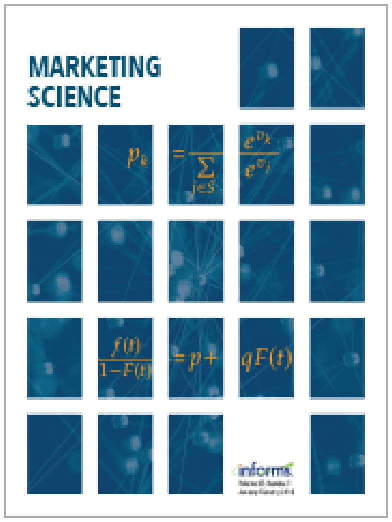

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Do “Little Emperors” Get More Than “Little Empresses”? Boy-Girl Gender Discrimination as Evidenced by Consumption Behavior of Chinese Households

Chen Lin, Yuxin Chen, Jeongwen Chiang, Yufei Zhang

MARKETING SCIENCE
Vol. 40, No. 6, November–December 2021, pp. 1123–1146
ISSN 0732-2399 (print), ISSN 1526-548X (online)
http://pubsonline.informs.org/journal/mksc

Do "Little Emperors" Get More Than "Little Empresses"? Boy-Girl Gender Discrimination as Evidenced by Consumption Behavior of Chinese Households

Chen Lin,[a] Yuxin Chen,[b] Jeongwen Chiang,[c] Yufei Zhang[d]

[a] School of Management, Fudan University, Shanghai 200433, China; [b] NYU Shanghai, Shanghai 200122, China; [c] China Europe International Business School, Shanghai 201206, China; [d] Collat School of Business, University of Alabama at Birmingham, Birmingham, Alabama 35233

Contact: linchen_fd@fudan.edu.cn, https://orcid.org/0000-0001-5666-5886 (CL); yc18@nyu.edu, https://orcid.org/0000-0002-1318-3294 (YC); jwchiang@ceibs.edu, https://orcid.org/0000-0003-4780-0817 (JC); zhangyf@uab.edu, https://orcid.org/0000-0003-0747-8915 (YZ)

Received: September 16, 2020
Accepted: November 11, 2020
Published Online in Articles in Advance: September 9, 2021

https://doi.org/10.1287/mksc.2021.1302

Abstract. This research aims to demonstrate that the abundant marketing data that companies are using to explore new business opportunities can be an equally fertile source for uncovering an undesirable social attitude or behavior that may be relevant to firms' business. Companies may benefit from this knowledge when developing innovative new programs that aim to benefit society, such as corporate social responsibility initiatives. In this study, we examine boy-girl gender discrimination in China as manifested in parents' purchase decisions on behalf of their children across different markets. Our study in itself is significant, because it is the first large-scale empirical work to clearly verify the phenomenon of boy-girl discrimination, taking advantage of e-commerce marketing data. Specifically, we compare the clothing expenditures on boys versus girls using a rich, household-specific data set obtained from two online retailers. We find that the patterns of gender inequality vary systematically across different geographic markets, as the relative expenditure difference on boys versus on girls is bigger in less developed areas as compared with metropolitan areas, and this relative expenditure difference is closely tied with socioeconomic conditions, education levels, and birth rates of a district. Managerial and social implications are discussed.

History: K. Sudhir served as the senior editor for this article.

Funding: C. Lin and J. Chiang thank the research grant support from China Europe International Business School [Grant 7OTSO]. Y. Chen also thanks the National Science Foundation of China [Grant 71972043] for the support.
Supplemental Material: Data and the online appendices are available at https://doi.org/10.1287/mksc.2021.1302.

Keywords: gender inequality • boy-girl discrimination • cause-related marketing • corporate social responsibility

1. Introduction

It has been well recognized that business practices and their outcomes reflect the social value of firms as well as consumers. In fact, this is the premise for the growing popularity of cause-related marketing and, in general, the corporate social responsibility (CSR) activities by firms. Companies across the globe implement various CSR initiatives that include corporate philanthropy, community support, equal opportunity hiring, diversified employment, eco-friendly manufacturing, and cause-related marketing. Companies understand that CSR is not only an ethical or ideological imperative, but also an economic one (Luo and Bhattacharya 2009); that is, through CSR initiatives, companies can project a better corporate image and gain customer support through positive word of mouth (Castaldo et al. 2009), loyalty, and purchase (Brown and Dacin 1997, Luo and Bhattacharya 2006). More importantly, a firm promoting positive social values can be perceived as being a responsible corporate citizen. The question, then, is how to uncover social attitude or behavior that may be relevant to a firm's social endeavor. This research aims to demonstrate the potential social implications gleaned from a firm's own organically-occurring marketing data, which may lead to opportunities for better managerial initiatives, such as CSR programs, without burdening the firm to collect outside survey or sociological/anthropological data.

The beauty-care brand Dove, for example, has been delivering a self-esteem campaign, "The Dove Campaign for Real Beauty," for more than 10 years. The cosmetics brand SK-II is another example. In 2016, it launched the "Change Destiny" and #INeverExpire campaigns in Asia, which aim to inspire women to challenge age-related social pressure.%[1] Increasing efforts have been made to empower young girls. The "Like A Girl" campaign by Always sets out to redefine the negative connotation of doing things "like a girl," while Barbie's "Imagine the Possibilities" campaign hopes to have a lasting positive impact on young girls, showing them that they can achieve anything they want in life. These powerful cause-related marketing campaigns not only strike a chord with women and young girls, they also promote a desirable social value that calls for a change in attitude and behavior across the entire society. There are many similar examples today, especially in emerging markets.%[2] For companies that aim to reach female consumers, the keen observation of the prevalence of undesirable social values and practices is a prerequisite to the success of such sensational cause-related marketing or CSR campaigns. Our research on using firm's transaction data for social implications would be particularly relevant to these companies.

Specifically, we examine the issue of parental boy-girl discrimination within households—a phenomenon that is often observed anecdotally yet is difficult to verify in a society. We explore girl-boy discrimination as manifested in parents' purchase decisions on behalf of their children, because we believe having a better understanding of this societal phenomenon is important for several reasons. First, girl-boy discrimination generates enduring but adverse social impact on societal development and growth (Suitor et al. 2008, Gilligan et al. 2013). In fact, discrimination against girls can begin as early as the prenatal stage by parental discretionary sex selection before birth. After birth, in many societies, boys are observed to get better treatment in nutrition, healthcare, and education opportunities (Hazarika 2000, Barcellos et al. 2014). This parental differential treatment (PDT) has a long-term ill effect on children's developmental experience that lasts into adulthood. Spears and Bigler (2005), for example, argued that children's perception of themselves, as the target of discrimination, is likely to affect their self-esteem, peer relations, academic achievement, occupational goals, and mental and physical well-being. When the practice is pervasive, it means that society is eventually affected and could potentially devolve into a female-unfriendly environment that further dampens women's performance (Jensen 2012).

Second, while the importance of boy-girl discrimination is well recognized by sociologists and economists, documenting and measuring this phenomenon is challenging due to the lack of detailed data. Nobel Prize Laureate Angus Deaton (1989) offers a novel approach to examine boy-girl discrimination within a household by comparing the estimates of expenditure elasticity on adult goods (e.g., alcohol or cigarettes) with respect to the change of family members in a gender group. Nevertheless, more expenditure to feed a boy, for example, is not necessarily an act of son-favoring. Biologically, boys need more calories than girls of the same age.%[3] Whereas the development economics literature offers strong evidence in different welfare-enhancing outcomes like education and nutrition, our research complements the related literature by directly examining the expenditures on children's nonessential goods—that is, consumption not linked to generating future income (i.e., nutrition or education) or human biology—with transaction data possessed by companies. As far as we know, no studies have yet looked into child gender discrimination from this angle.

Third, scholars in marketing have been noticing firms' questionable discriminatory practice by taking advantage of the knowledge of gender differences in product knowledge, attitude, or negotiation skill (i.e., Chen et al. 2008, 2014; Busse et al. 2016). This research looks into the gender discrimination from the opposite angle—discriminatory behavior by consumers instead of firms. This research aims to show evidence of discriminatory behavior across consumer segments; when equipped with this knowledge, firms will be able to launch proper cause-related marketing or CSR campaigns that aim to mitigate this behavior.

The data we utilize in this research come from two leading online children's clothing retail companies in China. Our data are especially useful for understanding the social phenomenon of boy-girl discrimination for the following reasons. First, China has been moving along the path of gender equality despite being still far from the ideal stage. The interactive effects of a traditional son-preferred culture, rapid economic growth, and the enforced family-planning policy make China an interesting context for examining child gender equality.

Second, as explained earlier, examining parents' expenditure on discretionary consumption for their children is an alternative angle to rigorously understand boy-girl discrimination. Among all possible nonessential goods spent on children, we examined specifically the expenditure on young children's clothing. This category was chosen for the following reasons: (1) Unlike expenditure on education or healthcare, parents' expenditure on children's apparel is not associated with children's survival or future earning power, yet differences in providing these goods to boys versus girls can still have deep impacts on their well-being. (2) Clothing is a necessity regardless of parents' socioeconomic conditions. In fact, it is the number one purchased category (25.4%) in the maternal and infant supply industry in China. (3) A prevalent perception is that parents spend more on girls than boys for their clothing needs in the absence of son or daughter favoritism, as reported in the New York City Department of Consumer Affairs (2015) report; in a setting where there is no cultural history of child gender discrimination, the report shows that parents spend 4% more on clothes and 7% more on toys, respectively, for a daughter than a son. Therefore, parents' expenditure on clothes is probably the most salient indicator of discrimination, if there is a reversed outcome.

Third, the online sales data, as compared with traditional household survey data in gender discrimination research, offer many benefits: (1) accessibility—customer purchases are not constrained by distance or location that normally apply to an off-line business; In addition, orders are all shipped to actual addresses, information that is not necessarily easily obtainable from off-line store sales; (2) availability—product displays, payment methods, delivery logistics, and customer service are uniformly presented to all shoppers through the largest Chinese e-commerce platform with no specific groups of customers targeted or excluded.

We estimate the degree of girl-boy discrimination against a set of socioeconomic variables. We find that the likelihood for consumers to spend more on boys' clothing relative to girls is higher among those who live in regions that are economically underdeveloped, are less educated, and have lower birth rates. In other words, rural parents are more likely to show favoritism toward boys as compared with urban parents. Since expenditure is driven by price and/or quantity, we further show that the relative expenditure gap between boys and girls in less developed areas versus that in bigger cities is mainly driven by the quantity. Such relative quantity difference (the quantity for boys relative to quantity for girls in rural vs. metropolitan areas) is not a result of boys going through clothes faster than girls, or brands charging higher prices on boy items, but a result of rural parents willing to buy more new clothes for their sons than for their daughters as compared with metropolitan parents, especially during festivals.

These findings generate substantial social implications for children's clothing companies, as it provides an evidentiary basis for them to use in designing CSR initiatives to have a positive impact on society by reducing parental boy-girl discrimination.

The rest of this paper is organized as follows. Section 2 encompasses a brief review of our research background regarding boy-girl discrimination. Section 3 describes how we organized the data and operationalized our dependent and controlled variables. Then we document and discuss the results of our analyses and related robustness checks. Lastly, we present our conclusions and implications, address limitations, and finally suggest areas for future research.

2. Research Background: Gender Discrimination

Economists and management scientists have studied gender inequality in adults, including workforce participation and performance, such as gender gap in wages and salary (Goldin and Polachek 1987, Blau and Kahn 1994, Ginther and Hayes 1999), promotions (McDowell et al. 1999), and job access (Gobillon et al. 2015). Even today in corporate America, "glass ceilings" persist in U.S. boardrooms (*Financial Times* 2010). One in 10 S&P 500 companies have no female directors, and women's participation on boards has barely moved since 2005. Ding et al. (2013) found that male scientists were almost twice as likely as females to serve on the corporate scientific advisory boards (SABs). Patterns held for the economics profession as well. Using data from American Economic Association members, McDowell et al. (1999) suggested that the promotion prospects for women were inferior to those of their comparable male colleagues.

