INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

IFPRI

IFPRI Discussion Paper 01540

June 2016

Does Female Labor Scarcity Encourage Innovation?

Evidence from China's Gender Imbalance

Zhibo Tan

Xiaobo Zhang

Development Strategy and Governance Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), established in 1975, provides evidence-based policy solutions to sustainably end hunger and malnutrition and reduce poverty. The Institute conducts research, communicates results, optimizes partnerships, and builds capacity to ensure sustainable food production, promote healthy food systems, improve markets and trade, transform agriculture, build resilience, and strengthen institutions and governance. Gender is considered in all of the Institute's work. IFPRI collaborates with partners around the world, including development implementers, public institutions, the private sector, and farmers' organizations, to ensure that local, national, regional, and global food policies are based on evidence. IFPRI is a member of the CGIAR Consortium.

AUTHORS

Zhibo Tan (tzb0905@fudan.edu.cn) is an assistant professor at the School of Economics at Fudan University, Shanghai.

Xiaobo Zhang (x.zhang@cgiar.org) is a senior research fellow in the Development Strategy and Governance Division of the International Food Policy Research Institute, Washington, DC. Xiaobo is also a professor of economics at the National School of Development, Peking University, Beijing.

Notices

[1.] IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by the International Food Policy Research Institute.

[2.] The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

Contents

Tables

Figures

ABSTRACT

Facing scarcity of a production factor, a firm can develop technologies to either substitute the scarce factor (price effect) or complement the more abundant factors (market size effect). Whether the market size effect or the price effect dominates largely depends on the elasticity of substitution among factors according to the theory of directed technical change. However, it is a great challenge to empirically test the theory because factor prices are often endogenously determined. In this paper, we use imbalanced sex ratios across Chinese provinces as a source of identification strategy to test how female labor scarcity affects corporate innovation based on the matched dataset of annual surveys of industrial firms in China and the national patent database. In regions with a large male population, female-intensive industries face more serious problems finding female workers than their male-intensive counterparts. We find that such female shortages have spurred firms in female-intensive industries to innovate more. The pattern is much more evident in industries with low substitution between female and male workers than in those with high substitution, consistent with the predictions of directed technical change theory.

Keywords: factor endowment, innovations, directed technical change, price effect, market size effect, elasticity of substitution

JEL Classification: O31, O32, J21

ACKNOWLEDGMENTS

This work was undertaken as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM), led by IFPRI and supported by CGIAR Fund Donors. This paper has not gone through IFPRI's standard peer-review process. The opinions expressed here belong to the author, and do not necessarily reflect those of PIM, IFPRI, or CGIAR.

1. INTRODUCTION

Historical evidence suggests that labor scarcity was an important driver of innovation in the United Kingdom and United States (Habakkuk 1962; Allen 2009; Hayami and Ruttan 1970).[1] Rising real wages induced the innovation and adoption of labor-saving technologies to overcome labor shortages. However, the pattern seems to reverse in the past several decades in the United States. For instance, college expansions after World War II have produced more skilled workers in the United States (Autor, Katz, and Krueger 1998), which depressed the skill premium in the short run and induced the innovation of technologies complementary to them in the long run.

To reconcile the puzzle, Acemoglu (1998, 2002a; 2002b; 2007) develops the directed technical change theory, expositing that the direction of technological change depends on two effects—price effect and market size effect. The price effect, which means that technologies are induced to substitute for the scarcer production factor, is essentially the induced innovation theory developed by Hicks (1932) and Hayami and Ruttan (1970). However, if a scarce factor can be replaced easily with a more abundant factor, it may be more profitable for firms to develop technologies that are complementary to the abundant factor rather than to replace the scarce factor because of a larger market for the technologies making use of the more abundant factor. This is the so-called market size effect. Whether the price effect or market size effect dominates the direction of technical change largely depends on the degree of substitution between the scarce and abundant factors. According to the theory of directed technical change, when the elasticity of substitution among factors is low, more innovations occur in industries intensive in the scarce factor thanks to the overwhelming price effect. When the elasticity is high, the market size effect dominates and innovations bias toward industries intensive in the abundant factor.

It is extremely difficult to empirically test the above predictions because factor prices, a measure of factor abundance and scarcity, are largely endogenous. Hanlon (2015) cleverly exploits the shock of the US cotton embargo on the British textile industry during the US Civil War to test the directed technical change theory. He shows that the shortage of cotton from the United States spurred firms in Britain to come up with technological innovations to process the more abundant Indian cotton. The rather high elasticity of substitution between US cotton and Indian cotton is a major reason why the market size effect overwhelms the price effect. The paper provides a powerful support for market size effect; that is, technical changes move in the direction of the abundant factor. However, due to the limitations of historical data, Hanlon (2015) focuses on only one industry (textile industry) and relies on imprecise estimates of the elasticity of substitution between US cotton and Indian cotton. Moreover, Hanlon does not test the direction of technical change when the elasticity of substitution is low.

Inspired by Hanlon's work, this paper tests the direction of technical change in Chinese industries with varying degrees of elasticity of substitution between male and female workers in the wake of rising female labor scarcity, using large regional variations in sex ratios as a source of identification strategy. China provides a good setting to test the directed technical change theory for two reasons.

First, real wages in China have appreciated rapidly since the mid-2000s (Zhang, Yang, and Wang 2011). A change in the relative scarcity of labor and enables us to investigate whether labor scarcity has induced more innovation. Developed countries experienced episodes of rapid wage appreciation a long time ago, making it hard for researchers to find high-quality historical data to empirically test this.

Second, China has experienced one of the most serious gender imbalances (118 boys per 100 girls at birth in 2010) in the world in the past several decades thanks to the combination of family-planning policy, son-preference culture, and availability of ultrasound technology (Bulte, Heerink, and Zhang 2011; Li, Yi, and Zhang 2011). It is estimated that there is an excess of 20 to 30 million men of marriageable age. Facing the shortage of female workers, firms in female-intensive industries (e.g., the textile industry) are more likely to struggle to find female workers than their counterparts in male-intensive industries. Because

[1] See Acemoglu (2002a, 2002b, 2010) for detailed discussions of this line of literature.

the wages of male and female workers are endogenous, we cannot directly use their relative wages to measure the degree of female labor scarcity. Instead, we use the interactions of dependence on female workers at the industry level (characterized by the US benchmark) and gender differences in labor supply across regions and over time in China as a proxy for female labor scarcity.

We use approved patents as a measure of firm innovation. Based on the matched dataset of Chinese annual surveys of industrial firms and the national patent database, we find that in provinces with a higher sex ratio of males to females, privately owned enterprises (POEs) in industries with a greater dependence on female workers invest more in research and development (R&D) and possess more patents, especially invention patents. By comparison, probably because state-owned enterprises (SOEs) shoulder more policy burdens, receive more subsidies, and are less sensitive to market competition and price signals, the pattern is not evident for SOEs.

We further divide industries into two groups according to the degree of substitution between male and female workers and examine the impact of female labor shortages on these industries, respectively. The effect varies between the two groups. In industries with a low degree of substitution, the price effect dominates. That is, female labor shortages have induced firms in female-labor-intensive industries to innovate more. However, in industries with a high substitutability between male and female workers, the market size effect is more evident. Our empirical results provide some evidence in support of the directed technical change theory. That is, the degree of substitution matters to the direction of technical change. Besides patents, we look into other dimensions of innovation, including total factor productivity (TFP) and new product value, and examine the channel of R&D expenditure. The results corroborate the baseline findings.

Our paper contributes to the broad literature on innovation. In the literature, competition (Aghion et al. 2005), financial systems (Benfratello, Schiantarelli, and Sembenelli 2008), foreign direct investment (FDI) spillovers (Xu 2000), and firm characteristics (Mairesse and Mohnen 2010) have been listed as important drivers of innovation. We control for these factors in our empirical analyses. Our main results are robust to the inclusion of these control variables.

The organization of the subsequent sections is as follows. In Section 2, we review the relevant literature. Section 3 describes data and empirical strategy. Section 4 reports the results of baseline specification and robustness checks. Section 5 tries to disentangle the price effect and market size effect. Section 6 discusses the underlying channels. Section 7 concludes the paper.

2. LITERATURE REVIEW

Apart from the aforementioned specific literature on induced innovation and directed technical change, our paper is closely related to the broad literature on the determinants of innovation. Numerous factors have been listed as contributors to innovation, including but not limited to market size, competition, financial development, FDI spillovers, and firm-specific characteristics. Each factor is too vast to be reviewed comprehensively. We briefly discuss those that are most relevant to this paper.

It is costly to make innovations. R&D, an essential input of innovations, often involves a large fixed cost. A firm cannot recoup the R&D cost unless it can sell its products to a large enough market. Therefore, market size has been regarded as a major contributing factor to the success of firm innovation. Drug development in the pharmaceutical sector is a classic example. Using changes in demographic structure as a proxy for market demand of different populations for various types of drugs, Acemoglu and Linn (2004) reveal that pharmaceutical companies have much greater incentives to develop new drugs that can be sold to a large population than to a narrow segment of a population because of the rather fixed R&D cost.

Hu and Jefferson (2009) point out that the rapid increase in investment in R&D and the number of patents in China is to a large extent driven by the huge demand for technology-intensive products. In the same vein, Beerli et al. (2012) find that an increase in market size in China by 1 percentage point is associated with an increase in R&D investment by 4.4 percent, labor productivity by 6.5 percent, and the probability of developing a new product by 1.1 percent. But a puzzle arises. Given that China's large population size has been around for a while, why did not the innovations rapidly take off until the early 2000s?[2] This question suggests additional forces at play.

Competition is also regarded as a key driving force for firm innovations. However, the empirical findings are largely mixed. Using data from publicly traded manufacturing firms in the United States, Hashmi (2013) finds that competition mildly inhibits innovation. In contrast, a few empirical studies based on UK data (Blundell, Griffith, and Van Reenen 1999; Bloom, Draca, and Van Reenen 2015) reveal that competition fosters innovation. To reconcile the conflicting findings, Aghion et al. (2005) expand the Schumpeterian model to allow an inverted-U-shape relationship between competition and innovation. In their model, whether the relationship is positive or negative depends on a product's distance to the world technology frontier. The literature on market competition focuses primarily on the product market, largely ignoring the role of the input market. Our paper complements the literature by investigating the effect of factor scarcity, a measure of input factor situations, on firm innovations.

Financial constraints are often a limiting factor to firms' investment in R&D. There is an active body of literature linking financial development and firm innovations. For example, based on firm-level data in Italy, Benfratello, Schiantarelli, and Sembenelli (2008) show that banking-sector development is beneficial for firms' process innovations, especially for small firms and high-technology firms, which tend to rely heavily on external finance. However, the impact on product innovation is much more muted. In a cross-country study, Ayyagari, Demirgüç-Kunt, and Maksimovic (2012) find that firms that can obtain external finance are more innovative. Interesting to note, such relationships apply only to POEs but not to SOEs. Following Ayyagari, Demirgüç-Kunt, and Maksimovic (2012), we distinguish between POEs and SOEs in our empirical analyses. In addition, we include financial development as a control variable in our analysis.

The spillover effect of foreign firms on local firms' innovation behavior has been studied. However, the empirical findings are rather mixed. For instance, some studies find that FDI brings about competition, fosters division of labor, boosts the host country's technology innovation, and promotes product quality upgrading (Xu 2000; Cheung and Lin 2004; Hatani 2009). However, a few studies (Hu and Jefferson 2002; Fan and Hu 2007) uncover a negative effect of FDI on Chinese firms' R&D efforts. Hale and Long (2011) fail to find any systematic positive evidence about the spillover of FDI on productivity.

[2] See Xie and Zhang (2015) for a description of innovation patterns over time.

Because firms are the major agents of innovation, firm characteristics, such as firm size, age, and ownership, matter to innovation (Crepon, Duguet, and Mairesse 1998; Mairesse and Mohnen 2010). Although our paper is about the effect of factor scarcity on innovation, we have controlled for the key factors mentioned in the literature, including market size, market competition, financial development, FDI, and firm characteristics.

3. DATA AND DESCRIPTIVE PATTERNS

Data

Since the patent system is one of the most important methods to encourage innovation in the modern economy, we use patent information to measure the outcome of corporate innovation. We purchased the patent data from the State Intellectual Property Office of China, which records the number of patents applied for by all firms, individuals, and research institutions as well as the number of patents finally granted to them from 1985 through 2009.

Chinese patents include three types—invention, utility model, and design. Invention patents normally have the most technical contents among the three types. According to the patent law of China, invention patents refer to any new technical solutions relating to a product or improvement. Utility model patents are defined as new technical solutions relating to the shape, the structure, or their combination of a product, which is fit for practical use. Design patents cover new designs of the shape, the pattern, or their combination or the combination of the color with shape or pattern of a product, which creates an aesthetic feeling and is fit for industrial application.

