the WUE, and 10 in the MSEP.

bilateral reciprocity agreements negotiated by neighbor states, though not those negotiated between pairs of schools in different states. Using microdata from the 5% sample of the 1960, 1970, 1980, 1990 and 2000 the United States Census of population, I match directed pairwise college migration flows observed between states with the states' reciprocity agreement membership, tuition data from the Integrated Postsecondary Education Data Service (IPEDS), state education expenditure data from the Annual Survey of State and Local Government Finances, unemployment data from the Bureau of Labor Statistics (BLS) and gross state product data from the Bureau of Economic Analysis (BEA).

I argue that the regional reciprocity agreements are unlikely to be correlated with these omitted college migration determinants: education quality and unobserved shocks. Most states joined the agreements all at once as soon as the agreement was established, making it unlikely each was responding to an unobserved shock. To check this, I repeat my analysis using a sample restricted to members of those agreements where 75% of members joined all at once and have more than two states in them, and to states that never joined an agreement. In addition, I also directly explore the pre-trends in migration before each state-pair joined the reciprocity agreement. Endogeneity is more likely in bilateral agreements with neighboring states, which is why I use these agreements only as a robustness check.

I find that the LATE effects estimated by my IV, in all specifications, are substantially larger (in absolute value) than the fixed effects results, which, if unbiased, represent average treatment effects (ATEs). A one percent decrease in the nonresident tuition the of destination state due to the regional reciprocity agreements increases nonresident students' inflow to the destination state by 0.4-0.5%. To learn more directly about the reciprocity agreement program, I also provide a reduced form evaluation of the regional reciprocity agreements program on college student migration using the fixed effects model. The reduced form shows that having a regional reciprocity agreement between states increases college migration between states by 29%.

My results may most directly be compared with those of Dwenger, Storck, and Wrohlich (2012), who study the introduction of tuition fees in certain German states.

They find that the introduction of home state tuition fees reduces the probability students apply to a university in their home state by 2 percentage points. The most similar American paper to mine is Knight and Schiff (2019). They compare attendance at institutions for students living close to state borders from 1997 to 2011 using a border discontinuity design. They find that a 1000 dollar increase in tuition is associated with 6 less students. And in addition, they compare borders of states in the same reciprocity agreement with the borders of states not in the same reciprocity agreement and find that borders under the same reciprocity agreement have 22 fewer students enrolling in-state. The effect found in my study is larger than theirs (a 1000 dollar increase is associated with 40 fewer students in my study).

The main differences between my study and theirs are identification strategy and data. Firstly, I use IV while they use a regression discontinuity (RD) to identify the effect of tuition on enrollment. RD must focus on students living near to the border, while the effect on students living close to border might be different from the effect on other students. Therefore, their RD result is less representative than my IV result. Besides, the comparison of in-state enrollment change crossing the borders between states in the same reciprocity agreement and states not in the same reciprocity agreement could be biased. Given students can study in any member state in the same regional reciprocity agreement, it is not necessary for them to just study in states directly across the border. Given the large regional area covered by each agreement, there are many other states than the adjacent states that are also available to students of member states. For example, if we think of two neighbor states that are in different reciprocity agreements, this is a pair of states which their paper assigns to the "control" groups. But the size of the border discontinuity actually reflects which side the reciprocity agreement has a larger power. In addition, they did not take into consideration bilateral agreements negotiated by neighbor states or neighbor schools. Lastly, the Higher Education Research Institute data used in their study is a survey data that covers only institutions that responded, which might have more measurement error problems and might be less representative than the Census data used in my study.

Although there is a literature examining the impact of tuition on enrollment, which finds mixed results,[4] there is surprisingly little research on the longstanding, geographically widespread regional tution agreements specifically. DesJardins (1999) studies the effect of the tuition reciprocity agreement between Minnesota and Wisconsin in 1997 and find that a 196 dollar decrease in tuition is associated with having 8 more students. Herzog and Stanley (2017) find that residents in states joining the WUE are 57% more likely to enroll than residents in other states. Rizzo and Ehrenberg (2004) study how the nonresident enrollment strategies at institutions react to changes in federal and state need-based student aid and state appropriations in 91 flagship public research institutions in the United States during the 1979 to 1998 period. They control for the share of undergraduates who are reciprocal to control for the enrollment pressure institutions face. They find a small negative relationship between share of reciprocal undergraduates and share of out-of-state undergraduates using cross-section analysis and no relationship in the panel analysis. Likewise, Marsicano (2015) studies the effect of a state's membership in a regional reciprocity agreement on its own nonresident enrollment. This study focuses on the out-of-state enrollment in institutions located in the border of states from 2003 to 2012 and finds that four-year institutions in a state participating a reciprocity agreement on average have 100% enrollment increase. Firstly, my study covers a longer period (1960-2000) and wider region (nationwide), which allow the membership variable in my study have more variations. The most volatile period of state membership change happened before 1994. Most previous studies either cover a later period or cover a shorter period when the membership barely changed over time. Therefore, the effect they captured mostly only come from difference from region to region. Rizzo and Ehrenberg (2004) measure membership via a survey of institutions with a low response rate. Secondly, given the outcome variable in these studies is stock variable: the share of out-of-state students in institution, they are not checking the matrix of flows between states as I do. Therefore, it is impossible for them to control for the source of state characteristics which are also very important to college migration.

I find larger impacts than papers studying the effect of merit-based scholarships at

[4] McHugh and Morgan (1984), Leslie and Kane (1994), Card and Lemieux (2000), Mixon and Hsing (1994), Groat (1964), Tuckman (1970), Morgan (1983), Mixon (1992) and Noorbakhsh and Culp (2002).

public universities, designed to keep top students in state. Dynarski (2004) shows that the effect of the HOPE scholarship on student enrollment might happen through retaining students who would have studied out of state. Zhang and Ness (2010) checked the brain drain phenomenon in the United States in STEM major and find that merit-based scholarship successfully attract more talent students with a 10% increase in in-state enrollment and a 10% decrease in students out migration. Cornwell, Mustard and Sridhar (2006) find that HOPE in Georgia increased its freshman enrollment by 5.9%. Cohodes and Goodman (2014) compare students just below and above the merit scholarship threshold of schools with relatively lower quality and finds that eligibility for the scholarship increases student enrollments in these schools by 4.8-6.9%. Kane (2007) finds that the DC TAG program increase enrollment in DC and a 1,000 dollars (in 2002 dollars) decrease in tuition is associated with 5.4% increase in enrollment. Given that my results show that 1% decrease in nonresident tuition, which is around 52 dollars, is associated with 0.36% increase in enrollment and my reduced form results show that the effect of the reciprocity agreement is around 25% increase in nonresident enrollment, overall, the effect found in my results is larger than most results find via merit-based scholarships.

A usual concern with the merit-based scholarship analysis is that it may have endogeneity problems because student with better academic performance might have some unobserved personal characteristics that determine the enrollment as well. In order to solve this endogeneity problem, many studies of merit-based scholarship have to restrict their sample to a specific group or a specific state, which limits their generality.

1.2 Data

I use four sources of data in this analysis. Firstly, I use the 5% sample of 1960-2000 Census data from the Integrated Public Use Microdata Series (IPUMS) for the migration and demographic information. I identify undergraduate students aged 30 or under who enroll in public universities and note their current state and their reported state five years previously. If these two states are not the same, then I count this student as a college migrant, and I define his or her current state as the destination state and the state five years prior as the source state. I aggregate the number of college migrants by each directed pair of states in the United States by each Census year. There are 12750 directed pairs (2550 per year*5 year) in total in my sample and 28% (3566) pairs among them have zero college migration.[5] I do not extend the sample period using the ACS because the question about past migration refers to 1 year prior rather than 5, introducing a break in the series.[6]

Secondly, I collect the state reciprocity agreement participation information by checking websites and contacting program directors. This dataset includes the membership of each state from 1958 to 2018 in the four major regional reciprocity agreements and bilateral reciprocity agreements negotiated by neighbor states. Thirdly, I get the tuition information data from the 1980 and 1984-2017 IPEDS dataset.[7] The unit of observation in the IPEDS dataset is institution. This dataset includes important variables such as out-of-state tuition and in-state tuition for each public institution. However, the data for DC in 1984 is missing. I take unweighted averages of the institution-level out-of-state tuitions and in-state tuitions separately by state and year to get the average annual state level tuitions for both nonresidents and residents.[8] Finally, I add more control variables by using gross state product (GSP) from 1962 to 2017 from the Bureau of Economic Analysis (BEA). I impute the 1960 value for use in my regressions.[9]

I get unemployment rates from 1976 to 2017 from the Bureau of Labor Statistics (BLS). The years before 1976 are not available in the BLS. Therefore, I also construct the state unemployment rate from 1960 to 2000 using the Census 1960-2000. In order

[5] I check the robustness of the results to recoding these zero flows as the migration of a single person and the results are similar.

[6] In future work, I will check the robustness of the results to multiplying ACS migration rates by five.

[7] The 1981-1983 IPEDS data is unavailable.

[8] The results does not change much whether using the weighted or unweighted average tuition.

[9] I impute using predicted values from:

$$log(GSP_{st}) = \beta_1 log(GDP_t^{US}) * Region_r + \beta_2 unemploymentrate_{st} + \beta_3 log(population_{st}) + \gamma_s + \beta_4 year + \epsilon_{st}$$

where s indexes state, t stands for census year. $Region_r$ stands for the census region dummy and γ_s stands for a set of state dummies. For this purpose, I ignore the break in the GSP series in 1997. I use decadal data, but the imputed results are very similar if I use yearly data, whose disadvantage is that I cannot control for annual state unemployment rate because the annual unemployment rate before 1976 is not available in the BLS.

to make sure the unemployment rate is comparable; I use state unemployment rate from the BLS in the IV regressions which only cover years after 1980 and use the state unemployment rate computed by the Census for the reduced form analysis. I collect government education expenditure ranging from 1960 to 2017 from the Annual Survey of State and Local Government Finances.[10]

I define a new variable named Log tuition gap to measure the log of the difference in tuition between source and destination state. It is defined as:

$$logtuitiongap_{sdt} = \begin{cases} log(residenttuition_{dt} - log(residenttuition_{st}), & Reciprocity_{sdt} = 1 \\ log(nonresidenttuition_{dt} - log(residenttuition_{st}), & Reciprocity_{sdt} = 0 \end{cases}$$

Where $Reciprocity_{sdt}$ is the reciprocity agreement dummy, it equals 1 if the pair of states (s, d) are in a common regional reciprocity agreement in year t; otherwise it equals 0.[11] $residenttuition_{st}$ is the average in-state tuition in source state s in year t. $residenttuition_{dt}$ is the average in-state tuition in destination state d in year t. $nonresidenttuition_{dt}$ is the average out-of-state tuition in destination state d in year t.[12]

Given the unavailability of certain variables in certain years, I define two samples. First, I define the shorter sample covering the period 1980-2000 for the IV analysis due to the tuition information being available only since 1980. Second, I define the longer sample covering 1960-2000, used in the reduced form analysis.

Table 1.1 shows the summary statistics for the shorter 1980-2000 sample used in the IV analysis. Column 1 shows there are 7650 potential source-destination pairs by year in my sample and among them, 17% (1365) pairs have zero college migration. Among the 6285 pairs of states by year with positive college student migration, 835 (12.3%) of them have a reciprocity agreement. Columns 2 and 3 show more details of how the number of college students migrating between pairs of states varies by

[10] The 2001 and 2003 data are missing, but my analysis currently ends in 2000.

[11] I only included the major four regional reciprocity agreements in the IV analysis, so Log tuition gap for those states with bilateral agreements are left as through there is no agreement.

[12] Note: Having a reciprocity agreement between A and B does not in general imply a tuition gap of zero. The existence of the reciprocity agreement only allows resident from A to pay B's in-state tuition (rather than B's out-of-state tuition when absent of agreement) when studying in B. The tuition gap then reflects any difference in in-state tuition.

whether the pair is in a common reciprocity agreement: the average Log tuition gap of pairs with a reciprocity agreement is 0.01 while the average Log tuition gap for pairs without any reciprocity agreement is 0.5, which is much higher than states with reciprocity agreements. The average number of students migrating between pairs with a reciprocity agreement is higher than pairs without any reciprocity agreement.

Table 1.2 Column 1 shows corresponding statistics for the longer sample covering 1960-2000, used in the reduced form regressions. There are 12750 potential pairs by year in this sample and among them, 28% (3566) pairs have zero college migration. Among 9184 pairs of states having positive college migration, 883 (9.6%) of them have a reciprocity agreement. Columns 2 and 3 show the average number of students migrating between pairwise states with any reciprocity agreement is 459, while the average number for those pairs of states without any reciprocity agreement is lower and only 346.

1.3 Methodology

My identification strategy consists of three parts: directed pair-wise state fixed effects (for each state pair i and j, a dummy for flows from state i to state j and another for flows from state j to state i), instrumental variables and reduced form regressions.

1.3.1 Fixed effects estimation

I begin by estimating a log-log fixed effect regression to estimate the effect of the tuition gap on college migration:

$$log(M_{sdt}) = \theta_1 logtuitiongap_{sdt} + \theta_2 x_{st} + \theta_3 x_{dt} + \alpha_{sd} + \tau_t + \epsilon_{sdt}$$

where s indicates source state, d indicates destination state, and t indicates census year. M_{sdt} is the number of college students moving from the source state s to the destination state d in census year t. x_{st} stands for control variables for source state. It includes log gross state product to control source state economics condition, state and local government education expenditure to control source state government's revenue and government policy, and the state unemployment rate to control source state labor market condition. x_{dt} stands for the same covariates for destination state. α stands for directed pair fixed effects to control for time invariant directed pair specific variables, such as location, climate. τ stands for time fixed effects to control for time-varying but common across directed pair factors, such as national policy, national business cycle. ϵ is the error term. In addition, I also run a specification with source state-specific trends and destination state-specific trends to control for other state-level linear trends in college student migration and run another specification with directed pair-specific trends to control for other directed pair level linear trends in college student migration. These specifications with trends also serve as a robustness check for the underlying assumption of the fixed effects model that the counterfactual trends of college migration between directed pairs of states are identical. All the standard errors are clustered by pairs of states. The $logtuitiongap_{sdt}$ defined before is the difference in log tuition between the source state s and the destination state d in year t.

1.3.2 Instrumental variables estimation

If the college migration determinants mentioned above, along with the tuition gap, capture all determinants of migration, then the fixed effects estimator will be unbiased. However, omitted variables such as education quality, unobserved migration shocks, could cause bias. In addition to the issue of state education quality which I explained in the introduction above, bias could come from state governments' changing tuition in response to unobserved college migration shocks. For example, there is a possibility that some state policymakers want to have more college students with the hope that these educated students would stay after college to increase the productivity of the local labor market. If an unobserved shock decreases the college student inflow to the state, the government might reduce nonresident tuition, and hence the tuition gap, to retain nonresidents. In this case, the effect of the tuition gap estimated from the fixed effects regressions will be biased up (less negative coefficient) because the tuition gap is positively correlated with the shock in the error term. Conversely, if an unobserved shock increases the college student inflow, a state government with limited education capacity and resources may response by raising nonresident tuition, increasing the tuition gap, to restrict the inflow of nonresident students. As before, the effect of the tuition gap estimated from the fixed effects regressions will be biased up (less negative coefficient) since tuition gap is positively correlated with the error term containing the shocks.

To address omitted variables bias, I apply an instrumental variable strategy, using a dummy for a pair of states' memberships in a common regional reciprocity agreement as an instrument. I argue that this instrumental variable satisfies the two requirements for a valid instrument. This instrument is highly correlated with the endogenous variable: the tuition gap. The policy of the tuition reciprocity agreement is to eliminate or shrink tuition gap between resident and nonresident tuition by reducing nonresident tuition. Thus, the tuition gap between destination state and source state would decrease if both states in the pair join the same regional reciprocity agreement.

More importantly, I argue that this instrument is unlikely to be correlated with these omitted college migration determinants mentioned above. The decision to join a regional reciprocity agreement is unlikely to be influenced by an education quality change within state, since joining the agreement will have no direct effect on quality (I assume that the net flows of students are small enough that changes in faculty-student ratios or laboratory crowding are minimal). More plausible is a correlation between joining and unobserved shocks to migration. Consider a state which values out-of-state student but faces a decreasing college inflow. Joining the reciprocity agreement could appear attractive to the state because a lower nonresident tuition brought by the agreement might help the state boost college inflows. Therefore, the effect of my IV would be biased up. However, I argue that this is not very likely. Firstly, note that the participation of most states happened together when the agreement was launched. It seems unlikely that all these states were experiencing college migration shocks in the same direction. For example, 12 of the 14 states in the ACM joined the ACM agreement at once when it was established in 1979. All states in NERSP joined the agreement at once in 1958 when it was established. In addition, almost all the states, except for North Carolina, never leave the agreements (North Carolina joined the ACM in 2001 and left in 2011). For example, all members in NERSP have remained members since joining the agreements since 1958. However, the participation of members in the other agreements has more variation, with some of them joining early and some later. Given that joining behavior that may be endogenous to migration is more likely in these other agreements, I do a robustness check using a restricted sample on those agreements with 75% of members joined all at once when the agreement was established and have more than two states in them, and on states that never joined an agreement. In addition, I also directly check the college student migration trend for each state before the pair of states joined the reciprocity agreement.

I interpret my instrumental variables results through the lens of heterogeneity. Since there is no state that would increase its nonresident tuition when joining the reciprocity agreement, my IV satisfies the monotonicity requirement. Therefore, the effect identified by my IV is a LATE (the local average treatment effect) among "compliers". Compliers in my study refer to those pairs of states who would not have lowered tuition gap without signing the reciprocity agreements but lower the tuition gap due to the reciprocity agreement.

1.3.3 Reduced form models

I use fixed effects regressions to evaluate the reduced form impact of regional reciprocity agreement on migration. The main regression equation is:

$$log(M_{sdt}) = \beta_0 + \beta_1 Reciprocity_{sdt} + \beta_2 x_{st} + \beta_3 x_{dt} + \alpha_{sd} + \tau_t + \epsilon_{sdt}$$

The new covariate compared to the previous regression is $Reciprocity_{sdt}$, which is the reciprocity agreement dummy, equal to 1 if the state pair (s, d) has any reciprocity agreement in year t; otherwise it equals 0. The coefficient β_1 on $Reciprocity_{sdt}$ is the effect of interest. All the standard errors are clustered by pairs, i.e. the number of groups is equal to the number of pair and a pair (s, d) is viewed as a different group from a pair (d, s). The coefficient resulting from this reduced form regression identifies the overall effect from the regional reciprocity agreement. I argue that this reduced form complements the IV results. The LATE identified by IV is of more scientific interest, while the effect from the reduced form of the policy is more straightforward and useful for policy.

1.4 Results

Table 1.3 presents the fixed effects regression results and the IV regression results. Four specifications with increasingly detailed controls for the fixed effects results are presented across Columns 1 to 4, with the corresponding specifications for the IV results in Columns 5 to 8. All the regressions include the log tuition gap variable, which is the main variable of interest. The base specification (Column 1 and Column 5) also controls for the directed pair fixed effects, and year fixed effects. In Columns 2 and Column 6, I also control for both source and destination state log gross state product, log state and local education expenditure and state unemployment rate. In Columns 3 and Column 7, I also include source state-specific trends and destination state-specific trends. In Columns 4 and Column 8, instead of source and destination state-specific trends, I add specific trends for each directed pair of states.

1.4.1 Fixed effects results

Columns 1 to 4 in Table 1.3 show the key coefficient from the fixed effects analysis – the full coefficients are shown in Table 1.A1. The coefficient on log tuition gap is small and statistically significant in all specifications. The Column 1 coefficient of 0.04 implies that a one percent increase in the nonresidents' tuition in the destination state of a directed pair of states would increase its nonresident undergraduate inflow by a negligible 0.04%. Adding more controls in Column 2 to 4 does not change the magnitude of the coefficient a lot.

1.4.2 IV results

To fix potential endogeneity problems of tuition, Table 1.3 Columns 5 to 8 present the IV results with the second stage results in the upper rows and the first stage results below. The coefficient on the reciprocity variable in the first stage in Column 5 in Table 1.3 is -0.36 and it is statistically significantly at 1% level (the full results from the first stage are shown in Table 1.A2). This coefficient shows that having a reciprocity agreement between a directed pair of states reduces nonresident tuition in destination state by 36%. Adding source state-specific trends and destination state-specific trends in Column 7 reduces the coefficient somewhat to -0.26. Adding directed pair-specific trends in Column 8 further reduces the coefficient to -0.17. Given what we know about how much the reciprocity agreements reduce the average tuition gap on paper – 50% - it seems unlikely that their effect is this small. Further, there are only three periods in this IV analysis, and while this is sufficient in theory to estimate a model with trends specific to the unit being followed as a panel (here, direct state pairs), Angrist and Pischke (2008) point out that "three periods is typically inadequate to pin down both the trends and the treatment effect in practice". In addition, given most reciprocity agreements were formed (treatment) before 1990, I have only one pre-treatment period. Pischke (2005) points out that when the effects of the treatment take place dynamically and limited pre-treatment periods are available, panel unit-specific trends would mainly rely on post-treatment periods and absorb the actual treatment effects. Thus, adding specific trends in this case could be problematic. Instead, this result suggests that including directed pair-specific trends constitutes over-controlling and is eliminating the genuine variation introduced by the agreements. Therefore, this specification will not be my preferred one.

The coefficient on log tuition gap in the second stage in Column 5 is -0.49 and it is statistically significant at 1% level (the full results from the second stage are shown in Table 1.A3). This coefficient shows that a one percent decrease in the nonresident tuition for the destination state in a directed pair of states would increase college inflow into that state by 0.49%, consistent with the prediction of theory. Adding the source and destination state characteristics in Column 6 does not change the sign but reduces the absolute value of the coefficient a little from 0.49 to 0.40. The coefficient is still statistically significant at 1% level. Given the standard error is 0.13, coefficient is probably not statistically significantly different from the coefficient in the previous column. Adding source and destination state-specific trends in Columns 7 reduces the absolute value of the coefficient to 0.36 and reduces its statistically significant level to 10%. Again, given the standard error, it shows that the coefficient does not change much. This effect is relatively large in the literature, compared with the effect of a 0.27% increase in enrollment in Kane (2007), where he checks the effect of the DC Tuition Assistant grant program.

In contrast to the stability of the coefficient in Columns 5-7, the coefficient changes greatly, from -0.36 to 0.80, when I add directed pair-specific trends in Column 8 .This seems to suggest that the increase in college migration that seemed to be due to the reduction in the tuition gap is actually due to the directed pair-specific trends. However, I do not prefer this specification, as noted above.

Table 1.A4 shows the results of the robustness check of using a sample restricted to members of those agreements where 75% of members joined all at once and have more than two states in them, and to states that never joined an agreement. These results are suggestively similar to the main analysis using the full sample. In addition, the results of pre-trends checking for each state are shown in Figures 1-5. I see no clear evidence of migration trending either up or down before joining. While this does not completely rule out the possibility that states were reacting to desires to change migration or forecasts of future migration changes, it does make us more confident that they were not reacting to concurrent trends in college migration.

1.4.3 Reduced form results

Table 1.4 presents the main reduced form results – the full coefficients are reported in Table 1.A5. The first four specifications of Table 1.4 correspond to those in the IV analysis. Column 1 in Table 1.4 shows that the coefficient on the reciprocity agreement dummy is 0.22 and it is statistically significant at 1% level. This suggests that the reciprocity agreement between a directed pair of states would increase the undergraduate flow from its source state to its destination state by 22%. Adding source and destination state covariates in Column 2 does not change the coefficient. Adding source and destination state-specific trends in Column 3 increases the coefficient from 0.22 to 0.29 and it is statistically significant. Therefore, the effect so far is quite robust across the different specifications. Adding directed pair-specific trends in Column 4 does not change the coefficient's sign but reduces the coefficient to a small and statistically insignificant 0.08. This suggests that there is almost no effect of the reciprocity agreement on college migration. For reasons stated previously, I do not prefer this specification.

In order to make these reduced form results (based on 1960-2000) consistent with the IV results (based on 1980-2000), in Column 5 in Table 1.4 I restrict the sample to 1980-2000 and return to the Column 3 specification. This cuts the coefficient in half and leaves it statistically significant only at the 10% level. This suggests that, had I been able to estimate the IV results for the longer period, the estimated effects might have been larger. The difference between Columns 3 and 5 in Table 1.4 lies more in the coefficient than the standard error, suggesting that the difference may be caused by heterogeneity in the effects of reciprocity agreements struck before and after 1980.

There are other bilateral reciprocity agreements negotiated by neighbor states or schools themselves. I also do another robustness check by adding a dummy representing these reciprocities. Table 1.5 shows that while the coefficient on the regional agreement dummy is unchanged, the coefficient on the bilateral agreement is much larger (the full coefficients are reported in Table 1.A6). This suggests that although fixed effects remove the direct effect of distance, the effect of reciprocity agreements weakens with distance.

