

Labor Market Discrimination against Family Responsibilities:

A Correspondence Study with Policy Change in China

Haoran He[†], Sherry Xin Li[‡], Yuling Han^{*}

Abstract

China's highly controversial one-child policy (1979 ~ 2015) was shifted to the two-child policy in January 2016. We use the heterogeneities in the pre-2016 policy environment and the unexpected policy change in 2016 as two sources of identification to investigate statistical discrimination against the expected family responsibilities. In a two-wave correspondence study before and after the policy change, about 9,000 fictitious resumes were sent to real online job advertisements. The information on the fictitious applicants' gender and whether they were an only- or sibling-child were systematically varied. We find that women, but not men, are subject to statistical discrimination for their expected family responsibilities. This discrimination worsens as the probability of maternity increases with age.

Keywords: labor market discrimination, parenthood penalty, correspondence study, field experiment, one-child policy, China

JEL Classification: C93, J71, M51

[†] Business School, Beijing Normal University, Beijing 100875, China; E-mail address: haoran.he@bnu.edu.cn; Tel.: 86-10-58807874.

[‡] School of Economic, Political and Policy Sciences, University of Texas at Dallas, 800 West Campbell Road, GR31, Richardson, TX 75080, USA. E-mail address: sherry.xin.li@utdallas.edu. Tel.: 1-972-8834987.

^{*} Business School, Beijing Normal University, Beijing 100875, China; E-mail address: summer.han@mail.bnu.edu.cn.

1 Introduction

Gary Becker's seminal work, *The Economics of Discrimination* (1957), has inspired numerous contributions to the economics literature on discrimination during the past six decades. Despite fruitful research in this literature (Riach and Rich, 2002; Pager, 2007; Rich, 2014; Bertrand and Duflo, 2017; Neumark *forthcoming*), one important area that has remained significantly understudied is labor market discrimination against family responsibilities or parenthood. That is, employers engage in statistical discrimination against workers of primary childbearing age, most likely women, out of the expectations that they may have higher likelihood of leaving their jobs due to childbearing and childcare responsibilities.

The main challenge in this literature is that family responsibilities or childbearing intentions are usually unobserved by employers at least in the early hiring stages. To bypass this difficulty, studies often take an indirect approach and use job applicants' other related characteristics (e.g., sexual orientation, age, and membership of parents' organizations) as proxies for their childbearing intentions and future childrearing responsibilities. Evidence in general supports the existence of discrimination against family responsibilities. For example, in France, single and childless women aged 25 are found to be less likely to receive job interviews than those who are 37 years old and have passed the primary child-bearing age (Duguet and Petit, 2005; Petit, 2007). Young heterosexual women receive lower interview callback rate than their homosexual counterparts in Belgium (Baert, 2014). In India, among job applicants of patrilineal origin, mothers are less likely to receive callback from employers compared to women or men without children. Mothers of matrilineal origin are not subject to such penalty (Bedi et al., 2018). In China, the (short- or long-term) unemployment history

reduces the chances of receiving job interviews for married women but not for single women (Maurer-Fazio and Wang, 2018). Hellester et al. (2016) find that in job ads with gender preferences younger women are preferred to younger men, and older men are preferred to older women. This “age-twist” in gender preferences is consistent with a potential discrimination against women for family responsibilities.¹ Studies also show that women are more vulnerable to parenthood penalty, relative to men (e.g., Anderson et al., 2002; 2003; Correll et al., 2007; Budig and Hodges, 2014; Gómez and Campos, 2014; Benard and Correll, 2015); men sometimes even enjoy a fatherhood premium due to a more responsible image that being a father may project (e.g., Glauber, 2008; Cooke, 2014).

Our study provides more direct evidence on statistical discrimination against family responsibilities by taking advantage of the heterogeneous treatments in China’s one-child policy before 2016 and its unexpected termination in 2016. The one-child policy, implemented between 1979 and 2015, was amended several times before its termination. The amendment in 2014 allowed two children for qualified families in which at least one parent was an only child himself/herself. This adjustment resulted in a higher expected childbearing possibility for the only-child workers, relative to their counterparts with a sibling(s) in 2014 and 2015 (see details in Section 2). On October 29, 2015, the government announced the shift toward a two-child policy and allowed *all* families to have up to two children, effective on January 1, 2016. The heterogeneities in the pre-2016 policy environment and the policy change in 2016 thus offer us a unique opportunity to study labor market equalities across men and women with different expected family responsibilities.

¹ There is a large literature on motherhood wage penalty and its determinants. See an excellent survey by Benard et al. (2008) and some recent studies Staff and Mortimer (2012), Budig and Hodges (2010, 2014), Killewald and Bearak (2014), Cooke (2014), Benard and Correll (2015), Jia et al. (2013), Yu and Xie (2014).

In this paper we use the same approach as in Agan and Starr (2017) that are the first to combine the field-experiment approach (i.e., correspondence study) with quasi-experimental analysis of policy changes (i.e., difference-in-difference) to draw *causal* inferences on labor market discrimination.² Specifically, we conducted a two-wave field experiment before and after the Chinese government unexpectedly ended the controversial one-child policy in January 2016. We sent about 9,000 fictitious resumes to real online job ads in three most economically advanced cities in China. On the four resumes sent to each job ad, the information on the fictitious applicants' gender and whether they were the only- or sibling-children was systematically varied. We found that *before* the policy change in 2016, the only-child women received a differentially lower callback rate than their sibling-child counterparts, relative to the male applicants. It suggested that the heterogeneous treatments in the pre-2016 policy which allowed the only-child adults to have up to two children adversely influenced the only-child women's labor market opportunities. We further applied a difference-in-difference approach and found that the disadvantage that the only-child women had faced disappeared after the birth policy was shifted to the two-child policy for *all* families in 2016. These findings showed strong evidence that women were negatively affected by the expected family responsibilities. We also found that compared to the younger women less than 25 years old, those women of age 25 or above were more likely to face the motherhood penalty due to their higher probability of maternity. Men, however, were not subject to this penalty, suggesting that the parenthood penalty was purely a female

² Agan and Starr (2017) conduct a two-wave correspondence study to investigate the impact of criminal records and race on individual's labor market performance before and after the policy change in 23 U.S. states that restricted employers from asking about applicants' criminal histories on job applications.

phenomenon. In addition, we found no evidence of the stereotype-based discrimination against the only-child workers.

Our study is an important extension to the literature on parenthood penalty. Compared to early studies that use *indirect* approaches based on groups (e.g., sexual orientation, age) with different probabilities of childbearing or childcare responsibilities, the unique contribution of this study is that we use two exogenous sources of variations – the differential treatments in the birth policy for the only- and sibling-child adults before 2016 *and* the unexpected termination of the one-child policy in 2016 – to draw *causal* inferences on statistical discrimination against family responsibilities. Moreover, we focus on how the *possibility* of having one or two children (rather than the fact of having children) may affect one’s labor market performance, which is an important but understudied research area in the previous literature (Bertrand and Duflo, 2017).³

This study is also the first to investigate potential discrimination against the only-child workers in the labor market, to the best of our knowledge. Since the 1960s the world has seen a sharp decline in birth rates in industrialized countries as well as many transitional and developing economies because of the increase in income, better education, and better access to family planning.⁴ As a result, only-child workers have become a non-negligible, continuously growing part of the labor force. Although they are often stereotyped as being spoiled, selfish, lonely, and maladjusted (Thompson, 1974; Falbo, 1979), little research has

³ “A topic of interest for future work would be to apply the correspondence method to measure the extent to which a bias exists against women with children, or against young women who may have children in the future.” (Bertrand and Duflo, 2017, p. 17)

⁴ In addition to China where the restrictive birth policy had been implemented since 1979, some developed countries have also experienced fast increase in the number of one-child families. In U.K. the number of families with one child increased by 16% from 3.1 million in 1996 to 3.6 million in 2014. In U.S. this number increased by 24% from 13.4 million in 1980 to 16.6 million in 2011, with single-child families making up nearly a quarter of the nation’s families.

been conducted on how they fare in the labor market relative to their counterparts with a sibling(s) primarily due to the methodological hurdle that one's only- or sibling-child status is in general unobservable by the employers in the early hiring process. With the largest number of only-child workers in its workforce, China's labor market makes an excellent testbed to study potential discrimination against them, and this study provides new empirical evidence.

The remainder of the paper is organized as follows. Section 2 provides a brief history of China's one-child policy and its amendments. We also discuss main issues related to the only children. Section 3 introduces the experimental design and procedure. In Section 4 we discuss the testing hypotheses and empirical strategies. The analysis and results are presented in Section 5. Section 6 concludes.

2 China's One-Child Policy and the Only Children

The one-child policy was implemented in China from 1979 to 2015. Figure 1 provides a brief timeline of this history. Before it was shifted to the two-child policy in January 2016 the one-child policy was amended several times, which resulted in differential treatments in the birth policy based on the parent(s)'s status of being an only-child or not.

The one-child policy was introduced in 1979. It required the Han ethnic majority, more than 90% of the nation's population, to limit to only one child per family.⁵ The

⁵ The one-child policy was ushered by multiple waves of family planning policies before 1979. For example, in 1975, the government started a family planning campaign under the slogan "late, long, and few" which advocated delaying marriage, increasing time gaps between births, and limiting family size. In 1978 the government introduced a voluntary program that urged couples to have no more than two children, preferably one. The start of the one-child policy was marked by an open letter issued by the Central Committee of the Communist Party to the public. The letter outlined the urgency of curbing China's population growth and set out a goal of capping the nation's total population at 1.2 billion by the end of the 20th century. The non-Han ethnic minorities were exempted from the one-child policy. See Liang (2014) and Wang et al. (2016) for more details on this history.

enforcement of the policy mostly relied on propaganda, incentives, and punishments. Those who delayed childbearing or received the “one-child certificate” were rewarded with subsidies, longer maternity leaves, better childcare or housing. The violators were subject to steep fines, demotion or loss of job if working in the public sectors. Forced sterilization or abortion was also used to enforce the policy.

Despite its efficacy on slowing down China’s population growth, the restrictive birth policy was subject to wide criticism for the brutality in its enforcement. It was also blamed for creating a skewed sex ratio of men to women. In 1984 only five years after the policy was enacted, it was amended to allow the second child for the families in rural areas if their firstborn was a girl and for the couples who were both only-children themselves.⁶ At the beginning of 2014, the same exemption was extended to include families in which one parent, rather than both, was an only child. These differential treatments resulted in higher expected family responsibilities for men and women who themselves were the only-children. Ironically, this seemingly preferential policy may exacerbate these only-child adults’ job perspectives and make them less preferable in the labor market compared to their counterparts with a sibling(s).