For gender discrimination on children and adolescents, evidence from literature has not been as rich. One stream of research is on the consequences of boy-favoring discrimination, while the other is the empirical investigations and findings.

Based on U.S. Census data, Ben-Porath and Welch (1976) found that when parents care about the gender of their children, it affects their fertility rate. In fact, child gender preference might lead to male-female ratio imbalance within a society, which over time had a negative impact on female labor force participation (Angrist 2002). In China, the long-time culture of son preference as a result of labor, ritual, inheritance, and old-age security practices, combined with the distorted impact of the government's one-child policy produced what may be the largest gender imbalance in the world (Bulte et al. 2011). The International Planned Parenthood Federation also revealed that more than 70% of aborted fetuses were female, citing the abortion of up to 750,000 female fetuses in China in 1999 (Baculinao 2004). As a result, figures from the National Bureau of Statistics showed that, at the end of 2014, the Chinese mainland population held 33.76 million more males than females. The sex ratio in China was 115.88 to 100, compared with the worldwide norm of about 107 to 100. Using a model of fertility choice when parents have access to a sex-selection technology and face a mandated fertility limit, Avraham (2011) found that a couple's first son was worth 1.42 years of income more than a first daughter, and the premium was highest among less-educated mothers and families engaged in agriculture. Needless to say, the imbalance of the male-female ratio caused many social and economic problems in China (Wei and Zhang 2011). *Prenatal* gender discrimination was not the focus of this study, but these findings demonstrated the grave consequences of pervasive gender discrimination within a society.

On *postnatal* matters, PDT was shown to have a long-lasting impact on a child into adulthood. Studies showed that PDT negatively affected children's relationships with siblings as well as parents continuing into their adulthood (Boll et al. 2003, Gilligan et al. 2013). In addition, research has suggested that the least-favored children experience lower levels of self-esteem and sense of social responsibility and higher levels of aggression, depression, and bad behavior as adults (Feinberg et al. 2001, Suitor et al. 2008).

The expression of PDT takes many forms, including day-to-day parent-child interactions (psychological aspects) to goods parents give their children (physical aspects). As parents tend to regard their children as possessions, children can be viewed as an extension of the self (Derdeyn 1979). From a parent's point of view, giving more (or fewer) material objects to their children is not only a way of conveying how much they care about their children, it is also a way of attempting to bring about desired behaviors in their children. Hence, the PDT of giving goods to children has just as much psychological impact as a parent's words or attitude (Garg and Morduch 1998, Lancaster et al. 2008).

As for the empirical findings of boy-girl discrimination, previous survey-based papers found some evidence in many emerging countries such as India (Behrman 1988, Lancaster et al. 2008), Bangladesh (Asadullah and Chaudhury 2009), Mexico (Antman 2011), Ghana (Garg and Morduch 1998), Côte d'Ivoire (Haddad and Hoddinott 1994), and Papua New Guinea (Gibson and Rozelle 2004). In these papers, children's consumption was often viewed as an intrafamily resource allocation or an intergenerational allocation matter. Early research in this area focused on uneven schooling or healthcare that favored boys, with the notion that favorable expenditure on boys' education and healthcare was viewed as an investment in the family's future income. Children who were expected to be more economically productive in the future would receive a larger share of family resources and had a greater propensity to survive. Several studies found pro-male biases regarding education (Lancaster et al. 2008), nutrition (Behrman 1988), and healthcare (Morduch and Stern 1997, Garg and Morduch 1998) across various countries. But Asadullah and Chaudhury (2009) found reverse gender gaps in education in Bangladesh. Note, however, that these studies were often conducted using small samples (i.e., a handful of villages).

Deaton (1989) proposed a new approach to check for child gender discrimination through intrahousehold expenditure reallocation, though he failed to find boy-girl discrimination in Côte d'Ivoire and Thailand. Many researchers followed his approach and examined the same issue in different countries. For example, using India panel survey data, Subramaniam (1996) found no gender differential in the intrahousehold allocation of resources when controlling for fixed effects of households. Using an experimental approach, Begum et al. (2014) explored parental attitude toward different gendered children. Results suggested that there was no systematic cultural bias in parental attitudes toward the gender of a child. In China, Gong et al. (2005) managed to collect a larger sample of data, including more than 5,000 families from 19 Chinese provinces, and analyzed expenditure patterns in rural China. Regarding the decision on education, they found that boys were more often sent to school, and expenditures on a boy that went to school were larger than those on a school-going girl of the same age. Table 1 summarizes all the related research.

In summary, prior social studies have conclusively found that boy-girl discrimination has long-lasting negative impacts on children and our society's development at large. If this boy-girl discrimination appears to be salient, companies, as good corporate citizens, can and should leverage CSR programs to bring public awareness toward this issue and advocate for solutions to reduce this discrimination in our society. However, among empirical investigations in which researchers have been examining whether boys get more favorable household resource allocations, the conclusion thus far is mixed. We would like to offer a clearer and more complete picture of boy-girl gender discrimination by looking at direct consumption measures. In what follows, we described our data, approach, and analyses.

3. Data

According to a recent report, sales of maternal and infant supplies in China reached RMB 637 billion in 2017 with a 27.3% growth rate compared with the sales in 2016.%[4] Among all categories in maternal and infant supplies, the children's apparel industry was number one (25.4%), followed by baby toys (14.8%). The children's apparel industry, as reported by the National Bureau of Statistics of China, reached RMB 300 billion (USD 47 billion) in sales volume in 2016 with a 25.3% compound annual growth rate (CAGR). The average expenditure on children's apparel was RMB 350 RMB ($55) per child in 2008, growing sharply to RMB 1,700 ($265) in 2017.

First, we introduced companies A and B, two pure e-commerce children's clothing companies in China

Table 1. Empirical Studies on Boy-Girl Discrimination

Reference	Country	Dependent Variable	Findings
Antman (2011)	Mexico	Intrahousehold resource allocation	Immigration of the head of the household affects resource allocation for boys vs. girls.
Asadullah and Chaudhury (2009)	Bangladesh	Education expenditure	Reverse gender gap is significant.
Begum et al. (2014)	Bangladesh	Parental attitudes toward children	No cultural bias in gender is found.
Behrman (1988)	India	Intrahousehold resource allocation of nutrients	Significant son bias is revealed.
Ben-Porath and Welch (1976)	United States and Bengali	Sex preference	Sex preference influences fertility.
Ben-Porath and Welch (1976)	India	Intrahousehold resource allocation	No significant findings regarding sex preference are found.
Bhalotra and Attfield (1998)	Pakistan	Intrahousehold resource allocation	Little evidence on gender differences among children is found.
Deaton (1989)	Côte d'Ivoire and Thailand	Intrahousehold resource allocation	No evidence on gender differences among children is found.
Garg and Morduch (1998)	Ghana	Health expenditure among siblings	Significant son bias is revealed.
Gibson (1997)	Papua New Guinea	Household expenditure	Pro-male bias on expenditure is found.
Gibson and Rozelle (2004)	Papua New Guinea	Intrahousehold resource allocation	Son bias is more prominent in rural areas but less prominent in regions of matrilineal descent.
Gong et al. (2005)	Rural China	Intrahousehold resource allocation	No gender differentials are found in food and alcohol expenditure, but significant son bias is revealed in education expenditure.
Haddad and Hoddinott (1994)	Côte d'Ivoire	Children's anthropometric status	Increases in the proportion of cash income accruing to women can increase boys' height-for-age relative to girls.
Haddad and Reardon (1993)	Burkina Faso	Intrahousehold resource allocation	No evidence on son bias is found.
Lancaster et al. (2008)	India	Education expenditure	Son bias reveals significant impact on education.
Li (2007)	China	Sex ratio at birth	Discrimination against girls has been demonstrated in both prenatal and postnatal periods.
Morduch and Stern (1997)	Bangladesh	Health treatment	Significant son bias is revealed.
Song (2008)	China	Intrahousehold resource allocation	Gender discrimination is found during the early age of children.
Subramaniam (1996)	India	Intrahousehold resource allocation	No evidence on gender differences among children is found.
Zimmermann (2012)	India	Education expenditure	Children's age has a positive impact on discrimination against girls.

for our research. Though these companies were two of many in this very low-concentrated market in China,%[5] they were ranked as the top brands in the children's apparel category on Taobao (the largest e-commerce platform in China), covering almost all geographic markets in the country. We obtained company A's SKU product-level sales data from August 2011 to August 2014 and company B's data from January 2015 to December 2015. Company A had annual sales of RMB 250 million (USD 40 million), whereas the figure for company B is RMB 500 million (USD 80 million). Per their management, the two companies did not have any off-line outlets or off-line advertising channels, nor did they differentiate their products, prices, or promotions across different regions.

Company A launched two primary brands, one exclusively for boys and the other for girls. Based on the purchase records, we selected customers who had *both girls and boys* in their household *in the same year*. The same-year purchase criterion was to control for the possibility that families might have purchases for young boys who did not grow up to appear in the data until the second year or that girls outgrew their clothes during the three-year window. In the latter case, we would observe seemingly higher spending for boys due to elder girls outgrowing their clothes instead of boys being favored. In our sample, if a customer happened to purchase both brands in two (three) years, then we treated this as two (three) separated sample units (this kind of customer constituted only 3% of our sample). For example, if a customer made purchases for both girls and boys in both 2011 and 2012, then we would have two sample units, one for 2011 and one for 2012. If a customer made purchases for both girls and boys in 2011 but only for boys in 2012, then we would only have one sample unit, which was from the purchases in 2011. We used this sample for within-subject comparisons to obtain our data for the main analyses.

Company B had a uniform brand for boys and girls but indicated whether a product was designed for boys or for girls in product names and descriptions. Similarly, we used the data containing customers who purchased both girl and boy clothing for our main analyses. We put the within-subject comparison sample from company A as sample A, and the sample from company B as sample B in what follows.

Samples A and B were both obtained from the company's enterprise resource planning (ERP) system containing information such as item price, discount, category, and shipping addresses. The shipping addresses were essential, because we matched those with district-level statistics obtained from the National Statistics Bureau of China. The latter included socioeconomic information for every administrative district in China. Hence, our research utilized data from multiple sources, which we elaborated in detail in Sections 3.1–3.3. This study was cross-sectional in nature and at the district/county level, since we did not have the socioeconomic information at the household level.

3.1. Sales Data from Sample A and Sample B

We removed records that were identified as institutional purchases (i.e., abnormally large orders) or gifts (i.e., shipping to multiple addresses). Sample A contained 250,664 product-level transactions from 108,291 orders purchased from 43 categories (for boys, for girls, or for both) by 43,506 customers during the three-year window. These sales data were from 2,721 counties and districts, roughly 75% of the total counties/districts in the country. Sample B contained 272,227 product-level transactions from 55,382 orders purchased from 11 categories (for boys, or girls, or both) by 41,158 customers during the one-year window. Sample B's sales data were from 2,480 districts. The complete category list is shown in Online Appendix A1. Since the analysis was at the district level, we further aggregated the data. For instance, to compute a customer's total expenditure on girls' clothing, we added up all the expenditure, quantities, and orders across all categories of this customer. The average price paid was around 90 RMB in sample A and 30 RMB in sample B. Sample A's price range was the price range for the top children's apparel brands—Balabala (92.4 RMB), Gap (104.2 RMB), and Zara (94.9 RMB)—sold in China, whereas sample B's was closer to local affordable brands, according to an industry report.%[6]

Again, these two samples came from two fairly representative companies targeting mainstream (both upper and lower) consumers. Summary statistics from the sales data are discussed in Section 4.1.