To further check how distance affects the effects. I do another robustness check by extending the model by adding an interaction term of the reciprocity agreement dummy with a dummy representing whether two states are neighboring states. Table 1.6 shows that the effect of neighboring states are much larger (the full coefficients are reported in Table 1.A7). This suggests that the effect of the reciprocity agreements is stronger for adjacent states.

1.5 Conclusion

In this paper, I study how interstate college migration responds to the inter-state tuition gaps in the United States. By instrumenting the tuition gap between the destination state's nonresident tuition and source state's resident tuition in a pair of states with a dummy for the pair of states' membership in a common regional reciprocity agreement, I provide evidence that a reduced tuition gap would increase college migration from the source state to the destination state. I find that a one percent decrease in nonresident tuition of destination state due to the regional reciprocity agreements would on increase nonresident students' inflow to the destination state by 0.4-0.5%. This effect is a LATE effect among the "complier" states who change tuition due to the agreement. The reduced form analysis shows that having a regional reciprocity agreement between states would increase college migration between states by 29%.

My study has several policy implications. Firstly, the results of my study can help policymakers have a better understanding and evaluation of regional reciprocity agreements and confirm their function in providing students with more options and increasing institutions' diversity. However, more evidence on the performance of the additional out-of-state students and their post-graduation geographic mobility, as well as data on the cost of educating a marginal student, is needed to complete the picture. Secondly, young and highly educated people are the most mobile demographic group in the United States, and among reasons of migration, going to another state for college ranks the first. (Raven, Smith, and Wozniak, 2011) This mobility could be even larger with a lower tuition gap according to my study. Thus, reducing the tuition gap could become one of the available options to arrest the long-term decline in interstate migration since 1980.

1.6 Figures and Tables

Table 1.1: Means: sample for the IV analysis (1980-2000)

	All	Reciprocity=1	Reciprocity=0
Panel A: Census& IPEDS 1980-2000			
Log tuition gap*	0.44	0.01	0.50
	(0.63)	(0.44)	(0.63)
College migration from source to destination	404	473	393
	(729)	(690)	(734)
State and local education expenditure (source)	1771	1419	1825
	(2281)	(2367)	(2262)
State and local education expenditure (destination)	1778	1425	1832
	(2280)	(2367)	(2262)
Unemployment rate (source)	5.36	5.10	5.40
	(1.73)	(1.57)	(1.75)
Unemployment rate (destination)	5.36	5.13	5.39
	(1.72)	(1.57)	(1.74)
Gross state product (source)	136919	99255	142690
	(186295)	(179449)	(186668)
Gross state product (destination)	136863	99565	142577
	(186182)	(179380)	(186560)
Observations	6285	835	5450

Notes: Unweighted means, standard deviations in parentheses. The sample is college students aged under 30 who study in public school from Census year 1980-2000. Gross state product and state and local government expenditure is in millions of current dollars. The years before 1997 is based on the Standard Industrial Classification (SIC) and years after 1997 is based on the North American Industry Classification System (NAICS).

$$*Log\ tuition\ gap_{sdt} = \begin{cases} log(resident\ tuition_{dt}) - log(resident\ tuition_{st}), & Reciprocity_{sdt} = 1 \\ log(nonresident\ tuition_{dt}) - log(resident\ tuition_{st}), & Reciprocity_{sdt} = 0 \end{cases}$$

Unemployment rate is collected from Bureau of labor statistics (BLS)

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.2: Means: sample for the reduced form analysis (1960-2000)

	All	Reciprocity=1	Reciprocity=0
Panel B: Census 1960-2000			
College migration from source to destination	358	460	347
	(654)	(677)	(650)
State and local education expenditure (source)	1271	1345	1263
	(2030)	(2323)	(1997)
State and local education expenditure (destination)	1277	1350	1269
	(2030)	(2322)	(1997)
Unemployment rate (source)	0.06	0.06	0.05
	(0.01)	(0.01)	(0.01)
Unemployment rate (destination)	0.06	0.06	0.06
	(0.01)	(0.01)	(0.01)
Gross state product (source)	99589	94307	100151
	(164106)	(175727)	(162822)
Gross state product (destination)	99336	94575	99843
	(164094)	(175680)	(162816)
Observations	9184	883	8301

Notes: Unweighted means, standard deviations in parentheses. The sample is college students aged under 30 who study in public school from Census year 1960 to 2000. Gross state product and state and local government expenditure is in millions of current dollars. The gross state product of the years before 1997 is based on the Standard Industrial Classification (SIC) and years after 1997 is based on the North American Industry Classification System (NAICS). 1960 gross state product is missing in the initial data and is predicted using Census data from 1960-2000.

**Unemployment rate is calculated using Census 1960-2000.*

Source: Census 1960-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.3: **Effects of tuition gaps on college interstate migration (1980-2000)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		OLS regressions				IV regressions		
Log tuition gap*	0.04	0.04	-0.02	0.02	-0.49***	-0.40***	-0.36*	0.80
	(0.03)	(0.03)	(0.05)	(0.08)	(0.13)	(0.12)	(0.19)	(0.67)
R-squared (within)	0.03	0.05	0.13	0.62	0.03	0.04	0.12	0.58
Log gross state product (source&destination)	-	YES	YES	YES	-	YES	YES	YES
Log state&local education expenditure (source&destination)	-	YES	YES	YES	-	YES	YES	YES
Unemployment rate (source&destination)	-	YES	YES	YES	-	YES	YES	YES
Source& destination state-specific trends?	-	-	YES	-	-	-	YES	-
Pair-specific trends?	-	-	-	YES	-	-	-	YES
P value: F-statistic of trends	-	-	0.00	0.00	-	-	0.00	0.00
Coefficient on reciprocity 1st stage	-	-	-	-	-0.36***	-0.39***	-0.26***	-0.17***
	-	-	-	-	(0.03)	(0.03)	(0.01)	(0.04)
Observations 1st stage	-	-	-	-	6285	6285	6285	6285
F-statistic 1st stage	-	-	-	-	166	248	429	18
R-squared 1st stage (within)	-	-	-	-	0.66	0.70	0.89	0.92

*Notes: The dependent variable is the log number of college students migrating between pairwise states. 6285 observations. The sample is college students aged under 30 who study in public school from Census 1980-2000. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The instrument in columns 5-8 is reciprocity, a dummy for a pair of states' memberships in a common regional reciprocity agreement. And the first stage uses census years. *** p<0.01, ** p<0.05, * p<0.1*

$$*Log\ tuition\ gap_{sdt} = \begin{cases} log(resident\ tuition_{dt}) - log(resident\ tuition_{st}), & Reciprocity_{sdt} = 1 \\ log(nonresident\ tuition_{dt}) - log(resident\ tuition_{st}), & Reciprocity_{sdt} = 0 \end{cases}$$

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.4: Effect of total reciprocity agreements on college interstate migration

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Total Reciprocity	0.22***	0.22***	0.29***	0.08	0.10*
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Observations	9184	9184	9184	9184	6285
R-squared (within)	0.39	0.42	0.47	0.68	0.13
Log gross state product (source&destination)	-	YES	YES	YES	YES
Log state&local education expenditure (source&destination)	-	YES	YES	YES	YES
Unemployment rate (source&destination)	-	YES	YES	YES	YES
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. The sample is college students aged under 30 who study in public school from Census year 1960-2000. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

Table 1.5: Effect of regional and bilateral reciprocity agreements on college interstate migration

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Regional Reciprocity	0.20***	0.19***	0.25***	0.06	0.09*
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Bilateral Reciprocity	0.80***	0.95***	1.04***	0.72***	0.52***
	(0.22)	(0.25)	(0.23)	(0.22)	(0.13)
Observations	9,184	9,184	9,184	9,184	6,285
R-squared (within)	0.40	0.42	0.47	0.68	0.13
Log gross state product (source&destination)	-	YES	YES	YES	YES
Log state&local education expenditure (source&destination)	-	YES	YES	YES	YES
Unemployment rate (source&destination)	-	YES	YES	YES	YES
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. The sample is college students aged under 30 who study in public school from Census year 1960-2000. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

Table 1.6: Effect of reciprocity agreements on college interstate migration with border interaction terms

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Total Reciprocity	0.10**	0.10**	0.15***	0.01	0.04
	(0.05)	(0.04)	(0.04)	(0.07)	(0.06)
Total Reciprocity*Bordering	0.40***	0.40***	0.44***	0.22**	0.24***
	(0.06)	(0.07)	(0.07)	(0.10)	(0.08)
Observations	9,184	9,184	9,184	9,184	6,285
R-squared (within)	0.40	0.42	0.48	0.68	0.13
Log gross state product (source&destination)	-	YES	YES	YES	YES
Log state&local education expenditure (source&destination)	-	YES	YES	YES	YES
Unemployment rate (source&destination)	-	YES	YES	YES	YES
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. The sample is college students aged under 30 who study in public school from Census year 1960-2000. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

1.7 Appendix. Additional Tables and Figures

Table 1.A1: Effects of tuition gaps on college interstate migration-fixed effects with detailed coefficients

	(1)	(2)	(3)	(4)
Log tuition gap	0.04	0.04	-0.02	0.02
	(0.03)	(0.03)	(0.05)	(0.08)
Log GSP (destination)		0.40***	0.68***	0.88***
		(0.09)	(0.19)	(0.26)
Log GSP*dummy1997 (destination)		-0.05***	-0.08***	-0.08**
		(0.02)	(0.03)	(0.04)
Log GSP (source)		0.02	-0.72***	-0.94***
		(0.08)	(0.21)	(0.28)
Log GSP*dummy1997 (source)		-0.01	-0.02	-0.02
		(0.02)	(0.03)	(0.04)
Log state&local education expenditure (destination)		0.16	0.33	0.18
		(0.12)	(0.20)	(0.26)
Log state&local education expenditure (source)		0.13	-0.45**	-0.45*
		(0.09)	(0.20)	(0.26)
Unemployment rate (destination)		-0.00	0.00	0.01
		(0.01)	(0.02)	(0.02)
Unemployment rate (source)		0.06***	0.04**	0.03
		(0.01)	(0.02)	(0.02)
R-squared (within)	0.03	0.05	0.13	0.62
Source& destination state-specific trends?	-	-	YES	-
Pair-specific trends?	-	-	-	YES
P value: F-statistic of trends	-	-	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. The sample is college students aged under 30 who study in public school from census year 1980-2000. 6285 observations. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.A2: Effects of tuition gaps on college interstate migration-first stage with detailed coefficients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	First stage using census years				First stage using annual data			
Coefficient on reciprocity 1st stage	-0.36***	-0.39***	-0.26***	-0.17***	-0.32***	-0.31***	-0.25***	-0.31***
	(0.03)	(0.03)	(0.01)	(0.04)	(0.01)	(0.01)	(0.01)	(0.02)
Log GSP (destination)		0.21***	0.38***	0.40***		-0.04**	0.23***	0.22***
		(0.04)	(0.07)	(0.09)		(0.02)	(0.03)	(0.03)
Log GSP*dummy1997 (destination)		0.00	0.01	0.01		0.01***	0.00	0.00
		(0.01)	(0.01)	(0.02)		(0.00)	(0.00)	(0.00)
Log GSP (source)		-0.70***	-0.39***	-0.37***		-0.37***	-0.53***	-0.53***
		(0.05)	(0.08)	(0.10)		(0.02)	(0.03)	(0.03)
Log GSP*dummy1997 (source)		-0.07***	0.11***	0.11***		-0.05***	0.02***	0.02***
		(0.01)	(0.01)	(0.02)		(0.00)	(0.00)	(0.00)
Log state&local education expenditure (destination)		-0.23***	0.40***	0.41***		-0.01	-0.14***	-0.15***
		(0.04)	(0.07)	(0.10)		(0.02)	(0.02)	(0.02)
Log state&local education expenditure (source)		0.67***	-0.32***	-0.28**		0.22***	0.26***	0.26***
		(0.06)	(0.09)	(0.12)		(0.02)	(0.02)	(0.02)
Unemployment rate (destination)		0.00	0.01*	0.01		0.01***	0.01***	0.01***
		(0.01)	(0.01)	(0.01)		(0.00)	(0.00)	(0.00)
Unemployment rate (source)		-0.03***	0.03***	0.02***		-0.02***	-0.02***	-0.02***
		(0.01)	(0.01)	(0.01)		(0.00)	(0.00)	(0.00)
Observations 1st stage	6285	6285	6285	6285	81,500	81,500	81,500	81,500
R-squared (within)	0.66	0.70	0.89	0.92	0.59	0.62	0.70	0.71
Source& destination state-specific trends?	-	-	YES	-	-	-	YES	-
Pair-specific trends?	-	-	-	YES	-	-	-	YES
F-statistic 1st stage	166	248	429	18	572	667	672	326

*Notes: The dependent variable is the Tuition Gap. The sample is college students aged under 30 who study in public school from census year 1980-2000. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The instrument is a dummy for a pair of states' memberships in a common regional reciprocity agreement. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.A3: Effects of tuition gaps on college interstate migration-Second stage with detailed coefficients

	(1)	(2)	(3)	(4)
Log tuition gap	-0.49***	-0.40***	-0.36*	0.80
	(0.13)	(0.12)	(0.19)	(0.67)
Log GSP (destination)		0.49***	0.80***	0.57
		(0.09)	(0.21)	(0.37)
Log GSP*dummy1997 (destination)		-0.05***	-0.08***	-0.08**
		(0.02)	(0.03)	(0.04)
Log GSP (source)		-0.28**	-0.85***	-0.64
		(0.12)	(0.23)	(0.40)
Log GSP*dummy1997 (source)		-0.03*	0.01	-0.11
		(0.02)	(0.03)	(0.08)
Log state&local education expenditure (destination)		0.06	0.49**	-0.17
		(0.12)	(0.22)	(0.41)
Log state&local education expenditure (source)		0.43***	-0.54***	-0.25
		(0.12)	(0.21)	(0.31)
Unemployment rate (destination)		-0.00	0.01	-0.00
		(0.01)	(0.02)	(0.02)
Unemployment rate (source)		0.04***	0.05***	0.01
		(0.01)	(0.02)	(0.03)
R-squared (within)	0.03	0.04	0.12	0.58
Source& destination state-specific trends?	-	-	YES	-
Pair-specific trends?	-	-	-	YES
P value: F-statistic of trends	-	-	0.00	0.00
Coefficient on reciprocity 1st stage	-0.36***	-0.39***	-0.26***	-0.17***
	(0.03)	(0.03)	(0.01)	(0.04)
Observations 1st stage	6285	6285	6285	6285
F-statistic 1st stage	166	248	429	18
R-squared 1st stage (within)	0.66	0.70	0.89	0.92

*Notes: The dependent variable is the log number of college students migrating between pairwise states. The sample is college students aged under 30 who study in public school from census year 1980-2000. 6285 observations. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The instrument is a dummy for a pair of states' memberships in a common regional reciprocity agreement. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.A4: Effects of tuition gaps on college interstate migration-restricted sample

	(1)	(2)	(3)	(4)
Log tuition gap*	-0.40***	-0.33***	-0.22	0.91
	(0.12)	(0.11)	(0.26)	(0.84)
R-squared (within)	0.01	0.01	0.13	0.57
Log gross state product (source&destination)	-	YES	YES	YES
Log state&local education expenditure (source&destination)	-	YES	YES	YES
Unemployment rate (source&destination)	-	YES	YES	YES
Source& destination state-specific trends?	-	-	YES	-
Pair-specific trends?	-	-	-	YES
P value: F-statistic of trends	-	-	0.00	0.00
Coefficient on reciprocity 1st stage	-0.39***	-0.42***	-0.23***	-0.14***
	(0.03)	(0.03)	(0.02)	(0.05)
Observations 1st stage	4084	4084	4084	4084
F-statistic 1st stage	158	240	224	11
R-squared 1st stage (within)	0.62	0.67	0.88	0.92

*Notes: The dependent variable is the log number of college students migrating between pairwise states. 4084 observations. The sample is college students aged under 30 who study in public school from Census 1980-2000. This sample is restricted to those agreements where 75% of members joined all at once and have more than two states in them. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The instrument in columns 5-8 is reciprocity, a dummy for pairwise states having a reciprocity agreement with each other since agreement starting years. And the first stage uses census years. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1980-2000& IPEDS 1980, 1984-2017& BLS 1976-2017& BEA 1962-2017.

Table 1.A5: Effect of total reciprocity agreements on college interstate migration with total reciprocity-detailed

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Total Reciprocity	0.22***	0.22***	0.29***	0.08	0.10*
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Log GSP (destination)		0.32***	0.26***	0.37***	0.63***
		(0.05)	(0.09)	(0.11)	(0.21)
Log GSP*dummy1997 (destination)		0.02	-0.11***	-0.12***	-0.09***
		(0.02)	(0.02)	(0.02)	(0.03)
Log GSP (source)		0.30***	-0.32***	-0.41***	-0.56**
		(0.05)	(0.08)	(0.11)	(0.22)
Log GSP*dummy1997 (source)		0.05***	-0.07***	-0.06***	-0.03
		(0.02)	(0.02)	(0.02)	(0.03)
Log state&local education expenditure (destination)		0.29***	0.11*	0.10	0.35*
		(0.05)	(0.06)	(0.07)	(0.21)
Log state&local education expenditure (source)		0.05	0.11**	0.11**	-0.47**
		(0.04)	(0.05)	(0.06)	(0.20)
Unemployment rate (destination)		0.84	-0.18	0.30	-0.51
		(1.00)	(1.08)	(1.26)	(1.82)
Unemployment rate (source)		2.96***	3.50***	2.90**	6.04***
		(0.95)	(1.09)	(1.29)	(1.81)
Observations	9,184	9,184	9,184	9,184	6,285
R-squared (within)	0.39	0.42	0.47	0.68	0.13
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. 9184 observations. The sample is college students aged under 30 who study in public school from census year 1960-2000. Besides the four regional reciprocity agreements, this regressions also include other smaller neighbor states' reciprocity agreements. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The 1960 gross state product is predicted using census years from 1970 to 2000. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

Table 1.A6: Effect of regional and bilateral reciprocity agreements on college interstate migration-detailed

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Regional Reciprocity	0.20***	0.19***	0.25***	0.06	0.09*
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Bilateral Reciprocity	0.80***	0.95***	1.04***	0.72***	0.52***
	(0.22)	(0.25)	(0.23)	(0.22)	(0.13)
Log GSP (destination)		0.32***	0.25***	0.37***	0.64***
		(0.05)	(0.09)	(0.11)	(0.21)
Log GSP*dummy1997 (destination)		0.02	-0.12***	-0.12***	-0.09***
		(0.02)	(0.02)	(0.02)	(0.03)
Log GSP (source)		0.31***	-0.32***	-0.41***	-0.55**
		(0.05)	(0.08)	(0.11)	(0.22)
Log GSP*dummy1997 (source)		0.05***	-0.07***	-0.06***	-0.03
		(0.02)	(0.02)	(0.02)	(0.03)
Log state&local education expenditure (destination)		0.30***	0.11*	0.10	0.35*
		(0.05)	(0.06)	(0.07)	(0.21)
Log state&local education expenditure (source)		0.05	0.11**	0.12**	-0.47**
		(0.04)	(0.05)	(0.05)	(0.20)
Unemployment rate (destination)		0.94	-0.13	0.32	-0.52
		(1.00)	(1.08)	(1.26)	(1.82)
Unemployment rate (source)		3.06***	3.54***	2.93**	6.04***
		(0.94)	(1.09)	(1.29)	(1.81)
Observations	9,184	9,184	9,184	9,184	6,285
R-squared (within)	0.40	0.42	0.47	0.68	0.13
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. 9184 observations. The sample is college students aged under 30 who study in public school from census year 1960-2000. Besides the four regional reciprocity agreements, this regressions also include other smaller neighbor states' reciprocity agreements. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The 1960 gross state product is predicted using census years from 1970 to 2000. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

Table 1.A7: Effect of reciprocity agreements on college interstate migration with border interaction terms-detailed

	(1)	(2)	(3)	(4)	(5)
		1960-2000			1980-2000
Total Reciprocity	0.10**	0.10**	0.15***	0.01	0.04
	(0.05)	(0.04)	(0.04)	(0.07)	(0.06)
Total Reciprocity*Bordering	0.40***	0.40***	0.44***	0.22**	0.24***
	(0.06)	(0.07)	(0.07)	(0.10)	(0.08)
Log GSP (destination)		0.32***	0.24***	0.36***	0.63***
		(0.05)	(0.09)	(0.11)	(0.21)
Log GSP*dummy1997 (destination)		0.02	-0.11***	-0.12***	-0.09***
		(0.02)	(0.02)	(0.02)	(0.03)
Log GSP (source)		0.30***	-0.33***	-0.41***	-0.56**
		(0.05)	(0.08)	(0.11)	(0.22)
Log GSP*dummy1997 (source)		0.05***	-0.07***	-0.06***	-0.02
		(0.02)	(0.02)	(0.02)	(0.03)
Log state&local education expenditure (destination)		0.29***	0.10*	0.10	0.35*
		(0.05)	(0.06)	(0.07)	(0.21)
Log state&local education expenditure (source)		0.05	0.11**	0.11**	-0.47**
		(0.04)	(0.05)	(0.06)	(0.20)
Unemployment rate (destination)		0.86	-0.25	0.26	-0.54
		(1.00)	(1.08)	(1.27)	(1.82)
Unemployment rate (source)		2.97***	3.44***	2.87**	6.02***
		(0.94)	(1.09)	(1.29)	(1.82)
Observations	9,184	9,184	9,184	9,184	6,285
R-squared (within)	0.40	0.42	0.48	0.68	0.13
Source& destination state-specific trends?	-	-	YES	-	YES
Pair-specific trends?	-	-	-	YES	-
P value: F-statistic of trends	-	-	0.00	0.00	0.00

*Notes: The dependent variable is the log number of college students migrating between pairwise states. 9184 observations. The sample is college students aged under 30 who study in public school from census year 1960-2000. Besides the four regional reciprocity agreements, this regressions also include other smaller neighbor states' reciprocity agreements. Standard errors clustered by state pair in parentheses. All regressions include year and state pair dummies. The 1960 gross state product is predicted using census years from 1970 to 2000. The coefficients of controls are in the appendix. *** p<0.01, ** p<0.05, * p<0.1*

Source: Census 1960-2000& BLS 1976-2017& BEA 1962-2017.