On October 29, 2015, the Communist Party leadership announced the termination of the one-child policy and a shift to the two-child policy, i.e., *all* married couples would be allowed to have two children starting on January 1, 2016. This drastic departure from the one-family-one-child core of the old policy was intended to address China’s severely skewed

⁶ This relaxation of the policy in rural areas was partly intended to accommodate a traditional preference for boys which has been more prominent in rural China, and partly due to the difficulty in enforcing the policy there. This exemption also applied to the couples who were *both* only child themselves to battle the so-called “4-2-1 problem”, i.e., at some point in his/her life, an adult only-child has to care for his/her two aging parents and four even older grandparents.

sex ratio which had resulted from the restrictive birth policy coupled with a traditional son preference in the Chinese culture. It was also driven by the nation's dreadful concerns over its vast aging population which, if continued to grow, would burden its health care system and cripple its hard-earned economic prosperity.

Many had expected that the extremely unpopular one-child policy would be ended eventually after it was gradually relaxed over time. The timing of its termination, however, was completely unanticipated by the general public especially because the policy had just gone through a major amendment merely two years ago in 2014. The unexpected timing can be seen by the search volume index on Baidu (www.baidu.com), the main Internet search engine and Google's counterpart in China, for the term "second child for all" (*quán miàn èr tāi*), a commonly used term by policy makers and public media to describe the final relaxation of the one-child policy before 2016. As shown in Appendix A, the Baidu search volume index for this term surged and reached its peak right after the summit of the Communist Party's policy-making Central Committee which was held from October 26 to 29, 2015.⁷ In sharp contrast, this index had been zero until February 2014, the month after the amendment that allowed two children if at least one parent was an only child. It had stayed at a negligible level of 3% of the peak volume between February 2014 and October 2015. This pattern, therefore, substantiated the unexpected timing of the policy change.

After the policy change, although government-issued birth permits are still required all families can now request to have two children. It was estimated that between 90 and 100 million couples in China have been affected (CNN, 2015). This radical change thus put an

⁷ The summit was known as the Fifth Plenum of the 18th Central Committee of the Communist Party of China. The formal announcement to shift the one-child policy to the two-child policy was made on the last day of the summit.

end to the almost four decades old, highly controversial one-child policy. It has also closed the gap in the expected family responsibilities between the only-child adults and those with a sibling(s) created by the policy amendment in 2014.

The 36-year implementation of the one-child policy has generated millions of only-children.⁸ Since they grow up without a biological or adopted siblings(s) and are constantly at the center of their families they are negatively stereotyped as being spoiled, selfish, lonely, and maladjusted (Thompson, 1974; Falbo, 1979). The high-frequency words associated with the only-children are mostly negative in news reports or media coverages (Bao, 2011; Feng, 2010). Some employers even included phrases such as “no single children” in their job recruiting advertisements (Hu, 2006; Chang, 2008; Cameron et al., 2013; Li, 2015).

Feng (2010) suggests, however, that the only-children are likely to be demonized. As a matter of fact, evidence from several-decade academic research is mixed and inconclusive. Some research shows that in childhood, only-children on average have better physical health and higher intelligence level but inferior personality traits and behavioral weaknesses, e.g., lack of independence and willingness to share (Xu, 1980; Xiao et al., 1981; Chen et al., 1985; Huang, 1994). A recent study by Cameron et al. (2013) reports that the only-children are less trusting and trustworthy, more risk-averse, more pessimistic, and lack the sense of responsibility. Other research, however, finds no or very small differences in their personalities and behaviors (e.g., Falbo and Polit 1986; Bai, 1992; Jing, 1997; Fan and Wang, 2001; Zhang et al., 2007; Huang and Wen, 2008), and these differences tend to shrink or

⁸ According to China’s Inter-Census Population Survey in 2005, the only-children to sibling-children ratio was 0.34 for the then 10-to-19 age group, i.e., the 20-to-29 age group in 2015. This ratio was 0.42 and 0.26 for men and women, respectively. The existence of a large number of sibling-children despite the one-child policy may be partially due to the challenges in the policy enforcement especially in the rural areas and partially due to the relaxations of the policy that allowed the second child for the qualified families. It is estimated that 61% of China’s population resided in the rural areas in 1980. It declined to 57% in 1990, 53% in 2000, and 49% in 2010 (World Bank Group, 2014).

disappear with age (Feng, 2000, 2006). While the debates continue little is known on how the only-child workers fare in the labor market relative to their counterparts with a sibling(s).

3 Experimental Design

We conducted a two-wave correspondence study *before* and *after* China shifted the one-child policy to the two-child policy on January 1, 2016. In the experiment, fictitious resumes were sent to real, online job ads in three major cities. In each set of four resumes for a job ad, we systematically varied the information on the fictitious applicant's gender and whether the applicant was the only child or had a sibling(s).

When our experiment started in May 2015 we, as the rest of the general public, did not anticipate the policy change. We initially planned to use the heterogeneous treatments for the only- and sibling-child adults in the pre-2016 policy environment to identify potential discrimination against parenthood. Luckily, the serendipitous timing of the policy change offered us another unique source of variation to obtain an even cleaner estimate on such discrimination. We will introduce the experimental design in this section and discuss the empirical strategies in detail in Section 4.

3.1 Timeline and Locations of the Experiment

As shown in the timeline in Figure 1, the first wave of our experiment was carried out from May to September in 2015 and the second wave from January to July in 2016. The study was conducted in China's three most economically advanced cities, Shanghai, Guangzhou, and Shenzhen. Shanghai, a global business and financial center, is China's most populous city with a population of over 24 million. Guangzhou (also known as Canton to westerners) is the capital city of Guangdong Province in the southeast and the third largest city behind Beijing

and Shanghai. It is a mighty manufacturing hub of clothing, electronics, plastic goods, and toys and has a population of more than 13 million. Shenzhen, another major city in Guangdong Province immediately adjacent to Hong Kong, is a manufacturing center for high technology products and has a population of more than 18 million.

The first wave of the experiment was also conducted in Beijing. However, halfway through the experiment in Beijing, the Chinese government announced on October 29, 2015, to end the one-child policy on January 1, 2016. Although we completed the data collection as initially planned, this *unexpected* policy change rendered the pre-policy sample in Beijing too small to be analyzed.⁹ Beijing is thus excluded from this paper.

The timing of our two-wave study is unique and crucial. Recall that families were allowed to have two children in 2014 and 2015 if *one* parent, rather than *both*, was an only child. Starting in January 2016, however, all couples are allowed to have two children for the first time in more than three decades. Therefore, the elimination of the one-child policy in 2016 should not influence the childbearing possibility of men and women who were the only children themselves because they have already been allowed to have up to two children since 2014. This policy change in 2016 should only have a direct impact on those adults who had a sibling(s).¹⁰ This unique timing renders our experiment a clean test on discrimination driven by possible childbearing cost, as will be discussed in Section 4.

⁹ We used the announcement date of the policy change as the cutoff rather than the actual implementation date since we expected that the mere announcement of the termination of the four-decade-long one-child policy would be sufficient to affect employers' perceptions on job applicants' family responsibilities. The data collection was not subject to any ambiguity in Shanghai, Guangzhou, and Shenzhen.

¹⁰ Strictly speaking, this policy change should only affect those who had a sibling(s) *and* were married to someone else who also had a sibling(s). In other words, a person who had a sibling(s) but the spouse was the only child would be allowed to have up to two children since 2014, hence would not be affected by the policy change in 2016. In our experiment, however, marital status was not revealed on the fictitious resumes because this information was not required on the resume template of the job board and was not commonly revealed on the real resumes. Therefore, we assume that in the resume review process, the fictitious applicants are treated based on their *own* only- or sibling-child status and the associated possibility of having the second child.

3.2 *Job Advertisements*

Our study was implemented on 51job.com (mimicking “I-want-job.com” in the Chinese pronunciation; Nasdaq: JOBS), a leading job ads website in China. According to the information published on their website, 51job.com has over 100 million registered users and an enormous database of 96 million resumes. Over 3.2 million job ads are posted online, and about 38 million applications are delivered to prospective employers every week.¹¹

We focused on the two industries that had the most number of job listings during the two-month period before our experiment in each city. These industries were internet and finance in Shanghai, internet and fast moving consumer goods in Guangzhou, internet and electronics in Shenzhen. Among all the occupations listed for these industries, we chose the three most popular job categories – sales, administrative assistant, and customer service.¹² We excluded those positions that were to be filled immediately, required photographs, required lower than an associate degree or higher than a bachelor degree, or targeted fresh college graduates or applicants with more than five years of work experience.¹³

3.3 *Fictitious Applicants and Design of Resumes*

Our fictitious job applicants were aged between 22 and 29. This age range met the 20-year-old marriage age requirement and fell in the primary childbearing time window. It also

¹¹ Source: http://www.51job.com/bo/AboutUs_e.php (accessed on June 15, 2016.)

¹² These occupation categories are also used in previous audit studies such as Bertrand and Mullainathan (2004) and Kroft et al. (2013).

¹³ We excluded the job ads that were to be filled immediately since our pilot study showed that the employers in these cases often offered interviews to all the resumes submitted. We also excluded the job ads that required applicants' photographs because most real resumes submitted on 51job.com did not provide photographs around the time of our experiment. The exclusion based on education and work experience was added since 68% of the publicly listed ads on 51job.com required college education. In addition, 82% needed some work experience, among which 76% preferred work experience below five years.

allowed us to focus on job applicants with college education and some but less than five years of work experience as stipulated in the majority of the job ads.

Many Chinese online job boards including 51job.com provide applicants with resume templates. Besides standard information such as name, date of birth, education background, and work experience, applicants are also required to report gender and complete a brief self-assessment. It allows us to implement a 2×2 within-subject factorial design by manipulating the fictitious applicants' gender and status of being an only child or having a sibling(s) on the resumes. This leads to four fictitious resumes for every job ad: female-only-child (FO), male-only-child (MO), female-with-sibling(s) (FS), and male-with-sibling(s) (MS). Our design, summarized in Table 1, was implemented *before* the government announced to end the one-child policy in October 2015 and *after* the policy ended on January 1, 2016.

[Table 1 about here]

Gender was explicitly specified on each fictitious resume as required by 51job.com. The fictitious name consisted of two Chinese characters – a common last name (Wu and Yang, 2014) and a popular male- or female-sounding first name. We avoided using gender-neutral first names to make the applicants' gender salient since as discussed in Section 1, men and women may be affected differently by their family responsibilities.

