2013

Essays in Experimental Labor Economics

Qiqi Wang
Iowa State University

Essays in experimental labor economics

by

Qiqi Wang

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:

Tanya Rosenblat, Co-major Professor

Markus Mobius, Co-major Professor

Joydeep Bhattacharya

Wallace Huffman

Peter Orazem

Iowa State University

Ames, Iowa

2013

DEDICATION

I would like to dedicate this thesis to my parents Qiaochun Wang and Ping Li for their love, encouragement, support and blessings.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my advisors Dr. Tanya Rosenblat and Dr. Markus Mobius. They gave me freedom to pursue interesting projects yet have been stood by me with meticulous guidance and insightful discussions. I particularly thank Tanya for her encouragement and support during hard times, and Markus for the effort he has made for running the experiments. It is my lifetime fortune to have them as mentors, as well as friends.

I am grateful to my committee members, Dr. Peter Orazem, Dr. Wallace Huffman, and Dr. Joydeep Bhattacharya for their advice, suggestions, and contribution to this work.

Special thanks to my dad Qiaochun Wang and his colleagues Dr. Jianping Cheng, Long Chen, ZhengFang Yin, Biao Wang, Chaohong Feng, Zhibo Zhang, Yanli Ma, without whose help I would have not be able to carry out the experiments.

The five years in Ames are colorful with the company of friends. I would like to particularly thank Juan Murguia, his wife Analia Roca, and their lovely daughter Julia Murguia. The first year was tough, but I was able to tackle every kind of difficulties with the encouragement and support of Juan. It is also because of his suggestion to take Experimental Economics I decided to be an experimental economist in the third year.

Last, but not least, I would like to thank my fiancee Xi Liu. She has been my driving force for completing this work, because every time I look into the future, I see happiness.

ABSTRACT

I use experimental methodology to study interactions in the labor market which are otherwise unobservable. In my experimental labor market, "workers" perform a real effort task by solving character puzzles. The worker first solves a single practice puzzle and then is paid to solve as many puzzles as possible in a 5-minute task period. I interpret the puzzles solved in the 5-minute task period as the worker's actual productivity, and the time to complete the single practice puzzle as a noisy signal of that productivity. Based on this noisy signal and other labor market characteristics such as ethnicity, gender, and urban/rural, "employers" are given incentives to estimate the productivity of workers.[1]

Compared to observational data, my experimental methodology has four key advantages. First, I precisely measure the workers' productivity through their ability to solve puzzles. This design eliminates the unobservable factors such as non-cognitive skills, which are likely to affect real labor market outcomes. Second, I build a direct link between the signal and the productivity using the uniform measure of puzzle-solving ability. Third, I explicitly measure the workers' self-confidence with a self-evaluation of their puzzle-solving ability. This measure allows me to study the role of self-confidence in the labor market. Finally, I can construct informative resumes for workers and observe how employers interpret this information when evaluating workers.

In Chapter 1, "Diversity and Discrimination in Experimental Labor Markets," I use this experimental framework to study how stereotyping discrimination against ethnic minorities depends on the shares of ethnic groups in the population. To this purpose, I conduct the experiment with university students in two Chinese provinces: (1) a diverse

[1]Because the experiments are conducted in China, the labor market characteristics I consider are those that are important in the Chinese labor market.

province, where 60% of the population is Han Chinese; and (2) a non-diverse province, where 99% of the population is Han Chinese. The stereotype against ethnic minorities is measured by the employer's estimate of minority workers' productivity.

I find that: (1) Han and minority workers are equally productive; (2) in the non-diverse province, Han employer productivity estimates are significantly lower for minority workers; (3) in the diverse population, a minority worker's productivity is equally estimated by Han and minority employers.

This research furthers our understanding of the economic effects of diversity. It establishes a negative relationship between labor market stereotypes and diversity. Such findings may also provide an explanation for why the inflows of immigrant workers in some US states, like California, have continuously increased. My work suggests that the immigrant workers are looking for diverse communities with lower stereotypes in the labor market.

In Chapter 2, "Self-confidence and Wage in Experimental Labor Markets," I study how signaling self-confidence to employers increases the worker's wage. Self-confidence is an example of a non-cognitive skill, that is likely to be important in the labor market.[2] My experimental framework provides an explicit measure for self-confidence: the worker's evaluation of their own productivity.

I find that for workers, being self-confident is a channel to signal high productivity to employers. Specifically, signaling 1% higher self-confidence to employers increases the employer estimate by 0.09%-0.21%, controlling for other labor market characteristics. The results establish the signaling value of self-confidence in wage negotiations, and highlight the importance of non-cognitive skills in the labor market.

Chapter 3 proposes a methodology to measure the value of worker characteristics. In the design, employers buy worker characteristics in the Becker-DeGroot-Marschak (BDM) market. Specifically, employers claim a willingness to pay (WTP) for a char-

[2]Heckman and Rubinstein [48].

acteristic. This characteristic is displayed on the resume if the WTP is higher than or equal to a randomly determined price. The value of a characteristic is measured by the magnitude of the WTP.

This methodology can be applied in pricing discrimination. The common method to do so is an ex post approach, in which we study discrimination with a wage regression. In such regressions we measure the discriminatory wage differential of a characteristic, which is not related to the actual productivity, by looking at its coefficient in the regression. In our design, the discriminatory wage differential is measured by the WTP of a characteristic.

Although Chapter 4 is not based on the experimental labor market, it serves as a complementary study to Chapter 1. To demonstrate how *indirect contact* can influence economic behavior, in this chapter, I study intergroup cooperation after *observing* in-group members interacting with out-group members.

In the control treatment, a student is matched with someone from the other major in a two-player public goods game. In another treatment, the game players watch intergroup contact prior to the public goods game. The intergroup contact is defined by playing a jigsaw puzzle with someone from the other major. I find that relative to the control treatment, the contribution to the public goods after observing intergroup contact is significantly higher. To distinguish intergroup contact effect from simply putting subjects in a cooperative mood, the game players in a third treatment watch random contact. The results show that it is important to have in-group members in the contact.

The results suggest that indirect contact can be applied when direct contact is restricted. When intergroup cooperation is desired, yet one or more groups are not available, we can select some members from each group and perform demonstrations on the rest. This is particularly useful for majority-minority intergroup cooperation, and for groups that are segregated in many dimensions. Indirect contact also implies financial freedom, as getting every group member involved in direct intergroup contact is very costly.

To summarize, my dissertation contributes to the growing experimental labor market literature. Relative to data from the real labor markets, the experimental labor markets allow us to study otherwise unobservable interactions. With such experimental labor markets, I study the relationship between stereotyping discrimination in the labor market and diversity, the signaling value of self-confidence in wage negotiations, and an alternative methodology to price worker characteristics. In addition, I study the application of indirect contact in raising intergroup cooperation.

CHAPTER 1. DIVERSITY AND DISCRIMINATION IN EXPERIMENTAL LABOR MARKETS

We analyze through experiments how stereotype-based discrimination against ethnic minorities depends on the shares of ethnic groups in the population. In our experimental labor market, "employers" estimate productivity of "workers" who perform a real-effort task. In some treatments, we provide subtle priming to employers about the ethnicity of workers, in addition to providing information on expected productivity. We conduct the experiment with university students in an ethnic non-diverse and an ethnic diverse province in China. We find that: (1) Han and minority workers are equally productive; (2) in the non-diverse population, Han employer estimates are significantly lower for minority workers; (3) in the diverse population, a minority worker is equally estimated by Han and minority employers. Our results establish a negative relationship between stereotype-based discrimination and the share of minorities in the population. It is further suggested that for discriminating employers, revealing the irrelevant ethnicity of workers is not only unhelpful to the labor market, but can deteriorate it by creating discriminatory income inequality between ethnic groups.

1.1 Introduction

Discrimination against ethnic minorities and women is often attributed to *group stereotypes*. According to this narrative, employers systematically believe that workers with certain ethnic or gender characteristics are less intelligent or less skilled than others. It is natural to assume that negative (or positive) stereotypes are more prevalent in communities where the minority represents a smaller share of the population because the majority lacks the chance of learning about the minority. This argument highlights the importance of the diversity-promoting policies that provide people with opportunities to communicate with each other to be able to achieve better understanding and appreciation between different social groups.

Additionally, because of less prevalent stereotypes, it is not surprising that minority workers may favor communities with higher shares of minority groups in the population, resulting in self-selection of minorities into diverse communities. This supports the observation by Borjas [16], that close to three quarters of immigrants aged $18-64$, compared to 50% in 1950, reside in six "immigrant states" in 1990: California, New York, Texas, Florida, New Jersey and Illinois. In particular, California's share of the foreign-born US population increased from 10% in 1950 to 34% in 1990 while its share of the native US population increased from 7% to only 10% during these years.

Finally, the economic gains of rapid industrialization and urbanization in developing countries are likely to be compromised by pre-existing stereotypes. For example, economic growth in China has induced many minority, rural, or low income workers to leave their hometown and migrate to large urban centers where Han and urban residents make up the vast majority of the local population. However, historical segregation by the Huji system may have caused serious stereotypes against migrants, farmers and minorities. These stereotypes are reflected in the labor market as discrimination against those workers, who can enjoy a greater economic welfare if the affiliated stereotypes are mitigated.

We examine how stereotype-based discrimination is related with the shares of ethnic groups in the population through a series of labor market experiments with students at two Chinese universities. One university is located in one of the most ethnically diverse provinces in the west of China, where only 60% of the population is Han Chinese. The second university serves an eastern province, where 99% of the population are Han Chinese.

Our experimental design closely follows Mobius and Rosenblat [68] where an employer sets a worker's wage by guessing the worker's productivity based on some signals. All subjects first play the role of the "worker," followed by the role of the "employer." The worker solves a single practice character puzzle and then is paid to solve as many character puzzles as possible in a 5-minute work period. The time that it takes the worker to complete the single practice character puzzle is named the "signal." The numbers of puzzles that the worker completes in the 5-minute period is referred as the "productivity." Subsequently, when subjects switch to the employer, each estimates 10 workers' productivity. All employers see a mini-"resume" of each worker, which always includes the worker's gender and signal. A random subset of employers can also see each the worker's ethnicity. By comparing the estimates of employers who see workers' ethnicity to the estimates of the control group, we can measure stereotypes.

Compared to observational data, our experimental methodology has two key advantages. First, because the signal and the productivity are measured by the same puzzle-solving ability in our experiment, we can precisely measure the skill level of workers who solve puzzles. Second, we can construct informative "resumes" for workers and observe how employers interpret this information when evaluating workers. These two advantages prevent employers in our experiment from considering unmeasured non-cognitive factors in setting a worker's wage. For example, personal preferences and interpersonal skills may confound empirical studies, using education as an proxy for the productivity.

We find that Han and non-Han subjects perform equally well on the puzzle-solving task. Nevertheless, minority workers in the non-diverse province are judged to perform 9%-10% lower than their Han peers. The source of this unequal treatment are Han employers who lower their ability estimates by up to 14% for minority workers. In contrast, Han employers in the diverse province do not discriminate against minorities.

These findings establish a negative relationship between stereotype-based discrimination and the share of ethnic minorities in the population, and enhance the importance of diversity that has been discussed in other studies. For example, Boisjoly et al. [15] show that white students are likely to show empathy toward all other ethnic minorities when they are randomly assigned one or more African-American roommates. Another study by Moody [69] shows that in American high schools, friendship segregation declines with school heterogeneity levels.

Another result of our study is that the ethnicity of the worker does not matter when it is not displayed even for discriminating employers. This suggests that revealing irrelevant information is not only unhelpful to the labor market, but instead creates discriminatory wage inequality between ethnic groups. We can eliminate the wage inequality by hiding irrelevant worker characteristics from discriminating employers.

Our paper is part of a growing experimental labor market literature. The most closely related work is Mobius and Rosenblat [68] who use a similar experiment to study the origins of the "beauty premium." Bertrand and Mullainathan [12] use a field experiment to study racial discrimination in the U.S. labor market. They construct synthetic resumes and respond to help-wanted advertisements in Boston and Chicago newspapers. Resumes are randomly assigned typical white or African-American names. They find that resumes with white names receive 50% more callbacks for interviews. The results provide explicit evidence for racial discrimination in the U.S. labor market. In the Chinese labor market, Maurer-Fazio [65] uses a similar approach and finds that Han Chinese are much more likely to receive a callback from jobs posted on the internet. Fershtman and Gneezy [36] use lab games to study types of discrimination in Israeli society. Gneezy, Leonard and List [41] compare men and women in a patriarchal society and a matrilineal society. They find that men in the patriarchal society compete more than women but the result is reversed in the matrilineal society. A number of papers have analyzed patterns of discrimination in the Chinese labor market. Zhang [97] compares gender differences in the willingness to enter a competition and finds no differences for Han Chinese but ethnic minority boys compete more than girls. Liu et al. [62] and Dong and Bowles [31] find that Chinese firms discriminate against female workers.

The rest of the paper is organized as follows. Section 2 gives an overview to the Chinese labor market and ethnic discrimination in China. Section 3 introduces the experimental design. Data analysis is discussed in section 4. Our experimental results are presented in section 5 and section 6 concludes.

1.2 Background

1.2.1 The *Huji* System

An important feature of the Chinese labor market is the *Hukou*, which is an official registration record issued to every resident on household basis. The record shows information about the individual, including the holder's name, gender, family status, place of origin, urban/rural status, ethnicity and employer. This system, called *Huji*, has been in use since the year 400 B.C. Its modern version has been applied in China since the year 1949.

Before the 1980s, the Chinese government imposed tough restrictions on population mobility. Most of the restrictions were realized by Hukou inspections. Rural residents were not allowed to migrate to cities; non-local workers were not permitted to look for local jobs; migration across provinces, towns and even counties were restricted. Salaries, infrastructure, commercial commodities and social welfare such as education, health care and pensions usually favor more developed urban centers. This system made urban Hukous highly desirable.

The economic reforms that started in the late 1970s removed many of the restrictions. Rural residents were suddenly free to leave their poor hometown for big cities, in an attempt to benefit from accelerating economic growth. For example, in the capital

Beijing, the share of non-local Hukous in the local population increased from 19% in 2000, to 36% in 2010. The migration to the cities led to rapid urbanization: at the beginning of the economic reforms, only 18% of the population lived in cities. In 2010, this ratio increased to 50%.

Despite these reforms, some Hukou restrictions are still in place to limit mobility. For example, non-local Hukou holders in cities are often limited in housing and automobile markets. Consequently, some Hukou characteristics are still considered inferior.[1] Some studies suggest that the "Hukou gap" is even increasing (Knight and Song [57]).

Within cities, Hukou characteristics are important for wage determination. Meng and Zhang [67] and Lu and Song [63] use case studies to show that ceteris paribus, local urban workers earn more than rural migrant workers. A recent laboratory study by Afridi et al. [2] demonstrates that inferior Hukou characteristics can even hurt self-performance. They ask rural migrant students and local Beijing students to perform a cognitive task and find that when the Hukou status is kept private, there is no difference in the performance between these two groups. However, when it is made salient, rural migrant students' performance decreases by 10%.

1.2.2 Ethnic Minorities

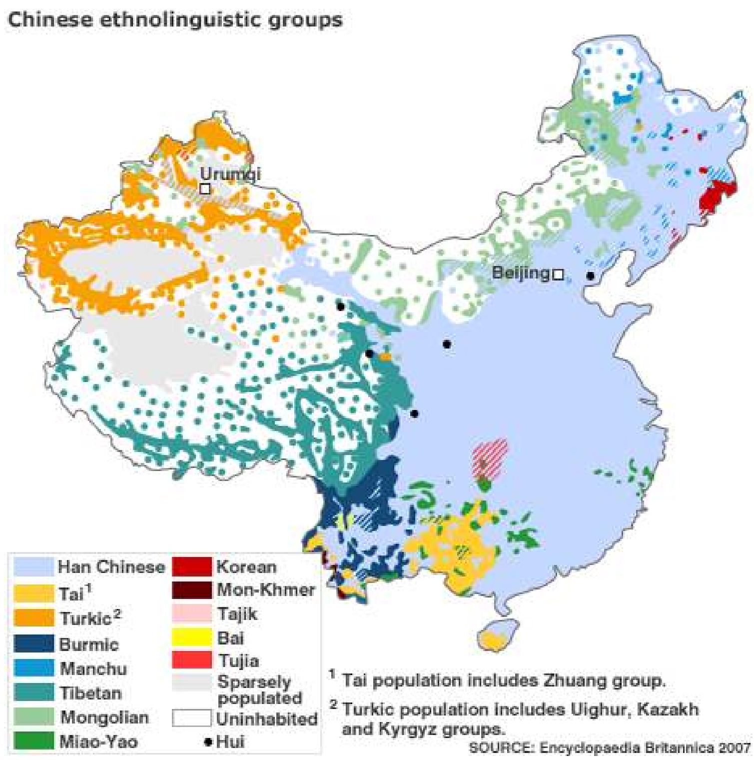
The Chinese government distinguishes between 56 ethnic groups. The largest ethnic group are the Han Chinese who account for 92% of the population. The remaining 114 million people belong to ethnic minorities. Most ethnic minorities are quite small: only 18 groups have populations above 1 million.

While ethnic minorities are distributed across all 31 Chinese provinces and districts, its majority inhabit the western and northeastern regions (see map). In particular, five provinces where about half of the local population are minorities are "ethnic autonomous districts:" Inner Mongolia, Xinjiang, Tibet, Guangxi and Ningxia. These autonomous

[1] See Fleisher and Chen [39], Kanbur and Zhang [55], Lin et al. [62], World Bank [93], and Yao and Zhang [96].

districts are authorized with self-regulation in some inner affairs under compliance with the country's constitution.

Figure 1.1 Map of China



Historically, native minorities are distinguished from Han Chinese in language, tradition, clothes, food, culture, and even physical appearance.[2] However, the economic growth has induced many minorities, like other social groups, to migrate to urban regions. Those migrants tend to mix up with local Han Chinese and look similar to them in various aspects.

1.2.3 Ethnic Discrimination in China

As in many other countries, discrimination against minorities exist in the Chinese labor market. Despite the fact that many migrant minorities look almost identical to

[2]Some minorities might be difficult to distinguish by physical appearance

Han Chinese, employers can easily tell whether a worker is a minority or not because ethnicity is a Hukou element. A substantial body of research has documented ethnic discrimination in China and the wage gap between minorities and Han Chinese.[3] Some of this gap can be attributed to educational attainment: in 2005, for example, 18.5% of Hans, but only 12.5% of minorities had high school diplomas. This education gap is reduced in higher education due to affirmative actions: 5.7% of Hans and 4.2% of minorities hold higher education diplomas. However, the empirical evidence seems to be ambiguous. Some papers find no direct evidence of ethnic discrimination.[4]

There are at least two reasons that can account for the existence of stereotype-based ethnic discrimination. First, the residential segregation between Han Chinese and minorities would contribute to the formation of stereotypes. As mentioned before, ethnic minorities used to inhabit the western and northeastern parts of the country. It is only in the last 30 years that people started migrating. Because most of the migrations happened when one moves to big urban centers, minorities are still rare in many Han inhabitants. Thus, the insufficient contact between Han Chinese and minorities is a primary cause of stereotypes.

Second, minorities benefit from certain affirmative action policies that apply to the labor market, the financial market, schools and universities, politics and even China's birth control policy. For example, minority high school students are subsidized with bonus credits in the national university entrance exam. However, according to Coate and Loury [29], affirmative actions may create negative stereotypes against minorities.

[3] See Gustafsson and Shi [45], Hasmath et al. [46], Johnson and Chow [54], and Li [59].

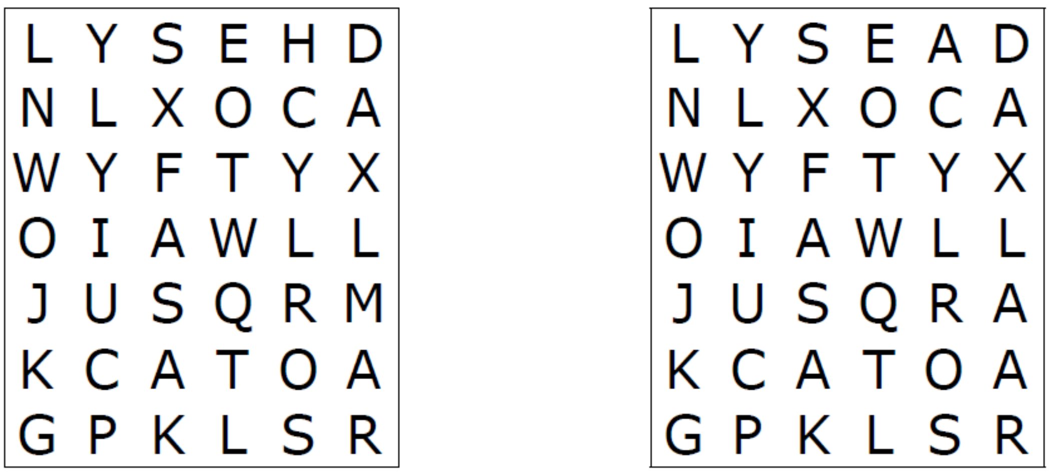
[4] See Appleton et al. [5] and Yang [95].

1.3 The Experiment

1.3.1 Experimental Design

We simulate a labor market where the "employer" determines the wage of the "worker." All subjects start the experiment as the worker who solves character puzzles on the computer. Figure 1.2 shows an example of the character puzzle. Each puzzle shows two

Figure 1.2 Character Puzzle



quadratic arrays of 7 times 7 characters of Latin alphabets. The two arrays are identical except for two random positions where the characters differ. Workers have to find these two locations and click them with their mouse.

In the first step as the worker, each subject is given two warm-up character puzzles. Afterwards, the worker is asked to solve one practice character puzzle. The time that it takes the worker to complete the practice character puzzle, which we will refer from now on as the "signal," is recorded by the experimenter. The signal and other personal information on gender, urban/rural status, ethnicity, and province of origin are used to construct the worker's "resume." In the last step, the worker has a 5-minute work period to solve as many puzzles as possible and is rewarded with 40 points for each solved puzzle.[5] The numbers of puzzles that the worker completes in the 5 minutes are referred to from now on as the worker's "productivity." The sequence of the puzzles in each step is identical for every subject, which means that the subjects are solving the same puzzles appearing in the same order.

The subjects are then switched to the employer, who estimates workers' productivity. The estimated productivity is referred to from now on as the "employer belief" on a worker. Each employer is randomly assigned to one out of four resume treatments. The treatment determines how a worker's resume is displayed to the employer:

TG: The employer sees the signal ("Practice time"on the example resume) and gender.

TGE: The employer sees ethnicity, in addition to TG.

TGEU: The employer sees urban/rural status, in addition of TGE.

TGPU: The employer sees origin of province and urban/rural status, in addition to TG.

Examples of each type of resume are given in Figure 1.3.

Figure 1.3 Examples of Resumes by Treatments

TG

Practice time:	18
Gender:	male
Evaluation:	_

TGE

Practice time:	20
Gender:	female
Ethnicity:	Han
Evaluation:	_

TGEU

Practice time:	22
Gender:	male
Ethnicity:	minority
Urban/Rural:	Urban
Evaluation:	_

TGPU

Practice time:	17
Gender:	female
Origin of Province:	Sichuan
Urban/Rural:	Rural
Evaluation:	_

[5] The experimental points are later converted to cash at a rate of 100 points = 1 Yuan $\simeq$ \$0.16.

Each employer evaluates ten other randomly selected workers and earns 150 points for each evaluated resume. However, if the employer belief is different from a worker's productivity by x puzzles, the earnings are reduced by $10 \cdot x$ points. For example, if a worker solved 20 puzzles in the 5 minutes and the employer's estimate is 18, the employer receives $150 - 10 \cdot |20 - 18| = 130$ points.

The worker receives a wage of 40 points per average employer belief. For example, if a worker is estimated by eight employers and the average employer belief is 20, the worker receives a wage of $40 \cdot 20$ points. Therefore, the worker has two sources of income: the productivity and the employer belief. This provides the worker with an incentive to achieve comparable performance in the timed character puzzle and the 5-minute work period.