Each patent observation in the patents dataset includes the application number, name of the patent, main classification code, patent classification code,[3] name of the applicant and the inventor,[4] application date, applicant's address, publication date and date of grant,[5] province code, and other information. However, the patent dataset does not provide any firm financial and performance information. By matching it with the dataset of the Annual Survey of Industrial Enterprises in China (ASIEC),[6] which includes financial statement information of all the industrial firms with annual sales greater than RMB 5 million, we are able to carry out in-depth analyses on innovations at the firm level.[7]

The sex ratio variable used in our empirical analysis is defined as the ratio of males to females at the province level for the cohort aged 15 to 19 in 2000, inferred from the China Population Census in 2000. The choice of this cohort at the province level is due to the following reasons. First, the size of the young local population largely represents the supply for new workers, thereby mattering to firms' labor cost. Our main sample covers 2006 and 2007. This cohort was 21 to 25 years old by 2006. As a robustness check, we also replace the sex ratio for the 15 to 19 cohort with the sex ratio for the 10 to 19 cohort in the 2000 population census. If the year of starting to work is earlier (say, 16 years old), this cohort will better capture the degree of relative scarcity of the new labor supply of female workers. Second, the cohort was relatively too young to migrate at the time being surveyed in the 2000 census.

Another key variable in our analysis is female intensity at the industry level. However, we cannot directly use the fraction of female workers in Chinese industries as a measure of industrial female intensity because it is largely endogenous.[8] Following the arguments of Rajan and Zingales (1998), we use the ratio

---

[3] While the classification codes for invention patents and utility model patents come from the International Patent Classification, the classification codes for design patents follow the International Classification for Industrial Designs. A patent can have several classification codes. The most important one is defined as the main classification code.

[4] The applicant and the inventor of a patent can be different. An applicant is the entity who actually owns the patent. It can be a firm, an institution, or a person. The inventor must be a person who has made creative contributions to the substantive features of the invention.

[5] The application date refers to the date when the applicant files the patent application. Only the invention patent includes the publication date. An invention patent has to go through three stages (preliminary examination, publication, and substantive examination) before being granted. The stage of substantive examination is not required for utility model and design patents. Thus, the publication date is not applicable to them, and only their dates of grant are revealed.

[6] For the details of the matching process, please refer to Xie and Zhang (2015).

[7] The dataset does not include service-sector or manufacturing firms which are not state-owned and whose annual sales are less than 5 million RMB (or 0.77 million US dollars).

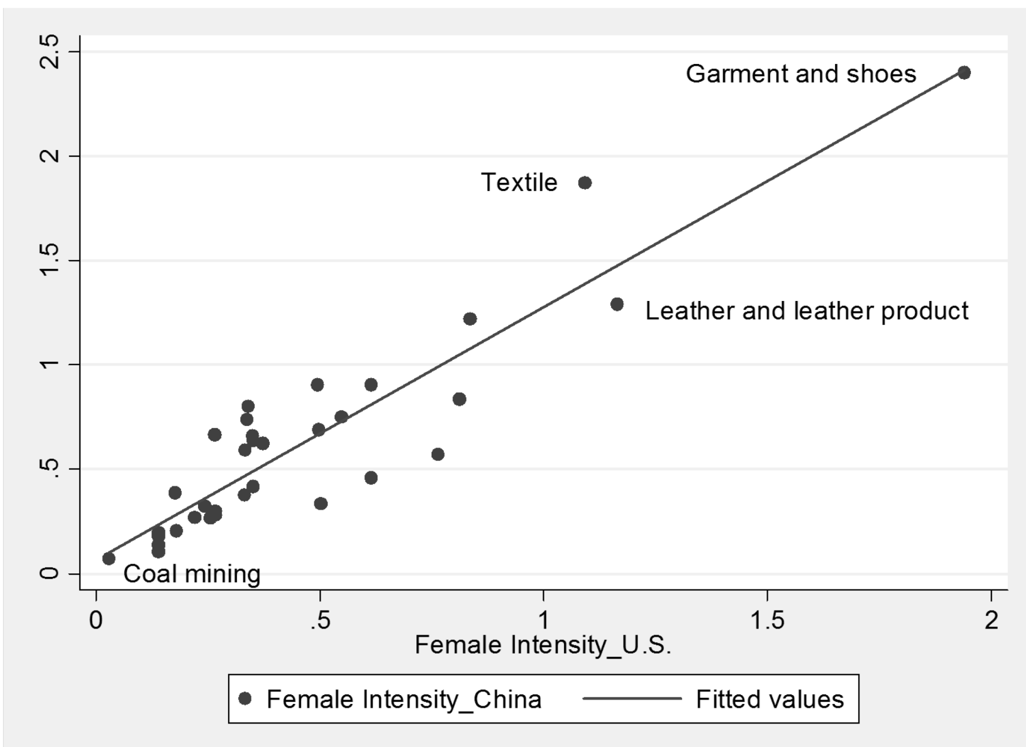
[8] As a robustness check, we use female intensity in China as an explanatory variable and use the female intensity in the United States as the instrument variable for it. The instrument variable regressions return similar results.

of female workers to male workers in the corresponding industry in the United States to construct the *FI* (female intensity) variable. The United States has a well-developed financial market and the financial constraint of firms is relatively small (Rajan and Zingales 1998). Thus, the construction of the female workers dependence variable based on US industry characteristics is least susceptible to other distortions, such as financial market frictions. Moreover, because the sex ratios in the United States are balanced, the characteristics of US industries primarily reflect the technological features of an industry and have nothing to do with sex ratio imbalances. Since the construction of the *FI* variable at the industry level is based on US data, the measure is exogenous to Chinese firms.

Descriptive Patterns

As shown in Figure 3.1, female intensity in US industries and Chinese industries mirrors closely, with a correlation coefficient .91. While in both countries the garment industry is the most female-intensive one, the coal mining industries are dominated by male workers. Thus, using the female intensity in US industries can largely capture the technical requirements of male and female workers in Chinese industries.

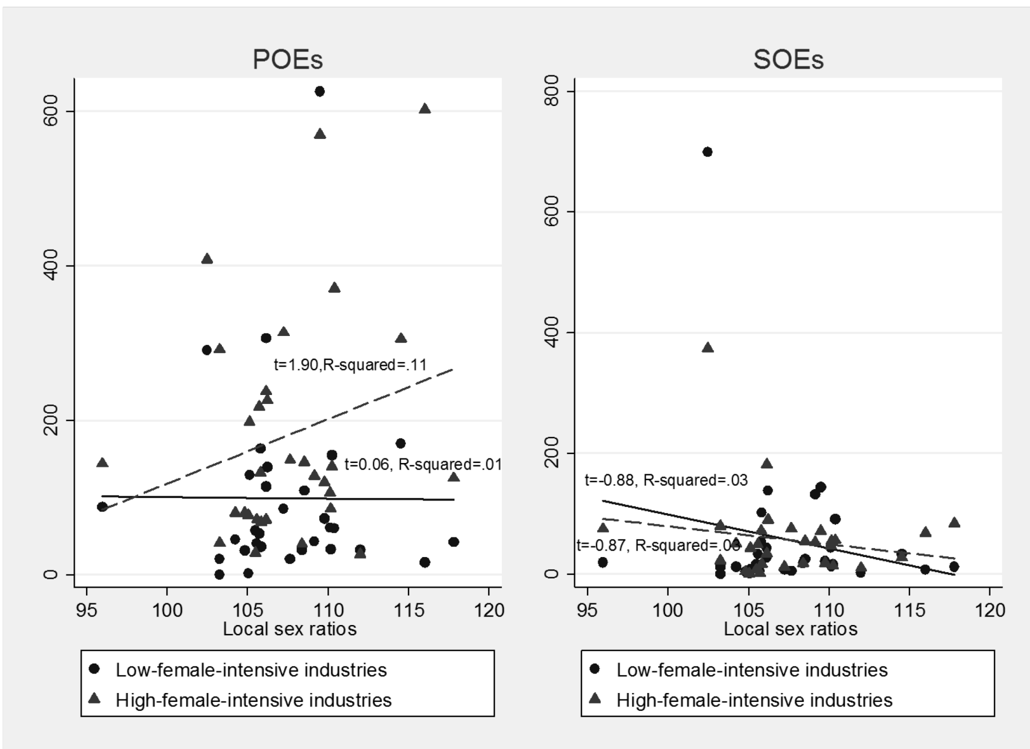
Figure 3.1 Female intensity at the industry level in the United States and China



Source: Data about the United States are from the Current Population Survey of the Bureau of Labor Statistics. Data about China are obtained from the National Bureau of Statistics.
Note: Female intensity is the ratio of female workers to male workers.

To visualize the relationship between production factor scarcity and innovation, in the left panel of Figure 3.2, we first report the number of patents per million workers for POEs across provinces in low- and high-female-intensive industries. The horizontal axis of Figure 3.2 stands for the sex ratio of males to females among the 15 to 19 cohort at the provincial level from the 2000 population census, and the vertical axis is the number of patents per million workers. We use the sample median of the *FI* variable to divide the sample into the high-female-intensive group (*FI* greater than the sample median) and low-female-intensive group (*FI* less than the sample median). In the figure, triangles (circles) represent the scatter plot of the number of patents per million workers versus local sex ratios in the high-female-intensive industries (low-female-intensive industries). The slope of the dotted line, that is, the fitted line of patents per million workers on local sex ratios for the high-female-intensive group, is highly positive, with a *t* value of 1.90. By comparison, the fitted line for the low-female-intensive group (the solid line) is almost flat, with a *t* value of 0.06. It is apparent from the panel that the slope of the two fitted lines significantly differs ($t = 2.03$). The positive relationship between the number of patents per million workers and local sex ratios is evident only in industries relying heavily on female workers. Similarly, we plot the number of patents per million workers for SOEs in different provinces and industries in the right panel of Figure 3.2. In neither the female-intensive nor the male-intensive industries does the number of patents per million workers reveal a clear relationship with local sex ratios. Moreover, there exists no significant difference between the two fitted lines ($t = 0.56$). The pattern exhibited in POEs no longer shows up in SOEs.

Figure 3.2 Sex ratios and patents per million workers in different industries



Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is the ratio of female workers to male workers in the corresponding US industry. Low- and high-female-intensive industries are categorized according to the median of female intensity. Sex ratio is the number of males per 100 females in the 15 to 19 cohort at the province level, calculated from the China Population Census 2000. While the dash line represents the linear fit of the number of patents on local sex ratios for the high-female-intensive industries, the solid line stands for the linear fit for the low-female-intensive industries. POEs = privately owned enterprises; SOEs = state-owned enterprises.

Although we have used predetermined sex ratios of the young cohort in earlier years as a proxy for the supply of the new labor force, the variable still may be subject to measurement error. One way to remedy the problem is to exclude provinces that are major destinations or sources of migration (Guangdong, Zhejiang, Shanghai, Beijing, Jiangsu, Sichuan, Henan, Anhui, Jiangxi, and Hunan) and repeat the analysis using the restricted sample.[9] As shown in Figure A.1 in the appendix, the patterns exhibited in Figure 3.2 remain robust: POEs in female-intensive industries are more sensitive to the scarcity of female workers and have stronger incentives to innovate than those in male-intensive industries. For SOEs, local sex ratios have little to do with firm innovations regardless of the degree of female intensity.

Admittedly, the bivariate scatter plots provide only suggestive evidence because other factors are not controlled for. We carry out regression analyses in the following sessions to tease out the effects of other factors.

[9] We largely follow Meng (2012).

4. OVERALL EFFECT OF FEMALE LABOR SCARCITY ON FIRM INNOVATIONS

Baseline Results

In the following regression analyses, the empirical specification is as follows:

$$y_{ijk}=\gamma FI_j \times Sexratio_k+\alpha_1 FI_j+\alpha_2 Sexratio_k+\beta_1 F_{ijk}+\beta_2 P_k+\beta_3 I_j+\varepsilon_{ijk,} \quad (1)$$

where the subscripts *i, j,* and *k* represent firms, industries, and provinces, respectively. γ, α_1, α_2, β_1, β_2, and β_3 are parameters to be estimated, and ε is the error term. y is the dependent variable. y_{ijk} refers to the number of patents granted to a firm. Since the dependent variable takes values of discrete integers, we employ the count model and decide whether to use the Poisson model or the negative binomial model by conducting the likelihood ratio test. *FI* characterizes an industry's female intensity, defined as the female to male ratio of the corresponding industry in the United States.

F, P, and *I* represent the control variables at the firm, province, and industry level, respectively. All the regressions employ the robust standard errors clustered at the province × industry level to allow for the common characteristics of firms in the same industry and province.