Figure 1.1: State migration trends in MSEP

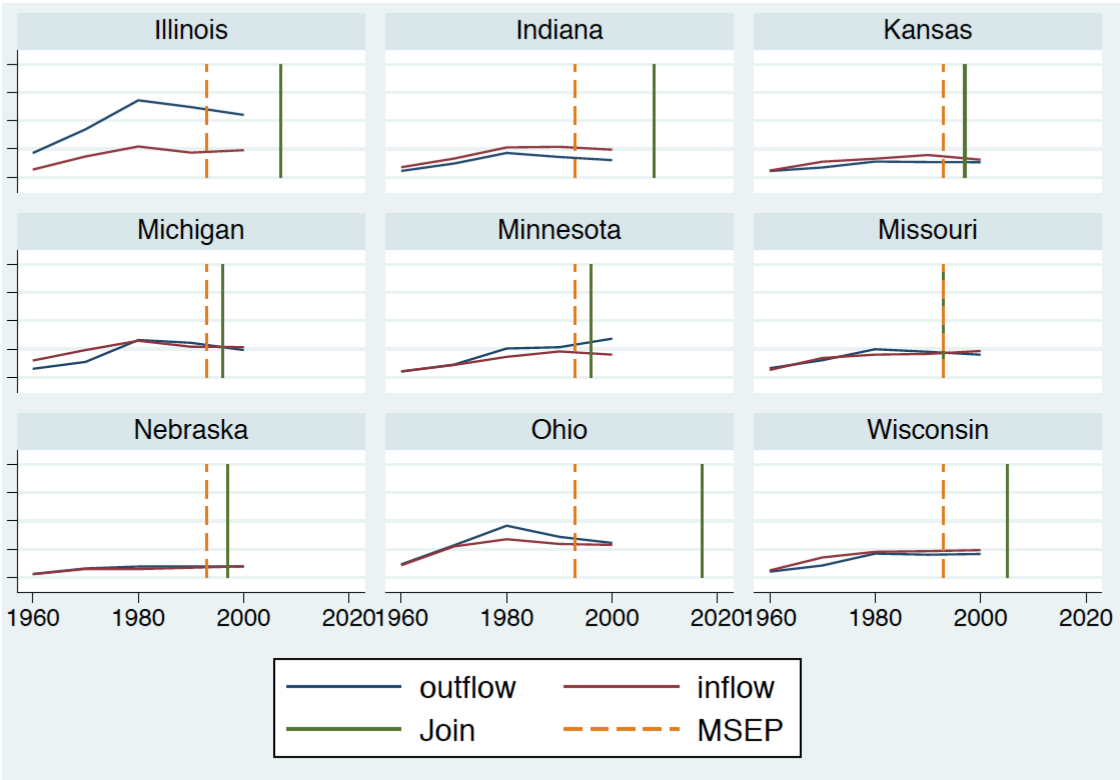


Figure 1.2: State migration trends in ACM

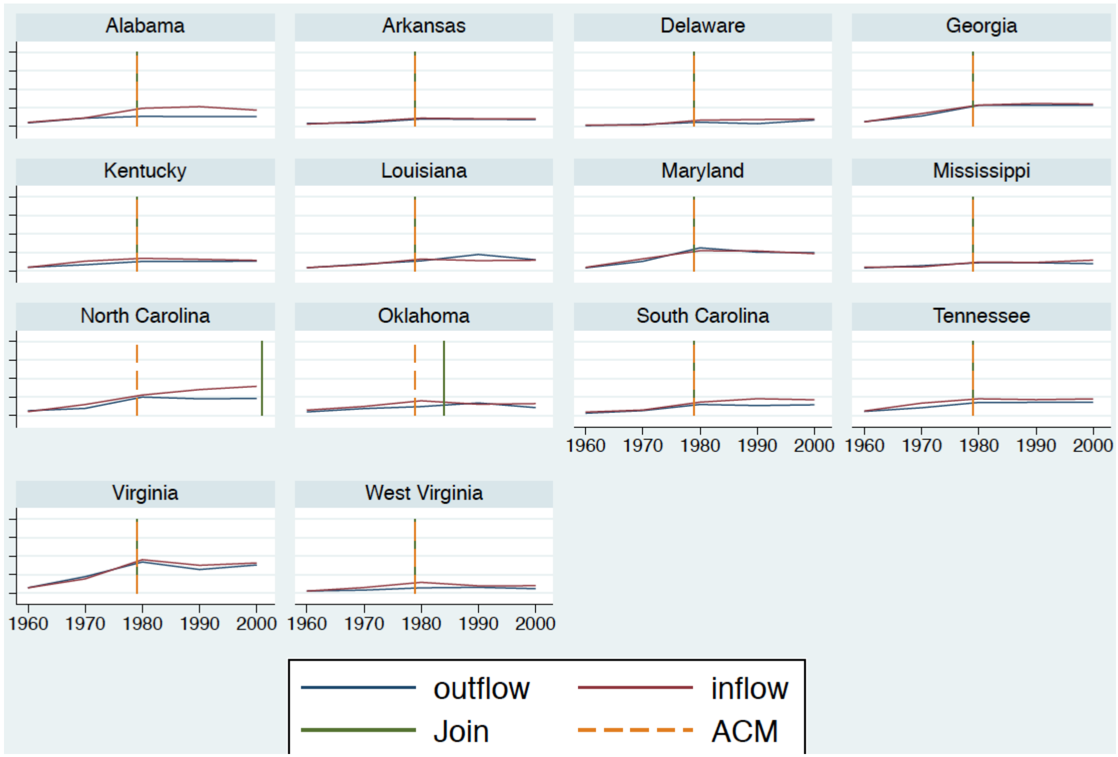


Figure 1.3: State migration trends in WUE

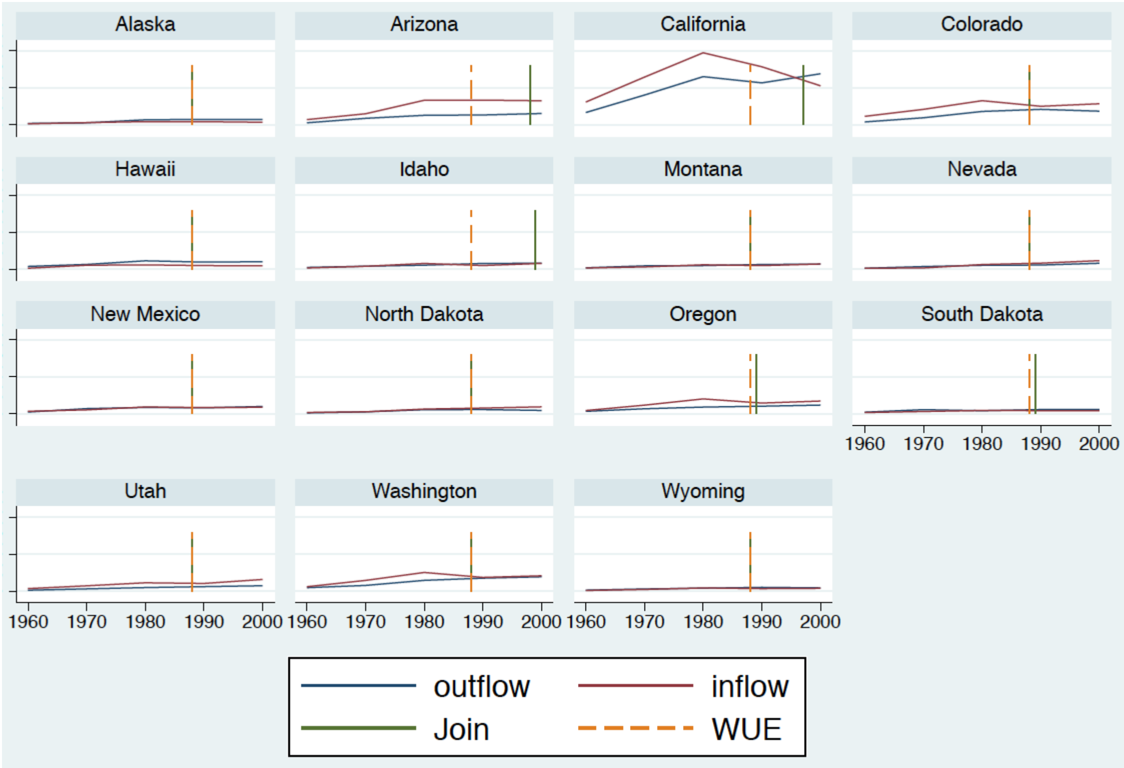


Figure 1.4: State migration trends in NERSP

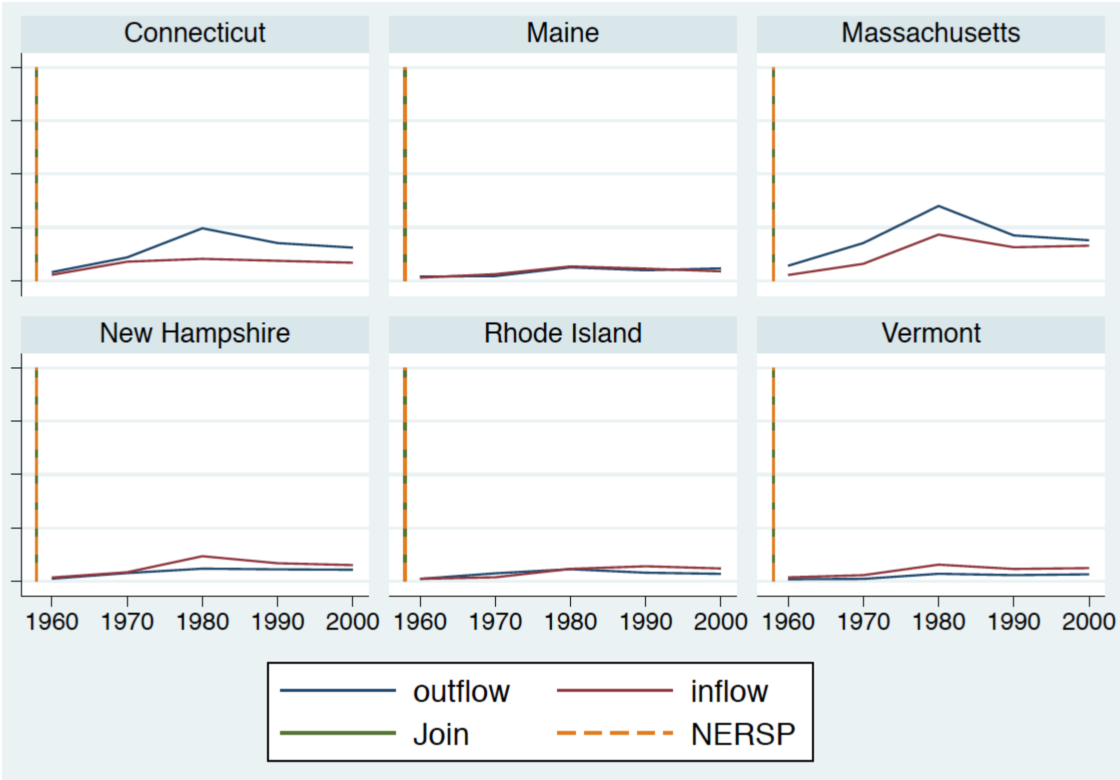
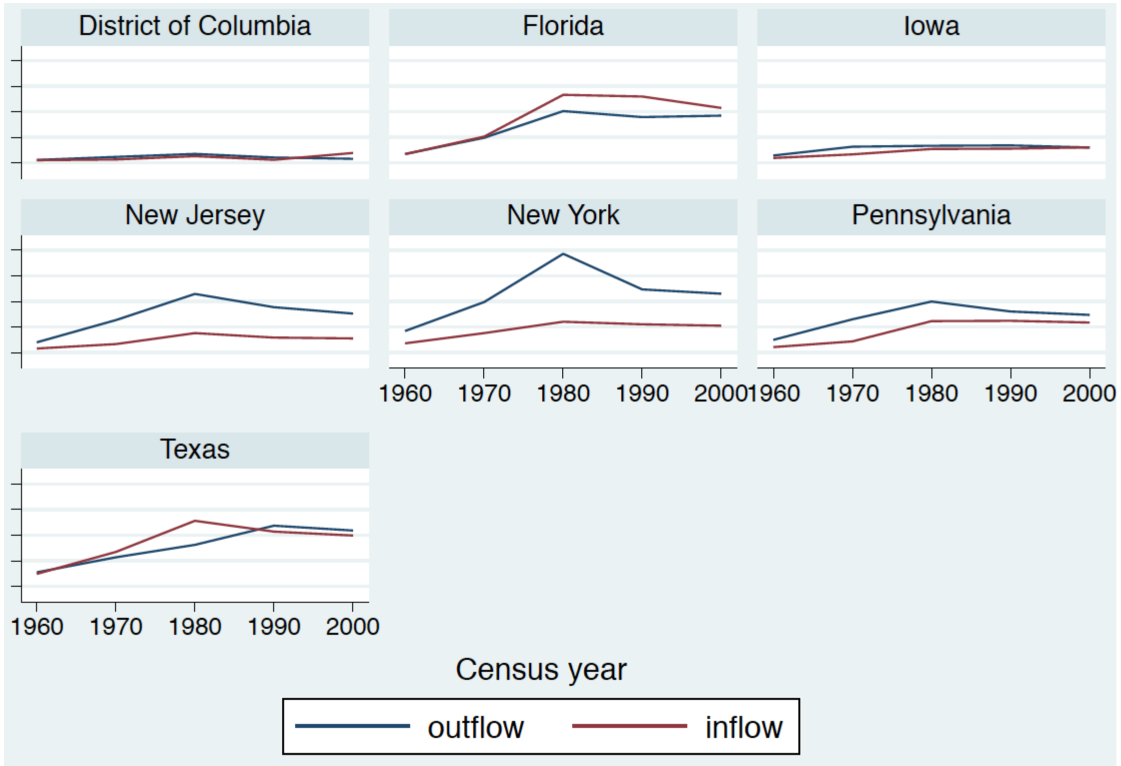


Figure 1.5: State migration trends in states without any agreement



Chapter 2

Does Customer or Coworker Discrimination Prompt Employers to Advertise for Attractive Employees?

2.1 Introduction

Unconstrained by laws governing labor market discrimination, Chinese employers are free to post ads specifying applicants should be beautiful, tall, short, young, old, male or female. Chinese data may therefore be used to study the sources of employer discrimination in a way not possible elsewhere, while still possibly shedding light on global phenomena. Employment discrimination occurs when employees or job applicants are treated differently on the basis of the group (female, Black, old, ugly, immigrants, etc.) with otherwise similar characteristics and in similar circumstances. Economists often classify discrimination into one of two major types: statistical discrimination and taste discrimination. Statistical discrimination occurs when in the presence of imperfect information, firms make group-based inference based on statistical information.[1] For example, American employers may infer that a Black applicant is more likely than a White applicant to have a criminal record based on disproportionate criminal records among Blacks.[2] Statistical discrimination can help firms take advantage of the group characteristic information to reduce search costs and find bet-

[1] Arrow, K. (1973); Phelps, E. S. (1972)
[2] Agan and Starr (2017)

ter matched employees faster and more easily. However, it would also cause some problems like discouraging the disfavored group from participating in the market, therefore lowering the market efficiency, and may be viewed as unfair.[3]

Taste discrimination occurs when members of one group have a taste or prejudice against another group.[4] Taste discrimination can be further divided into employer taste discrimination, customer taste discrimination and coworker taste discrimination. Employer taste discrimination in a perfectly competitive market with no frictions should be inefficient because employers do not employ disfavored workers to the point where their wage equals their marginal product, meaning that prejudice reduces profits.[5] However, customer or coworker discrimination can be profitable for employers. If many customers or employees are members of the favored group and dislike interacting with employees from the disfavored group, restricting hiring to the favored can increase productivity and sales.[6] Customer and coworker discrimination can thus exist even in the competitive equilibrium, which renders ineffective solutions, such as, enhancing competition or reducing frictions in markets, that are effective for fixing employer taste discrimination.[7] Therefore, different types of discrimination have very different economic consequences. Figuring out the source of discrimination is not only important for fairness but also necessary for market efficiency.

Disentangling types of discrimination has proved difficult empirically. In this paper, I focus on employer demand for workers who are beautiful or tall, and estimate the degree to which customer or coworker discrimination is responsible for this demand. I use data created by Kuhn and Shen (2012) from a Chinese online job board, which contains rich information about both the explicit demographic requirements, such as, age, gender, etc. and relevant skills asked in job ads, such as, education, experience, and the firm information such as size, ownership, etc. I merge these with

---

[3] Kenneth Arrow (1972) and George Borjas and Matthew Goldberg (1978) their models both assume that those firms taking fully use of the productivity information indicated by group are more efficient than those firms ignoring the information.

[4] Becker (1957)

[5] Becker (1971) shows that when the supply of workers in the favored group is less than the demand from prejudiced firms, the wage received by the disfavored group in a prejudiced firm cannot fully compensate the productivity, therefore, it is inefficient for both the firms and the labor market.

[6] Goldberg (1982)

[7] Becker (1971); Nardinelli and Simon (1990); Black and Strahan (2002)

four indices representing customer and coworker contact based on occupation tasks description data from the American O*NET website.

I assume that if the employer appearance requirements are motivated by customer or coworker discrimination, firms are more likely to worry about the attractiveness of workers for jobs that have more contact with customers or coworkers. Thus, they are more likely to post beauty or height requirements in these job ads.[8] If employer prejudice is the only reason for appearance attractiveness requirements, employers would favor attractive workers for all jobs, no matter whether face-to-face contact tasks are required. Thus, the posting of beauty requirements in job ads would not vary a lot across jobs with different degree of contact with customers or coworkers. Furthermore, if part of the beauty or height requirements is caused by statistical discrimination, which means that firms like attractive workers because firms take beauty or height as an indicator of high productivity, we can assume that the more information about the productivity of applicants revealed to firms, the less emphasis they would put on their "beauty" or "height" indicator and thus, thus less likely to post beauty requirements in job ads. In other words, those job ads with more education and experience requirements are less likely to have beauty or height requirements. The importance of customer versus coworker contact can be judged either by comparing the size of the coefficients on the customer related indices versus the coefficients on the coworker related indices or by comparing their contributions to the R-squared, and I assess both.

The main result of this paper is that employers are more likely to require beauty for jobs with more direct contact with customers and require height for jobs with more contact with coworkers. And customer-related contact indices contribute more to both the beauty requirement variance and the height requirement variance. These results suggest that the employer advertisements requiring attractive employees are principally motivated by customer discrimination and coworker discrimination is the principal motivation of the height requirement. In addition, the beauty requirement of ads targeting females are mostly prompted by customer discrimination while the

[8] Holzer and Ihlanfeildt (1998) find that the customer racial composition has a large effect on the race of hired employees, Specifically for those jobs with direct contact with customers.

beauty requirement of ads targeting males are mostly prompted by coworker discrimination. But there is not much difference between genders for the height requirement.

I contribute to the literature in two understudied areas: testing types of discrimination and exploring the determinants of employer advertisements. Firstly, most studies in the literatures focus on distinguishing statistical discrimination from taste discrimination. Some (List, 2004; Knowles, Persico and Todd, 2005; Zussman, 2013) find discrimination is explained by statistical discrimination, some (Sanga, 2009; Kerwin and Guryan, 2013) find that it is explained by taste discrimination model, and some (Agan and Starr, 2017) find it is a combination of both. However, so far economists have done little work on breaking down the sources of taste discrimination. My paper does this, disentangling coworker discrimination and customer discrimination. Kuhn and Shen (2009), the most closely related paper to mine, find that cross-sectional patterns in job ads suggest some role for customer discrimination. They test for customer discrimination by establishing a dummy variable equal to one if they subjectively consider the respondent's occupation to involve high customer contact. The contact indices created in my paper are more objective and accurate and can provide more information. Furthermore, my indices make it possible for me to test for coworker discrimination in addition to customer discrimination and compare the importance of these two discrimination types. Stinebrickner, Stinebrickner and Sullivan (2018), another study that is close to mine, tease out the employer taste discrimination from others in beauty wage premium by using job tasks information from the Berea Panel Study, and they find that the wide variation of beauty premia across jobs which cannot be explained by employer taste discrimination might be explained by customer or coworker discrimination, but since the survey lacks information about the detailed sources of interaction in jobs, they admit that their paper is unable to further disentangle customer discrimination from coworker discrimination.

Secondly, my work can help to better understand employer advertisements and establish some new facts about the explicit beauty requirements in job ads.[9] Kuhn

[9] Hamermesh and Biddle (1994) find a large beauty premia exists by studying employer advertisements from two household surveys: The 1977 Quality of Employment Survey (QES) and the 1971 Quality of American Life survey (QAL).

and Shen (2012) and Helleseter, Kuhn and Shen (2016) also study the employer explicit preferences in advertisements. They find that employers' valuations are highly specific to detailed jobs and occupations and they focus on job skill level to explore employer advertisements, while my study focuses on the degree of contact in jobs to explore the employer advertisements.

A caveat is that the impact of discrimination found in the job posting stage could be quite different from the wage offering stage. Kuhn and Shen (2012) point out that when posting a requirement for a particular feature in a job ad, employers are excluding the whole group of people with that feature from being considered for the position even without seeing their resumes, which reflects a strong and conscious prejudice towards that group. Therefore, it would be different from the discrimination detected in other stages, which is a combination of both unconscious and conscious choices of the employers.

2.2 Data

The main data source is the Chinese online job advertisement data collected by Kuhn and Shen (2012). Kuhn and Shen (2012) crawled and scraped all the unique job ads from zhaopin.com, the third largest Chinese online job board, during four observation periods: May 19th, 2008 – June 22th, 2008; Jan 19th, 2009 – Feb 22th, 2009; May 18th, 2009 – June 21th, 2009; Jan 18th, 2010 – Feb 21th, 2010 and built the ads dataset of 1,322,671 observations.[10] This dataset includes several key variables reflecting the physical requirements in job ads, for example, "Any beauty requirement": is there any beauty requirement in the ad; "Any height requirement": is there any height requirement in the ad. It is worth noting that the beauty dummy refers to the overall image of a person rather than just "facial attractiveness", while the "Any height requirement" dummy in the data refers to whether or not an ad specifies of height. (e.g. at least 170cm). Since the main focus is on occupation information, I exclude observations whose occupation classification is "student" or "others", which counts for 1.17% of

[10] The original dataset and more information about the procedures of collecting data can be found on Professor Kuhn's personal website: https://sites.google.com/view/peter-kuhn/

the sample, and study the remaining 1,300,118 observations.

Table 2.1 shows means of requirements in job ads. Column 1 shows means of requirements for whole sample. Among 1,300,118 ads studied in this paper, 105,078 (8.2%) of them have a beauty requirement and 2.9% have a height requirement. Columns 2 and 3 show more details of how different kinds of requirements in job ads vary by beauty requirement. Firstly, ads with a beauty requirement are 18.6 percentage points more likely to also have a height requirement. Secondly, ads with a beauty requirement are 19.9 percentage points more likely to also have any age requirement, and the preferred age is around 3.8 years younger than in ads without a beauty requirement. Thirdly, Ads with a beauty requirement are 20.9 percentage points more likely to also require workers to be female, however, there is not much difference for the requirement to be male. Fourthly, ads with a beauty requirement are more likely to also ask for relatively less education, such as no school restrictions, junior middle school or below, high school and post-secondary, while ads with a beauty requirement are less likely to also ask for a higher education degree, such as undergraduate, master and Ph.D. Ads with a beauty requirement require around 1.16 more years of working experience than ads with no beauty requirement. Finally, ads with a beauty requirement are less like to be part-time jobs.

Columns 4 and 5 in Table 2.1 show details of different kinds of requirements on job ads vary by height requirement. Most statistics for height are similar to those for beauty. The only differences lie in the gender and part-time jobs requirements. Firstly, ads with a height requirement are more like to ask workers to be male, 8.4 percentage points more than ads without a height requirement, while there is no such difference as for beauty requirement. Besides, ads with a height requirement are more likely to be part-time jobs.

To measure the degree of contact for each occupation category in the Chinese job ads data, I adopt the following three steps.

The first step is to merge the Chinese job ads data with the American O*NET database by manually creating a crosswalk between these two datasets as shown in Appendix Table 2.A1. Note that since the unit of observation in the O*NET data is the industry-specific job title, each job category in the Chinese job ads data is matched to a set of O*NET job titles.[11] For each job title in the O*NET database, the Bureau of Labor Statistics (BLS) reports 1000 standardized,occupation-specific descriptors. I use these descriptors to generate the contact indices measures as described in the second step.[12]

The second step is to generate the contact indices for each job title in the O*NET data. For each job title, I use four descriptors under "work context" content that are relevant to this study.[13] The first descriptor is "Communicating with Persons Outside the Organization", which is defined as "Communicating with people outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, or by telephone or e-mail." The second descriptor is "Performing for or Working Directly with the Public", which is defined as "Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores and receiving clients or guests." The third descriptor is "Communicating with Supervisors, Peers or Subordinates", which is defined as "Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person." The fourth descriptor is "Face-to-Face Discussions", which is defined as "How often do you have to have face-to-face discussions with individuals or teams in this job." Each descriptor is rated by two dimensions: level and importance. Since the level and importance dimensions reflect two different aspects of the same descriptor, these two dimensions are combined to give a complete measure of the contact. I use the Cobb-Douglas function to combine them:[14]

---

[11] The O*NET data is created and maintained by the U.S. Bureau of Labor Statistics (BLS). It contains 969 industry-specific U.S. job titles. In contrast, the Chinese job ads data has 37 aggregated job categories. More information about the O*NET database can be found at https://www.onetonline.org/help/onet/

[12] The occupation data in American O*NET mainly focus on U.S. labor market. However, since there is no such data for the Chinese labor market and most occupations have similar characteristics and tasks around the world, the American O*NET data is a relatively good choice.

[13] Work context is defined as physical and social factors that influence the nature of work in the O*NET.

[14] I have tried different functions to combine these two dimensions: sum, product and Cobb-Douglas. The results have not much different, so I finally decide to use the Cobb-Douglas function. The parameters of Cobb Douglas function come from Firpo, Sergio, Nicole Fortin, and Thomas Lemieux (2011).

$$Degree\ of\ customer\ contact_j = \ Level_j^{1/3} Importance_j^{2/3}$$

which is also commonly used in the literature (Firpo, Sergio, Nicole Fortin, and Thomas Lemieux, 2011; Goos, 2011; Lee and Shin, 2017). The two indices based on "Communicating with Persons Outside the Organization" and "Performing for or Working Directly with the Public" are treated as a potential motivation for customer discrimination, and the two indices based on "Communicating with Supervisors, Peers or Subordinates" and "Face-to-Face Discussions" are treated as a potential motivation for coworker discrimination. To facilitate interpretation, I standardize the indices.