1.3.2 Features of Our Design

1.3.2.1 The puzzle

English is a mandatory class in China's education system. It is one of the test subjects of the university entrance exam. Students formally start the course in the first year of middle school. Nevertheless, many parents pay for early English education to get their children well prepared before middle school. This implies that by the time of entrance to the university, the students must have a decent level of English. Thus, the puzzle containing basic Latin characters is simple enough for any university student. Minority, gender, or other characteristics are not supposed to make any difference in puzzle-solving ability. Moreover, the same sequence of puzzles shown to subjects guarantees that the measured ability in solving puzzles is comparable.

1.3.2.2 The signal and the productivity

In our settings, the employer is provided with monetary incentives and a signal to reveal a precise belief on a worker's productivity. The signal thus has to be as informative as possible in predicting the productivity. We interpret the performance on the timed character puzzle as the signal, because the timed character puzzle and the 5-minute work period are highly comparable: (1) the task is the same: solving puzzles from the same type; and (2) the worker is induced to perform identically in both.

Compared to studies with observational data that usually use years of schooling as the signal, ours has the following advantages. First, it better links to the true productivity. For example, college major and performance are frequently unavailable, making years of schooling a less convincing index for the quality of education. Because the worker solves the same type of puzzles for the timed character puzzle and the 5-minute work period, our design provides the employer with a precise measure of the worker's skill.

Second, the lack of unmeasurable non-cognitive factors, such as preference, that are as important as cognitive skills in the labor market according to Heckman and Rubinstein [48], may bias the estimation. For example, personal characteristics like interpersonal skills and personality traits may affect a worker's wage. In addition, certain types of workers may value other things more than their work.[6] Some women for example, may place childbearing in priority instead of training or promotion opportunities. Other workers may receive lower wages reflected by their lower willingness to negotiate.[7] To the contrary, the worker's wage in our experiment is designed to be solely related to productivity, leaving no space for the employer to interpret a worker's wage as a function of those unmeasurable factors.

Finally, the exclusion of workers belonging to specific social groups, like minorities, in the early hiring and promotion stage may further bias the estimation. By design our settings avoid such a problem.

[6] Fortin [40]

[7] Babcock and Laschever [7]

1.3.2.3 The dual roles

Of course, the discussed benefits are based on the fact that the employer has good information on what the worker does in the experiment. This is another feature of our design: the dual roles of the subject. Self-experience could provide the employer with the better information on the worker's task than any other descriptive words would do. In the real world, many human resource officers are themselves employees. For instance, university search committee members are mostly professors by themselves.

1.3.3 Data

We conduct our experiment with students at two Chinese universities. One of the universities is located in a non-diverse province and we will refer to it from now on as the "Non-diverse" location. In this province, fewer than 1% of the population are members of ethnic minorities, which is below the national average. The second university is located in a diverse western province, which is considered one of China's most ethnically diverse provinces, and 40% of the population belongs to ethnic minorities. We will refer to this university as the "Diverse "location.

We contacted students through their class supervisors and obtained their consent to participate in an experiment.[8] Due to different lab sizes, subjects at Non-diverse are divided into two sessions whereas subjects in Diverse are divided into four sessions. Each session lasted about 50 minutes and sessions were conducted back to back in order to reduce communication between students about the nature of the experiment.

Table 1.1 presents descriptive statistics of subjects' demographic characteristics. We recruited 281 students from agronomy, forestry and horticulture majors and 276 students from agricultural products, agronomy, Chinese medicinal herbs, environment and resources, horticulture and plant protection majors at Non-diverse and Diverse, respec-

[8] In Chinese universities, the department is divided into classes. Each class has a supervisor whose duty is to oversee a student's curriculum and extracurricular activities.

tively. Participants are all of similar ages as they are all freshmen. We over-sampled ethnic minority students at Non-diverse to obtain a reasonable mix of minority and Han students in our experiment. At Diverse, the share of ethnic minorities among university students reflects the population shares in the province because most students are local. The table also shows the average earnings of subjects from (1) solving puzzles in the 5-minute work period, (2) worker wages averaged on employer estimates, and (3) the employer's earnings from evaluating workers. Combined earnings at both locations were similar: Non-diverse participants earned 21.5 Yuan ($3.3) and Diverse participants earned 22.1 Yuan ($3.4) during the course of the experiment.[9]

1.4 Measuring the Minority Stereotype

1.4.1 Analysis Strategy

Since the minority stereotype is defined as the wrong belief regarding minorities, the criterion to judge its existence in our settings is the comparison between a minority worker's productivity and the employer belief about the minority worker. There is no stereotype if the worker's ethnic status has the same impact on the productivity and the employer belief. To this purpose, we first check to see if ethnic status is a good predictor to a worker's productivity. Then we look at how an employer interprets a worker's ethnic status when evaluating the worker.

The stereotype that can be found from the above strategy is the population mean at each experimental location. However, the different degrees of diversity at the two locations have two effects on the mean minority stereotyping. One is the exposure effect we are interested in: the higher chance of inter-ethnicity interactions at Diverse eliminates minority stereotyping. It is possible that there is no such an exposure effect but we still find a lower mean minority stereotype at Diverse, simply because the share of persons

[9] The opportunity cost of one hour for a university student, e.g. tutoring a school kid, is about 20 Yuan in the two provinces.

with stereotypes in the population is low. A natural way of separating these two effects is a division between Han employers and minority employers. Comparing Hans with Hans at the two experimental locations allows us to focus on the exposure effect.

1.4.2 Econometric Analysis

Following the above identification strategy, we construct the following model:

$$\ln(Productivity^j) = \alpha + \beta \ln(300/Signal^j) + \delta Minority^j + \gamma X^j + \mu^j \quad (1.1)$$

$$\ln(Employer\ Belief^{ij}) = \zeta + \eta \ln(300/Signal^j) + \theta Minority^j + \lambda X^j + E^i + \nu^{ij} \quad (1.2)$$

where $Productivity^j$ is worker j's productivity, $300/Signal^j$ is worker j's signal converted to an equivalent number of puzzles in 5 minutes, $Minority^j$ is a dummy variable with the value of 1 if worker j is a minority and 0 otherwise, X^j is a vector of worker j's other labor market characteristic dummies: female, rural, province of origin, $Employer\ Belief^{ij}$ is employer i's estimate on worker j's productivity, and E^i is a dummy variable with the value of 1 for employer i, and 0 otherwise.

Equation 1.1 is referred to as the "Productivity Regression." It looks for good predictors to a worker's productivity. For example, β measures the additional percents of puzzles a worker would complete in the 5-minute work period if he/she performs 1% better in the converted timed character puzzle. The productivity difference between minorities and Hans is measured by the coefficient δ on the dummy variable $Minority$, of which a positive (negative) value implies a higher (lower) puzzle-solving ability by minorities.

Equation 3.2 is the "Belief Regression," which tells us what an employer is looking for when evaluating a worker. Compared to the Productivity Regression, the explanatory variable E^i is added to capture the employer's personal tastes or preferences, such as generosity or jealousy. The coefficient η means that if a worker performs 1% better in the converted timed character puzzle, the employer would give an η% higher estimate to the numbers of puzzles the worker would complete in 5 minutes. The minority stereotype is captured by the coefficient θ on the variable $Minority$, which is interpreted as compared to a Han worker, the employer would give a θ percent higher/lower, depending on the sign of θ, estimate to the numbers of puzzles a minority worker would complete in the 5-minute work period.

1.5 Results

1.5.1 Overview

In order to have a general picture on Han and minority workers' productivity, we display distributions of numbers of puzzles solved in the 5-minute task period in the following two figures. Figure 1.4 draws the performance of Han and minority workers at Non-diverse. To identify any difference in the performance, we conduct three tests: Bartlett's test for equal variances, t test for equal means, and Kolmogorov-Smirnov test for equal distributions. The statistics reported at the right side of the graph suggest that at the significance level of 5%, the null hypothesis that Han and minority workers have similar performance is not rejected. Likewise, Figure 1.5 suggests that Han and minority workers at Diverse performed similarly in the 5-minute task period as well.

1.5.2 Productivity Regression

Table 1.2 reports the results of the Productivity Regression. The strongest predictor of productivity is the converted timed character puzzle: a one percent increase in ($300/Signal$) significantly raises the productivity by 0.31% at Non-diverse and 0.32 percent at Diverse. Ethnic status and urban/rural have no significant impact on a worker's productivity. The only other resume variable which has any affect is the gender dummy at Diverse.

The results verified that once the performance on the timed character puzzle and

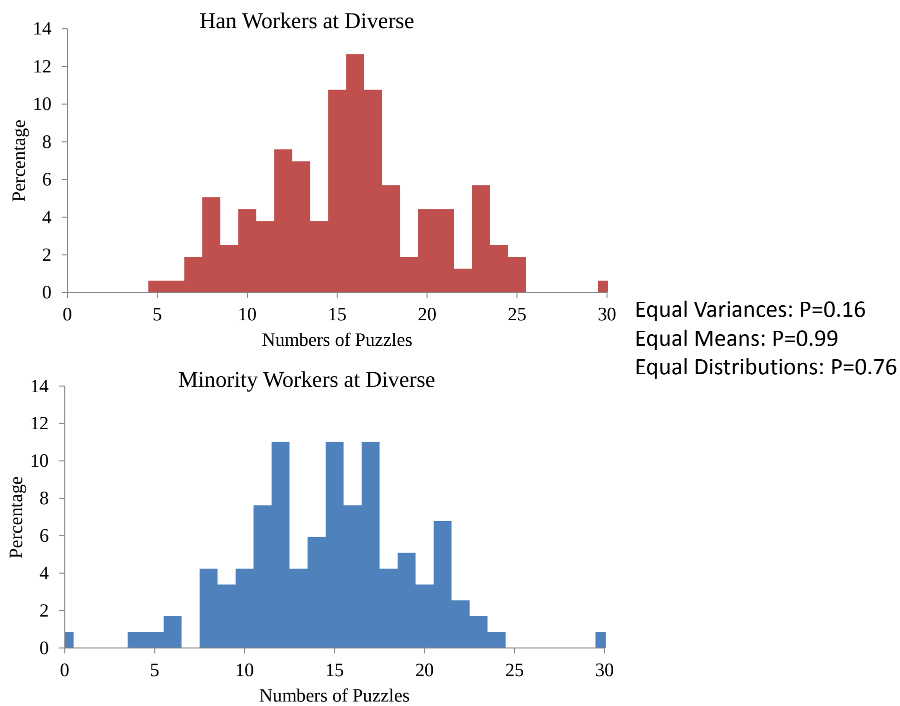
Figure 1.4 Numbers of Puzzles Solved in the 5-Minute Task Period

other characteristics are controlled for, there is no significant difference in the puzzle-solving ability between Hans and minorities at both experimental locations. The next step is to check if ethnic status matters in employer estimates.

1.5.3 Belief Regression

The results for each treatment and experimental locations are shown in Table 1.3. The employers interpret the converted timed character puzzle pretty well: $ln(300/Signal)$ is significantly positive in every treatment and experimental location. For example, 0.35 in Column (1) implies that if a worker performs one percent better in the converted timed character puzzle, the employer gives a 0.35 percent higher estimate to the numbers of puzzles the worker would complete in the 5-minute work period. In particular, employers do a good job of interpreting the signal: the coefficients of $ln(300/Signal)$

Figure 1.5 Numbers of Puzzles Solved in the 5-Minute Task Period



(except Column (3)) are close to their counterpart in the Productivity Regression. The good performance of employers in interpreting the signal implies that the nature of the character puzzle is well understood by subjects.

The variable $Minorty$ is only significantly negative in columns (5) and (7). The two significant coefficients -0.11 and -0.08 suggest that employers at Non-diverse believe that, compared to Han workers, the minority workers would complete 11% and 8% fewer puzzles in the 5-minute work period. However, for employers at Non-diverse who do not see the worker's ethnicity (treatments TG and TGPU) and employers at Diverse (all treatments), minority status does not matter for employer estimates. Therefore, the average minority stereotype at Non-diverse is around 10% while there is no such a stereotype at Diverse.

Comparing the impact of the variable $Minority$ on employer estimates at Non-diverse across the four treatments, its insignificance in treatments where the worker's ethnicity is not displayed (TG and TGPU) implies that it is better to hide this irrelevant information from employers. For minority workers, once their ethnicity is revealed, they receive a lower estimate than Han workers by about 10%. In this case, irrelevant information such as the worker's ethnicity is not only unhelpful to the labor market but deteriorates it by creating discriminatory wage inequality between ethnic groups.

Next, as suggested by the identification strategy, we separate the employers into Han employers and minority employers to focus on the exposure effect. Tables 1.4 and 1.5 are the results of Belief Regression for treatments TG and TGPU, where employers do not see the worker's ethnicity. Except the Han employers in treatment TGPU at Non-diverse, all employers use the converted Signal as a very important predictor to evaluating workers. The coefficients on $ln(300/Signal)$, with the exception in the first column, ranging from 0.21 to 0.57, are all significant at 1%. The ethnic status of a worker does not matter for employer estimates: $Minority$ is not significant across all types of employers and experimental locations. This is straightforward since employers do not see the ethnic status of the worker.

We turn to treatments TGE and TGEU, where employers do see the ethnic status of the worker. Results are displayed in Tables 1.6 and 1.7. The coefficient on $ln(300/Signal)$ is significantly positive across all treatments, experimental locations and types of employers, ranging from 0.15 to 0.35. The coefficients of $Minority$ are now significantly negative only for Han employers at Non-diverse (see -0.14 and -0.08 in the first columns of Table 1.6 and Table 1.7, respectively). This suggests that minority workers are discriminated against only by Han employers at Non-diverse. Minority employers at Non-diverse behave like all employers at Diverse: they give equal estimates to both Hans and minorities. Therefore, the pure exposure effect of diversity eliminates minority stereotypes by 8% and 14%. In addition, revealing the irrelevant ethnicity of a worker to discriminating employers creates an income inequality of 8% and 14% between ethnic groups.

Note that because the coefficients -0.14 and -0.08 are stereotypes of the discriminating Han employers, they serve as a lower bound of the mean stereotype found in Column (5) and Column (7) in Table 1.3, respectively. In other words, the latter should not be significantly lower than the former. This can be verified by checking that the latter coefficients fall in the confidence intervals of the former ones: $(-0.21, -0.01)$ and $(-0.14$, -0.02) respectively.

1.5.4 Checking Structural Differences between Treatments

Because the results are obtained from separating Belief Regression into four treatments, one concern is that the treatments might be structurally different. If this is the case, we have obtained the results simply because of the treatment effect. To check this, we pool across all treatments and include treatment dummies in Belief Regression. The labor market characteristic dummies are modified: a variable is now 1 if it can be observed by employers and its previous dummy value is 1, and 0 otherwise. For example, a minority worker in treatment TG has the value of 0 for the variable $Minority$, because ethnicity is not observed by employers in this treatment. Results are shown in Table 1.8. As no treatment has a significant effect on employer estimates, we conclude that the treatments are structurally identical. Other results are consistent with previous ones. Minority workers at Non-diverse are underestimated by 7%, particularly by Han employers. The minority status does not matter for minority employers at Non-diverse and all employers at Diverse.

1.5.5 Employer Profit

We find in the previous section that minority workers' productivity is underestimated by Han employers at Non-diverse. In this section, we examine from the employer's point of view the next question: do discriminating employers make more (less) profit from evaluating workers' productivity?

The profit that an employer makes from evaluating a worker is defined before as $150 - 10 \cdot |mistake|$ points, where *mistake* is the difference between the worker's actual productivity and the employer belief. Consider the following equation:

$$Employer\ Profit^{ij} = \pi + \rho Minority^j + \phi X^j + \tau^i + \sigma^{ij} \tag{1.3}$$

where $Employer\ Profit^{ij}$ is the profit that employer i makes from evaluating worker j, $Minority^j$ is a dummy variable with the value of 1 if worker j is a minority and 0 otherwise, X^j is a vector of worker j's other labor market characteristic dummies: female, rural, province of origin, and τ^i is the employer fixed effect. In this equation, the change of employer profit between evaluating a Han worker and a minority worker is measured by the coefficient ρ.

Like before, we sort the employers to four types: Han at Non-diverse, minority at Non-diverse, Han at Diverse, and minority at Diverse, and run this regression for each treatment. Tables 1.9, 1.10, 1.11, 1.12 report the results for treatment TG, TGPU, TGE, and TGEU, respectively. Except Han employers at Diverse in treatments TG and TEGU, all employers make the same profit from evaluating Han and minority workers. Thus, despite their stereotype against minority workers, Han employers at Non-diverse do not make lower profits than other types of employers from evaluating worker productivity.

1.5.6 Comparing the Quality of Employer Estimation between Diverse and Non-diverse

We adopt the methodology by Granger [44] and Mankiw and Shapiro [64] to evaluate the quality of employer estimation. Specifically, we examine who are the better forecasters, the Han employers or the minority employers, at Diverse or Non-diverse. Consider the following model:

$$\ln(Productivity^j) = a + b ln(Employer\ Belief^j) + cX^j + d^i + \phi^{ij}, \tag{1.4}$$

and the following hypotheses:

1. a=0;

2. b=1;

3. c=0.

The first two hypotheses state that the estimates are unbiased, because high quality of estimation means that the plots of *Employer Belief* and *Productivity* would look similar. The third hypothesis tests the efficiency of the estimates. Since good employer estimation should have contained all useful information that can predict the actual productivity, the coefficients of these information should be 0 once *Employer Belief* is controlled for. Employers are rational if both unbiasedness and efficiency are satisfied.

Results for treatment TG are displayed in Table 1.13. The constants are significantly higher than 0 in every column. Hence Hypothesis 1 is rejected for the two types of employers at both experimental locations. Although we reject Hypothesis 2 for all four groups of employers as well, employers at Diverse perform better than those at Non-diverse because the coefficient b is significantly higher than 0 and closer to its null hypothesis value of 1.

To test Hypothesis 3, we conduct a joint F test. The first two rows of Table 1.17 report the F statistics and the corresponding P values in the prentices. At the significance level of 5%, Hypothesis 3 is rejected for all four groups of employers.

Regression results for treatments TGPU, TGE, and TGEU are shown in Table 1.14, Table 1.15, and Table 1.16, respectively. Like in treatment TG, Hypothesis 1 is rejected for every type of employers and experimental location, because the constants are all significantly higher than 0. Hypothesis 2 is rejected for all four groups of employers too. Nevertheless, Han employers at Diverse perform better than others in the estimation, as their coefficient b is significantly higher than 0, and closer to 1 than other type of employers.

The last three rows of Table 1.17 show the F statistics for the three treatments. Hypothesis 3 is not rejected only for minority employers at Non-diverse in treatment TGE. Therefore, except those employers, all other employers are not efficient enough.

Combining the results for the three hypotheses, we conclude that the quality of the employer estimation can be easily improved in terms of unbiasedness and efficiency.

1.6 Conclusions

We find that Han employers at Non-diverse systematically underestimate minority workers by about 10%. To the contrary, Han employers at Diverse give equal estimates to both Hans and minorities. The results provide some evidence that experience and social proximity can reduce stereotypes. Our experiment does not imply that social proximity eliminates discrimination: stereotypes are only one source of discrimination in the labor market. Taste-based discrimination can be present even in the absence of stereotypes.

A valuable comment on the result is that, we would consider the self-selection issue. Since randomization in this case is almost impossible, the best way to completely solve the self-selection problem is to look for a difference-by-difference approach. Thus we conduct the experiment with university freshmen the first time at the entrance to the campus, and the second time one or two years later. By looking at the difference between these two time spots, we can know the impact that the diverse location has on individuals.

Another comment on the paper relates the results to the Hawthorne effect, which refers the phenomenon in which subjects modify their behavior when knowing they are experimentally studied. We propose three reasons to argue that there is hardly the Hawthorne effect in our results. First, discriminating against someone is publicly viewed as a negative image. Hence the reasonable modification to the behavior would be hiding the discrimination from instead of showing it to the experimenter. Second, because of the identical experiment manipulation, the Hawthorne effect would show up in both experimental locations. But since employers at Diverse do not discriminate, it is doubtful that only employer at Non-diverse exhibit the Hawthorne effect. Finally, other characteristics like gender and urban/rural are also displayed on the resume. As subjects do not know the real objective of the experiment, which is about ethnicity, the Hawthorne effect would apply to all other characteristics as well. However, we did not find a systematic overestimation or underestimation of other characteristics. Therefore, we conclude that the Hawthorne effect is ignorable in our experiment.

We can consider some extensions to this paper. For example, we ask employers to make hiring decisions instead of estimating productivity. This is a reasonable modification since discrimination is easier to happen in the hiring stage than the wage setting stage. In addition, to study the relationship between economic cycles and intergroup discrimination, we ask employers to select workers to limited vacant positions. During depressions, there are fewer vacant jobs and we check if there is more or less discrimination.

Table 1.1 Summary statistics

Experimental location	Non-diverse		Diverse		
Variable	Mean	Std.Dev		Mean	Std.Dev
Age	19.7	1.1		20.2	1.3
College major					
—Agronomy	0.34	0.5	—Agronomy	0.24	0.4
—Forestry	0.38	0.5	—Agricultural products	0.16	0.4
—Horticulture	0.28	0.5	—Chinese medicinal herbs	0.11	0.3
			—Environment and resources	0.17	0.4
			—Horticulture	0.15	0.4
			—Plant protection	0.16	0.4
Minority province	0.13	0.3		0.01	0.1
Minority	0.18	0.38		0.4	0.5
Local	0.26	0.4		0.85	0.4
Female	0.52	0.5		0.4	0.5
Urban	0.47	0.5		0.3	0.5
Points earned as the worker from the 5-minute work period	642	233		611	190
Points earned as the worker from employer evaluations	750	142		696	130
Points earned as the employer	759	634		900	485

a. Minority province refers to any from the five ethnic autonomous provinces according to China's constitution: Guangxi, Innermongolia, Ningxia, Tibet, Xinjiang.

b. Local refers to students whose province of origin is the experimental province

Table 1.2 Productivity Regression

Dependent variable: ln(Productivity)		
Experimental location	Non-diverse	Diverse
ln(300/Signal)	0.31**	0.32**
	(0.06)	(0.05)
Minority	0.09	-0.02
	(0.07)	(0.04)
Female	0.07	0.09*
	(0.04)	(0.04)
Urban	0.06	0.07
	(0.04)	(0.04)
N	275	272
R^2	0.30	0.26

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 1.3 Belief Regression by Treatments and Experimental Locations

Dependent variable: ln(Employer Belief)								
Resume type	TG		TGPU		TGE		TGEU	
Experimental location	Non-diverse	Diverse	Non-diverse	Diverse	Non-diverse	Diverse	Non-diverse	Diverse
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(300/Signal)	0.38^{**}	0.39^{**}	0.12^{**}	0.38^{**}	0.29^{**}	0.28^{**}	0.27^{**}	0.34^{**}
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)
Minority	-0.00	-0.03	-0.02	0.00	-0.11^{*}	0.00	-0.08^{**}	0.01
	(0.03)	(0.02)	(0.05)	(0.03)	(0.05)	(0.02)	(0.03)	(0.01)
Female	-0.04	0.00	-0.09^{**}	0.02	-0.02	-0.03	-0.04^{*}	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.01)
Urban	0.03	0.00	-0.01	0.08^{**}	-0.03	-0.04	0.01	0.03
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
N	47	48	48	45	48	47	138	137
R^2	0.35	0.33	0.10	0.30	0.20	0.22	0.19	0.26

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.4 Belief Regression: Han vs. Minority Employers in Treatment TG

Dependent variable: ln(Employer Belief)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(300/Signal)	0.38**	0.35**	0.28**	0.57**
	(0.04)	(0.09)	(0.04)	(0.05)
Minority	-0.01	-0.02	-0.02	-0.07
	(0.04)	(0.14)	(0.03)	(0.04)
Female	-0.05	0.10	-0.01	0.03
	(0.03)	(0.08)	(0.03)	(0.04)
Urban	0.04	-0.10	0.01	-0.01
	(0.03)	(0.07)	(0.03)	(0.04)
N	39	8	32	16
R^2	0.35	0.55	0.22	0.59

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.5 Belief Regression: Han vs. Minority Employers in Treatment TGPU

Dependent variable: ln(Employer Belief)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(300/Signal)	0.07	0.53**	0.48**	0.21**
	(0.05)	(0.14)	(0.04)	(0.05)
Minority	-0.00	0.01	0.01	-0.02
	(0.06)	(0.15)	(0.04)	(0.04)
Female	-0.09*	-0.14	0.02	0.03
	(0.04)	(0.10)	(0.03)	(0.04)
Urban	0.01	-0.12	0.07	0.08*
	(0.04)	(0.09)	(0.04)	(0.04)
N	39	9	27	18
R^2	0.11	0.45	0.41	0.19

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.6 Belief Regression: Han vs. Minority Employers in Treatment TGE

Dependent variable: ln(Employer Belief)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(300/Signal)	0.30**	0.33**	0.35**	0.20**
	(0.05)	(0.10)	(0.04)	(0.05)
Minority	-0.14**	-0.01	-0.02	0.02
	(0.05)	(0.13)	(0.04)	(0.03)
Female	-0.02	-0.06	-0.01	0.06
	(0.03)	(0.09)	(0.04)	(0.04)
Urban	-0.04	0.02	-0.05	-0.05
	(0.03)	(0.09)	(0.04)	(0.04)
N	39	9	24	23
R^2	0.21	0.45	0.32	0.17

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.7 Belief Regression: Han vs. Minority Employers in Treatment TGEU

Dependent variable: ln(Employer Belief)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(300/Signal)	0.30**	0.15**	0.35**	0.32**
	(0.02)	(0.06)	(0.03)	(0.03)
Minority	-0.08**	-0.05	0.01	0.01
	(0.03)	(0.06)	(0.02)	(0.02)
Female	-0.03	-0.04	-0.01	-0.01
	(0.02)	(0.05)	(0.02)	(0.02)
Urban	0.00	0.03	0.05*	-0.00
	(0.02)	(0.05)	(0.02)	(0.02)
N	114	24	76	61
R^2	0.23	0.14	0.27	0.25

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.8 Belief Regression: Pooled Across Treatments

Dependent variable: ln(Employer Belief)						
Experimental location	Non-diverse			Diverse		
Employers	All	Han	Minority	All	Han	Minority
ln(300/Signal)	0.28**	0.29**	0.25**	0.34**	0.36**	0.31**
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.02)
Minority_observed	-0.07**	-0.07**	-0.02	-0.01	-0.01	-0.01
	(0.01)	(0.02)	(0.04)	(0.01)	(0.02)	(0.02)
Female	-0.04**	-0.04**	-0.05	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)
Urban_observed	0.00	0.00	0.00	0.04**	0.06**	0.02
	(0.01)	(0.02)	(0.04)	(0.01)	(0.02)	(0.02)
TGE	0.03	0.00	0.15	-0.02	0.05	-0.09
	(0.07)	(0.09)	(0.13)	(0.07)	(0.08)	(0.1)
TGEU	-0.05	-0.07	0.06	-0.01	-0.02	0.00
	(0.06)	(0.07)	(0.11)	(0.06)	(0.06)	(0.11)
TGPU	-0.04	-0.06	0.10	0.20	0.27	-0.05
	(0.10)	(0.11)	(0.22)	(0.15)	(0.17)	(0.28)
N	281	231	50	277	159	118
R^2	0.17	0.18	0.17	0.25	0.28	0.23

a. Baseline treatment: TG.

b. Significance levels of 5% and 1% are denoted by * and **, respectively.

c. Province of origin dummies are included in all columns.