Following Crepon, Duguet, and Mairesse (1998) and Mairesse and Mohnen (2010), *F* includes a series of firm-level factors that may affect innovation, such as firm size (log of total assets), year of establishment, ratio of fixed assets (plant and equipment) to total assets, leverage ratio (the ratio of liability to assets), and ratio of operating profits to total assets.

P and *I* include a series of province-level and industry-level variables. Following Benfratello, Schiantarelli, and Sembenelli (2008) and Ruan and Zhang (2010), *P* and *I* contain log of gross domestic product (GDP) per capita, log of population density, share of urban population in total population, log of FDI per capita, share of secondary industry GDP in total GDP, ratio of employees in SOEs to total employees, financial development (the sum of the ratio of stock market total transaction value to GDP and the ratio of private credit to GDP), financial structure (the log ratio of the stock market total transaction value to private credit), and share of sales of the top three largest firms in an industry (market concentration index) to control for the impact of economic development, market demand, urbanization, FDI's spillovers, industrial structure, ownership structure, financial development, financial structure (relative development of the direct finance and indirect finance), and market competition on innovation. Most of these indicators are from various issues of *China Statistical Yearbook of China* (China National Bureau of Statistics) and Wind Database. The industry concentration index is computed from the annual survey of manufacturing firms. Table 4.1 presents the summary statistics of these variables.

As a robustness check, we also control for the industry and province dummy variables (no longer control for the level term of *FI* and *Sexratio*) so as to further exclude the intrinsic differences among industries and provinces.

Since expenditure on R&D, a key input of innovations, is available for only 2006 and 2007, and sex ratio variable is only available at the five-year interval, our main sample covers these two years. In the following empirical analyses, we not only conduct the panel regressions but also check whether the baseline patterns hold year by year in the cross sectional analyses.

Table 4.1 Summary statistics of key variables

Variable	Mean	Standard deviation	Minimum	Maximum
Number of invention patents	0.14	1.52	0.00	67.00
Number of utility model patents	0.48	5.68	0.00	253.00
Number of design patents	0.11	2.30	0.00	351.00
Female intensity	0.48	0.40	0.03	2.19
Sex ratio	104.69	4.62	101.01	124.08
Log cohort size	15.25	0.59	12.54	15.82
Log of gross domestic product per capita	9.94	0.35	8.71	10.87
Log of population density	5.30	1.43	0.74	7.88
Share of urban population in total population	0.39	0.10	0.16	0.73
Log of foreign direct investment per capita	4.35	1.02	0.13	6.31
Share of secondary industry in total industry	0.52	0.05	0.28	0.60
State-owned firm employees to total employees	0.54	0.11	0.33	0.94
Financial development	1.19	1.06	0.11	3.51
Financial structure	–0.19	0.88	–2.20	1.71
Top three sales share in an industry (%)	1.31	3.59	0.00	100
Log of assets	9.40	1.18	4.74	16.51
Age	5.85	6.37	0	107
Fixed assets to total assets ratio	0.42	0.23	0	1.00
Total liability to total assets ratio	0.48	0.27	0	0.99
Operating profit to total assets ratio	0.19	0.29	–0.16	1.29

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

In Table 4.2, we first report the baseline regression results based on equation (1). Columns (1) through (3) and columns (4) through (6) of Table 4.2 present the results about POEs and SOEs, respectively. For each type of firm, we use the number of invention, utility model, and design patents as the dependent variable, respectively. Since the likelihood ratio test rejects the null hypothesis of no significant difference between the Poisson count model and the negative binomial distribution count model, we report the regression results of panel analysis based on negative binomial distribution count model.[10]

Table 4.2 reveals that for POEs, the interaction term of female intensity and sex ratio is significantly positive for all three types of patents. The coefficient for the interaction term is the largest for the invention patents. That is, in provinces with more skewed sex ratios (more men than women), industries with greater dependence on female workers possess more patents, especially invention patents. In comparison, for SOEs, all the interaction terms for the three types of patents are not significant. These results further corroborate the findings from the scatter plots. Unlike SOEs, which receive government subsidies and are somewhat isolated from market competition, POEs face intense market competition and are sensitive to changes in production factor scarcity.[11]

[10] The panel analysis of this paper is based on the random effect model. The identification and consistent estimation of the fixed effect count model require a number of assumptions and conditions (Greene 2002), one of which is the change in dependent variable. This will result in a loss of observations and information. The estimations based on pooled regressions with the control of year dummies return similar results. Due to space limitations, they are not reported in separate tables.

[11] The different patterns between POEs and SOEs may be that their distributions are not the same across industries. For

Table 4.2 Female worker intensity, sex ratio, and patents

	Privately owned enterprises			State-owned enterprises		
	Invention	Utility	Design	Invention	Utility	Design
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Female intensity × Sex ratio	0.035**	0.016*	0.024**	0.037	0.035	0.040
	(0.013)	(0.010)	(0.009)	(0.030)	(0.027)	(0.034)
Female intensity	–4.197**	0.295	–2.698**	–3.621	–4.407	–3.409
	(1.317)	(0.980)	(0.910)	(3.195)	(2.858)	(3.569)
Sex ratio	–0.046***	–0.025***	–0.069***	–0.018	–0.005	–0.056*
	(0.009)	(0.007)	(0.008)	(0.021)	(0.017)	(0.026)
Cohort size (log)	–0.032	–0.080	–0.075	0.199	0.131	–0.139
	(0.090)	(0.064)	(0.087)	(0.163)	(0.119)	(0.207)
Log of gross domestic	–1.333	–0.987	–1.297	–0.733	–0.685	–1.885
product per capita	(0.946)	(0.972)	(0.934)	(0.510)	(0.679)	(1.690)
Population density (log)	7.487***	6.041***	7.311***	13.820***	5.983**	8.688*
	(1.127)	(0.812)	(1.112)	(3.080)	(2.265)	(3.946)
Share of urban population	2.436***	2.907***	1.937***	3.420**	3.312***	4.639**
	(0.513)	(0.366)	(0.487)	(1.267)	(0.895)	(1.572)
Log of foreign direct	0.422***	0.187**	–0.013	–0.125	–0.197*	–0.082
investment per capita	(0.083)	(0.060)	(0.076)	(0.116)	(0.082)	(0.157)
Share of secondary	–2.067**	–2.136***	–5.358***	–1.013	0.973	2.418
industry	(0.698)	(0.504)	(0.701)	(1.184)	(0.864)	(1.540)
State-owned firm employees	–0.014	–0.821*	–3.624***	–1.459	–2.850***	–3.321*
to total employees ratio	(0.543)	(0.380)	(0.546)	(0.952)	(0.746)	(1.308)
Financial development	0.017*	0.010*	0.008	–0.001	0.001	–0.004
	(0.007)	(0.005)	(0.007)	(0.009)	(0.006)	(0.012)
Financial structure	0.055	–0.571**	–0.559*	–0.483	–0.303	–0.000
	(0.257)	(0.186)	(0.279)	(0.374)	(0.275)	(0.606)
Top three sales share in an	0.130	0.023	0.898	–0.783	0.326	–0.269
industry	(1.126)	(0.927)	(1.096)	(0.476)	(0.376)	(0.710)
Log of assets	0.947***	0.768***	0.809***	0.867***	0.677***	0.638***
	(0.022)	(0.016)	(0.020)	(0.036)	(0.024)	(0.044)
Age	0.017***	0.019***	0.015***	0.006*	0.009***	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)
Fixed assets to total assets	–1.305***	–1.549***	–0.797***	–1.547***	–2.434***	–2.333***
ratio	(0.136)	(0.099)	(0.126)	(0.299)	(0.247)	(0.441)
Total liability to total assets	–0.937***	–0.383***	–0.176	–1.164***	–0.233	–0.921*
ratio	(0.114)	(0.081)	(0.107)	(0.282)	(0.215)	(0.371)
Operating profit to total	0.022	–0.550***	–0.423**	0.115	–1.479**	–0.558
assets ratio	(0.156)	(0.121)	(0.160)	(0.571)	(0.523)	(0.780)
Effect of FI at 75p	–0.34	0.80	–0.04	0.10	–0.03	0.16
Effect of FI at 25p	–0.44	0.77	–0.06	0.03	–0.05	0.11
Observations	29,8434	298,434	298,434	13,592	13,592	13,592

Source: Merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is defined as the ratio of female workers to male workers in the corresponding industry in the United States. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Robust standard errors clustered at the Province × Industry level are in parentheses. All regressions control for the year dummies. Effect of FI at 75p (25p) = change in the number of standard deviations of the dependent variable in response to a one standard deviation change in the female intensity variable by holding other variables at their 75th (25th) percentile (ascending order).

instance, SOEs are more likely to locate in industries with strategic significance or related to national security. We tabulate the percentage of POEs and SOEs in each industry, calculate and rank the differences, and exclude those industries whose differences between POEs and SOEs rank in the top and bottom 5 percentiles. The empirical results remain robust in the restricted sample.

To interpret the results more intuitively, we divide the sample into two subsamples according to the median of female intensity and regress the number of patents on the sex ratio and other control variables in the two subsamples, respectively. For a private firm in a female-intensive industry (*FI* greater than the median), a rise in sex ratio by one standard deviation will increase the number of invention patents by 0.74 standard deviation. But for a private firm in a male-intensive industry (*FI* less than the median), the effect is –0.01 standard deviation.

With respect to other significant control variables, log of population density and share of urban population exert positive effects on the number of patents, which comply with the literature emphasizing the impact of market size (Acemoglu and Linn 2004; Hu and Jefferson 2009): a large market is beneficial to corporate innovation. In addition, larger firms own more patents probably because they have more financial resources and human capital to innovate. The coefficient for the ratio of fixed assets to total assets is negative. With a larger ratio of fixed assets it is hard to satisfy the liquidity requirements of highly innovative activities, which are full of uncertainty.[12]

To check whether the baseline results are robust to different sets of control variables, we also add the control variables stepwise. We report the regression results about invention patents in Appendix Table A.1. The results about utility model patents and design patents are similar. To save space, they are not reported in separate tables. Columns (1)-(4) and (5)-(8) of Table A.1 present the regression results based on POEs and SOEs, respectively. In columns (1) and (5), we only include the key explanatory variables. In columns (2) and (6), we add two control variables about market size. In columns (3) and (7), we further control for other factors that may affect corporate innovation, such as FDI effect, industrial structure, financial market characteristics and degree of competition. Columns (4) and (8) further include firm size and age as control variables. As revealed in Table A.1, regardless of the set of control variables, the pattern is the same as that the baseline results in Table 4.2.

Robustness Checks

In this subsection, we conduct a series of robustness checks on the baseline results. The first concern is about migration. The actual degree of female labor scarcity measured by sex ratios may not be as serious as it appeared due to labor migration across regions. In other words, the sex ratio variable may embody measurement errors for female labor scarcity. To deal with this problem, we follow the analysis of Meng (2012) and exclude the 10 largest destinations and source provinces of migration (Guangdong, Zhejiang, Shanghai, Beijing, Jiangsu, Sichuan, Henan, Anhui, Jiangxi, and Hunan) and perform the same regression analysis as those in Table 4.2. The regression results are reported in Table 4.3.

As shown in the table, baseline results still hold.[13] While for private firms, the magnitude of the coefficients for the interaction terms becomes larger, the interaction terms in regressions on SOEs are not significant.

---

[12] Patents may be counted as intangible assets of a firm. From this perspective, the number of patents and the ratio of fixed assets may be simultaneously determined. If we drop the fixed assets ratio or lag the variable by one period to make it predetermined, the baseline results are robust.

[13] As a matter of fact, the studies on regional adjustment to labor market shocks suggest that mobility across regions is slow and incomplete in developed countries (Topel 1986; Blanchard and Katz 1992; Glaeser and Gyourko 2005; Autor, Dorn, and Hanson 2013). Labor mobility across regions and industries in China is further compounded by the household registration (*Hukou*) system, booming housing prices, nontransferability of government's social insurance programs across regions, care for old parents, and education of children. In addition, Bound and Holzer (2000) show that mobility is lowest for non-college workers, who are overrepresented in the manufacturing sector. Since our data cover only the manufacturing sector, the concern on mobility is further alleviated.