3.2. Data on Socioeconomic Information

The 2010 census data from the National Bureau of Statistics of China%[7] covered all of the 3,640 districts across the entire country. It included information such as average education level (years), birth rate, male-female ratio, and percentage of fertile women, minorities, and children. We further collected 2016 district-level GDP data from the International Data Group (IDG, a leading data, marketing services and venture capital organization) as a proxy for economic development. However, IDG provided only GDP for 2,533 counties, and we missed data for a few hundred small districts especially in the rural area as compared with the sales data. To control all other systematic differences across regions, we constructed regional dummy variables (West, East, South, and North) and city-level dummy variables (metropolitan cities, other cities, and rural counties). Table 2 contained the descriptive statistics of the socioeconomic variables. The total number of unique districts is 2,866 once we combine samples A and B.

3.3. Other Data Sources and Variables

3.3.1. Off-Line Shopping. Children's clothing could also be purchased from off-line retail outlets. Though there was no reason to speculate that parental attitude would be different when parents shopped online versus off-line, the concern about the potential effect of shopping formats should be attenuated. Unfortunately, we did not have information about the distribution of children's apparel stores across the country. Instead, we used the information of off-line store locations obtained from Balabala, Gap, and Zara in 2011 to form the proxy and to mitigate the impact of off-line children's clothing purchases. These three brands were among the top five children's apparel brands in China.

3.3.2. Survey Results from Off-Line Competitors and Channel Partners. We conducted a survey of Balabala senior executives, as well as 74 leading national channel partners of off-line children's apparel brands (including Balabala). Some of these channel partners were publicly listed companies that carried a large variety of brands all across China.

Table 2. Descriptive Statistics of Part 2 Data (Number of Unique Districts)

Panel A

Variable	Percentage
Region	
North	35.80%
South	28.58%
West	9.63%
East	25.99%
City level	
Metropolitan cities	2.06%
Other cities	50.56%
Rural counties	47.38%

Panel B

Variable	*N*	Mean	Standard deviation	Min	Max
Log (GDP)	2,533	0.89	0.90	−2.33	4.01
Average education (years)	2,866	9.01	1.32	2.42	13.11
Birth rate	2,866	10.26%	3.60%	1.79%	25.24%
Male-female ratio	2,866	1.05	0.06	0.73	1.57
Percentage of minority	2,866	11.99%	23.86%	0.00%	98.92%
Percentage of children	2,866	16.47%	4.98%	1.05%	35.93%
Percentage of fertile women	2866	28.45%	2.37%	21.33%	38.67%
E-commerce index	2,866	12.08	3.95	2.91	52.59

Note. $N = 2{,}866$, except for log (GDP), which is 2,533.

Highlights from the survey include the following:

- An overwhelming majority (85%) responded that there were no systematic differences in marketing strategies for boys versus girls from the brands as well as the channel partners. Prices between boys' clothing and girls' clothing were very similar.
- Among all the product categories that channel partners operated in, clothing was mostly purchased online (60%). Infant formula was around 30%.
- About half of the respondents thought there was no discrimination in buying clothes for boys versus girls, but for those who thought there was, 76% responded that discrimination was more likely to happen in rural areas and smaller cities.

3.3.3. E-Commerce Development.
Another concern is about the degree of e-commerce penetration and competition across different regions. Shoppers in some areas might be more receptive to e-commerce than others. To control for this potentially compounding factor, we included the 2015 E-Commerce Development Index, a continuous variable created by Alibaba (aEDI),%[8] in which they gathered both information about online shopping and online retailing (including customer expenditure, frequency and vendor density, and competitiveness) to define the level of e-commerce development in a given district.

3.4. Dependent Variable
Following Deaton (1989), the dependent variable (DV) was the ratio of boys' clothing expenditures to girls' clothing expenditures. If the ratio was higher than 1, then we called it *boy-girl expenditure discrimination* by definition. We were primarily interested in how this ratio varied across urban versus rural areas and how it varied with the socioeconomic conditions. We constructed this ratio for each district by computing the aggregate expenditure on boys over the aggregate expenditure on girls in all categories in that district. The analysis of total expenditure over all categories would be more meaningful and logically sound, as there were potential interactions among categories (substitution and complementarity). Combing samples A and B,%[9] we found that the average ratio across all districts was 1.90, much higher than the benchmark of 1 (i.e., equal expenditure). We also tried constructing the DVs with respect to *Quantity* (the ratio of boys' clothing total quantities to girls' clothing quantities) and *Number of orders* (the ratio of boys' clothing number of orders to girls' clothing number of orders) and found consistent patterns (1.48 and 1.15, respectively). The descriptive statistics of those ratios are shown in Table 3. As a preliminary check, we correlated our dependent variable ratios with some macro province-level indicators on gender equality and women's rights in Online Appendix A2.

3.5. Independent Variables
Previous research has suggested that socioeconomic conditions might influence gender discrimination. For example, the *Science* paper of Guiso et al. (2008) empirically showed that the gender differences in math

Table 3. Descriptive Statistics of Various Operationalization of Gender Discrimination (Combined Samples A and B)

				95% Confidence interval[a]	
Operationalization	City level	Mean	Standard deviation	Lower	Upper
Ratio of gender discrimination (*Expenditure* on boys' clothing vs. *Expenditure* on girls' clothing)	Metropolitan cities	1.41[b]	0.91		
	Other cities	1.75	1.77	−0.28	0.97
	Rural counties	2.09	3.59	0.06	1.31
	Total	**1.90**	**2.78**		
Ratio of gender discrimination (*Quantity* of boys' clothing vs. *Quantity* of girls' clothing)	Metropolitan cities	1.16[b]	0.58		
	Other cities	1.39	0.98	−0.08	0.54
	Rural counties	1.59	1.74	0.12	0.74
	Total	**1.48**	**1.39**		
Ratio of gender discrimination (*Number of orders* for boys' clothing vs. *Number of orders* for girls' clothing)	Metropolitan cities	1.06[b]	0.19		
	Other cities	1.14	0.43	−0.02	0.19
	Rural counties	1.16	0.50	−0.00	0.21
	Total	**1.15**	**0.46**		

[a]The total number of districts or sample size used for main regression analyses was 5,201, which was the sum of the number of districts in Sample A and Sample B. We implemented a Tukey test (ANOVA) to examine whether the differences among city levels were statistically significant at a 95% confidence interval.

[b]The ratio of metropolitan cities was the reference group.

scores disappeared in countries with a more gender-equal culture and better economic, political, and educational opportunities for women. Jensen (2012) also argued that if the market environment improved, then women would be able to develop better capabilities that eventually reduce the performance gap.

Therefore, we sorted independent variables into two groups, socioeconomic characteristics and other controlled variables. The socioeconomic characteristics (the primary interest of this research) are as follows.

3.5.1. GDP. Although income was not reported in the census, local GDP was used as a proxy for the economic development of the district. When a certain area was more economically developed, we speculate that people would be more likely to be open and progressive, and, hence, there was less likelihood of son preference. We used GDP instead of GDP per capita, as the latter was calculated as GDP over number of household registrations (referred to as "hukou") and thus often less accurate and less reliable in China.%[10] GDP was also highly correlated with whether a district was rural.

3.5.2. Average Education Level. This was the number of years of education on average in a certain district. Similarly, when parents were more educated, they were less likely to be bound by traditional mindsets.

3.5.3. Birth Rate. This was defined as the average birth rate of a district as a proxy for how many infants were born in a given district. The one-child policy drastically reduced the average fertility rate in urban households from about three in 1970 to just over one by 1982. Gupta and Bhat (1997) showed that one consequence of fertility decline in East Asian countries was the increased manifestation of sex bias, including prenatal gender selection, excessive mortality rate of young girls, and continuous gender discrimination in adulthood. Therefore, we conjectured a negative relationship between birth rate and gender discrimination.

3.5.4. Other Control Variables. We included the male-female ratio (gender balance in the district), minority percentage (percentage of residents who are minorities), region (geographic location dummy variables), percentage of fertile women (percentage of residents who are female and in their child-bearing years), children percentage (percentage of residents who are children), off-line shopping (whether a district has a Balabala, Gap, or Zara store), and e-commerce development index defined by Alibaba. The correlation matrix of all continuous variables included in the analyses is shown in Online Appendix A3.

To summarize, our efforts to collect multisource, multitype, and multicompany data allowed us to examine the gender discrimination on a scale that previous research could not achieve. In Sections 4, 5, and 6, we present how we used these data and what the results were.

4. Empirical Strategy, Analyses, and Results

We show the relative differences of gender discrimination across different city tiers (urban vs. rural areas) in Section 4.1 and across different socioeconomic conditions in Section 4.2.

4.1. Discrimination Across City Tiers

We obtained city-level information from the State Council%[11] (whether a district was located in a metropolitan city, other city, or rural county). Using our combined sales data for samples A and B, we compared discrimination ratios across city levels. We contrasted these parameters within each level of districts. We found that the expenditure ratio in rural counties (2.09) was significantly larger than that in metropolitan cities (1.41) at 95% confidence interval using the Tukey test (analysis of variance [ANOVA]). Similar patterns held for the ratios using quantities. For *Number of orders*, we still found that the ratio in rural counties was larger than that in metropolitan cities, although not statistically significant (confidence interval [CI] [−0.00, 0.21]). The descriptive statistics are shown in Table 3.

For sample A, we compare in Table 4 the number of items and orders, total expenditure, and average paid price between the boy brand and the girl brand that a given customer bought by using a paired-sample *t*-test. We found that people, in general, spent more (i.e., more items, more orders, more expenditure, and more expensive products) on boys than on girls. What's more remarkable was the relative differences in expenditure between boys and girls across city tiers. The expenditure gap was significantly bigger in smaller cities and in rural areas than in metropolitan areas, though the price gap was much smaller. Whereas the differences in price paid were not significantly different among city tiers, the relative higher expenditure ratio of boys versus girls in rural counties was, in fact, driven by the relative *quantity* difference between boys and girls (*Item Quantities* in Table 4) in rural areas versus cities.

Using the difference between the boy-brand expenditure and the girl-brand expenditure of metropolitan cities as the reference group (control group) and given the CIs derived from the difference-in-difference (DID) Tukey test (ANOVA), we found that the difference between boys and girls in rural counties was significantly larger than that of metropolitan cities across the total expenditure, total quantities, and number of orders. Furthermore, the results indicated that the difference between boys and girls in nonmetropolitan cities was significantly larger than that of metropolitan cities across the same parameters.

For sample B, we compare in Table 5 the same variables between the boy clothing and the girl clothing that a given customer bought by using a paired-sample *t*-test. The results indicated that people spent more on boys than on girls. We found consistent patterns, as in sample A's data, that the expenditure difference between boys and girls was the largest in rural counties.

Unlike sample A, sample B reported larger total quantities and number of orders for girls, yet the difference was smallest in rural areas.

Although the prices paid for boys were higher than for girls in sample B, as in sample A, the price difference did not significantly vary between cities and rural areas. In addition, the sample B brand was lower-end, and hence customers might be more price-sensitive. As a result, consumers bought fewer boy's clothes than girl's, but that *quantity* difference was significantly smaller in rural areas than in cities, as shown in the first section (under "Item quantities") of Table 5. Therefore, we still saw that the relative expenditure on boys versus girls was higher in rural areas than in cities. Consistent with Table 4, the relative higher expenditure ratio of boys versus girls in rural counties was in fact again driven by the relative *quantity* differences in rural areas versus cities.