Table 1.9 Profit Regression: Treatment TG

Dependent variable: Employer Profit				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
Minority	-8.28	-16.61	9.15*	-8.03
	(8.04)	(20.61)	(4.12)	(7.30)
Female	0.42	8.36	3.97	-8.61
	(6.11)	(12.44)	(4.24)	(7.80)
Urban	8.16	-8.67	-7.25	-0.58
	(5.94)	(10.81)	(4.68)	(8.32)
N	39	8	32	16
R^2	0.17	0.48	0.05	0.02

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.10 Profit Regression: Treatment TGPU

Dependent variable: Employer Profit				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
Minority	21.71	-8.58	2.46	-4.45
	(14.70)	(20.92)	(3.26)	(5.79)
Female	13.01	8.03	-1.20	-12.67
	(9.42)	(13.65)	(3.21)	(5.96)
Urban	10.98	4.33	0.09	2.13
	(9.23)	(12.16)	(3.50)	(6.18)
N	39	9	27	18
R^2	0.11	0.33	0.04	0.12

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.11 Profit Regression: Treatment TGE

Dependent variable: Employer Profit				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
Minority	4.96	3.89	2.80	0.36
	(8.69)	(30.92)	(6.51)	(4.59)
Female	7.29	24.40	-5.01	6.69
	(6.17)	(20.85)	(6.59)	(4.83)
Urban	4.46	-22.03	5.89	5.48
	(5.94)	(21.12)	(6.98)	(5.00)
N	39	9	24	23
R^2	0.12	0.33	0.03	0.06

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.12 Profit Regression: Treatment TGEU

Dependent variable: Employer Profit				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
Minority	5.99	-9.07	-7.05**	-0.77
	(6.02)	(14.35)	(2.59)	(2.93)
Female	9.64	14.69	-1.30	-5.50
	(4.40)	(10.43)	(2.64)	(2.94)
Urban	2.84	-26.16	-1.93	10.21**
	(4.25)	(10.15)	(2.90)	(3.21)
N	114	24	76	61
R^2	0.05	0.15	0.01	0.03

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.13 Rational Expectation Regression: Treatment TG

Dependent variable: ln(Productivity)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(Employer Belief)	0.08	0.23	0.21**	0.31**
	(0.10)	(0.15)	(0.08)	(0.09)
Minority	-0.09	0.10	0.04	0.01
	(0.07)	(0.15)	(0.04)	(0.06)
Female	0.03	-0.17	0.13**	0.11
	(0.05)	(0.09)	(0.04)	(0.06)
Urban	0.11*	0.12	0.12**	0.00
	(0.05)	(0.08)	(0.04)	(0.07)
Constant	2.27**	2.09**	1.93**	2.21**
	(0.30)	(0.42)	(0.30)	(0.34)
N	39	8	32	16
R^2	0.19	0.58	0.17	0.27

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.14 Rational Expectation Regression: Treatment TGPU

Dependent variable: ln(Productivity)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(Employer Belief)	0.13	0.24	0.24**	0.14
	(0.07)	(0.13)	(0.06)	(0.09)
Minority	0.19*	0.22	0.03	-0.01
	(0.07)	(0.15)	(0.04)	(0.04)
Female	0.06	0.08	0.07	0.04
	(0.05)	(0.10)	(0.04)	(0.04)
Urban	0.14**	0.13	0.08	0.18**
	(0.05)	(0.09)	(0.04)	(0.05)
Constant	2.51**	1.94**	2.10**	2.68**
	(0.23)	(0.42)	(0.26)	(0.37)
N	39	9	27	18
R^2	0.27	0.51	0.22	0.35

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.15 Rational Expectation Regression: Treatment TGE

Dependent variable: ln(Productivity)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(Employer Belief)	0.06	0.00	0.26**	0.15
	(0.07)	(0.16)	(0.08)	(0.09)
Minority	-0.08	-0.13	0.01	-0.05
	(0.06)	(0.16)	(0.05)	(0.05)
Female	0.00	-0.12	0.14**	0.11*
	(0.04)	(0.11)	(0.05)	(0.05)
Urban	0.05	-0.13	0.21**	0.15**
	(0.04)	(0.11)	(0.05)	(0.05)
Constant	2.41**	2.62**	1.77**	2.79**
	(0.21)	(0.50)	(0.28)	(0.35)
N	39	9	24	23
R^2	0.27	0.31	0.38	0.22

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.16 Rational Expectation Regression: Treatment TGEU

Dependent variable: ln(Productivity)				
Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
ln(Employer Belief)	0.04	0.03	0.25**	0.10
	(0.04)	(0.11)	(0.04)	(0.05)
Minority	-0.08*	-0.14	-0.06*	-0.04
	(0.04)	(0.10)	(0.02)	(0.03)
Female	-0.05	-0.04	0.07**	0.09**
	(0.03)	(0.07)	(0.03)	(0.03)
Urban	0.09**	0.06	0.03	0.08*
	(0.03)	(0.07)	(0.03)	(0.03)
Constant	2.55**	2.56**	1.78**	2.48**
	(0.13)	(0.34)	(0.17)	(0.20)
N	114	24	76	61
R^2	0.17	0.26	0.19	0.15

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. Province of origin dummies are included in all columns.

Table 1.17 F Tests for Hypothesis 3

Experimental location	Non-diverse		Diverse	
Employers	Han	Minority	Han	Minority
TG	2.43	2.52	3.49	2.22
	(0.00)	(0.00)	(0.00)	(0.01)
TGPU	3.80	1.84	2.37	4.94
	(0.00)	(0.03)	(0.00)	(0.00)
TGE	3.81	0.96	6.79	3.34
	(0.00)	(0.53)	(0.00)	(0.00)
TGEU	6.39	2.33	6.30	5.45
	(0.00)	(0.00)	(0.00)	(0.00)

a. Corresponding p-values are indicated in the prentices.

CHAPTER 2. SELF-CONFIDENCE AND WAGE IN EXPERIMENTAL LABOR MARKETS

Being self-confident is a channel to signal high ability. We analyze through experiments how signalling higher self-confidence to employers can increase the worker's wage. In our experimental labor market, "employers" estimate productivity of "workers" who perform a real-effort task. To help with the estimation, we provide employers with "resumes" containing worker characteristics like the expected productivity, ethnicity, gender, and urban/rural. In addition, the worker sends to employers a self-evaluation of the productivity. We find that controlling for other factors, more confident workers are predicted to be more productive by employers. Specifically, signalling 1% higher self-confidence increases the employer estimate by 0.14%-0.31%, depending on a treatment. Our results establish the signalling value of self-confidence in wage negotiations, and highlight the importance of non-cognitive skills in the labor market.

2.1 Introduction

Non-cognitive skills are as important as cognitive skills in many dimensions of social performance (Heckman et al. [49]). In some cases, the former is even found to be more important. For example, Duckworth and Seligman [32] find that for school students, self-discipline accounts for more than twice as much variance as IQ in final grades, high school selection, school attendance, hours spent doing homework, hours spent watching television (inversely), and the time of day students began their homework.

As social performance is closely related to productivity, the labor market is a place where the non-cognitive skills matter significantly. Heckman and Rubinstein [48] first pointed out the importance of non-cognitive skills by observing the wage differentials between two groups of high school dropouts: the General Educational Development (GED) recipients and others. The GED is a program that administers exams equating high school dropouts psychometrically to high school graduates. The data reveals that the GED recipients earn more than other high school dropouts but this wage gap is reversed if controlling for schooling factors. Heckman and Rubinstein [48] argue that this is due to the lack of non-cognitive skills of the GED recipients who are cognitively bright. Another evidence comes from Persico et al. [75], who investigate why taller workers receive a wage premium. They find that what matters is the adolescent height but not the adult height. They attribute this height difference in wage to different non-cognitive skills developed in school non-academic activities between tall and short students.

We study the impact of a specific element in the set of non-cognitive skills, the self-confidence, on labor market outcomes. In particular, we examine how signalling higher self-confidence to employers can increase the worker's wage, in the process of wage negotiation. This is identified by Benabou and Tirole [10] as the "signalling value" of demanding self-confidence. Individuals obtain a value from being self-confident, because it is a channel to signal high ability. In simple words, in order to convince others that one has high ability, oneself needs to be convinced. Self-confidence also creates a consumption value, when thinking of oneself favorably makes a person happier. In this case, it is another consumption good that affects one's utility. The motivation value exists because self-confidence improves the individual's motivation to tackle difficulties in the pursuit of his goals, and thus brings more successes.

Our experimental design closely follows Mobius and Rosenblat [68] where an employer sets a worker's wage by guessing the worker's productivity, based on the worker's labor market characteristics. All subjects first play the role of the "worker," followed by the role of the "employer." The worker solves a single practice character puzzle and then is paid to solve as many character puzzles as possible in a 5-minute work period. The time that it takes the worker to complete the single practice character puzzle is named the "signal." The numbers of puzzles that the worker completes in the 5-minute period is referred as the "productivity." Subsequently, workers are informed about their *expected* productivity, which is calculated as $300/signal$, and are asked to give a self-evaluation to their *actual* productivity. The difference between the worker's self-evaluation and his/her expected productivity is referred to as the worker's "self-confidence." When subjects switch to the employer, each estimates 10 workers' productivity. All employers see a mini-"resume" of each worker, which includes the worker's labor market characteristics. Afterwards, the employer has a chance of revising the estimation. For the revision, the employer is provided with the same resume, the previous estimate, plus the worker's self-evaluation.

Compared to observational data, our experimental methodology has four key advantages. First, we precisely measure the skill level of workers with the puzzle-solving ability. Second, with this uniform measure of ability, we build a direct link between the signal and the productivity. Third, we explicitly measure the worker's self-confidence. Finally, we can construct informative resumes for workers and observe how employers interpret this information when evaluating workers.

We find that when self-confidence is explicitly signalled to the employer, more confident workers receive higher employer estimates. The magnitude of the impact is not small: signalling 1% higher self-confidence increases the employer estimate by 0.14%-0.31%, controlling for all other labor market characteristics. Out results provide explicit evidence to how signalling self-confidence can convince employers about one's high ability. Thus we establish the signalling value of self-confidence identified by Benabou and Tirole [10]. It also supports the study by Burks et al. [19], who claim that "overconfidence is induced by the desire to send positive signals to others about one's own skill."

The income gap between some social groups can be explained by our results partly. For example, as men are shown to be more confident than women, our results suggest that men would make more in the labor market.

Out paper is part of the literature that links self-confidence to labor market outcomes. In a similar experimental labor market, Mobius and Rosenblat [68] decompose the beauty premium into oral and visual stereotypes, and self-confidence. Since physical attractiveness is positively associated with self-confidence, more beautiful workers are more confident in the labor market and thus receive a higher wage. Niederle and Vesterlund [70] explain the gender gap in high paid executive positions by investigating men and women's preference for competition. They find that in spite of being equally capable, men are more overconfident and like competition more than women. Falk et al. [34] use an experiment to show that more confident people engage more actively in job search. Empirically, Goldsmith et al. [43] show that a person's wage is more sensitive to changes in self-esteem than to comparable alterations in human capital.

Our paper is also part of a growing experimental labor market literature. Bertrand and Mullainathan [12] use a field experiment to study racial discrimination in the U.S. labor market. They construct synthetic resumes and respond to help-wanted advertisements in Boston and Chicago newspapers. Resumes are randomly assigned typical white or African-American names. They find that resumes with white names receive 50% more callbacks for interviews. The results provide explicit evidence for racial discrimination in the U.S. labor market. In the Chinese labor market, Maurer-Fazio [65] uses a similar approach and finds that Han Chinese are much more likely to receive a callback from jobs posted on the internet.

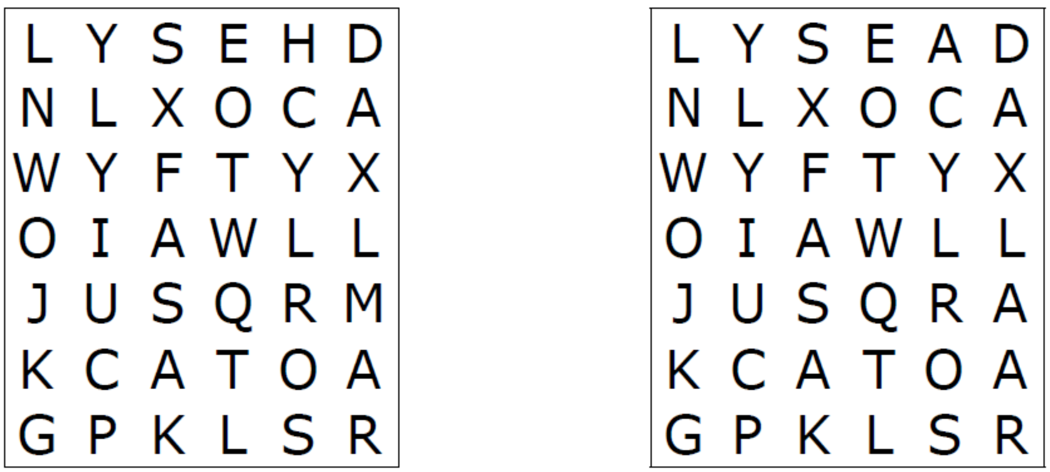
The rest of the paper is organized as follows. Section 2 introduces the experimental design. Results are reported in section 3. Section 4 concludes.

2.2 The Experiment

2.2.1 Experimental Design

We simulate a labor market where the "employer" determines the wage of the "worker." All subjects start the experiment as the worker who solves character puzzles on the computer. Figure 2.1 shows an example of the character puzzle. Each puzzle shows two

Figure 2.1 Character Puzzle



quadratic arrays of 7 times 7 characters of Latin alphabets. The two arrays are identical except for two random positions where the characters differ. Workers have to find these two locations and click them with their mouse.

In the first step as the worker, each subject is given two warm-up character puzzles. Afterwards, the worker is asked to solve one practice character puzzle. The time that it takes the worker to complete the practice character puzzle, which we will refer to from now on as the "signal," is recorded by the experimenter. The signal and other personal characteristics of ethnicity, gender, urban/rural, and origin of province are used to construct the worker's "resume."

In the last step, the worker has a 5-minute work period to solve as many puzzles as possible and is rewarded with 40 points for each solved puzzle.[1] The numbers of puzzles that the worker completes in the 5 minutes are referred to from now on as the worker's "productivity." The sequence of the puzzles in each step is identical for every subject, which means that the subjects are solving the same puzzles appearing in the same order.

Following the practice puzzle but before proceeding to the 5-minute task period, the worker is informed of his/her *expected* productivity, which is calculated as 300 seconds divided by the signal. The worker then is asked to give an estimate to his/her *actual* productivity. We interpret the difference between the self-evaluation and the expected productivity as the worker's self-confidence. When the self-evaluation is higher (lower) than the expected productivity, the worker is over-confident (underconfident). A correct report of the self-evaluation is rewarded with 50 points.

The subjects are then switched to the employer, who estimates workers' productivity. The estimated productivity is referred to from now on as the "employer belief" on a worker. Each employer is randomly assigned to one out of four resume treatments. The treatment determines how a worker's resume is displayed to the employer:

TG: The employer sees the signal ("Practice time"on the example resume) and gender.

TGE: The employer sees ethnicity, in addition to TG.

TGEU: The employer sees urban/rural status, in addition of TGE.

TGPU: The employer sees origin of province and urban/rural status, in addition to TG.

Examples of each type of resume are given in Figure 2.2.

Next, the employers is given a chance of revising the estimation. In addition to what is previously displayed on the resume, the worker's self-confidence is displayed as well. Examples are given in Figure 2.3. We refer to the revised estimation as the "revised employer belief."

[1]The experimental points are later converted to cash at a rate of 100 points = 1 Yuan $\simeq$ \$0.16.

Figure 2.2 Examples of Resumes by Treatments

TG

Practice time:	18
Gender:	male
Evaluation:	_

TGE

Practice time:	20
Gender:	female
Ethnicity:	Han
Evaluation:	_

TGEU

Practice time:	22
Gender:	male
Ethnicity:	minority
Urban/Rural:	Urban
Evaluation:	_

TGPU

Practice time:	17
Gender:	female
Origin of Province:	Sichuan
Urban/Rural:	Rural
Evaluation:	_

Each employer evaluates 10 other randomly selected workers and earns 150 points for each evaluated resume. However, if the employer belief is different from a worker's productivity by x puzzles, the earnings are reduced by $10 \cdot x$ points. For example, if a worker solved 20 puzzles in the 5 minutes and the employer's estimate is 18, the employer receives $150 - 10 \cdot |20 - 18| = 130$ points. For each employer, the actual payoff is determined by a random draw between the employer belief and the revised employer belief.

The worker receives a wage of 40 points per average employer belief, or revised employer belief. For example, if a worker is estimated by eight employers and the average estimation is 20, the worker receives a wage of $40 \cdot 20$ points. Therefore, the worker has two sources of income: the productivity and the employer belief (revised employer belief). This provides the worker with an incentive to achieve comparable performance in the timed character puzzle and the 5-minute work period.

2.2.2 Data

We conduct our experiment with university students in a western province of China. We contacted students through their class supervisors and obtained their consent to par-

Figure 2.3 Examples of Resumes by Treatments

TG	
Practice time:	18
Gender:	Male
Worker's self-evaluation:	12
Your previous evaluation:	15
Your new evaluation:	_

TGE	
Practice time:	20
Gender:	Female
Ethnicity:	Han
Worker's self-evaluation:	20
Your previous evaluation:	18
Your new evaluation:	_

TGEU	
Practice time:	22
Gender:	Male
Ethnicity:	minority
Urban/Rural:	Urban
Worker's self-evaluation:	15
Your previous evaluation:	16
Your new evaluation:	_

TGPU	
Practice time:	17
Gender:	Female
Origin of Province:	Sichuan
Urban/Rural:	Rural
Worker's self-evaluation:	10
Your previous evaluation:	10
Your new evaluation:	_

ticipate in an experiment.[2] Subjects are divided into 4 sessions. Each session lasted about 50 minutes and sessions were conducted back to back in order to reduce communication between students about the nature of the experiment.

Table 2.1 presents descriptive statistics of subjects' labor market characteristics. We recruited 276 students from agricultural products, agronomy, Chinese medicinal herbs, environment and resources, horticulture and plant protection majors. Participants are all of similar ages as they are all freshmen. As many students are local and the province is ethnic-diverse, the share of ethnic minorities among university students reflects the population shares. The female ratio is lower than the national average which is about 48%. The urban ratio is also lower than the national average because the province is a less developed one. The table also shows the average earnings of subjects from (1) solving puzzles in the 5-minute work period, (2) worker wages averaged on employer estimates, and (3) the employer's earnings from evaluating workers. Combining these

[2] In Chinese universities, the department is divided into classes. Each class has a supervisor whose duty is to oversee a student's curriculum and extracurricular activities.

earnings, the subjects received a payoff 22.1 Yuan (\$3.4) in average during the course of the experiment.[3]

2.3 Results

We first verify that self-confidence does not affect productivity. It does, however, increase the employer belief when it is explicitly signalled to the employer.

2.3.1 What Determines the Worker's Self-confidence?

Before proceeding to econometric analysis, we display the distribution of workers' self-confidence in Figure 2.4, of which the horizontal axis draws the difference between the self-evaluation and the expected productivity. By construction, over-confident (under-confident) workers are plotted to the right (left) side of 0. In general, the self-confidence is distributed like a normal distribution, but the center seems to be at 1 instead of 0. Close to 20% of workers believe that they can do 1 more puzzle than their expected productivity. To the right, about 15% of workers are confident that they will do 2 more puzzles than their expected productivity. To the left, about 12% of workers indicated a self-evaluation equal to the expected productivity. Moreover, most of the self-evaluations are within the range of 5 more and fewer puzzles than the expected productivity.