The information on whether a fictitious applicant was an only- or sibling-child was included and manipulated in the self-assessment. Specifically, the assessment started with a sentence “[a]s the only child (or one of the children) of my family ...,” followed by a four-sentence gender-neutral personal statement on the applicant's personality traits (e.g.,

responsible, confident, motivated, eloquent, team-player, etc.)¹⁴ We adopted these keywords or phrases from real job resumes on 51job.com. In the real-life case scenarios it is not unusual to reveal one's only- or sibling-child status and portray it in a positive light on the resume. In fact, among the real resumes publicly posted on 51job.com and updated during the six months before our experiment, about 7% explicitly included this information.^{15, 16}

As explained above, we focused on the job ads that required college education and between 0 (exclusive) and five (inclusive) years of work experience, and the fictitious resumes were designed accordingly. For the education background, two local and two non-local public universities were used as the graduating schools on the four fictitious resumes. The fictitious applicants held either a three-year associate degree or a four-year bachelor degree depending on the requirements in the job ads. Date of birth and hence age was specified to match the number of years of education and work experience. The possible college majors included history, economics, accounting, finance, business administration, marketing, international economics, and trade which were the employers' most preferred college majors for the occupations that we targeted in this experiment. To make our fictitious applications appear authentic, fictitious applicants' work experiences and job skills were

¹⁴ These qualities and personality traits are in general perceived positively in the Chinese culture. On the resumes for the sibling-child applicants, no information was given on the number of siblings or the applicants' birth order.

¹⁵ Among the real resumes that revealed the only- or sibling-child information, 62% were done by the only-child applicants. Both the only- and sibling-child cases were presented positively. We also conducted a manipulation check to confirm the salience of fictitious applicants' only- or sibling-child status. Details are in Appendix B.

¹⁶ Since only 7% of the real resumes on 51job.com mentioned one's only- or sibling-child status, one may ask whether the mere fact of revealing this information on the resumes would affect the likelihood of receiving callbacks. We investigated this question by conducting an additional experiment using a separate sample. Details are provided in Appendix C. We find that revealing one's only- or sibling-child status on the resumes does not influence the likelihood of receiving callbacks.

adapted from real resumes that had been used for the corresponding industries and occupations.¹⁷

Overall, the four fictitious resumes represented four fake job applicants who had unique identities and contact information, differed in gender and only- or sibling-child status, but otherwise had comparable job credentials.

3.4 Procedure

For each randomly selected job ad that was newly published within one day and met our selection criteria discussed above, four fictitious resumes were sent in a random order, one on each day within four consecutive days. Following Bertrand and Mullainathan (2004), we defined the success of an application as receiving a callback from the recruiter within two weeks of resume submission. A callback could be a phone call, email or text message from the recruiter requesting an interview with the fictitious applicant. In our experiment, more than 98% of callbacks, if received at all, were made within two weeks, and 97% within ten days of resume submission. Our research assistants answered the callbacks and told the recruiters that the (fictitious) applicants were no longer available for their job openings.

For each job ad, we collected information on job requirements (e.g., education, work duration, skills, etc.), the recruiting firm's size (i.e., the number of employees), type of ownership, and the number of people who followed this firm on 51job.com. Some of this information will be used as covariates to control for the job's or the firm's specific characteristics in the empirical analysis.

4 Identification Strategies and Hypotheses

¹⁷ To minimize the chance that our fictitious resumes may coincide with some real applications, we adapted the work experiences and job skills from some real resumes that were active *before* 2015.

The primary goal of this study is to investigate potential statistical discrimination against expected family responsibilities, especially for women. As explained in Section 2, however, in 2014 and 2015 this type of discrimination was intertwined with stereotype-based discrimination. That is, compared to the sibling-child women, the only-child women may be treated unfavorably in the labor market because of the negative stereotypes against them *or/and* their potential higher childbearing possibility since they were allowed to have two children starting in January 2014.¹⁸ Therefore, the challenge that we faced when designing this experiment in 2014 was how to parse these two aspects out and identify separately the statistical discrimination against expected family responsibilities.

Since we did not anticipate the termination of the one-child policy our initial identification strategy was to use the male applicants as the baseline group and the females as the treatment group in the old policy environment. Recall that the early literature shows that men are not discriminated based on family responsibilities (e.g., Anderson et al., 2002; 2003; Correll et al., 2007; Budig and Hodges, 2014; Benard and Correll, 2015). In addition, no evidence is found on gender difference in the stereotype-based discrimination against the only-children.¹⁹ Under these assumptions, we could apply a difference-in-difference approach to the pre-2016 data across gender groups to obtain a clean measure on discrimination against expected family responsibilities for women. This leads to Hypothesis 1.

Hypothesis 1 (Discrimination against family responsibilities: pre policy change)

¹⁸ A sibling-child woman could have up to two children in 2014 and 2015 if her husband was an only child. An only-child woman, however, could have up to two children for sure because of her own only-child status.

¹⁹ This assumption will be substantiated by Result 2 based on the analysis on the post-policy callback rates.

The only-child women receive differentially lower callback rate than the sibling-child women, relative to men, before the policy change in 2016. That is, $(FOB - FSB) < (MOB - MSB)$.

The unanticipated policy change in 2016 and its serendipitous timing offered us a cleaner strategy to identify statistical discrimination against family responsibilities. It is reasonable to believe that the stereotype-based discrimination against the only-children was not affected by the termination of the one-child policy in 2016. Under this assumption, we could obtain the differences in callback between the only- and sibling-child applicants, conditional on gender, before and after the policy change and then apply a difference-in-difference approach to eliminate any potential confounds related to negative stereotypes against the only-child applicants. In addition, this difference-in-difference measure, conditional on gender, can get rid of other potential differences between the only- and sibling-child applicants that are unaffected by the change of birth policy. For example, in the traditional Chinese culture, boys are preferred to girls, and this son preference is more prominent in poor households from rural areas or in rural migrant families (Qian, 2008; Chew et al., 2017). Rural families are more likely to have an additional child, either legally with government approval or illegally, if the firstborn is a girl. Therefore, sibling-child status for women may be perceived in the labor market that they are more likely to be (originally) from the rural areas (thereby facing greater intra-household resource competition with their sibling(s) and being more disadvantaged in education), or they are more likely to be the firstborn of the family (thereby being more responsible and a natural leader). Since these perceptions are unlikely to be affected, at least immediately, by the change of China's birth policy in 2016, the difference-in-difference approach discussed above could also eliminate

these confounds and hence serve as a clean estimate on the discrimination against family responsibilities. We hypothesize that women, but not men, are subject to the discrimination against family responsibilities.

Hypothesis 2 (Discrimination against family responsibilities: policy change in 2016)

2a) (Women) *The only-child women receive a differentially lower callback rate than their sibling-child counterparts before relative to after the termination of the one-child policy in January 2016. That is, $(FOB - FSB) < (FOA - FSA)$.*

2b) (Men) *The gaps in callback between the only-child and sibling-child men are not different before and after the termination of the one-child policy in January 2016. That is, $(MOB - MSB) = (MOA - MSA)$.*

As discussed in Section 2, although the debates on the comparisons of labor market performance between the only- and sibling-child groups are inconclusive in the academic research, the only-children are widely subject to negative stereotyping in personalities and social skills in Chinese society. If the recruiters use this negative stereotyping to evaluate the applications we expect to observe discrimination against the only-child job applicants.

Fortunately, the policy change in January 2016 also provides an excellent opportunity for us to test this type of discrimination.²⁰ Because both the only- and sibling-child applicants have been allowed to have up to two children since the termination of the one-child policy in January 2016, the only difference between them, conditional on gender, is their only- or sibling-child status. Therefore, the comparison in the callback between these two categories

²⁰ Stereotypes-related discrimination against only-child workers can be taste-based if the employers dislike hiring them due to their (alleged) socially undesirable personalities. It can also relate to statistical discrimination if such personalities as the lack of independence or responsibility (Huang, 1994; Cameron et al., 2013) impede cooperation with coworkers and obstruct team productivity. Differentiating these two channels, however, is beyond the scope of this study.

of applicants within each gender group could yield a clean test on the stereotype-based discrimination against the only-child job applicants. We hypothesize that the only-child applicants are subject to the stereotype-based discrimination and are less likely to receive callbacks compared to their sibling-child, same-sex counterparts. In addition, we hypothesize that this stereotype-based discrimination is equally likely to occur for men and women. Note that this was an assumption that we used to derive Hypothesis 1. The policy change in 2016 allows us to test it directly. These discussions lead to the third hypothesis.

Hypothesis 3 (Stereotype-based discrimination against only-child applicants)

3a) *After the termination of the one-child policy in January 2016, the callback rate in each gender group is lower for the only-child applicants than for their sibling-child counterparts due to the negative stereotypes against the former, that is, $FOA < FSA$, $MOA < MSA$.*

3b) *The degree of the stereotype-based discrimination is the same across the gender groups, that is, $FOA - FSA = MOA - MSA$.*

5 Empirical Analysis and Results

In this section, we test the hypotheses introduced in Section 4 and investigate the determinants of the interview callbacks.

5.1 Descriptive Statistics

The overall average callback rate in our study is 32.3% for the 8,848 resumes sent to the 2,212 job ads.²¹ The average callback rate is 31.8% before the policy change and 32.8%

²¹ Our callback rate is higher than some earlier correspondence studies conducted on China's online job boards, e.g., Maurer-Fazio (2012) and Maurer-Fazio et al. (2015). The higher callback rate may be due to the fact that the work experiences on the fictitious resumes in our experiment were tailored to better fit the job positions. It could also be attributed to a tighter labor market. To verify this conjecture, we calculated the expected recruitment rate (9.3%) using the total number of applications received and the intended number of recruits for the 1,562 job ads in our sample for which such information was publicly available on the job board. The average callback rate did not seem unreasonable for such expected recruitment rate.

afterwards, and the difference is statistically insignificant ($p = 0.340$, test of proportions). Table 2 presents the summary statistics on callback for each category of fictitious applications (upper-left panel), the differences in callback across categories conditional on the policy environment (McNemar's Chi-square test for the paired binomial dataset, upper-right panel), the comparisons before and after the policy change (unpaired test of proportions, bottom-left panel), and the difference-in-difference comparisons (OLS with the standard errors clustered at the job ad level, bottom-right panel).