The third step is to create a weight for each O*NET job title and construct the contact index for each job category in the Chinese job ads data by calculating the weighted average contact index value of job titles under the same Chinese job category. Specifically, for each Chinese job category j, I denote the set of US job titles matched to j to be I_j. Assuming that the number of job ads posted online is proportionate to the number of jobs actually occupied, the weight of each job title i matched to Chinese job category j is given by its employment share:[15]

$$Weight_{i(j)} = \frac{Employment_i}{\sum_{k \in I_j} Employment_k}$$

Then, the contact index for job category j is given by:

$$ContactIndex_{ChineseJobCategory_j} = \sum_{k \in I_j} Weight_{k(j)} \times ContactIndex_{USJobTitle_k}$$

where $ContactIndex_{USJobTitle_k}$ is the person contact indices for the US job titles calculated in step 2.

The means of the standardized job contact indices by beauty and height requirement are presented in Table 2.2. Column 1 and 2 in Table 2.2 show that most of the contact indices have a higher score for the ads with a beauty requirement than those

[15] The U.S. employment information is from the U.S. Occupational Employment Statistics (OES) dataset: https://www.bls.gov/oes/tables.htm

ads without a beauty requirement, and the gap is larger for "working with the public". For example, the difference between ads with a beauty requirement and ads without a beauty requirement for "working with the public" (0.6) is larger than the difference for "communication with coworker" and "face-to-face discussion" (0.1 and 0.2 for each). While column 3 and 4 in Table 2.2 show that most of the contact indices (except for "working with the public") have a lower score for ads with a height requirement than those ads without, but the gap is not large.

The means of the component parts are shown in Appendix Table 2.A2. In addition, I also checked the correlations among all the contact indices and all other variables, and the result is presented in Appendix Table 2.A3. It shows that most correlations among the four contact indices are less than 0.5 and the correlation of the contact indices with other variables are even lower. For example, a typical job with a high "communication with customer" but a low "working with the public" would be "manager of sales worker", while a typical job with the reversed case is "judges". And a typical job with a high "communication with coworker" but a low "face-to-face discussion" would be "manufactured building installer", while the reversed case is "skin care specialist".

2.3 Methodology

I use linear probability regressions to test the effect of different types of interpersonal contact. The main regression equation is:

$$
\begin{aligned}
Y_{ijf} &= \beta_0 + \beta_1 \textit{Communication with customer}_j + \beta_2 \textit{Working with the public}_j \\
&+ \beta_3 \textit{ Communication with coworker}_j + \beta_4 \textit{Face-to-face discussion}_j + \ \beta_5 x_{ijf} + \gamma_f + \varepsilon_{ijf}
\end{aligned}
$$

where Y is either a beauty or a height requirement dummy, i indexes the advertisement, j indexes the occupation of the position in the advertisements and f indexes the firm posting the advertisements. x are control variables, includes: the age dummy, the gender dummy, the education dummy, the experience years, the part-time jobs dummy, period dummies and number of ads. γ stands for firm fixed effects. ε is the error term. Since the first two variables "Communication with customer" and "Working with the public" are treated as a potential motivation for customer discrimination, the coefficients $\beta_1, \ \beta_2$ together can be viewed as the effects of customer discrimination on the probability of having a beauty requirement in job ads. The latter two variables "Communication with coworker" and "Face-to-face discussion" are treated as a potential motivation for coworker discrimination, so coefficients $\beta_3, \ \beta_4$ together can be viewed as the effects of coworker discrimination on the probability of having a beauty requirement in job ads.

I will assess the importance of customer and coworker discrimination in two ways, considering first the coefficients associated with the four indices and second the contribution of the indices to the R-squared. The comparison of coefficients of both customer indices and both coworker indices reflect which one has a larger effect on the probability of having a beauty requirement in job ads and the comparison of their ANOVA results reflect which one contributes more to the variation of the probability of having a beauty requirement in job ads.

Considering all the requirements in job ads are highly correlated with each other and could be caused by a same factor, I also do seemingly unrelated regression (SUR) in order to test formally whether the contact indices have the same effects on the beauty, height, age, and gender requirements. The SUR regressions take the same form as the main equation, but the estimation allows the error terms to be correlated.

2.4 Results

2.4.1 Main regression

To check the relationship of the beauty requirement in job ads and occupations' contact indices and then further uncover whether employer advertisements for attractive employees is prompted by customer or coworker discrimination, Table 2.3 presents regression results for the determinants of the beauty requirement. Nine specifications with increasingly detailed controls are presented across columns from left to right. All the regressions include all four job contact indices, which are the main variables of interest in this paper. The first specification (column 1) controls for only these variables of interest. In columns 2 through 9, I gradually add first the age requirement dummy; next the education requirement dummy and experience years requirement; then the part-time jobs requirement dummy, period dummies, number of ads, and firm fixed effects; then the gender requirement dummy and finally the interaction terms.

Column 1 in Table 2.3a shows that among the four indices of interest, only the coefficient of the index “working with the public” is statistically significant, it is positive and shows that a one standard deviation increase in “working with the public” would increase the probability of having a beauty requirement in a job ad by 0.06 percentage points (on average, 8% of ads have a beauty requirement, so this represents a 0.75% increase on the mean). The other three indices are small and statistically insignificant. Overall, referring to the difference of the definitions in the four indices, a comparison of their coefficients in this column indicates that only the type of direct contact with the public in the job requirement matters when employer deciding whether to post a beauty requirement in the job ad.

Successive controls in columns 2-5 render the coefficients of all the indices become smaller, which shows that the degree of contact requirements are correlated with other job requirements and some of the beauty requirement can be explained by other job requirements in ads. The index “Working with the public” drops relatively larger after controlling for the firm fixed effects and the gender requirement dummy (from 0.06 to 0.04 percentage points after controlling for the firm fixed effects and from 0.04 to 0.03 percentage points after controlling for the gender requirement dummy, though these differences are not statistically significant given the standard errors for both are 0.01) . This change indicates that this index is correlated with firm fixed effects (e.g. firms with some specific culture are more likely to work directly with the public) and is positively correlated with the job gender requirement. Therefore, once these requirements get controlled for this index no longer picks up the over presented firm fixed effects and gender requirements, therefore it appears to have less positive force.

Table 2.3b shows the coefficients for specification interaction terms with each of the contact indices controlled in Column 6, 7, 8 and 9 in Table 2.3a to get a richer picture of the role of customer and coworker discrimination. Column 6 shows that the coefficients of interaction terms of the contact indices "working with the public" and "communication with the coworker" with any age requirement are positive and statistically significant, which means that the effects of these two contact indices on the probability of having a beauty requirement are overall larger for ads with any age requirement than ads without. Since the average preferred age is rather young (31.3 years old), ads with any age requirement can be further explained as ads targeting younger workers.[16] In other words, employers of ads targeting younger workers are in general more sensitive to the degree of contact required by jobs. Specifically, the magnitude of the coefficient of the index "working with the public" is slightly larger: one standard deviation increase in it would increase the probability of having a beauty requirement by 0.02 percentage points (a 0.25% on the mean) more in ads with any age requirement than ads without.

Column 7 in Table 2.3b shows the coefficients of the interaction terms of the contact indices and the gender requirement dummies. The signs of the coefficients of the male interaction terms are inconsistent among different indices. Specifically, in terms of customer-based degree of person contact, the coefficient of "communication with customer" is positive and statistically significant at 1% level but the coefficient of "working with the public" is negative and statistically significant at 10% level; in terms of coworker-based degree of person contact, the coefficient of "communication with coworker" is negative and statistically significant at 5% level while the coefficient of "face-to-face discussion" is zero and insignificant. Overall, whether the relationship of job contact requirements and the beauty requesting in job ads is stronger among ads targeting the male than ads without the specification of gender depends on the type of contact. Again, the signs of the coefficients of the female interaction terms are inconsistent among different indices. However, among them, only the coefficient of the interaction term of the index "communication with coworker" is positive and statistically significant at 5% level. This indicates that the relationship of ads asking beauty and job requirement of coworker contact are stronger among ads targeting the

[16] From Table 2.1.

female than ads without the specification of gender.

Column 8 in Table 2.3b shows the coefficients of the interaction terms of the contact indices and the dummy of whether a job ad asking for at least a bachelor's degree. Two coefficient with statistically significant results are the coefficient of the index "working with the public" and the index "communication with coworker" and both of them are negative. These results indicate that the relationship between job contact indices and beauty request in ads becomes weaker with more education required by the job.

In addition, I put all these interaction terms all together in column 9. The coefficients are quite similar as the ones I put them one by one.

In Table 2.4, I calculate the contributions of the two types of discrimination in whether firms have a beauty requirement by summing the coefficients. Column 1 shows that a one standard deviation increase in the two customer-based degree of person contact would increase the probability of having a beauty requirement in job ads by 3.32 percentage points (41.5%) while a one standard deviation increase in the two coworker-based degree of person contact would increase the probability of having a beauty requirement in job ads by 2.56 percentage points (32%). Only the customer-based degree of person contact is statistically significant at 10% level. In the columns 2 to 4, the results become gradually smaller and more statistically significant after adding more controls. The larger coefficient of the customer discrimination suggests that customer discrimination plays a more important role than coworker discrimination in whether employers determine to have a beauty requirement in job ads, but the p-values of the equality test in Table 2.4 show that the two effects are not statistically significantly different. Employer advertisements for good-looking workers are therefore prompted substantially similarly by customer and coworker discrimination.

Column 6 to column 8 in Table 2.4 show the contributions of the two types of discrimination action in different subgroups of ad types: those with an age requirement, those preferring women and those preferring men in order to get a richer picture of the role of customer and coworker discrimination. Column 6 shows that one standard deviation increase in customer-based degree of person contact would increase the probability of having a beauty requirement by 0.03 percentage points (0.38%) among ads with any age requirement and the coefficient is statistically significant at 5% level, while the coefficient of coworker discrimination is the same but with a larger statistic significant level. Column 7 shows that one standard deviation increase in customer-based degree of person contact would increase the probability of having a beauty requirement by 0.03 percentage points (0.38%) among ads preferring female and the coefficient is statistically significant at 1% level, while the coefficient of coworker discrimination is much smaller and not statistically significant. And the p-value of the equality test in Column 7 shows that the customer-based degree of person contact' effects are statistically significantly different from the coworker discrimination. This result means that it is customer interaction really driving the beauty requirements for jobs where employers want to hire female workers, therefore, customer discrimination plays a more important role. Column 8 shows similar results of customer discrimination for ads preferring male to ads preferring female, but with a much smaller magnitude, and the result of coworker discrimination increase to 0.07 and is statistically significant at 10% level, however, two effects are still not significantly different from each other.

In addition to the beauty requirement, the height requirement also reflects employer preferences towards attractive workers. Table 2.5 shows the regression results for the determinants of the height requirement. The six specifications of Table 2.5 are the same as Table 2.3 except that the dependent variable in Table 2.5 is an indicator for a height requirement.

Column 1 in Table 2.5a shows similar results to those of the beauty requirement in Table 2.3, but with a smaller magnitude. From column 2 to column 5 in Table 2.5a, successively controlling more various requirements in ads renders the coefficients of all the indices become slightly smaller, which shows that some of the height requirement can be explained by other requirements in ads but not much. A worth noting point is that the coefficient of the "communication with customer" moves closer to zero and statistically significant at 1% level when controlling for the firm fixed effects and other controls (the job posting periods, the number of positions), which indicates that this index is not strongly correlated with the height requirements in job ads although it keeps having a negative sign. Column 6, 7, 8 and 9 in Table 2.5a and Table 2.5b show the results of age, gender and education interaction terms with each of the contact indices. The signs of the coefficients of these interaction terms with the contact indices are mixed and the magnitudes are quite small.

In Table 2.6, I calculate the contributions of the two types of discrimination action by summing the coefficients of the height requirement (the same as Table 2.4). Table 2.6 shows some flipped results to those of the beauty requirement in Table 2.4. The larger and positive coefficients of the coworker-based degree of person contact suggests that coworker discrimination plays a more important role than customer discrimination in employers' determining whether or not having a height requirement in job ads. I am thinking this might have something due to the way of constructing the beauty dummy and height dummy variables. As I mentioned in the data part, the beauty dummy in the dataset refers to the overall image of a person which includes height, weight, face as a whole rather than just facial attractiveness, in this way, the beauty requirements in ads with more contact with customers are more likely to refer to the overall attractiveness, therefore, they will not bother by adding another height requirement. While, height requirement are typical height requirement, jobs with more contact with coworker are more likely to have height requirement either because the job itself needs some specific height (e.g. subway director) or because jobs with more coworker contact are more likely to ask for female workers as well (as indicated by the results of column 7 in Table 2.5b) and female workers care more about height than beauty. However, the two effects are not statistically significantly different in almost all column. Employer advertisements for tall workers is therefore prompted approximately similarly by customer and coworker discrimination.

Column 6 to column 8 in Table 2.6 show the contributions of the two types of discrimination action in different subgroups for the height requirement. Column 6 shows that the coefficients are rather similar, thus the two effects are approximately equal among ads targeting younger workers. Column 7 and column 8 show the coefficients of both customer and coworker based degree of person contact are small and not statistically significant.

2.4.2 ANOVA

To check which type of the discrimination explains more variation of the beauty and height requirements among occupations, Table 2.7 presents the ANOVA results. Two specifications each for both the beauty and height requirements are presented across columns from left to right, the first column of which includes the ANOVA results of all four job contact indices and the second one adds other covariates, including the age, gender, education, experience, part-time jobs, periods dummies and number of ads.

Column 1 in Table 2.7 shows that customer-based degree of person contact can explain 3.97% (0.51% for "Communication with customer" and 3.46% for "Working with the public") of the variation in the beauty requirement among job ads, while the number for coworker-based degree of person contact is 0.74% (0.71% for "Communication with coworker" and 0.04% for "Face-to-face discussion"). Therefore, customer discrimination can explain more variation in the beauty requirement. Customer discrimination has a larger size of effects from the previous analysis in Table 2.4. And it turns out that customer discrimination also accounts more for the variation in the beauty requirement among job ads. By adding the covariates in column 2, both the results become smaller, but customer discrimination can still explain more for the beauty requirement variation than coworker discrimination.

Column 3 in Table 2.7 shows that customer-based degree of person contact can explain 2.99% of the variation in the height requirement among job ads, but the number for coworker-based degree of person contact is smaller (0.47%). Therefore, there are more variation of customer discrimination across occupations and customer discrimination explains more variation in the height requirement. By adding the covariates in column 4, the results become smaller and coworker discrimination keeps explaining less variation than customer discrimination. Therefore, customer discrimination is more important for height in terms of variance as well though the gap between these two discrimination is smaller for the height requirement than for the beauty requirement.

2.4.3 SUR regression

Table 2.8 shows that the correlations of the residuals of equation "any beauty requirement" and equation "any height requirement", of equation "any height requirement" and "any gender requirement" and of equation "any gender requirement" and "any age requirement" are relatively high, which is 0.28, 0.23 and 0.27 respectively. This result means that the error terms of these four equations are related with each other.

The coefficients and their standard errors for each regression are similar to the main regression results and are presented in Appendix Table 2.2. Table 2.9 shows various joint tests across different contact indices.

Panel A in Table 2.9 shows the joint tests for equation "any beauty requirement" and "any height requirement". Row 1 shows the joint tests including all four contact indices. Its p value is zero suggests that the hypothesis that the overall effect of four contact indices on the beauty requirement is the same as the height requirement is rejected. Row 2 and row 3 further checks whether this difference comes from customer-based degree of person contact or coworker-based degree of person contact. Row 2 shows that the p value of customer-based degree of contact is zero, which suggests that the hypothesis that the customer discrimination has the same effect on the beauty requirement and the height requirement is rejected. However, the p values of coworker-based degree of contact in row 3 is relatively larger (0.45). Therefore, customer discrimination contributes more for the different effects of indices on the beauty requirement and the height requirement. From row 4 to row 7, I further check the p value of each index and show that the index "working with the public" contributes the most while other indices seem have similar effect on beauty and height requirements.

Panel B in Table 2.9 shows the joint tests for equation "any beauty requirement" and "any age requirement". Row 1 shows the joint tests including all four contact indices. Its p value is around 0.02. This large p value suggests that the hypothesis that the overall effect of four contact indices on the beauty requirement is the same as the gender requirement can be rejected at 5% level. And a checking in row 2 and row 3 of which discrimination in general contributes more the beauty and gender requirements finds some evidence that the hypothesis that coworker discrimination has the same effect on the beauty and gender requirements is rejected at 1% level with a zero P-value. And a further checking for all four indices from row 4 to row 7 reveals that the "communication with coworker" plays the most important role.

Panel C in Table 2.9 shows the joint tests for equation "any beauty requirement" and "any gender requirement". Row 1 shows the joint tests including all four contact indices. Its p value is zero suggests that the hypothesis that the overall effect of four contact indices on the beauty requirement is the same as the gender requirement is rejected. Row 2 and row 3 further checks whether this difference comes from customer-based degree of person contact or coworker-based degree of person contact. Row 2 shows that the p value of customer-based degree of person contact is less than 1%, which suggests that the hypothesis that the customer discrimination has the same effect on the beauty requirement and the gender requirement is rejected. And the p value of coworker-based degree of person contact in row 3 is much larger (0.49). Therefore, customer discrimination contributes more for the different effects of indices on the beauty requirement and the gender requirement. From column 4 to column 7, I further check the p value of each index and show that the index "communication with customer" and "working with the public " contribute more to the difference.

2.5 Conclusion

I provide a new and effective method to test the type of discrimination by Chinese employers. Using limited data, I create the contact indices, categorize them into customer discrimination and coworker discrimination and compare their effects. I find that customer-based degree of person contact are positively related with the probability of having beauty requirement and coworker-based degree of person contact are positively related with the probability of having height requirement, but the magnitude is small and not very significant. A one standard deviation increase in customer-based degree of person contact increases the chance of having a beauty requirement in job ads by 3.32 percentage points (41.5%) and decreases the chance of having a height requirement by 0.35 percentage points (12%). A one standard deviation increase in the two coworker-based degree of person contact increases the probability of having a beauty requirement by 2.56 percentage points (32%) and the probability of having a height requirement by 1.16 percentage points (40%). The variance analysis further finds that customer discrimination can explain more for the variance of both having a beauty requirement and having a height requirement. These results suggest that both customer discrimination and coworker discrimination play some role in prompting employer advertisements of attractive employees and relatively, customer discrimination is more present.

In addition, by analyzing different subgroups, I find that it is customer interaction driving both the beauty and height requirements for jobs targeting younger workers, but the magnitude is small. Therefore, customer discrimination plays a slightly more important role for employers who want to hire younger workers.

However, from the SUR regression, I find that the results found for beauty and height requirements cannot be further generalized to other job requirements, such as, age requirement, gender requirement.

2.6 Tables

Table 2.1: Means of Requirements in Ads

	All	Without beauty requirement	With beauty requirement	Without height requirement	With height requirement
	(1)	(2)	(3)	(4)	(5)
Any beauty requirement	0.081	0	1	0.067	0.552
	(0.273)	/	/	(0.249)	(0.497)
Any height requirement	0.029	0.014	0.200	0	1
	(0.169)	(0.119)	(0.400)	/	/
Any age requirement	0.240	0.224	0.423	0.231	0.557
	(0.427)	(0.417)	(0.494)	(0.421)	(0.497)
Ad specifies a minimum age	0.167	0.155	0.300	0.160	0.400
	(0.373)	(0.362)	(0.458)	(0.366)	(0.490)
Ad specifies a maximum age	0.200	0.184	0.375	0.190	0.512
	(0.400)	(0.388)	(0.484)	(0.393)	(0.500)
Male preferred	0.055	0.057	0.040	0.053	0.137
	(0.228)	(0.231)	(0.196)	(0.224)	(0.344)
Female preferred	0.053	0.036	0.245	0.042	0.413
	(0.224)	(0.187)	(0.430)	(0.201)	(0.492)
No school restrictions	0.197	0.194	0.232	0.194	0.303
	(0.398)	(0.396)	(0.422)	(0.396)	(0.459)
Junior middle school or below	0.008	0.007	0.022	0.006	0.057
	(0.089)	(0.082)	(0.147)	(0.078)	(0.233)
High school	0.095	0.089	0.164	0.090	0.268
	(0.293)	(0.284)	(0.370)	(0.286)	(0.443)
Post-secondary	0.367	0.363	0.413	0.369	0.273
	(0.482)	(0.481)	(0.492)	(0.483)	(0.445)
Undergraduate	0.321	0.335	0.164	0.328	0.096
	(0.467)	(0.472)	(0.370)	(0.469)	(0.294)
Master	0.012	0.012	0.006	0.012	0.004
	(0.108)	(0.111)	(0.074)	(0.109)	(0.060)
Ph.D. and above	0.001	0.001	0.000	0.001	0.000
	(0.029)	(0.030)	(0.014)	(0.030)	(0.007)
Experience years required	2.726	2.819	1.662	2.764	1.446
	(2.813)	(2.858)	(1.956)	(2.827)	(1.936)
Job is part-time	0.015	0.015	0.012	0.014	0.025
	(0.120)	(0.120)	(0.111)	(0.118)	(0.158)
Number of ads	1,300,118	1,195,040	105,078	1,262,090	38,028
Preferred age*	31.317	31.861	28.034	31.679	26.393
	(7.592)	(7.537)	(7.082)	(7.552)	(6.318)
Number of ads	317,777	272,619	45,158	296,016	21,761

Notes: Data refer to ads downloaded from Professor Kuhn's website. Unweighted means, standard deviations in parentheses." Any age requirement" is a dummy variable: it equals 1 if the job ads require any age requirement, 0 if else. "Male preferred" ("Female preferred") is a dummy variable valuing 1 if the job ads require male (female) workers and 0 if else. All education variables are dummy variables, for example, the dummy variable "High school" equals 1 if the education requirement in ads is "high school or above". Ads without specifying education requirements entry the category "No school restrictions" and its dummy variable "No school restrictions" equals 1. Variable "experience years required" describe the experience year required in job ads and its value of ads with no experience requirement is 0.

**"preferred age" is the mean of max and min age when either is specified in Ads. The default min age is 16 when max age is only specified and the default max age is 62.5 when min age is only specified.*

Table 2.2: Means of the Standardized Weighted Job Contact Indices

	Without beauty requirement	With beauty requirement	Without height requirement	With height requirement
	(1)	(2)	(3)	(4)
Communication with customer	0.001	-0.014	0.017	-0.557
	(0.989)	(1.122)	(0.987)	(1.241)
Working with the public	-0.051	0.581	-0.016	0.529
	(1.002)	(0.764)	(1.000)	(0.827)
Communication with coworker	-0.007	0.078	0.005	-0.163
	(0.987)	(1.133)	(0.993)	(1.207)
Face-to-face discussion	-0.015	0.173	0.002	-0.060
	(1.009)	(0.874)	(0.998)	(1.067)
N	1,195,040	105,078	1,262,090	38,028

Notes: Means of standardized indices. Standard deviations in parentheses. Communication with customer: Communicating with people outside organization measured by ; Working with the public: Performing for or working directly with the public measured by ; Communication with coworker: Communicating with supervisors, peers, or subordinates measured by ; Face-to-face discussion: The frequency of Face-to-Face Discussions.

Table 2.3a: Effects of Job's Degree of Contact Demands on the Probability an Ad Is Beauty-targeted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Communication with customer/100	-0.02	-0.02	-0.02*	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Working with the public/100	0.06***	0.05***	0.05***	0.04***	0.03***	0.03***	0.03***	0.04***	0.04***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Communication with coworker/100	0.03	0.03	0.03*	0.03	0.03	0.02	0.02	0.03*	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Face-to-face discussion/100	-0.01	-0.01	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ad specifies a minimum age		0.08***	0.06***	0.05***	0.04***	0.04***	0.04***	0.04***	0.04***
		(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ad specifies a maximum age		0.09***	0.08***	0.07***	0.06***	0.06***	0.06***	0.06***	0.05***
		(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Preferred age*Any age requirement/100		-0.23***	-0.18***	-0.16***	-0.12***	-0.12***	-0.12***	-0.12***	-0.12***
		(0.07)	(0.06)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
Male preferred					-0.02***	-0.02***	-0.02***	-0.02***	-0.02***
					(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female preferred					0.21***	0.21***	0.19***	0.21***	0.19***
					(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
Education & Experience	NO	NO	YES	YES	YES	YES	YES	YES	YES
Other covariates	NO	NO	NO	YES	YES	YES	YES	YES	YES
Firm fixed effects	NO	NO	NO	YES	YES	YES	YES	YES	YES
Interaction terms	NO	NO	NO	NO	NO	YES	YES	YES	YES
R-squared	0.04	0.05	0.07	0.33	0.35	0.35	0.35	0.35	0.36

Notes: OLS estimates. Dependent variable in all columns equals 1 if the ad explicitly requests beauty, and 0 otherwise. I have standardized the job contact indices and their coefficients are multiplied by 100. There are 1,300,118 observations.
**** p<0.01, ** p<0.05, * p<0.1.*

Table 2.3b: Effects of Job's Degree of Contact Demands on the Probability an Ad Is Beauty-targeted — Interaction Terms

	(6)	(7)	(8)	(9)
Any age requirement*Communication with customer/100	-0.01			-0.01*
	(0.01)			(0.00)
Any age requirement*Working with the public/100	0.02***			0.02***
	(0.01)			(0.01)
Any age requirement*Communication with coworker/100	0.01**			0.01*
	(0.01)			(0.00)
Any age requirement*Face-to-face discussion/100	-0.00			0.00
	(0.01)			(0.00)
Male preferred*Communication with customer/100		0.02***		0.02***
		(0.01)		(0.01)
Female preferred*Communication with customer/100		-0.04		-0.04
		(0.03)		(0.03)
Male preferred*Working with the public/100		-0.01*		-0.02**
		(0.01)		(0.01)
Female preferred*Working with the public/100		0.02		0.01
		(0.02)		(0.02)
Male preferred*Communication with coworker/100		-0.02**		-0.02**
		(0.01)		(0.01)
Female preferred*Communication with coworker/100		0.06**		0.06**
		(0.03)		(0.02)
Male preferred*Face-to-face discussion/100		0.00		-0.00
		(0.01)		(0.01)
Female preferred*Face-to-face discussion/100		0.00		-0.00
		(0.02)		(0.02)
Bachelor's degree*Communication with customer/100			0.01	0.01
			(0.01)	(0.00)
Bachelor's degree*Working with the public/100			-0.02***	-0.02***
			(0.00)	(0.00)
Bachelor's degree*Communication with coworker/100			-0.01**	-0.01**
			(0.01)	(0.00)
Bachelor's degree*Face-to-face discussion/100			0.00	0.00
			(0.00)	(0.00)

Notes: OLS estimates. Dependent variable in all columns equals 1 if the ad explicitly requests beauty, and 0 otherwise. I have standardized the job contact indices and their coefficients are multiplied by 100. Other covariates are the same as the column 6,7,8,9 in the Table 3a. There are 1,300,118 observations.