Table 4.3 Female worker intensity, sex ratio, and patents: excluding large migration provinces

	Privately owned enterprises			State-owned enterprises		
	Invention	Utility	Design	Invention	Utility	Design
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Female intensity ×	0.140***	0.092**	0.095***	0.048	–0.008	0.134
Sex ratio	(0.032)	(0.030)	(0.028)	(0.050)	(0.047)	(0.082)
Female intensity	–14.697***	–11.036***	–9.882**	–4.459	0.340	–13.442
	(3.357)	(3.153)	(3.015)	(5.389)	(5.107)	(8.779)
Sex ratio	–0.051	–0.081***	–0.024	–0.014	0.000	–0.085
	(0.028)	(0.021)	(0.025)	(0.035)	(0.028)	(0.053)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Effect of FI at 75p	0.13	–0.84	0.04	0.15	–0.03	0.16
Effect of FI at 25p	–0.66	–1.06	–0.05	0.11	–0.02	0.08
Observations	112,220	112,220	112,220	7,905	7,905	7,905

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Notes: The table excludes the 10 largest destination and source provinces of migration (Guangdong, Zhejiang, Shanghai, Beijing, Jiangsu, Sichuan, Henan, Anhui, Jiangxi, and Hunan). Other control variables are the same as those in Table 4.2. Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Robust standard errors clustered at the Province × Industry level are in parentheses. All regressions control for the year dummies. Effect of FI at 75p (25p) = the change in the number of standard deviations of the dependent variable in response to a one standard deviation change in the female intensity variable by holding other variables at their 75th (25th) percentile (ascending order). **Significant at the 5 percent level. ***Significant at the 1 percent level.

Besides migration across provinces, if the flow of labor across industries is costless, female labor scarcity can be attenuated. However, as shown in Artuc, Chaudhuri, and McLaren (2010), even in the United States, because many industries require specific skills and experience that can be accumulated only during a long period of time, it is costly for workers to move across industries. Dix-Carneiro (2014) provides such evidence for Brazil as well. To check whether this is a great challenge to the baseline results, we use the estimates of an industry's dependence on specific skills from Tang (2012) and divide the sample into two subsamples according to the median of this indicator.[14] Regressions in the two subsamples show that in the industries with higher industry-specific skills intensity (the cost to move across these industries is higher), the magnitude of the coefficient for the interaction term is larger. This is in accordance with the predictions of the baseline results. Due to space limitations, we do not report these results in tables.

Firms may move to other places in the face of female labor scarcity. To deal with this concern, we exclude from our sample the industries that do not depend a great deal on natural resources and are relatively easy to relocate (textile; manufacturing of apparel, shoes, and caps; manufacturing of culture, education, and sporting goods; manufacturing of electrical equipment and instruments, communication devices, computers, and other electronics appliances; printing) or take a more conservative approach and restrict the industries to those that depend a lot on natural resources and hence are hard to relocate (coal mining and dressing; petroleum and natural gas extraction; ferrous metals mining and dressing; nonferrous metals mining and dressing; petroleum refining; coking and nuclear fuel processing; ferrous and nonferrous metals smelting and rolling processing; wood, bamboo, cane, and grass products processing; and furniture) and rerun the regressions. The estimation results are reported in panel A and panel B of Table A.2 in the appendix. Table A.2 reveals that after we consider the possibility of relocation of industries, the baseline results continue to hold.

[14] The idea is to use the estimated wage return to tenure to proxy for the industry-specific skill intensity.

When firms are confronted with the female labor shortage, they have the option to exit the market. Since the paper exploits mainly the cross-sectional variations, we also carry out the analysis in the subsample of 2006 or 2007 to attenuate the impact of the exit of firms (focus on the incumbent firms). The year-by-year regression results are reported in Table A.3 in the appendix. In each year's cross-sectional regression analysis, the interaction term of female intensity and sex ratio is still significantly positive for POEs but not for SOEs. To further deal with the issue of exit, we restrict the sample to those firms that are existent in the sample for two periods. The baseline results continue to hold. Due to space limitations, we do not report these results in tables.

Apart from the above analyses, we conduct a few more robustness checks. Due to space limitations, we report only the regression results with the number of invention patents as the dependent variable. The results about other types of patents are qualitatively similar, but the magnitude of the coefficient before the interaction term is smaller. To begin with, the sex ratio variable in the baseline regressions is defined as the 15 to 19 age cohort at the province level. There is a concern that the variable is defined too narrowly for age and too widely for region. First, we check whether the results remain robust if we exploit a cohort with a wider range by replacing the sex ratio of the 15 to 19 cohort with the 10 to 19 cohort from the 2000 population census. In 2006 or 2007, the cohort was between 16 and 26. It is likely that some young people start their jobs at an earlier age. In this case, this ratio may better capture the actual situations of new labor supply. In Table A.4 in the appendix, we report the regression results with the use of this new cohort for sex ratio. As shown in the table, no matter the panel regressions or the pooled regressions, the baseline results continue to hold when using the sex ratio defined as a wider cohort. The interaction term of female intensity and sex ratio is still significantly positive for regressions on POEs but not on SOEs.

Second, we replace the province-level sex ratio variable with the city-level sex ratio variable and control for the province dummies to better exclude the impact of confounders at the province level. The larger variation of sex ratio at the city level allows us to take fuller advantage of the variations across regions. The regression results based on 2006 and 2007 are listed in columns (1) and (2) of Table A.5 in the appendix.

It is well known that human capital is a key engine of innovation. The education level of employees in a firm may play an important role in shaping firm innovation. However, since data about the education levels of employees in 2006 and 2007 are not available, we do not control for this variable in baseline regressions. Data about this variable are only available for 2004. As a robustness check, we add the share of employees with a bachelor's degree and higher in total employees in 2004 to control for the impact of human capital on innovation. As shown in column (3) of Table A.5, the results are robust to the inclusion of worker education levels at the firm level.

To cope with the potential reverse causality problem of including many current firm-level variables, we lag all the firm-level variables and present the results in column (4) of Table A.5. The main findings continue to hold.

Fourth, we replace control variables at the province and industry level with the province and industry fixed effects so as to systematically control for province- and industry-specific factors on corporate innovations. The regression results based on 2006 and 2007 are reported in columns (5) and (6) of Table A.5. The main findings remain robust.

As a robustness check, we also try alternative measurements for some indicators. The first one is a competition measure. Following Aghion et al. (2005), we first compute the ratio of operating profit to sales for each firm li and then calculate the competition index for industry *j* according to $1-\frac{1}{N_j}\sum_{i\in j} l_i$, where *i* stands for a firm, *j* an industry, and *Nj* the number of firms in industry *j*. The fiercer the competition, the lower is a firm's profit ratio and the larger is the value of the index. Column 7 of Table A.5 reports the results based on this alternative competition measure.[15]

In addition, we use the reduction in tariff rate in 2006 or 2007 relative to 2000 of each industry to proxy for the competition effect. China joined the World Trade Organization in 2001 and reduced its import tariff significantly after that. The decrease in tariff rate poses more competition to domestic manufacturers. From the World Trade Organization's database on tariffs, we calculate the reduction in tariff rate. The baseline results continue to hold when we use this variable to proxy for the competition effect.

Second, we use the log of the number of employees instead of the log of total assets to measure firm sizes in column 8 of Table A.5. From these two columns, the regression results are robust to alternative measures.

Considering that many firms have zero patents, we also exploit the zero inflated negative binomial model to perform the regression analysis, and the results are presented in column 9 of Table A.5.

Last, we use the data from 1999 and 2009 as a robustness check based on city-level sex ratio of the 15 to 19 cohort inferred from the 1990 and 2000 population census, respectively. With such specification, we can employ the variation of sex ratio across time as well. The drawback is that the R&D information is not available in these two years, and hence it's impossible to investigate the R&D channel if we use these years' data. Column 10 of Table A.5 reports the regression results based on the 1999 and 2009 data, with sex ratio constructed in the above way. From columns 1 to 10 of Table A.5, it is clear that the baseline regressions remain robust.

[15] In addition, we separate the sample by the median of the competition index proposed by Aghion et al. (2005) and carry out the baseline regressions in the two subsamples. For POEs, in the subsample with a higher degree of competition, the coefficient for the interaction term is more significant and greater. Even in the SOEs sample, as competition turns fiercer, the interaction term becomes significantly positive. However, it is not significant in the subsample with a lower degree of competition. Therefore, competition matters, even to SOEs. Since SOEs shoulder more policy burdens, receive more subsidies, and generally operate in a less competitive market, the overall impact on them is not evident.

5. DISENTANGLE THE PRICE AND MARKET SIZE EFFECTS

As shown in the previous section, overall, the scarcity of female workers has induced female-intensive firms to innovate more. However, the total effect may mask the opposite price and market size effects as predicted by directed technical change theory. Facing the scarcity of female workers, if female workers are highly substitutable by male workers, firms may choose to replace the scarce female workers with more abundant male workers and develop technologies complementary to the abundant factor because of a larger market making use of such technologies. This is the market size effect in the theory of directed technical change. To identify the market size effect, we first need to define the elasticity of substitution between male and female workers because whether price effect or market size effect dominates crucially hinges on the degree of substitution. However, due to lack of wage data by gender and exogenous shocks in China, we cannot estimate the elasticity of substitution between female and male workers directly in China as previously done in the literature.[16] As a compromise, we calculate the ratio of female to male workers in each industry (female intensity index) for China and the United States and rank all the industries according to the index. We argue that if the absolute difference in rank between China and the United States is small, such industry exhibits low substitution between female and male workers and vice versa. Given large differences in many dimensions, including sex ratio imbalances, between China and the United States, if an industry shows similar female intensity in both countries, the female intensity likely reflects industry-specific technological features related to workers' composition. For example, the ready-to-wear garment industry requires that workers be patient, be meticulous, and have slender fingers. Because males usually perform worse than females in these aspects, it is hard for them to replace female workers in the industry. According to the distribution of the absolute rank difference in our sample, if such a difference is smaller than or equal to 2 (57.6 percent of industries), we define it as a low-substitution industry.[17] Otherwise, it is a high-substitution industry.

We then repeat the baseline regressions using the two subsamples. Panels A and B in Table 5.1 report the estimates on industries with low and high substitutability between female and male workers, respectively. As shown in panel A, the interaction term between local sex ratios and industrial female intensity is positive in all the regressions for both POEs and SOEs. The impact of sex ratio imbalances on the innovation of female-intensive firms is stronger for POEs than for SOEs, particularly for invention patents by POEs (0.078, significant at the 1 percent level). When the degree of substitution is low, the price effect dominates—firms try to overcome the limitation of the scarce factor through innovations. By comparison, for industries with high degrees of substitution between female and male workers (panel B), the coefficient for the interaction terms in all six regressions is negative. While it is statistically significant in the first three regressions based on the subsample of POEs, it is not significant in the regressions on SOEs. In other words, in industries where female workers can be replaced easily by male workers, POEs tend to hire more male workers rather than to innovate when facing the scarcity of female workers. Due to lack of market competition, SOEs are less sensitive to the scarcity of female workers. The findings in this table are consistent with the theoretical predictions of Acemoglu (2002a).

As a robustness check, we gather each industry's male and female workers data from Australia (a country with few institutional distortions and where industry-level employment data are publicly available) and use the absolute rank difference between China and Australia or between Australia and the United

[16] Acemoglu, Autor, and Lyle (2004) exploit the military mobilization of World War II and estimate the elasticity of substitution between female and male workers. It is around 3 in the long run for the whole economy. However, they do not differentiate age cohorts or industries, and the substitution is usually larger in the long run than in the short run. De Giorgi, Paccagnella, and Pellizzari (2013) utilize the abolishment of compulsory military service in Italy and provide estimates of the short-run elasticity of substitution between female and male workers. The elasticity is around 0.7 to 1.6 for the 20 to 24 cohort.

[17] Since we use a dichotomous method, the threshold should be around the median. If we use the absolute rank difference of 1 as the threshold, the results are qualitatively similar.

States to split the sample and conduct similar analyses. The results, which are not reported here, remain robust.