[Table 2 about here]

Before the policy change, the callback rates are 30.8%, 31.0%, 35.4%, and 30.1% for men with a sibling(s) or without a sibling(s) and their female counterparts, respectively. After the policy change, the callback rates are 30.2%, 29.9%, 34.4%, and 36.6% for these four categories of job applicants, respectively. When the comparisons are made conditional on gender, we find no significant differences in callback between the only- and sibling-child men regardless of policy environment ($p > 0.10$, McNemar's test); no difference in callback is found before and after the policy change for either only- or sibling-child men ($p > 0.10$, unpaired test of proportions). In contrast, several main observations emerge for women. Under the old policy, the callback rate 30.1% for only-child women is significantly lower than 35.4% for sibling-child women (FOB vs. FSB, $p < 0.001$, McNemar's test). The callback rate for only-child women increases significantly from 30.1% to 36.6% after the policy change (FOB vs. FOA, $p = 0.001$, unpaired test of proportions). Under the new policy, however, the preference between the only- and sibling-child women is reversed and the

difference in callback is marginally significant (36.6% for FOA v. 34.4% for FSA, $p = 0.063$, McNemar's test).²²

5.2 Regression Results

Regression analyses on the determinants of callback are reported in Table 3 with coefficients of OLS in Columns [1]-[3] and marginal effects of Probit in Columns [4]-[6]. The dependent variable is whether the fictitious applicant receives a callback from the recruiter for a job interview. The independent variables in Columns [1] and [4] include the dummy variables for the experimental treatments (with the female only-child group before the policy change FOB being omitted) and the city fixed effects. Columns [2] and [5] further add other characteristics of the applicants, e.g., age, education, years of work experience, the number of previous jobs before the application, whether the academic degree is granted by a local university, college major, and university fixed effects. The job- and firm-related characteristics are added in Columns [3] and [6]. They include occupations, the number of applications for a job ad (as a proxy of the competitiveness of this job position), the number of people who followed this firm on the job board (as a proxy of the popularity of this firm), firm size (a categorical variable for the number of the firm's employees specified in each job ad), the type of firm's ownership, and the type of industry to which the firm belonged. In all the analysis standard errors are clustered at the job ad level. Table 4 provides the Wald tests for the equality of estimated coefficients in OLS (or marginal effects in Probit) of Table 3. Panel I of Table 4 contains the pairwise comparisons of the coefficients (or marginal effects)

²² In this study we refrain from comparing the callback rates across the gender line and drawing inferences on gender discrimination above and beyond the gender difference in statistical discrimination against the family responsibilities. We will return for more discussions on potential gender segregations in the occupations in our experiment at the end of Section 5.

across treatments, Panel II the difference-in-difference (DD) analysis, and Panel III the difference-in-difference-in-difference (DDD) analysis.

[Tables 3 and 4 about here]

Since the OLS and Probit results are very similar to the non-parametric analysis in Table 2, our discussions will focus on the OLS results reported in Columns [3] of Tables 3 and 4 which contain the complete set of covariates.

We first investigate the possible statistical discrimination against family responsibilities by comparing the differences in callback between the only- and sibling-child applicants across the gender line before the 2016 policy change. We find that the callback rate is 5.24 percentage points lower for the only-child women (FOB) than for the sibling-child women (FSB) ($p < 0.001$ in Column [3] of Table 4). This difference is significantly *lower* than the difference of 0.09 percentage points between the only- and sibling-child men ((FOB – FSB) – (MOB – MSB) = -5.34%, $p = 0.005$). This finding, in support of Hypothesis 1, indicates that when being allowed to have more than one child under the old policy in 2014 and 2015, only-child women receive a differentially *lower* rate of callback due to their potential family responsibilities compared to their male counterparts. This serves as the first piece evidence of discrimination against family responsibilities for women.

For further investigation on discrimination against family responsibilities, we take advantage of the unexpected policy change in 2016. Recall from the discussions in Section 4, a clean measure for this type of discrimination can be obtained by applying a difference-in-difference approach, conditional on gender, to the callback rates between the only- and sibling-child applicants before and after the policy change. According to Table 2, the callback rate is 30.1% for the only-child women (FOB) and 35.4% for the sibling-child

women (FSB) before the policy change. After the policy change, the callback rate becomes 36.6% and 34.4% for these two groups, respectively. The OLS analysis in Tables 3 and 4 show that the difference between FOB and FSB -5.24%, net of the post-policy difference 2.24% between FOA and FSA, yields a difference-in-difference estimate of -7.49% ((FOB – FSB) – (FOA – FSA), $p < 0.001$ in Column [3] of Table 4). This finding supports (FOB – FSB) < (FOA – FSA) in Hypothesis 2a). This result serves as the second and cleaner evidence of the discrimination against family responsibilities for *women* since the difference-in-difference approach eliminates any potential confound from stereotype-related discrimination against the only-child applicants as well as any other unobservable differences between the only-child and sibling-child applicants that are invariant before and after the policy change. Similarly, we apply this approach to men. The 0.09% difference between the only-child and sibling-child men before the policy change ($p = 0.943$ for MOB – MSB in Column [3] of Table 4), combined with the -0.37% difference after the policy change (MOA – MSA, $p = 0.772$), yields a difference-in-difference of 0.46% ((MOB – MSB) – (MOA – MSA), $p = 0.789$) in the callback rate for men.²³ This finding supports Hypothesis 2b) and indicates no discrimination based on family responsibilities against *men*.

The discussions above lead to Result 1.

Result 1 (Discrimination based on family responsibilities): *Women but not men are negatively affected by the expected family responsibilities.*

Result 1 shows that the parenthood penalty is a female phenomenon. As discussed at the beginning of this section, evidence indicated a fairly tight labor market in the selected

²³ It is worth noting that the discrimination against family responsibilities is significantly more severe for women than for men because (FOB – FSB) – (FOA – FSA) < (MOB – MSB) – (MOA – MSA) ($p = 0.002$), as shown in the difference-in-difference-in-difference estimate in the bottom row of Table 4.

industries when our experiment was conducted. In contrast, in a slack labor market, competitions may heighten due to the limited number of jobs, and the employers may be more selective, which may cause more severe discrimination against women with family responsibilities. Therefore, it is possible that our estimate on the motherhood penalty is a conservative estimate compared to that in a slack labor market.

Next, we investigate Hypothesis 3. As discussed in Section 4, to obtain evidence of the stereotype-based discrimination against only-child job applicants, we compare their callback to their sibling-child counterpart *within* each gender group *after* the one-child policy ended in 2016. We find that the callback rate is -0.37% lower for only-child men than for sibling-child men after the policy change (MOA – MSA, $p = 0.772$ in Column [3] of Table 4) whereas it is 2.24% higher for only-child women than for sibling-child women (FOA – FSA, $p = 0.077$). These observations about men and women suggest that the only-child status and the associated negative stereotypes do not adversely affect one’s opportunity of receiving an interview callback relative to the same gender, sibling-child applicants. Therefore, we fail to find supportive evidence for Hypothesis 3a). Moreover, the gap in the callback rate between the only- and sibling-child applicants is not significantly different across the gender line ((FOA – FSA) – (MOA – MSA) = 2.61%, $p = 0.160$ in Column [3] of Table 4), which supports Hypothesis 3b). It is worth noting that this result substantiates the assumption we used to derive Hypothesis 1. These findings are summarized in Result 2.

Result 2 (Stereotype-based discrimination against only-child applicants): *For both men and women, we find no evidence of discrimination against only-child applicants based on the negative stereotypes. In addition, no gender difference is found in the gaps of the callback rates between the only- and sibling-child applicants.*

In our investigation on the stereotype-based discrimination above, we find that after the policy change, the callback rate is marginally higher for only-child women than for sibling-child women (FOA – FSA = 2.24%, $p = 0.077$ in Column [3] of Table 4) while it is similar between only-child and sibling-child men (MOA – MSA = -0.37%, $p = 0.772$). Since the 2016 policy change has eliminated the differential child-birth treatments for the only-child and sibling-child women the 2.24% difference in moderate favor of the *only-child women* is surprising at the first glance. One possible explanation is the short-term behavioral response of the employers in reaction to the unexpected policy shock. On the supply side of the policy, the termination of the one-child policy has *only* affected the sibling-child applicants (because they were not allowed to have the second child until the policy change in 2016) and released their long suppressed demand to have more children. On the demand side, some sibling-child women may take immediate actions by either starting a family or having the second child in response to this new policy, or at least are expected to do so. Therefore, the policy change in 2016 has resulted in a sudden increase in the sibling-child women's child-bearing possibilities, relative to the only-child women who had already been exempted from the one-child restriction as early as in 2014 and hence were *not* affected by the 2016 policy change.²⁴ Consequently, we would expect that the employers respond by substituting

²⁴ The evidence in the literature generally supports greater fertility desire of sibling-child women than that of only-child women in China before and after the policy change in 2016. Chen and Deng (2007) study the factors that influence Chinese women's desired births by reanalyzing the data from the 2004 China Health and Nutrition Survey. They find that the number of desired births is positively associated with the number of siblings a woman has. In 2013 China's National Health and Family Planning Commission conducted a large scale survey among 63,451 households in 29 provinces except Tibet and Xinjiang Uygur Autonomous Regions. Their statistics showed the average preferred number of children was 1.79, 1.83, and 1.95 for the three groups of couples – both being only-children, one being an only-child, or neither being an only-child, respectively (Zhuang et al., 2014). In April 2016 after the change to the two-child policy, Jin et al. (2016) conducted a survey among more than 3,000 married women in urban areas of six provinces who were between 20 and 49 years old and already had one child. They find that similar to Chen and Deng (2007), the number of siblings of the mother has a significant positive impact on her desirable family size. These pieces of evidence is in line with the employers' expectations on the higher childbearing possibilities of the sibling-child working women, relative to the only-child women, as a result of the 2016 policy change.

the only-child women for the sibling-child ones in order to alleviate the sudden increase in the uncertainty associated with the latter group's family planning.²⁵

If our conjecture above is correct we would also expect that such substitution effect to be more pronounced for the slightly older women since the urgency to begin motherhood may increase with women's age given the fact that female fertility starts to decline in early 30s. Although many women attend college, especially in urban areas, and marry later than traditionally, social norm remains and younger brides are preferred in the modern Chinese society. Unmarried women after their mid-twenties are unpleasantly called "leftover women" or ("leftover ladies") and are subject to enormous social pressure to find spouses. Since fictitious applicants' age spans from 22 to 29 in our study we split the *female* group into two age subsamples – younger than 25 versus 25 or older – and compare the impact of the policy change on the callback rates for the sibling- and only-child women within each age range. For completeness, we conduct the same exercise for men. The callback rates by gender and age range are reported in Table 5. We find that before the policy change, the difference in callback between the only- and sibling-child women is -5.7% in the younger age group (FOB – FSB, $p = 0.008$, McNemar's test in Table 5A), similar to the -5.0% difference in the older group (FOB – FSB, $p = 0.003$, McNemar's test in Table 5B). Therefore, for both the younger and the older age groups, the only-child women are significantly less favored by the employers, relative to their same-sex sibling-child counterparts in the pre-2016 policy environment because the former was exempted from the one-child restriction and hence had higher childbearing possibility before 2016.