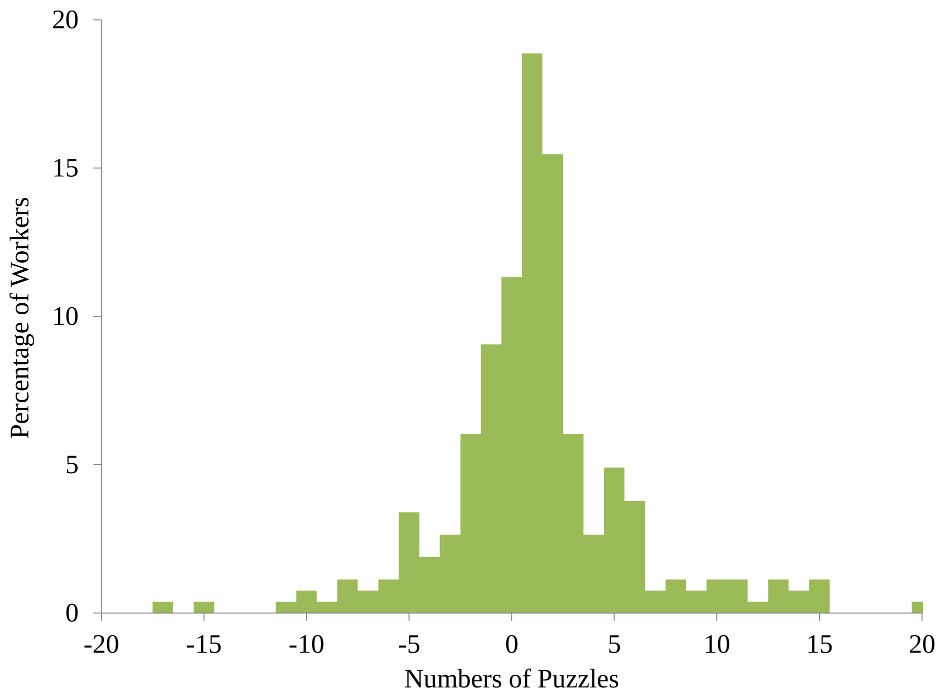
We then analyze what determines the worker's self-confidence in the following equation:

$$\ln(\textit{Self-confidence1}^j) = \alpha + \beta X^j + \mu^j \tag{2.1}$$

where $\textit{Self-confidence1}^j$ is the ratio of worker j's self-evaluation to his/her expected productivity, and X^j is a vector of dummy variables for worker j's labor market characteristics: ethnicity, gender, urban/rural, and origin of province. By definition, worker j is over-confident (under-confident) if $\textit{Self-confidence1}^j$ is higher (lower) than 1.

[3] The opportunity cost of one hour for a university student, e.g. tutoring a school kid, is about 20 Yuan in the two provinces.

Figure 2.4 Distribution of Workers' Self-confidence



We report the results in the first column of Table 2.2. As the coefficients suggest, self-confidence is not determined by the worker's identity characteristics. In the next column, we add the log of the expected productivity, $ln(300/Signal)$. Note that the coefficient -0.26 does not mean that a worker who performs better in the practice puzzles has lower self-confidence. It means that if the expected productivity is increased by 1%, the worker's self-evaluation will still increase, but by only 0.74%. Its minus sign implies, however, that the workers' *over-confidence* (*under-confidence*) is decreasing (increasing) in the expected productivity.

When we add to the regression the actual productivity, which has no impact on self-confidence, the impact of the expected productivity stays significant. However, workers are less over-confident: 1% increase (decrease) in the expected productivity raises (reduces) the self-evaluation by 0.69%.

2.3.2 What Determines Worker Productivity?

To check how the actual productivity is determined, we construct the following model:

$$\ln(Productivity^j) = \delta + \gamma X^j + \nu^j \tag{2.2}$$

where $Productivity^j$ is worker j's productivity, and X^j is the same vector of characteristics in Equation 2.1.

Results are shown in the first column of Table 2.3. For the actual productivity, a female performs 8% better than a male, while a urban person outperforms a rural person by 11%. The difference between urban and rural becomes insignificant when we add the control $ln(300/Signal)$. The better performance of females, however, stays significant. The strongest predictor to worker productivity is how the worker performed in the practice puzzle: a 1% increase in the expected productivity, $(300/Signal)$, significantly raises the productivity by 0.32%. The results stay very close when we include the variable $ln($*Self-confidence1*$)$, which does not affect the actual productivity.

2.3.3 How Employers Make Estimates?

In order to see how employers estimate worker productivity, we look at the following equation:

$$\ln(Employer\ Belief^{ij}) = \eta + \theta X^j + \zeta^i + \epsilon^{ij} \tag{2.3}$$

where $Employer\ Belief^{ij}$ is employer i's belief on worker j's productivity, X^j is the same vector of characteristics in Equation 2.1, and ζ^i is the employer fixed effect controlling for employers' individual characteristics.

In Table 2.4, the first four columns report the results for each of the four treatments, respectively. The results suggest that regardless of the treatment, the most important resume characteristic for estimating a worker is the signal. Increase of 1% in the expected productivity increases the employer estimate by 0.28-0.39%.

In the last column, we pool across all treatments and include treatment dummies. The dummy variables are now with the value of 1 if the original value of 1 is observed, and 0 otherwise. As no treatment has a significant effect on employer estimates, there is no structural difference between the treatments. Results are consistent with the other four columns.

2.3.4 Does Worker's Self-confidence Change Employer Estimates?

We first check the impact of the worker's self-confidence on employer estimates when it is not signalled to the employer. To this purpose, we add the variable *ln*(*Self-confidence1*) as a control in Equation 2.3 and report the results in Table 2.5.

As the coefficient is insignificant in every column, self-confidence does not matter for employer estimation. This is a straightforward result because what the employers face are silent resumes. In other words, there is no way that the employer can either explicitly or implicitly observe the worker's self-confidence.

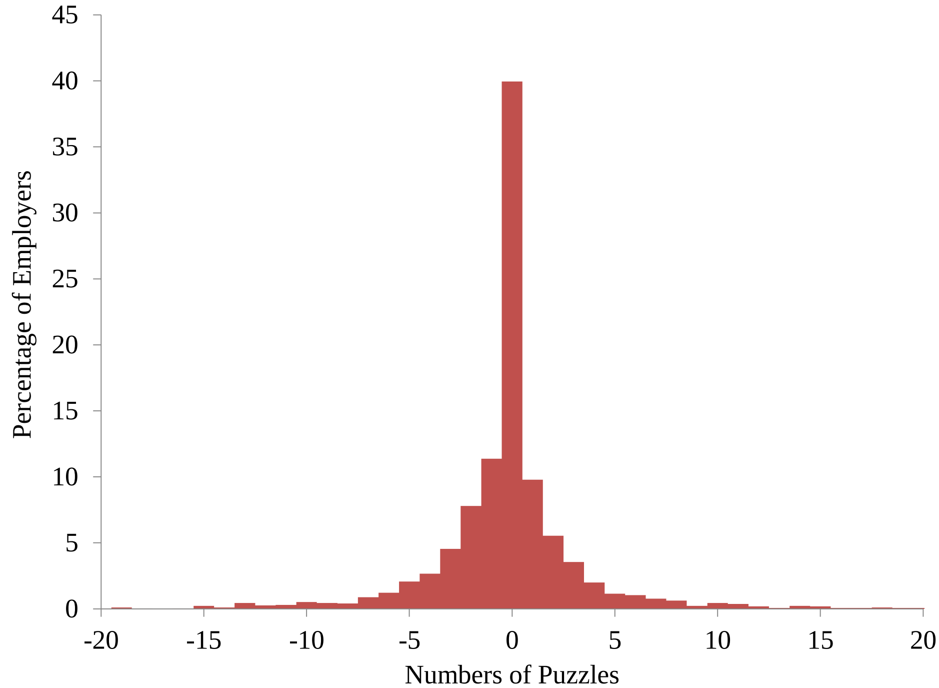
Now we check how self-confidence changes employer estimation when it is explicitly signalled to the employer. We first look at Figure 2.5 which draws the percentages of employers who changed their estimates when observing the worker's self-confidence. As the graph suggests, about 60% of employers made a positive or negative change to their estimates. The distribution looks close to a normal one as the positive side and the negative side are balanced distributed.

To closely examine how employers revise their estimates, we construct the following model:

$$\ln(Revised\,Employer\,Belief^{ij}) = \kappa + \omega X^j + \lambda^i + \upsilon^{ij} \tag{2.4}$$

where $Revised\,Employer\,Belief^{ij}$ is employer i's revised belief on worker j's productivity, X^j is the same vector of characteristics in Equation 2.1, and λ^i is the employer fixed effect controlling for employers' individual characteristics.

Figure 2.5 Distribution of Differences of Employer Belief



We report the results in Table 2.6. The coefficients of the variable $ln(Self-confidence1)$ reveal that for two workers with the same signal, the more confident worker receives a higher wage. Specifically, signalling 1% higher self-evaluation to employers increase the employer estimate by 0.14%-0.31%. Comparing with the impact of the signal, this impact is not small: 1% difference in the signal can be compensated by about 1.5%-4% change in the over-confidence (under-confidence).

In Table 2.7, we replace the identity characteristics and the expected productivity with the employer's previous estimate. The results suggest that even when the wage decision has been made, workers still have the chance to raise his/her wage by signalling higher self-confidence to the employer.

Finally, we combine the explanatory variables from Table 2.6 and Table 2.7 in Table 2.8. Even though we control for all labor market characteristics and the employer's previous estimate, the variable $ln(Self-confidence1)$ remains significant in every treatment. The signalling value of self-confidence is again confirmed.

2.3.5 Do Employers Benefit from Knowing Workers' Self-confidence?

In this section, we study the impact of workers' self-confidence on employer profits. As introduced in the section of experimental design, an employer earns $150-10\cdot|mistake|$ points for each evaluated worker, where $mistake$ is defined by the difference between the worker's actual productivity and the employer belief. To check the change of employer profits from the first offer to the second offer, Table 2.9 displays a comparison of the two profits. At the significance level of 5%, the profit from the second offer is significantly higher than the profit from the first offer. Moreover, there is less variation of profit in the second offer. The distributions of the two profits indicate that to a large degree of extent, the difference between the two profits is due to their difference in the first quartile.

The statistics reveal that employers make more profit from revising their estimates for worker productivity. To distinguish the impact of self-confidence from having the chance of rethinking about the estimation, we construct to following equation:

$$Employer\ Profit^{ij} = \xi + \pi \textit{Second Offer}^{i} + \rho \textit{Self-confidence2}^{j} + \tau^{i} + \sigma^{ij} \tag{2.5}$$

where $Employer\ Profit^{ij}$ is the profit of employer i makes from evaluating worker j, *Second Offer*^{i} is a dummy variable with the value of 0 (1) if the profit is made from first (second) offer, *Self-confidence2*^{j} is the difference between worker j's self-evaluation and his/her expected productivity, and τ^{i} is the employer fixed effect.

Table 2.10 reports the results. Except in treatment TG, employers make significantly higher profits when having the second chance of thinking about their estimation. The negative coefficient of the variable *Self-confidence2* implies that employers make lower profits from evaluating more confident workers. The impact is significant in treatments TG and TGPU, and when pooled across all treatments. This is not a surprising result.

We have seen from the previous section that worker self-confidence can easily affect employer belief. But since the expected productivity is the only factor that matters for predicting the worker's actual productivity, higher self-confidence leads to larger estimation mistakes.

2.3.6 Does Workers' Self-confidence Improve the Quality of Employer Estimation?

We adopt the methodology by Granger [44] and Mankiw and Shapiro [64] to evaluate whether observing workers' self-confidence improves the quality of employer estimation. Consider the following model:

$$\ln(Productivity^j) = a + b\ln(Employer\ Belief^i) + cX^j + d^i + \phi^{ij}. \tag{2.6}$$

If *Employer Belief* are good estimates of *Productivity*, their plots would look similar. We then form the following hypotheses:

1. a=0;
2. b=1;
3. c=0.

The first two hypotheses state that the estimates are unbiased, because high quality of estimation means that the plots of *Employer Belief* and *Productivity* would look similar. The third hypothesis tests the efficiency of the estimates. Since good employer estimation should have contained all useful information that can predict the actual productivity, the coefficients of these information should be 0 once *Employer Belief* is controlled for. Employers are rational if both unbiasedness and efficiency are satisfied.

Results are displayed in the odd number of columns in Table 2.11. The coefficients suggest that all the hypotheses above can be rejected at significance level of 5%. Therefore, the quality of the employer estimates can be easily improved.

To see whether the quality has improved after workers' self-confidence is observed, we modify Equation 2.6 to the following one:

$$\begin{aligned}\ln(Productivity^j) = e \quad &+ \quad fln(Revised\ Employer\ Belief^j) + gln(Self\text{-}confidence1^j) \\ &+ \quad hX^j + k^i + \psi^{ij}. \qquad (2.7)\end{aligned}$$

The even numbers of columns report the results. By comparing the coefficients c and h, the efficiency of employer estimation did not improve. However, the unbiasedness did improve, though not significantly. This can be seen by checking that f is closer to to its null hypothesis value of 1 than b, and e is closer to 0 than a.

2.4 Conclusions

As identified by Benabou and Tirole [10], there are three values of demanding self-confidence: the consumption value, the motivation value, and the signalling value. The first two regard how one directly benefits from being self-confident. Our results provide explicit evidence to the signalling value, which concerns about how to gain a utility from convincing others that one has high ability. Indeed, as discussed by Burks et al. [19], overconfidence is more likely to be induced by the desire to send positive signals to others about one's own skill.

Since it is shown that some social groups differ in self-confidence, our results can help explain the income gap between different social groups. For example, as men are shown to be more confident than women, our results suggest that men would earn more in the labor market, just by signalling more self-confidence. We can also interact self-confidence with ethnic minorities. We did not find any interaction in the current paper, most likely because the location is a diverse province, with 40% of the population being minorities. Thus minorities and Hans are equal in many social dimensions. It is likely, however, to be different in non-diverse provinces, where minorities make up only a minor part of the population. In that case, they might be different from Hans in some social

Table 2.1 Summary statistics

Variable	Mean	Std.Dev
Age	20.2	1.3
College major		
—Agronomy	0.24	0.4
—Agricultural products	0.16	0.4
—Chinese medicinal herbs	0.11	0.3
—Environment and resources	0.17	0.4
—Horticulture	0.15	0.4
—Plant protection	0.16	0.4
Minority	0.4	0.5
Female	0.4	0.5
Urban	0.3	0.5
Points earned as the worker from the 5-minute work period	611	190
Points earned as the worker from employer evaluations	734	134
Points earned as the employer	956	472

dimensions, due to their very low representation in the population. It is then useful to conduct the experiment in a non-diverse Chinese province, and examine the interaction of self-confidence, ethnicity, and income.

We can also consider field experiments. Does self-confidence affect the wage in the field? Some works have answered this question from a more general angle, the non-cognitive skills. With our design, we can measure self-confidence in the lab, and look at the wage in the field. The ideal subjects would be university students on the job market. We will not be able to identify the specific value of self-confidence leading to the wage differential though. It can be that a person feels happier by being confident and thus performs better in the labor market. It is also possible that more confident individuals are more motivated and are more capable when facing difficulties. Or just like in the paper, more confident people signal more confidence, and hence convince employers about the high ability.

Table 2.2 Determinants of Self-confidence

Dependent variable: ln(Self-confidence1)			
	(1)	(2)	(3)
Minority	-0.01	-0.01	-0.01
	(0.05)	(0.05)	(0.05)
Female	-0.08	-0.08	-0.09
	(0.05)	(0.05)	(0.05)
Urban	0.05	0.08	0.07
	(0.05)	(0.05)	(0.05)
ln(300/Signal)		-0.26**	-0.31**
		(0.06)	(0.07)
ln(Productivity)			0.13
			(0.08)
N	270	270	269
R^2	0.06	0.13	0.14

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.3 Determinants of Productivity

Dependent variable: ln(Productivity)			
	(1)	(2)	(3)
Minority	-0.02	-0.02	-0.02
	(0.04)	(0.04)	(0.04)
Female	0.08*	0.09*	0.09*
	(0.04)	(0.04)	(0.04)
Urban	0.11*	0.07	0.07
	(0.04)	(0.04)	(0.04)
ln(300/Signal)		0.32**	0.34**
		(0.05)	(0.05)
ln(Self-confidence1)			0.08
			(0.05)
N	275	272	269
R^2	0.13	0.26	0.27

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.4 Determinants of Employer Belief

Dependent variable: ln(Employer Belief)					
	TG	TGE	TGEU	TGPU	All
Minority	-0.03	0.00	0.01	0.00	0.00
	(0.02)	(0.02)	(0.01)	(0.03)	(0.01)
Female	0.00	-0.03	-0.01	0.02	-0.00
	(0.02)	(0.03)	(0.01)	(0.03)	(0.01)
Urban	0.00	-0.04	0.03	0.08**	0.02
	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)
ln(300/Signal)	0.39**	0.28**	0.34**	0.38**	0.34**
	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)
TGE					-0.02
					(0.07)
TGEU					-0.00
					(0.05)
TGPU					0.02
					(0.07)
N	48	47	137	45	277
R^2	0.33	0.22	0.26	0.30	0.25

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.5 Does Worker Self-confidence Affect Employer Belief?

Dependent variable: ln(Employer Belief)	TG	TGE	TGEU	TGPU	All
Minority	-0.03	0.00	0.01	0.00	0.00
	(0.02)	(0.02)	(0.01)	(0.03)	(0.01)
Female	0.01	-0.03	-0.01	0.02	-0.00
	(0.02)	(0.03)	(0.01)	(0.03)	(0.01)
Urban	-0.00	-0.04	0.03	0.07**	0.02
	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)
ln(300/Signal)	0.40**	0.27**	0.33**	0.39**	0.34**
	(0.03)	(0.03)	(0.02)	(0.03)	(0.01)
ln(Self-confidence1)	0.06	-0.02	-0.02	0.04	0.00
	(0.03)	(0.03)	(0.02)	(0.04)	(0.01)
TGE					-0.02
					(0.07)
TGEU					-0.00
					(0.06)
TGPU					0.02
					(0.07)
N	48	47	137	45	277
R^2	0.33	0.22	0.26	0.30	0.25

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.6 Does Worker Self-confidence Affect Revision of Employer Belief?

Dependent variable: ln(Revised Employer Belief)	TG	TGE	TGEU	TGPU	All
Minority	0.02	0.02	0.02	-0.00	0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Female	-0.00	-0.04*	-0.01	0.01	-0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Urban	-0.02	0.00	0.00	0.03	0.00
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
ln(300/Signal)	0.51**	0.50**	0.46**	0.51**	0.48**
	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)
ln(Self-confidence1)	0.17**	0.14**	0.17**	0.31**	0.18**
	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)
TGE					-0.00
					(0.04)
TGEU					-0.03
					(0.04)
TGPU					0.01
					(0.04)
N	48	47	137	45	277
R^2	0.46	0.58	0.39	0.51	0.44

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.7 Does Worker Self-confidence Affect Revision of Employer Belief?

Dependent variable: ln(Revised Employer Belief)	TG	TGE	TGEU	TGPU	All
ln(Belief)	0.80**	0.57**	0.79**	0.70**	0.68**
	(0.03)	(0.04)	(0.02)	(0.03)	(0.01)
ln(Self-confidence1)	0.07**	0.04	0.13**	0.20**	0.11**
	(0.02)	(0.03)	(0.01)	(0.03)	(0.01)
TGE					0.02
					(0.03)
TGEU					-0.03
					(0.03)
TGPU					-0.01
					(0.03)
N	48	47	137	45	277
R^2	0.61	0.31	0.58	0.53	0.52

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.8 Does Worker Self-confidence Affect Revision of Employer Belief?

Dependent variable: ln(Revised Employer Belief)	TG	TGE	TGEU	TGPU	All
Minority	0.03*	0.02	0.01	-0.00	0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Female	-0.01	-0.03	-0.01	0.00	-0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Urban	-0.02	0.02	-0.01	-0.00	-0.00
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
ln(Employer Belief)	0.61**	0.30**	0.63**	0.48**	0.52**
	(0.03)	(0.03)	(0.02)	(0.04)	(0.01)
ln(300/Signal)	0.26**	0.42**	0.25**	0.32*	0.31**
	(0.03)	(0.02)	(0.01)	(0.03)	(0.01)
ln(Self-confidence1)	0.13**	0.14**	0.18**	0.29**	0.18**
	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
TGE					0.01
					(0.03)
TGEU					-0.03
					(0.02)
TGPU					-0.00
					(0.03)
N	48	47	137	45	277
R^2	0.70	0.65	0.67	0.67	0.65

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

Table 2.9 Comparison of Employer Profits From Two Offers

	Mean	Std.Error	Std.Dev	P25	P50	P100
Profit from first offer	913	26	427	840	1020	1140
Profit from second offer	1002	14	229	930	1060	1150

Table 2.10 Does Worker Self-confidence Raise Employer Profit?

Dependent variable: Employer Profit					
	TG	TGE	TGEU	TGPU	All
Second Offer	4.14	17.09**	9.70**	10.11**	10.08**
	(2.92)	(3.62)	(1.79)	(2.85)	(1.28)
Self-confidence2	-0.92*	-0.10	-0.61**	-0.53	-0.58**
	(0.40)	(0.50)	(0.20)	(0.34)	(0.15)
TGE					-4.45
					(6.48)
TGEU					-0.79
					(5.3)
TGPU					2.29
					(6.56)
N	48	47	137	45	277
R^2	0.01	0.02	0.01	0.02	0.01

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

Table 2.11 Rational Expectation

Dependent variable: ln(Productivity)	TG		TGPU		TGE		TGEU		All	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln(Employer Belief)	0.25**		0.21**		0.18**		0.21**		0.10**	
	(0.06)		(0.06)		(0.03)		(0.05)		(0.01)	
ln(Revised Employer Belief)		0.32**		0.40**		0.28**		0.32**		0.22**
		(0.06)		(0.06)		(0.03)		(0.05)		(0.02)
ln(Self-confidence1)		-0.02		-0.04		0.00		-0.01		-0.00
		(0.05)		(0.04)		(0.02)		(0.04)		(0.02)
Minority	0.03	0.02	-0.02	-0.03	-0.05**	-0.06**	0.02	0.02	-0.07**	-0.05**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)
Female	0.13**	0.13**	0.12**	0.13**	0.08**	0.08**	0.06*	0.06*	0.07**	0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)	(0.01)
Urban	0.08*	0.07*	0.18**	0.15**	0.06**	0.05**	0.10**	0.11**	0.11**	0.12**
	(0.04)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)
TGE									0.04	0.01
									(0.02)	(0.02)
TGEU									-0.00	0.01
									(0.02)	(0.02)
TGPU									0.31	0.29
									(0.17)	(0.16)
Constant	2.10**	1.90**	2.15**	1.64**	2.08**	1.82**	2.26**	1.97**	2.38**	2.01**
	(0.22)	(0.22)	(0.22)	(0.21)	(0.13)	(0.12)	(0.21)	(0.21)	(0.04)	(0.05)
N	48	48	47	47	137	137	145	45	277	277
R^2	0.18	0.21	0.28	0.34	0.16	0.19	0.23	0.28	0.07	0.14

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

CHAPTER 3. EMPLOYER DEMAND FOR WORKER CHARACTERISTICS IN EXPERIMENTAL LABOR MARKETS

We use the Becker-DeGroot-Marschak (BDM) mechanism to study the ex ante employer demand for worker characteristics. In our experimental labor market, "employers" estimate productivity of "workers" based on worker characteristics. In some treatments, we provide employers with worker resumes containing a subset of information on expected productivity, ethnicity, gender, and urban/rural characteristics. In another treatment, a worker characteristic is displayed on the resume if the employer claims a willing to pay (WTP) higher than or equal to a randomly determined price. The ex ante demand for a characteristic is determined by the WTP for it. We find the ex ante demand for the expected productivity is correctly calibrated to be the highest, while other characteristics are overdemanded. We then compare the ex ante demand with the ex post demand, which is determined by a characteristic's explaining strength in employer estimation. We find that a combination of both mechanisms gives the best labor market outcome. That is, the ex post demand of employers who buy worker characteristics is the most accurate demand.

3.1 Introduction

The labor market is a platform where job applicants provide employers with a set of relevant worker characteristics based on which the hiring and wage decisions are made.

However, job applicants are usually offered with restricted time length or physical space to present these characteristics. For instance, this question of "can you tell us about yourself in 5 minutes?" is known to be common during job interviews. The dilemma of the job applicant who confronts such limitations is, from the set of available characteristics, what are the most useful ones to present to the employer. This question becomes even more important when presenting irrelevant characteristics can possibly hurt the outcome. In Chapter 1 for example, employers at Non-diverse discriminate against ethnic minorities even though ethnicity is not related to worker productivity. Hence a very important but usually puzzling question in the labor market is: what information is really important for employers?

When the hiring or the wage decision is known, we can answer this question by studying the outcome with an equation that has worker characteristics as the explanatory variables. In experimental labor markets, this can be done using the framework in Chapter 1. In that labor market, the "employer" estimates the productivity of the "worker" based on worker characteristics. The ex post employer demand for a worker characteristic is determined by how it affects employer estimation. In other words, if employer estimation is the dependent variable and worker characteristics are the explanatory variables, then the ex post demand for a characteristic is determined by its coefficient in the regression equation.