Table 5.1 Female worker intensity, sex ratio, and patents: Split sample by substitutability

	Privately owned enterprises			State-owned enterprises		
	Invention	Utility	Design	Invention	Utility	Design
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Industries with low substitutability between female and male workers						
Female intensity	0.078***	0.024	0.036***	0.038	0.061*	0.051
× Sex ratio	(0.020)	(0.015)	(0.012)	(0.033)	(0.035)	(0.040)
Female intensity	–8.414***	–1.567	–3.707***	–3.173	–6.727*	–4.354
	(2.026)	(1.572)	(1.178)	(3.509)	(3.718)	(4.211)
Sex ratio	–0.059***	–0.028**	–0.068***	–0.032	0.009	–0.051
	(0.017)	(0.013)	(0.013)	(0.030)	(0.025)	(0.037)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,2361	14,2361	142,361	7,237	7,237	7,237
Panel B: Industries with high substitutability between female and male workers						
Female intensity	–0.065**	–0.121***	–0.045**	–0.013	–0.059	–0.008
× Sex ratio	(0.025)	(0.018)	(0.022)	(0.072)	(0.051)	(0.079)
Female intensity	7.021***	11.360***	5.755***	0.384	3.888	0.624
	(2.582)	(1.776)	(2.222)	(7.612)	(5.409)	(8.260)
Sex ratio	0.000	0.011	–0.040***	0.017	0.015	–0.048
	(0.014)	(0.009)	(0.012)	(0.037)	(0.026)	(0.045)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	156,073	156,073	156,073	6,355	6,355	6,355

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

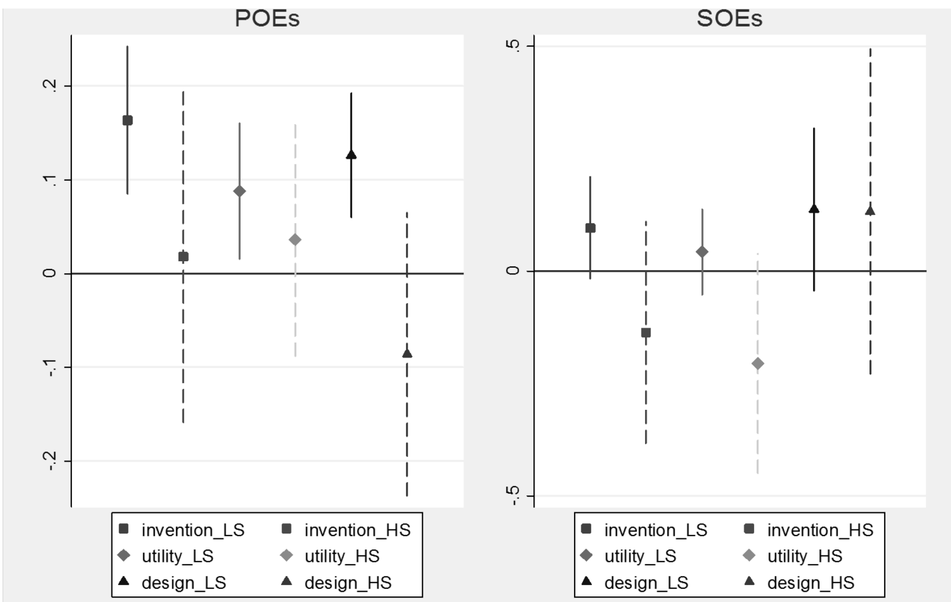
Notes: Other control variables are the same as those in Table 4.2. Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Robust standard errors clustered at the Province × Industry level are in parentheses. All regressions control for the year dummies. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

The drawback of the above method is that the rank difference may reflect other differences (for example, education level of workers) between the United States and China rather than the elasticity of substitution. Since the US population census includes detailed wage information by gender and industry and there is a an exogenous shock to female labor supply due to mobilization of World War II, we can estimate the elasticity of substitution between female and male workers by industry with the use of the US population census data in 1940 and 1950 following the methodology of Acemoglu et al. (2004), excluding the usual determinants of wages. The drawback of this method is that in different stages of development, the elasticity of substitution may change. According to Penn World Table 8.0, in 1950, GDP per capita at constant 2005 national prices in the United States is US$12,725 and China's GDP per capita at constant 2005 national prices is US$6,711. Since the income of two countries is not on a similar scale, using the US elasticity of substitution in 1950s may not capture the situations in China in the 2000s accurately. Furthermore, the sample size of some industries is small, which may also lead to an imprecise estimate of the elasticity of substitution. Nonetheless, we also use a measure based on this method as a robustness check.

We estimate the elasticity of substitution by industry and for all the industries in the sample as a whole. If an industry's elasticity of substitution is smaller (or larger) than that for all the industries, it is classified as having a low (or high) substitution (LS or HS). We then carry out separate baseline analyses in LS and HS industries and report them in Appendix Figure A.2. As shown in the figure, the patterns in Table 5.1 still hold. In LS industries, the price effect dominates and innovations bias towards female-scarce industries. By contrast, in HS industries, the market size effect plays a leading role and the opposite pattern emerges. Moreover, such patterns are more evident for POEs than for SOEs consistent with previous findings.

We next check whether migration poses a threat to above results. When workers are mobile, local sex ratios for the 15 to 19 cohort may not capture the actual labor market situations felt by firms. To address the concerns about measurement error due to migration, we drop the 10 largest destination and source provinces of migration, as in Table 4.3, and repeat the exercises in Table 5.1. Figure 5.1 plots the estimated coefficient for the interaction term and its confidence intervals. The left panel is for POEs and the right panel is for SOEs. The solid lines represent estimates from industries with low substitutability between male and female workers, while the dashed lines stand for industries with high substitutability. As revealed in the left panel, the coefficient is positive and significantly different from zero in regressions for all three types of patents in the subsample of POEs in industries with low female-male substitutability. Yet among the industries with high degrees of substitution, the coefficient is not statistically different from zero, indicating a limited market size effect. By comparison, none of the coefficients are significant in regressions on SOEs. Their innovation behavior is not responsive to the scarcity of factors of production.

Figure 5.1 Splitting sample by substitutability: Excluding provinces with large numbers of immigrants and out-migrants

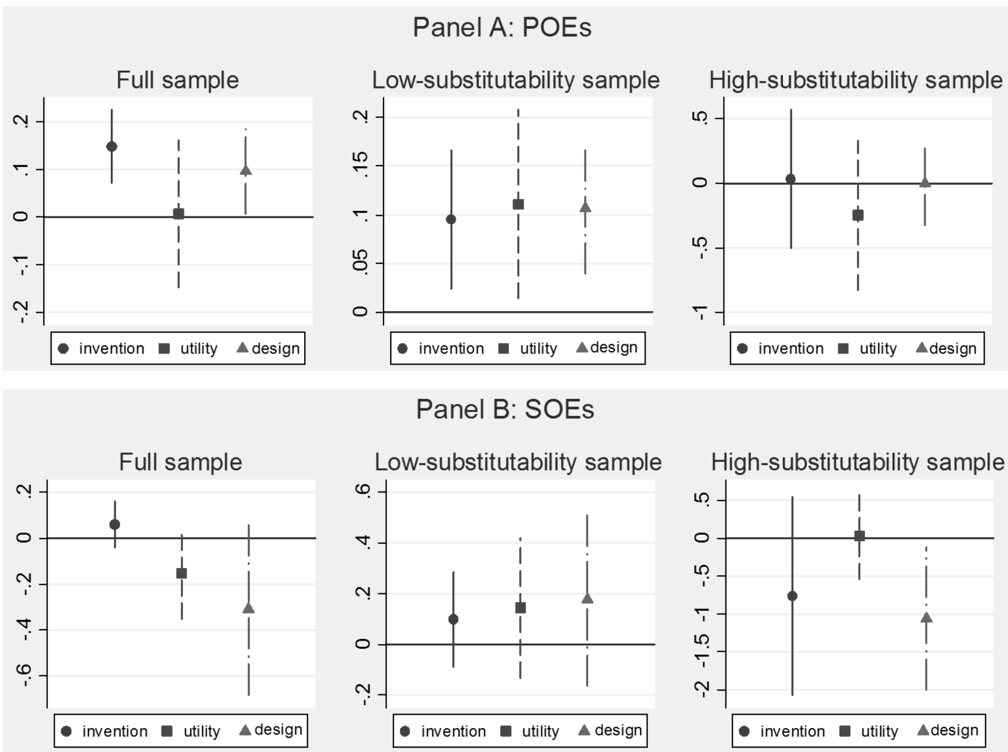


Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure reports the coefficient for the interaction term of female intensity and sex ratio and its 95 percent confidence interval in regressions similar to those in Table 5.1. One key difference from Table 5.1 is that the analysis excludes the 10 largest destination and source provinces of migration (Guangdong, Zhejiang, Shanghai, Beijing, Jiangsu, Sichuan, Henan, Anhui, Jiangxi, and Hunan). Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort at the provincial level, calculated from the China Population Census 2000. Other control variables are the same as those in Table 4.2. While LS in the legend and solid line in the figure indicate regression results based on the subsample of industries with low substitutability between female and male workers, HS and the dashed line imply regression results based on the subsample of industries with high substitutability between female and male workers. HS = high substitutability; LS = low substitutability; POEs = privately owned enterprises; SOEs = state-owned enterprises.

Figure 5.2 reports the instrument variable regression results, with the top panel for POEs and the bottom panel for SOEs.[18] Within each panel, the coefficient for the interaction term between local sex ratio and industrial female intensity and its 95 percent confidence interval in regressions based on the full sample, subsample of low female-male substitution, and subsample of high female-male substitution are separately plotted in three graphs. In the full-sample regressions on POEs, female labor shortage exerts a positive effect on invention patents and design patents. When restricting the sample to industries with low degrees of female-male substitution, the coefficient for the interaction term is significantly positive for all three types of patents, suggesting a strong price effect for POEs. However, none of the coefficients are statistically different from zero in the subsample of high female-male substitutability, indicating that the price effect has been largely offset by market size effect.

Figure 5.2 Splitting sample by substitutability: Instrument variable generalized method of moments (GMM) estimations



Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure reports the coefficient for the interaction term of female intensity and sex ratio and its 95 percent confidence interval in the instrument variable regressions based on the count model. GMM estimations of the model are used. The dependent variable is the number of granted invention/utility model/design patents. Female intensity is the ratio of female workers to male workers in corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Instruments for the sex ratio include fines for violating one-child policy in terms of yearly income and a dummy indicating whether there is a premium for higher-order births. Other control variables are the same as those in Table 4.2. The Hansen overidentification test does not reject the null hypothesis of no overidentification in all the cases. Robust standard errors clustered at the Province × Industry level are used to calculate the confidence interval. POEs = privately owned enterprises; SOEs = state-owned enterprises.

[18] The Hansen overidentification test does not reject the null hypothesis of no overidentification.

The bottom panel for SOEs can serve as a falsification test. No matter whether the full sample, low-substitutability subsample, or high-substitutability subsample is used, female labor shortage does not seem to have any effect on the innovation of SOEs.

Another challenge to the above results is that firms may exit an industry in response to the scarcity of female workers, creating a selection issue in the first place. It is difficult to estimate the nonlinear model in the presence of selection using firm-level data. Instead we aggregate the firm-level data to the industry level (four digits) for each province and year. At the aggregate level, the model becomes linear. We can then apply instrument variable regressions on the aggregate dataset. The regression results about invention patents are displayed Table 5.2. Estimates about utility model patents and design patents are qualitatively similar and are not reported to save space.

Table 5.2 Female worker intensity, sex ratio, and patents: Industry-level instrument variable GMM (generalized method of moments) analysis

	Privately owned enterprises				State-owned enterprises			
	US female intensity (FI) as explanatory variable		China FI instrumented with US FI		US female intensity (FI) as explanatory variable		China FI instrumented with US FI	
	Low ES	High ES	Low ES	High	Low ES	High ES	Low ES	High ES
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Female	0.232***	-0.855*	0.209***	-0.496	0.092	0.124	0.063	0.138
× Sex ratio	(0.089)	(0.500)	(0.078)	(0.399)	(0.281)	(1.298)	(0.166)	(0.844)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,066	9,970	6,066	9,970	2,804	3,470	2,804	3,470

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Notes: The table reports the results about invention patents. The pattern about utility model patents and design patents are similar. Due to space limits, they are not reported. In each year, firm patents are aggregated to each province's each 4 digits industry. Low (High) ES indicates low (high) elasticity of substitution industries. If the estimated elasticity of substitution of an industry is lower (larger) than that estimated based on all the industries in the sample, it is classified as a low (high) ES industry. The estimates of the elasticity of substitution follow the methodology of Acemoglu et al. (2004). GMM estimations of the endogenous count model are used. In columns (1), (2), (5), (6), female intensity is the ratio of female workers to male workers in corresponding U.S. industry. In columns (3), (4), (7), (8), female intensity is the ratio of female workers to male workers in each industry of China and is instrumented with the female intensity of the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Instruments for the sex ratio include fines for violating one-child policy in terms of yearly income and a dummy indicating whether there is a premium for higher-order births. Other control variables are the same as those in Table 4.2. The Hansen overidentification test does not reject the null hypothesis of no overidentification in all the cases. Robust standard errors clustered at the province × industry level are in parentheses. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

Moreover, to check whether the baseline results are robust to the use of female intensity in Chinese industries directly, we exploit female intensity in China as an explanatory variable and instrument it with the female intensity of the United States. The results are presented in columns (3)-(4) and (7)-(8) of Table 5.2. In Table 5.2, the classification of low (or high) elasticity of substitution industries follows the estimation methodology of Acemoglu et al. (2004) and is the same as that used in Appendix Figure A.2. The first stage regression results are listed in Appendix Table A.6.