[Table 5 about here]

²⁵ Note that this substitution did not occur across the gender line, i.e., through an increase in callback for men and a decrease in callback for women. More discussions will follow after Result 3 below.

However, the gap in callback between the sibling- and only-child women exhibits different patterns for the two age groups after the policy change. On the one hand, the gap shrinks from -5.7% (FOB – FSB, $p = 0.008$) to merely 0.8% (FOA – FSA, $p = 0.596$, Table 5A) for the younger subgroup, suggesting that as expected, the policy change has closed the gap in callback between the sibling- and only-child, *younger* women. Since many professional women in the urban areas do not start a family until after 25 years old regardless, the employers may expect that the termination of the one-child policy in 2016 is unlikely to have an immediate, noticeable impact on family planning by the sibling-child, *younger* women. On the other hand, for the cohort of age 25 or above, the initial pre-policy gap of -5.0% (FOB – FSB, $p = 0.003$) in favor of the sibling-child women has been *reversed* to a sizable post-policy gap of 5.7% in favor of the *only-child* women (FOA – FSA, $p = 0.009$, Table 5B). This reversal could occur if the sibling-child women of age 25 or above decide to expedite their childbearing plan (or are believed to do so by the employers), either immediately or in the foreseeable future, in response to the sudden removal of the one-child restriction.²⁶ This would make these older sibling-child women less desirable employees compared to their same-age only-child counterparts who are not subject to the unexpected policy shock. The *reversal* in employers' preferences for the only-child, *older* women is in sharp contrast to the *convergence* in their preferences for the sibling- and only-child, *younger* women after the policy change. It lends support to our conjecture on employers' short-term

²⁶ Consider a sibling-child woman of age 25 or older who had a preference for two children but was constrained by the one-child policy before 2016. If she has had one child she may respond to the removal of the one-child restriction by considering to have the second child soon before her biological clock is running out of time. If she has had no children she may want to expedite her childbearing plan to have the first child sooner in order to have two children within her biological time window.

behavioral response to the policy change and shows that this response is driven by the substitution between the only-child and sibling-child women in the *older* age cohort.

As for men, no significant difference in callback is found between the only-child and sibling-child men before and after the policy change, and this applies to both the younger and older age cohorts ($p = 0.447$ for men younger than 25; 0.219 for men of age 25 or above, difference-in-difference estimates in the lower-right panels of Tables 5A and 5B).

These findings in Table 5 provide the third piece of evidence that further corroborates Result 1 that women but not men are negatively affected by the expected family responsibilities. They also extend Result 1 and indicate that women's higher probability of maternity, reflected by age, may exacerbate statistical discrimination against them because of their expected family responsibilities. These discussions lead to Result 3.

Result 3 (Motherhood penalty and age): Statistical discrimination against family responsibilities for women worsens as their probability of maternity increases with age.

One observation remains puzzling in our analysis above. That is, despite the average increase in childbearing possibilities by women due to the universal removal of one-child restriction in 2016, the overall callback rate did not increase for men (30.9% before the policy change and 30.1% afterwards, $p = 0.873$, OLS with standard errors clustered at the job ad level) but it increased moderately for women from 32.7% to 35.5% after the policy change ($p = 0.125$). It suggests possible gender segregation in the occupations used in our study. In Tables 6A-6C, therefore, we report the average callback rates by occupation, i.e., sales, administrative assistance, and customer service. We find that although men and women in sales receive comparable callback (Table 6A) women receive a much higher callback rate

than do men in administrative assistance and customer service (Tables 6B-6C) under both the old and new policy environments.²⁷ Note that these occupations are not high-end or high-salary jobs. Instead, they are often considered as jobs with decent pay and high stability. The employers may prefer women for these occupations to minimize employee turnover since women may (be perceived to) value job stability more than the pay, relative to men. Hence, this higher tendency to recruit female applicants (at least in the stage of callback) in our study should not be interpreted as reverse gender discrimination in favor of women in China's labor market.²⁸

Most importantly, we find that despite the potential caveat of gender segregation, our main results hold across occupations in Tables 6A-6C. The difference-in-difference estimates between the only- and sibling-child *men* before and after the policy change are all small in size and statistically insignificant ($(MOA - MSA) - (MOB - MSB) = -0.4\%$, 1.4% , and -1.5% for sales, administrative assistance, and customer service, $p > 0.10$ in all cases). These estimates for *women*, however, are all economically sizable and statistically significant for two out of the three occupations ($(FOA - FSA) - (FOB - FSB) = 8.3\%$, $p < 0.001$ for sales; 3.0% , $p = 0.556$ for administrative assistance, maybe due to the lack of statistical power; 7.9% , $p = 0.043$ for customer service). These observations further corroborate our main result, that is, women, but not men, are negatively affected by their family responsibilities.

²⁷ In sales, the average callback rates are 37.8% and 36.8% for men and women before the policy change ($p = 0.666$, OLS with standard errors clustered at the job ad level), and 36.2% and 39.1% after the policy change ($p = 0.016$). In administrative assistance, the average callback rates are 5.6% and 14.7% for men and women before the policy change ($p < 0.001$), and 7.1% and 16.2% after the policy change ($p < 0.001$). In customer service, the average callback rates are 25.3% and 30.9% for men and women before the policy change ($p = 0.010$), and 25.3% and 35.6% after the policy change ($p < 0.001$).

²⁸ Our research should be differentiated from other recent studies that focus on gender discrimination but fail to find discrimination against women in China's labor market, for example, correspondence studies by Zhou et al. (2013) and Maurer-Fazio and Lei (2015), and an empirical study by Kuhn and Shen (2013) based on a large sample of job ads downloaded from an online job board.

6 Conclusion

We conducted a two-wave resume correspondence study in China before and after the government shifted the 36-year-long unpopular one-child policy to the two-child policy in January 2016. On about 9,000 fictitious resumes sent to real online job ads, the information on applicants' gender and whether they were an only- or sibling-child was systematically varied. Using the differential treatments in the birth policy before 2016 and the policy change in 2016 as two exogenous sources of variation in the analysis, we find strong evidence on statistical discrimination against expected family responsibilities. Such discrimination applies only for women, and it worsens as their probability of maternity increases with age.

The shift to the two-child policy in 2016 was part of Chinese government's effort to combat declining births, shrinking workforce, and aging population. Unfortunately, relaxing the birth limits to two children per family has failed to boost the number of new births to the intended level during the past few years. More aggressive measures, therefore, may be taken by the government to entirely abolish the population-control policy. Ongoing discussions on new policies include proposals to replace the current two-child policy with "independent fertility" as early as in 2019, giving back to families their freedom to decide how many children to have (Bloomberg News, 2018a, par. 3). Our study, however, provides an important caveat to possible new policies that primarily rely on loosening the birth limits. The past few decades have witnessed the declining fertility desire by married couples, particularly in urban areas of China, mainly due to factors such as skyrocketing childcare, education, and medical costs. Labor market discrimination against women for their (expected) family responsibilities also joins these factors to stifle their fertility desire. As a matter of fact, many are skeptical about whether merely relaxing the birth limits can successfully stimulate the

fertility desire since as shown in this study, such policy would further increase childbearing probabilities by female workers and exacerbate statistical discrimination against them by employers.

Our findings on statistical discrimination against working women are confirmed in recent media coverage. For example, “married women without children and women in prime child-bearing years were much more likely to face discrimination than their peers” (Bloomberg News, 2018b, par. 6). Therefore, a direct implication of our study is that new policies may not successfully reverse the declining birth rate without alleviating the labor market discrimination against women for their (expected) family responsibilities. One crucial step is to actively enforce the anti-discrimination laws, as successfully being done in Sweden in 1930s, so that women of childbearing age will truly have equal employment opportunities and equal pay.²⁹ Policy makers may also need to reevaluate the policies that were intended to help working women but unfortunately ended up hurting them. An example was that the expanded parental-leave policies launched in 30 provinces in 2017 caused even more discrimination against women of child-bearing age as employers shunned hiring or promoting them. Complementary measures for these well-intended policies may include reducing employers’ costs to hire or retain these working women. Chinese government may also follow other countries’ (e.g., Australia, Canada, France, Russia, and Singapore) examples by offering financial incentives (e.g., stipends, tax breaks, and expanded childcare options) to offset the expenses associated with childbearing for working women/families with newborn children.

²⁹ The Employment Promotion Law passed in 2007 prohibits discrimination on numerous groups including gender. The Law on Labor Dispute Mediation and Arbitration, also passed in 2007, provides workers with the means to enforce and protect their labor rights. See Burnett (2010) for more details on the history of women’s employment rights in China.

Our study, like other resume correspondence studies, only focuses on the very first stage of recruitment – the interview callback. Hence, we are unable to observe any potential labor market inequalities in the hiring outcome, salary payment, job promotion, employers' layoff or firing decisions. One direction for fruitful future research is to combine correspondence studies with empirical data that match the employers with their employees (e.g., Hellerstein et al., 1999). This approach will allow researchers to investigate potential labor market inequalities that may occur in the later stages of recruitment and employment.

Acknowledgments

We are grateful to Amy Farmer, Fabio Galeotti, Elizabeth Hoffman, Andrew Horowitz, Raja Kali, Chuliang Luo, David Neumark, Marie Claire Villeval, Qian Weng, and seminar participants at the Beijing Normal University, Central University of Finance and Economics, Renmin University of China, Tsinghua University, University of Arkansas, Iowa State University, the 6th SEBA-GATE international workshop, the 2016 ESA international meeting, the 2017 North America ESA meeting, the 2018 Asia-Pacific ESA meeting for very useful comments. We thank Wanyu Jiang, Lunwen Qi, Huikun He, Zhenglin Sun, Xingyu Zhong, Lijia Wang, Yumei Peng, Han Huang, Wenqi Guan, Bin Wei, Xuehan Yang, Yuqi He, Jiaqian Song, Wenxin Yan, Zhen Huang, Tenghui Wang, Jing Pan, and Chao Wang for excellent research assistance. Financial support from the National Natural Science Foundation of China (Projects No. 71303022&71503255) and the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project No. 13YJC790039) is gratefully acknowledged.

