DOI: 10.1111/fima.12280

ORIGINAL ARTICLE

Financial Management WILEY

The effects of rural–urban migration on corporate innovation: Evidence from a natural experiment in China

Deqiu Chen[1] | Huasheng Gao[2] | Jiang Luo[3] | Yujing Ma[4]

[1]Business School, University of International Business and Economics, Beijing, China

[2]Fanhai International School of Finance, Fudan University, Shanghai, China

[3]Nanyang Business School, Nanyang Technological University, Singapore

[4]Business School, University of International Business and Economics, Beijing, China

Correspondence
Huasheng Gao, Fanhai International School of Finance, Fudan University, 318 South Zhongshun Road, Shanghai, China.
Email: huashenggao@fudan.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 71790604, 71872045, 71572035; Shanghai Pujiang Program, Grant/Award Number: 18PJC007; University of International Business and Economics, Grant/Award Number: 841612

Abstract

We show that the migration of low-skilled, rural workers to urban centers has a negative causal effect on innovation of firms in such urban centers. Our tests exploit the staggered relaxation of city-level household registration system in China, which facilitates rural residents to migrate to cities. We find a significant decrease in innovation for firms headquartered in cities that have adopted such policies relative to firms headquartered in cities that have not. Overall, our results support the view that an abundant supply of low-skilled workers increases the benefit of using existing low-skilled technology and thus reduces firms' incentive to innovate.

KEYWORDS

innovation, patents, migration, low-skilled worker, household registration, hukou

1 | INTRODUCTION

In this paper, we examine the effects of rural–urban migration on innovation in the host areas. This research question is important for at least two reasons. First, in the last few decades, hundreds of millions of people have moved from rural to urban centers. For example, China has experienced a massive rural–urban migration in the last 40 years; the percentage of rural population among total population in China has decreased steadily from approximately 80% in 1978 to approximately 40% in 2018 (Figure 1). This phenomenon makes it important to study the consequences of such rural–urban migration on economic growth. Our study focuses on innovation, because it is one key aspect of growth policy. Second, most of this rural–urban migration has been taking place in China, India, and other developing countries, which are also important players in the geography of innovation, and engines of world economic growth. These countries have an urgent need to absorb their excess rural workforce that has been steadily moving into cities,

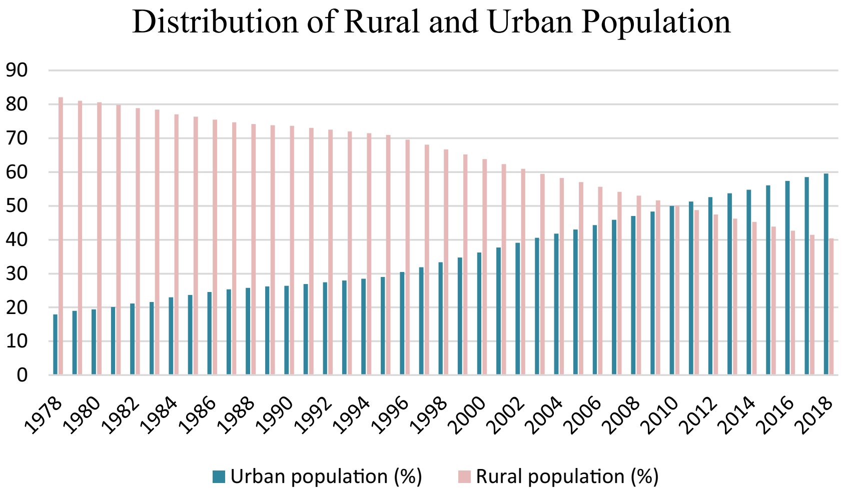

FIGURE 1 Distribution of rural and urban population by year. This figure plots the percentage distribution of rural and urban population in China during 1978–2018. Data are obtained from the National Bureau of Statistics of China [Color figure can be viewed at wileyonlinelibrary.com]

while a sizable part of innovation in these countries is focused on automation. This naturally begs the question whether policies facilitating rural–urban migration are at loggerheads with those promoting innovation.

In this paper, we shed light on this issue and document a negative effect of rural–urban migration on firms' innovation in the migrant host areas, using a quasi-natural experiment in China. Our empirical identification strategy is based on the staggered relaxation of China's city-level household registration system (also known as the *hukou* system), which reduces the restriction for rural residents (who are largely undereducated and low-skilled) to migrate to nearby cities. We use these policy changes to capture an exogenous increase in the inflow of low-skilled migrant workers and examine the subsequent changes in corporate innovation in the host areas.

This setting is highly appealing from an empirical standpoint for two reasons. First, the motivation behind such changes in the household registration system is to provide rural-to-urban migrants equal access to the urban welfare system and reduce the rural–urban divide. As these policy changes were not made with the intention of hindering innovation, potential effects on innovation are likely to be an unintended consequence. Second, the staggered policy changes in several Chinese cities provide a set of counterfactuals for how corporate innovation would have evolved in the absence of such policy changes, and enable us to identify their effects in a difference-in-differences framework. Because multiple shocks affect different firms exogenously at different times, we can avoid the common identification difficulty faced by studies with a single shock: the potential biases and noise coinciding with the shock that directly affects corporate innovation (Roberts & Whited, 2013).

We expect the rural–urban migration to decrease corporate innovation because companies are less likely to adopt new technologies or innovate when there is an abundant supply of low-skilled labor (Lewis, 2011; Peri, 2012). Suppose a firm is currently using a preexisting, low-skilled technology operated by low-skilled workers and is considering to invest in some risky research and development (R&D) projects to develop a high-quality technology operated by high-skilled workers. The likelihood of making such an R&D investment depends on the cost of the R&D expenditure and the relative profit of using the new technology versus the existing one. An abundant supply of low-skilled workers in the labor market would increase the benefit of using the existing low-skilled technology and thus would enhance the hurdle for the firm to pursue the new high-quality technology, which in turn would hinder corporate innovation. Anecdotal evidence supports this view. For example, Habakkuk (1962) claims that technological progress was slower in Britain than in the United States in the 19th century because Britain had a large supply of cheap, low-skilled workers.

Elvin (1972) suggests that a sophisticated spinning wheel used for hemp in 14th-century China was later abandoned and was not used for cotton largely because an abundance of cheap Chinese labor made it unprofitable relative to existing low-skilled technologies.

Using a panel of 30,216 public Chinese firm-year observations from 1997 to 2016 and a difference-in-differences approach, we show that an exogenous increase in the inflow of low-skilled migrant workers subsequently leads to a significant decrease in innovation outputs. On average, firms headquartered in cities that made such a policy change experienced a decrease in the number of patents by 14%, relative to firms headquartered in cities that did not adopt such a policy.

The identifying assumption central to a causal interpretation of the difference-in-differences estimation is that treated and control firms share parallel trends prior to the policy changes. Our tests show that their pretreatment trends are indeed indistinguishable. Moreover, most of the impact of the household registration policy on innovation occurs 3 years after the policy's enactment, which suggests a causal effect.