As shown in Table A.6, in the first stage regressions about sex ratio, fines for violating the one-child policy in terms of yearly income and the dummy indicating the existence of premium for higher order birth are significantly positive, suggesting that the stricter the one-child policy is, the more imbalanced the sex ratio is. This is consistent with the current literature (Bulte, Heerink, and Zhang 2011; Li, Yi, and Zhang 2011) about the causes of sex ratio imbalance in China. In the first-stage regressions about female intensity in China, the coefficient for the US female intensity is significantly positive. In all the first-stage regressions, the F statistics are highly significant and larger than 10, implying that the IVs are not weak.

Table 5.2 demonstrates that the instrument variable regressions based on aggregate data yield the same results as firm-level regressions. For industries with low elasticity of substitution, the price effect dominates, while for industries with high elasticity of substitution, the market size effect is more evident. Overall, the selection effect (exit of firms) is not a great challenge to the baseline results.

6. ALTERNATIVE MEASURES FOR INNOVATION AND UNDERLYING CHANNELS

Some have argued that the quality of domestically approved patents has been diluted along with quantity explosions, casting a doubt about using them as a measure of firm innovation (Long and Wang, 2015). One option to address this concern is to count only patents filed by Chinese firms and approved by the United States Patent and Trademark Office (USPTO), which presumably have higher quality. However, these patents are highly skewed in a handful of Chinese companies with the top 10 assignees accounting for over 85 percent of total USPTO patents (Eberhardta, Helmers, and Yu 2011). The lack of variation among most firms prevents us from using USPTO patents as the major outcome variable in our empirical analyses. Alternatively, we use TFP, the ratio of new product value to total product value, and R&D to characterize innovations. Figure 6.1 reports the coefficient for the interaction term between sex ratio at the province level and female intensity at the industry level in regressions on the three alternative outcome variables. While the upper panel is for POEs, the lower panel is for SOEs. As shown in the upper panel, in the full sample and the subsample of industries with low female-male substitutability, female labor shortages induce firms in female-intensive industries to improve their TFP and develop more new products. These results indicate a strong price effect. In the subsample of high substitutability, the impact on TFP and new product value becomes negative. Because firms can easily find more abundant male workers to replace female workers in high-substitutability industries, the price effect is weaker and largely canceled out by the market size effect.

In the lower panel for SOEs, none of the coefficients are significant. Unlike POEs, SOEs in industries with greater dependence on female workers and in regions with imbalanced sex ratios do not respond to the scarcity of female workers in terms of TFP and new product development.

Firms may use different strategies to cope with female labor shortages. Investing more in R&D is one important option. They can also offer higher wages to female workers to retain them, expand fixed-asset investment, or import capital goods (such as machines) to reduce reliance on scarce female workers.

We first check if R&D expenditure, an essential input to firm innovation, also responds to change in factor scarcity. Since some firms have zero R&D expenditure, we use the inverse hyperbolic sine function to transform the ratio of R&D expenditure to sales. As a robustness check, we exploit the Tobit model by setting the lower limit as zero. Since the results are qualitatively similar, we report only the former in Figure 6.1 (the explanatory variables are the same as the baseline regressions). The results for R&D mirror those for patents. Overall, POEs are more responsive to the scarcity of female workers than the SOEs. Female labor shortages induce POEs to invest more in R&D, in particular in industries with lower substitutability between female and male workers. In industries in which female workers can be easily replaced with male workers, the price effect is largely offset by the market size effect.

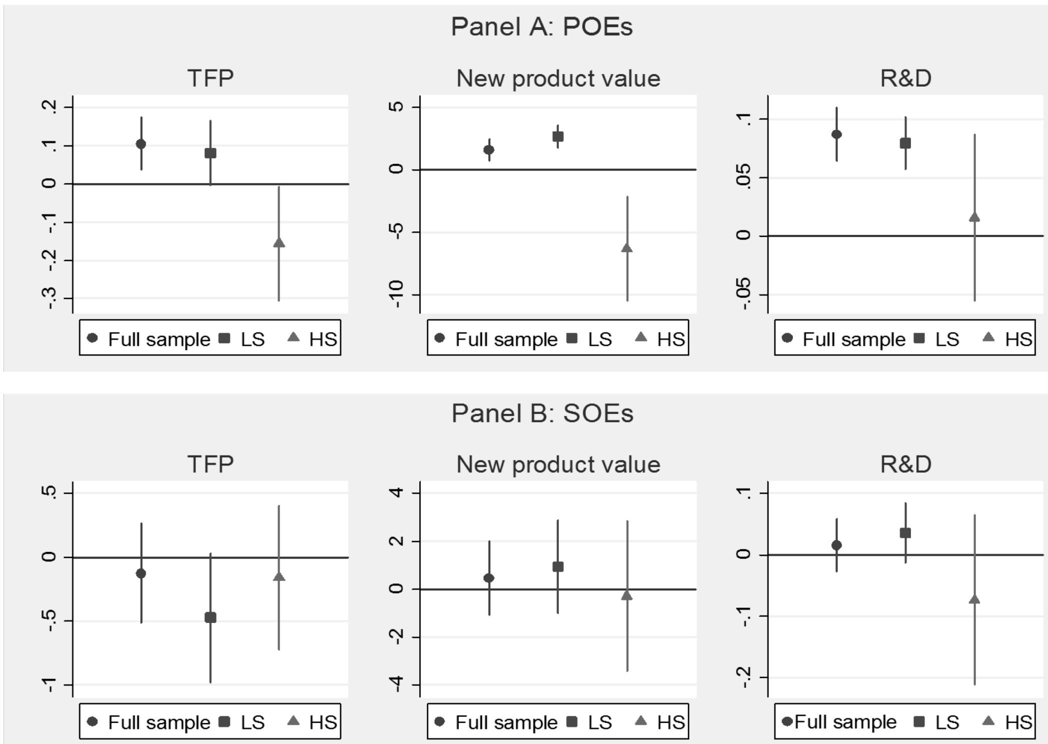
We next explore the wage channel by replacing the dependent variable with percent changes in wages (all the control variables are the same with baselines). The coefficient for the interaction term in regressions is displayed in Figure 6.2. For both POEs and SOEs, female labor shortage is not a major driver behind changes in wages.

Third, we examine the channel of capital investment. We employ the percent change in fixed capital as the dependent variable and rerun the baseline regressions. Admittedly, the expansion of fixed capital may not accurately characterize firms' investment in female labor-saving machines. Nonetheless, the variable provides some hints about this. As shown in Figure 6.2, female labor shortages do not affect female-intensive industries' capital investment, suggesting that substituting scarce female workers through fixed capital expansion is not the main channel driving the baseline results.

Last but not least, we scrutinize the channel of importing capital goods. Firms may import capital goods (machines) to save scarce female workers. From the China Customs Dataset, we first obtain value and varieties of imported capital goods at the firm level before merging them with the Annual Survey of Industrial Enterprises in China following the methodology described in Yu (2014). The coefficient for the interaction term in regressions with the ratio of imported capital goods value to sales (take the inverse hyperbolic sine transformation) and the number of capital goods variety as the dependent variable are displayed in the bottom panel of Figure 6.2. No matter whether we use the import value or the variety of import goods as an outcome measure, the interaction term between female intensity and sex ratio is insignificant.

Collating the above evidence, R&D investment is shown to be the main strategy for female-intensive firms to cope with female labor shortages, which ultimately affects firms' productivity, new product share, and patents.

Figure 6.1 Estimation results for productivity, new product value, and R&D expenditure



Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure reports the coefficient before the interaction term of female intensity and sex ratio and its 95 percent confidence interval in the regressions. The dependent variable is TFP, which is estimated using the Olley-Pakes method, the ratio of n product value to total product value, and the inverse hyperbolic sine transformation of the ratio of R&D expenditure to sales. Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Province and industry dummies are controlled for in all the regressions. Other firm-level control variables are the same as those in Table 4.2. Full sample, LS, and HS in the legend indicate that regression results are based on the full sample, the subsample of industries with low degree of substitution between female and male workers, and the subsample of industries with high degree of substitution between female and male workers, respectively. HS = high substitutability; LS = low substitutability; POEs = privately owned enterprises; R&D = research and development; SOEs = state-owned enterprises; TFP = total factor productivity.

Figure 6.2 Alternative channels in response to production factor scarcity

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure reports the coefficient before the interaction term of female intensity and sex ratio and its 95 percent confidence interval in the regressions. The dependent variable is percent change of wages, percent change of capital, the inverse hyperbolic sine transformation of the ratio of imported capital goods to sales, and the number of imported capital goods variety. Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Other control variables are the same as those in Table 4.2. POEs = privately owned enterprises; SOEs = state-owned enterprises.

7. CONCLUSION

This paper employs variations in dependence on female workers across industries and regional differences in sex ratio to investigate whether and how factor endowment shape corporate innovations. Empirical results demonstrate that in female-intensive industries, POEs in provinces with a higher sex ratio of males to females spend more on R&D and possess more patents, especially invention patents. This pattern is more evident in industries with a lower degree of substitution between female and male workers. In industries with high substitutability, this pattern is less pronounced or even reversed, indicating that the price effect is weakened by the market size effect. In comparison, these patterns do not apply to SOEs at all, which are less sensitive to the scarcity of factors of production due to lack of competition. In a competitive environment, factor scarcity is the mother of firm innovation, in particular when the scarce factor cannot be easily substituted. The findings lend support to the directed technical change theory.

APPENDIX: SUPPLEMENTARY TABLES AND FIGURES

Table A.1 Female worker intensity, sex ratio, and patents: Add control variables stepwise

	Privately owned enterprises				State-owned enterprises			
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Female intensity× Sex	0.049***	0.047***	0.046***	0.036***	0.040	0.039	0.033	0.032
	(0.014)	(0.014)	(0.014)	(0.013)	(0.039)	(0.038)	(0.037)	(0.030)
Female intensity	-5.763***	-	-	-	-5.086	-5.101	-4.430	-2.989
	(1.432)	(1.419)	(1.392)	(1.333)	(4.100)	(4.043)	(3.932)	(3.227)
Sex ratio	-0.075***	-	-	-	-0.024	-0.025	-0.016	-0.015
	(0.008)	(0.009)	(0.010)	(0.009)	(0.020)	(0.025)	(0.025)	(0.022)
Cohort size (log)	-0.789***	-	-0.213**	-0.043	-0.073	0.337**	0.383**	0.178
	(0.055)	(0.074)	(0.095)	(0.090)	(0.114)	(0.157)	(0.173)	(0.164)
Log of gross domestic		-0.520	-0.969	-1.216		-1.285	-1.217	-0.809
product per capita		(0.357)	(0.785)	(0.993)		(0.925)	(0.974)	(0.537)
Population density		6.441***	7.985***	7.520***		16.441***	16.402***	13.991***
		(1.037)	(1.215)	(1.138)		(3.206)	(3.476)	(3.121)
Share of urban		2.992***	2.915***	2.255***		2.691**	3.653***	2.998**
		(0.387)	(0.533)	(0.513)		(1.277)	(1.345)	(1.273)
Log of foreign direct			0.154*	0.436***			-0.100	-0.104
investment per capita			(0.083)	(0.083)			(0.133)	(0.115)
Share of secondary			-1.229*	-			0.812	-1.547
			(0.743)	(0.706)			(1.275)	(1.182)
Ratio of SOE			-0.037	-0.006			-0.543	-1.271
to total employees			(0.551)	(0.543)			(1.075)	(0.946)
Financial development			0.024***	0.016**			-0.005	-0.003
			(0.007)	(0.007)			(0.009)	(0.009)
Financial structure			0.035	0.159			-0.651	-0.430
			(0.269)	(0.259)			(0.437)	(0.373)
Top 3 sales share in			1.018	0.263			1.885***	-0.561
industry			(1.214)	(1.137)			(0.565)	(0.475)
Log of assets				0.958***				0.870***
				(0.022)				(0.037)
Age				0.018***				0.004
				(0.003)				(0.003)
Observations	29,8434	298,434	298,434	298,434	13,592	13,592	13,592	13,592

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is defined as the ratio of female workers to male workers in the corresponding industry in the United States. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Robust standard errors clustered at the Province × Industry level are in parentheses. All regressions control for the year dummies. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

Table A.2 Female worker intensity, sex ratio, and patents: Excluding industries easy to move

	Privately owned enterprises			State-owned enterprises		
	Invention	Utility	Design	Invention	Utility	Design
Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Exclude industries that are easy to move					
Female intensity	0.151*	0.097*	0.092**	–0.024	0.018	0.094
× Sex ratio	(0.081)	(0.054)	(0.048)	(0.083)	(0.118)	(0.137)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	90,100	90,100	90,100	6,628	6,628	6,628
	Panel B: Keep industries that are hard to move					
Female intensity	0.171*	0.132	0.150**	–0.064	0.053	–0.071
× Sex ratio	(0.098)	(0.127)	(0.078)	(0.128)	(0.121)	(0.204)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,780	51,780	51,780	3,524	3,524	3,524

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: In panel A, the table excludes industries that are easy to move (textile; manufacturing of apparel, shoes, and caps; manufacturing of culture, education, and sporting goods; manufacturing of electrical equipment and instruments; communication devices, computers, and other electronics appliances; printing). In panel B, we keep industries that depend a lot on natural resources and are hard to relocate (coal mining and dressing; petroleum and natural gas extraction; ferrous metals mining and dressing; nonferrous metals mining and dressing; petroleum refining; coking and nuclear fuel processing; ferrous and nonferrous metals smelting and rolling processing; wood, bamboo, cane, and grass products processing; and furniture), which are more conservative than the subsample in panel A. Other control variables and notes are the same as those in Table 4.2. *Significant at the 10 percent level. **Significant at the 5 percent level.