References

- Agan, Amanda, & Starr, Sonja. (2017). Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment. *The Quarterly Journal of Economics*, 133(1), 191–235.
- Anderson, D.J., Binder, M., Krause, K., 2002. The motherhood wage penalty: which mothers pay it and why? *American Economic Review, Papers and Proceedings* 92, 354–358.
- Anderson, D.J., Binder, M., Krause, K., 2003. The motherhood wage penalty revisited: experience, heterogeneity, work effort, and work-schedule flexibility. *Industrial and Labor Relations Review* 56, 273–294.
- Baert, S. (2014). Career lesbians. Getting hired for not having kids? *Industrial Relations Journal*, 45(6), 543-561.
- Bai, Y. (1992). A comparative study on the social interaction of only child and non-only child between 5 to 12 years old. *Journal of Inner Mongolia Normal University (Philosophy & Social Science, in Chinese)*, 4, 111-117.
- Bao, L. (2011). An empirical study on stereotypes of the only child in China. *Contemporary Youth Research (in Chinese)*, 9, 13-19.
- Becker, G. (1957). *The Economics of Discrimination*. Chicago, IL: The University of Chicago Press.
- Bedi, A. S., Majilla, T., & Rieger, M. (2018). Gender Norms and the Motherhood Penalty: Experimental Evidence from India. *IZA Discussion Papers*.
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, 94(4), 991-1013.
- Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. *Handbook of Economic Field Experiments*, 1, 309-393.
- Benard, S., & Correll, S. J. (2015). The Motherhood Wage Penalty and Status Discrimination. *Working in America: Continuity, Conflict, and Change in a New Economic Era*, 271.
- Benard, S., Paik, I., Correll, S.J. (2008). Cognitive bias and the motherhood penalty. *Hastings Law Journal* 59, 1359–1387.
- Bloomberg News (2018a). China Considers Ending Birth Limits as Soon as This Year. 2018-05-21. <https://www.bloomberg.com/news/articles/2018-05-21/china-said-to-consider-ending-birth-limits-as-soon-as-this-year>, accessed on May 23, 2018.
- Bloomberg News (2018b). China’s Baby Boom Needs to Start in the Workplace: The Key to Increasing Fertility to Empower Women. 2018-05-22. <https://www.bloomberg.com/view/articles/2018-05-22/to-increase-births-china-needs-to-empower-working-women>, accessed on May 31, 2018.
- Budig, M. J., & Hodges, M. J. (2010). Differences in disadvantage variation in the motherhood penalty across white women’s earnings distribution. *American Sociological Review*, 75(5), 705-728.
- Budig, M. J., & Hodges, M. J. (2014). Statistical models and empirical evidence for differences in the motherhood penalty across the earnings distribution. *American Sociological Review*, 79(2), 358-364.
- Burnett, J. (2010). Women’s employment rights in China: creating harmony for women in the workplace. *Indiana Journal of Global Legal Studies*, 17(2), Article 8.
- Cameron, L., Erkal, N., Gangadharan, L., & Meng, X. (2013). Little emperors: behavioral impacts of China’s One-Child Policy. *Science (New York, NY)*, 339(6122), 953-957.
- Chen, K. (1985). Discussion about gregariousness of only child—based on an investigation on 1000 children in Beijing. *Psychological Science (in Chinese)*, 4, 45-47.
- Chen, Z., & Deng, C. (2007). Study on desired birth of Chinese women. *Chinese Journal of Population Science (in Chinese)*, 2007(6), 75-81.
- Chew, S. H., Yi, J., Zhang, J., & Zhong, S. (2017). Risk aversion and son preference: experimental evidence from Chinese twin parents. *Management Science*, forthcoming.
- CNN. (2015). “China unveils two-child policy.” December 27, 2015. <https://www.cnn.com/2015/12/27/asia/china-two-child-policy/index.html> (accessed April 20, 2016).
- Cooke, L. P. (2014). Gendered parenthood penalties and premiums across the earnings distribution in Australia, the United Kingdom, and the United States. *European Sociological Review*, 30(3), 360-372.

- Correll, S.J., Benard, S.S., Paik, I., 2007. Getting a job: is there a motherhood penalty? *American Journal of Sociology* 112, 1297–1338.
- Duguet, Emmanuel, and Pascale Petit. 2005. Hiring discrimination in the French financial sector: An econometric analysis on field experiment data." *Annales d'Économie et de Statistique* 78: 79-102.
- Falbo, T. (1979). Only children, stereotypes, and research. In M. Lewis & L. A. Rosenblum (Eds.), *The child and its family* (pp. 127-142). New York: Plenum.
- Falbo, T., & Polit, D. F. (1986). Quantitative review of the only child literature: Research evidence and theory development. *Psychological Bulletin*, 100(2), 176.
- Fan, F., & Wang, J. (2001). Psychological quality and mental health of college students in Beijing. *Research on Education Tsinghua University*, 22(4), 26-32.
- Feng, X. (2000). The growing up of a new generation: the socialization of Chinese only-child teenagers and its outcome. *Social Sciences in China* (in Chinese), 6, 118-131.
- Feng, X. (2006). The one child in China: the size, difference and evaluation. *Theory Monthly* (in Chinese), 4, 5-10.
- Feng, X. (2010). The Construction and Demonstration of the Only Child's Negative Image in Media. *Sociological Studies* (in Chinese), 3, 177-198.
- Feng, X. (2010). Bearing the Second Child: The Analysis on the Desire of the Both Only Child Couples and Related Factors. *Social Science* (in Chinese) 2015 (5), 58-66.
- Glauber, R., 2008. Gender and race in families and at work: the fatherhood wage premium. *Gender & Society* 22, 8–30.
- Hellerstein, J.K., Neumark, D. and Troske, K.R. (1999). Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations. *Journal of Labor Economics*, Vol. 17, No. 3, pp. 409-446
- Helleseter, M. D., Kuhn, P., & Shen, K. (2016). Age and gender profiling in the Chinese and Mexican labor markets: Evidence from four job boards (No. w22187). National Bureau of Economic Research.
- Hu, X. (2006). Urban only-child suffered employment discrimination. *China Youth Daily*. 2006-12-11. http://zqb.cyol.com/content/2006-12/11/content_1604766.htm, accessed on May 15, 2016.
- Huang, L. (1994). No significant difference of personality characteristics between only child and non-only child. *Journal of Anhui University (Philosophy & Social Sciences, in Chinese)*, 3, 90-95.
- Huang, L., & Wen, D. (2008). A comparative study between only child and non-only child college students. *Education Research Monthly* (in Chinese), 2, 36-39.
- Jia, N., Gan, L., and Zhang, J., (2013). Wage Rate, Fertility Trap and Unobserved Types. *Economics Journal* (in Chinese) 2013 (5), 61-72.
- Jing, H. (1997). A comparative study on social psychological quality between college students with and without siblings. *Supplement to the Journal of Sun Yatsen University* (in Chinese), 6, 147-154.
- Killewald, A., & Bearak, J. (2014). Is the motherhood penalty larger for low-wage women? A comment on quantile regression. *American Sociological Review*, 79(2), 350-357.
- Kroft, K., Lange, F., & Notowidigdo, M. J. (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment. *The Quarterly Journal of Economics*, 128(3), 1123-1167.
- Kuhn, P., & Shen, K. (2013). Gender discrimination in job ads: Evidence from China. *The Quarterly Journal of Economics*, 128(1), 287-336.
- Kuhn, P., & Shen, K. (2015). Do employers prefer migrant workers? Evidence from a Chinese job board. *IZA Journal of Labor Economics*, 4(1), 1.
- Li, F. (2015). Campus recruitment of individual companies refused to only-child to provoke controversy. *Wuhan Daily* (in Chinese) 2015-11-05. http://news.cnr.cn/native/gd/20151105/t20151105_520408801.shtml, accessed on May 15, 2016.
- Liang, Z., 2014a, *China Population Problem* (in Chinese), China Development Press
- Jin, Y., Song, J., & Chen, W. (2016). Women's fertility preference and intention in urban China: An empirical study on the nationwide two-child policy. *Population Research* (in Chinese). 40 (6): 22-37.
- Maurer-Fazio, M. (2012). Ethnic discrimination in China's internet job board labor market. *IZA Journal of Migration*, 1(1), 1-24.

- Maurer-Fazio, M., & Lei, L. (2015). "As rare as a panda" How facial attractiveness, gender, and occupation affect interview callbacks at Chinese firms. *International Journal of Manpower*, 36(1), 68-85.
- Maurer-Fazio, M., & Wang, S. (2018). Does Marital Status Affect How Firms Interpret Job Applicants' Un/Employment Histories? *IZA Discussion Papers*.
- Neumark, D. (Forthcoming). Experimental Research on Labor Market Discrimination. *Journal of Economic Literature*.
- Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques and directions for the future. *The Annals of the American Academy of Political and Social Sciences*, vol. 609(1), pp. 104-133.
- Petit, P. (2007). The effects of age and family constraints on gender hiring discrimination: A field experiment in the French financial sector. *Labour Economics*, 14(3), 371-391.
- Qian, N. (2008). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *The Quarterly Journal of Economics*, 123(3), 1251-1285.
- Riach, P. A. and Rich, J. (2002). Field experiments of discrimination in the market place. *Economic Journal*, vol. 112(483), pp. F480-F518.
- Rich, J., (2014). What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted since 2000, *IZA Discussion Papers*, No. 8584.
- Staff, J., & Mortimer, J. T. (2012). Explaining the motherhood wage penalty during the early occupational career. *Demography*, 49(1), 1-21.
- Thompson, V. D. (1974). Family size: Implicit policies and assumed psychological outcomes. *Journal of Social Issues*, 30(4), 93-124.
- Wang, F., Zhao, L., & Zhao, Z. (2016). China's Family Planning Policies and Their Labor Market Consequences. *IZA Discussion Paper No. 9746*
- World Bank Group. (2014). World Bank Staff Estimates Based on the United Nations Population Divisions World Urbanization Prospects. <https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS>, accessed on May 23, 2017.
- Wu, J., & Yang, J. (2014). Zhang, Wang, Li, Zhao, Who is the most common - Structure and distribution of surnames in China 2010 Population Census. *China Statistics (in Chinese)*, 6, 21-22.
- Xu, B. (1980). The only child education—based on an investigation in Fengsheng Street, Xicheng District in Beijing. *Population & Economics (in Chinese)*, 3, 5-10.
- Yu, J., and Xie, Y., (2014). The Effect of Fertility on Women's Wages in China (In Chinese). *Population Research*, 38(1), 18-29.
- Zhang, X., Yu, S., Zhao, J., Li, J., & Xiao, R. (2007). Comparison of mental health state and psychological capacities between college students with and without siblings. *Journal of Southern Medical University (in Chinese)*, 27, 482-484.
- Zhou, X., Zhang, J., & Song, X. (2013). Gender discrimination in hiring: Evidence from 19,130 resumes in China. Available at SSRN 2195840.
- Zhuang, Y., Jiang, Y., Wang, Z., Li, C., Qi, J., Wang, H., Liu, H., Li, B., & Qin, M. (2014). Fertility intention of rural and urban residents in China. *Population Research (in Chinese)*, 2014 (5), 3 -13.