Alternatively, we propose an experimental labor market to reveal employer demand for worker characteristics *ex ante*. It is the same labor market where employers estimate worker productivity. It is different, however, that instead of being assigned to different worker resumes, employers *buy* worker characteristics in the Becker-DeGroot-Marschak (BDM) market. To do so, they claim a willingness to pay (WTP) for a specific worker characteristic. If the WTP is higher than or equal to a randomly determined price, then this characteristic is bought by the employer and is displayed on the worker's resume. Otherwise it is not displayed to the employer. The ex ante demand for a characteristic is determined by the magnitude of the WTP.

We compare the two experimental labor markets by analyzing the true values, the ex ante demand, and the ex post demand for the worker characteristics. We find that the ex ante demand for the expected productivity is the highest among all characteristics. This is justified by its highest explaining power to the actual productivity. However, employers systematically overbid for other three characteristics. We attribute this overbidding behavior to curiosity satisfaction. In other words, employers spend money to know more about the worker just to satisfy their curiosity. We further find that a combination of both markets gives the best labor market outcome. That is, relative to employers who are given worker characteristics for free, employers who are asked to buy worker characteristics display the most accurate ex post demand. Two reasons can account for this result. First, employers who buy worker characteristics have two chances of thinking about how to use the characteristics in productivity estimation. Second, a costly characteristic is supposed to be used more carefully than a free one.

Our paper is part of a growing experimental labor market literature. The framework is similar to Chapter 1 that studies how stereotype is related with diversity by comparing the belief of employers in a diverse and a non-diverse community about minorities. It is also related to Mobius and Rosenblat [68] who study the origins of the "beauty premium". Bertrand and Mullainathan [12] use a field experiment to study racial discrimination in the U.S. labor market. They construct synthetic resumes and respond to help-wanted advertisements in Boston and Chicago newspapers. Resumes are randomly assigned typical white or African-American names. They find that resumes with white names receive 50 percent more callbacks for interviews. The results provide explicit evidence for racial discrimination in the U.S. labor market. In the Chinese labor market, Maurer-Fazio [65] uses a similar approach and finds that Han Chinese are much more likely to receive a callback from jobs posted on internet. The BDM mechanism used in this paper has been extensively applied in studying reservation prices for products like petrol

(Bohm [14]), goods with induced values (Irwin et al. [53], Keller et al. [56]), and food safety (Rozan [77]). Another commonly used demand revealing mechanism is the second price auction. For example, it has been used to study food safety (Buzby et al. [20] and Hayes et al. [47]) and Europeans' willingness to pay for U.S. beef (Alfnes [3]). The random n-th price auction is a combination of the BDM mechanism and the second price auction in which the winner pays the n-th bidder's price. Huffman et al. [51]) have used this mechanism to study consumers' willingness to pay for genetically modified food.

The rest of the paper is organized as follows. Section 2 introduces the experimental design, the BDM mechanism, and data collection process. Section 3 describes the analysis strategy. Experimental results are reported in section 4 and section 5 concludes.

3.2 The Experiment

3.2.1 Experimental Design

The experimental design is based on the framework of Chapter 1 in which employers set wages of workers who do a real effort task. The task is the same character puzzles from Chapter 1 (shown in Figure 3.1). Each puzzle shows two quadratic arrays of 7 times 7 characters of Latin alphabets. The two arrays are identical except for two random position where the characters differ. Workers have to find these two locations and click them with their mouse.

In the first step as the worker, each subject is given two warm-up character puzzles. Afterwards, the worker is asked to solve one practice character puzzle. The time that takes the worker to complete the practice character puzzle, which we will refer from now on as the "signal", is recorded by the experimenter. The signal and other personal information on gender, urban/rural status, ethnicity, and province of origin, are used to construct the worker's "resume". In the last step, the worker has a 5-minute work period to solve as many puzzles as possible and is rewarded with 40 points for each solved

Figure 3.1 Character Puzzle

L	Y	S	E	H	D
N	L	X	O	C	A
W	Y	F	T	Y	X
O	I	A	W	L	L
J	U	S	Q	R	M
K	C	A	T	O	A
G	P	K	L	S	R

L	Y	S	E	A	D
N	L	X	O	C	A
W	Y	F	T	Y	X
O	I	A	W	L	L
J	U	S	Q	R	A
K	C	A	T	O	A
G	P	K	L	S	R

puzzle.[1] The numbers of puzzles that the worker completes in the 5 minutes are referred from now on as the worker's "productivity". The sequence of the puzzles in each step are identical for every subject, which means that the subjects are solving the same puzzles appearing in the same order.

The subjects are then switched to the employer who estimates worker productivity. The estimated productivity is referred from now on as the "employer belief" on a worker. Each employer evaluates 10 other randomly selected workers and earns 150 points for each evaluated resume. However, if the employer belief is different from a worker's productivity by x puzzles, the earnings are reduced by $10 \cdot x$ points. For example, if a worker solved 20 puzzles in the 5 minutes and the employer's estimate is 18, the employer receives $150 - 10 \cdot |20 - 18| = 130$ points.

The worker receives a wage of 40 points per average employer belief. For example, if a worker is estimated by 8 employers and the average employer belief is 20, the worker receives a wage of $40 \cdot 20$ points. Therefore, the worker has two sources of income: the productivity and the employer belief. This provides the worker with an incentive to

[1] The experimental points are later converted to cash at a rate of 100 points = 1 Yuan $\simeq$ \$0.16.

achieve comparable performance in the timed character puzzle and the 5-minute work period.

Employers are randomly assigned to one out of five resume treatments. In four of them, the treatment determines how a worker's resume is displayed to the employer. This is identical to the experimental design in Chapter 1 except that the TGPU resume is replace by the TGU one. Examples of each type of resume are given in Figure 3.2.

TG: The employer sees the signal ("Practice time" on the example resume) and gender.

TGE: The employer sees ethnicity in addition to TG.

TGEU: The employer sees urban/rural in addition of TGE.

TGU: The employer sees urban/rural in addition to TG.

Figure 3.2 Examples of Resumes in Control Experiment

TG

Practice time:	18
Gender:	male
Evaluation:	_

TGE

Practice time:	20
Gender:	female
Ethnicity:	Han
Evaluation:	_

TGEU

Practice time:	22
Gender:	male
Ethnicity:	minority
Urban/Rural:	Urban
Evaluation:	_

TGU

Practice time:	17
Gender:	female
Urban/Rural:	Rural
Evaluation:	_

Some employers are assigned to the WTP treatment where they can buy worker characteristics. In this treatment, an employer is given 150 points for every estimated worker, and is asked to claim a WTP for each of the four worker characteristics: signal, ethnicity, gender, urban/rural, in the following question.

WTP for practice time

For seeing each worker's practice time, the highest price you are willing to pay is _____ credits.

WTP for ethnicity

There are x Han workers, y Tibetan workers, z Mongolian workers, ... For seeing each worker's ethnicity, the highest price you are willing to pay is _____ credits.

WTP for gender

There are x male workers and y female workers. For seeing each worker's gender, the highest price you are willing to pay is _____ credits.

WTP for urban/rural

There are x urban workers and y rural workers. For seeing each worker's urban/rural status, the highest price you are willing to pay is _____ credits.

When the decisions are made, three characteristics are randomly chosen by the computer to be displayed to the employer for free. The remaining one characteristic is hidden and a price p between 0 and 150 points is randomly determined for it. If the corresponding WTP is higher than or equal to p, then the employer pays p and the hidden information is displayed. Otherwise the employer keeps 150 points and sees the free characteristics. Figure 3.3 is an example in which the ethnicity characteristic is chosen to be hidden. The left side resume has no information on the worker's ethnicity because WTP is lower than p but the right side one does show that the worker is Han because WTP is higher than or equal to p.

3.2.2 The BDM Mechanism in Revealing True WTP

The BDM mechanism guarantees that revealing true WTPs is the dominant strategy for employers. To see why, consider an employer with the true WTP w for a certain

Figure 3.3 When ethnicity is chosen to be hidden

($If\ WTP^{ethnicity} < p$)		($If\ WTP^{ethnicity} \geq p$)	
Practice time:	18	Practice time:	20
Gender:	male	Gender:	female
Urban/Rural:	urban	Urban/Rural:	rural
Evaluation:	_	Ethnicity:	Han
		Evaluation:	_

characteristic. Suppose that the employer claims a WTP w' higher than w. If the randomly determined price p falls between w and w', then the employers pays p. But since the maximum price the employer is willing to pay is w, he/she loses $(p-w)$. Hence the employer will not claim a WTP higher than w. Now suppose that the employer claims a WTP w'' lower than w. If p falls between w'' and w, the characteristic will not be displayed. But the employer is willing to see the characteristic at price p because he/she can gain $w - p$. Hence the employer will increase the claim until $w'' = w$. Therefore, the employer will always claim a WTP equals the true WTP w.

3.2.3 Data

The data collection process is similar to that in Chapter 1. We conduct our experiment with students at Changan University, China. The university is located in the province Shaanxi where fewer than 1% of the population are members of ethnic minorities. We contacted students through their class supervisors and obtained their consent to participate in the experiment. As Table 3.1 shows, the students are from Math, Engineering, and Computer Sciences majors. A total of 284 students are recruited, and 145 students are assigned to treatment WTP Experiment and 139 students are assigned to other treatments. The table reports the mean and standard deviations statistics of demographic characteristics like ethnicity, gender, urban/rural, local/non-local, minority/non-minority province. It also shows the average earnings of subjects from (1) solving puzzles in the 5-minute work period, (2) worker wages averaged on employer estimates, (3) the employer's earnings from evaluating workers, and (4) the employer's earnings from buying worker characteristics. Combined earnings are 33 Yuan for WTP Experiment participants and 23 Yuan for Control Experiment participants.[2]

3.3 Results

3.3.1 True Values of Worker Characteristics

The true value of a characteristic is determined by its strength in explaining worker productivity, which can be determined in the following equation:

$$\ln(Productivity^{ij}) = \alpha + \beta X^j + \epsilon^{ij} \tag{3.1}$$

where $Productivity^j$ worker j's productivity,X^j is a vector of worker j's characteristics: $300/Signal$, $Minority$, $Female$, $Urban$. The variable $300/Signal$ is the signal converted to equivalent numbers of puzzles in 5 minutes. The other three are dummy variables with the value of 1 if the worker is a minority, female, urban person, respectively, and 0 otherwise.

Table 3.3 reports the results in the first column. The signal characteristic is a very important explanatory factor to worker productivity. A one percent increase in $(300/Signal)$ significantly raises worker productivity by 0.16 percent. The urban/rural characteristic is also a good factor in explaining worker productivity. A change from rural to urban increases the productivity by 8%. Note that from the maximum feasible productivity of 50 puzzles, we can infer that the maximum value for urban/rural characteristic is 10 credits. To see this, suppose an employers see a worker resume with signal but without urban/rural. Based on the signal, the employer gives an estimate of x puzzles. Knowing the urban/rural characteristic would change the estimation by 8%, which

[2]The opportunity cost of one hour for a university student, e.g. tutoring a school kid, is about 20 Yuan in the two provinces.

is $x \cdot 8\%$ puzzles. Since the maximum the x can be is 50, the estimate could change by 4 puzzles at most. But because wrong estimation is punished by 10 credits for each unit of deviation, the employer will not spend more than 40 credits on knowing the urban/rural characteristic. Similarly, because the mean worker productivity is 16 puzzles, we can also infer that the mean value of the urban/rural characteristic is 13 credits. The other two characteristic, ethnicity and gender, have no impact on productivity and hence have the value of 0. Let V denote the true value of a characteristic, we can summarize the findings as follows:

1. Signal has the highest value among all characteristics;
2. $Max(V^{urban/rural}) = 40$ credits;
3. $Mean(V^{urban/rural}) = 13$ credits;
3. $V^{ethnicity} = 0$ credit;
4. $V^{gender} = 0$ credit.

3.3.2 Ex Ante Demand for Worker Characteristics

The ex ante demand for a characteristic is determined by the WTP for it. We investigate the WTPs by looking at the distributions as well as performing statistical tests for equal means and variances. Figure 3.4 draws the distribution of WTP for ethnicity. Taking the WTP of 50 as the midpoint, the graph is inclining to the left side toward 0. The highest percentage, 15%, of employers claimed a WTP of 0. As indicated on the graph, half of employer claimed a WTP below 20. To be in the upper quartile, the WTP is above 45 which is about one third of the total endowment. The distribution shapes of WTPs for gender and urban/rural are similar to WTP for ethnicity. Highest percentages of employers happen when WTP is 0. Half of employers are not willing to pay a price above 20. In the contrary, the graph of WTP for signal is more balanced on the two sides of 50, which is claimed by the highest percentage of employers. The median WTP of 20 for ethnicity, gender and urban/rural is the lower quartile in the distribution of WTP for signal. Moreover, the median of 40 is twice of other three WTPs. In particular, it is close to the 75th percentile of the other three.

Figure 3.4 WTP for Ethnicity

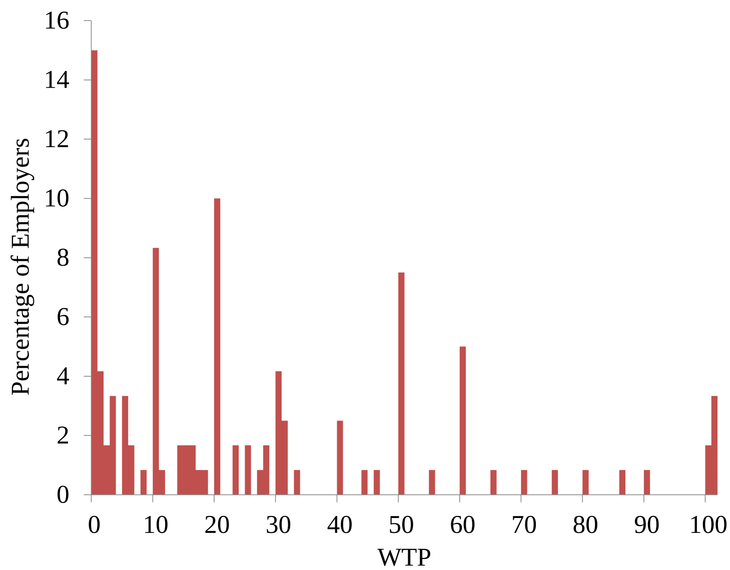
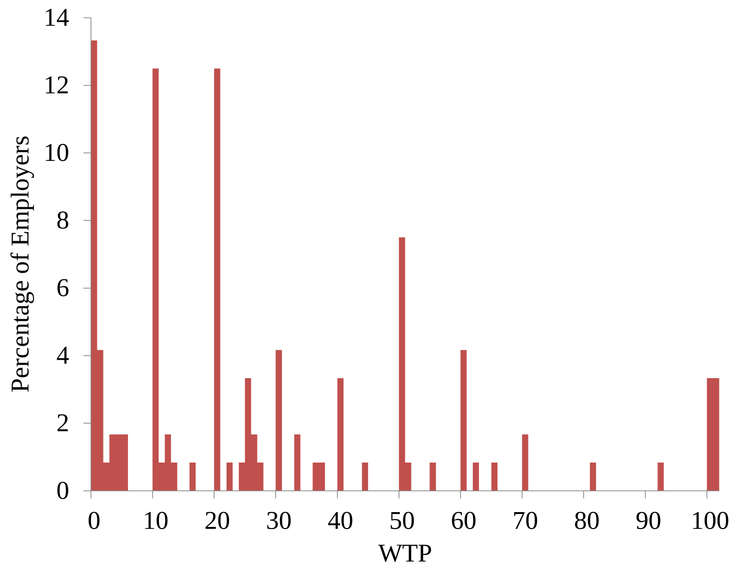


Table 3.2 reports the specific statistics of means and variances. The Bartlett's test suggests that the four WTPs have equal variances but the F test rejects the null hypothesis of equal means. By comparing the means and the 95% confidence intervals, WTPs for the ethnicity, gender, and urban/rural characteristics are not significantly different from each other. The null hypothesis of equal means is rejected because that WTP for signal is significantly higher than other three WTPs.

According to these figures and statistics, we have the following findings. First, the WTP for signal, 43 credit, is the highest among all characteristics. Moreover, it is about 50% higher than the other three characteristics. Second, there is systematically overbidding for the other three characteristics. Employer even spend a significant amount of money on knowing ethnicity and gender, which are valued at 0 credit. We explain

Figure 3.5 WTP for Gender



graphwtpgender3

this over-bidding behavior by curiosity satisfaction. In other words, employers want to know about the worker to satisfy their curiosity. The average spending on this curiosity is about 30 credits, which is one fifth of the total endowment.

3.3.3 Ex Post Demand for Worker Characteristics

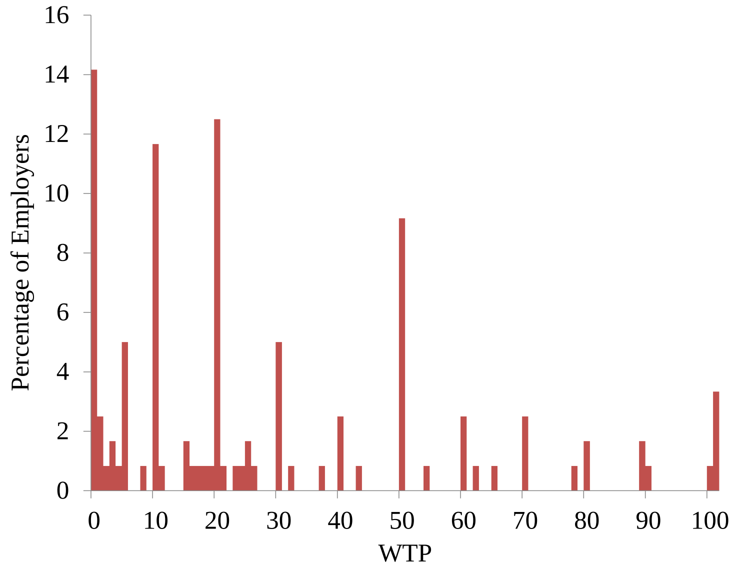
The ex post demand for worker characteristics can be examined in the following equation:

$$\ln(Employer\ Belief^{ij}) = \alpha + \beta X^j + \upsilon^{ij} \tag{3.2}$$

where $Employer\ Belief^{ij}$ is employer i's estimate on worker j's productivity, and X^j is the same vector of worker j's characteristics in Equation (1).

We report the results for each of the five treatments in the last four columns of Table 3.3. For every treatment, the signal is an important characteristic to estimate worker productivity. If a worker performs one percent better in the converted timed character puzzle, employers give a 0.15-0.43 percent higher estimate to the productivity. In par-

Figure 3.6 WTP for Urban/Rural



graphwtpurbanrural3

ticular, its impact on employer estimation is closer to its impact on actual productivity of 0.16, in treatments WTP, TG, and TGE. This suggests that the ex post demand for signal is more accurate in these three treatments.

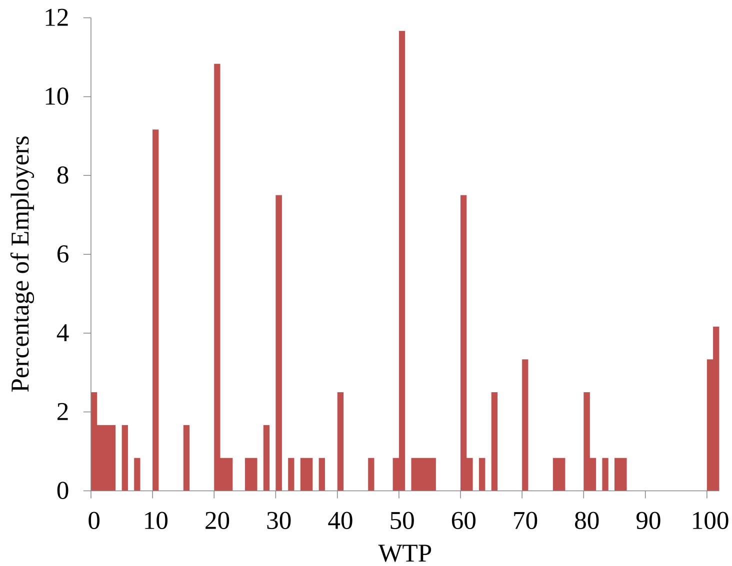
The urban/rural characteristic is significant only in treatment WTP. The magnitude of the coefficient, 0.05, is close to its impact on actual productivity of 0.08. This implies that urban/rural is demanded accurately only in treatment WTP. Using the same logic before, we can infer that the maximum and mean values of the urban/rural characteristic in the ex post demand is 25 and 10 credits, respectively. Let $\widetilde{V}$ denote the value of a characteristic in the ex post demand, we can summarize the findings as follows.

For treatment WTP:

1. Signal has the highest value among all characteristics;

2. $max(\widetilde{V}^{urban/rural}) = 25$ credits; (in treatment WTP)

Figure 3.7   WTP for Signal



graphwtpsignal3

3. $mean(\widetilde{V}^{urban/rural}) = 10$ credits; (in treatment WTP)

4. $\widetilde{V}^{ethnicity} = 0$ credit;

5. $\widetilde{V}^{gender} = 0$ credit.

For treatment TG:

1. Signal has the highest value among all characteristics;

2. $\widetilde{V}^{urban/rural} = 0$ credit;

3. $\widetilde{V}^{ethnicity} = 0$ credit;

4. $\widetilde{V}^{gender} = 0$ credit.

3.3.4 Comparing the Two Mechanisms

In terms of closeness to the true values, a combination of both mechanisms gives the best labor market outcome. In other words, the ex post demand of employers who buy worker characteristics is the most accurate one.

The results make sense for two reasons. First, relative to employers who are given characteristics for free, employers in treatment WTP are using the information they have spent money on. Hence they are likely to be more careful in using worker characteristics for the estimation. Second, the WTP treatment provides employer with two chances of thinking about how to use the characteristics.

3.3.5 Employer Profit

In this section, we examine the relationship between the WTPs and the employers' profits. The profit that an employer makes from evaluating a worker is defined before as $150 - 10 \cdot |mistake|$ points, where $mistake$ is the difference between the worker's actual productivity and the employer belief. Consider the following equation:

$$Employer\ Profit^{ij} = \zeta + \eta X^j + \theta WTP^i + \nu^{ij} \tag{3.3}$$

where $Employer\ Profit^{ij}$ is the profit that employer i makes from evaluating worker j, X^j is a vector of worker j's characteristic dummies: minority, female, rural, province of origin.

Tables 3.4 reports the results. As the coefficients of the WTPs are not significant, it is suggested that the WTPs are not associated with the employers' profits from evaluating workers.

3.3.6 The Quality of Employer Estimation

We adopt the methodology by Granger [44] and Mankiw and Shapiro [64] to evaluate the quality of employer estimation. Specifically, we examine whether the employers' estimation accuracy is associated with their WTP. Consider the following model:

$$\ln(Productivity^j) = a + b\ln(Employer\ Belief^j) + cX^j + dWTP^i + \phi^{ij}, \tag{3.4}$$

and the following hypotheses:

1. a=0;

2. b=1;

3. c=0 and d=0.

The first two hypotheses state that the estimates are unbiased, because high quality of estimation means that the plots of *Employer Belief* and *Productivity* would look similar. The third hypothesis tests the efficiency of the estimates. Since good employer estimation should have contained all useful information that can predict the actual productivity, the coefficients of these information should be 0 once *Employer Belief* is controlled for. Employers are rational if both unbiasedness and efficiency are satisfied.

Results are displayed in Table 3.5. The coefficients suggest that all the hypotheses above can be rejected at significance level of 5%. Therefore, the quality of the employer estimates can be easily improved.

3.4 Conclusion

We propose a methodology that measures the ex ante employer demand for worker characteristics, and compare it with the ex post demand. We find that the ex post demand of employers who buy worker characteristics is the most accurate one.

Our methodology can serve as an alternative to measure labor market discrimination. The wage differentials in the ex post demand would be then proportional to the WTP differentials in our mechanism.