Whereas boy's clothes were 25%–30% more expensive than girl's, the quantities were just 6%–18% less for boys in sample B. By contrast, when the price difference was only 6% in sample A, the quantity consumption for boys was 35%–50% higher than that for girls. These facts further explain that the main driver for the total expenditure on boys to be higher than that on girls was the quantity difference rather than the price difference.

We would like to show that our expenditure gap results were not an artifact of higher pricing for boys' clothing. We developed a measure that was independent of the distribution of prices in the market (i.e., supply side). We denoted this measure as λ, to represent the sales-weighted average percentile of price paid given the market price distribution. It measured the propensity of consumers to purchase at the high end of price distribution (i.e., more expensive items). The detailed specification and discussion are included in Online Appendix A4.

We found that this measure was consistently higher for boys than girls for both samples, as shown in Tables 6 and 7. The difference did not significantly vary between cities and rural areas in sample A, whereas rural families were more likely to buy in the relatively higher end of the price distribution for boys than for girls in sample B. We believed this was an interesting insight, even though, as we discussed earlier, price was not the main driver for discrimination measured by relative expenditure.

Table 4. Customers Who Bought Both Boy and Girl Brands: Expenditure on Boy Brand vs. Girl Brand (Sample A Data—All Categories)

Panel A

	Item quantities (total)				95% Confidence interval[a]		Number of orders				95% Confidence interval[a]		Total expenditure				95% Confidence interval[a]	
City level	Boy	Girl	t-value[b]	D[c]	Lower	Upper	Boy	Girl	t-value[b]	D[c]	Lower	Upper	Boy	Girl	t-value[b]	D[c]	Lower	Upper
Metropolitan cities	3.34	2.47	12.58*	0.87[d]			1.57	1.38	8.61*	0.19[d]			309.04	212.31	14.07*	96.73[d]		
Other cities	3.53	2.41	39.02*	1.11	0.06	0.42	1.73	1.42	30.47*	0.31	0.06	0.19	347.19	222.20	41.32*	124.99	9.07	47.45
Rural counties	3.40	2.26	21.98*	1.13	0.05	0.47	1.72	1.38	17.77*	0.34	0.07	0.22	345.96	218.12	23.09*	127.85	8.78	53.46

Panel B

	Paid average price				95% Confidence interval[a]	
City level	Boy	Girl	t-value[b]	D[c]	Lower	Upper
Metropolitan cities	93.95	87.59	5.88*	6.36[d]		
Other cities	99.19	93.74	13.14*	5.45	−3.61	1.80
Rural counties	102.80	98.91	4.75*	3.89	−5.62	0.68

[a]A difference-in-differences (DID) significant test was used to examine if the differences were significantly distinct among city levels, with the difference in metropolitan cities (i.e., 0.87 for item quantities) as the reference. We implemented a Tukey test in post hoc tests. There was a significant DID if the confidence intervals (CIs) derived from the DID Tukey test did not contain zero. difference_quantity_metropolitan = 0.87; difference_quantity_rural = 1.13 with CI of the DID test = [0.05 to 0.47]; difference_orders_metropolitan = 0.19; difference_orders_rural = 0.34 with CI of the DID test = [0.07 to 0.22]; difference_expenditure_metropolitan = 96.73; difference_expenditure_rural = 127.85 with CI of the DID test = [8.78 to 53.46]; difference_quantity_metropolitan = 0.87; difference_quantity_cities = 1.11 with CI of the DID test = [0.06 to 0.42]; difference_orders_metropolitan = 0.19; difference_orders_cities = 0.31 with CI of the DID test = [0.06 to 0.19]; difference_expenditure_metropolitan = 96.73; difference_expenditure_cities = 124.99 with CI of the DID test = [9.07 to 47.45]. The t-values were derived from paired sample t-tests between the boy brand and girl brand.

[c]D is the difference between boy brand and girl brand.

[d]The difference between boy brand and girl brand for people in metropolitan cities was the reference group.

*$p < 0.05$, one-tailed test; $N = 43{,}506$.

Table 5. Customers Who Bought Both Boy Clothing and Girl Clothing: Expenditure on Boy Clothing vs. Girl Clothing (Sample B Data—All Categories)

Panel A

	Item quantities (total)				95% Confidence interval[a]		Number of orders				95% Confidence interval[a]		Total expenditure				95% Confidence interval[a]	
City level	Boy	Girl	t-value[b]	D[c]	Lower	Upper	Boy	Girl	t-value[b]	D[c]	Lower	Upper	Boy	Girl	t-value[b]	D[c]	Lower	Upper
Metropolitan cities	3.28	4.01	−8.86*	−0.72[d]			1.12	1.19	−9.40*	−0.08[d]			102.79	97.11	2.33*	5.69[d]		
Other cities	3.26	3.64	−6.02*	−0.38	0.04	0.64	1.12	1.18	−14.78*	−0.05	0.00	0.04	109.98	94.39	7.72*	15.59	0.41	19.39
Rural counties	3.34	3.57	−3.08*	−0.23	0.13	0.85	1.10	1.15	−9.27*	−0.05	0.00	0.05	113.81	93.54	9.37*	20.27	3.31	25.85

Panel B

	Paid average price				95% Confidence interval[a]	
City level	Boy	Girl	t-value[b]	D[c]	Lower	Upper
Metropolitan cities	32.77	25.04	19.74*	7.73[d]		
Other cities	34.96	28.34	19.41*	6.63	−2.83	0.62
Rural counties	36.09	28.81	12.14*	7.28	−2.49	1.60

[a]A difference-in-differences (DID) significant test was used to examine if the differences were significantly distinct among city levels, with the difference in metropolitan cities (i.e., −0.72 for item quantities) as the reference group. We implemented a Tukey test in post hoc tests. There was a significant DID if the confidence intervals (CIs) derived from the DID Tukey test did not contain zero: difference_quantity_metropolitan = −0.72; difference_quantity_rural = −0.23 with CI of the DID test = [0.13 to 0.85]; difference_orders_metropolitan = −0.08; difference_orders_rural = −0.05 with CI of the DID test = [0.00 to 0.05].
[b]The t-values were derived from paired sample t-tests between boy clothing and girl clothing.
[c]D is the difference between boy clothing and girl clothing.
[d]The difference between boy clothing and girl clothing for people in metropolitan cities was the reference group.
*$p < 0.05$, one-tailed test. $N = 41{,}158$.

Table 6. Customers Who Bought Both Boy Brand and Girl Brand: Boy Brand Lambda vs. Girl Brand Lambda (Sample A)

	Lambda				95% Confidence Interval[a]	
City level	Boy	Girl	*t*-value[b]	D[c]	Lower	Upper
Metropolitan cities	0.493	0.452	7.77*	0.04[d]		
Other cities	0.523	0.490	16.57*	0.03	−0.02	0.00
Rural counties	0.549	0.520	7.58*	0.03	−0.03	0.00

[a] A difference-in-differences (DID) significant test was used to examine if the differences were significantly distinct among city levels, with the difference in metropolitan cities (i.e., 0.04 for lambda) as the reference group. We implemented a Tukey test in post hoc tests with 95% confidence intervals (CIs). There was a significant difference in differences if the CIs derived from the DID Tukey test did not contain zero.
[b] The *t*-values were derived from paired sample *t*-tests between boy brand and girl brand at each city level.
[c] D is the difference between boy brand lambda and girl brand lambda.
[d] Difference between boy-brand lambda and girl-brand lambda of people in metropolitan cities was the reference group.
*$p < 0.05$, one-tailed test; $N = 43{,}506$.

For the relative quantity difference, one might argue that gender differences in wearing clothes could be the driving factor for the higher expenditure on boys. For example, boys were naturally more active, so they wore out clothes faster and required more purchases than girls did.

However, in today's world, where waste from clothing is a globally trending topic, the actual wearing-out of children's clothes rarely happens (more likely to grow out instead). Industry experts and the International Textile Fair Claims Consumer Guide further confirmed that children's clothing is designed to last more than three years, and boys' clothes often use more long-lasting fabrics. As for the potential grow-out argument, growth chart statistics show that boys and girls grow at about the same rate, with the girls' rate rising faster in adolescence. But we observed a larger difference in relative quantities for rural boys in our data.

In addition, if, under the usage rate assumption of boys being more active than girls and rural children being more active than city children, quantity consumption would be boys > girls and rural > city, then we should expect that the quantities for rural boys > (city boys or rural girls) > city girls. But our results suggest that rural girls < city girls in both item quantity and number of orders among our sample of boy-girl families (Table 4, sample A, and Table 5, sample B). In fact, the quantity for city girls was even higher than for rural boys in sample B. This pattern could not be explained by the wear-out rate difference between rural areas and cities but could be explained by gender discrimination. In other words, rural parents tended to allocate more clothing budget and buy more new clothes for their sons as compared with urban parents among our boy-girl families.

We would like to highlight that the *relative quantity differences* for boy versus girl between rural area and cities were not due to the usage rate difference and the *need* to replace worn-out/outgrown clothes.

To further tackle this issue, we looked at occasions for new purchases—specifically, the festival purchases. Major festivals in China include New Year, Chinese New Year (usually in late January or early February, but logistics and delivery would be closed 15–20 days before), Children's Day (June 1), Mid-Autumn Festival (usually in late September or October), and National Day (October 1, followed by a seven-day holiday period). We examined purchases before the major festivals in five months—January, May, September, October, and December—and compared them with nonfestival purchases in the other

Table 7. Customers Who Bought Both Boy Clothing and Girl Clothing: Boy Clothing Lambda vs. Girl Clothing Lambda (Sample B)

	Lambda				95% Confidence interval[a]	
City level	Boy	Girl	*t*-value[b]	D[c]	Lower	Upper
Metropolitan cities	0.491	0.480	2.79*	0.01[d]		
Other cities	0.535	0.497	16.34*	0.04	0.02	0.04
Rural counties	0.551	0.503	18.68*	0.05	0.03	0.05

[a] A difference-in-differences (DID) significant test was used to examine if the differences were significantly distinct among city levels, with the difference in metropolitan cities (i.e., 0.01 for lambda) as the reference group. We implemented a Tukey test in post hoc tests with 95% confidence intervals (CIs). There was a significant DID if the CIs derived from the DID Tukey test did not contain zero. The *t*-values were derived from paired sample *t*-tests between boy clothing and girl clothing at each city level.
[c] D is the difference between boy clothing lambda and girl clothing lambda.
[d] Difference between boy clothing lambda and girl clothing lambda of people in metropolitan cities was the reference group.
*$p < 0.05$, one-tailed test; $N = 41{,}158$.

seven months (which included the biggest shopping festival on November 11).

As shown in Table 8, we found that the *relative quantity difference* between city levels was mainly driven by festival purchases and was not significant during other times of the year. Consistent with the results in Tables 4 and 5, we observed larger quantities for rural boys than for rural girls, relative to purchases for the metropolitan boys versus girls in sample A, and a significantly smaller *quantity difference* in rural areas as compared with that in city areas in sample B for festival purchases.

Wearing new clothes for the festivals had historical, customary, and symbolic significance in China, yet we observed families in socioeconomically less-developed areas *want* to purchase more new clothes for boys during festivals as compared with families in more developed areas. Fundamentally, the expenditure discrimination in the form of *relative quantity difference* was more tied to the tradition and culture, as opposed to the gender difference in the usage rates of clothes.

In summary, we found convergent evidence that the average expenditure difference between boys and girls in rural counties was larger than the expenditure difference in metropolitan cities, contrary to previous media reports in the United States (New York City Department of Consumer Affairs 2015) and the United Kingdom (*Daily Mail* 2016) that girls should have higher expenditure on clothing. Difference in quantity was the key driver of the expenditure gap in gender in rural counties.