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

Table 2.4: Relative Importance of Customer and Coworker Discrimination Coefficients for Targeting Beauty in Ads

			All ads			Ads with any age requirement	Ads preferring female	Ads preferring male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Customer discrimination	3.32*	3.11*	2.73*	2.95**	2.54**	0.03**	0.03***	0.002
	(1.73)	(1.65)	(1.41)	(1.31)	(1.19)	(0.02)	(0.01)	(0.05)
Coworker discrimination	2.56	2.55	2.67	2.59*	2.19	0.03*	-0.01	0.07*
	(1.64)	(1.54)	(1.36)	(1.76)	(1.32)	(0.02)	(0.01)	(0.04)
P-Value of equality test	0.799	0.850	0.991	0.888	0.884	0.804	0.001	0.386
N	1,300,118	1,300,118	1,300,118	1,300,118	1,300,118	312,028	68,906	71,506

Notes: Dependent variable in all columns equals 1 if the ad explicitly requests beauty, and 0 otherwise. Each column refers to the sum of corresponding column's coefficients in Table 3 and columns 6-8 refer to the coefficients for each subgroup based on column 6 in Table 3. Standard errors are in parentheses. There are 1,300,118 observations.

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

Table 2.5a: Effects of Job's Degree of Contact Demands on the Probability an Ad Is Height-targeted

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Communication with customer/100	-0.03**	-0.03**	-0.02**	-0.01***	-0.01**	-0.00*	-0.00	-0.01**	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Working with the public/100	0.02***	0.02***	0.02***	0.01***	0.01***	0.00*	0.01*	0.01***	0.01*
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Communication with coworker/100	0.02*	0.02**	0.02**	0.01*	0.01*	0.01*	0.01	0.01*	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Face-to-face discussion/100	-0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ad specifies a minimum age		0.06***	0.05***	0.05***	0.04***	0.04***	0.04***	0.04***	0.04***
		(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Ad specifies a maximum age		0.06***	0.06***	0.05***	0.04***	0.03***	0.04***	0.04***	0.03***
		(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Preferred age*Any age requirement/100		-0.18**	-0.15**	-0.16***	-0.14***	-0.15***	-0.14***	-0.14***	-0.14***
		(0.07)	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Male preferred					0.04***	0.04***	0.03***	0.04***	0.03***
					(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female preferred					0.15***	0.15***	0.13***	0.15***	0.13***
					(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Education & Experience	NO	NO	YES	YES	YES	YES	YES	YES	YES
Other covariates	NO	NO	NO	YES	YES	YES	YES	YES	YES
Firm fixed effects	NO	NO	NO	YES	YES	YES	YES	YES	YES
Interaction terms*	NO	NO	NO	NO	NO	YES	YES	YES	YES
R-squared	0.03	0.04	0.06	0.35	0.38	0.38	0.38	0.38	0.38

Notes: OLS estimates. Dependent variable in all columns equals 1 if the ad explicitly requests height, and 0 otherwise. I have standardized the job contact indices and their coefficients are multiplied by 100. There are 1,300,118 observations.

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

Table 2.5b: Effects of Job's Degree of Contact Demands on the Probability an Ad Is Height-targeted — Interaction Terms

	(6)	(7)	(8)	(9)
Any age requirement*Communication with customer/100	-0.02**			-0.01**
	(0.01)			(0.00)
Any age requirement*Working with the public/100	0.02***			0.01***
	(0.00)			(0.00)
Any age requirement*Communication with coworker/100	0.01			0.01
	(0.01)			(0.00)
Any age requirement*Face-to-face discussion/100	0.00			-0.00
	(0.00)			(0.00)
Male preferred*Communication with customer/100		-0.04**		-0.03**
		(0.01)		(0.01)
Female preferred*Communication with customer/100		-0.06		-0.06
		(0.04)		(0.04)
Male preferred*Working with the public/100		0.02*		0.01
		(0.01)		(0.01)
Female preferred*Working with the public/100		0.03*		0.02
		(0.02)		(0.02)
Male preferred*Communication with coworker/100		-0.01		-0.01
		(0.01)		(0.01)
Female preferred*Communication with coworker/100		0.03		0.03
		(0.03)		(0.03)
Male preferred*Face-to-face discussion/100		0.03		0.03
		(0.02)		(0.02)
Female preferred*Face-to-face discussion/100		-0.02		-0.02
		(0.02)		(0.02)
Bachelor's degree*Communication with customer/100			0.01***	0.01***
			(0.00)	(0.00)
Bachelor's degree*Working with the public/100			-0.01***	-0.01***
			(0.00)	(0.00)
Bachelor's degree*Communication with coworker/100			-0.01*	-0.00*
			(0.00)	(0.00)
Bachelor's degree*Face-to-face discussion/100			-0.00	-0.00
			(0.00)	(0.00)

Notes: OLS estimates. Dependent variable in all columns equals 1 if the ad explicitly requests height, and 0 otherwise. I have standardized the job contact indices and their coefficients are multiplied by 100. Other covariates are the same as the column 6,7,8,9 in the Table 5a. There are 1,300,118 observations.

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

Table 2.6: Relative Importance of Customer and Coworker Discrimination Coefficients for Targeting Height in Ads

			All ads			Ads with any age requirement	Ads preferring female	Ads preferring male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Customer discrimination	-0.35	-0.50	-0.51	0.04	-0.14	0.004	-0.02	-0.03
	(0.96)	(0.92)	(0.75)	(0.45)	(0.39)	(0.01)	(0.02)	(0.05)
Coworker discrimination	1.16*	1.15*	1.20**	1.16**	0.84**	0.01*	0.02	0.02
	(0.65)	(0.58)	(0.47)	(0.47)	(0.38)	(0.01)	(0.01)	(0.03)
P-Value of equality test	0.245	0.174	0.080	0.102	0.109	0.473	0.227	0.472
N	1,300,118	1,300,118	1,300,118	1,300,118	1,300,118	312,028	68,906	71,506

Notes: Dependent variable in all columns equals 1 if the ad explicitly requests height, and 0 otherwise. Each column refers to the sum of corresponding column's coefficients in Table 5 and columns 6-8 refer to the coefficients for each subgroup based on column 6 in Table 5. Standard errors are in parentheses. There are 1,300,118 observations.

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

Table 2.7: The Percent of the Sum of Squares(SS) Explained by Job's Degree of Contact Covariate (ANOVA)

	Any beauty requirement		Any height requirement	
	(1)	(2)	(3)	(4)
Communication with customer	0.51%	0.31%	1.56%	0.80%
	(6,900)	(4,467)	(20,837)	(11,761)
Working with the public	3.46%	1.77%	1.42%	0.39%
	(46,710)	(25,749)	(19,003)	(5,724)
Sum of customer discrimination	*3.97%*	*2.08%*	*2.99%*	*1.19%*
Communication with coworker	0.71%	0.55%	0.43%	0.27%
	(9,550)	(8,055)	(5,749)	(3,907)
Face-to-face discussion	0.04%	0.02%	0.03%	0.02%
	(505)	(218)	(461)	(267)
Sum of coworker discrimination	*0.74%*	*0.57%*	*0.47%*	*0.28%*
Other covariates	NO	YES	NO	YES
R-squared	0.04	0.11	0.03	0.12

Notes: ANOVA results. Dependent variable in column 1 and 2 (3 and 4) equals 1 if the ad explicitly requests beauty (height), and 0 otherwise. The job contact indices are defined in Table 2. "Other covariates" refer to additional controls: gender, age, education, experience, part time, periods, and numbers of ads. F values are in parentheses. There are 1,300,118 observations.

Table 2.8: Correlation Matrix of Residuals

	Any beauty requirement (1)	Any height requirement (2)	Any age requirement (3)	Any gender requirement (4)
Panel A: no covariates				
Any beauty requirement	1.00			
Any height requirement	0.28	1.00		
Any age requirement	0.11	0.11	1.00	
Any gender requirement	0.15	0.23	0.27	1.00
Panel B: with covariates				
Any beauty requirement	1.00			
Any height requirement	0.27	1.00		
Any age requirement	0.10	0.10	1.00	
Any gender requirement	0.14	0.21	0.26	1.00

Notes: Breusch-Pagan test of independence: chi2(6) = 3.31e+05, P = 0.0000. There are 1,300,118 observations.

Table 2.9: Tests of Equality of Coefficients

Coefficient comparisons of beauty requirement and	P-Value	Covariate coefficients tested			
		Communication with customer	Working with the public	Communication with coworker	Face-to-face discussion
	(1)	(2)	(3)	(4)	(5)
Height requirement	0.0001	Y	Y	Y	Y
	0.0012	Y	Y	-	-
	0.4194	-	-	Y	Y
	0.7985	Y	-	-	-
	0.0008	-	Y	-	-
	0.2226	-	-	Y	-
	0.7672	-	-	-	Y
Age requirement	0.0185	Y	Y	Y	Y
	0.9616	Y	Y	-	-
	0.0033	-	-	Y	Y
	0.7800	Y	-	-	-
	0.9651	-	Y	-	-
	0.0010	-	-	Y	-
	0.1695	-	-	-	Y
Gender requirement	0.0002	Y	Y	Y	Y
	0.0062	Y	Y	-	-
	0.4911	-	-	Y	Y
	0.0934	Y	-	-	-
	0.0258	-	Y	-	-
	0.2437	-	-	Y	-
	0.4950	-	-	-	Y

Notes: Based on SUR. Tests of whether contact indices have the same effects on the beauty, height, age and gender requirement. There are 1,300,118 observations.

2.7 Appendix. Additional Tables

Table 2.A1: The Crosswalk of the O*NET Job Titles and the Chinese Job Categories

	Job family name (1)	Occupation Code (2)
Sales	Sales and related	41.2022.00-41.9091.00
IT	Computer and mathematical	15.1111.00-15.2091.00
Marketing	Business and financial operations	13.1011.00-13.1199.06
Accounting	Office and administrative support	43.3011.00-43.4151.00
	Sales and related	41.2011.00-41.2021.00
Administration	Management	11.9031.00-11.9033.00
Construction	Construction and extraction	47.1011.00-47.5081.00
Manufacturing	Production	51.4051.00-51.4072.00
Human resource	Office and administrative support	43.1011.00-43.9081.00
Finance	Business and financial operations	13.1031.01-13.2099.04
	Office and administrative support	43.9041.01-43.9041.02
Customer service	Food preparation and serving related	35.3022.00-35.9031.00
	Personal care and service	39.9041.00
	Office and administrative support	43.2011.00-43.5111.00
Commerce trade	Sales and related	41.3031.03
Tourism	Personal care and service	39.7011.00-39.7012.00
	Food preparation and serving related	35.1011.00-35.2015.00
Communication logistics	Management	11.3031.01-11.9121.01
High management	Management	11.1011.00-11.9199.11
	Protective service	33.1011.00-33.1021.02
	Food preparation and serving related	35.1012.00
	Building and grounds cleaning and maintenance	37.1011.00-37.1012.00
	Personal care and service	39.1011.00-39.1021.01
	Sales and related	41.1011.00-41.1012.00
	Production	51.1011.00
Electronics	Architecture and engineering	17.1011.00-17.3031.02
Design	Arts, design, entertainment, sports, and media	27.1011.00-27.3043.05
Education training	Education, training, and library	25.1011.00-25.9041.00
	Personal care and service	39.9031.00-39.9032.00
Media	Arts, design, entertainment, sports, and media	27.1014.00-27.4032.00
Retail	Sales and related	41.2031.00, 41.9012.00
Quality control	Computer and mathematical	15.1151.00-15.1199.01
Construction machinery	Production	51.3093.00-51.9199.01
Consultancy	Business and financial operations	13.1041.01-13.1199.05
Technical workers	Office and administrative support	43.9051.00-43.9111.01
	Installation, maintenance, and repair	49.1011.00-49.9099.01
	Production	51.2011.00-51.6011.00
biomedical and pharmaceutical	Healthcare Practitioners and Technical	29.1011.00-29.9099.01
Healthcare	Healthcare support	31.1011.00-31.9099.02
	Personal Care and Service	39.9011.00-39.9021.00
Law	Legal	23.1011.00-23.2093.00
Communication tech	Computer and mathematical	15.1142.00-15.1143.01
Translation	Arts, design, entertainment, sports, and media	27.3091.00
Energy	Life, physical, and social science	19.1031.01-19.3099.01
Chemical	Life, physical, and social science	19.4031.00
Labor domestic service	Food preparation and serving related	35.2021.00-35.9021.00
	Building and grounds cleaning and maintenance	37.2011.00-37.3013.00
	Personal care and service	39.2011.00-39.6012.00
Environment protection	Protective service	33.2011.01-33.9099.02
Research	Life, physical, and social science	19.1011.00-19.4099.03
Agriculture, forest, fishery	Farming, fishing, and forestry	45.1011.05-45.4023.00
Public servant	Community and social service	21.1011.00-21.2021.00
Textile	Production	51.6021.00-51.6092.00
Vehicle	Transportation and material moving	53.1011.00-53.7121.00

*Notes: Job information downloaded from O*NET website.*

Table 2.A2: Summary Statistics of the Component Parts of the Weighted Job Contact Indices

	All			Without beauty requirement	With beauty requirement	Without height requirement	With height requirement
	Min (1)	Max (2)	Mean (3)	Mean (4)	Mean (5)	Mean (6)	Mean (7)
Communicating with persons outside the organization:							
Importance	5.536	27.806	21.133	21.137	21.082	21.209	18.607
	/	/	(4.482)	(4.443)	(4.904)	(4.430)	(5.370)
Level	5.709	26.509	18.598	18.605	18.511	18.665	16.381
	/	/	(4.265)	(4.205)	(4.897)	(4.206)	(5.466)
Performing for or working directly with the public:							
Importance	3.447	32.182	15.573	15.274	18.982	15.476	18.792
	/	/	(6.103)	(6.116)	(4.767)	(6.105)	(5.074)
Level	3.554	25.610	13.984	13.747	16.686	13.916	16.239
	/	/	(4.400)	(4.407)	(3.260)	(4.401)	(3.683)
Communicating with supervisors, peers or subordinates:							
Importance	16.833	29.848	24.985	24.960	25.274	24.990	24.845
	/	/	(2.314)	(2.298)	(2.465)	(2.304)	(2.599)
Level	14.108	26.978	20.698	20.693	20.757	20.723	19.867
	/	/	(3.094)	(3.042)	(3.638)	(3.064)	(3.886)
N	1,300,118	1,300,118	1,300,118	1,195,040	105,078	1,262,090	38,028

*Notes: Summary statistics of job descriptors downloaded from O*NET website. Standard deviations in parentheses.*

Table 2.A3: Correlation Matrix of the Job Contact Indices

	Communication with customer (1)	Working with the public (2)	Communication with coworker (3)	Face-to-face discussion (4)	Preferred age (5)	Experience (6)	Education (7)
Communication with customer	1.00						
Working with the public	0.14	1.00					
Communication with coworker	0.56	-0.16	1.00				
Face-to-face discussion	0.44	0.24	0.54	1.00			
Preferred age	-0.04	0.12	-0.07	0.01	1.00		
Experience	-0.08	-0.06	-0.05	-0.04	0.03	1.00	
Education	0.10	-0.08	0.10	0.04	-0.03	0.33	1.00

Notes: There are 1,300,118 observations.

Table 2.A4: Effects of Job's Degree of Contact Demands on the Probability an Ad Is beauty (Height, Age, Gender)-targeted (SUR)

	Any beauty requirement (1)	Any height requirement (2)	Any age requirement (3)	Any gender requirement (4)
Communication with customer/100	-0.02	-0.03**	-0.03*	-0.04***
	(0.02)	(0.01)	(0.01)	(0.01)
Working with the public/100	0.06***	0.02***	0.06***	0.04***
	(0.02)	(0.01)	(0.01)	(0.01)
Communication with coworker/100	0.03	0.02*	-0.01	0.02
	(0.02)	(0.01)	(0.02)	(0.02)
Face-to-face discussion/100	-0.01	-0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
R-squared	0.04	0.03	0.02	0.02

Notes: SUR regressions. Dependent variable in column 1 (2,3,4) equals 1 if the ad explicitly requests beauty (height, age, gender), and 0 otherwise. The job contact indices are defined in Table 2. Standard errors (in parentheses) are robust and clustered at the firm level. There are 1,300,118 observations.

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

Table 2.A5: Effects of Job's Degree of Contact Demands on the Probability an Ad Is Beauty (Height, Age, Gender)-targeted (SUR with Controls)

	Any beauty requirement (1)	Any height requirement (2)	Any age requirement (3)	Any gender requirement (4)
Communication with customer/100	-0.03*	-0.02**	-0.02**	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
Working with the public/100	0.05***	0.02***	0.05***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Communication with coworker/100	0.03*	0.02**	-0.00	0.03*
	(0.02)	(0.01)	(0.01)	(0.02)
Face-to-face discussion/100	-0.01	-0.00	0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Other covariates	YES	YES	YES	YES
R-squared	0.06	0.05	0.04	0.04

Notes: SUR regressions. Dependent variable in column 1 (2,3,4) equals 1 if the ad explicitly requests beauty (height, age, gender), and 0 otherwise. The job contact indices are defined in Table 2. "Other covariates" refer to additional controls: education, experience, part time, periods, and numbers of ads. Standard errors (in parentheses) are robust and clustered at the firm level. There are 1,300,118 observations.

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

Chapter 3

Interpersonal Contact in Jobs and the Gender Wage Gap

3.1 Introduction

The last century has seen women make a lot of progress in education, labor market participation, work related training and experience, and this has been accompanied by a large decline in the gender wage gap from a 0.59 female-male annual earnings ratio in the 1970s to 0.77 in the 2000s (Goldin, 2014). However, a 20% gap has persisted since the 1990s. This gender wage gap increases with age: Goldin (2014) finds in almost all the cohorts that women's annual earnings start at a similar level to men's at early ages and drops to below 70% level when the workers are in their 40s.

A worker's wage is important not only because it reflects a person's productivity and financial resources but also because it reflects how the person is valued socially and economically (Goldin, 2014). Researchers have studied how large the gap is (Goldin, 1989; Polachek, 1993; Altonji and Blank, 1999), what causes it (Blau and Kahn, 2006) and how policy could reduce or eliminate it (Goldin, 2014). There have been several potential explanations considered for the remaining gender wage gap and how it grows with age. Lazear and Rosen (1990) point out that the difference in job ladders and the higher standard for women's promotion might contribute to the increasing gender wage gap. Goldin (2014) explains the growing gender wage gap from the personnel perspective and attributes it to the workplace flexibility. Keller and Utar (2018) find that childbearing and family have negative effects on the employment and wage of females. Flory et al. (2015) explain the growing wage gap by the lack of competition among females.

In this paper, I consider whether the amount of contact workers in an occupation have with coworkers, the public and customers affects the widening of the gender gap with age. Contact could matter if there is beauty-related discrimination against older, female workers. The link between age and perceived beauty seems to work differently for females versus males. Based on a social psychology experiment on 60 subjects, Korthase and Trenholme (1982) find that perceived physical attractiveness is negatively correlated with perceived age and that age is perceived as reducing females' beauty more than males'. Along with the fact found by Mathes et al. (2010) that "physical attractiveness is more important to males in selecting partners than females", we can infer that females are more likely judged by physical attractiveness and females are more likely to be perceived as less attractive at older ages. Therefore, if there is beauty-related discrimination in labor market, it would reduce females' relative wage, and increasingly so with age.

Furthermore, the gender wage gap would be expected to grow faster in jobs where physical attractiveness matters more. Physical attractiveness is likely to matter more in jobs involving more interpersonal contact with others. Helleseter et al. (2020) use job ads in several countries to document employers' preference for younger females and older males and find that around one third of this preference can be explained by the ads for jobs involving more contact, such as managerial roles, customer contact and helping roles. However, the authors do not quantify contact and do not distinguish types of contact.

In this paper, I investigate the effect of interpersonal contact on wages by comparing how the gender wage gap evolves with age for jobs with different degrees of contact, and I distinguish between types of contact in order to distinguish the types of discrimination possibly at work. I hypothesize that any interpersonal contact effects are likely to reflect beauty-based discrimination, although there might exist other factors that also contribute to the differences. To do this, I merge the O*NET occupation characteristics with worker data from the merged outgoing rotation groups (MORGs) of the Current Population Survey (CPS).

Becker (2010) identifies three sources of discrimination: customer discrimination, coworker discrimination, and employer discrimination. Distinguishing these sources of discrimination empirically is often challenging due to limited data. I investigate the source of potential beauty discrimination by comparing how the gender wage gap evolves with age for jobs with different levels of contact with different groups of people, namely, customers, the public and coworkers. Specifically, I construct four contact indices "Working with the public", "Communication with customers", "Communication with coworkers", and "Face-to-face discussion" using occupation information from the O*NET as Wu (2020) did. If the beauty discrimination comes from customers, we would expect the gender wage gap to grow more quickly with age in jobs with more customer contact, while if the prejudice comes from coworkers, then the gender wage gap would grow faster in jobs with more contact with coworkers.

In addition, I split the sample and do a comparative analysis for college graduates and non-college graduates. According to Kuhn and Shen (2013), when a job ad asks for more skills, the challenge of finding a good fit increases. Therefore, the employer may need to give up their own stereotype or the stereotype that they think their customers or employees have towards less attractive people to find the right match. Therefore, the effect of beauty might be different among college graduates. Furthermore, the pattern of the gender wage gap might also look different between college graduates and non-college graduates because the types of jobs college graduates work and the types of people college graduates have interaction with at work are different from those of non-college graduates.

I find that the gender wage gap in jobs with more overall interpersonal contact rises faster with age than in jobs with less interpersonal contact, though the difference is statistically insignificant. In terms of contact type, the gender wage gap rises statistically significantly faster with age in jobs with a high "Working with the public" requirement. A one-year increase in age increases the gender wage gap in jobs involving more contact than jobs involving less contact by 0.21%. By contrast, the relative increase in the gap is much smaller and statistically insignificant in jobs with high levels of other types of contact. To the extent that beauty discrimination explains this pattern, this suggests that the source of the beauty discrimination is customers. Comparing college graduates and non-graduates, I find that non-graduates have a statistically significantly faster growing gender wage gap with age than college graduates.

My study is the first to link quantitatively interpersonal contact, age and the origin of the discrimination. Biddle and Hamermesh (1998) find evidence for beauty discrimination among graduates in law. However, they do not classify the beauty effects by gender and how they change with age. Some existing papers suggest that the degree of contact with others might play a role in discrimination, but few actually measure the contact level in job and relate it to the level of discrimination suffered by workers in jobs as I do.[1] The papers that do measure contact are Kenney and Wissoker (1994), Kuhn and Shen (2009), Holzer and Inlanfeldt (1998), Combes et al. (2016) and Wu (2020). Kenney and Wissoker (1994) measure interpersonal contact in jobs based on their a priori impression and find the job application outcomes of Hispanic job seekers are slightly worse in high contact jobs relative to low contact jobs. Using a similar method, Kuhn and Shen (2009) find evidence supporting customer discrimination against the disfavored group in terms of age, gender, beauty and height. I extend this by measuring contact in a more objective way using the O*NET. Holzer and Inlanfeldt (1998) and Combes et al. (2016) measure workers' contact level with customers by asking employers relevant questions in surveys. Both papers find customer discrimination in the hiring process, especially for jobs involving contact with customers. I improve upon this by studying contact with both customers and coworkers.