As evidence suggests that irrelevant information can hurt the labor market outcome, we can also apply the mechanism to avoid the labor market being affected by noisy factors. In that case, if it is costly to reveal worker characteristics, employers would not be willing to pay for something not really related to the worker's productivity.

Table 3.1 Summary statistics

Treatment	WTP (N=145)			TG, TGU, TGE, TGEU (N=139)	
Variable	Mean	Std.Dev		Mean	Std.Dev
College major					
—Math	0.24	0.43	—Math	0.27	0.44
—Engineering	0.44	0.50	—Engineering	0.47	0.50
—Computer Sciences	0.32	0.47	—Computer Sciences	0.26	0.44
Minority	0.04	0.20		0.11	0.30
Female	0.21	0.41		0.23	0.42
Urban	0.41	0.49		0.45	0.50
Minority province	0.07	0.25		0.05	0.22
Local	0.17	0.37		0.23	0.42
Points earned as the worker from the 5-minute work period	610	311		667	274
Points earned as the worker from employer evaluations	802	143		803	147
Points earned as the employer from evaluating workers	582	915		830	510
Points earned as the employer from buying worker information	1309	364			

a. "Minority province" refers to any from the five ethnic autonomous provinces according to China's constitution: Guangxi, Innermongolia, Ningxia, Tibet, Xinjiang.
b. "Local" refers to students whose province of origin is the experimental province.

Table 3.2 WTP by Worker Characteristics

	Mean	Standard Error	95% Confidence Interval
Signal	43	3	(37, 49)
Ethnicity	28	3	(22, 34)
Gender	30	3	(24, 36)
Urban/Rural	29	3	(24, 35)

F test for equal means: P=0.0007

Bartlett's test for equal variances: P=0.99

Table 3.3 Regression Results

Dependent variable	Productivity	ln(Employer Belief)				
Treatment		WTP	TG	TGU	TGE	TGEU
	(x.1)	(x.1)	(x.2)	(x.3)	(x.4)	(x.5)
ln(300/Signal)	0.16**	0.15**	0.16**	0.22**	0.19**	0.43**
	(0.05)	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)
Minority	0.04	0.00	-0.04	0.01	-0.02	-0.08
	(0.08)	(0.02)	(0.03)	(0.04)	(0.02)	(0.04)
Female	0.07	-0.02	0.01	-0.07	0.03	-0.04
	(0.05)	(0.02)	(0.03)	(0.04)	(0.02)	(0.04)
Urban	0.08*	0.05**	-0.03	-0.01	0.03	-0.02
	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.04)
N	247	119	21	19	62	21
R^2	0.2	0.07	0.08	0.17	0.11	0.36

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

b. In Column (2), the explanatory variables are observed ones, which are equal to the original value of 1 if observed, and 0 otherwise.

Table 3.4 Profit Regression

Dependent variable: Employer Profit	
Minority	-6.92
	(3.74)
Female	-1.84
	(3.98)
Urban	-8.08
	(3.46)
WTP_signal	-0.77
	(0.35)
WTP_ethnicity	-1.23
	(0.63)
WTP_gender	-0.62
	(0.63)
WTP_urban/rural	-0.76
	(0.67)
N	116
R^2	0.01

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

Table 3.5 Rational Expectation Regression

Dependent variable: ln(Productivity)	
ln(Employer Belief)	0.02
	(0.02)
Minority	-0.01
	(0.02)
Female	0.07**
	(0.02)
Urban	0.12**
	(0.02)
WTP_signal	0.00
	(0.00)
WTP_ethnicity	-0.00
	(0.00)
WTP_gender	0.00
	(0.00)
WTP_urban/rural	-0.00
	(0.00)
Constant	2.72**
	(0.05)
N	116
R^2	0.05

a. Significance levels of 5% and 1% are denoted by * and **, respectively.

CHAPTER 4. INDIRECT CONTACT AND SOCIAL COOPERATION

Intergroup contact is the key to improving intergroup relationship and reducing intergroup bias. Recent literature has demonstrated its impact on various settings of economics like cooperation, trust, and altruism. However, intergroup contact is often limited by scarce resources and availability of the involved groups. Indirect contact provides solutions to these limitations. We study how intergroup cooperation is increased by simply *observing* in-group members being in contact with out-group members. Our subject pool consists of students recruited from two different college majors at a Chinese university. In our control treatment, a student is matched with someone from the other major in a two-player public goods game. In a second treatment, the game players watch intergroup contact prior to the public goods game. We find that subsequent contribution to the public goods is significantly higher compared to the control treatment. To distinguish intergroup contact effect from simply putting subjects in a cooperative mood, the game players in a third treatment watch random contact. The results show that it is important to have in-group members in the contact.

4.1 Introduction

Intergroup contact is shown to be the key to improving intergroup relationships and reducing prejudice against out-group members (Allport [4] and Pettigrew and Tropp [75]). Its impact on economic settings like trust (Fiedler and Haruvy [37]), cooperation

(Eckel and Grossman [33]), reciprocity and altruism (Buchan et al. [18]) has been studied extensively. For example, Eckel and Grossman [33] show that random individuals cooperate more after interacting with each other. However, most of the works have focused on direct interactions of which the feasibility may be restricted in many cases. In the case of inter-racial/ethnic interactions, they are often limited by the shortage of the minority population. Likewise, contact between domestics and immigrants is usually blocked by residential, educational, and occupational segregation. Abundant financial and other resources are also required for direct intergroup contact. International cooperation is one of the primary goals of international agencies like the United Nations. Yet, because of the physical distance, having every international citizen involved in direct interactions is very costly, if not impossible.

Indirect intergroup contact, defined as learning, observing, and imagining in-group members in contact with out-group members, provides us with a solution to these limitations. These three types of indirect contact are shown to be effective in improving intergroup relationship and reducing intergroup prejudice. In the pioneer work of Wright et al. [94], two groups of individuals watch an in-group member interacting with an out-group member in a puzzle task. When the interaction process is viewed as friendly, the observers expressed lower in-group-out-group bias in a subsequent survey. Liebkind and McAlister [60] conduct distribution, reading, and discussion of stories of inter-ethnic friendship in Finnish middle schools. They find that compared to the control group, students exhibited improved attitudes toward all ethnic minorities. In an ethnically diverse island of Cyprus, Husnu and Crisp [52] asked Turkish Crypiot and Greek Crypiot subjects to imagine contact with someone from the other group. This imagination process increased subjects' intentions to engage in future contact with the out-group members.

According to Pettigrew and Tropp [75] and Pettigrew et al. [76], there are two powerful features of the indirect contact. First, the effect from the contact of an in-group member with an out-group member will spill over to that out-group as a whole.

In economics, this feature implies resource savings. In the example of international cooperation, diplomatic activities with representatives from a foreign nation that are watched, learned, and discussed by domestic citizens can build a positive image and friendship with all citizens from that foreign nation. Second, although the theory was originally developed for ethnic and racial groups, it can be applied to an extensive range of social groups, and it is universal. This feature is particularly useful in managing different diversities. Indirect contact provides a uniform solution to reduce intergroup bias for these different diversities.

In the present work, we demonstrate that intergroup cooperation can be increased by observing in-group members interacting with out-group members as an example of how indirect contact can improve intergroup relationship, and in turn influence economic behavior. Subjects in our experiment play a two player public goods game with out-group members. In the control treatment, two students from different college majors are matched in the game. In the second treatment, subjects watch intergroup contact prior to the public goods game. To this purpose, a puzzle-solving team is formed by a classmate of one player and a classmate of the other player. The players are asked to watch the puzzle-solving team working together to solve a jigsaw puzzle. To distinguish the effect of indirect intergroup contact from simply being emotionally cooperative due to the jigsaw puzzle, players in the third treatment watch random contact, where the puzzle-solving team consists of two random students.

We find that subjects who watched intergroup contact contribute about eight of the ten endowment credits in the public fund, while subjects in the control treatment contribute only six and half credits. In terms of percentages increased, the former contribute 25%-30% more than the latter. On the other hand, the impact of watching random contact on increasing public contribution is limited and significantly lower than watching intergroup contact. The findings suggest that having in-group members in the contact is important.

The rest of the paper is organized as follows. The next section reviews previous studies on direct and indirect intergroup contact in both psychology and economics literature. Section 3 introduces the experimental design. Results are reported in Section 4. Section 5 concludes.

4.2 Literature Review

In this chapter, we review works related to intergroup relationship, from direct contact to indirect contact, and from psychology to economics.

4.2.1 Direct Contact and Intergroup Relations (Psychology)

The literature on intergroup relations is so enormous that we may need a handbook to cover it. Likewise, there are more than 600 studies on intergroup contact (Pettigrew [76]) which makes it impossible to cover every one in this study. We present here the most classical and representative examples.

We begin the introduction to the development of intergroup relationship with Adorno et al's work, *The Authoritarian Personality* [1]. The authors claim that prejudice and discrimination against the out-group evolves from motivational sequences of interpersonal interactions. In the example of Hogg and Abrams [50], children who have overly harsh and restrictive parents are usually required to follow and execute strict convention, duty, rules and authority. When the emotions are repressed, an out-group with lower status is selected as an ideal target to express these repressed parts of personality. These expressions are often aggressive as the depressed emotions are mostly with aggression.

Another approach is coined by Campbell [24] as the Realistic Group Conflict Theory, where intergroup conflict is driven by real conflict of intergroup interests. In the well-known Robers Cave experiment of Sherif [84], 22 boys who never met each other before and participated in a summer camp in the Robbers Cave State Park, Oklahoma were separated into two groups. The two groups were assigned identity building activities within the group. Afterwards, the boys were brought together in competitive activates with prizes given to the winners. Intergroup tension, hostility, and aggressive actions were found in the process of the competition.

An approach that is widely studied and applied today in psychological and economic research is the “minimal group paradigm ”(Tajfel [86]). The theory states that any arbitrary distinction is sufficient to trigger intergroup bias. In Tajfel et al.'s experiment [87], subjects showed in-group-out-group bias following categorization according to underestimation or overestimation of the number of dots on the screen, or preference to paintings by Klee or Kandinsky. This pattern of behavior persists, even when the random division process was explicitly told to the subjects (Billig and Tajfel [13]).

The first work that proposed the contact theory is Allport's *The Nature of Prejudice* [4]. Since then, the vast amount of studies have shown that putting groups in direct contact is the most efficient way to reduce intergroup bias. Allport characterized four conditions in which the contact would work effectively: equal status of the groups in the situation, intergroup cooperation, common goals, and authority support. Pettigrew et al. [75], in a meta-analysis, indicate that the “contact effects typically generalize to the entire outgroup, and they emerge across a broad range of outgroup targets and contact settings”. It is further found that the contact effect generalizes even to out-groups not involved, and it is universal (Pettigrew [76]).

4.2.2 Direct Contact and Intergroup Relationship (Economics)

Intergroup bias has been studied extensively in economics. In cooperative behaviors, for example, Solow and Kirkwood [85] compare people with pre-existing group affiliations to randomly selected ones, and find that the former cooperate more with each other in public goods experiments. Ruffle and Sosis [78] find that in Israeli society, Kibbutz residents are more likely to cooperate with other Kibbutz residents, than to cooperate with city residents. Social norm is also affected by intergroup bias. It is shown that when individuals act as a judge who punishes people violating the norm, they put a heavier punishment on those who belong to the out-group (Bernhard, Fehr and Fischbacher [11] and Goette, Huffman, and Meier [42]). In altruism, Ben-Ner et al. [9] showed that the intergroup bias exists in a wide range of social identities, from family to music preference, religion to sports-team loyalty, and kinship to political views. Yan Chen and Sherry Xin Li [27] generalized the different behaviors based on group identity into social preference that include charity, envy, norm enforcement and efficiency.

The different cooperative behavior might be a mixed consequence of group membership with culture, social norms and so on. Goette, Huffman, Meier [42] use random assignment from the Swiss army to attribute the higher cooperation toward in-group members to group membership. To a similar purpose, Charness et.al [25] argue that when an audience is present, a person's behavior toward in-group members and out-group members diverges further as the group affiliation to the audience increases.

Following these works that have documented intergroup bias, researchers started looking for ways to overcome it. Building a common identity is shown to be useful. Eckel and Grossman [33] show that cooperation between random individuals can be raised by building a common identity between them. Moreover, the cooperation level increases as the identity-building activities are intensified. Roy Chen and Yan Chen [26] find that building a common identity with communication between random people can help the equilibrium identification in multi-equilibria games. Buchan et al. [18] demonstrated that even irrelevant communication has a powerful influence on participants' behavior. In a novel design of Fiedler and Haruvy [37], participants interact through a virtual world where social distance is varied but social identity is kept private. This increases both the amount sent and the percentage returned in the trust games, relative to the control group.

4.2.3 Indirect Contact and Intergroup Relationships (Psychology)

Indirect contact through an in-group member with an out-group member takes three forms: learning, observing, and imagining. We give a separate introduction and literature review to each of these forms.

4.2.3.1 Extended contact

Extended contact is defined as learning that an in-group member interacts with an out-group member. Several works have documented its effect in intergroup relationships. Wright et al. [94] shows that if an in-group member has a friend in a particular out-group, the individual expresses lower prejudice toward that out-group. Pettigrew et al. [74] surveyed German adults on how many of their German friends have friends who are foreigners, and their attitudes toward foreigners. They found that the number of friends who have foreign friends are negatively related with prejudice against foreigners. Even for groups with intensive intergroup conflicts like the Catholics and Prostestants in Northen Ireland, individuals expressed reduced prejudice toward the other group if he/she had an in-group friend engaged in a friendship with a person from that out-group (Paolini et al. [72]). This mechanism works also for the in-group members who are unknown. Liebkind and McAlister [60] collect stories of friendship with ethnic minorities from middle school students in Finland. They conduct distribution, reading, and discussion of these stories in experimental schools. Compared to the schools that did not go through this intervention process, subjects' attitudes toward all minorities improved. It is found that not only adults, but also children, are influenced by indirect contact. In a work by Cameron et al. [23], the experimenters read stories of friendship between English and refugees to elementary school kids. Subsequent surveys show a drop in the negative attitude toward refugees. Likewise, Cameron and Rutland [22] found that reading stories of friendship between non-disabled and disabled children increased children's positive attitude and intended behavior toward disabled people.

Turner et al. [91] conducted a study on inter-group contact between white British and Asians, and looked for the reason behind the indirect contact effect. They found that, among the four mediators, reduced anxiety, in-group norms, out-group norms, and inclusion of the out-group itself, each had an independent role in the improved intergroup relationship.

4.2.3.2 Observed contact

Another form of the indirect contact is observing the friendship of in-group members with out-group members. In another study of Wright et al. [94], subjects whose group identity was induced with minimal group paradigm were asked to observe an in-group member and an out-group member interacting in a puzzle task. The observers expressed a lower in-group-out-group bias when the relationship between the puzzle-solvers was perceived as a close friendship. Perhaps the most powerful application of the observed contact is the media. For example, Schiappa et al. [79] found that viewing television programs that portrayed positive intergroup contact was associated with lower levels of prejudice. Compared to explicit expressions and behaviors, implicit racial statements and actions can also influence intergroup relationships. Weisbuch et al. [92] demonstrated that television shows which exhibited negative nonverbal behavior toward blacks increase the viewers' racial bias.

4.2.3.3 Imagined contact

Probably the most appealing form of the indirect contact is the imagined contact. According to Crisp and Turner [30], it is the "mental simulation of a social interaction with a member or members of an outgroup category. " This is probably the cheapest intervention among all kinds of social interactions. Nevertheless, it is shown to be effective in reality. In a recent work by Husnu and Crisp [52], for example, Turkish Cypriots and Greek Cypriots, on the ethnically diverse island of Cyprus, were repeatedly asked to imagine contact with someone from the other group. This imagined intervention alone increased subjects' intentions to engage in future contact with the out-group members. There might be some argument on what drives this mechanism, the person in the imagined contact, or the contact itself. To answer this question, Turner et al. [89] compared three experiments where young participants imagined an outdoor scene, an elderly person, and talking with an elderly person, respectively. They found that subjects in the last experiment showed lower levels of intergroup bias than the former. In a similar work (Turner and Crisp [88]), non-Muslims showed a more positive attitude toward Muslims after imagining conversation with a Muslim stranger.

4.2.4 Indirect Contact and Intergroup Relationships (Economics)

There are a few works related to the indirect contact effect in economics. Boisjoly et al. [15] show that if a white student is randomly assigned a black roommate, he/she is more likely to feel empathy toward all ethnic minorities. In an experiment studied by Senen and Schram [83], the donor's decision as to whether or not to provide costly help to a recipient depends on the history of the recipient's behavior with third parties.

4.3 The Experiment

4.3.1 Experimental Design

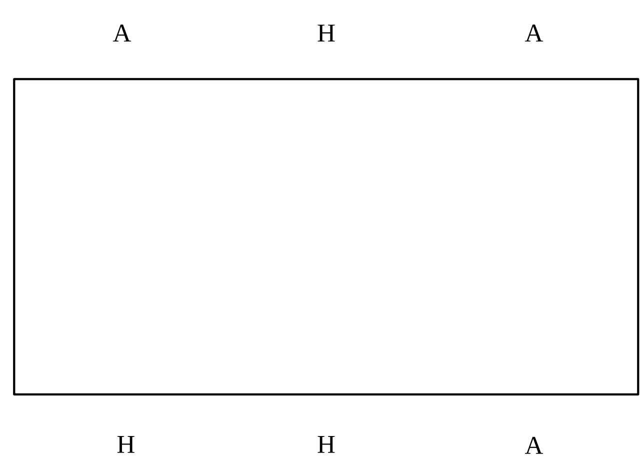
Two students, each from a different college major, are randomly matched to play a two-player public goods game of ten rounds. In the beginning of every round, each player is given an endowment of ten experimental credits.[1] The players need to invest the credits in a public fund and a private fund. After investment decisions are made, 50% of the public investment is added to the public fund by the experimenter. The players then share the public fund equally with their game partners. Therefore, a player's earnings

[1] 1 experimental credit = 1 Yuan $\simeq$ \$0.16

by the end of each round are what he/she has invested in the private fund plus half of the public fund. The investment earnings are not cumulative to subsequent rounds.

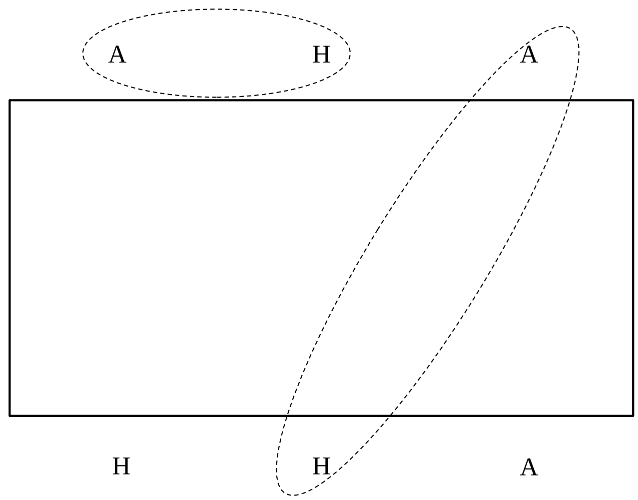
In each session, a total of $6 - 12$ subjects are randomly seated around a big square desk. The seating arrangement is illustrated in Figure 4.1. This arrangement allows the subjects to easily identify in-group members and out-group members. An instruction sheet is then distributed to every subject and is read aloud by the experimenter. Afterwards, subjects are given a game credit sheet to record investment decisions. When investment decisions are made, the experimenter collects the game credit sheets, calculates the earnings, and returns them to subjects. When the game credit sheet is returned, each player sees on the sheet their own earnings and their game partner's earnings and proceeds to the next round.

Figure 4.1 Seating Arrangement



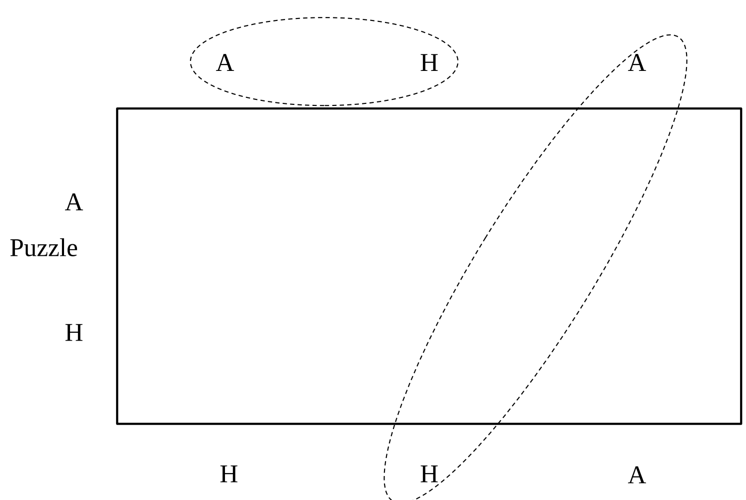
There are three treatments. The first one is referred as Out-group Game Partner (OGP) treatment. In this treatment, subjects are told explicitly that their game partners will be someone from the other major. As shown in Figure 4.2, a participant from major A can only be matched with someone from major H. This treatment is designed to simulate intergroup cooperation and serves as the baseline treatment. In the second treatment,

Figure 4.2 Seating Arrangement of Treatment OGP



referred as Out-group Game Partner with Observed Intergroup Contact (OGPOIC), the matching procedure is the same as in treatment OGP. In addition, the experimenter pairs a classmate of one game player and a classmate of the other game player in solving a jigsaw puzzle. The seating of the puzzle-solving team is illustrated in Figure 4.3. Prior

Figure 4.3 Seating Arrangement of Treatment OGPOIC

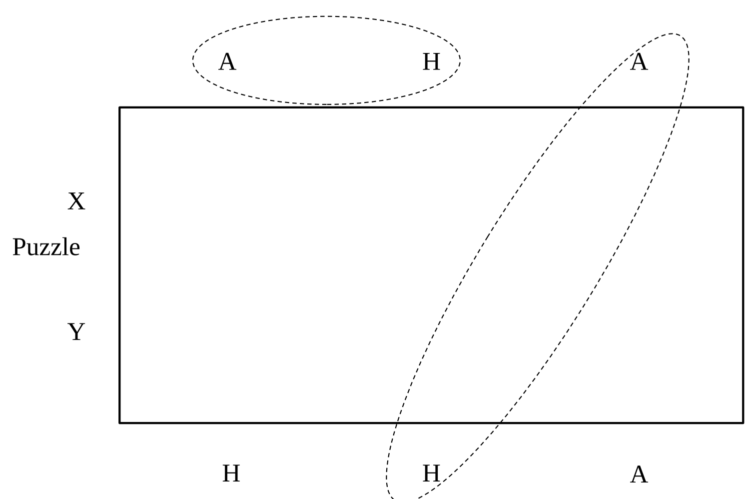


to the public goods game, the players are asked to watch the whole process of the puzzle-solving team working together in solving the jigsaw puzzle. There is no time limit for completing the puzzle and no evaluation on the puzzle-solving team. This treatment is designed to test the public contribution of subjects after going through indirect contact. Comparing treatment OGP and treatment OGPOIC motivates the following hypothesis:

Hypothesis 1: Watching intergroup contact increases intergroup cooperation.

Watching in-group members in contact with out-group members has two effects. One is the indirect contact effect, and the other is the cooperative mood the contact itself brings to subjects. If the latter is the main force, then there is no need to involve in-group members in the contact. To separate these two effects, we design the third treatment, referred as Out-group Game Partner with Observed Random Contact (OGPORC), where the puzzle-solving team consists of two random students recruited in the campus (Figure 4.4). In this case, the difference of public contribution between treatment OGP

Figure 4.4 Seating Arrangement of Treatment OGPORC



and treatment OGPORC is the effect of watching random contact. The importance of having an in-group member in the contact is the difference of public contribution between treatment OGPOIC and treatment OGPORC. Here we propose the second hypothesis:

Hypothesis 2: Watching random contact has limited effect on increasing intergroup cooperation and is not as effective as watching intergroup contact.