However, it is possible that the changes of household registration policies are triggered by local business conditions that in turn influence firms' innovation. To mitigate this concern, we additionally control for local business conditions such as city-level gross domestic product (GDP), population, education, income per capita, investment on science and technology, and housing price. Our inferences are largely unchanged. In further tests, we exploit the fact that economic conditions are likely to be similar in neighboring cities, whereas the effects of these city-level policies stop at city borders. This discontinuity in the household registration policy allows us to difference away any unobserved confounding factors as long as they affect both the treated cities and their neighbors. By comparing treated firms to their immediate neighbors, we can better identify how much of the observed innovation change is due to the household registration policy rather than other shocks to local business conditions. When we difference away changes in local business conditions by focusing on treated and control firms closely located on either side of a city border, we continue to find a significant decrease in firms' innovation after their cities loosen their household registration policy, relative to their neighboring firms. These results indicate that the observed decrease in innovation following the relaxation of household registration system is not driven by local economic shocks.

To provide further evidence that the effects of the household registration relaxation on innovation are indeed tied to low-skilled migrant workers, we apply a triple difference-in-differences approach to examine heterogeneous treatment effects. We find that the treatment effects are stronger for firms that operate in labor-intensive industries and for firms in cities with stronger enforcement of *hukou* relaxation (measured by the percentage of people who have obtained local urban *hukou* following *hukou* relaxation). These cross-sectional variations in the treatment effects further increase our confidence that the impact of household registration policy changes on innovation is indeed related to the rural–urban migration of low-skilled workers.

Our study sheds light on the real consequences of rural–urban migration, which has recently been at the center of many governments' political and economic agendas. Economists have studied extensively the impact of migration on several economic and social indicators in migrant host areas, such as natives' wages (Borjas, 2003; Ottaviano & Peri, 2012), employment opportunities (Card, 2005; Pischke & Velling, 1997), firm productivity (Peri, 2012), crime rate (Bell, Fasani, & Machin, 2013; Bianchi, Buonanno, & Pinotti, 2012), and so forth. Complementing this strand of literature, we provide evidence that such migration has a negative causal effect on corporate innovation, a key element for economic growth.

Our paper is closely related to Lewis (2011), who shows that the automation machinery adopted by U.S. manufacturing plants complements middle-skilled workers relative to low-skilled workers. Our paper differs from Lewis (2011) along the following three dimensions. First, the research question is different. The purpose of Lewis (2011) is to verify technology-skill complementarity, whereas the purpose of our paper is to understand whether policies facilitating rural–urban migration hinder innovation. Second, Lewis (2011) studies U.S. manufactures over the sample 1988–1993, whereas our paper studies all China public firms in the recent 20 years. Rural–urban migration is much more prominent in China than in United States, and innovation is increasingly more important for economic growth in recent years, especially for developing countries. Thus, our setting is more economically relevant. Third, the identification strategy of Lewis (2011) is to construct an instrumental variable based on the flow of immigrants, whereas our identification strategy is to exploit the staggered *hukou* relaxation across different Chinese cities. The difference-in-differences estimation in our paper enables us to better investigate the before versus after effects of rural–urban migration.

The remainder of the paper is organized as follows. Section 2 reviews the background on China's household registration policy. Section 3 describes our sample and key variable construction. Section 4 presents our main empirical results. We implement additional robustness check in Section 5 and conclude in Section 6.

2 | BACKGROUND ON CHINA'S HOUSEHOLD REGISTRATION POLICY

The household registration system was established in China in the 1950s, following the issuance of Soviet-style internal passports to all Chinese citizens. This system served as an invisible wall to prevent the rural labor force from moving into urban areas. It still broadly divides Chinese citizens into two classes: those living in rural areas versus those living in urban areas. Such classification is largely based on a resident's place of birth and the household registration status of a resident's parents. Under this system, some 800 million rural residents are treated as inferior citizens, deprived of the right to settle in cities and the basic welfare services enjoyed by urban residents, ranging from small benefits such as buying a city bus pass to enrolling their children in public schools in cities where they work.

The original purpose of this system was to promote the development of heavy industry, a high priority at the time, and to speed up industrialization generally. In order to accumulate capital for investment, the system kept the rural labor force in agricultural sectors. It also limited the number of people who had access to low-priced food, guaranteed nonagricultural employment, and subsidized urban social benefits on housing, education, medical care, and so on. It is widely documented that the *hukou* system significantly restricts rural-to-urban migration, denies China's rural population the access to quality education and urban employment, and is a major factor contributing to China's rural–urban inequality (see, e.g., Afridi, Li, & Ren, 2015; Liu, 2005; Wu & Treiman, 2004).

The *hukou* system was applied stringently. Public security bureaus controlled place-to-place migration, and it was virtually impossible to move from a rural area to an urban area without authorized plans or official agreement. Because the "reform and opening-up" policy was instituted in the 1980s, controls over urban-to-rural labor mobility started to relax. In the mid-1980s, the Chinese government introduced a system of temporary residence permits that allowed people with an agricultural *hukou* to move to urban areas as long as they could provide for their food and lodging. This policy unleashed a massive flow of migrants into cities, with more than 60 million migrants moving to cities within the first 10 years of its application. Starting in the late 1990s, China experimented with a variety of reforms to further relax the restriction of the *hukou* system; for example, in 1997, the State Council began permitting families of migrant workers to alter their *hukou* status (i.e., spouses and children).

Reform further accelerated in the early 2000s when the State Council gave local governments the power to decide their own *hukou* policies. Several Chinese cities have since adopted some *hukou* relaxation, such as abolishing the distinction between rural and urban *hukou* or lowering the hurdle for migrant workers to obtain local urban *hukou*.

We mainly collect our information regarding cities loosening their *hukou* policies from the *China City Statistical Yearbooks*. We record the events of such policy changes as the first year in which the city abolished the distinction between rural and urban *hukou*, or when the city lowered the criteria for migrant workers from rural areas to obtain local urban *hukou*. In our sample, 30 cities relaxed their *hukou* policies (Table 1).[1] The Chinese government decides to relax the *hukou* system policies for various reasons, including improving human rights for migrant workers, enhancing labor mobility, facilitating urbanization, rebalancing the economy by increasing the consumption share of GDP, gaining international legitimacy, and more.

[1] In Table 1, we focus on *hukou* relaxation occurring during 2000–2011 so that for each relaxation, we can track firms' innovation from several years before the event to several years afterward.

TABLE 1 List of *hukou* relaxation

City	Year of the policy becoming effective
Tonghua (通化)	2000
Urumqi (乌鲁木齐)	2001
Beijing (北京)	2002
Fuzhou (福州)	2002
Jiaxing (嘉兴)	2002
Jincheng (晋城)	2002
Haining (海宁)	2003
Nanning (南宁)	2003
Taizhou (泰州)	2003
Tianjin (天津)	2003
Xiamen (厦门)	2003
Zhengzhou (郑州)	2003
Changde (常德)	2004
Nanjing (南京)	2004
Shanghai (上海)	2004
Shenzhen (深圳)	2004
Chengdu (成都)	2005
Harbin (哈尔滨)	2006
Xian (西安)	2006
Yunchen (运城)	2006
Taiyuan (太原)	2007
Anshan (鞍山)	2008
Dalian (大连)	2008
Kunming (昆明)	2008
Shenyang (沈阳)	2008
Zhuhai (珠海)	2008
Guangzhou (广州)	2009
Qiqihaer (齐齐哈尔)	2009
Chongqing (重庆)	2010
Yinchuan (银川)	2011

Note: This table reports the year in which each city implemented its *hukou* relaxation, which relaxes the restrictions for migrant workers from rural areas to obtain local urban *hukou*. Chinese names of the cities are reported in parentheses.