4.2. Discrimination Associated with Socioeconomic Variables

The ordinary least squares (OLS) regression was the primary estimation method we employed in this study after testing on homoscedasticity. Results from the combined sample (district-level) are shown in Table 9.%[12] The results revealed that families in more economically advanced areas ($B = -0.12$, $p < 0.05$), in districts with higher education level ($B = -0.06$, $p < 0.05$), and in the areas with higher birth rates ($B = -0.03$, $p < 0.05$) were less discriminatory toward their girls. Similar patterns held when using *Quantity* and *Number of orders* as the DVs. All of these results aligned with our conjectures. Online Appendix A5 contains the full results of our main regression analyses. To ensure the validity and reliability of our analyses, we conducted a series of robustness checks, which are shown in Section 5.

5. Robustness Checks

Our measures were potentially subject to confounding factors that might not truly reflect gender discrimination. Hence, in Section 5.1, we first provide a discussion of our empirical strategies to address these concerns. The details of the robustness checks discussed in Section 5.1 are then reported in Sections 5.2–5.7.

5.1. Discussion on Potential Confounding Factors

In this section, we listed potential confounding factors to our gender discrimination measure and explained how we would try to rule them out. First, one might argue that consumer brand preference could potentially confound our measure. For example, even though the company did not deliberately implement gender-specific marketing strategies, the boy brand might be better received by rural customers, whereas the girl brand might be better received by city customers. To rule out this, we conducted robustness check 1, in which we used customer-level analysis in rural counties only, combining both samples. The patterns of discrimination across socioeconomic conditions still held in this more homogeneous subsample.

Second, to eliminate the concern that districts with few customer representatives might bias the obtained results, we reran our main regression analysis using data with the bottom 10%, 20%, and 30% samples trimmed accordingly (based on the number of customers aggregated to a district). The consistent results we gained in robustness check 2 helped enhance the reliability of our findings.

Third, although the company executives mentioned that they did not implement any district-specific marketing strategy, it was possible that the availability of off-line options and competitive landscape in each region was different. Although we tried our best to control for off-line competitions, one may question (1) whether, as compared with boys, urban girls had more options than rural girls in the off-line space, and therefore clothing for urban girls was bought off-line (substitution effect); and (2) whether, as compared with boys, urban girls could return more easily than rural girls (return effect). For the first case, intuitively, urban girls, as compared with boys, should have more variety, more expensive options, and more try-on opportunities than rural girls in the off-line space, meaning that the relative urban girls/urban boys' online consumption should not be higher than their rural counterparts—but that is not what we observed in our results. Therefore, considering the possibility of off-line options, our results are strengthened even more. For the return argument, unfortunately, we did not have return information in this data set. However, China is probably one of the most advanced countries in logistics covering rural areas with speedy delivery and flexible return policy. A working paper on online product returns (Zhang et al. 2019) used data from a leading women's clothing company and found no

Table 8. Purchase Patterns During Festival Months vs. Nonfestival Months[a]

Sample A												
	Item quantities (total) during festival months				95% Confidence interval[b]		Item quantities (total) during nonfestival months				95% Confidence interval[b]	
City level	Boy	Girl	t-value[c]	D[d]	Lower	Upper	Boy	Girl	t-value[c]	D[d]	Lower	Upper
Metropolitan cities	2.34	1.80	8.85*	0.54[e]			2.81	2.07	10.76[e]	0.74[e]		
Other cities	2.40	1.67	30.40*	0.73	0.04	0.34	2.93	2.06	29.54*	0.87	−0.05	0.31
Rural counties	2.26	1.53	18.00*	0.73	0.02	0.37	2.76	1.85	17.60*	0.91	−0.05	0.38
Sample B												
	Item quantities (total) during festival months				95% Confidence interval[b]		Item quantities (total) during nonfestival months				95% Confidence interval[b]	
City level	Boy	Girl	t-value[c]	D[d]	Lower	Upper	Boy	Girl	t-value[c]	D[d]	Lower	Upper
Metropolitan cities	3.02	3.61	−6.08*	−0.59[e]			2.90	3.62	−7.12*	−0.72[e]		
Other cities	3.10	3.24	−3.24*	−0.14	0.21	0.69	2.81	3.30	−10.02*	−0.49	−0.05	0.50
Rural counties	3.32	3.30	0.25	0.02	0.33	0.90	2.78	3.27	−4.52*	−0.49	−0.11	0.56

[a] Festivals include Children's Day, Mid-Autumn Festival, National Day, New Year, and Chinese New Year. Festival months include purchases occurred in January, May, September, October, and December. Nonfestival months are the rest of the months in a year.

[b] A difference-in-differences (DID) significant test was used to examine if the differences were significantly distinct among city levels, with the difference in metropolitan cities (i.e., 0.54 for item quantities (total) during festivals) as the reference group. We implemented a Tukey test in post hoc tests. There was a significant DID if the CIs derived from the DID Tukey test did not contain zero.

[c] Difference between boy clothing and girl clothing of people in metropolitan cities was the reference group.

[d] The t-values were derived from paired sample t-tests between boy brand and girl brand.

[e] D is the difference between boy clothing and girl clothing.

*$p < 0.05$.

Table 9. OLS Results for Main Regression Analyses (Combined Samples A and B)

Variables	Main regression analysis with ratio of gender discrimination (*Expenditure*) as DV (district-level data)		Main regression analysis with ratio of gender discrimination (*Quantity*) as DV (district-level data)		Main regression analysis with ratio of gender discrimination (*Number of orders*) as DV (district-level data)	
	Within-subject (families with both boys and girls) comparison: Expenditure on boy clothing vs. expenditure on girl clothing					
	B	*t*-value	*B*	*t*-value	*B*	*t*-value
Log (GDP)	−0.12*	−4.08	−0.08*	−3.94	−0.02*	−2.04
Average education (years)	−0.06*	−2.15	−0.04[a]	−1.94	−0.01	−0.97
Birth rate	−0.03*	−3.08	−0.02*	−3.11	−0.01*	−2.05
Sample[b]	0.43*	11.65	0.61*	23.80	0.30*	29.77
Covariates[c]						
R^2	4.84%		12.21%		16.05%	

[a] $p < 0.10$; *$p < 0.05$. $N = 4{,}647$; the reduced sample size was because we were unable to obtain some small counties' GDP information.
[b] Sample B was the reference group.
[c] Covariates consisted of city levels (other cities and rural cities with metropolitan cities as the reference group), male-female ratio, percentage of minority, region, off-line shopping (Balala Children's Clothing Company), e-commerce development index, percentage of fertile women, and percentage of children.

systematic differences between urban return rates and rural return rates.

To further control for the unobserved local demand and supply factors that might muddle our results, we performed robustness check 3, which proposed an incremental measure of gender discrimination, that is, the relative favoritism toward boys in families with both a boy and a girl compared to the favoritism toward boys rather than girls among families with children of a single gender in the same location. This comparison controlled for unobserved local demand and supply factors. We found even stronger evidence of gender discrimination in families with children of mixed genders.

Fourth, the inherent family composition and birth order might also affect the relative expenditure on boys versus girls. From the sampling perspective, the distribution of a BG (boy-then-girl) family and a GB (girl-then-boy) family might be unbalanced.%[13] Then, if there was favoritism toward younger or older kids, then this might confound our results on gender discrimination. Thus, we implemented robustness check 4, which compared the expenditure ratios of the second child versus the first child in the GB, GG (girl-then-girl), and BG families and further confirmed that the favoritism was indeed toward the boy instead of the younger kid. Again, we computed the ratios by first aggregating across households and then taking the ratio of the second child over the first one. We used the clothing size as a proxy for child age to determine if a family is a BG, GB, or GG family. Here, we utilized one additional variable that we had from sample A: size. Sizes across different clothing categories could be very sparse. For example, a child may need size 120 for a T-shirt, but size 130 for outerwear. Children (both boys and girls) usually grow (at least) one size up every year. Although size was probably not a clean proxy for age, we could use size as a screening variable for two-children families, which we defined as families who purchased clothes that were more than two sizes apart within the year of study. For instance, if a family purchased boys' clothing and then purchased boys' clothing two sizes larger within the same year, then this family is deemed a BB (boy-then-boy) family. This robustness check used only sample A, since only sample A has clothing size information.

Lastly, given that China is a large country with substantial climate variation across regions, one concern might be that popular items purchased could be different across regions (E.G., coat vs. T-shirt). If, for example, coats are more expensive than T-shirts and there are more rural areas in North China (where coats are more popular) than in South China (where T-shirts are more popular), then our main results could be confounded. We controlled for this with the inclusion of regional dummies in the regression. In addition, we conducted robustness check 5, which examines the gender discrimination ratio at each region that has metropolitan cities.

5.2. Robustness Check 1: Samples from Rural Counties Only

China is a unique market, where, in 2014, rural areas had an extraordinarily high mobile Internet penetration rate of 84.6%; moreover, in 2014, 84.4% of rural residents preferred to shop online and spent RMB 2,000 (USD 300) on average.%[14] Families in rural counties might rely more on Internet shopping to purchase children's clothing, because physical children's apparel stores in rural areas are less accessible and convenient. Of course, street bazaars in villages are a common off-line option, but their selection was not comparable to the online offerings. One may wonder if rural families might happen to like the boy brand more or if obtaining girls' clothes was relatively easier from the street bazaars.

To address this concern, we applied the same analysis at the customer level in rural counties, and the results from this subsample (combined samples A and B with a sample size of 16,798 customers) are shown in Table 10. Again, when using expenditure as the dependent variable, we found that families from more educated areas were less likely to be discriminatory toward girls ($B = -0.17$, $p < 0.05$), and we found the same for the districts with higher birth rates ($B = -0.05$, $p < 0.05$). Although, for this particular robustness check, the economic development variables (as GDP was correlated with rural counties) were not significant using the incremental measure, the signs were consistent with our predictions.

5.3. Robustness Check 2: Removing Bottom Districts with Fewer Customer Representatives

We tried to address the concern that some districts with few customer representatives may bias the obtained results. The average number of households in each district was about 16, with 2% of the districts having more than 100 customers. We trimmed our data by removing the bottom 10%, 20%, and 30% samples based on the number of customers aggregated to a district and then ran the main regression analysis three times, one for each subsample. The results are shown in Table 11 with all three ratios of gender discrimination (*Expenditure*, *Quantity*, and *Number of orders*) as the dependent variables, as we did in the main regression analysis (shown in Table 9). The completely aligned results that we obtained, shown in Tables 9 and 11, helped enhance the reliability of our findings.

5.4. Robustness Check 3: Incremental Measure of Gender Discrimination Controlling Unobserved Local Factors

We constructed an additional and incremental gender discrimination ratio, with the nominator being the expenditure for boys from families with both boy and girl, over the expenditure for boys from boy-only families; and the denominator being the expenditure for girls from families with both boy and girl over the expenditure for girls from girl-only families. We also created similar measures for quantity and number of orders following this operationalization. In that way, we got a cleaner and tighter measure of the incremental gender effect in families with both boy and girl over the families with children of the same gender, while controlling for unobserved local factors.