Wu (2020) uses the same O*NET contact indices to investigate direct requests for beautiful applicants in a Chinese on-line job board. I expand on this in this paper by merging the U.S. O*NET contact indices with workers in the United States and extend my earlier analysis by investigating the impact of different levels and types contact indices on the change of gender wage gap with age and relating this to beauty

[1] Kahn and Shearer (1988); Nardinelli and Simon (1990); Ihlanfeldt and Sjoquist (1991); Kenney and Wissoker (1994); Young and Inlanfeldt (1994); Neumark (1996); Holzer and Inlanfeldt (1998); Leonard et al. (2010); Bar and Zussman (2017)

discrimination.

3.2 Data

I use two sources of data in this analysis. The merged outgoing rotation groups (MORGs) from The U.S. Current Population Survey (CPS) data from 2010 to 2019 for information on wage and the core for workers' demographics. And the Occupational Information Network (O*NET) database, a database containing hundreds of job descriptors that encapsulate the distinguishing features of an occupation in the United States from various domains. Each occupation in the O*NET database has a unique O*NET ID code whose first 6 digits are the job's Standard Occupational Classification (SOC) code. Occupations in the CPS also have SOC code. I merge the O*NET database with the CPS data by this SOC code so that each occupation in CPS has its correspondent job descriptors. I restrict my sample to workers 18-64 years old workers who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status. I construct the hourly wage variables by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below 2*orabove*200 with weekly working hours less or equal to 15 hours.

There are hundreds of variables describing the working activities of a job in the O*NET. To measure contact levels, I use the same four variables used in Wu (2020) from the working context domain, which are the only variables that specify both the type of the interaction the job requires and the group of people towards whom the interaction is made. They are: 1) "Working with the public" defined in the O*NET as "Performing for people or dealing directly with the public. This includes serving customers in restaurants and stores, and receiving clients or guests." 2) "Communication with the customer" defined as "Communicating with people outside the organization, representing the organization to customers, the public, government, and other external sources. This information can be exchanged in person, in writing, or by telephone or e-mail." Although both variables refer to the contact toward people outside the firms, the difference between this variable and the previous one is that this variable measures the overall contact that includes both direct in-person contact and indirect contact via phones and emails rather than just direct in-person contact in "Working with the public". 3) "Communication with the coworker" defined as "Providing information to supervisors, co-workers, and subordinates by telephone, in written form, e-mail, or in person." This variable is the same as "Communication with the customer" except that the group of people involved in contact are coworkers. 4) "Face-to-face discussion" defined as "How often do you have to have face-to-face discussions with individuals or teams in this job?". This variable gives a measure of direct face-to-face contact specifically rather than an overall measure of any kinds of contact and it refers to the interaction with coworkers. Appendix Table 3.1A gives a summary of these four contact variables.

Except for "Face-to-face discussion", each contact variable mentioned above has two dimensions: "Level" and "Importance". I integrate these two dimensions into a single value metric using the method from Firpo et al. (2011). Table 3.A2 shows that the correlation among these four contact indices in not high, which eliminates the concern of collinearity when putting them all at once in the regression.

The "Contact Index" is constructed by averaging the four contact indices such that this index summarizes the overall contact level of a job. To simplify the analysis, I recode each continuous contact indices into binary variables. Specifically, I define high contact for a particular contact index as above the 50th percentile of its distribution equally weighted across workers in the CPS.[2] For example, I define high contact as above 67 for the combined index, 67 is the 50th percentile of the distribution.

Table 3.1 is the table of means. Column 1 shows the means for the full sample. Column 2 and Column 3 are the means by level of the "Contact Index". A comparison between these columns shows that jobs with a high contact have a slightly higher hourly wage than low contact jobs but the difference is insignificant. Most of the other variables used in this paper, such as gender, age and union status look similar among

2 The distributions of these contact indices can be found in appendix.

low and high contact jobs. The proportion of college graduates are larger in jobs with a high contact and this is almost always true across four component contact indices, suggesting controlling for education will be important in the analysis.

3.3 Methodology

The main regression equation for this study is:

$$\log\left(hourly\ wage_{ijt}\right) = \alpha_0 + \beta_1 Female_i + \beta_2 High\ Contact_j + \beta_3 Female_i * High\ Contact_j + \beta_4 Age_{it} + \beta_5 Female_i * Age_{it} + \beta_6 High\ Contact_j * Age_{it} + \beta_7 Female_i * High\ Contact_j * Age_{it} + \beta_8 X_{ijt} + \gamma_t + \epsilon_{ijt}$$

where i indexes the individual, j indexes the job, and t indexes the year. The binary variable Female is equal to 1 for female workers and 0 for male workers. The binary variable High Contact measure the contact level for its occupation and it is 1 for jobs with a high contact level and 0 for jobs with a low contact level defined as above or below the median of the distribution within the CPS. X are control variables, including: education, union status, state fixed effects, industry fixed effects. γ are year fixed effects. ϵ is the error term. In my main specification, I include age as linear to ease interpretation. In robustness checks, I also include age as quadratic, including in the interaction terms.

The coefficient β_7 on the interaction term "Female*High Contact*Age" is the coefficient of interest in this study. It shows how the slope of the gender wage gap trend in jobs with a high contact level is different from that in jobs with a low contact level. To the extent that beauty discrimination explains some of the trend in the gender wage gap, perceived less attractive older female workers may be more likely to be exposed to discrimination in jobs with more contact with others. In this case the gender wage gap would grow faster in jobs with a high contact than jobs with a low contact, and this coefficient β_7 would be negative.

In addition to the main analysis using the overall "Contact Index" to check how the gender wage gap pattern changes across synthesized contact levels of jobs, I also compares across different types of contact in jobs. The regression equation is the same as the main regression but I replace "High Contact" with its four binary components "High Working with the Public", "High Communication with the customer", "High Communication with the coworker", and "High Face-to-face discussion". If most of the beauty discrimination towards older female workers comes from customers, then we would see a faster growing gender wage gap for jobs with more contact with customers, that is to say we would expect negative coefficients on both the interaction terms "Female*High Working with the public*Age" and "Female*High Communication with customer*Age" and similarly, if coworkers are more likely to be prejudiced, then we would the coefficients of "Female*High Communication with coworker*Age" and "Female*High Face-to-face discussion*Age" are negative. And a further comparison between these two coefficients would allow us to know whether the effect comes from both direct and indirect contact or direct in-person contact only.

I also do a comparative analysis for college graduates and non-college graduates. Kuhn and Shen (2013) points out that employers tend to give up their own prejudice or less cater to their customers' prejudice for positions requiring more skills, which can be reflected in positions asking for more education. Therefore, I would expect that the gender wage gap would grow faster with age for non-college graduates than college graduates.

I also perform two main robustness checks. Since the gender wage gap trend might look different across cohorts, which attributes some unobserved cohort specific factors that are consist for the same cohort over time but is different from cohort to cohort, I would like to control for cohort fixed effects. However, because of the collinearity issue, I cannot control for both the year and cohort fixed effects in a regression with age as one of the independent variables. To solve this problem, I use the method in Hall et al. (2007). I try both year fixed effects and cohort fixed effects in the regression respectively and I find that the results do not change much. The regression results with cohort fixed effects are presented in Appendix Table 3.A3.

I also extend the main regression by adding quadratic age terms to allow more flexibility in the shape of the trend.

3.4 Results

3.4.1 Contact measure based on contact index

Table 3.2 presents regression results with the binary variable "High Contact" as the variable of interest. Five specifications with increasingly detailed controls are presented across columns. All the specifications include the interaction term "Female*High Contact*Age" and its main effects terms. The first specification in Column 1 controls for only the year fixed effects. In Columns 2 through 5, I gradually add first the education; next the union status; then the state fixed effects, and finally the industry fixed effects.

In most columns, coefficients on the interaction term of interest "Female*High Contact*Age" are negative as expected but the magnitude is quite small and insignificant. The coefficient of -0.011 in Column 1 implies that a one-year increase in age increases the gender wage gap more in jobs involving more contact than jobs involving less contact by 0.11% and the result is insignificant. Given the females' annual earnings is on average around 20% less than males' annual earnings, this effect of contact on the gender wage gap trend is relatively small. Adding education as controls in Column 2 renders the coefficient increase to -0.001 and it is insignificant. This tentatively indicates that workers working at high contact jobs are more likely workers with relatively less education among whom the difference of the slope of the gender wage gap between working in low contact jobs and high contact jobs is larger. Adding more controls in Column 3 to Column 5 does not change the magnitude a lot.

Figure 1 presents a graphical version of the predicted gender wage gap trend by contact level drawn from Table 3.2 Column 1 and Column 5. The predicted trend without any controls are presented in the left and the predicted trend with all controls are presented in the right. Figure 1 shows that the trend for jobs with a high contact level is slightly steeper than the trend for jobs with a low contact level and the 95% confident interval shown by the shaded area for each trend overlapping with each other indicates that this difference is not statistically significant.

The fact that gender wage gap in jobs with a high contact grows only slightly faster than that in jobs with a low contact and this difference is not statistically significant, suggests several potential implications for the gender wage gap. It might be against the beauty discrimination explanation while in favor of the workplace flexibility explanation or less promotion opportunities for females explanation. However, the overall contact index is a mixed measure of various sources of contact in jobs and the three sources of discrimination might interact with each other (Partridge, 2001), it is possible that the effects of different sources of discrimination offset each other. Therefore, in the following section, I break down the encapsulated contact index into four individual contact indices that indicate different sources of potential discrimination.

3.4.2 Contact measure based on four types of contact indices

Table 3.3 presents the results of the regressions with the four individual contact indices. The five specifications in Table 3.3 are the same as Table 3.2 except that the variables of interest are other four individual contact indices representing different contact types. The coefficients on the interaction term "Female*High Working with the public*Age" are negative and statistically significant in all specifications. The coefficient -0.021 in Column 1 means that a one-year increase in age increases the gender wage gap more in jobs involving more contact than jobs involving less contact by 0.21% and the result is statistically significant at 5% level. Adding more controls in Column 2 to Column 4 does not change the magnitude a lot but further reduce the standard errors and increase the significant level into 1%. This suggests that older females in jobs with a high in-person contact with customers experiences a faster growing wage gap than those in jobs with a low in-person contact with customers. To the extent that beauty discrimination explains the gender wage gap, this suggests the presence of customer discrimination. Adding industry fixed effects in Column 5 reduces the coefficient to -0.008, while part of the trend may be explained by industry effects, we still see a marginally significant increase in the gap with age for jobs with more direct customer contact.

The coefficient on the interaction terms of other three contact types "Female*High communication with customer*Age", "Female*High communication with coworker*Age", and "Female*High face-to-face discussion*Age" are relatively small and insignificant in all specifications. This means that more indirect contact with customers and more contact with coworkers does not significantly lower older females' wage, which suggests that indirect customer discrimination and coworker discrimination might not serve as a main source of beauty discrimination in jobs.

3.4.3 Comparative analysis for college graduates and non-college graduates

Table 3.4 presents the results of a comparison analysis between college graduates and non-graduates. Panel A presents the regression results for college graduates and Panel B presents the regression results for non-graduates. All the specifications are the same as Table 3.2. The coefficients on the interaction term of interest "Female*High Contact*Age" are positive and insignificant in almost all the specifications in Panel A, while they are negative and insignificant in Panel B. The coefficient -0.009 in Column 1 in Panel B means that a one-year increase in age increases the gender wage gap more in jobs involving more contact than jobs involving less contact by 0.09%, which indicates that the gender wage gap grows faster in high contact jobs than low contact jobs but the magnitude is considered quite small.

And I also do a t-test to test the difference. The p values of the t-test of the coefficients of the interaction term "Female*High Contact*Age" in Panel A and Panel B are presented in the bottom row of Table 3.4. The p value is 0.2 in Column 1 without controls. Adding more controls in Column 2 to Column 5 reduce the p value to 0.03. This shows that the effect of contact on the growth rate of the gender wage gap for non-college graduates is statistically significantly different that from college graduates. And the gender wage gap grows faster in jobs with a high contact among non-college graduates.

To further identify the source of difference, I extend the college graduates and non-college graduates comparison by checking the effects of the four individual contact indices instead of the single contact index. Table 3.5 and Table 3.6 present the results of the regressions with the four individual contact indices for college graduates and non-college graduates respectively. None of the coefficients on the interaction terms with the contact indices is statistically significant in Table 3.5 and the magnitudes are quite small. This shows that the trend of the gender wage gap of college graduates does not look quite different across different levels of each type of contact. The coefficients on the interaction term "Female*High working with the public*Age" in Table 3.6 are negative and statistically significant. The coefficient -0.01 in Column 5 means that a one-year increase in age would increase the gender wage gap more by 0.1% in jobs with a high direct contact with customers versus low for non-college graduates. This indicates that for non-college graduates, from the perspective that perceived beauty-related discrimination explains the gender wage gap, the discrimination might come from customers. The coefficients on the interaction terms of other three indices are small and insignificant, indicating that under the explanation of the perceived beauty-related discrimination towards older females without a college degree for the gender wage gap, the source of the discrimination might not be coworkers.

These results show that the impact of customer contact levels on the gender wage gap trend is concentrated on non-college graduates. If beauty discrimination is the explanation, this implies that beauty related discrimination is relevant for lower-skill population but not higher. This is consistent with the wide search theory for high-skill jobs in Kuhn (2014).

3.4.4 Robustness check with quadratic age term

All the previous analysis are based on using only a linear term for age, constraining the gender wage gap trend to be linear. It is worth noting that the shape of the trend is not necessarily linear. I check the robustness of the results by expanding the base model to include the quadratic age term. Table 3.7 presents the results. Since the results with quadratic age is hard to explain, I present the corresponding margins by gender for both low contact jobs and high contact jobs for several age values in Table 3.8. It shows that high contact jobs on average the gender wage gaps grows faster but the differences are insignificant, which is consistent with the results of regressions with only linear age term.

Figure 2 illustrates the quadratic results by graphing the predicted gender wage gap by contact level. The shape of the gender wage gap trend with quadratic age in Figure 2 looks pretty linear and similar to the predicted gender wage gap with linear age term in Figure 1. This shows that the results are robust when extending the model to include quadratic age term.

3.5 Conclusion

In this paper, I use the U.S. Current Population Survey (CPS) 2010-2019 and O*NET to compare and document the difference in the pattern of the gender wage gap trend over age between jobs with different levels and types of contact. I find that the gender wage gap in jobs with more interaction with others grows faster than jobs with less interaction, but the difference is statistically insignificant. For the four components of the contact index, the gender wage gap in jobs with more interaction with customers grows faster than that in jobs with less customer interaction and the result is statistically significant, while jobs with more indirect contact with customer or both direct and indirect contact with coworkers do not see this pattern. I believe perceived beauty-related discrimination fits these patterns. And my results suggest that customers are the source of the beauty discrimination and in-person contact matters more. Comparing between college graduates and non-college graduates, I find that the gender wage gap grows faster in jobs with more contact among non-college graduates, but the result is insignificant.

There are several caveats in my analysis. I am able to measure contact but I cannot directly measure beauty or beauty-related discrimination. Though I do believe that customer beauty-related discrimination is a likely explanation for my results, I cannot prove this directly. In addition, the composition of people that select into high contact jobs might be different from those that select into low contact jobs across age. For example, it might be the case that high contact jobs where older females work are more likely to be jobs with relatively low wage, such as sales or service occupations, while high contact jobs where older males work are more likely to be jobs with relatively high wage, such as managers or chief executives, which would weaken the results.

3.6 Tables

Table 3.1: Means

	All	Contact		Working with the public		Communication with customer		Communication with coworker		Face-to-face discussion	
		Low	High	Low	High	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log(hourly wage)	3.002	2.857	3.129	3.109	2.879	2.776	3.159	2.843	3.198	2.966	3.055
	(0.634)	(0.580)	(0.652)	(0.626)	(0.620)	(0.557)	(0.637)	(0.585)	(0.637)	(0.617)	(0.655)
Female	0.470	0.450	0.488	0.408	0.542	0.436	0.494	0.504	0.429	0.499	0.428
	(0.499)	(0.497)	(0.500)	(0.491)	(0.498)	(0.496)	(0.500)	(0.500)	(0.495)	(0.500)	(0.495)
Age	40.479	39.604	41.247	41.872	38.882	38.736	41.694	39.360	41.861	40.265	40.792
	(12.733)	(13.094)	(12.358)	(12.306)	(13.024)	(13.317)	(12.164)	(13.166)	(12.036)	(12.786)	(12.650)
College graduates	0.311	0.198	0.410	0.342	0.275	0.149	0.423	0.218	0.426	0.288	0.343
	(0.463)	(0.398)	(0.492)	(0.474)	(0.447)	(0.357)	(0.494)	(0.413)	(0.494)	(0.453)	(0.475)
Married	0.680	0.638	0.716	0.723	0.630	0.606	0.730	0.635	0.734	0.670	0.693
	(0.467)	(0.481)	(0.451)	(0.448)	(0.483)	(0.489)	(0.444)	(0.481)	(0.442)	(0.470)	(0.461)
Not member or covered	0.901	0.882	0.918	0.893	0.910	0.889	0.910	0.899	0.903	0.883	0.927
	(0.299)	(0.322)	(0.275)	(0.309)	(0.286)	(0.315)	(0.287)	(0.301)	(0.296)	(0.321)	(0.260)
Union member	0.089	0.107	0.073	0.096	0.081	0.101	0.080	0.091	0.086	0.106	0.064
	(0.284)	(0.309)	(0.260)	(0.294)	(0.273)	(0.302)	(0.271)	(0.288)	(0.281)	(0.308)	(0.245)
Union coverage	0.010	0.011	0.009	0.011	0.009	0.010	0.010	0.010	0.011	0.011	0.009
	(0.100)	(0.104)	(0.097)	(0.105)	(0.095)	(0.100)	(0.101)	(0.099)	(0.102)	(0.105)	(0.093)
Observations (%)	100%	47%	53%	53%	47%	41%	59%	55%	45%	59%	41%
Observations (N)	836,283	390,819	445,464	446,726	389,557	343,287	492,996	461,924	374,359	496,209	340,074

Notes: Standard deviations in parentheses. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019."hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below $2 or above $200 with weekly working hours less or equal to 15 hours. "College graduates" is a binary variable: it equals 1 if the worker has a college degree or above, 0 if else.

Source: CPS 2010-2019.

Table 3.2: Effects of Interpersonal Contact on the Widening of the Gender Wage Gap with Age

	(1)	(2)	(3)	(4)	(5)
Female*High Contact*Age/10	-0.011	-0.001	-0.003	-0.002	0.005
	(0.013)	(0.009)	(0.009)	(0.008)	(0.004)
Female	-0.039	-0.105**	-0.102**	-0.101**	0.014
	(0.035)	(0.045)	(0.044)	(0.045)	(0.025)
High Contact	-0.028	-0.088	-0.100	-0.103	0.063**
	(0.127)	(0.075)	(0.072)	(0.069)	(0.028)
Female*High Contact	0.100	0.077*	0.081*	0.081*	-0.026
	(0.075)	(0.046)	(0.044)	(0.043)	(0.017)
Age/10	0.208***	0.175***	0.169***	0.171***	0.150***
	(0.019)	(0.016)	(0.016)	(0.017)	(0.006)
Female*Age/10	-0.061***	-0.045***	-0.044***	-0.045***	-0.046***
	(0.009)	(0.008)	(0.008)	(0.009)	(0.005)
High Contact*Age/10	0.050**	0.024	0.029*	0.029*	0.024***
	(0.022)	(0.015)	(0.015)	(0.015)	(0.007)
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.165	0.337	0.341	0.350	0.416

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 836,283 observations. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019."hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below $2 or above $200 with weekly working hours less or equal to 15 hours. "High Contact" is a binary variable: it equals 1 if the job the worker has involves high overall interpersonal contact, 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

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

Table 3.3: Effects of Different Types of Interpersonal Contact on the Widening of the Gender Wage Gap with Age

	(1)	(2)	(3)	(4)	(5)
Female*High Working with the Public*Age/10	-0.021**	-0.021***	-0.020***	-0.020***	-0.008*
	(0.009)	(0.007)	(0.007)	(0.007)	(0.004)
Female*High Communication with Customer*Age/10	-0.006	0.009	0.007	0.007	0.001
	(0.007)	(0.005)	(0.006)	(0.006)	(0.005)
Female*High Communication with Coworker*Age/10	-0.014	-0.011*	-0.011	-0.009	-0.003
	(0.010)	(0.006)	(0.007)	(0.006)	(0.006)
Female*High Face-to-face Discussion*Age/10	-0.004	0.001	0.000	0.001	0.004
	(0.007)	(0.004)	(0.005)	(0.005)	(0.005)
Female	-0.094***	-0.159***	-0.156***	-0.156***	-0.039**
	(0.034)	(0.036)	(0.035)	(0.036)	(0.016)
High Working with the Public	-0.498***	-0.441***	-0.435***	-0.439***	-0.155***
	(0.125)	(0.103)	(0.093)	(0.094)	(0.033)
High Communication with Customer	0.382***	0.214***	0.207***	0.199***	0.144***
	(0.055)	(0.041)	(0.041)	(0.040)	(0.031)
High Communication with Coworker	-0.182	-0.135**	-0.144**	-0.135**	-0.099**
	(0.115)	(0.065)	(0.061)	(0.059)	(0.041)
High Face-to-face Discussion	-0.173**	-0.076*	-0.073*	-0.067*	-0.035
	(0.078)	(0.042)	(0.040)	(0.039)	(0.034)
Female*High Working with the Public	0.145**	0.170***	0.170***	0.173***	0.068***
	(0.058)	(0.048)	(0.044)	(0.045)	(0.024)
Female*High Communication with Customer	-0.059*	-0.048**	-0.046**	-0.044**	-0.054***
	(0.032)	(0.022)	(0.022)	(0.022)	(0.019)
Female*High Communication with Coworker	0.117**	0.086***	0.089***	0.083***	0.035*
	(0.052)	(0.027)	(0.026)	(0.025)	(0.020)
Female*High Face-to-face Discussion	0.117***	0.051**	0.048**	0.046**	0.035*
	(0.035)	(0.021)	(0.020)	(0.020)	(0.020)
Age/10	0.146***	0.135***	0.129***	0.131***	0.128***
	(0.010)	(0.008)	(0.008)	(0.009)	(0.006)
Female*Age/10	-0.044***	-0.035***	-0.034***	-0.034***	-0.037***
	(0.007)	(0.005)	(0.006)	(0.006)	(0.004)
High Working with the Public*Age/10	0.047***	0.036***	0.032**	0.032***	0.003
	(0.016)	(0.013)	(0.013)	(0.012)	(0.007)
High Communication with Customer*Age/10	0.021	-0.001	0.003	0.004	0.019**
	(0.014)	(0.011)	(0.011)	(0.011)	(0.009)
High Communication with Coworker*Age/10	0.049**	0.040***	0.041***	0.039***	0.030***
	(0.020)	(0.013)	(0.013)	(0.013)	(0.010)
High Face-to-face Discussion*Age/10	0.017	0.004	0.007	0.006	0.005
	(0.015)	(0.008)	(0.009)	(0.009)	(0.009)
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.244	0.374	0.379	0.387	0.424
P-Value of equality test	0.622	0.861	0.895	0.771	0.544

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 836,283 observations. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019."hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. "High Working with the Public" ("High Communication with Customer", "High Communication with Coworker", "High Face-to-face Discussion") is a binary variable: it equals 1 if the job the worker has involves high direct contact with the public (high contact with customers, high contact with coworkers, high direct contact with coworkers), 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

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

Table 3.4: Comparative Analysis on Effects of Interpersonal Contact on the Widening of the Gender Wage Gap with Age (College Graduates versus Non-Graduates)