It is worth pointing out that local economic conditions are an important driver of the relaxation of *hukou* policies. Local governments are more likely to relax *hukou* restrictions when local economic conditions are good (i.e., when there is a greater demand for labor supply). In contrast, when local economic conditions are bad, the local government is less likely to relax such restrictions to avoid the financial burden of providing social welfare to new migrants and to secure employment opportunities for incumbent urban residents (Cai, 2011). This in fact works against our finding a negative effect of household registration relaxation on firm innovation, considering that good economy conditions are likely positively associated with innovation outputs. Later in the paper (Table 11), we conduct a formal test and find that relaxing *hukou* is indeed positively related to local economic conditions. We also find that after controlling for local economic conditions, the implementation of such policies is unrelated to local firms' preexisting innovation activities, supporting the exogeneity of such policy changes to corporate innovation.

3 | SAMPLE FORMATION AND VARIABLE CONSTRUCTION

We start with all Chinese public companies in the Shanghai and Shenzhen stock exchanges during 1997–2016, obtained from the China Stock Market & Accounting Research (CSMAR) database, from which we collect the firms' financial information. We start in 1997 because Chinese patent information is widely available only from that year onward.

We use the number of patents to measure a firm's success of R&D investment in corporate innovation, which has been widely used in the literature since Scherer (1965) and Griliches (1981). Information about patent grants is from the State Intellectual Property Office of China (SIPO). For each patent, SIPO provides information on the patent application date, application identification, publication identification, granting date, and patent identification, along with inventors and application institutions. We extract patent applications filed by the sample firms, including those filed by their subsidiaries, from the SIPO database to construct measures for a firm's innovative outcomes. The Chinese patent system classifies patents into three types: invention patents, utility model patents, and design patents. Invention patents refer to those granted for a new technical solution to a product or an industrial process. Utility model patents are for new and practical technical solutions relating to certain characteristics of a product, such as the product's shape and structure. This type of patent demonstrates new functional aspects of a product. Design patents are for a product's new shape, pattern, or color that makes the product more attractive and industrially applicable. It is worth noting that the SIPO database does not provide reliable information on patent citations; thus, we are unable to use patent citations to capture the quality of each patent. As pointed out by Tan, Tian, Zhan, and Zhao (2015), invention patents represent the most original inventions among all three types of patents; thus, the number of invention patents can also measure the quality of the patents produced by a firm.

We control for a vector of firm and industry characteristics that may affect a firm's innovation productivity. These variables include firm size, firm age, asset tangibility, leverage, cash holding, R&D expenditures, capital expenditures, return on assets (ROA), and Tobin's *Q*. All explanatory variables are lagged by 1 year. To minimize the effect of outliers, we winsorize all variables at the 1st and 99th percentiles. Detailed variable definitions are provided in the Appendix. Our final sample consists of 30,216 firm-year observations from 1997 to 2016.

Table 2 provides summary statistics. On average, firms in our sample have filed 16 patents (which were subsequently granted) per year. Out of these patents, three are invention patents and 13 are utility and design patents. Our average sample firms have book value assets of RMB 8.68 billion (or approximately USD 1.3 billion) and are 13 years old. They hold a sizable amount of cash with a cash ratio of 19% of total assets. The average R&D and capital expenditure account for 0.8% and 5.5% of total assets, respectively. The average firms are levered with a book leverage ratio of 47.3%, and tangible assets account for 25.2% of total assets. In terms of performance, sample firms perform well with an average ROA of 6.2% and Tobin's *Q* of 2.57.

4 | EMPIRICAL RESULTS

4.1 | Effect of *hukou* relaxation on migration, R&D, and employment

To provide evidence that *hukou* relaxation indeed leads to a greater inflow of low-skilled migration, we conduct a standard difference-in-differences test at the city level. Specifically, our specification is as follows:

$$Migration_{i,t} = \alpha + \beta_1 Relaxation_{s,t-1} + \beta_2 Other\ City\ Characteristics_{i,t-1} + City\ FE + \ Year\ FE + \varepsilon_{i,t}. \quad (1)$$

For the treated cities, the indicator variable *Relaxation* equals 1 for the period after the relaxation of the city-level household registration system, and 0 otherwise. For the control group, the indicator variable *Relaxation* always takes the value of 0. The year fixed effects enable us to control for intertemporal migration shocks. The city fixed effects

TABLE 2 Summary statistics

Variable	Mean	SD	P25	Median	P75
All patent	16.145	44.204	0.000	2.000	12.000
Invention patent	3.429	9.940	0.000	0.000	4.000
Utility model and design patent	12.717	36.872	0.000	0.000	9.000
Total assets (RMB billion)	8.680	28.317	0.952	1.992	4.746
Cash	0.190	0.150	0.083	0.148	0.252
Leverage	0.473	0.241	0.302	0.462	0.620
R&D	0.008	0.014	0.000	0.000	0.011
Capex	0.055	0.056	0.014	0.037	0.078
ROA	0.062	0.045	0.032	0.051	0.077
Firm age	13.056	5.922	8.000	13.000	17.000
Tobin's *Q*	2.566	1.946	1.355	1.978	3.079
Tangible	0.252	0.178	0.112	0.219	0.363

Note: The sample consists of 30,216 firm-year observations from 1997 to 2016. We obtain patent information from the State Intellectual Property Office (SIPO) of China and financial information from the China Stock Market & Accounting Research (CSMAR) database. Definitions of all variables are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles.

allow us to control for time-invariant differences across cities. Given that our treatment is defined at the city level, we cluster standard errors by city (Chen, Li, Xiao, & Zou, 2014; Moulton, 1986).

The coefficient of interest in this model is the β_1 coefficient. As explained by Imbens and Wooldridge (2009), the employed fixed effects lead to β_1 being estimated as the *within-city* differences before and after the *hukou* policy change as opposed to similar before–after differences in cities that did not experience such a change during the same period.

It is helpful to consider an example. Suppose we want to estimate the effect of the relaxation of household registration in Beijing in 2002 on migration. We can subtract the migration measure before the policy change from the migration measure after the policy change in Beijing. However, economy-wide shocks may occur at the same time and affect migration in 2002. To difference away such factors, we calculate the same difference in the migration measure in a control city that does not adopt such a policy change. Finally, we calculate the difference between these two differences, which represents the incremental effect of the policy change on migration in Beijing compared to other cities.

The results are reported in Table 3, Panel A. In columns (1) and (2), the dependent variable is *ln(number of people newly obtaining urban hukou)* in a given city in a given year, which measures the number of people who were previously holding rural *hukou* and have newly obtained their local urban *hukou*. We find that the coefficients on the *Relaxation* indicator are positive and significant in both columns. Taking column (2) for example, the coefficient on the *Relaxation* indicator is 0.052 and is significant at the 5% level, which indicates that *hukou* relaxation leads to a significant increase in the number of rural-to-urban migration by approximately 5% ($= e^{0.052} - 1$), relative to the cities that did not implement such policies.

As a robustness check, in columns (3) and (4) we further normalize the number of people newly obtaining urban *hukou* by the total number of people with an urban *hukou* in the city. We continue to find positive and significant coefficients on the *Relaxation* indicator.

Further, if *hukou* relaxation reduces a firm's incentive to innovate because it raises the profitability of using low-skilled technology, we expect that firms cut R&D expenditure and hire more employees (especially low-skilled employees) following *hukou* relaxation. We investigate this prediction in Table 3, Panel B.