Note that the new measure could also be written as [(the expenditure for boys from families with both boy and girl)/(the expenditure for girls from families with both boy and girl)] × [(the expenditure for girls from girl-only families/the expenditure for boys from boy-only families)], which is also equal to our

Table 10. OLS Results for Robustness Check 1: Customers from Rural Counties (Combined Samples A and B)

	Robustness Check 1 with ratio of gender discrimination (*Expenditure*) as DV (customer-level data[a] for rural counties only)		Robustness Check 1 with ratio of gender discrimination (*Quantity*) as DV (customer-level data[a] for rural counties only)		Robustness Check 1 with ratio of gender discrimination (*Number of orders*) as DV (customer-level data[a] for rural counties only)	
	Within-subject (families with both boys and girls) comparison: Expenditure on boy clothing vs. expenditure on girl clothing		Within-subject (families with both boys and girls) comparison: Quantity of boy clothing vs. quantity of girl clothing		Within-subject (families with both boys and girls) comparison: Orders of boy clothing vs. orders of girl clothing	
Variables	*B*	*t*-value	*B*	*t*-value	*B*	*t*-value
Log (GDP)	−0.02	−0.30	−0.04	−1.18	−0.02	−1.19
Average education (years)	−0.17*	−2.59	−0.01	−0.19	0.01	0.45
Birth rate	−0.05*	−3.52	−0.02*	−1.99	−0.01*	−3.16
Sample[b]	0.16	1.27	0.15*	2.36	0.06*	2.41
Covariates[a]						
R^2	6.89%		14.99%		20.43%	

[a]Covariates consisted of male-female ratio, percentage of minority, region, off-line shopping (Balala Children's Clothing Company), e-commerce development index, percentage of fertile women, and percentage of children. In this analysis, since we used customer-level data, we also included promotion intensity, number of orders, average product quantity per order, and average order price as covariates.

[b] Sample B was the reference group.

*$p < 0.05$; $N = 16{,}798$.

Table 11. Robustness Check 2: OLS Results for Main Regression Analyses (Combined Samples A and B) with Bottom 10%, 20%, and 30% of Districts with Fewer Customer Representatives Removed

	Main regression analysis with ratio of gender discrimination (*Expenditure*) as DV (district-level data)		Main regression analysis with ratio of gender discrimination (*Quantity*) as DV (district-level data)		Main regression analysis with ratio of gender discrimination (*Number of orders*) as DV (district-level data)	
	Within-subject (families with both boys and girls) comparison: Expenditure on boy clothing vs. expenditure on girl clothing					
Variables (bottom 10% removed: N = 4,299)	*B*	*t*-value	*B*	*t*-value	*B*	*t*-value
Log (GDP)	−0.13*	−4.45	−0.09*	−4.40	−0.02*	−2.52
Average education (years)	−0.06*	−2.43	−0.05*	−2.64	−0.01	−1.42
Birth rate	−0.04*	−3.67	−0.02*	−3.20	−0.01*	−2.07
Sample[a]	0.41*	11.22	0.63*	24.08	0.31*	30.42
Covariates[b]						
R^2	5.20%		13.85%		18.82%	
Variables (bottom 20% removed: N = 3,934)	*B*	*t*-value	*B*	*t*-value	*B*	*t*-value
Log (GDP)	−0.11*	−3.91	−0.08*	−3.73	−0.02*	−2.21
Average education (years)	−0.08*	−2.94	−0.06*	−3.24	−0.02*	−2.43
Birth rate	−0.03*	−3.26	−0.02*	−3.46	−0.01*	−2.30
Sample[a]	0.41*	11.61	0.63*	24.81	0.32*	30.96
Covariates[b]						
R^2	5.65%		15.67%		21.55%	
Variables (bottom 30% removed: N = 3,586)	*B*	*t*-value	*B*	*t*-value	*B*	*t*-value
Log (GDP)	−0.10*	−3.88	−0.07*	−3.43	−0.02*	−2.38
Average education (years)	−0.04	−1.81	−0.04*	−2.30	−0.01	−1.41
Birth rate	−0.02*	−2.41	−0.02*	−2.54	−0.00	−1.67
Sample[a]	0.41*	12.10	0.63*	26.12	0.31*	31.03
Covariates[b]						
R^2	5.93%		17.91%		23.60%	

[a] Sample B was the reference group.
[b] Covariates consisted of cities levels (other cities and rural cities with metropolitan cities as the reference group), male-female ratio, percentage of minority, region, off-line shopping (Balala Children's Clothing Company), e-commerce development index, percentage of fertile women, and percentage of children.
*$p < 0.05$.

previous DV weighted by the inverse of expenditure ratio from same-gender families. The same applied to the quantity and number of orders measures.

As shown in Table 12, the overall ratios (expenditures) were 1.89 for sample A and 2.17 for sample B. The expenditure ratios for metropolitan cities, other cities, and rural counties were 1.56, 1.80, and 2.01, respectively for sample A, and 1.51, 1.95, and 2.54, respectively, for sample B. The correlation between the new measure and the original one was also significant and positive (sample A: 0.93; sample B: 0.94). We also found similar patterns for quantity and number of orders, as shown in Table 12. All the summary statistics of our incremental measures suggested good validity.

Once we combined samples A and B, we performed the district-level analyses again and found that gender discrimination was negatively correlated with economic development ($B = -0.16$, $p < 0.05$) as well as

Table 12. Robustness Check 3: Eliminating Local Confounding Factors—Descriptive Statistics of Incremental Gender Discrimination Ratio

Operationalization	City level	Mean	Difference	t-value[a]	95% Confidence interval of the difference: Lower	95% Confidence interval of the difference: Upper
Company A: Alternative gender discrimination ratio (*Expenditure*)[c]	Metropolitan cities	1.56[b]				
	Other cities	1.80	0.23	1.92	−0.01	0.48
	Rural counties	2.01	0.45	3.33	0.18	0.71
	Total	**1.89**				
Company A: Alternative gender discrimination ratio (*Quantity*)[c]	Metropolitan cities	1.33[b]				
	Other cities	1.48	0.15	2.12	0.01	0.30
	Rural counties	1.59	0.26	3.33	0.11	0.42
	Total	**1.53**				
Company A: Alternative gender discrimination ratio (*Order*)[c]	Metropolitan cities	1.17[b]				
	Other cities	1.23	0.06	1.96	−0.00	0.12
	Rural counties	1.25	0.08	2.35	0.01	0.14
	Total	**1.24**				
Company A[d]: Correlations between alternative operationalization and original ratio of gender discrimination (*Expenditure*)				0.93**		
Company A[d]: Correlations between alternative operationalization and original ratio of gender discrimination (*Quantity*)				0.90**		
Company A[d]: Correlations between alternative operationalization and original ratio of gender discrimination (*Number of orders*)				0.93**		
Company B: Alternative gender discrimination ratio (*Expenditure*)[c]	Metropolitan cities	1.51[b]				
	Other cities	1.95	0.44	3.00	0.15	0.73
	Rural counties	2.54	1.03	3.49	0.45	1.61
	Total	**2.17**				
Company B: Alternative gender discrimination ratio (*Quantity*)[c]	Metropolitan cities	1.43[b]				
	Other cities	1.82	0.39	3.60	0.18	0.60
	Rural counties	2.20	0.77	5.37	0.49	1.05
	Total	**2.05**				
Company B: Alternative gender discrimination ratio (*Number of orders*)[c]	Metropolitan cities	0.99[b]				
	Other cities	1.01	0.02	0.89	−0.02	0.05
	Rural counties	1.02	0.03	1.42	−0.01	0.06
	Total	**1.02**				
Company B[e]: Correlations between alternative operationalization and original ratio of gender discrimination (*Expenditure*)				0.94**		
Company B[e]: Correlations between alternative operationalization and original ratio of gender discrimination (*Quantity*)				0.88**		
Company B[e]: Correlations between alternative operationalization and original ratio of gender discrimination (*Order*)				0.92**		

[a] Two-tailed independent sample t-test. The tests did not assume equal variances.
[b] The ratio of metropolitan cities was the reference group.
[c] The operationalization: (expenditure or quantity or number of orders for boys from families with both boys and girls/expenditure or quantity or number of orders for boys from boy-only families)/(expenditure or quantity or number of orders for girls from families with both boys and girls/expenditure or quantity or number of orders for girls from girl-only families). We aggregated the data to the district level. Company A sample = 2,453.
[e] Company B sample = 2,479.
**Correlation was significant at the 0.01 level (two-tailed).

birth rate ($B = -0.02$, $p < 0.05$), as shown in Table 13. We were unable to replicate the result for education this time; however, its sign was consistent with our main results. Similar patterns were found when using quantity and number of orders as the dependent variable, as shown in Table 14. Overall, we believe that the incremental measure provided additional robustness to the main regression analyses.

5.5. Robustness Check 4: Gender Discrimination vs. Birth Order Favoritism

In order to address the birth order concern, we need to answer the following question: What is the impact of having a second child on the firstborn child? Some might argue that there is favoritism toward the younger child (or older child) rather than toward the boy.

Using size as a screening variable for two-children families, we were left with 887 districts for the GG versus GB comparison, and 723 districts for the GB versus BG comparison. Note that the GB families in the two sets of samples were slightly different, as the number of BG families was smaller than that of the GG families.

We compared the ratio in *Expenditure*, *Quantity*, and *Number of orders* of the second-born over the first-born in the GG, GB, and BG families at the district level. As shown in Table 14, the paired-sample *t*-tests were all significant. The ratios of expenditure for the second born versus first born in the GB family were larger than those in the GG family (B/G in GB vs. G/G in GG) and larger than those in the BG family (B/G in GB vs. G/B in BG), indicating a stronger level of favoritism toward the boys, regardless of the birth order. In spite of the potential pass-on in the GG families, all the robustness checks so far consistently suggested compelling evidence that there was stronger favoritism toward boys. Unfortunately, clothing size information was only available to us in sample A and not in Sample B. Thus, we only implemented this robustness check using sample A, and we were able to replicate our results.

5.6. Robustness Check 5: Demand Across Regions

Given the wide landscape and climate variations of China, certain product categories might be purchased differently across regions. The concern here was whether there were regional demand side factors that drove the differences in category popularity in rural areas versus urban areas. For example, expensive coats might be more popular in the North, where there are more rural areas, whereas cheap T-shirts might be more popular in the South, where there are more urban areas.

We controlled for this in the main regression analysis by including regional dummies. In addition, we conducted an additional robustness check to further control for demand difference in rural cities versus metropolitan cities across regions.

Using customer-level data, we split the sample based on a district's region and whether it was a metropolitan city. For example, Beijing, as a metropolitan city, would be compared with the regional average of the North, Shanghai with the East, and Guangzhou and Shenzhen with the South. Then, we calculated the average expenditure, quantity, and order ratio for boy clothing and girl clothing and took the gender discrimination ratios. As shown in Table 15, regardless of regions, gender discrimination ratios were consistently lower in metropolitan cities as compared with their regional average except for Guangzhou, which was well known in China to have a long cultural history of

Table 13. Robustness Check 3: Eliminating Local Confounding Factors—OLS Results (Combined Samples A and B) Using Incremental Gender Discrimination Ratio

	The alternative operationalization as DV: *Expenditure* (district-level data)		The alternative operationalization as DV: *Quantity* (district-level data)		The alternative operationalization as DV: *Number of orders* (district-level data)	
Variables	B	*t*-value	B	*t*-value	B	*t*-value
Log (GDP)	−0.16*	−4.69	−0.14*	−4.67	−0.01	−1.36
Average education (years)	−0.05	−1.54	−0.01	−0.25	−0.01	−1.09
Birth rate	−0.02*	−2.11	−0.01	−1.14	−0.00	−1.48
Sample[a]	−0.02	−0.47	−0.22*	−5.88	0.20*	18.85
Covariates[b]						
R^2	2.24%		2.15%		8.23%	

[a] Sample B was the reference group.
[b] Covariates consisted of city levels (other cities and rural cities, with metropolitan cities as the reference group), male-female ratio, percentage of minority, region, off-line shopping (Balala Children's Clothing Company), e-commerce development index, percentage of fertile women, and percentage of children.
*$p < 0.05$; $N = 4{,}407$. The sample size was reduced from 4,932 to 4,407, because we included log (GDP) in the model and we were unable to retrieve information on GDP for some cities.