Table 1. Experimental Design

		Female	Male
Before policy change	Only child	FOB	MOB
	With a sibling(s)	FSB	MSB
After policy change	Only child	FOA	MOA
	With a sibling(s)	FSA	MSA

Table 2. Callback Rate (%)

	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 30.8 (46.2)	MOB 31.0 (46.3)	FSB 35.4 (47.8)	FOB 30.1 (45.9)	0.2 [0.884]	-5.3*** [<0.001]
After policy change	MSA 30.2 (45.9)	MOA 29.9 (45.8)	FSA 34.4 (47.5)	FOA 36.6 (48.2)	-0.3 [0.823]	2.2* [0.063]
After - Before	-0.6 [0.747]	-1.1 [0.579]	-1.0 [0.624]	6.5*** [0.001]	<u>Difference in difference (DID)</u> -0.5 [0.794]	7.5*** [<0.001]

Notes: This table reports the callback rates (the standard deviations in parentheses) for the four categories of fictitious applicants before and after the policy change. The number of resumes sent in each cell is 1,106. The differences in callback rates across categories conditional on policy are reported in the upper-right panel with the p values [in brackets] of McNemar's Chi-square test for the paired binomial data. The differences in callback due to the policy change are reported in the lower-left panel with the p values of the unpaired test of proportions. The difference-in-differences (DID) are reported in the lower-right panel. The p values for DID are based on OLS in which the standard errors are clustered at the job ad level. * $p < 10\%$, ** $p < 5\%$, and *** $p < 1\%$.

Table 3. Determinants of the Callbacks (OLS and Probit)

	OLS			Probit		
	[1]	[2]	[3]	[4]	[5]	[6]
Male only-child before (MOB)	0.009 (0.014)	0.004 (0.015)	0.004 (0.015)	0.008 (0.014)	0.002 (0.015)	0.001 (0.015)
Female with a sibling(s) before (FSB)	0.052*** (0.013)	0.052*** (0.013)	0.052*** (0.013)	0.051*** (0.013)	0.050*** (0.013)	0.051*** (0.013)
Male with a sibling(s) before (MSB)	0.007 (0.014)	0.003 (0.014)	0.003 (0.014)	0.006 (0.014)	0.002 (0.014)	-0.000 (0.014)
Female only-child after (FOA)	0.065*** (0.020)	0.061** (0.024)	0.070*** (0.023)	0.063*** (0.020)	0.059** (0.024)	0.067*** (0.024)
Male only-child after (MOA)	-0.002 (0.019)	-0.011 (0.024)	-0.001 (0.023)	-0.004 (0.020)	-0.014 (0.025)	-0.006 (0.024)
Female with a sibling(s) after (FSA)	0.043** (0.020)	0.038 (0.023)	0.048** (0.023)	0.041** (0.020)	0.036 (0.024)	0.046* (0.024)
Male with a sibling(s) after (MSA)	0.001 (0.019)	-0.007 (0.023)	0.002 (0.023)	-0.000 (0.020)	-0.009 (0.024)	-0.002 (0.024)
Applicant characteristics	No	Yes	Yes	No	Yes	Yes
Employer and job characteristics	No	No	Yes	No	No	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,848	8,848	8,848	8,848	8,848	8,848
P-value of F test/Chi-square test	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
R ² /Pseudo R ²	0.030	0.034	0.086	0.024	0.027	0.074

Notes: The dependent variable is an indicator variable for whether a fictitious applicant receives callback or not.

Applicant's individual characteristics include age, education, years of work experience, the number of previous jobs before the application, whether the academic degree was granted by a local university, college major, and university fixed effects. Job and firm characteristics include occupations, the number of applications submitted to this job ad, the number of people who followed this firm on the job board, firm size, firm's ownership, and the type of industry to which the firm belongs. Marginal effects are reported for Probit in columns [4]-[6]. Standard errors clustered at the job ad level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4. Wald Tests on the Equality of Estimates in Table 3

	Difference in Callback Rate (%)					
		OLS		Probit		
	[1]	[2]	[3]	[4]	[5]	[6]
<i>Panel I: Pairwise Comparisons</i>						
FOB-FSB	-5.24*** [<0.001]	-5.21*** [<0.001]	-5.24*** [<0.001]	-5.08*** [<0.001]	-5.04*** [<0.001]	-5.09*** [<0.001]
FOA-FSA	2.26* [0.063]	2.30* [0.071]	2.24* [0.077]	2.23* [0.057]	2.27* [0.063]	2.17* [0.079]
MOB-MSB	0.18 [0.884]	0.06 [0.962]	0.09 [0.943]	0.22 [0.862]	0.07 [0.956]	0.09 [0.942]
MOA-MSA	-0.27 [0.823]	-0.39 [0.761]	-0.37 [0.772]	-0.33 [0.790]	-0.50 [0.703]	-0.41 [0.757]
<i>Panel II: Difference-in-Difference (DD)</i>						
(FOB-FSB)-(MOB-MSB)	-5.42*** [0.003]	-5.27*** [0.006]	-5.34*** [0.005]	-5.30*** [0.004]	-5.11*** [0.007]	-5.18*** [0.006]
(FOA-FSA)-(MOA-MSA)	2.53 [0.150]	2.69 [0.150]	2.61 [0.160]	2.57 [0.143]	2.78 [0.134]	2.57 [0.165]
(FOB-FSB)-(FOA-FSA)	-7.50*** [<0.001]	-7.50*** [<0.001]	-7.49*** [<0.001]	-7.31*** [<0.001]	-7.31*** [<0.001]	-7.25*** [<0.001]
(MOB-MSB)-(MOA-MSA)	0.45 [0.794]	0.45 [0.795]	0.46 [0.789]	0.55 [0.756]	0.57 [0.746]	0.50 [0.778]
<i>Panel III: Difference-in-Difference-in-Difference (DDD)</i>						
[(FOB-FSB)-(FOA-FSA)] - [(MOB-MSB)-(MOA-MSA)]	-7.96*** [0.002]	-7.96*** [0.002]	-7.95*** [0.002]	-7.86*** [0.002]	-7.88*** [0.002]	-7.76*** [0.002]

Notes: This table provides the Wald tests for the equality of estimated coefficients in OLS (or marginal effects in Probit) of Table 3. The p values are reported in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5. Callback Rate by Age Group (%)

5A: Age < 25

	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 32.0 (46.7)	MOB 30.7 (46.2)	FSB 37.2 (48.4)	FOB 31.5 (46.5)	-1.3 [0.515]	-5.7*** [0.008]
After policy change	MSA 29.7 (45.7)	MOA 30.4 (46.0)	FSA 36.3 (48.1)	FOA 37.1 (48.3)	0.6 [0.652]	0.8 [0.596]
After - Before	-2.3 [0.423]	-0.4 [0.889]	-0.9 [0.766]	5.6* [0.061]	Difference in difference (DID) 1.9 [0.447]	6.5** [0.010]

5B: ≥ Age 25

	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 30.2 (45.9)	MOB 31.2 (46.3)	FSB 34.4 (47.5)	FOB 29.3 (45.6)	1.0 [0.538]	-5.0*** [0.003]
After policy change	MSA 31.3 (46.4)	MOA 29.0 (45.4)	FSA 29.9 (45.8)	FOA 35.5 (47.9)	-2.3 [0.285]	5.7*** [0.009]
After - Before	1.1 [0.725]	-2.2 [0.470]	-4.5 [0.148]	6.2** [0.044]	Difference in difference (DID) -3.3 [0.219]	10.7*** [<0.001]

Notes: This table reports the callback rates (the standard deviations in parentheses) by age group before and after the policy change. The statistical tests are the same as in Table 2. * $p < 10\%$, ** $p < 5\%$, and *** $p < 1\%$.

Table 6. Callback Rate by Occupation (%)

6A: Sales

	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 37.4 (48.4)	MOB 38.2 (48.6)	FSB 39.7 (49.0)	FOB 34.0 (47.4)	0.8 [0.620]	-5.7*** [<0.001]
After policy change	MSA 36.0 (48.0)	MOA 36.4 (48.1)	FSA 37.8 (48.5)	FOA 40.4 (49.1)	0.4 [0.792]	2.6* [0.086]
After - Before	-1.4 [0.581]	-1.8 [0.474]	-1.9 [0.478]	6.4** [0.013]	Difference in difference (DID) -0.4 8.3*** [0.860] [<0.001]	

6B: Administrative Assistance

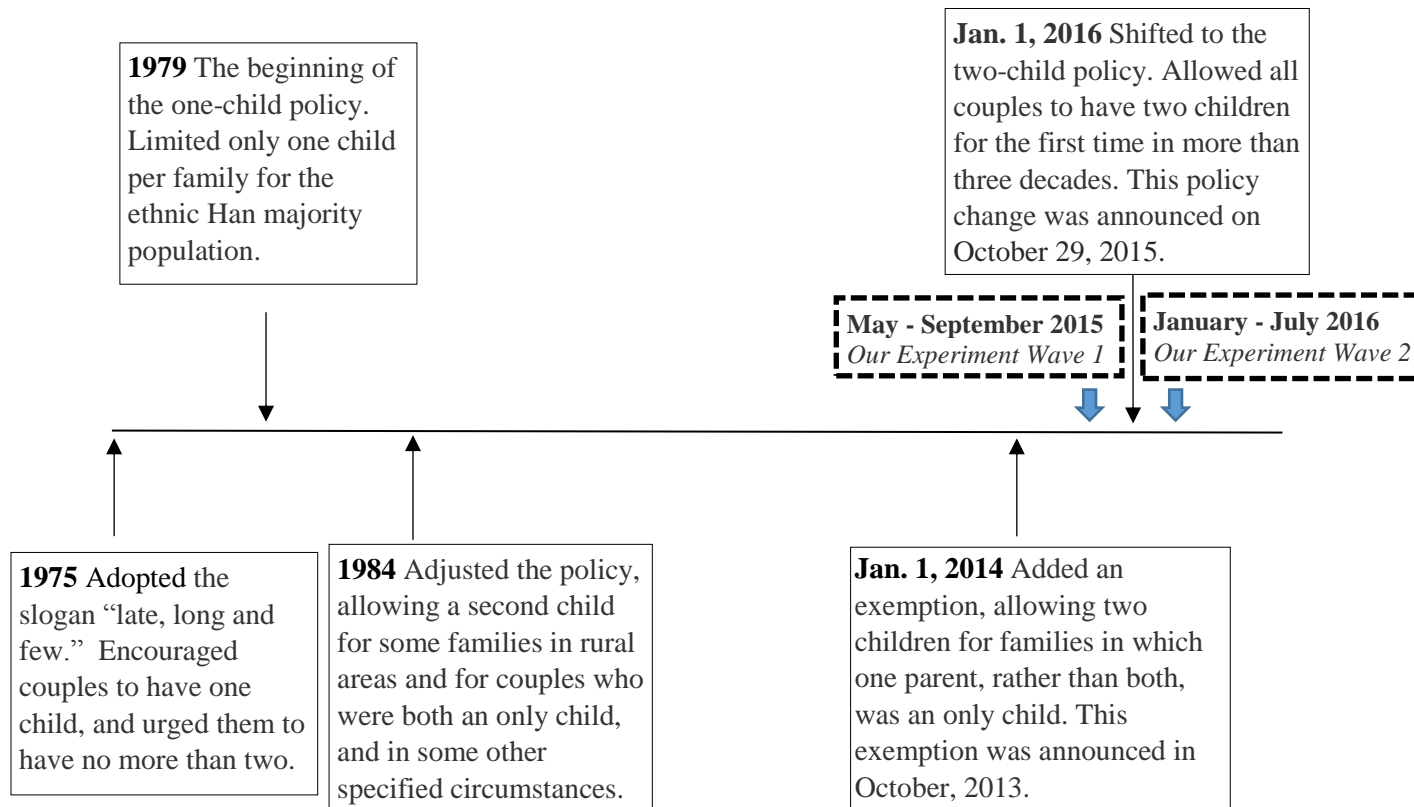
	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 6.0 (23.9)	MOB 5.3 (22.4)	FSB 16.5 (37.3)	FOB 12.8 (33.5)	-0.7 [0.706]	-3.7 [0.275]
After policy change	MSA 6.8 (25.2)	MOA 7.5 (26.5)	FSA 16.5 (37.3)	FOA 15.8 (36.6)	0.7 [0.706]	-0.7 [0.827]
After - Before	0.8 [0.802]	2.2 [0.452]	0.0 [1.000]	3.0 [0.483]	Difference in difference (DID) 1.4 3.0 [0.625] [0.556]	