In column (1) of Panel B, the regression specification is the same as that in Equation (1), except that the dependent variable is *R&D* expenditure and that the regression is at the firm-year level. The coefficient on the *Relaxation* indicator

TABLE 3 Effects of *hukou* relaxation on migration, R&D, and employment

Panel A: City-level analysis on migration				
	(1)	(2)	(3)	(4)
	ln(Number of people newly obtaining urban *hukou*)		Percentage of people newly obtaining urban *hukou*	
Relaxation	0.072**	0.052**	0.036**	0.014**
	(.031)	(.046)	(.027)	(.048)
ln(City GDP)		0.017**		0.008**
		(.016)		(.012)
ln(City population)		0.042**		0.026**
		(.011)		(.013)
ln(Number of universities in the city)		−0.005		−0.004
		(.453)		(.342)
City income per capita		0.001		−0.001
		(.823)		(.782)
City expenditure on science and technology		0.289*		0.052*
		(.071)		(.067)
City housing price index		−0.008		−0.005
		(.531)		(.327)
Constant	0.245***	−0.394**	0.128***	−0.289**
	(.000)	(.047)	(.000)	(.023)
Observations	5,789	5,789	5,789	5,789
Year FEs	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes
R^2	.226	.246	.238	.254

Panel B: Firm-level analysis on R&D and employment					
	(1)	(2)	(3)	(4)	(5)
	R&D	ln(Number of employees)	Labor intensity	Percentage of technicians	Percentage of employees with a bachelor's degree
Relaxation	−0.004**	0.112**	0.006*	−0.007**	−0.026**
	(.024)	(.037)	(.082)	(.041)	(.042)
Cash	−0.002***	−0.213**	−0.014**	−0.006	0.113**
	(.008)	(.013)	(.042)	(.711)	(.026)
Firm size	0.002	0.521***	0.023***	−0.011	−0.142***
	(.114)	(.002)	(.003)	(.331)	(.002)
Leverage	−0.002	−0.213**	−0.014**	0.017	−0.006
	(.217)	(.038)	(.025)	(.137)	(.426)
Capex	0.002	0.223	0.025**	0.004	−0.034
	(.312)	(.145)	(.012)	(.482)	(.492)

(Continues)

TABLE 3 (Continued)

Panel B: Firm-level analysis on R&D and employment					
	(1)	(2)	(3)	(4)	(5)
	R&D	ln(Number of employees)	Labor intensity	Percentage of technicians	Percentage of employees with a bachelor's degree
ROA	−0.001	−0.312**	−0.036***	0.002	0.085*
	(.217)	(.011)	(.012)	(.452)	(.076)
Firm age	0.001***	−0.051***	−0.004***	−0.001	0.014***
	(.000)	(.000)	(.000)	(.259)	(.000)
Tobin's *Q*	0.001	−0.002	0.001	0.001	−0.001
	(.253)	(.715)	(.624)	(.822)	(.713)
Tangible	0.001	0.824***	0.036***	−0.005	−0.138***
	(.275)	(.000)	(.000)	(.703)	(.008)
Constant	−0.001	−4.236***	0.051	0.217***	3.114***
	(.136)	(.000)	(.114)	(.002)	(.000)
Observations	30,216	30,216	30,216	30,216	30,216
Year FEs	Yes	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes	Yes
R^2	.489	.712	.706	.351	.726

Note: This table reports the difference-in-differences tests that examine the impacts of *hukou* relaxation on migration, R&D, and employment. Panel A uses city-year-level regressions. The dependent variable in columns (1) and (2) is *ln(number of people newly obtaining urban hukou)*; the dependent variable in columns (3) and (4) is *Percentage of people newly obtaining urban* hukou. Panel B uses firm-year-level regressions. The dependent variables are *R&D*, *ln(number of employees)*, *Labor intensity*, *Percentage of technicians*, and *Percentage of employees with a bachelor's degree* in columns (1)–(5), respectively. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

is negative and significant at the 5% level, indicating that firms reduce R&D expenditure after their cities relax their *hukou* policies. This result indicates that *hukou* relaxation leads to a reduction in the input of the innovation process.

We further examine *ln(number of employees)* as the dependent variable in column (2). We find a positive and significant (at the 5% level) coefficient on the *Relaxation* indicator. This result indicates that while firms cut their R&D expenditure, they are actually expanding by hiring more employees. Examining firms' labor intensity as the dependent variable in column (3), we find that the coefficient on the *Relaxation* indicator is significantly positive, indicating that firms rely more on labor after *hukou* relaxation.

Last, we investigate the composition of employees in columns (4) and (5). The dependent variable in column (4) is the percentage of technician employees among all employees; the dependent variable in column (5) is the percentage of employees with a bachelor's degree among all employees. These two variables measure the skill and education level of the firm's workforce. The coefficients on the *Relaxation* indicator are negative and significant at the 5% level in both columns, indicating that the skill and education level of the employees decrease following *hukou* relaxation. This result indicates that firms hire more low-skilled labor after their cities relax *hukou* restrictions.

In summary, Table 3 shows that *hukou* relaxation indeed leads to a greater inflow of low-skilled migrant workers from rural areas into the host cities. Moreover, following the relaxation of the *hukou* system, firms in the host areas decrease their R&D expenditure, hire more employees (especially low-skilled employees), and become more labor intensive. These findings indicate that *hukou* relaxation leads to significant rural–urban migration and has important effects on firm policies.

4.2 | Baseline results

Several Chinese cities relaxed their *hukou* policies in different years during the sample period. Thus, we can examine the before–after effect of *hukou* relaxation in affected cities (the treatment group) compared to the before–after effect in cities in which such a relaxation was not effected (the control group). This is a difference-in-differences test design in multiple treatment groups and multiple time periods as employed by Bertrand, Duflo, and Mullainathan (2004), Imbens and Wooldridge (2009), Atanassov (2013), and Gao and Zhang (2017). We implement this test through the following regression:

$$Innovation_{i,t} = \alpha + \beta_1 Relaxation_{s,t-1} + \beta_2 Other\ firm\ characteristics_{i,t-1} + Firm\ FE + \ Year\ FE + \varepsilon_{i,t}, \tag{2}$$

where *i* indexes firms, *s* indexes the city in which the firms' headquarters are located, and *t* indexes the year. The dependent variable is a proxy for innovation performance.

Table 4 presents the regression results. The coefficient estimates on the *Relaxation* indicator are negative and statistically significant in all columns. The dependent variable in column (1) is *ln(1+all patent)*, and we find that the coefficient on the *Relaxation* indicator is –0.132 and significant at the 5% level, suggesting a negative effect of the policy change on corporate innovation. The economic magnitude is also sizable: the relaxation of the household registration system leads to a decrease in the number of patents by approximately 14% ($= e^{0.132} - 1$).

Examining *ln(1+invention patent)* as the dependent variable in column (2), we find that the coefficient on the *Relaxation* indicator is –0.087 and is significant at the 5% level, which implies that *hukou* relaxation leads to a decrease in the number of invention patents by approximately 9% ($= e^{0.087} - 1$). We examine *ln(1+ utility and design patent)* in column (3). The coefficient on the *Relaxation* indicator is –0.107 and is significant at the 5% level, indicating a decrease in the number of utility and design patents by 11% ($= e^{0.107} - 1$).

Taken together, these results indicate a negative effect of *hukou* relaxation on innovation outputs, therefore supporting our hypothesis.