Table 14. Robustness Check 4 (Sample A): Gender Discrimination vs. Birth Order Favoritism—Ratio Comparisons[a] for Second-Born vs. First-Born Between Girl-Girl (GG) Families[b] and Girl-Boy (GB) Families,[b] and Between GB Families and Boy-Girl (BG) Families[b]

					95% Confidence interval	
Gender discrimination ratio	GB families	GG families	Mean difference	t-value	Lower	Upper
Expenditure[c]	1.57	1.09	0.48	6.47	0.34	0.63
Quantity[c]	1.54	1.11	0.43	7.87	0.33	0.54
Number of orders[c]	1.55	1.10	0.45	8.04	0.34	0.56
					95% Confidence interval	
Gender discrimination ratio	GB families	BG families	Mean difference	t-value	Lower	Upper
Expenditure[c]	1.59	1.04	0.55	6.44	0.38	0.71
Quantity[c]	1.56	1.12	0.44	6.25	0.30	0.57
Number of orders[c]	1.56	1.04	0.52	8.27	0.40	0.64

[a]We first aggregated the ratio of expenditure between the second born vs. first born for GG families, GB families, and BG families to the district level and then conducted two paired-sample t-tests using these district-level data. We used one to compare GB families and GG families (N = 887) and the other to compare GB families and BG families (N = 723).

[b]GG families were those with the first-born child being a girl and the second-born child being a girl as well. GB families were those with the first-born child being a girl and the second-born child being a boy. BG families were those with the first-born child being a boy and the second-born child being a girl.

[c]Expenditure of GB families: expenditure for boy (the second born)/expenditure for girl (the first born). Expenditure of GG families: expenditure for girl (the second born)/expenditure for girl (the first born). Expenditure of BG families: expenditure for girl (the second born)/expenditure for boy (the first born). Quantity of GB families: quantity purchased for boy (the second born)/quantity purchased for girl (the first born). Quantity of GG families: quantity purchased for girl (the second born)/quantity purchased for girl (the first born). Quantity of BG families: quantity purchased for girl (the second born)/quantity purchased for boy (the first born). Order of GB families: orders purchased for boy (the second born)/orders purchased for girl (the first born). Order of GG families: orders purchased for girl (the second born)/orders purchased for girl (the first born). Order of BG families: orders purchased for girl (the second born)/orders purchased for boy (the first born).

favoritism toward boys%[15] and showed a slightly higher expenditure ratio than that in the South region. The relative difference in metropolitan cities versus the rest of the region seemed to be largest in the North, which were less developed than the East and South regions of China. In Beijing, a northern capital city in China, we even observed in sample B that parents were spending more on girls than on boys.

6. Additional Analysis: Implications from the One-Child Policy in Policy-Restricted Areas vs. Nonrestricted Areas

The one-child policy was imposed across China from the late 1970s to 2015; however, there were a few exceptions. Four areas in Mainland China (i.e., rural counties in Chengde, Jiuquan, Linfen, and Enshi) and two special administrative regions, Hong Kong (HK) and Macau, were not subject to the one-child policy. Specifically, for the four areas in Mainland China selected by the Chinese State Family Planning Commission, regardless of the first child's gender, families could bear a second child (the two-child policy). As for the special administrative regions, families did not have any restrictions on the number of children that they could have. For our analysis, we also added Taiwan (TW) to the latter group (no restrictions), which also enjoyed a higher level of economic development as compared with most parts of Mainland China. We anticipated lower ratios of child gender discrimination in the nonrestricted regions than in the policy-restricted areas. In fact, we contrasted the ratios for these three types of regions (shown in Table 16), combining both samples A and B, and found the expected results: the ratios of boy-girl discrimination (*Expenditure*) in the nonrestricted areas were significantly lower than in the policy-restricted areas ($\text{Mean}_{\text{policy-restricted areas}} = 1.98$, $\text{Mean}_{\text{non-policy-restricted areas in mainland}} = 1.43$, t-value = 3.21, $\text{Mean}_{\text{HK,Macau,TW}} = 1.51$, t-value = 2.30). Similar results were revealed when *Quantity* and *Number of orders* were tested. Thus, we concluded that the one-child policy (low birth rate) was a salient factor associated with gender discrimination.

As a final remark, the fact that the four experimental rural areas, as compared with other rural areas, did not show higher ratios further proved that usage rate in rural areas was not a main driver for gender discrimination.

With these analyses and robustness checks, we are convinced that boy-girl discrimination still existed in China during the time period studied and that the degree of gender bias varied across socioeconomic factors. Our results, complementing Guiso et al. (2008)'s finding, show that better economic conditions, better education, and higher birth rates were some of the factors that diminished boy-girl discrimination in consumption.

Table 15. Robustness Check 5: Gender Discrimination Ratios by Regions and by Metropolitan Cities[a] (Samples A and B)

Sample	North (excluding Beijing)			Beijing			South (excluding Guangdong and Shenzhen)			Guangzhou			Shenzhen		
	Expenditure	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*
Sample A	1.68	1.55	1.26	1.39	1.23	0.99	1.36	1.28	1.10	1.38	1.35	1.13	1.32	1.25	1.08
Sample B	1.11	0.86	0.96	0.97	0.76	0.93	1.11	0.84	0.95	1.02	0.78	0.95	1.11	0.90	0.93

Sample	East (excluding Shanghai)			Shanghai			West		
	Expenditure	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*	*Expenditure*	*Quantity*	*Number of orders*
Sample A	1.50	1.42	1.21	1.41	1.36	1.16	1.50	1.43	1.21
Sample B	1.19	0.92	0.96	1.08	0.85	0.93	1.31	0.98	0.95

[a]Using customer-level data, we split the sample based on a district's region and whether it was a metropolitan city. Then, we calculated the average expenditure, quantity, and order for boy clothing and girl clothing, and took the gender discrimination ratios across the aforementioned segments using the calculated averages.

7. Discussion and Conclusion

Discrimination against girls is universally regarded as socially unacceptable, and yet it is still very prevalent worldwide. As stated in a recent report by Save the Children, 30% of countries are characterized by discrimination against girls (55 of 185 countries).%[16] Sociologists worry that pervasive girl discrimination within households could potentially transcend to a female-unfriendly society and create further gender frictions in the workplace. Business communities certainly cannot ignore this threat, as they have been working hard to promote and comply with gender-equal work environments.

The actual acts of discrimination against girls are, unfortunately, hard to detect, because they are done behind closed doors and unobservable to outsiders. Also, as Deaton (1989) mentioned, the ability to detect boy-girl child discrimination is hampered by a lack of data on actual intrahousehold resource allocations. Hence, our study in itself is significant, because it is the first large-scale empirical work to clearly show the phenomenon of boy-girl discrimination, taking advantage of e-commerce data.

On boy-girl discrimination, the following is from *The Book of Songs*, a collection of ancient Chinese poetry (1000–700 BC) (McNaughton 1971):

When a son is born,
Let him sleep on the bed,
Wrap him with fine clothes,
And give him jade to play …
When a daughter is born,
Let her sleep on the ground,
Clothe her in plain swaddle,
And give her cotton spinning wheel to play …

Fortunately, our study shows that the degree of discrimination diminishes as economic development, community openness, and level of education increase. In other words, as socioeconomic conditions of a society continue to improve, discrimination will likely gradually subside and hopefully disappear altogether.

In summary, we found the following:

- Families in economically less-developed areas and rural areas were more likely to show boy-girl discrimination tendency compared with those living in more prosperous cities.
- The expenditure difference was largely due to the fact that rural parents were more likely to buy more new clothes for boys than for girls compared with their peers in urban areas.
- Higher education and birth rate could reduce this discrimination.

Table 16. Additional Analysis: Ratio of Gender Discrimination Between Policy-Restricted Areas and Non-Policy-Restricted Areas (Combined Samples A and B)

	City level	Mean	Difference	t-value[a]	95% Confidence interval of the difference: Lower	Upper
Ratio of gender discrimination (*Expenditure*)	Policy-restricted areas	1.98[b]				
	Non-policy-restricted areas in Mainland China[c]	1.43	0.55	3.21	0.20	0.90
	HK, Macau, and TW	1.51	0.46	2.30	0.05	0.88
Ratio of gender discrimination (*Quantity*)	Policy-restricted areas	1.46[a]				
	Non-policy-restricted areas in Mainland China[c]	1.07	0.39	3.59	0.17	0.62
	HK, Macau, and TW	1.24	0.22	1.10	−0.19	0.63
Ratio of gender discrimination (*Number of orders*)	Policy-restricted areas	1.16[a]				
	Non-policy-restricted areas in Mainland China[c]	1.15	0.01	0.14	−0.13	0.15
	HK, Macau, and TW	0.95	0.21	5.78	0.13	0.28

[a] Two-tailed independent sample *t*-test. The tests did not assume equal variances.
[b] Policy-restricted areas were the reference group.
[c] Non-policy-restricted areas in mainland china included Chengde, Jiuquan, Linfen, and Enshi.

- The newly less-restricted population-control policy is expected to reduce the degree of discrimination, if it can indeed promote higher birth rates.

Our analysis of marketing data related to e-commerce purchases of children's clothing reveals the existence of the undesirable social behavior of parental discrimination against girls, particularly in less-developed rural areas of China. This may have practical implications for companies looking to design corporate initiatives, such as CSR programs, that can help educate the public and mitigate this problem.

Like their western counterparts, many Chinese companies are now aware of the importance of CSR, as the Chinese government is also putting pressure on businesses and society to comply with responsible and ethical business policies. Between 2010 and 2018, China dropped from 61st (among 134 countries) to 103rd (among 149 countries) in the World Economic Forum's Gender Gap Report.%[17] Economic disparities between the sexes tend to narrow as countries grow richer (*The Economist* 2019). To market in these rapid-developing emerging markets, companies should seek opportunities to carry out cause-related marketing or CSR initiatives to educate families about the importance and benefits of treating children of both genders equally. China's geographically widespread provinces and regions display cultural differences, even while sharing some cultural roots. Combining these local cultural variations with the different organizational cultures of companies, it is understandable that the notion of CSR in China faces more challenges; companies probably need to embrace a tailored approach based on the interface of three dimensions: customer segmentation, regional idiosyncrasy, and economic development—as illustrated by our study.

Echoing the recommendations made by the #SaveTheChirdren report, our results suggest that companies should do the following: (1) Invest in achieving gender equality, including increasing expenditures and monitoring budgets designed to close gender gaps and increase access to basic services and empowerment programs, especially in rural, marginalized, and vulnerable populations. For example, #UnitedbyHalf, is a campaign promoting gender equality in India, the second largest market for United Colors of Benetton. The company's long-term Benetton Women Empowerment Program quickly opened its previously male customer-targeted brand to female consumers. (2) Raise awareness in advertising campaigns. The issue of gender equality has been a key theme at the Cannes Lions Festival for several years, with a focus on eliminating the objectification of women and girls portrayed in advertising and increasing the number of women in the higher echelons of the greater advertising and marketing workplace. For example, the Cannes Glass Lion Award winner, Whisper's "Touch the Pickle" sanitary napkins campaign, aims to break menstruation taboos of "not touching the pickle" in India. According to AdAge, more than 2.9 million women pledged to "touch the pickle jar" after seeing the ad, and Whisper's share of voice grew from 21% to 91% in its category.

In summary, the contributions of this study are twofold. First, as noted in marketing communities, the strategy of customer segmenting and targeting, which has worked well for exploring new business opportunities, can be equally useful when developing innovative CSR campaigns. Our study demonstrates that today's abundant marketing data obtained by companies through online and mobile e-commerce and other activities can be a fertile source for uncovering social causes that would otherwise remain subtle or hidden. Second, on the issue of discrimination against girls, though it is universally considered unacceptable, it is difficult to document, let alone verify, its presence. This study is the first to investigate the phenomenon on a large scale and statistically substantiate its existence in China.