Table A.3 Female worker intensity, sex ratio, and patents: Year-by-year analysis

	2006			2007		
	Invention	Utility	Design	Invention	Utility	Design
Variable	(1)	(2)	(3)	(4)	(5)	(6)
	Privately owned enterprises					
Female intensity	0.167***	0.160**	0.222**	0.158***	0.124**	0.090
× Sex ratio	(0.055)	(0.082)	(0.087)	(0.054)	(0.058)	(0.075)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,100	51,100	51,100	61,120	61,120	61,120
	State-owned enterprises					
Female intensity	0.040	0.009	0.249	0.027	–0.013	0.080
× Sex ratio	(0.059)	(0.057)	(0.223)	(0.059)	(0.063)	(0.113)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,403	4,403	4,403	3,502	3,502	3,502

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the 2000 population census. Other control variables are the same as those in Table 4.2. Due to space limitations, they are not reported here. Robust standard errors clustered at the Province × Industry level are in parentheses. **Significant at the 5 percent level. ***Significant at the 1 percent level.

Table A.4 Female worker intensity, sex ratio, and patents: Robustness check by replacing the cohort of sex ratio

	Privately owned enterprises					
	2006 and 2007 R.E.			Pooled 2006 and 2007		
	Invention	Utility	Design	Invention	Utility	Design
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Female intensity × Sex ratio	0.149***	0.109**	0.109***	0.204***	0.171**	0.207*
of the 10–19 age cohort	(0.043)	(0.048)	(0.042)	(0.072)	(0.083)	(0.115)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	112,220	112,220	112,220	112,220	112,220	112,220
	State-owned enterprises					
	2006 and 2007 R.E.			Pooled 2006 and 2007		
	Invention	Utility	Design	Invention	Utility	Design
	(1)	(2)	(3)	(4)	(5)	(6)
Female intensity × Sex ratio	0.051	–0.012	0.101	0.010	0.012	0.129
of the 10–19 age cohort	(0.071)	(0.076)	(0.112)	(0.071)	(0.075)	(0.142)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,905	7,905	7,905	7,905	7,905	7,905

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio of the 10–19 age cohort is the number of males per 100 females in the 10 to 19 cohort, calculated from the 2000 population census. Other control variables are the same as those in Table 4.2. Due to space limitations, they are not reported. Robust standard errors clustered at the Province × Industry level are in parentheses. R.E. = random effects estimations. *Significant at the 10 percent level. **Significant at the 5 percent level. ***Significant at the 1 percent level.

Table A.5 Female worker intensity, sex ratio, and patents: Additional robustness checks

Variable	City sex ratio, province dummies 2006	City sex ratio, province dummies 2007	Control employees' education structure	Lag firm controls	Province and industry dummies 2006	Province and industry dummies 2007	Substitute competition	Substitute size	Zero inflated model	City sex ratio, panel of 1999 and 2009
Privately owned enterprises										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female intensity ×	0.105**	0.105***	0.136***	0.165***	0.169***	0.103***	0.104***	0.114***	0.325***	0.051***
Sex ratio	(0.041)	(0.029)	(0.041)	(0.056)	(0.065)	(0.039)	(0.029)	(0.026)	(0.081)	(0.019)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51,100	61,120	51,525	46,039	51,100	61,120	112,220	112,220	112,220	66,183
State-owned enterprises										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female intensity ×	0.009	-0.007	0.059	0.039	0.042	0.034	0.042	0.055	-0.125	-0.027
Sex ratio	(0.062)	(0.043)	(0.056)	(0.063)	(0.054)	(0.042)	(0.047)	(0.047)	(0.111)	(0.043)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,403	3,502	5,389	3,043	4,403	3,502	7,905	7,905	7,905	20,840

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the 2000 population census. In column (10), in 1999 and 2009, sex ratio is calculated from the 1990 and 2000 population censuses, respectively. Other control variables are the same as those in Table 4.2. Due to space limitations, they are not reported. Robust standard errors clustered at the Province × Industry level are in parentheses. **Significant at the 5 percent level. ***Significant at the 1 percent level.

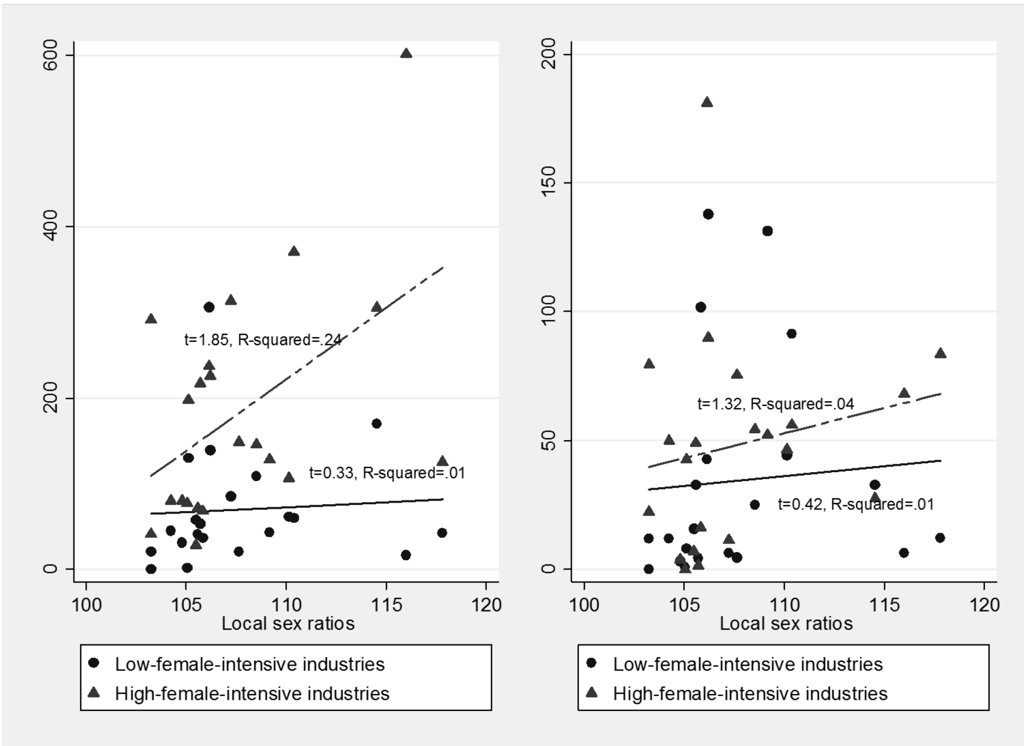
Table A.6 Female worker intensity, sex ratio, and patents: Industry-level first-stage regression results

	Privately owned enterprises						State-owned enterprises					
	US female intensity (FI) as explanatory variable		China female intensity (FI) instrumented with US female intensity (FI)				US female intensity (FI) as explanatory variable		China female intensity (FI) instrumented with US female intensity (FI)			
	Low ES	High ES	Low ES		High ES		Low ES	High ES	Low ES		High ES	
	Sex ratio	Sex ratio	Sex ratio	China FI	Sex ratio	China FI	Sex ratio	Sex ratio	Sex ratio	China FI	Sex ratio	China FI
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Fines for violating	1.385**	0.523	1.385**	-0.015	0.523	-0.009	3.990***	1.259	3.990***	-0.033	1.259	0.098
one-child policy	(0.678)	(0.705)	(0.678)	(0.038)	(0.705)	(0.044)	(1.158)	(1.367)	(1.158)	(0.044)	(1.367)	(0.078)
Premium for higher	3.091***	3.474***	3.091***	-0.007	3.474***	-0.031	3.248***	2.212***	3.248***	-0.018	2.212***	0.019
order birth	(0.362)	(0.383)	(0.362)	(0.020)	(0.383)	(0.024)	(0.558)	(0.713)	(0.558)	(0.021)	(0.713)	(0.041)
US female intensity			0.196	1.164***	-0.594	0.591***			0.923	1.195***	0.519	0.771****
			(0.887)	(0.049)	(1.565)	(0.097)			(1.664)	(0.063)	(3.104)	(0.177)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage F statistic	55.99	115.95	45.30	6096.31	93.09	234.73	13.13	10.52	10.52	3885.76	10.43	81.87
Observations	6,066	9,970	6,066	6,066	9,970	9,970	2,804	3,470	2,804	2,804	3,470	3,470

Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The table reports the results about first-stage instrument variable regressions at the industry level. In each year, firm patents are aggregated to each province's each 4 digits industry. Low (High) ES stands for low (high) elasticity of substitution industries, which are classified according to the estimates of the elasticity of substitution of each industry relative to the elasticity of substitution among all the industries following the methodology of Acemoglu et al. (2004). Sex ratio is the number of males per 100 females in the 15 to 19 cohort, calculated from the China Population Census 2000. Instruments for the sex ratio include fines for violating one-child policy in terms of yearly income and a dummy indicating whether there is a premium for higher-order births. Other control variables are the same as those in Table 4.2. Robust standard errors clustered at the province × industry level are in parentheses. **Significant at the 5 percent level. ***Significant at the 1 percent level.

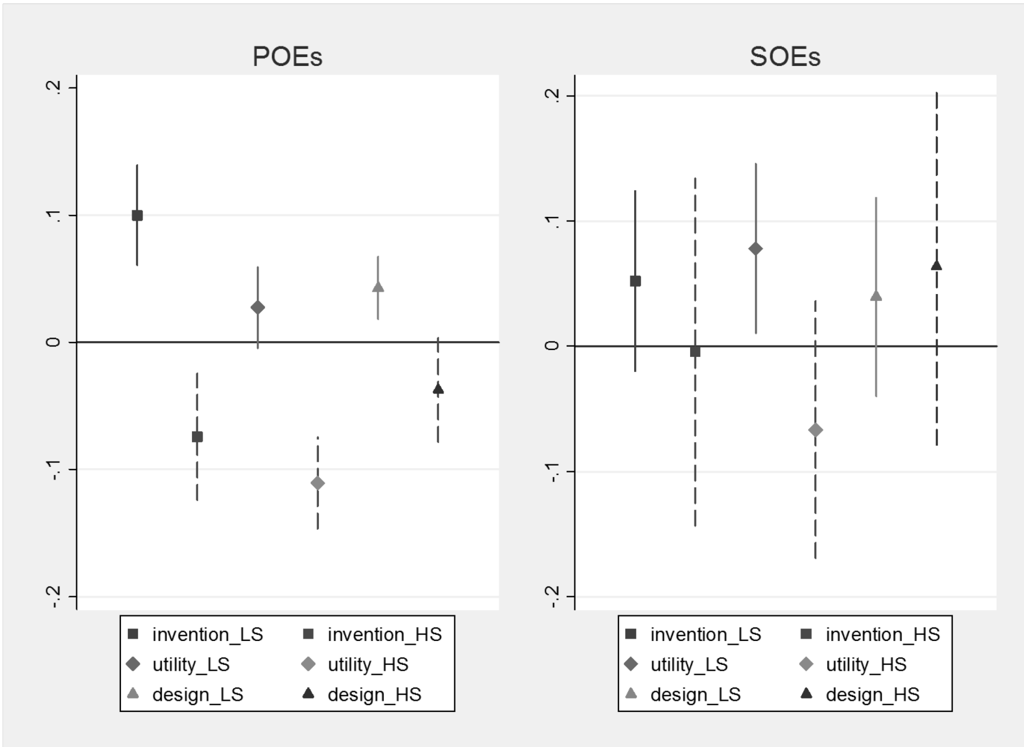
Figure A.1 Sex ratios and patents per million workers in different industries: Excluding large migration provinces



Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure excludes the 10 largest destination and source provinces of migration (Guangdong, Zhejiang, Shanghai, Beijing, Jiangsu, Sichuan, Henan, Anhui, Jiangxi, and Hunan). Female intensity is the ratio of female workers to male workers in the corresponding US industry. Low- and high-female-intensive industries are categorized according to the median of female intensity. Sex ratio is the number of males per 100 females in the 15 to 19 cohort at the province level, calculated from the China Population Census 2000. While the dashed line represents the linear fit of the number of patents per million workers on local sex ratios for the high-female-intensive industries, the solid line stands for the linear fit for the low-female-intensive industries.