4.3 | The pretreatment trends

The validity of difference-in-differences estimation depends on the parallel trends assumption: absent the treatment, treated firms' innovation would have evolved in the same way as that of control firms. Table 5 investigates the pretreatment trend between the treated group and control group. In particular, we define seven dummy variables, *Year −2*, *Year −1*, *Year 0*, *Year 1*, *Year 2*, *Year 3*, and *Year 4*$^{+}$, to indicate the year relative to the relaxation of the *hukou* system. For example, *Year 0* indicates the year in which *hukou* relaxation is implemented; *Year −2* indicates that it is 2 years before the relaxation; and *Year 2* indicates that it is 2 years after the relaxation. Then we reestimate Equation (2) by replacing the *Relaxation* indicator with the seven indicators above.

The coefficients on *Year −2* and *Year −1* indicators are especially important because their significance and magnitude indicate whether there is any difference in innovation between the treatment group and the control group prior to the policy change. The coefficients on both indicators are close to zero and not statistically significant across all three columns, suggesting that the parallel trends assumption of the difference-in-differences approach is not violated. The absence of any significant lead effects has at least three implications. First, the implementation of *hukou* relaxation does not seem to be anticipated by the treated firms. Second, even if some treated firms anticipated such policy changes, the actual rural-to-urban migration did not change until the policies took effect. Third, the negative effect of *hukou* relaxation on innovation is not the result of policymakers simply responding to past innovation activities, mitigating the reverse causality concern. (This result is also consistent with the evidence in Table 11, which shows that *hukou* relaxation policies are indeed unrelated to the pre-event corporate innovation activities.)

TABLE 4 Effect of *hukou* relaxation on innovation

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Relaxation	−0.132**	−0.087**	−0.107**
	(.025)	(.012)	(.013)
Cash	−0.142	−0.098	−0.063
	(.148)	(.164)	(.413)
Firm size	0.161***	0.112***	0.146***
	(.000)	(.000)	(.000)
Leverage	−0.052	−0.021	−0.039
	(.216)	(.512)	(.326)
R&D	19.253*	16.423*	12.128
	(.056)	(.047)	(.147)
Capex	−0.054	−0.071	−0.012
	(.558)	(.524)	(.917)
ROA	−0.102	−0.073	−0.108
	(.213)	(.142)	(.156)
Firm age	0.042***	0.026***	0.028***
	(.000)	(.000)	(.000)
Tobin's *Q*	0.006	0.003	0.009
	(.215)	(.317)	(.163)
Tangible	0.163*	0.118	0.119
	(.078)	(.106)	(.127)
Constant	−3.287***	−2.156***	−3.243***
	(.000)	(.000)	(.000)
Observations	30,216	30,216	30,216
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R^2	.612	.557	.589

Note: This table reports the difference-in-differences tests that examine the impacts of *hukou* relaxation on corporate innovation. For cities that ever relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 otherwise. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change, and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The coefficients on *Year 0*, *Year 1*, and *Year 2* indicators are also small in magnitude and insignificant in all three columns. The impact of the policy change starts to show up 3 years after the enactment: the coefficients on *Year 3* indicator become significantly negative in all three columns. The coefficients on *Year 4+* are more than twice as large as the coefficients on the *Year 3* indicator for all three innovation measures, indicating that it takes a few years to reveal the full impact of *hukou* relaxation on corporate innovation. This is understandable given that labor migration and innovation are usually a long-term process.

Overall, Table 5 shows that the treated group and the control group share a similar trend in innovation prior to the treatment, thus supporting the parallel trends assumption associated with the difference-in-differences estimation.

TABLE 5 Testing for pretreatment trends and reversals

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Year −2	0.045	0.038	0.041
	(.312)	(.178)	(.345)
Year −1	−0.023	0.012	0.017
	(.723)	(.713)	(.758)
Year 0 (event year)	−0.021	0.003	−0.014
	(.524)	(.623)	(.589)
Year 1	−0.057	−0.038	−0.043
	(.157)	(.312)	(.287)
Year 2	−0.112	−0.078	−0.093
	(.142)	(.191)	(.266)
Year 3	−0.136**	−0.065*	−0.092*
	(.032)	(.086)	(.079)
Year 4+	−0.212***	−0.186***	−0.196***
	(.000)	(.007)	(.006)
Cash	−0.132	−0.068	−0.049
	(.142)	(.198)	(.347)
Firm size	0.145***	0.102***	0.136***
	(.000)	(.000)	(.000)
Leverage	−0.052	−0.038	−0.041
	(.162)	(.313)	(.297)
R&D	20.168*	18.135*	15.279
	(.059)	(.062)	(.125)
Capex	−0.041	−0.038	0.025
	(.612)	(.549)	(.723)
ROA	−0.121	−0.106	−0.114
	(.213)	(.126)	(.157)
Firm age	0.053***	0.041***	0.042***
	(.000)	(.000)	(.000)
Tobin's Q	0.006	0.003	0.009
	(.312)	(.423)	(.215)
Tangible	0.127	0.109	0.113
	(.124)	(.134)	(.142)
Constant	−3.127***	−2.235***	−3.089***
	(.000)	(.000)	(.000)
Observations	30,216	30,216	30,216
Year FE	Yes	Yes	Yes

(Continues)

TABLE 5 (Continued)

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Firm FE	Yes	Yes	Yes
R^2	.624	.538	.596

Note: This table investigates the pretreatment trends between the treated group and the control group. The indicator variables *Year −2*, *Year −1*, *Year 0*, *Year 1*, *Year 2*, *Year 3*, and *Year 4*$^+$ indicate the year relative to *hukou* relaxation. For example, the *Year 1* indicator takes the value of 1 if it is one year after a city adopts such a policy, and 0 otherwise. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Moreover, Table 5 also indicates that most of the *hukou* relaxation's impact on innovation occurs 3 years after it is implemented, which suggests a causal effect.

4.4 | Confounding local business conditions

Location is one important common factor that likely induces an association between *hukou* policies and corporate innovation. In this section, we implement two tests to address this issue. In our first test, we additionally control for a set of observable city characteristics in the regression. In our second test, we difference away unobservable local business conditions by focusing on treatment firms and their neighboring control firms. In both tests, we continue to find a significant decrease in innovation after *hukou* relaxation.

Table 6 presents our first test. In addition to our usual set of explanatory variables used in Table 4, we also account for various time-varying, city-level variables in our regressions. Given that richer and larger cities may have the resources to provide a higher level of innovation, we include the logarithm of GDP and per capita income in a city. We additionally control for the logarithm of city population. Further, investment in education and R&D is another factor that may lead to differences in patenting. Therefore, we also control for a city's intellectual resources using the number of universities and the city's expenditure for science and technology. Local housing price may be correlated to *hukou* policies and corporate innovation; therefore, we also include housing price index as a control variable. These city-level data are collected from the *China Statistical Yearbooks*.

We find that the relaxation of *hukou* policies continues to have a negative and (statistically and economically) significant impact on corporate innovation. Compared to Table 4, the coefficient on the *Relaxation* dummy is largely unchanged. We find that city-level variables have no significant impact on corporate innovation, probably because we have already controlled for firm fixed effects in the regression.

Although the above test accounts for "observable" local business conditions, some unobservable local economic shocks may be associated with both the relaxation of *hukou* restriction and corporate innovation. In our second test, we exploit the discontinuity of *hukou* policy and examine the innovation change in the treatment firms relative to their neighboring control firms. The logic is as follows. Suppose that a *hukou* relaxation is driven by unobserved changes in local business conditions and that it is these changes (not the *hukou* relaxation) that influence corporate innovation. Then both firms in treated cities and their neighbors in untreated cities just across the city border would spuriously appear to react to the policy changes, because economic conditions, unlike the city-level *hukou* policy, have a tendency to spill across city borders (Heider & Ljungqvist, 2015). In this case, the change in innovation in treated firms should be no different from that in the neighboring control firms.