There are a few caveats to address. Boy-girl discrimination is a complex issue. Discretionary parental actions on behalf of their children are motivated by both self-interest and altruism. What we discovered in the children's clothing category is just a piece of corroborating evidence for such actions. Ideally, other discretionary expenditure categories, children's toys or books, for example, should be examined concurrently. Unfortunately, these data were not readily accessible to the authors. Furthermore, though the purchase data that we examined are at the household level, we do not have household-specific data. To take advantage of the statistics gathered from the Chinese Bureau of Statistics, data were aggregated, and analyses were carried out at the district level. Thus, based on our findings, we cannot infer or suggest any possible reasons or motives for parental boy-girl discrimination on nonessential expenditures. Moreover, though we tried our best to control for off-line options, obtaining complete information for the competitive landscape is always a challenge. One future research direction is to model and analyze the behavior at the household level, provided that household-specific information is available or can be properly inferred through other measurable proxies.

Acknowledgments

The authors thank seminar participants at China Europe International Business School, NYU Shanghai, Fudan University, and Tsinghua University for helpful comments.

Endnotes

[1] See https://www.prnewswire.com/news-releases/sk-iis-ineverexpire-campaign-inspires-women-to-challenge-age-related-pressure-300639059.html.

[2] See https://econsultancy.com/blog/67626-17-marketing-campaigns-with-a-positive-message-for-women.

[3] See https://www.nhlbi.nih.gov/health/educational/wecan/downloads/calreqtips.pdf.

[4] See http://www.zhongbangshuju.com/viewdoc?eid=4248C8491C09A07A.

[5] The number one company, Balabala, has only a 3.1% market share. For comparison, the number one brand in the United States, Carters, has a 12% market share.

[6] http://www.100ec.cn/detail–6438314.html

[7] Census data in China is collected every 10 years.

[8] See http://www.aliresearch.com/html/stopic/aedi/about.html.

[9] The results did not change when separating samples A and B.

[10] In that case, the migrant population was not included in the denominator, even though the contribution of migrants is included in the numerator.

[11] Metropolitan cities, or first-tier cities, are administrative districts of Beijing, Shanghai, Guangzhou, and Shenzhen, whereas rural counties are county-governed districts. Other cities are the rest of the city-governed districts in China.

[12] Sample B was the reference group. Our results did not change with two separate regressions for samples A and B.

[13] See Online Appendix A6 for the detail composition of family types across city tiers.

[14] See https://www.forbes.com/sites/ceibs/2014/11/10/mobile-and-rural-dual-engines-for-alibabas-future/.

[15] Guangzhou is predominantly Cantonese speaking with deep roots in local culture and tradition, whereas the newly industrialized areas (many are in rural counties in the South) are populated with Mandarin-speaking migrant workers and young professionals.

[16] See "The many faces of exclusion" (https://www.savethechildren.org/content/dam/global/reports/2018-end-of-childhood-report.pdf).

[17] See https://www.livescience.com/18573-countries-gender-equality-ranking.html.

References

Angrist J (2002) How do sex ratios affect marriage and labor markets? Evidence from America's second generation. *Quart. J. Econom.* 117(3):997–1038.

Antman FM (2011) International migration and gender discrimination among children left behind. *Amer. Econom. Rev.* 101(3):645–649.

Asadullah MN, Chaudhury N (2009) Reverse gender gap in schooling in Bangladesh: Insights from urban and rural households. *J. Development Stud.* 45(8):1360–1380.

Avraham YE (2011) Estimating a dynamic model of sex selection in China. *Demography* 48(2):783–811.

Baculinao E (2004) China grapples with legacy of its "missing girls." *NBC News* (September 14), http://www.nbcnews.com/id/5953508/ns/world_news/t/china-grapples-legacy-its-missing-girls/.

Barcellos SH, Carvalho LS, Lleras-Muney A (2014) Child gender and parental investments in India: Are boys and girls treated differently? *Amer. Econom. J. Appl. Econom.* 6(1):157–189.

Begum L, Grossman PJ, Islam A (2014) Parental attitude and investment in children's education and health in developing countries. Discussion paper 30/14, Department of Economics, Monash University, Melbourne, Australia.

Behrman JR (1988) Intra-household allocation of nutrients in rural India: Are boys favored? Do parents exhibit inequality aversion? *Oxford Econom. Papers* 40(1):32–54.

Ben-Porath Y, Welch F (1976) Do sex preferences really matter? *Quart. J. Econom.* 90(2):285–307.

Bhalotra S, Attfield C (1998) Intra-household resource allocation in rural Pakistan: A semiparametric analysis. *J. Appl. Econometrics* 13(5):463–480.

Blau FD, Kahn LM (1994) Rising wage inequality and the US gender gap. *Amer. Econom. Rev.* 84(2):23–28.

Boll T, Ferring D, Filipp SH (2003) Perceived parental differential treatment in middle adulthood: Curvilinear relations with individuals' experienced relationship quality to sibling and parents. *J. Family Psych.* 17(4):472–487.

Brown TJ, Dacin PA (1997) The company and the product: Corporate associations and consumer product responses. *J. Marketing* 61(1):68–84.

Bulte E, Heerink N, Zhang X (2011) China's one-child policy and "the mystery of missing women": Ethnic minorities and male-biased sex ratios. *Oxford Bull. Econom. Statist.* 73(1):21–39.

Busse MR, Israeli A, Zettelmeyer F (2016) Repairing the damage: The effect of gender and price knowledge on auto-Repair price quotes. *J. Marketing Res.* 54(1):75–95.

Castaldo S, Perrini F, Misani N, Tencati A (2009) The missing link between corporate social responsibility and consumer trust: The case of fair trade products. *J. Bus. Ethics* 84(1):1–15.

Chen D, Li X, Lai F (2017) Gender discrimination in online peer-to-peer credit lending: Evidence from lending platform in China. *Electronic Commerce Res.* 17(4):553–583.

Chen Y, Yang S, Zhao Y (2008) A simultaneous model of consumer brand choice and negotiated price. *Management Sci.* 54(3):538–549.

Daily Mail (2016) September 11.

Deaton A (1989) Looking for boy-girl discrimination in household expenditure data. *World Bank Econom. Rev.* 3(1):1–5.

Derdeyn AP (1979) Adoption and the ownership of children. *Child Psychiatry Human Development* 9(4):215–226.

Ding WW, Murray F, Stuart TE (2013) From bench to board: Gender differences in university scientists' participation in corporate scientific advisory boards. *Acad. Management J.* 56(5):1443–1464.

Feinberg M, Neiderhiser J, Howe G, Hetherington EM (2001) Adolescent, parent, and observer perceptions of parenting: Genetic and environmental influences on shared and distinct perceptions. *Child Development* 72(4):1266–1284.

Financial Times (2010) October 18.

Garg A, Morduch J (1998) Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *J. Population Econom.* 11(4):471–493.

Gibson J (1997) Testing for boy-girl discrimination with household expenditure data: Results for Papua New Guinea. *Appl. Econom. Lett.* 4(10):643–646.

Gibson J, Rozelle S (2004) Is it better to be a boy? A disaggregated outlay equivalent analysis of gender bias in Papua New Guinea. *J. Development Stud.* 40(4):115–136.

Gilligan M, Suitor JJ, Seoyoun K, Pillemer K (2013) Differential effects of perceptions of mothers' and fathers' favoritism on sibling tension in adulthood. *J. Gerontology B Psych. Sci. Soc. Sci.* 68(4):593–598.

Ginther DK, Hayes KJ (1999) Gender differences in salary and promotion in the humanities. *Amer. Econom. Rev.* 89(2):397–402.

Gobillon L, Meurs D, Roux S (2015) Estimating gender differences in access to jobs. *J. Labor Econom.* 33(2):317–363.

Goldin C, Polachek S (1987) Residual differences by sex: Perspectives on the gender gap in earnings. *Amer. Econom. Rev.* 77(2): 143–151.

Gong X, van Soest A, Zhang P (2005) The effects of the gender of children on expenditure patterns in rural China: A semiparametric analysis. *J. Appl. Econometrics* 20(4):509–527.

Guerrera F and Rappeport A (2018) Glass ceiling' persists in US boardrooms. *Financial Times* (October 18), https://www.ft.com/content/942177a2-da1a-11df-bdd7-00144feabdc0.

Guiso L, Monte F, Sapienza P (2008) Culture, gender, and math. *Science* 320(5880):1164–1165.

Gupta MD, Bhat PNM (1997) Fertility decline and increased manifestation of sex bias. *Population Stud.* 51(3):307–315.

Haddad L, Hoddinott J (1994) Women's income and boy-girl anthropometric status in the Côte d'Ivoire. *World Development* 22(4):543–553.

Haddad L, Reardon T (1993) Gender bias in the allocation of resources within households in Burkina Faso: A disaggregated outlay equivalent analysis. *J. Development Stud.* 29(2):260–276.

Hazarika G (2000) Gender differences in children's nutrition and access to healthcare in Pakistan. *J. Development Stud.* 37(1):73–92.

Jensen R (2012) Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *Quart. J. Econom.* 127(2):753–792.

Lancaster G, Maitra P, Ray R (2008) Household expenditure patterns and gender bias: Evidence from selected Indian states. *Oxford Development Stud.* 36(2):133–157.

Li S (2007) Imbalanced sex ratio at birth and comprehensive interventions in China. *Fourth Asia Pacific Conf. Reproductive Sex Health Rights*, *Hyderabad, India*, October 29–31.

Luo X, Bhattacharya CB (2006) Corporate social responsibility, customer satisfaction, and market value. *J. Marketing* 70(4):1–18.

Luo X, Bhattacharya CB (2009) The debate over doing good: Corporate social performance, strategic marketing levers, and firm-idiosyncratic risk. *J. Marketing* 73(6):198–213.

McDowell JM, Singell LD Jr, Ziliak JP (1999) Cracks in the glass ceiling: Gender and promotion in the economics profession. *Amer. Econom. Rev.* 1999(1):392–396.

McNaughton W (1971) *The Book of Songs* (Twayne Publishers, Inc., New York).

Morduch JJ, Stern HS (1997) Using mixture models to detect sex bias in health outcomes in Bangladesh. *J. Econometrics* 77(1): 259–276.

New York City Department of Consumer Affairs (2015) https://www1.nyc.gov/site/dca/partners/gender-pricing-study.page.

Song L (2008) In search of gender bias in household resource allocation in rural China. Working paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1136265.

Spears CB, Bigler RS (2005) Children's perceptions of discrimination: A developmental model. *Child Development* 76(3):533–553.

Subramaniam R (1996) Gender-bias in India: The importance of household fixed-effects. *Oxford Econom. Papers* 48(2):280–299.

Suitor JJ, Sechrist J, Plikuhn M, Pardo ST, Pillemer K (2008) Within-family differences in parent-child relations across the life course. *Current Directions Psych. Sci.* 17(5):334–338.

The Economist (2019) May 18.

Wei SJ, Zhang XB (2011) The competitive saving motive: Evidence from rising sex ratios and savings rates in China. *J. Political Econom.* 119(3):511–564.

Zhang Y, Voorheesb CM, Lin C , Chiang J, Hult T, Calantonee RJ (2021) Information search and product returns across mobile and traditional online channels. *J. Retailing*. Forthcoming.

Zhuhai (2019) The daughter rose in the East. *The Economist* (May 18), https://www.economist.com/special-report/2019-05-18.

Zimmermann L (2012) Reconsidering gender bias in intrahousehold allocation in India. *J. Development Stud.* 48(1):151–163.