Figure A.2 Female worker intensity, sex ratio and patents: Splitting sample by estimated elasticity of substitution



Source: Calculated by authors based on the merged firm patent database between the national patent database and ASIEC database.

Note: The figure reports the coefficient for the interaction term between female intensity and sex ratio and its 95 percent confidence interval in regressions similar to those in Table 5.1. One key difference from Table 5.1 is that the division of the full sample is according to the estimated elasticity of substitution between female and male workers following the methodology of Acemoglu et al. (2004). An industry is classified as a low (high) substitution (abbreviated as LS or HS) industry if the estimate of its elasticity of substitution is lower (larger) than the estimate of the elasticity of substitution based on all the industries in the sample. Female intensity is the ratio of female workers to male workers in the corresponding US industry. Sex ratio is the number of males per 100 females in the 15 to 19 cohort at the provincial level, calculated from the China Population Census 2000. Other control variables are the same as those in Table 4.2. While LS in the legend and solid line in the figure indicate regression results based on the subsample of industries with low elasticity of substitution between female and male workers, HS and the dashed line imply regression results based on the subsample of industries with high elasticity of substitution between female and male workers. HS = high elasticity of substitution; LS = low elasticity of substitution; POEs = privately owned enterprises; SOEs = state-owned enterprises.

REFERENCES

Acemoglu, D. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics* 113 (4): 1055–1089.

———. 2002a. "Directed Technical Change." *Review of Economic Studies* 69 (4): 781–809.

———. 2002b. "Technical Change, Inequality, and the Labor Market." *Journal of Economic Literature* 40: 7–72.

———. 2007. "Equilibrium Bias of Technology." *Econometrica* 75 (5): 1371–1409.

———. 2010. "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy* 118 (6): 1037–1078.

Acemoglu, D., D. Autor, and D. Lyle. 2004. "Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Mid Century." *Journal of Political Economy* 112 (3): 497–551.

Acemoglu, D., and J. Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3): 1049–1090.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt. 2005. "Competition and Innovation: An Inverted-U Relationship." *Quarterly Journal of Economics* 120 (2): 701–728.

Allen, R. C. 2009. *The British Industrial Revolution in Global Perspective.* New York: Cambridge University Press.

Artuc, E., S. Chaudhuri, and J. McLaren. 2010. "Trade Shocks and Labor Adjustment: A Structural Empirical Approach." *American Economic Review* 100 (3): 1008–1045.

Autor, D. H., D. Dorn, and G. H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103 (6): 2121–2168.

Autor, D. H., L. F. Katz, and A. B. Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics* 113 (4): 1169–1213.

Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic. 2012. "Firm Innovation in Emerging Markets: The Role of Finance, Governance, and Competition." *Journal of Financial and Quantitative Analysis* 46 (6): 1545–1580.

Beerli, A., F. Weiss, F. Zilibotti, and J. Zweimüller. 2012. *Demand Forces of Technical Change: Evidence from Chinese Manufacturing Industry.* Working Paper No. 206, Department of Economics, University of Zurich.

Benfratello, L., F. Schiantarelli, and A. Sembenelli. 2008. "Banks and Innovation: Microeconometric Evidence on Italian Firms." *Journal of Financial Economics* 90 (2): 197–217.

Blanchard, O. J., and L. F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1992 (1): 1–75.

Bloom, N., M. Draca, and J. Van Reenen. 2015. "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity." *Review of Economic Studies* 83 (1): 87–117.

Blundell, R., R. Griffith, and J. Van Reenen. 1999. "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms." *Review of Economic Studies* 66 (3): 529–554.

Bound, J., and H. J. Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics* 18 (1): 20–54.

Bulte, E., N. Heerink, and X. Zhang. 2011. "China's One-child Policy and 'the Mystery of Missing Women': Ethnic Minorities and Male-biased Sex Ratios." *Oxford Bulletin of Economics and Statistics* 73 (1): 21–39.

Cheung, K., and P. Lin. 2004. "Spillover Effects of FDI on Innovation in China: Evidence from the Provincial Data." *China Economic Review* 15 (1): 25–44.

Crepon, B., E. Duguet, and J. Mairesse. 1998. "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level." *Economics of Innovation and New Technology* 7 (2): 115–158.

De Giorgi, G., M. Paccagnella, and M. Pellizzari. 2013. *Gender Complementarities in the Labor Market.* Bank of Italy Occasional Paper 183. Rome: Bank of Italy.

Dix-Carneiro, R. 2014. "Trade Liberalization and Labor Market Dynamics." *Econometrica* 82 (3): 825–885.

Eberhardt, M., C. Helmers, and Z. Yu. 2011. "Is the Dragon Learning to Fly? An Analysis of the Chinese Patent Explosion." CSAE Working Paper WPS/2011-15. Cambridge, UK: University of Cambridge.

Fan, C. S., and Y. Hu. 2007. "Foreign Direct Investment and Indigenous Technological Efforts: Evidence from China." *Economics Letters* 96 (2): 253–258.

Glaeser, E. L., and J. Gyourko. 2005. "Urban Decline and Durable Housing." *Journal of Political Economy* 113 (2): 345–375.

Greene, W. 2002. *The Bias of the Fixed Effects Estimator in Nonlinear Models.* New York University Working Paper. New York: New York University.

Habakkuk, H. J. 1962. *American and British Technology in the Nineteenth Century.* London: Cambridge University Press.

Hale, G., and C. Long. 2011. "Are There Productivity Spillovers from Foreign Direct Investment in China?" *Pacific Economic Review* 16 (2): 135–153.

Hanlon, W. 2015. "Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change." *Econometrica* 83 (1): 67–100.

Hashmi, A. 2013. "Competition and Innovation: The Inverted-U Relationship Revisited." *Review of Economics and Statistics* 95 (5): 1653–1668.

Hatani, F. 2009. "The Logic of Spillover Interception: The Impact of Global Supply Chains in China." *Journal of World Business* 44 (2): 158–166.

Hayami, Y., and V. W. Ruttan. 1970. "Factor Prices and Technical Change in Agricultural Development: The United States and Japan, 1880–1960." *Journal of Political Economy* 78 (5): 1115–1141.

Hicks, J. R. 1932. *The Theory of Wages.* New York: Macmillan.

Hu, A. G., and G. H. Jefferson. 2002. "FDI Impact and Spillover: Evidence from China's Electronic and Textile Industries." *World Economy* 25 (8): 1063–1076.

———. 2009. "A Great Wall of Patents: What Is behind China's Recent Patent Explosion?" *Journal of Development Economics* 90 (1): 57–68.

Li, H., J. Yi, and J. Zhang. 2011. "Estimating the Effect of the One-child Policy on Sex Ratio Imbalance in China: Identification based on the Difference-in-differences." *Demography* 48 (4): 1535–1557.

Long, C., and J. Wang. 2015. "China's Patent Explosion and Its Quality Implications." *Journal of World Economy* (6): 115-142.

Mairesse, J., and P. Mohnen. 2010. *Using Innovations Surveys for Econometric Analysis.* NBER Working Paper No. 15857. Cambridge, MA, US: National Bureau of Economic Research.

Meng, X. 2012. "Labor Market Outcomes and Reforms in China." *Journal of Economic Perspectives* 26 (4): 75–102.

Rajan, R. G., and L. Zingales. 1998. "Financial Dependence and Growth." *American Economic Review* 88 (3): 559–586.

Ruan, J., and X. Zhang. 2010. *Made in China: Crisis Begets Quality Upgrade.* IFPRI Discussion Paper 01025. Washington, DC: International Food Policy Research Institute.

Tang, H. 2012. “Labor Market Institutions, Firm-specific Skills, and Trade Patterns.” *Journal of International Economics* 87 (2): 337–351.

Topel, R. H. 1986. “Local Labor Markets.” *Journal of Political Economy* 94 (3): 111–143.

Wei, S., and X. Zhang. 2011. “The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China.” *Journal of Political Economy* 119 (3): 511–564.

Xie, Z., and X. Zhang. 2015. “The Patterns of Patents in China.” *China Economic Journal* 8 (2): 122–142.

Xu, B. 2000. “Multinational Enterprises, Technology Diffusion, and Host Country Productivity Growth.” *Journal of Development Economics* 62 (2): 477–493.

Yu, M. 2014. “Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms.” *Economic Journal* 125 (585): 943–988.

Zhang, X., J. Yang, and S. Wang. 2011. “China Has Reached the Lewis Turning Point.” *China Economic Review* 22 (4): 542–554.

RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to www.ifpri.org/pubs/pubs.htm#dp.
All discussion papers can be downloaded free of charge.

1539. *Linking smallholder farmers to commercial markets: Evidence from nongovernmental organization training in Nicaragua*. Ayako Ebata and Manuel A. Hernandez, 2016.

1538. *Can labor market imperfections explain changes in the inverse farm size–productivity relationship?: Longitudinal evidence from rural India.* Klaus Deininger, Songqing Jin, Yanyan Liu, and Sudhir K. Singh, 2016.

1537. *Labor adaptation to climate variability in eastern Africa.* Xiaoya Dou, Clark Gray, Valerie Mueller, and Glenn Sheriff, 2016.

1536. A Dynamic Spatial Model of Agricultural Price Transmission: Evidence from the Niger Millet Market. Anatole Goundan and Mahamadou Roufahi Tankari, 2016.

1535. *Qualitative methods for gender research in agricultural development*. Deborah Rubin, 2016.

1534. *Anchoring bias in recall data: Evidence from Central America*. Susan Godlonton, Manuel A. Hernandez, and Mike Murphy, 2016.

1533. *Contracting by small farmers in commodities with export potential: Assessing farm profits of lentil growers in Nepal.* Anjani Kumar, Devesh Roy, Gaurav Tripathi, P. K. Joshi, and Rajendra P. Adhikari, 2016.

1532. *Rent dispersion in the US agricultural insurance industry.* Vincent Smith, Joseph Glauber, and Robert Dismukes, 2016.

1531. *Long-term drivers of food and nutrition security.* David Laborde, Fahd Majeed, Simla Tokgoz, and Maximo Torero, 2016.

1530. *Understanding compliance in programs promoting conservation agriculture: Modeling a case study in Malawi.* Patrick S. Ward, Andrew R. Bell, Klaus Droppelmann, and Tim Benton, 2016.

1529. *A model of reporting and controlling outbreaks by public health agencies.* Alexander E. Saak and David A. Hennessy, 2016.

1528. *Boserupian pressure and agricultural mechanization in modern Ghana.* Frances Cossar, 2016.

1527. *Agricultural mechanization and agricultural transformation.* Xinshen Diao, Jed Silver, and Hiroyuki Takeshima, 2016.

1526. *Delegation of quality control in value chains*. Alexander E. Saak, 2016.

1525. *Structural transformation and intertemporal evolution of real wages, machine use, and farm size–productivity relationships in Vietnam.* Yanyan Liu, William Violette, and Christopher B. Barrett, 2016.

1524. *Can contract farming increase farmers' income and enhance adoption of food safety practices?: Evidence from remote areas of Nepal.* Anjani Kumar, Devesh Roy, Gaurav Tripathi, P. K. Joshi, and Rajendra P. Adhikari, 2016.

1523. *Effectiveness of food subsidies in raising healthy food consumption: Public distribution of pulses in India.* Suman Chakrabarti, Avinash Kishore, and Devesh Roy, 2016.

1522. *Findings across agricultural public expenditure reviews in African countries*. Stephen D. Mink, 2016.

1521. *Risk and sustainable crop intensification: The case of smallholder rice and potato farmers in Uganda.* Bjorn Van Campenhout, Emmanuel Bizimungu, and Dorothy Birungi, 2016.

1520. *Varietal integrity, damage abatement, and productivity: Evidence from the cultivation of Bt cotton in Pakistan.* Xingliang Ma, Melinda Smale, David J. Spielman, Patricia Zambrano, Hina Nazli, and Fatima Zaidi, 2016.

1519. *Institutional arrangements to make public spending responsive to the poor—(where) have they worked?: Review of the evidence on four major intervention types.* Tewodaj Mogues and Alvina Erman, 2016.

1518. *A poverty-sensitive scorecard to prioritize lending and grant allocation: Evidence from Central America.* Manuel A. Hernandez and Maximo Torero, 2016.

1517. *Can information help reduce imbalanced application of fertilizers in India?: Experimental evidence from Bihar.* Ram Fishman, Avinash Kishore, Yoav Rothler, Patrick S. Ward, Shankar Jha, and R. K. P. Singh, 2016.

1516. *Pakistan's fertilizer sector: Structure, policies, performance, and impacts.* Mubarik Ali, Faryal Ahmed, Hira Channa, and Stephen Davies, 2016.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org