To examine this possibility, we match each treated firm to a control firm that is in the same industry, is in an adjacent city that has not relaxed its *hukou* policies, and is closest to the treated firm in distance. Obviously, treated firms may not necessarily share the same local economic condition with its "closest" control firm if the treated firm is in the middle of a large city. To alleviate this concern, we further require that the distance between the treated firm and its matched

TABLE 6 Controlling for city-level characteristics

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Relaxation	−0.112**	−0.062**	−0.076**
	(.025)	(.013)	(.052)
Cash	−0.163	−0.089	−0.112
	(.131)	(.217)	(.327)
Firm size	0.165***	0.123***	0.147***
	(.000)	(.000)	(.000)
Leverage	−0.032	−0.008	−0.026
	(.372)	(.542)	(.493)
R&D	17.228*	15.168*	13.892
	(.081)	(.062)	(.129)
Capex	−0.113	−0.082	−0.046
	(.423)	(.414)	(.623)
ROA	−0.123	−0.093	−0.117
	(.213)	(.165)	(.198)
Firm age	0.042***	0.022***	0.032***
	(.000)	(.000)	(.000)
Tobin's *Q*	0.007	0.005	0.006
	(.481)	(.326)	(.353)
Tangible	0.152**	0.117*	0.125*
	(.043)	(.062)	(.052)
ln(City GDP)	0.023	0.008	0.015
	(.214)	(.639)	(.211)
ln(City population)	−0.042	−0.038	−0.042
	(.559)	(.218)	(.548)
ln(Number of universities in the city)	0.041	0.037	0.047
	(.203)	(.193)	(.123)
City income per capita	0.009	0.004	0.007
	(.117)	(.163)	(.212)
City expenditure on science and technology	−1.424	−1.128	−1.173
	(.214)	(.269)	(.225)
City housing price index	−0.017	−0.012	−0.019
	(.628)	(.517)	(.669)
Constant	−3.269***	−2.163***	−3.155***
	(.000)	(.000)	(.000)
Observations	30,216	30,216	30,216
Year FEs	Yes	Yes	Yes

(Continues)

TABLE 6 (Continued)

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Firm FEs	Yes	Yes	Yes
R^2	.626	.538	.574

Note: This table reports the difference-in-differences tests that examine the impacts of *hukou* relaxation on corporate innovation, controlling for city-level characteristics. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change, and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

untreated firm be within 100 miles.[2] If the distance between the treated firm and its closest control firm is more than 100 miles, we drop this pair from our sample. By doing so, we increase our confidence that our treated firm and control firm are truly close to each other geographically and thus face similar local economic shocks. Then, we reestimate Equation (2) by focusing on this subsample of firms across the city border.

Table 7 presents the results. Restricting our sample to the pairs of neighboring treated and control firms reduces the sample to 18,661 firm-year observations; yet, we still find negative and significant coefficients (at the 5% level) on the *Relaxation* indicator in all three columns. Under the identifying assumption that the control firms are exposed to similar local economic conditions and hence the change in innovation output of the treated firms should be no different from that of their control firms, our findings suggest that any unobservable confounding local economic conditions cannot be driving the observed impact of *hukou* relaxation on corporate innovation.

Overall, the results in Tables 6 and 7 indicate that our main findings are unlikely driven by confounding local business conditions.

4.5 | Heterogeneous treatment effects

To provide further evidence that the effects of *hukou* policies on innovation are indeed tied to the migration of low-skilled workers, in this subsection we examine the cross-sectional variation of the treatment effects. Examining heterogeneous treatment effects can further help to alleviate the concern that some omitted firm or city variables are driving our results, because such variables would have to be uncorrelated with all the control variables we include in the regression model; further, they would also have to explain the cross-sectional variation of the treatment effects. As pointed out by Claessens and Laeven (2003) and Raddatz (2006), it is less likely to have an omitted variable correlated with the interaction term than with the linear term.

First, considering that the *hukou* policy affects productivity associated with labor, not physical capital, the treatment effects should be stronger for firms that rely more on labor. We measure *Labor intensity* as employee wage as a proportion of the firm's sales. Then we reestimate Equation (2) by adding the *Labor insensitivity* variable and its interaction with the *Relaxation* indicator. Table 8 presents the results. We find that the coefficients on *Relaxation* × *Labor intensity* are negative and significant in all three columns, indicating that the treatment effect is stronger when the firms rely more on labor.

Furthermore, if the decreased innovation is due to the inflow of low-skilled migrant workers triggered by *hukou* relaxation, we expect this treatment effect to be stronger in cities that have a stronger enforcement of such *hukou* relaxation. We measure the extent of enforcement using the percentage of residents who have newly obtained local urban *hukou* in response to the relaxation of the *hukou* system, as stronger enforcement is expected to lead to more

[2] As a robustness check, in untabulated results, we also require the distance between the treated firm and control firm to be within 50, 80, or 120 miles, and our inferences are unchanged.

TABLE 7 Treated firms and neighboring control firms

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Relaxation	−0.202**	−0.125**	−0.163**
	(.012)	(.023)	(.014)
Cash	−0.094	−0.035	−0.042
	(.427)	(.623)	(.771)
Firm size	0.124***	0.063***	0.127***
	(.000)	(.000)	(.000)
Leverage	−0.042	−0.008	−0.029
	(.369)	(.621)	(.523)
R&D	22.126**	16.142**	17.103*
	(.041)	(.022)	(.072)
Capex	−0.102	−0.049	−0.051
	(.578)	(.591)	(.672)
ROA	−0.149	−0.062	−0.151*
	(.118)	(.273)	(.092)
Firm age	0.051***	0.032***	0.043***
	(.000)	(.000)	(.000)
Tobin's *Q*	0.007	−0.004	0.012
	(.408)	(.423)	(.125)
Tangible	0.153	0.102	0.126
	(.119)	(.212)	(.133)
Constant	−2.563***	−1.453***	−2.163***
	(.000)	(.000)	(.000)
Observations	18,661	18,661	18,661
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R^2	.612	.539	.572

Note: This table examines whether the effect of *hukou* relaxation on corporate innovation is confounded by unobserved changes in local business conditions. For each treated firm, we match to a control firm that is in the same industry, in a city that has not relaxed its *hukou* restrictions, and closest in distance. To ensure that the treated firm and its "closest" control firm are truly close to each other, we further require that the distance between the treated firm and its "closest" control firm be within 100 miles. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

people obtaining local urban *hukou*. We define the variable *Percentage of people newly obtaining urban hukou* as the number of residents who were previously holding rural *hukou* and have newly obtained their local urban *hukou* normalized by total residents with a local urban *hukou* in a city. Then we reestimate Equation (2) by adding *Percentage of people newly obtaining urban hukou* and its interaction with *Relaxation*, *Relaxation* × *Percentage of people newly obtaining urban hukou*. Table 9 presents the results.