	(1)	(2)	(3)	(4)	(5)
Panel A: College Graduates					
Female*High Contact*Age/10	0.005	0.011	0.011	0.011	0.016**
	(0.012)	(0.011)	(0.011)	(0.011)	(0.007)
Female	-0.137***	-0.145***	-0.146***	-0.148***	-0.029
	(0.043)	(0.040)	(0.040)	(0.038)	(0.026)
High Contact	-0.018	0.002	0.006	-0.002	0.131***
	(0.076)	(0.072)	(0.071)	(0.068)	(0.049)
Female*High Contact	0.112**	0.077	0.076	0.079	-0.022
	(0.056)	(0.051)	(0.050)	(0.049)	(0.028)
Age/10	0.175***	0.160***	0.161***	0.166***	0.160***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.010)
Female*Age/10	-0.050***	-0.046***	-0.045***	-0.045***	-0.049***
	(0.011)	(0.011)	(0.010)	(0.010)	(0.007)
High Contact*Age/10	0.006	-0.002	-0.003	-0.002	-0.001
	(0.017)	(0.017)	(0.017)	(0.016)	(0.011)
R-squared	0.110	0.148	0.148	0.172	0.265
Panel B: Non-college Graduates					
Female*High Contact*Age/10	-0.009	-0.010	-0.011	-0.010	0.001
	(0.009)	(0.009)	(0.008)	(0.008)	(0.004)
Female	-0.071	-0.095*	-0.091*	-0.091*	0.026
	(0.049)	(0.054)	(0.052)	(0.053)	(0.020)
High Contact	-0.143	-0.167*	-0.178**	-0.175**	0.033
	(0.097)	(0.086)	(0.080)	(0.078)	(0.035)
Female*High Contact	0.089	0.094*	0.098**	0.096**	-0.022
	(0.055)	(0.050)	(0.046)	(0.045)	(0.018)
Age/10	0.183***	0.179***	0.169***	0.170***	0.148***
	(0.018)	(0.018)	(0.017)	(0.018)	(0.005)
Female*Age/10	-0.049***	-0.046***	-0.043***	-0.044***	-0.044***
	(0.009)	(0.008)	(0.009)	(0.009)	(0.004)
High Contact*Age/10	0.047***	0.045***	0.050***	0.048***	0.031***
	(0.016)	(0.014)	(0.014)	(0.014)	(0.007)
R-squared	0.160	0.197	0.211	0.218	0.295
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
P-Value of equality test	0.237	0.069	0.040	0.047	0.030

Notes: OLS estimates on college graduates and non-college graduates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019. There are 259,898 college graduates and 576,385 non-college graduates. "hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below $2 or above $200 with weekly working hours less or equal to 15 hours. "High Contact" is a binary variable: it equals 1 if the job the worker has involves high overall interpersonal contact, 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

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

Table 3.5: Comparative Analysis of Effects of Different Types of Interpersonal Contact on the Widening of the Gender Wage Gap with Age (College Graduates)

	(1)	(2)	(3)	(4)	(5)
Female*High Working with the Public*Age/10	-0.000	-0.003	-0.003	-0.003	0.003
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Female*High Communication with Customer*Age/10	-0.017*	-0.007	-0.007	-0.009	-0.005
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)
Female*High Communication with Coworker*Age/10	-0.002	-0.011	-0.011	-0.007	0.004
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)
Female*High Face-to-face Discussion*Age/10	0.003	0.008	0.008	0.008	0.008
	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)
Female	-0.155***	-0.154***	-0.154***	-0.151***	-0.031
	(0.045)	(0.044)	(0.044)	(0.044)	(0.031)
High Working with the Public	-0.234**	-0.236**	-0.236**	-0.228**	-0.045
	(0.107)	(0.099)	(0.100)	(0.103)	(0.064)
High Communication with Customer	0.176*	0.173*	0.173*	0.144	0.086
	(0.104)	(0.100)	(0.100)	(0.099)	(0.068)
High Communication with Coworker	-0.058	-0.058	-0.058	-0.037	0.017
	(0.082)	(0.086)	(0.086)	(0.083)	(0.064)
High Face-to-face Discussion	0.025	0.038	0.039	0.054	0.091**
	(0.052)	(0.047)	(0.046)	(0.046)	(0.044)
Female*High Working with the Public	0.026	0.030	0.031	0.030	-0.018
	(0.045)	(0.043)	(0.043)	(0.044)	(0.035)
Female*High Communication with Customer	0.073	0.031	0.031	0.042	-0.002
	(0.051)	(0.047)	(0.047)	(0.048)	(0.038)
Female*High Communication with Coworker	0.025	0.058	0.058	0.038	-0.015
	(0.040)	(0.038)	(0.039)	(0.038)	(0.031)
Female*High Face-to-face Discussion	0.081**	0.044	0.044	0.042	0.024
	(0.034)	(0.030)	(0.030)	(0.030)	(0.028)
Age/10	0.124***	0.110***	0.110***	0.116***	0.125***
	(0.017)	(0.016)	(0.016)	(0.017)	(0.014)
Female*Age/10	-0.031***	-0.027**	-0.027**	-0.028**	-0.039***
	(0.012)	(0.011)	(0.011)	(0.012)	(0.009)
High Working with the Public*Age/10	0.008	0.010	0.010	0.009	-0.008
	(0.018)	(0.017)	(0.017)	(0.018)	(0.012)
High Communication with Customer*Age/10	0.025	0.014	0.014	0.018	0.027**
	(0.016)	(0.015)	(0.015)	(0.016)	(0.013)
High Communication with Coworker*Age/10	0.042**	0.052***	0.052***	0.047**	0.029*
	(0.018)	(0.018)	(0.018)	(0.018)	(0.015)
High Face-to-face Discussion*Age/10	-0.011	-0.016	-0.016	-0.017	-0.013
	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.174	0.212	0.212	0.231	0.282
P-Value of equality test	0.315	0.635	0.633	0.444	0.344

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 259,898 observations. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019. High Working with the Public" ("High Communication with Customer", "High Communication with Coworker", "High Face-to-face Discussion") is a binary variable: it equals 1 if the job the worker has involves high direct contact with the public (high contact with customers, high contact with coworkers, high direct contact with coworkers), 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

*** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Comparative Analysis of Effects of Different Types of Interpersonal Contact on the Widening of the Gender Wage Gap with Age (Non-college Graduates)

	(1)	(2)	(3)	(4)	(5)
Female*High Working with the Public*Age/10	-0.028***	-0.027***	-0.025***	-0.025***	-0.010**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.004)
Female*High Communication with Customer*Age/10	0.005	0.004	0.002	0.003	-0.002
	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)
Female*High Communication with Coworker*Age/10	-0.003	-0.002	-0.001	-0.001	0.002
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Female*High Face-to-face Discussion*Age/10	0.001	0.001	0.001	0.001	0.006
	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Female	-0.145***	-0.169***	-0.164***	-0.163***	-0.033**
	(0.037)	(0.043)	(0.041)	(0.041)	(0.016)
High Working with the Public	-0.459***	-0.478***	-0.461***	-0.465***	-0.161***
	(0.092)	(0.096)	(0.088)	(0.088)	(0.032)
High Communication with Customer	0.168***	0.150***	0.138***	0.142***	0.115***
	(0.050)	(0.048)	(0.046)	(0.045)	(0.034)
High Communication with Coworker	-0.113*	-0.088*	-0.096**	-0.092**	-0.082**
	(0.062)	(0.052)	(0.048)	(0.046)	(0.033)
High Face-to-face Discussion	-0.103**	-0.085**	-0.085**	-0.081**	-0.043
	(0.045)	(0.041)	(0.037)	(0.037)	(0.035)
Female*High Working with the Public	0.197***	0.209***	0.203***	0.205***	0.079***
	(0.043)	(0.046)	(0.043)	(0.043)	(0.020)
Female*High Communication with Customer	-0.034	-0.034	-0.029	-0.032	-0.042**
	(0.025)	(0.024)	(0.023)	(0.022)	(0.018)
Female*High Communication with Coworker	0.057**	0.049**	0.050**	0.050**	0.022
	(0.025)	(0.022)	(0.021)	(0.020)	(0.020)
Female*High Face-to-face Discussion	0.050**	0.040*	0.038*	0.036*	0.022
	(0.023)	(0.022)	(0.020)	(0.020)	(0.021)
Age/10	0.145***	0.142***	0.133***	0.135***	0.132***
	(0.008)	(0.009)	(0.009)	(0.009)	(0.005)
Female*Age/10	-0.039***	-0.035***	-0.033***	-0.034***	-0.037***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.004)
High Working with the Public*Age/10	0.042***	0.042***	0.037***	0.037***	0.006
	(0.011)	(0.012)	(0.012)	(0.012)	(0.007)
High Communication with Customer*Age/10	0.018	0.015	0.019	0.017	0.026**
	(0.014)	(0.013)	(0.012)	(0.012)	(0.010)
High Communication with Coworker*Age/10	0.025*	0.021*	0.021*	0.021*	0.017*
	(0.013)	(0.011)	(0.011)	(0.011)	(0.010)
High Face-to-face Discussion*Age/10	0.005	0.004	0.008	0.007	0.003
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.207	0.238	0.250	0.258	0.303
P-Value of equality test	0.172	0.119	0.165	0.168	0.159

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 576,385 observations. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019. "High Working with the Public" ("High Communication with Customer", "High Communication with Coworker", "High Face-to-face Discussion") is a binary variable: it equals 1 if the job the worker has involves high direct contact with the public (high contact with customers, high contact with coworkers, high direct contact with coworkers), 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

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

Table 3.7: Effects of Interpersonal Contact on the Widening of the Gender Wage Gap with Age (Quadratic)

	(1)	(2)	(3)	(4)	(5)
Female	0.311**	0.264***	0.262**	0.261**	0.319***
	(0.140)	(0.101)	(0.101)	(0.106)	(0.059)
High Contact	-0.497***	-0.433***	-0.414***	-0.411**	-0.286***
	(0.176)	(0.154)	(0.159)	(0.173)	(0.107)
Female*High Contact	-0.002	0.034	0.024	0.017	-0.078
	(0.101)	(0.076)	(0.078)	(0.084)	(0.069)
Age/10	0.976***	0.864***	0.849***	0.850***	0.679***
	(0.138)	(0.070)	(0.073)	(0.075)	(0.041)
Female*Age/10	-0.256***	-0.251***	-0.247***	-0.247***	-0.220***
	(0.073)	(0.045)	(0.045)	(0.047)	(0.027)
High Contact*Age/10	0.249**	0.165**	0.155*	0.152*	0.172***
	(0.102)	(0.079)	(0.082)	(0.089)	(0.054)
Female*High Contact*Age/10	0.063	0.040	0.045	0.049	0.050
	(0.055)	(0.040)	(0.041)	(0.045)	(0.033)
Age^2/100	-0.096***	-0.086***	-0.085***	-0.085***	-0.066***
	(0.015)	(0.007)	(0.008)	(0.008)	(0.005)
Female*Age^2/100	0.025***	0.026***	0.026***	0.025***	0.022***
	(0.008)	(0.005)	(0.005)	(0.005)	(0.003)
High Contact*Age^2/100	-0.020	-0.013	-0.012	-0.011	-0.015**
	(0.012)	(0.009)	(0.009)	(0.010)	(0.006)
Female*High Contact*Age^2/100	-0.011	-0.007	-0.007	-0.008	-0.007*
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
Year Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.201	0.357	0.361	0.370	0.428

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 836,283 observations.The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019."hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below $2 or above $200 with weekly working hours less or equal to 15 hours. "High Contact" is a binary variable: it equals 1 if the job the worker has involves high overall interpersonal contact, 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.

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

Table 3.8: Gender Wage Margins by Contact Level at Ages

	Low Contact		High Contact	
	Male	Female	Male	Female
Age	(1)	(2)	(3)	(4)
18	2.440	2.370	2.403	2.409
	(0.053)	(0.030)	(0.057)	(0.028)
23	2.653	2.506	2.709	2.648
	(0.039)	(0.031)	(0.054)	(0.029)
28	2.830	2.618	2.965	2.843
	(0.031)	(0.033)	(0.052)	(0.033)
33	2.972	2.707	3.169	2.994
	(0.030)	(0.036)	(0.051)	(0.036)
38	3.078	2.772	3.323	3.100
	(0.031)	(0.038)	(0.049)	(0.037)
43	3.148	2.814	3.425	3.163
	(0.032)	(0.039)	(0.046)	(0.038)
48	3.182	2.833	3.477	3.182
	(0.031)	(0.038)	(0.043)	(0.036)
53	3.180	2.827	3.477	3.156
	(0.029)	(0.036)	(0.038)	(0.034)
58	3.143	2.798	3.426	3.086
	(0.027)	(0.033)	(0.034)	(0.031)
63	3.069	2.746	3.324	2.973
	(0.029)	(0.029)	(0.030)	(0.029)

Notes: the Gender Wage Margins at different ages for Table 7.

Figure 3.1: Linear Prediction of the Wage Gap Trends

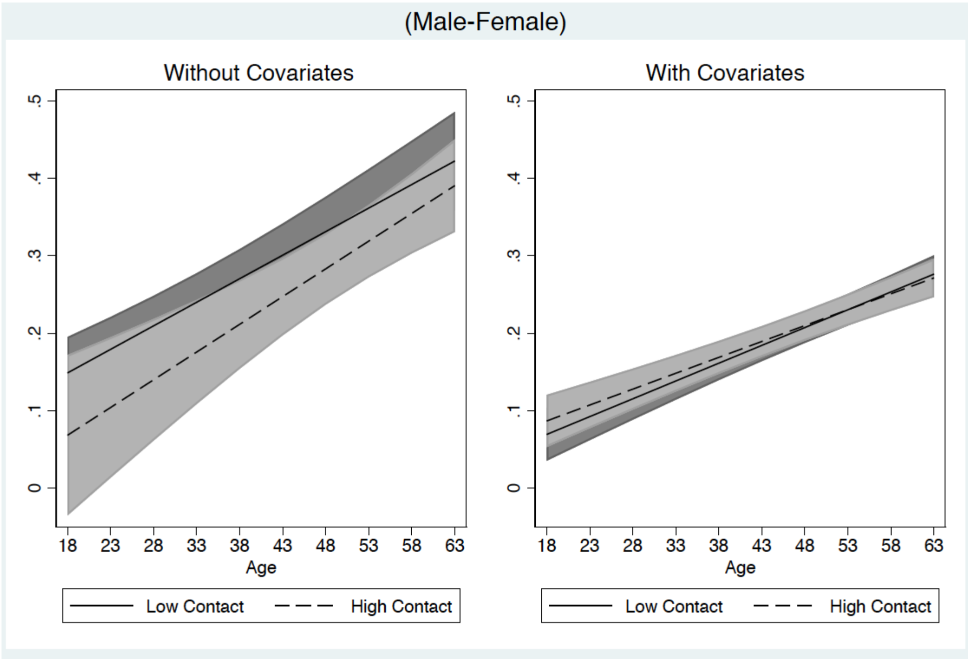
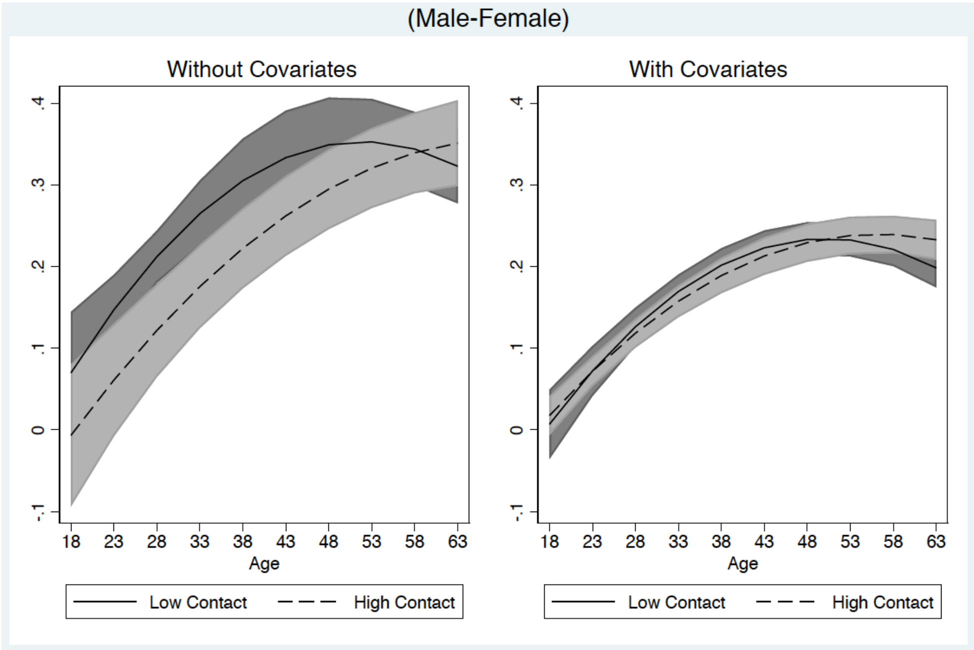


Figure 3.2: Quadratic Prediction of the Wage Gap Trends



3.7 Appendix. Additional Tables and Figures

Table 3.A1: A Summary of the Job Contact Indices

By how / *With whom*	In-person	Overall
Customer	Working with the Public	Communication with Customer
Coworker	Face-to-face Discussion	Communication with Coworker

*Source: the O*NET dataset: https://www.onetcenter.org/database.html#overview*

Table 3.A2: Correlation Matrix of the Job Contact Indices

	Working with the Public (1)	Communication with Customer (2)	Communication with Coworker (3)	Face-to-face Discussion (4)
Working with the Public	1.00			
Communication with Customer	0.29	1.00		
Communication with Coworker	-0.14	0.59	1.00	
Face-to-face Discussion	0.23	0.40	0.21	1.00

Notes: CPS 2010-2019. There are 576,385 observations.

Table 3.A3: Effects of Interpersonal Contact on the Widening of the Gender Wage Gap with Age (Cohort)

	(1)	(2)	(3)	(4)	(5)
Female*High Contact*Age/10	-0.011	-0.001	-0.003	-0.002	0.005
	(0.013)	(0.009)	(0.009)	(0.008)	(0.004)
Female	-0.039	-0.105**	-0.102**	-0.101**	0.014
	(0.035)	(0.045)	(0.044)	(0.045)	(0.025)
High Contact	-0.028	-0.088	-0.100	-0.103	0.063**
	(0.127)	(0.075)	(0.072)	(0.069)	(0.028)
Female*High Contact	0.100	0.077*	0.081*	0.081*	-0.026
	(0.075)	(0.046)	(0.044)	(0.043)	(0.017)
Age/10	0.274***	0.213***	0.210***	0.222***	0.201***
	(0.023)	(0.019)	(0.018)	(0.019)	(0.010)
Female*Age/10	-0.061***	-0.045***	-0.044***	-0.045***	-0.046***
	(0.009)	(0.008)	(0.008)	(0.009)	(0.005)
High Contact*Age/10	0.050**	0.024	0.029*	0.029*	0.024***
	(0.022)	(0.015)	(0.015)	(0.015)	(0.007)
Cohort Fixed Effects	YES	YES	YES	YES	YES
Education	NO	YES	YES	YES	YES
Union	NO	NO	YES	YES	YES
State	NO	NO	NO	YES	YES
Industry	NO	NO	NO	NO	YES
R-squared	0.165	0.337	0.341	0.350	0.416

Notes: OLS estimates. Dependent variable in all columns is "log(hourly wage)". Standard errors clustered by industry are in parentheses. There are 836,283 observations. The sample is workers aged 18-64 who are not self-employed and do not have missing values on age, gender, education, industry, occupation, state, and union status from CPS year 2010-2019."hourly wage" is constructed by dividing the weekly earnings by usual weekly working hours or hours worked last week if usual hours vary as the same way as Hunt and Nunn (2019) do. I adjust the wages to 2019 dollars by applying the CPI-U-RS deflation index. I multiply top coded wages by 1.5 and drop the workers whose weekly wage is below $2 or above $200 with weekly working hours less or equal to 15 hours. "High Contact" is a binary variable: it equals 1 if the job the worker has involves high overall interpersonal contact, 0 if else. There are 16 education categories, 3 union statuses (union member, union coverage, not member or coverage), 51 states, and 271 industries.
**** p<0.01, ** p<0.05, * p<0.1.*