4.3.2 Studying Cooperation with the Public Goods Game

The public goods game is the workhorse for studying cooperation. It simulates a social situation where it is socially optimal for people to contribute to the society. However, motivated by pursuing individual interests, the society will end up with an inferior outcome. In order to see the mechanism in our experiment, suppose a social planner who maximizes total payoffs. The optimal plan is to invest all endowment in the public fund for both players. The players will each have a payoff of $(10 - 10) + 1.5(10 + 10)/2 = 15$ credits, with a profit of 5 credits in every round. However, the unique subgame perfect Nash equilibrium is to make zero public investment in each round. To see why, first suppose a static one-shot public goods game with the game parameters. A player's total earnings from investing x credits in the public fund is $(10 - x) + \frac{1.5(x+y)}{2}$, where y is the public investment of the other player. Because this expression is decreasing in x, the optimal public investment is 0. The two players will end the game with their endowment of 10 credits. If this is the unique Nash equilibrium for the one-shot game, then the unique subgame perfect Nash equilibrium in a 10 round game is zero public contribution in every round.

4.3.3 The Jigsaw Puzzle in the Indirect Contact

The jigsaw puzzle is the assembly of small, often oddly shaped, interlocking and tessellating pieces which complete a picture. A natural question is whether this puzzle is the right tool for intergroup contact. In Allport's *The Nature of Prejudice* [4], the author categorized four conditions that would make the contact effective in improving intergroup relations. The first is the equal status of the groups in contact. In our design, the two majors have equal status because they are with similar backgrounds: same university, same campus, close entrance grades, the same year of entrance. The second condition is having a common goal for the groups. In the jigsaw puzzle it is clear that the common goal of the puzzle-solving team is to successfully solve the puzzle. Intergroup cooperation is also required. The jigsaw puzzle is viewed as a more cooperative than a competitive or neutral task. Similar types of puzzles have been used in previous studies. For example, a three-dimensional puzzle was used in Wright et al.'s [94] work on indirect contact effect. The last condition is support of authority, law, and customs. In the experiment, the puzzle solving is supported because it was part of the experiment and has no conflict with any other law, rules, and convention.

Empirical studies show that these conditions are not essential in reducing intergroup bias. Each one of the conditions is sufficient to have an impact on improving intergroup relations (Pettigrew and Tropp [75]). Even contact that does not meet any of these conditions can be effective in reducing prejudice (Pettigrew et al. [76]).

Combining theories and empiric evidence, the jigsaw puzzle in our design is an effective tool in intergroup contact.

4.3.4 Data

Players of the public goods game are undergraduate students from College of Horticulture and College of Agronomy, Northwest A&F University, China. We first contacted major supervisors who are in charge of students' curriculum and non-curriculum activities about conducting an experiment with the students. When the request is approved, we received a list of 114 students from each college. The 114 students at the College of Agriculture consisted of students from three classes of the agronomy major and one class of the plant sciences major. Their game partners at the College of Horticulture belong to students from three classes of the horticulture sciences major and one class of the horticulture facility major. The students are assigned into one of 18 sessions based on their curriculum schedule. Each session is composed of $3-6$ classmates of each major.

Two students did not show up at the experiment. As a result, their partners did not participate in the game as well. The two students who showed up but did not participate in the experiment were compensated with 150 credits (15 yuan in cash). One

Table 4.1 Summary Statistics

College/Treatment	Subjects	Sessions	Female	Minority	Rural
Agriculture	105		39%	7%	60%
Horticulture	105		52%	10%	54%
OGP	64	6	55%	8%	50%
OGPOIC	86	6	44%	8%	58%
OGPORC	60	5	38%	8%	63%

session was delayed by nearly one hour because of one student's late arrival. This waiting process, especially for those from another college, had a significant impact on cooperation behaviors. The public contribution of this session was obviously much lower than other sessions. This session was removed from our data analysis.

After excluding problematic subjects and sessions, there are in total 210 valid subjects and 17 sessions. Table 4.1 displays the descriptive statistics of the subjects by college and treatment.

4.4 Results

In this section, we analyze the treatment effect on the cooperative behavior of the subjects. We first compare the average public investment of the three treatments. To control for individual characteristics that might be related with the cooperative behavior, we consider individual demographic factors and group features in a cross sectional regression where the dependent variable is the subject's average public investment over all 10 rounds. Because the game is repeated 10 rounds, we also look at individual public investment in a panel data regression.

Table 4.2 presents the mean, the standard errors, and the 95% confidence intervals of subjects' overall public contribution by treatment. The highest average public investment is in treatment OGPOIC. Subjects contribute in average 80% of their endowment in the public fund. The public contribution of treatment OGP is the lowest among all

Table 4.2 Mean Percentage Public Investment by Treatment

	OGP	OGPOIC	OGPORC
Mean Percentage Contribution	63%	80%	70%
Standard Error	3%	2%	4%
95% Confidence Interval	(56%, 70%)	(75%, 84%)	(63%, 77%)

treatments: subjects in average contribute only 63% of their endowment. In other words, watching intergroup contact raises the average percentage public investment by 17%. By comparing the confidence intervals at the 95% level, this increase is statistically significant. The average contribution of 70% in treatment OGPORC implies that observing random contact has some impact on subjects' cooperative behaviors, but this impact is not statistically significant at the 5% level.

Because the decision of public investment may be related with individual characteristics and session features, we consider the following equation:

$$Meanpercentagecontribution_i = \alpha + \beta \cdot D_i + \gamma \cdot X_i + \epsilon_i, \tag{4.1}$$

where $Meanpercentagecontribution_i$ is subject i's average percentage contribution to the public fund over all rounds, D_i is a vector of dummy variables for treatments OGPOIC and OGPORC, X_i is a vector of explanatory variables, and ϵ_i is an error term. The magnitude and significance level of the coefficient β_{OGPOIC} on the variable $OGPOIC$ test Hypothesis 1. Compared with a subject in treatment OGP who did not go through the indirect intergroup contact, subjects in treatment OGPOIC contribute from their endowment, in average, β_{OGPOIC} percent more in the public goods game. Hypothesis 2 can be tested by checking the significance of the difference between β_{OGPOIC} and β_{OGPORC}. Because β_{OGPORC} is the increase of public investment after watching random contact, the difference between β_{OGPOIC} and β_{OGPORC} is the importance of having an in-group member in the contact.

Table 4.3 Mean Percentage Public Contribution

Dependent variable: Meanpercentagecontribution$_i$			
Baseline treatment: OGP			
	(1)	(2)	(3)
OGPOIC	16.67**	16.53**	27.19**
	(4.13)	(4.42)	(9.10)
OGPORC	7.16	5.47	9.34
	(4.47)	(4.75)	(8.16)
Female		6.95	7.71
		(3.97)	(4.27)
Minority		-5.16	0.53
		(7.96)	(9.17)
Rural		2.01	-0.53
		(4.01)	(4.18)
Puzzletime			-0.00
			(0.02)
Gendercomposition			0.02
			(0.09)
Groupsize			-3.26**
			(1.10)
N	192	192	171
R^2	0.08	0.21	0.28

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

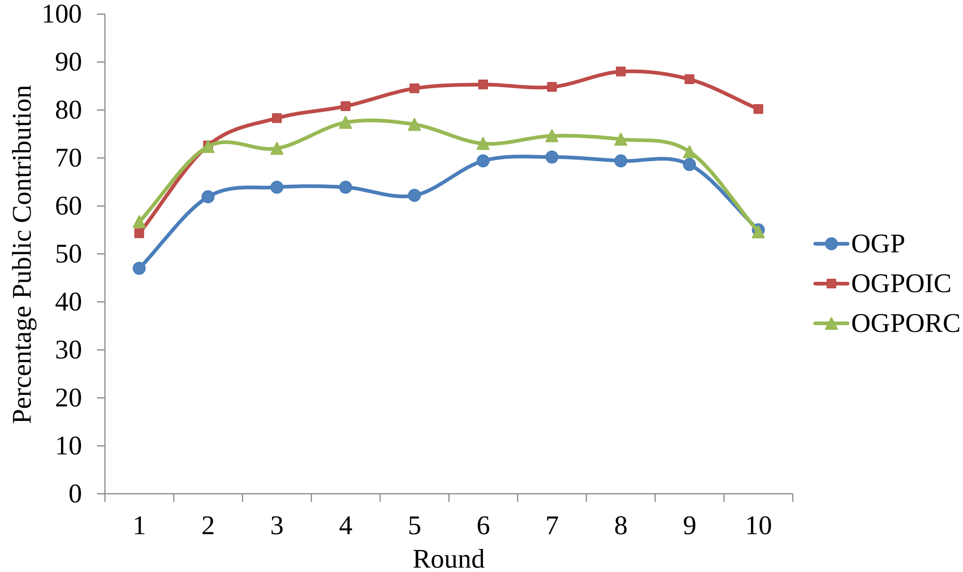
The first specification is reported in the first column of Table 4.3. The coefficient on the variable *OGPOIC* suggests that compared with subjects in treatment OGP, watching intergroup contact pulls public contribution up by 16.67% of the endowment. At the significance level of 1%, Hypothesis 1 is confirmed. Watching random contact has a positive but insignificant effect of 7.16% on increasing intergroup cooperation. On the other hand, the 95% confidence intervals of the two β's suggest that their difference is significant. Therefore, having an in-group member in the indirect intergroup contact is important, and hence Hypothesis 2 is confirmed.

When we consider individual characteristics, the results are similar. Watching intergroup interaction involving in-group members can significantly increase the percentage public contribution by 16.5%. The insignificance of the coefficient β_{OGPORC} and its significant difference from β_{OGPOIC} again confirms Hypothesis 2. Looking at the individual characteristics, females contribute 6.95% more of their endowment than males. This is consistent with the recent public goods experiments by Nowell and Tinkler [71] and Seguino et al. [80].[2] Minorities are less cooperative than Han Chinese by 0.52 units. The rural/urban status seems to have little impact on cooperative behavior.

In the third specification, group features are added. The first group feature is *Puzzletime*, the time that it takes the puzzle solving team to complete the puzzle. It measures the strength and length of the exposure to the indirect contact, and hence may have two opposite effects. A shorter *Puzzletime* may imply that the intergroup contact is more successful, while a longer one may put observers in greater exposure to the contact. The coefficient suggests that the exposure of the intergroup contact does not matter for intergroup cooperation. It is possible, however, that the two effects have cancelled each other. Further investigation is needed. The variable *Groupgendercomposition* is the ratio of females in the out-group in a particular session. This variable represents the expected probability that a subject will be paired with an out-group female in the public goods game. The results that this ratio does not affect public investment decisions. The variable *Groupsize* is also added to control for the possible size effect of the out-group. The coefficient implies that one more member in the out-group significantly decreases the public investment by 0.33 credits.

Because the public goods game lasts 10 rounds, it is important to analyze the dynamic behavior of the subjects. This provides us with more insight into the treatment effect on public contribution in a particular round. Figure 4.5 compares the percentage public contribution of the three treatments. It shows a clear pattern of sharply declining public investment of all treatments in the last round. The reason is straightforward: there

[2] However, the gender difference in cooperation is not conclusive. For example, Brown-Kruse and Hummels [17] and Sell et al. [82] find that females contribute significantly less than males. Cadsby and Maynes [21] and Sell [81] report mixed evidence.



is no incentive for future cooperation. Moreover, the three treatments share a similar shape of public investment that takes a jump of about 15% from the first round to the second round. The line with solid circles draws the average public fund contribution of treatment OGP, which is the lowest of all treatments in every round. The contribution starts from 47% in the first round, jumps to about 62% in the second round, and then stabilizes between 60% to 70% before reaching the last round. The solid squares are the average public investment in treatment OGPOIC. Subjects, on average, contribute 54% of their endowment in the first round and gradually increase the contribution to a maximum of 88% in round 8. Comparing treatments OGP and OGPOIC, for every round, cooperation is enhanced after watching in-group members in intergroup contact. The line with solid triangles shows the trend of average public contribution in treatment OGPORC. The cooperation levels stay between treatment OGPOIC and treatment OGP. The contribution is 10% to 15% higher than in treatment OGP in the first five rounds and converges to the latter in the last five rounds. The gap between OGPOIC and OGPORC is initially small, but gets larger as the game evolves. This justifies Hypothesis 2 that watching an in-group member is the main force of the intergroup contact effect.

Like in the mean public contribution analysis, we add individual and group characteristics as explanatory variables to the dynamic individual cooperative behavior:

$$Percentagecontribution_{i,j} = \delta + \theta \cdot D_i + \eta \cdot X_i + \kappa R + \mu_{i,j}, \tag{4.2}$$

where $Percentagecontribution_{i,j}$ is subject i's percentage public contribution in round j, R is a vector of dummy variables for $j = 1, 2, \ldots, 10$ rounds.

Results are shown in Table 4.4. The indirect intergroup contact effect on raising cooperation is about 17%-27%, and it is statistically significant. Although this time watching random groups in the contact has a significantly positive effect on increasing cooperation by 5.5%-9.3%, the difference between β_{OGPOIC} and β_{OGPORC} is still significant. These results again confirm Hypothesis 1 and Hypothesis 2. The higher public contribution of female subjects is now significant. A female subject puts 6.9%-7.7% more in the public fund than male subjects. The effect of the group features are like before. The time to complete the puzzle and the female ratio in the out-group have little impact on the cooperation. Having one more member in a session reduces the public investment significantly by 0.33 credits.

4.5 Conclusions

We find that, relative to the the control treatment, subjects who observe indirect contact contribute 16.5%-27.2% more of their endowment to the public fund. The effect of random intergroup contact on raising cooperation is limited and significantly lower than the indirect intergroup contact. These findings confirm the hypothesis that indirect contact can improve intergroup relationship, and in turn raise intergroup cooperation.

Our results suggest that indirect contact can be applied when direct contact is restricted. When intergroup cooperation is desired but one or more groups are not available, we can select some members from each group and do demonstrations to the rest. This is particularly useful for majority-minority intergroup cooperation, and for groups

Table 4.4 Dynamic Individual Percentage Public Contribution

Dependent variable: $\text{Percentagecontribution}_{i,j}$ Baseline treatment: OGP	(1)	(2)	(3)	(4)
OGPOIC	16.67**	16.53**	27.19**	27.19**
	(1.90)	(1.98)	(4.11)	(8.80)
OGPORC	7.16**	5.47**	9.34*	9.34
	(2.06)	(2.13)	(3.69)	(7.92)
Female		6.95**	7.71**	7.71*
		(1.78)	(1.93)	(3.60)
Minority		-5.16	0.53	0.53
		(3.57)	(4.14)	(7.93)
Rural		2.01	-0.53	-0.53
		(1.80)	(1.89)	(3.85)
Puzzletime			-0.00	-0.00
			(0.01)	(0.02)
Gendercomposition			0.02	0.02
			(0.04)	(0.09)
Groupsize			-3.27**	-3.27**
			(0.50)	(1.13)
N	1920	1920	1710	1710
R^2	0.08	0.14	0.18	0.18

a. Significance levels of 5% and 1% are denoted by * and **, respectively.
b. Province of origin dummies are included in all columns.

that are segregated in many dimensions. Indirect contact also implies financial freedom, as getting every group member involved in direct intergroup contact is very costly.

Some questions remain open to future studies. First, is indirect contact as effective as direct contact? If the answer is yes, the implication is magnificent. For a firm manager who manages a diversity of employees, direct interaction is often used to eliminate intergroup bias and to promote cooperation. In this case, indirect intergroup interaction can save a lot of resources when selected employees demonstrate intergroup contact to the rest of the employees. Second, because indirect contact has three forms and only one is tested in the present work, it is desired to test the other two and conduct a comparison of these 3 forms. The comparison may include questions like: (1) which is the most effective intervention and (2) which is the cheapest one. It is also interesting to check other types of contact. The jigsaw puzzle in our experiment is viewed as a positive and cooperative task. Do the results stay the same if it is replaced by a competitive or even negative task?

APPENDIX A. EXPERIMENTAL INSTRUCTIONS FOR CHAPTER 1

Page 1

There are two roles in this experiment: *workers* and *employers*. In this section, you will play the role of the worker. In the next section we will invite you to play the role of the employer.

Workers have the task to solve as many character puzzles as possible within a 5 minute period. You will be able to perform a few practice puzzles on the next page to familiarize yourself with the task. For each puzzle that you solve during the 5 minute period, you will receive 40 credits. For example, if you solve 5 puzzles, you will receive 200 credits.

As a worker, you will be evaluated by several employers who set your wages. Each employer will see your performance in a timed practice game and might also see your gender, Hukou, ethnicity or major. The employer's task is to estimate as precisely as possible how many puzzles you are able to perform during the 5 minute period. The employer's earnings will be higher the better he/she predicts your performance.

The employers' estimates of your puzzle-solving skills can increase your earnings as a worker. For each employer, you earnings will increase by the employers' average estimate of your puzzle-solving skills times 40 credits. For example, if the employers estimate

on average that you can solve 5 puzzles, then you would receive 5 times 40 = 200 credits additionally.

Page 2

On this page, you have the opportunity to solve two example puzzles to familiarize yourself with your task. The square with characters on the right differs from the square of characters on the left in two letters. You have to find those letters and click on them to solve the puzzle.

CHARACTER PUZZLE

Page 3

On the next page, you are asked to solve a timed practice game. You see a running clock that measures your time until you solve the puzzle. This practice time will be visible to employers who later estimate your puzzle-solving ability.

Remember, that the higher each employers' average estimate of your puzzle-solving ability, the higher are your earnings, as the employers' average estimate will be multiplied by 40 credits and added to your earnings.

Only go to the next page when you are ready. The practice game will start immediately.

Page 4

Please solve this time practice puzzle as quickly as possible.

CHARACTER PUZZLE

Page 5

On the next page, you are asked to solve as many puzzles as possible within a 5 minute period. You will receive 40 credits for each solved puzzle.

Only go to the next page when you are ready. The game will start immediately.

Page 6

Please solve as many puzzles as possible within the next 5 minutes.

CHARACTER PUZZLE

Page 7

In this section of the experiment, you will play the role of the employer.

On the next page, we will ask you to evaluate 10 workers who just completed their 5 minute puzzle-solving task.

As an employer, you have to estimate the performance of each worker. We will provide you with some basic information about each worker, such as worker's performance in the timed practice puzzle. For each worker, you will receive 150 credits if you predict the worker's performance in the 5 minute task precisely. If your estimate is off by X puzzles for this worker, then you will receive 150 credits minus X times 10 credits. For example: If you predict that the worker can solve 5 puzzles and he or she solves 3, then your earnings are 130 credits (150 credits minus 2 times 10 credits). Similarly, if a worker solves 8 puzzles and you predict that he or she can solve 5 puzzles, then your estimate is off by 3 and you earn earns 120 credits (150 credits minutes 3 times 10 credits).

Your estimates of a worker's puzzle-solving skills can increase that worker's earnings. Each worker will be evaluated by several employers, and the worker's earnings will increase by the average estimate of all employers times 40 credits.

Page 8

EVALUATION OF WORKERS' PRODUCTIVITY

APPENDIX B. EXPERIMENTAL INSTRUCTIONS FOR CHAPTER 2

Page 1

There are two roles in this experiment: *workers* and *employers*. In this section, you will play the role of the worker. In the next section we will invite you to play the role of the employer.

Workers have the task to solve as many character puzzles as possible within a 5 minute period. You will be able to perform a few practice puzzles on the next page to familiarize yourself with the task. For each puzzle that you solve during the 5 minute period, you will receive 40 credits. For example, if you solve 5 puzzles, you will receive 200 credits.

As a worker, you will be evaluated by several employers who set your wages. Each employer will see your performance in a timed practice game and might also see your gender, Hukou, ethnicity or major. The employer's task is to estimate as precisely as possible how many puzzles you are able to perform during the 5 minute period. The employer's earnings will be higher the better he/she predicts your performance.

The employers' estimates of your puzzle-solving skills can increase your earnings as a worker. For each employer, you earnings will increase by the employers' average estimate of your puzzle-solving skills times 40 credits. For example, if the employers estimate

on average that you can solve 5 puzzles, then you would receive 5 times 40 = 200 credits additionally.

Page 2

On this page, you have the opportunity to solve two example puzzles to familiarize yourself with your task. The square with characters on the right differs from the square of characters on the left in two letters. You have to find those letters and click on them to solve the puzzle.

CHARACTER PUZZLE

Page 3

On the next page, you are asked to solve a timed practice game. You see a running clock that measures your time until you solve the puzzle. This practice time will be visible to employers who later estimate your puzzle-solving ability.

Remember, that the higher each employers' average estimate of your puzzle-solving ability, the higher are your earnings, as the employers' average estimate will be multiplied by 40 credits and added to your earnings.

Only go to the next page when you are ready. The practice game will start immediately.

Page 4

Please solve this time practice puzzle as quickly as possible.

CHARACTER PUZZLE

Page 5

Your performance in the timed practice puzzle was __ seconds.

Your next task is to solve as many puzzle as possible within a 5 minute period. You will receive 40 credits for each solved puzzle. If you take the same amount of time for each puzzle as you took for the practice puzzle, you are projected to solve __ puzzles.

Before starting this task, please give your best estimate on how many puzzles you will be able to solve during these 5 minutes. We will pay you 150 credits if your estimate is exactly correct. If your estimate is off by X puzzles, then you will receive 150 credits minus X times 10 credits.

My best estimate for the number of puzzles that I am able to solve in a 5 minute period is __

Page 6

On the next page, you are asked to solve as many puzzles as possible within a 5 minute period. You will receive 40 credits for each solved puzzle.

Only go to the next page when you are ready. The game will start immediately.

Page 7

Please solve as many puzzles as possible within the next 5 minutes.

CHARACTER PUZZLE

Page 8

In this section of the experiment, you will play the role of the employer.

On the next page, we will ask you to evaluate 10 workers who just completed their 5 minute puzzle-solving task.

As an employer, you have to estimate the performance of each worker. We will provide you with some basic information about each worker, such as worker's performance in the timed practice puzzle. For each worker, you will receive 150 credits if you predict the worker's performance in the 5 minute task precisely. If your estimate is off by X puzzles for this worker, then you will receive 150 credits minus X times 10 credits. For example: If you predict that the worker can solve 5 puzzles and he or she solves 3, then your earnings are 130 credits (150 credits minus 2 times 10 credits). Similarly, if a worker solves 8 puzzles and you predict that he or she can solve 5 puzzles, then your estimate is off by 3 and you earn earns 120 credits (150 credits minutes 3 times 10 credits).

Your estimates of a worker's puzzle-solving skills can increase that worker's earnings. Each worker will be evaluated by several employers, and the worker's earnings will increase by the average estimate of all employers times 40 credits.

Page 9

EVALUATION OF WORKERS' PRODUCTIVITY

Page 10

On this page, we give you the opportunity to revise your evaluations of the same 10 workers. This time we also provide you with the worker's own estimate on how many puzzles he or she thinks can solve. For each worker, we remind you of the evaluation you provided on the previous page.

RE-EVALUATION OF WORKERS' PRODUCTIVITY

APPENDIX C. EXPERIMENTAL INSTRUCTIONS FOR CHAPTER 3

Page 1

There are two roles in this experiment: *workers* and *employers*. In this section, you will play the role of the worker. In the next section we will invite you to play the role of the employer.

Workers have the task to solve as many character puzzles as possible within a 5 minute period. You will be able to perform a few practice puzzles on the next page to familiarize yourself with the task. For each puzzle that you solve during the 5 minute period, you will receive 40 credits. For example, if you solve 5 puzzles, you will receive 200 credits.

As a worker, you will be evaluated by several employers who set your wages. Each employer will see your performance in a timed practice game and might also see your gender, Hukou, ethnicity or major. The employer's task is to estimate as precisely as possible how many puzzles you are able to perform during the 5 minute period. The employer's earnings will be higher the better he/she predicts your performance.

The employers' estimates of your puzzle-solving skills can increase your earnings as a worker. For each employer, you earnings will increase by the employers' average estimate of your puzzle-solving skills times 40 credits. For example, if the employers estimate

on average that you can solve 5 puzzles, then you would receive 5 times 40 = 200 credits additionally.

Page 2

On this page, you have the opportunity to solve two example puzzles to familiarize yourself with your task. The square with characters on the right differs from the square of characters on the left in two letters. You have to find those letters and click on them to solve the puzzle.