TABLE 8 heterogeneous treatment effects based on labor intensity

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Relaxation × labor intensity	−0.367**	−0.212**	−0.503**
	(.027)	(.043)	(.012)
Relaxation	−0.052**	−0.037**	−0.041
	(.012)	(.046)	(.172)
Labor intensity	0.152	0.179*	0.145
	(.152)	(.069)	(.138)
Cash	−0.102*	−0.052	−0.027
	(.091)	(.152)	(.368)
Firm size	0.156***	0.102***	0.134***
	(.000)	(.000)	(.000)
Leverage	−0.042	−0.027	−0.034
	(.174)	(.279)	(.328)
R&D	20.126***	17.247***	13.503***
	(.000)	(.000)	(.000)
Capex	−0.038	−0.027	0.017
	(.645)	(.503)	(.812)
ROA	−0.062	−0.051	−0.042
	(.189)	(.187)	(.192)
Firm age	0.042***	0.021***	0.037***
	(.000)	(.000)	(.000)
Tobin's *Q*	0.014**	0.009*	0.017**
	(.042)	(.081)	(.025)
Tangible	0.132**	0.074**	0.082*
	(.034)	(.028)	(.074)
Constant	−2.217***	−1.143***	−2.089***
	(.000)	(.000)	(.000)
Observations	30,216	30,216	30,216
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R^2	.612	.542	.597

Note: This table reports the cross-sectional variation of the treatment effects based on the firm's labor intensity. *Labor intensity* is the expenditure on employee wages normalized by total revenue. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

The coefficients on *Relaxation* × *Percentage of people newly obtaining urban hukou* are negative and significant across all three columns. This result indicates that the negative effect of *hukou* relaxation on corporate innovation is more pronounced for firms in cities that have a stronger enforcement of *hukou* relaxation.

Taken together, the effects of *hukou* relaxation on corporate innovation are stronger for firms that rely more on labor and for firms in cities that have a stronger enforcement of *hukou* relaxation. These results suggest that the impact

TABLE 9 Heterogeneous treatment effects based on the enforcement of *hukou* relaxation

	(1) ln(1+all patent)	(2) ln(1+invention patent)	(3) ln(1+utility model and design patent)
Relaxation × percentage of people newly obtaining urban *hukou*	−0.623**	−0.279*	−0.424**
	(.043)	(.071)	(.039)
Relaxation	−0.028	−0.023	−0.008
	(.412)	(.308)	(.712)
Percentage of people newly obtaining urban *hukou*	0.127	−0.102	0.198
	(.313)	(.208)	(.267)
Cash	−0.102	−0.051	−0.062
	(.142)	(.224)	(.312)
Firm size	0.123***	0.098***	0.121***
	(.000)	(.000)	(.000)
Leverage	−0.036	−0.011	−0.022
	(.287)	(.425)	(.314)
R&D	13.172*	11.025*	10.116
	(.072)	(.042)	(.136)
Capex	−0.041	−0.029	0.016
	(.258)	(.213)	(.529)
ROA	−0.123	−0.041	−0.114
	(.162)	(.128)	(.136)
Firm age	0.019	−0.004	0.017
	(.315)	(.423)	(.251)
Tobin's *Q*	0.004	0.002	0.005
	(.491)	(.389)	(.423)
Tangible	0.139**	0.122**	0.132*
	(.041)	(.048)	(.062)
Constant	−2.152***	−1.927***	−2.138***
	(.005)	(.004)	(.003)
Observations	30,216	30,216	30,216
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R^2	.592	.513	.538

Note: This table reports the cross-sectional variation of the treatment effects based on the enforcement of a city's *hukou* relaxation, measured by *Percentage of people newly obtaining urban* hukou. This variable is computed as the number of people who have newly obtained their local urban *hukou* normalized by the city's total number of residents with a local urban *hukou*. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

of *hukou* relaxation on corporate innovation is indeed tied to migrant workers and is unlikely to be spuriously driven by unobserved heterogeneity.

5 | ADDITIONAL ANALYSIS

5.1 | City-level aggregate innovation

Considering that each city has a different number of firms in our sample, a potential concern is that our main results could be driven by a few cities with a large number of firms. To mitigate this concern, we conduct a robustness check by investigating the city aggregate level of patents. Based on 5,789 city-year observations, we implement our difference-in-differences estimation using the following regression:

$$City\ aggregate\ innovation_{i,t} = \alpha + \beta_1 Relaxation_{s,t-1} + \beta_2 City\ characteristics_{i,t-1} + City\ FE + \ Year\ FE + \varepsilon_{i,t}. \quad (3)$$

Table 10 reports the results. The coefficient on the *Relaxation* indicator is negative and significant at the 5% level across all three columns. Taking column (1) for example, the dependent variable is the city-level aggregate patent number, including both invention patents and utility model and design patents, which is defined as the total number of patent in all firms in a city normalized by the total number of firms in the city. The coefficient on the *Relaxation* indicator is –0.126 and is significant at the 5% level, indicating a decrease in the aggregate number of patents by approximately 13% ($=e^{0.126} - 1$). As shown in columns (2) and (3), the number of city-level invention patents and utility model and design patents decreases by 11% ($=e^{0.108} - 1$) and 12% ($=e^{0.113} - 1$), respectively.

It is also worth noting that new firms may be started by the influx of cheap labor following *hukou* relaxation. If these new firms are sufficiently more productive than existing firms, our inference may be different (considering that these new firms will drop out of our baseline estimations because of firm fixed effects). The city-level analysis in Table 10 indicates that our results will still hold even if new firms starting up after *hukou* relaxation are more productive.

In summary, we find a significant decrease in the city-level aggregate number of patents following *hukou* relaxation. This result is consistent with our baseline results using firm-level data.

5.2 | Validating tests on the timing of *hukou* relaxation

Our empirical tests are based on the assumption that the cross-city timing of a *hukou* relaxation is unrelated to the innovation of firms in these event cities. To investigate the validity of this assumption, we employ a hazard model that is similar to the one used by Beck, Levine, and Levkov (2010) to study the U.S. state-level banking deregulation.

In particular, we run a city-level regression where the dependent variable is *ln(expected time to* hukou *relaxation implementation)*, based on the 30 event cities. Cities are dropped from the sample once they implement the policy change. The independent variables are the averages and changes of innovation outputs of all firms in the event cities. We also control for various city-level variables used in Table 6.

The estimated results of the hazard model are reported in Table 11. None of the coefficients on the level or the change of innovation is significant, and the magnitudes of these coefficients are also close to zero. These results indicate that the timing of the *hukou* relaxation is not related to the level or the change of the preexisting innovation activities, supporting the exogeneity of such city-level policy changes.

It is also worth noting that the coefficients on *ln(city GDP)* are significantly negative across all columns, indicating that cities with strong economic growth are more likely to relax their *hukou* policies. This result is consistent with the view that good economic conditions mitigate the financial burden of providing social welfare to new migrants, help to secure employment opportunities for incumbent urban residents, and thus increase the likelihood of local governments relaxing their *hukou* policies (Cai, 2011). Considering that innovation outputs may be positively related to local

TABLE 10 City-level aggregate innovation

	1 ln(City-level average number of all patent)	2 ln(City-level average number of invention patent)	3 ln(city-level average number of utility model and design patent)
Relaxation	−0.126**	−0.108**	−0.113**
	(.029)	(.031)	(.042)
ln(city GDP)	0.151***	0.098***	0.136***
	(.000)	(.000)	(.000)
ln(city population)	0.061	0.032	0.049
	(.163)	(.317)	(.224)
ln(Number of universities in the city)	0.051	0.042	0.038
	(.120)	(.114)	(.213)
City income per capita	0.054**	0.042**	0.039***
	(.017)	(.021)	(.013)
City expenditure on science and technology	−0.428	−0.519	−0.448
	(.521)	(.313)	(.428)
City housing price index	−0.024	−0.019	−0.023
	(.279)	(.314)	(.285)
Constant	−2.127***	−1.843***	−2.029***
	(.000)	(.000)	(.000)
Observations	5,789	5,789	5,789
Year FEs	Yes	Yes	Yes
City FEs	Yes	Yes	Yes
R^2	.515	.494	.503

Note: This table reports the difference-in-differences tests that examine the impacts of *hukou* relaxation on corporate innovation, using city-level aggregate innovation. For cities that relaxed their *hukou* policies, the indicator variable *Relaxation* takes the value of 1 for the period after the policy change and 0 for the period prior to the policy change. For cities that never relaxed their *hukou* policies in our sample period, *Relaxation* always takes the value of 0. Variable definitions are provided in the Appendix. City-level average number of patents is computed as the total number of patents of all firms in the city normalized by the number of firms in the city. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-Values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

economic conditions, this positive relation between the relaxation of *hukou* policies and economic conditions is likely to work against our finding a negative effect of *hukou* relaxation on innovation.