6C: Customer Service

	Men		Women		MO - MS	FO - FS
	With a sibling(s) (MS)	Only child (MO)	With a sibling(s) (FS)	Only child (FO)		
Before policy change	MSB 25.8 (43.9)	MOB 24.7 (43.2)	FSB 33.3 (47.2)	FOB 28.5 (45.2)	-1.1 [0.612]	-4.9* [0.080]
After policy change	MSA 26.6 (44.3)	MOA 24.0 (42.8)	FSA 34.1 (47.5)	FOA 37.1 (48.4)	-2.6 [0.360]	3.0 [0.322]
After - Before	0.8 [0.844]	-0.7 [0.840]	0.8 [0.855]	8.6** [0.034]	Difference in difference (DID) -1.5 [0.634]	7.9** [0.043]

Notes: Tables 6A-6C report the callback rates, with the standard deviations in parentheses, by occupation before and after the policy change. The number of resumes sent in each cell is 706 for sales, 133 for administrative assistance, and 267 for customer service. The statistical tests are the same as in Table 2. * $p < 10\%$, ** $p < 5\%$, and *** $p < 1\%$.

Figure 1: Timeline of China's One-Child Policy

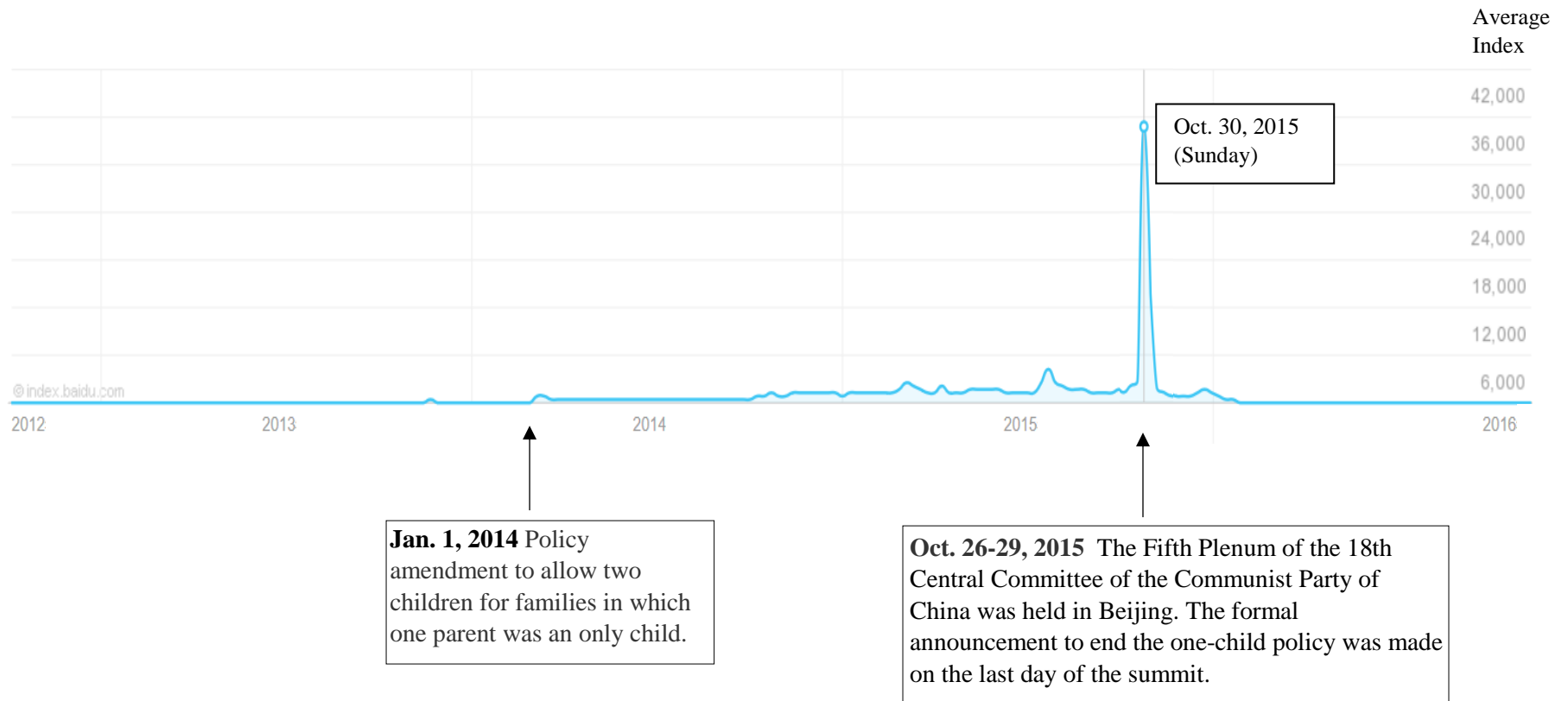


Source: The Guardian
 (URL: <https://www.theguardian.com/world/2013/nov/15/china-one-child-family-policy-timeline>)

Online Appendices

Appendix A. Evidence on the Unexpected Termination of the One-China Policy

Figure A1: Search Volume Index on Baidu for “Second Child for All” (*quán miàn èr tāi*)



Appendix B. Manipulation Check on the Salience of Only- or Sibling-Child Status

We conducted a manipulation check on the salience of the job applicants' status of being an only- or sibling-child by following Kroft *et al.* (2013). In the manipulation check, 99 MBA students at the Beijing Normal University with specialization or previous work experience in human resource management were given a job ad and two fictitious resumes that we used in the experiment. Each of them was asked to act as a human resource manager to choose one resume for job interview. We found that without referring to the resumes, more than two-thirds of our MBA-student-evaluators were able to recall whether the applicants were the only child or had a sibling(s). The likelihood of correctly recalling this information did not significantly differ from that of recalling other important information such as gender, education, and years of work experience. When being asked to comment on the resumes in the post-task survey, no evaluators mentioned anything strange or unusual about the resumes, e.g., showing the only- or sibling-child information. When being asked about which applicant would better fit the job, 20 out of the 90 evaluators referred to the fictitious applicant's only- or sibling-child status in their evaluations.

Appendix C. Did Revealing the Only- or Sibling-Child Status Affect the Callbacks? An Additional Experiment in Shanghai

One question in our study was whether revealing a job applicant's only- or sibling-child status *per se* could affect the callback, relative to not revealing this information at all on the resume. We investigated this question by conducting an additional experiment right after the second-wave of the main study in Shanghai in April and May 2016. In this additional experiment, 948 resumes were sent to 237 job ads in Shanghai. For 120 jobs, we sent four resumes for each ad including male (or female) with the *sibling-child* status, just as in the main study, and male (or female) without such information being revealed. For 117 jobs, we sent four resumes for each ad including male (or female) with the *only-child* status, just as in the main study, and male (or female) without such information being revealed.³⁰ The selection criteria for the job ads and the submission procedure were identical to that in the main study.

Table A1 reports the OLS regression analysis. Results suggest that revealing the applicants' status of being an only- or sibling-child on the resumes does not influence the callback rates. Results from the Probit analysis are very similar.

³⁰ Dividing the job ads into two bins was to avoid an alternative design which was to send to each job ad six resumes including male (or female) with the sibling-child status, male (or female) with the only-child status, and male (or female) without such information being revealed. We did not choose the alternative design since it would change the number of resumes for each job ad.

Table A1. Determinants of the Callbacks in the Companion Experiment in Shanghai (OLS)

Variables	No Info. vs. With-Sibling(s) Status			No Info. vs. Only-Child Status		
	[1]	[2]	[3]	[4]	[5]	[6]
Status info.	-0.025 (0.020)	-0.025 (0.030)	-0.021 (0.031)	0.021 (0.028)	0.017 (0.038)	0.006 (0.041)
Male		-0.092** (0.041)	-0.085* (0.043)		-0.034 (0.042)	-0.084 (0.052)
Status info*Male		0.000 (0.053)	-0.001 (0.059)		0.009 (0.051)	0.049 (0.055)
Constant	0.488*** (0.041)	0.533*** (0.046)	1.840 (1.749)	0.410*** (0.041)	0.427*** (0.046)	-1.765 (1.606)
Control for other variables	No	No	Yes	No	No	Yes
Observations	480	480	480	468	468	468
P-value of F test/Chi-square test	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
R-squared	0.001	0.009	0.141	0.000	0.001	0.154

Notes: The dependent variable is an indicator variable for whether a fictitious applicant receives a callback or not. Standard errors clustered at the job ad level are in parentheses. The control variables (i.e., applicant, employer, and job characteristics) in columns [3] and [6] are the same as in columns [3] and [6] in Table 3. * $p < 10\%$, ** $p < 5\%$, and *** $p < 1\%$.