Figure 3.A1: Contact Index Distribution

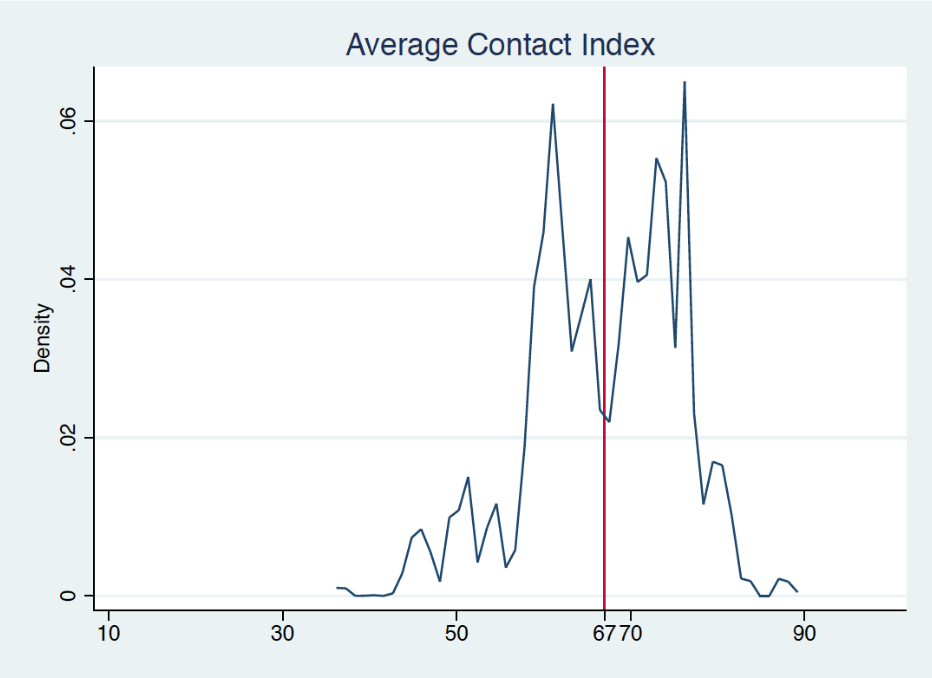


Figure 3.A2: Working with the Public Distribution

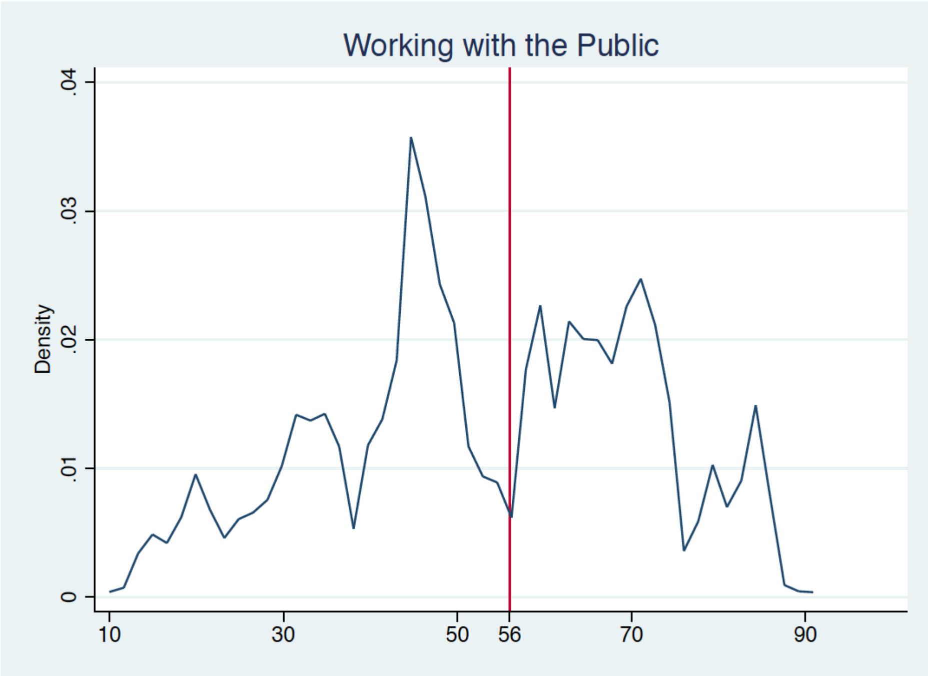


Figure 3.A3: Communication with Customers Distribution

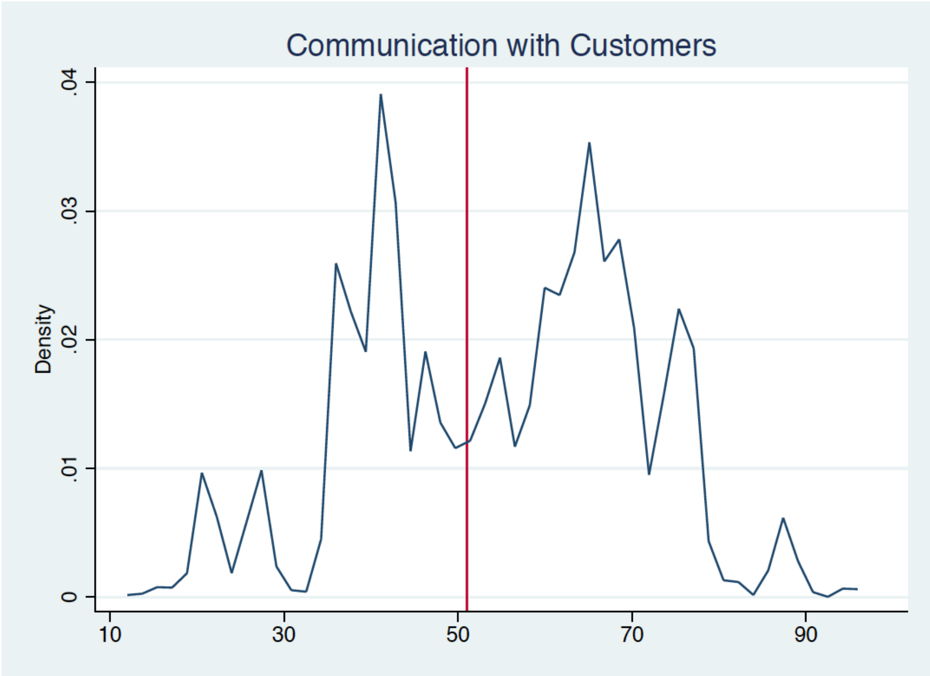


Figure 3.A4: Communication with Coworkers Distribution

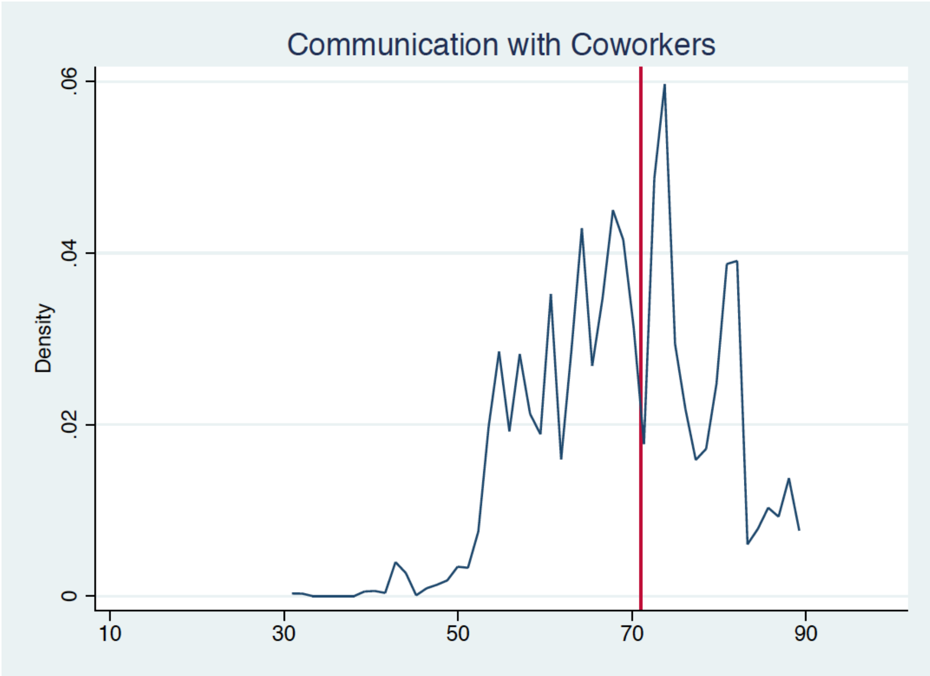
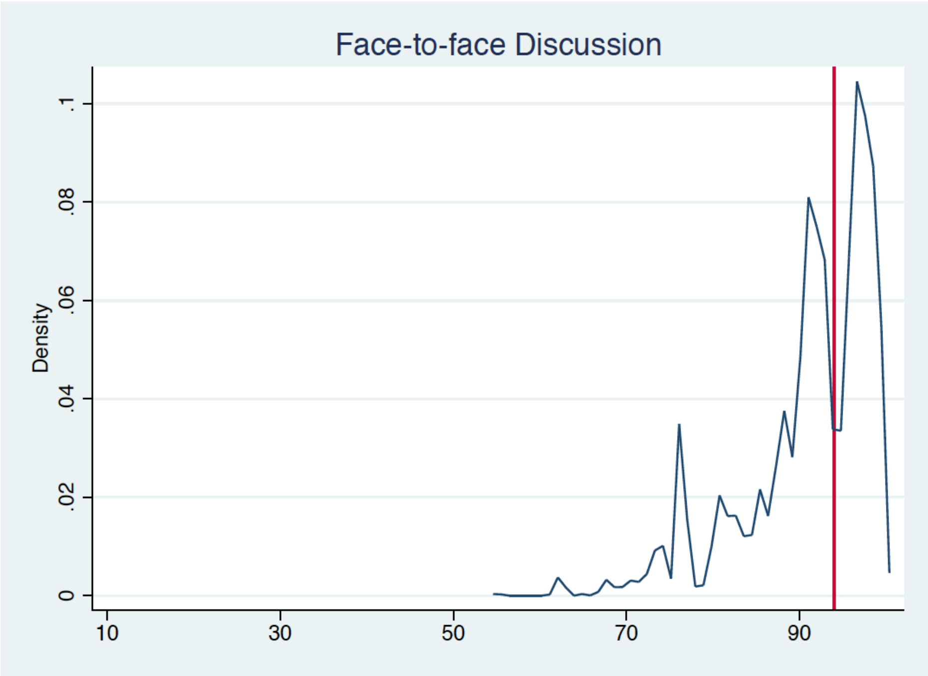


Figure 3.A5: Face-to-face Discussion Distribution



Bibliography

Abraham, Katharine G, and Melissa A Clark. 2006. "Financial aid and students' college decisions evidence from the District of Columbia Tuition Assistance Grant Program". *Journal of Human Resources* 41 (3): 578–610.

Agan, Amanda, and Sonja Starr. 2018. "Ban the box, criminal records, and racial discrimination: A field experiment". *The Quarterly Journal of Economics* 133 (1): 191–235.

Aigner, Dennis J, and Glen G Cain. 1977. "Statistical theories of discrimination in labor markets". *Industrial and Labor Relations Review Review* 30 (2): 175–187.

Akerlof, George A, and Rachel E Kranton. 2008. "Identity, supervision, and work groups". *American Economic Review* 98 (2): 212–17.

Altonji, Joseph G, and Charles R Pierret. 2001. "Employer learning and statistical discrimination". *The Quarterly Journal of Economics* 116 (1): 313–350.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Arrow, Kenneth J. 1972. "Models of job discrimination". *Racial Discrimination in Economic Life* 83.

Autant-Bernard, C, N Massard, and J Mairesse. 2007. "Identifying Age, Cohort and Period Effects in Scientific Research Productivity: Discussion and Illustration Using Simulated and Actual Data on French Physicists". *Papers in Regional Science* 86 (3): 341–350.

Baker, Michael, and Marie Drolet. 2010. "A new view of the male/female pay gap". *Canadian Public Policy* 36 (4): 429–464.

Barron, John M, and John Bishop. 1985. "Extensive search, intensive search, and hiring costs: New evidence on employer hiring activity". *Economic Inquiry* 23 (3): 363–382.

Beam, Emily A, Joshua Hyman, and Caroline Theoharides. 2020. "The Relative Returns to Education, Experience, and Attractiveness for Young Workers". *Economic Development and Cultural Change* 68 (2): 391–428.

Becker, Gary S. 2010. *The economics of discrimination*. University of Chicago press.

Beller, Andrea, George Borjas, Marta Tienda, David Bloom, and Gilles Grenier. 1994. "Beauty and the labor market". *The American Economic Review* 84 (5): 1174–1194.

Bertrand, Marianne, and Esther Duflo. 2017. "Field experiments on discrimination". In *Handbook of Economic Field Experiments*, 1:309–393. Elsevier.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz. 2010. "Dynamics of the gender gap for young professionals in the financial and corporate sectors". *American Economic Journal: Applied Economics* 2 (3): 228–55.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination". *American Economic Review* 94 (4): 991–1013.

Biddle, Jeff E, and Daniel S Hamermesh. 1998. "Beauty, productivity, and discrimination: Lawyers' looks and lucre". *Journal of labor Economics* 16 (1): 172–201.

Black, Sandra E, and Philip E Strahan. 2002. "Entrepreneurship and bank credit availability". *The Journal of Finance* 57 (6): 2807–2833.

Blau, Francine D, and Lawrence M Kahn. 2017. "The gender wage gap: Extent, trends, and explanations". *Journal of Economic Literature* 55 (3): 789–865.

— . 2003. "Understanding international differences in the gender pay gap". *Journal of Labor Economics* 21 (1): 106–144.

Bodvarsson, Örn B, and Mark D Partridge. 2001. "A supply and demand model of co-worker, employer and customer discrimination". *Labour Economics* 8 (3): 389–416.

Borjas, George J, and Matthew S Goldberg. 1978. "Biased screening and discrimination in the labor market". *The American Economic Review* 68 (5): 918–922.

Brenčič, Vera. 2012. "Wage posting: evidence from job ads". *Canadian Journal of Economics* 45 (4): 1529–1559.

Card, David, and Thomas Lemieux. 2001. "Dropout and enrollment trends in the postwar period: What went wrong in the 1970s?" In *Risky Behavior among Youths: An Economic Analysis*, 439–482. University of Chicago Press.

Castleman, Benjamin L, and Bridget Terry Long. 2016. "Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation". *Journal of Labor Economics* 34 (4): 1023–1073.

Charles, Kerwin Kofi, and Jonathan Guryan. 2008. "Prejudice and wages: an empirical assessment of Becker's The Economics of Discrimination". *Journal of Political Economy* 116 (5): 773–809.

Cohodes, Sarah R, and Joshua S Goodman. 2014. "Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy". *American Economic Journal: Applied Economics* 6 (4): 251–85.

Combes, Pierre-Philippe, Bruno Decreuse, Morgane Laouenan, and Alain Trannoy. 2016. "Customer discrimination and employment outcomes: theory and evidence from the french labor market". *Journal of Labor Economics* 34 (1): 107–160.

Cornwell, Christopher, David B Mustard, and Deepa J Sridhar. 2006. "The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program". *Journal of Labor Economics* 24 (4): 761–786.

Daniel, Fernández-Kranz, Aitor Lacuesta, and Núria Rodríguez-Planas. 2013. "The motherhood earnings dip: Evidence from administrative records". *Journal of Human Resources* 48 (1): 169–197.

Delgado Helleseter, Miguel, Peter Kuhn, and Kailing Shen. 2016. "Age and gender profiling in the Chinese and Mexican labor markets: Evidence from four job boards".

Deming, David J. 2017. "The growing importance of social skills in the labor market". *The Quarterly Journal of Economics* 132 (4): 1593–1640.

DesJardins, Stephen L. 1999. "Simulating the enrollment effects of changes in the tuition reciprocity agreement between Minnesota and Wisconsin". *Research in Higher Education* 40 (6): 705–716.

Dwenger, Nadja, Johanna Storck, and Katharina Wrohlich. 2012. "Do tuition fees affect the mobility of university applicants? Evidence from a natural experiment". *Economics of Education Review* 31 (1): 155–167.

Dynarski, Susan. 2000. "Hope for whom? Financial aid for the middle class and its impact on college attendance". National Bureau of Economic Research.

Eeckhout, Jan, and Philipp Kircher. 2010. "Sorting versus screening: Search frictions and competing mechanisms". *Journal of Economic Theory* 145 (4): 1354–1385.

Firpo, Sergio, Nicole M Fortin, and Thomas Lemieux. 2011. "Occupational tasks and changes in the wage structure".

Flory, Jeffrey A, Andreas Leibbrandt, and John A List. 2015. "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions". *The Review of Economic Studies* 82 (1): 122–155.

Friedman, Milton, and Gary S Becker. 1957. "A statistical illusion in judging Keynesian models". *Journal of Political Economy* 65 (1): 64–75.

Giuliano, Laura, and Michael R Ransom. 2013. "Manager ethnicity and employment segregation". *Industrial and Labor Relations Review* 66 (2): 346–379.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in competitive environments: Gender differences". *The Quarterly Journal of Economics* 118 (3): 1049–1074.

Goldberg, Matthew S. 1982. "Discrimination, nepotism, and long-run wage differentials". *The Quarterly Journal of Economics* 97 (2): 307–319.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter". *American Economic Review* 104 (4): 1091–1119.

Goldin, Claudia, and Lawrence F Katz. 2016. "A most egalitarian profession: pharmacy and the evolution of a family-friendly occupation". *Journal of Labor Economics* 34 (3): 705–746.

Goos, Maarten, Alan Manning, and Anna Salomons. 2011. "Explaining job polarization: the roles of technology, offshoring and institutions". *Offshoring and Institutions (December 1, 2011)*.

Greenwood, Michael J. 1997. "Internal migration in developed countries". *Handbook of Population and Family Economics* 1:647–720.

Groen, Jeffrey A. 2004. "The effect of college location on migration of college-educated labor". *Journal of Econometrics* 121 (1-2): 125–142.

Guryan, Jonathan, and Kerwin Kofi Charles. 2013. "Taste-based or statistical discrimination: the economics of discrimination returns to its roots". *The Economic Journal* 123 (572): F417–F432.

Hall, Robert E, and Alan B Krueger. 2010. "Evidence on the determinants of the choice between wage posting and wage bargaining". National Bureau of Economic Research.

Hamermesh, Daniel S, and Jeff E Biddle. 1993. "Beauty and the labor market". National Bureau of Economic Research.

Harper, Barry. 2000. "Beauty, stature and the labour market: A British cohort study". *Oxford Bulletin of Economics and Statistics* 62:771–800.

Helleseter, Miguel Delgado. 2013. "English Skills and Wages in a Non English Speaking Country: Findings from Online Advertisements in Mexico". *Unpublished Manuscript, University of California*.

Helleseter, Miguel Delgado, Peter Kuhn, and Kailing Shen. 2020. "The Age Twist in Employers' Gender Requests Evidence from Four Job Boards". *Journal of Human Resources* 55 (2): 428–469.

Herzog, Serge, and John C Stanley. 2017. "11 Researching the Different Effects of Tuition Policies that Impact Enrollment". *Institutional Research Initiatives in Higher Education*: 187.

Holzer, Harry J, and Keith R Ihlanfeldt. 1998. "Customer discrimination and employment outcomes for minority workers". *The Quarterly Journal of Economics* 113 (3): 835–867.

Hoxby, Caroline M. 2007. *College choices: The economics of where to go, when to go, and how to pay for it*. University of Chicago Press.

— . 2009. "The changing selectivity of American colleges". *Journal of Economic Perspectives* 23 (4): 95–118.

Hunt, Jennifer. 2006. "Staunching emigration from East Germany: Age and the determinants of migration". *Journal of the European Economic Association* 4 (5): 1014–1037.

Hunt, Jennifer, and Ryan Nunn. 2019. "Is Employment Polarization Informative About Wage Inequality and Is Employment Really Polarizing?" National Bureau of Economic Research.

Ihlanfeldt, Keith R, and David L Sjoquist. 1991. "The effect of job access on black and white youth employment: A cross-sectional analysis". *Urban Studies* 28 (2): 255–265.

Ihlanfeldt, Keith R, and Madelyn V Young. 1994. "Intrametropolitan variation in wage rates: the case of Atlanta fast-food restaurant workers". *The Review of Economics and Statistics*: 425–433.

Ingels, Steven J, and Ben Dalton. 2013. "High School Longitudinal Study of 2009 (HSLS: 09) First Follow-Up: A First Look at Fall 2009 Ninth-Graders in 2012. NCES 2014-360." National Center for Education Statistics.

Kahn, Lawrence M. 1991. "Customer discrimination and affirmative action". *Economic Inquiry* 29 (3): 555–571.

Kahn, Lawrence M, and Peter D Sherer. 1988. "Racial differences in professional basketball players' compensation". *Journal of Labor Economics* 6 (1): 40–61.

Kane, Thomas J. 2007. "Evaluating the impact of the DC tuition assistance grant program". *Journal of Human Resources* 42 (3): 555–582.

Kane, Thomas J, and Cecilia Elena Rouse. 1995. "Labor-market returns to two-and four-year college". *The American Economic Review* 85 (3): 600–614.

— . 1999. "The community college: Educating students at the margin between college and work". *Journal of Economic Perspectives* 13 (1): 63–84.

Keller, Wolfgang, and Hâle Utar. 2018. "Globalization, gender, and the family". National Bureau of Economic Research.

Kenney, Genevieve M, and Douglas A Wissoker. 1994. "An analysis of the correlates of discrimination facing young Hispanic job-seekers". *The American Economic Review* 84 (3): 674–683.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and gender inequality: Evidence from Denmark". *American Economic Journal: Applied Economics* 11 (4): 181–209.

Knight, Brian, and Nathan Schiff. 2019. "The out-of-state tuition distortion". *American Economic Journal: Economic Policy* 11 (1): 317–50.

Korenman, Sanders, and David Neumark. 1991. "Does marriage really make men more productive?" *Journal of Human Resources*: 282–307.

— . 1990. "Marriage, motherhood, and wages". National Bureau of Economic Research.

Korthase, Kathleen M, and Irene Trenholme. 1982. "Perceived age and perceived physical attractiveness". *Perceptual and Motor Skills* 54 (3_suppl): 1251–1258.

Kroft, Kory, Fabian Lange, and Matthew J Notowidigdo. 2013. "Duration dependence and labor market conditions: Evidence from a field experiment". *The Quarterly Journal of Economics* 128 (3): 1123–1167.

Kuhn, Peter, and Kailing Shen. 2013. "Gender discrimination in job ads: Evidence from China". *The Quarterly Journal of Economics* 128 (1): 287–336.

LaFave, Allison, Emily Kelly, and Jacob Ford. 2018. "Factors That Influence Student College Choice. Data Point. NCES 2019-119." National Center for Education Statistics.

Lazear, Edward P, and Sherwin Rosen. 1990. "Male-female wage differentials in job ladders". *Journal of Labor Economics* 8 (1, Part 2): S106–S123.

Leonard, Jonathan S, David I Levine, and Laura Giuliano. 2010. "Customer discrimination". *The Review of Economics and Statistics* 92 (3): 670–678.

Leslie, Larry L, and Paul T Brinkman. 1988. *The Economic Value of Higher Education. American Council on Education/Macmillan Series on Higher Education.* ERIC.

Levitt, Steven D. 2004. "Testing theories of discrimination: evidence from Weakest Link". *The Journal of Law and Economics* 47 (2): 431–452.

List, John A. 2004. "The nature and extent of discrimination in the marketplace: Evidence from the field". *The Quarterly Journal of Economics* 119 (1): 49–89.

Loughran, David S, and Julie M Zissimopoulos. 2009. "Why wait? The effect of marriage and childbearing on the wages of men and women". *Journal of Human Resources* 44 (2): 326–349.

Malamud, Ofer, and Abigail Wozniak. 2012. "The impact of college on migration evidence from the Vietnam generation". *Journal of Human Resources* 47 (4): 913–950.

Marsicano, Christopher Russell. 2015. "Regional state tuition exchanges and out-of-state enrollment at public universities and community colleges".

Mathes, Eugene W, Susan M Brennan, Patricia M Haugen, and Holly B Rice. 1985. "Ratings of physical attractiveness as a function of age". *The Journal of Social Psychology* 125 (2): 157–168.

Michelacci, Claudio, and Javier Suarez. 2006. "Incomplete wage posting". *Journal of Political Economy* 114 (6): 1098–1123.

Mobius, Markus M, and Tanya S Rosenblat. 2006. "Why beauty matters". *American Economic Review* 96 (1): 222–235.

Molloy, Raven, Christopher L Smith, and Abigail Wozniak. 2011. "Internal migration in the United States". *Journal of Economic Perspectives* 25 (3): 173–96.

Nardinelli, Clark, and Curtis Simon. 1990. "Customer racial discrimination in the market for memorabilia: The case of baseball". *The Quarterly Journal of Economics* 105 (3): 575–595.

Neumark, David, Roy J Bank, and Kyle D Van Nort. 1996. "Sex discrimination in restaurant hiring: An audit study". *The Quarterly Journal of Economics* 111 (3): 915–941.

Neumark, David, Ian Burn, and Patrick Button. 2019. "Is it harder for older workers to find jobs? New and improved evidence from a field experiment". *Journal of Political Economy* 127 (2): 922–970.

Olivetti, Claudia, and Barbara Petrongolo. 2008. "Unequal pay or unequal employment? A cross-country analysis of gender gaps". *Journal of Labor Economics* 26 (4): 621–654.

Parrett, Matthew. 2011. "Customer discrimination in restaurants: Dining frequency matters". *Journal of Labor Research* 32 (2): 87–112.

Persico, Nicola, and Petra E Todd. 2005. "Passenger profiling, imperfect screening, and airport security". *American Economic Review* 95 (2): 127–131.

Petersen, Trond, Andrew Penner, and Geir Hogsnes. 2006. "The male marital wage premium: sorting versus differential pay".

Phelps, Edmund S. 1972. "The statistical theory of racism and sexism". *The American Economic Review* 62 (4): 659–661.

Pischke, Jorn-Steffen. 2005. "Empirical methods in applied economics: Lecture notes". *Downloaded July* 24:2015.

Reuben, Ernesto, Pedro Rey-Biel, Paola Sapienza, and Luigi Zingales. 2012. "The emergence of male leadership in competitive environments". *Journal of Economic Behavior and Organization* 83 (1): 111–117.

Rizzo, Michael, and Ronald G Ehrenberg. 2004. "Resident and nonresident tuition and enrollment at flagship state universities". In *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*, 303–354. University of Chicago Press.

Robins, Philip K, Jenny F Homer, and Michael T French. 2011. "Beauty and the labor market: Accounting for the additional effects of personality and grooming". *Labour* 25 (2): 228–251.

Rouse, Cecilia Elena. 1994. "What to do after high school: The two-year versus four-year college enrollment decision". *Choices and Consequences: Contemporary Policy Issues in Education* 1994:59–88.

Saks, Raven E, and Abigail Wozniak. 2011. "Labor reallocation over the business cycle: New evidence from internal migration". *Journal of Labor Economics* 29 (4): 697–739.

Sanga, Sarath. 2009. "Reconsidering racial bias in motor vehicle searches: Theory and evidence". *Journal of Political Economy* 117 (6): 1155–1159.

Scholz, John Karl, and Kamil Sicinski. 2015. "Facial attractiveness and lifetime earnings: Evidence from a cohort study". *Review of Economics and Statistics* 97 (1): 14–28.

Shi, Shouyong. 2006. "Wage differentials, discrimination and efficiency". *Contributions to Economic Analysis* 275:83–115.

Shimer, Robert. 2008. "The probability of finding a job". *American Economic Review* 98 (2): 268–73.

Sjoquist, David L, and John V Winters. 2015. "State merit aid programs and college major: A focus on STEM". *Journal of Labor Economics* 33 (4): 973–1006.

Snyder, Thomas D, Cristobal De Brey, and Sally A Dillow. 2019. "Digest of Education Statistics 2017, NCES 2018-070." National Center for Education Statistics.

Stanley, Tom D, and Stephen B Jarrell. 1998. "Gender wage discrimination bias? A meta-regression analysis". *Journal of Human Resources*: 947–973.

Stevenson, Betsey, and Justin Wolfers. 2006. "Bargaining in the shadow of the law: Divorce laws and family distress". *The Quarterly Journal of Economics* 121 (1): 267–288.

Stewart, Gregory, Dianne Brown Wright, and Angelica Kennedy. 2008. "Tuition Reciprocity in the United States". *College and University* 84 (2): 65.

Stigler, George J. 1961. "The economics of information". *Journal of Political Economy* 69 (3): 213–225.

Stinebrickner, Ralph, Todd Stinebrickner, and Paul Sullivan. 2019. "Beauty, job tasks, and wages: A new conclusion about employer taste-based discrimination". *Review of Economics and Statistics* 101 (4): 602–615.

Symons, Donald. 1995. "Beauty is in the adaptations of the beholder: The evolutionary psychology of human female sexual attractiveness". *Sexual Nature, Sexual Culture*: 80–118.

Topel, Robert H, and Michael P Ward. 1992. "Job mobility and the careers of young men". *The Quarterly Journal of Economics* 107 (2): 439–479.

Weidmann, Ben, and David J Deming. 2020. "Team Players: How Social Skills Improve Group Performance". National Bureau of Economic Research.

Wozniak, Abigail. 2010. "Are college graduates more responsive to distant labor market opportunities?" *Journal of Human Resources* 45 (4): 944–970.

Wu, Yuanting. 2020. "Does Customer or Coworker Discrimination Prompt Employer advertisements for Attractive Employees?"

Zhang, Liang, and Erik C Ness. 2010. "Does state merit-based aid stem brain drain?" *Educational Evaluation and Policy Analysis* 32 (2): 143–165.

Zussman, Asaf. 2013. "Ethnic discrimination: Lessons from the Israeli online market for used cars". *The Economic Journal* 123 (572): F433–F468.