CHARACTER PUZZLE

Page 3

On the next page, you are asked to solve a timed practice game. You see a running clock that measures your time until you solve the puzzle. This practice time will be visible to employers who later estimate your puzzle-solving ability.

Remember, that the higher each employers' average estimate of your puzzle-solving ability, the higher are your earnings, as the employers' average estimate will be multiplied by 40 credits and added to your earnings.

Only go to the next page when you are ready. The practice game will start immediately.

Page 4

Please solve this time practice puzzle as quickly as possible.

CHARACTER PUZZLE

Page 5

On the next page, you are asked to solve as many puzzles as possible within a 5 minute period. You will receive 40 credits for each solved puzzle.

Only go to the next page when you are ready. The game will start immediately.

Page 6

Please solve as many puzzles as possible within the next 5 minutes.

CHARACTER PUZZLE

Page 7

In this section of the experiment, you will play the role of the employer.

As an employer, you have to estimate the performance of each worker. We will provide you with some basic information about each worker, such as worker's performance in the timed practice puzzle. For each worker, you will receive 150 credits if you predict the worker's performance in the 5 minute task precisely. If your estimate is off by X puzzles for this worker, then you will receive 150 credits minus X times 10 credits. For example: If you predict that the worker can solve 5 puzzles and he or she solves 3, then your earnings are 130 credits (150 credits minus 2 times 10 credits). Similarly, if a worker solves 8 puzzles and you predict that he or she can solve 5 puzzles, then your estimate is off by 3 and you earn earns 120 credits (150 credits minutes 3 times 10 credits).

Your estimates of a worker's puzzle-solving skills can increase that worker's earnings. Each worker will be evaluated by several employers, and the worker's earnings will increase by the average estimate of all employers times 40 credits.

Page 8

For each worker you evaluate, we will provide you with 3 of the following 4 characteristics for free: practice time, gender, ethnicity, urban/rural.

One of these 4 characteristics will not be free to view. We will ask you for your willingness to pay for each of these 4 characteristics. Then we will randomly select one characteristic and a price you have to pay a price between 0 and 150 credits per worker to see this information. A computer program has already selected a random price between 0 and 150 for each piece of information. We will not tell you this price. Instead, we will ask you - for each characteristic - how much you would be willing to pay at most per student to see that piece of information.

You will see the hidden characteristic only when your willingness to pay is higher than or equal to the price that is randomly determined by the computer. Please indicate your willingness to pay for:

- Practice time
- Gender
- Ethnicity
- Urban/rural

Page 9

EVALUATION OF WORKERS' PRODUCTIVITY

APPENDIX D. EXPERIMENTAL INSTRUCTIONS FOR CHAPTER 4

(Treatment OGP)

Participants in this experiment are students from the College of Agriculture and the College of Horticulture. The experiment will be in groups of two students. You will be in a group with someone from **the other** college. For example, if you are from the College of Agriculture, you will be in a group with someone from the College of Horticulture, and vice versa.

Each of you is requested to make a decision on how to invest 10 credits. There is a private fund and a public fund. For each credit invested in the private fund, you will receive one credit. Each credit invested in the public fund will yield 1.5 credits for the group. Each person in the group will receive half of the money in the public fund. Therefore, your earnings from the investment are the money in your personal fund and half of the public fund.

The game will repeat 10 rounds. Credits cannot be accumulated to the next round. In other words, you will start with 10 credits in every round. Your final earnings are the sum of credits of all 10 rounds.

You will be given a Game Credit Form, on which you record your investment decisions.

In the end of each round, the experimenter will collect this form, calculate and report the followings on the form: the other player's investment decision, the amount of money in the public fund, your earnings, the other player's earnings. Afterwards, the form will be returned to you. When you get the form back, make your investment decision for the next round.

In the end of the experiment, you can cash in your earnings at the rate of 10 credits to 1 RMB.

Thank you for your participation!

(Treatments OGPOIC and OGPORC)

Participants in this experiment are students from the College of Agriculture and the College of Horticulture. The experiment will be in groups of two students. You will be in a group with someone from **the other** college. For example, if you are from the College of Agriculture, you will be in a group with someone from the College of Horticulture, and vice versa.

Each of you is requested to make a decision on how to invest 10 credits. There is a private fund and a public fund. For each credit invested in the private fund, you will receive one credit. Each credit invested in the public fund will yield 1.5 credits for the group. Each person in the group will receive half of the money in the public fund. Therefore, your earnings from the investment are the money in your personal fund and half of the public fund.

The game will repeat 10 rounds. Credits cannot be accumulated to the next round. In other words, you will start with 10 credits in every round. Your final earnings are the sum of credits of all 10 rounds.

You will be given a Game Credit Form, on which you record your investment decisions. In the end of each round, the experimenter will collect this form, calculate and report the followings on the form: the other player's investment decision, the amount of money in the public fund, your earnings, the other player's earnings. Afterwards, the form will be returned to you. When you get the form back, make your investment decision for the next round.

In the end of the experiment, you can cash in your earnings at the rate of 10 credits to 1 RMB.

Thank you for your participation!

Before the experiment starts, a group of two will be assembling a jigsaw puzzle. Please stay quiet and patient, and wait until that group completes the jigsaw puzzle.

Thank you for your participation!

BIBLIOGRAPHY

[1] **Adorno, T.W.; Frenkel-Brunswik, Else; Levinson, Daniel, and Sanford, Nevitt (1950)**. *The Authoritarian Personality.* New York: Harper & Row.

[2] **Afridi, Farzana; Li, Sherry Xin, and Ren, Yufei (2012)**. "Social Identity and Inequality: The Impact of China's Hukou System." *IZA Discussion Paper*, 6417.

[3] **Alfnes, Frode and Rickertsen, Kyrre (2003)**. "European Consumers' Willingness to Pay for U.S. Beef in Experimental Auction Markets." *American Journal of Agricultural Economics*, 85(2), 396-405.

[4] **Allport, Gordon (1954)**. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

[5] **Appleton, Simon; Song, Lina, and Xia, Qingjie (2005)**. "Has China Crossed the River? The Evolution of Wage Structure in Urban China During Reform and Retrenchment." *Journal of Comparative Economics*, 33(4), 644-663.

[6] **Arrow, Kenneth (1998)**. "What Has Economics to Say about Racial Discrimination?" *Journal of Economic Perspectives*, 12(2), 91-100.

[7] **Babcock, Linda and Laschever, Sara (2003)**. *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press.

[8] **Becker, Gordon; Degroot, Morris, and Marschak, Jacob (1964)**. "Measuring Utility by a Single-response Sequential Method." *Behavioral Science*, 9(3), 226-232.

[9] **Ben-Ner, Avner; McCall, Brian; Stephane, Massoud, and Wang, Hua (2009)**. "Identity and In-group/out-group Differentiation in Work and Giving Behaviors: Experimental Evidence." *Journal of Economic Behavior and Organization*, 72(1), 153-170.

[10] **Benabou, Roland and Tirole, Jean (2002)**. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics*, 117(3), 871-915.

[11] **Bernhard, Helen; Ernst, Fehr, and Fischbacher Urs (2006)**. "Group Affiliation and Altruistic Norm Enforcement." *American Economic Review*, 96(2), 217-221.

[12] **Bertrand, Marianne and Mullainathan, Sendhil (2004)**. "Are Emily and Greg more Employable than Lakisha and Jamal?" *American Economic Review*, 94(4), 991-1013.

[13] **Billig, Michael and Tajfel, Henri (1973)**. "Social Categorization and Similarity in Intergroup Behaviour." *European Journal of Social Psychology*, 3(1), 27-52.

[14] **Bohm, Peter; Linden, Johan, and Sonnegard, Joakim (1997)**. "Eliciting Reservation Prices: Becker-DeGroot-Marschak Mechanism vs. Markets." *The Economic Journal*, 107(443), 1079-1089.

[15] **Boisjoly, Johanne; Duncan, Greg; Kremer, Michael; Levy, Dan, and Eccles, Jacque (2006)**. "Empathy or Antipathy? The Impact of Diversity." *American Economic Review*, 96(5), 1890-1905.

[16] **Borjas, George (1999)**. “The Economic Analysis of Immigration.” Handbook of Labor Economics, Volume 3.

[17] **Brown-Kruse, Jamie and Hummels, David (1993)**. “Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is?“ *Journal of Economic Behavior and Organization*, 22(3), 255-267.

[18] **Buchan, Nancy; Johnson Eric, and Croson Rachel (2006)**. “Let’s Get Personal: An International Examination of the Influence of Communication, Culture and Social Distance on Other Regarding Preferences.” *Journal of Economic Behavior and Organization*, 60(3), 373-398.

[19] **Burks, Stephen; Carpenter, Jeffrey; Goette, Lorenz, and Rustichini, Aldo (2013)**. “Overconfidence and Social Signalling.” *Review of Economic Studies*, 0, 1-35.

[20] **Buzby, Jean; Fox, John; Ready, Richard, and Crutchfield, Stephen (1998)**. “Measuring Consumer Benefits of Food Safety Risk Reductions.” *Journal of Agricultural and Applied Economics*, 30(1), 69-82.

[21] **Cadsby, Bram and Maynes, Elizabeth (1998)**. “Gender and free riding in a threshold public goods game: Experimental evidence.“ *Journal of Economic Behavior and Organization*, 34(4), 603-620.

[22] **Cameron, Lindsey and Rutland, Adam (2006)**. “Extended Contact through Story Reading in School: Reducing Childrens Prejudice toward the Disabled.” *Journal of Social Issues*, 62(3), 469-488.

[23] **Cameron, Lindsey; Rutland, Adam; Brown, Rupert, and Douch, Rebecca (2006)**. “Changing Childrens Intergroup Attitudes Toward Refugees: Testing Different Models of Extended Contact.” *Child Development*, 77(5), 1208-1219.

[24] **Campbell, D.T. (1965)**. "Ethnocentric and Other Altruistic Motives." *Nebraska Symposium on Motivation*, 13. Lincoln: University of Nebraska Press.

[25] **Charness, Gary; Rigotti, Luca, and Rustichini, Aldo (2007)**. "Individual Behavior and Group Membership." *American Economic Review*, 97(4), 1340-1352.

[26] **Chen, Roy and Chen, Yan (2011)**. "The Potential of Social Identity for Equilibrium Selection." *American Economic Review*, 101(6), 2562-2589.

[27] **Chen, Yan and Li, Sheery Xin (2009)**. "Group Identity and Social Preferences." *American Economic Review*, 99(1), 431-457.

[28] **China Statistical Yearbook (various years)**. National Bereau of Statistics of China.

[29] **Coate, Stephen and Loury, Glenn (1993)**. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, 83(5), 1220-1240.

[30] **Crisp, Richard and Turner, Rhiannon (2009)**. "Can Imagined Interactions Produce Positive Perceptions?" *American Psychologist*, 64(4), 231-240.

[31] **Dong, Xiao Yuan and Bowles, Paul (2002)**. "Segmentation and Discrimination in China's Emerging Industrial Labor Market." *China Economic Review*, 13(2-3), 170-196.

[32] **Duckworth, Angela and Seligman, Martin (2005)**. "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents." *Psychological Science*, 16(12), 939-944.

[33] **Eckel, Catherine and Grossman, Philip (2005)**. "Managing Diversity by Creating Team Identity." *Journal of Economic Behavior and Organization*, 58(3), 371-392.

[34] **Falk, Armin; Huffman, David, and Sande, Uwe (2006)**. "Self-confidence and Search." *IZA Discussion Paper*, 2525.

[35] **Fehr, Urnst and Fischbacher Urs (2003)**. "The Nature of Human Altruism."*Nature*, 425, 785-791.

[36] **Fershtman, Chaim and Gneezy, Uri (2001)**. "Discrimination in a Segmented Society: An Experimental Approach." *Quarterly Journal of Economics*, 116(1), 351-377.

[37] **Fiedler, Marina and Haruvy, Ernan (2009)**. "The Lab Versus The Virtual Lab and Virtual FieldAn Experimental Investigation of Trust Games with Communication." *Journal of Economic Behavior and Organization*, 72(2), 716-724.

[38] **Fiedler, Marina; Haruvy, Ernan, and Li, Sherry Xin (2011)**. "Social Distance in a Virtual World Experiment." *Games and Economic Behavior*, 72(2), 400-426.

[39] **Fleisher, Belton and Chen, Jian (1997)**. "The Coast-noncoast Income Gap, Productivity, and Regional Economic Policy in China." *Journal of Comparative Economics*, 25(2), 220-236.

[40] **Fortin, Nicole (2008)**. "The Gender Wage Gap Among Young Adults in the United States: The Importance of Money Versus People." *Journal of Human Resources*, 43(4), 884-918.

[41] **Gneezy, Uri; Leonard, Kenneth, and List, John (2009)**. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica*, 77(5), 1637-1664.

[42] **Goette, Lorenz; Huffman, David, and Meier, Stephan (2006)**. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups. "*American Economic Review*, 96(2), 212-216.

[43] **Goldsmith, Arthur; Veum, Jonathan, and Darity, William (1997)**. "The Impact of Psychological and Human Capital on Wages." *Economic Inquiry*, 35(4), 815-829.

[44] **Granger, C.W.J. (1989)**. *Forecasting in Business and Economics*. San Diego, CA: Academic Press.

[45] **Gustafsson, Bjorn and Shi, Li (2003)**. "The Ethnic Minority-Majority Income Gap in Rural China During Transition." *Economic Development and Cultural Changes*, 51(4), 805-822.

[46] **Hasmath, Reza; Ho, Benjamin, and Liu, Elaine (2012)**. "Ethnic Minority Disadvantages in Chinas Labour Market?" *University of Oxford China Growth Centre Discussion Paper*, 16, 1-35.

[47] **Hayes, Dermot; Shogren, Jason; Shin, Seung Youll, and Kliebenstein, James (1995)**. "Valuing Food Safety in Experimental Auction Markets." *American Journal of Agricultural Economics*, 77(1), 40-53.

[48] **Heckman, James and Rubinstein, Yona (2001)**. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review Papers and Proceedings*, 91(2), 145-149.

[49] **Heckman, James; Stixrud, Jora, and Urzua, Sergio (2006)**. "The Effect of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *NBER Working Paper*, 12006.

[50] **Hogg, Michael and Abrams, Dominic (1988)**. *Social Identifications: A Social Psychology of Intergroup Relations and Group Processes.* New York: Routledge.

[51] **Huffman, Wallace; Shogren, Jason; Rousu, Matthew, and Tegene, Abebayehu (2003)**. "Consumer Willingness to Pay for Genetically Modified Food Labels in a Market with DiverseInformation: Evidence from Experimental Auctions." *Journal of Agricultural and Resource Economics*, 28(3), 481-502.

[52] **Husnu, Senel and Crisp, Richard (2010)**. "Imagined Intergroup Contact: A New Technique for Encouraging Greater Inter-Ethnic Contact in Cyprus." *Peace and Conflict*, 16(1), 97-108.

[53] **Irwin, Julie; McClellan, Gary; McKee, Michael; Schulze, William, and Nordan, Elizabeth (1998)**. "Payoff Dominance vs. Cognitive Transparency in Decision Making." *Economic Inquiry*, 36(2), 272-285.

[54] **Johnson, Emily and Chow, Gregory C (1997)**. "Rates of Return to Schooling in China." *Pacific Economic Review*, 2(2), 101-113.

[55] **Kanbur, Ravi and Zhang, Xiaobo (1999)**. "Which Regional Inequality? The Evolution of Rural-Urban and Inland-Coastal Inequality in China from 1983-1995." *Journal of Comparative Economics*, 27(4), 686-701.

[56] **Keller, Robin; Segal, Uzi, and Wang, Tan (1993)**. "The Becker-DeGroot-Marschak Mechanism and Generalized Utility Theories: Theoretical Predictions and Empirical Observations." *Theory and Decision*, 34(2), 83-97.

[57] **Knight, John and Song, Lina (2003)**. "Increasing Urban Wage Inequality in China." *Economics of Transition*, 11(4), 597-619.

[58] **Kollock, Peter (1998)**. "Transforming Social Dilemmas: Group Identity and Co-operation." in Peter A. Danielson, ed., *Modeling Rationality, Morality and Evolution*. Oxford: Oxford University Press, 186-210.

[59] **Li, Haizheng (2003)**. "Economic Transition and Returns to Education in China." *Economics of Education Review*, 22(3), 317-328.

[60] **Liebkind, Karmela and McAlister, Alfred (1999)**. "Extended contact through peer modelling to promote tolerance in Finland." *European Journal of Social Psychology*, 29(5-6), 765-780.

[61] **Lin, Justin Y.; Wang, Gewei, and Zhao, Yaohui (2004)**. "Regional Inequality and Labor Transfers in China." *Economic Development and Cultural Change*, 52(3), 587-603.

[62] **Liu, Pak-Wai; Meng, Xin, and Zhang, Junsen (2000)**. "Sectoral Gender Wage Differentials and Discrimination in the Transitional Chinese Economy." *Journal of Population Economics*, 13(2), 331-352.

[63] **Lu, Zhigang and Song, Shunfeng (2006)**. "Rural-urban migration and Wage Determination: The Case of Tianjin, China." *China Economic Review*, 17(3), 337-345.

[64] **Mankiw, Gregory and Shapiro, Matthew (1986)**. "Do We Reject Too Often? Small Sample Properties of Tests of Rational Expectations Models." *Economics Letters*, 20(2), 139-145.

[65] **Maurer-Fazio, Margaret (2012)**. "Ethnic Discrimination in China's Internet Job Board Labor Market." *IZA Journal of Migration*, 1-12.

[66] **Mayo, Elton (1933)**. *The Human Problems of an Industrial Civilization*. New York: MacMillan.

[67] **Meng, Xin and Zhang, Junsen (2001)**. "The Two-Tier Labor Market in Urban China: Occupational Segregation and Wage Differentials between Urban Residents and Rural Migrants in Shanghai." *Journal of Comparative Economics*, 29(3), 485-504.

[68] **Mobius, Markus and Rosenblat, Tanya (2006)**. "Why Beauty Matters?" *American Economic Review*, 96(1), 222-235.

[69] **Moody, James (2001)**. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology*, 107(3), 679-716.

[70] **Niederle, Muriel and Vesterlund, Lisa (2007)**. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, 122(3), 1067-1101.

[71] **Nowell, Clifford and Tinkler, Sarah (1994)**. "The Influence of Gender on the Provision of A Public Good." *Journal of Economic Behavior and Organization*, 25(1), 25-36.

[72] **Paolini, Stefania; Hewstone, Miles; Cairns, Ed, and Voci, Alberto (2004)**. "Effects of Direct and Indirect Cross-Group Friendships on Judgments of Catholics and Protestants in Northern Ireland: The Mediating Role of an Anxiety-Reduction Mechanism." *Personality and Social Psychology Bulletin*, 30(6), 770-786.

[73] **Persico, Nicola; Postlewaite, Andrew, and Silverman, Dan (2004)**. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Political Economy*, 112(5), 1019-1053.

[74] **Pettigrew, Thomas; Christ, Oliver; Wagner, Ulrich, and Stellmacher, Jost (2007)**. "Direct and Indirect Intergroup Contact Effects on Prejudice: A Normative Interpretation." *International Journal of Intercultural Relations*, 31(4), 411-425.

[75] **Pettigrew, Thomas and Tropp, Linda (2006)**. "A Meta-Analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology*, 90(5), 751-783.

[76] **Pettigrew, Thomas; Tropp, Linda; Wagner, Ulrich, and Christ, Oliver (2011)** "Recent Advances in Intergroup Contact Theory." *International Journal of Intercultural Relations*, 35(3), 271-280.

[77] **Rozan, Anne; Stenger, Anne, and Willinger, Marc (2004)**. "Willingness-to-pay for Food Safety: An Experimental Investigation of Quality Certification on Bidding Behaviour." *European Review of Agricultural Economics*, 31(4), 409-425.

[78] **Ruffle, Bradley and Sosis Richard (2006)**. "Cooperation and The In-group-out-group Bias: A Field Test on Israeli Kibbutz Members and City Residents." *Journal of Economic Behavior and Organization*, 60, 147-163.

[79] **Schiappa, Edward; Gregg, Peter, and Hewes, Dean (2005)**. "The Parasocial Contact Hypothesis." *Communication Monographs*, 72(1), 92-115.

[80] **Seguino, Stephanie; Stevens, Thomas, and Lutz, Mark (1996)**. "Gender and Cooperative Behavior: Economic Man Rides Alone." *Feminist Economics*, 2(1), 1-21.

[81] **Sell, Jane (1997)**. "Gender, Strategies, and Contributions to Public Goods." *Social Psychology Quarterly*, 60(3), 252-265.

[82] **Sell, Jane; Griffith, W., and Wilson, Rick (1993)**. "Are Women More Cooperative Than Men in Social Dilemmas?" *Social Psychology Quarterly*, 56(3), 211-222.

[83] **Senen, Ingrid and Schram, Arthur (2006)**. "Social Status and Group Norms: Indirect Reciprocity in a Repeated Helping Experiment." *European Economic Review*, 50(3), 581-602.

[84] **Sherif, Muzafer; Harvey, O.J.; White, Jack; Hood, William, and Sherif, Carolyn (1961)**. *Intergroup Conflict and Cooperation: The Robbers Cave Experiment.* Norman, OK: University Book Exchange.

[85] **Solow, John and Kirkwood, Nicole (2002)**. "Group Identity and Gender in Public Goods Experiments." *Journal of Economic Behavior and Organization*, 48(4), 403-412.

[86] **Tajfel, Henri (1970)**. "Experiments in Intergroup Discrimination." *Scientific American*, 223(5), 96-102.

[87] **Tajfel, Henri; Billig, Michael; Bundy, R.P., and Flament, Claude (1971)**. "Social Categorization and Intergroup Behaviour. " *European Journal of Social Psychology*, 1(2), 149-178.

[88] **Turner, Rhiannon and Crisp, Richard (2010)**. "Imagining Intergroup Contact Reduces Implicit Prejudice." *British Journal of Social Psychology*, 49(1), 129-142.

[89] **Turner, Rhiannon; Crisp, Richard, and Lambert, Emily (2007)**. "Imagining Intergroup Contact Can Improve Intergroup Attitudes." *Group Processes Intergroup Relations*, 10(4), 427-441.

[90] **Turner, Rhiannon; Hewstone, Miles, and Voci, Alberto (2007)**. "Reducing Explicit and Implicit Outgroup Prejudice Via Direct and Extended Contact: The Mediating Role of Self-Disclosure and Intergroup Anxiety." *Journal of Personality and Social Psychology*, 93(3), 369-388.

[91] **Turner, Rhiannon; Hewstone, Miles; Voci, Alberto, and Vonofakou, Christiana (2008)**. "A Test of the Extended Intergroup Contact Hypothesis: The Mediating Role of Intergroup Anxiety, Perceived Ingroup and Outgroup Norms, and Inclusion of the Outgroup in the Self." *Journal of Personality and Social Psychology*, 95(4), 843-860.

[92] **Weisbuch, Max; Pauker, Kristin, and Ambady, Nalini (2009)**. "The Subtle Transmission of Race Bias via Televised Nonverbal Behavior." *Science*, 326(18), 1711-1714.

[93] **World Bank (1997)**. *Sharing Rising Incomes: Disparities in China*. Washington, DC: World Bank.

[94] **Wright, Stephen; Aron, Arthur; McLaughlin-Volpe, Tracy, and Ropp, Stacy (1997)**. "The Extended Contact Effect: Knowledge of Cross-Group Friendships and Prejudice." *Journal of Personality and Social Psychology*, 73(1), 73-90.

[95] **Yang, Danis (2005)**. "Determinants of Schooling Returns During Transition: Evidence from Chinese Cities." *Journal of Comparative Economics*, 33(2), 244-264.

[96] **Yao, Shujie and Zhang, Zongyi (2001)**. "On Regional Inequality and Diverging Clubs: A Case Study of Contemporary China." *Journal of Comparative Economics*, 29(3), 466-484.

[97] **Zhang, Jane (2012)**. "Narrowing the gender gap in Competitive Inclination: Evidence from the Communist Experiment in China." *Working Paper*, Hong Kong University of Science and Technology.