In summary, we show that the relaxation of city-level *hukou* policies is likely to be exogenous to local firms' innovation activities prior to the policy changes.

5.3 | Placebo tests

Finally, to ensure that our main results are not purely driven by chance, we conduct a placebo test to check whether our results disappear when we randomly pick a relaxation year other than the actual one. Specifically, for each city that relaxed *hukou*, we assign a pseudo-relaxation year randomly chosen from our sample period 1997–2016. We further require the pseudo-event year to be either at least 5 years before or 5 years after the actual event year, so that

TABLE 11 Timing of *hukou* relaxation and preexisting corporate innovation

	1	2	3	4	5	6
	ln(Expected time to *hukou* relaxation implementation)					
ln(City-level average number of all patent)	−0.112					
	(.131)					
ln(City-level average number of invention patent)		−0.179				
		(.145)				
ln(City-level average number of utility model and design patent)			−0.081			
			(.157)			
Change in ln(City-level average number of all patent)				0.034		
				(.422)		
Change in ln(City-level average number of invention patent)					0.023	
					(.612)	
Change in ln(City-level average number of invention patent)						0.031
						(.312)
ln(City GDP)	−0.916**	−0.922**	−0.954**	−0.713**	−0.726**	−0.739**
	(.014)	(.010)	(.011)	(.013)	(.014)	(.012)
ln(City population)	−0.071	−0.074	−0.077	0.243	0.235	0.261
	(.248)	(.223)	(.252)	(.422)	(.437)	(.412)
ln(Number of universities in the city)	−0.161**	−0.145**	−0.171**	−0.123	−0.121	−0.133
	(.023)	(.041)	(.027)	(.136)	(.125)	(.144)
City income per capita	−0.029	−0.034	−0.031	0.211**	0.216**	0.223**
	(.305)	(.312)	(.227)	(.026)	(.032)	(.024)
City expenditure on science and technology	−0.517	−0.426	−0.528	0.434	0.402	0.506
	(.709)	(.717)	(.706)	(.712)	(.703)	(.616)
City housing price index	−0.014	−0.014	−0.014	−0.0136	−0.016	−0.015
	(.134)	(.126)	(.157)	(.1975)	(.193)	(.182)
Constant	15.781**	15.613**	14.226**	8.023**	8.127**	8.005**
	(.013)	(.011)	(.015)	(.029)	(.025)	(.026)
Observations	377	377	377	323	323	323
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes	Yes
R^2	.812	.812	.834	.828	.828	.829

Note: The dependent variable is *Ln(expected time to* hukou *relaxation implementation)*. The sample consists of 30 cities that relaxed their *hukou* policies in our sample. Cities are dropped from the sample once they have relaxed their *hukou* policies. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1st and 99th percentiles. *p*-values based on standard errors clustered by city are in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

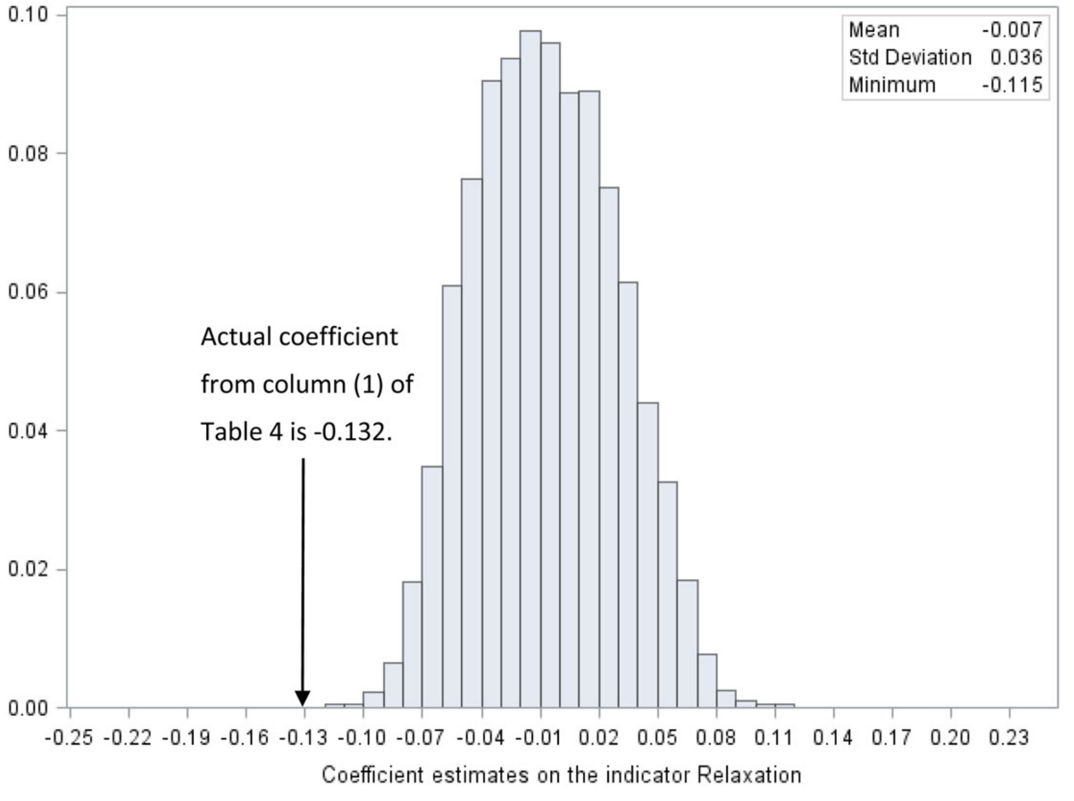

FIGURE 2 Placebo tests. This figure plots the histogram of coefficient estimates on the indicator "Relaxation" from 5,000 bootstrap simulations of the baseline model in column 1 of Table 4. For each city that relaxed its *hukou* system, we assign a pseudo-relaxation year randomly chosen from the sample period 1997–2016, and at least either 5 years before or 5 years after the actual event year. We then reestimate the baseline regression based on those pseudo-relaxation years and save the coefficient estimates on the indicator *Relaxation*. We repeat this procedure 5,000 times [Color figure can be viewed at wileyonlinelibrary.com]

the pseudo-event year is not confounded with the actual one. We then reestimate the baseline regression in Equation (2) based on those pseudo-event years and save the coefficient on *Relaxation*. We repeat this procedure 5,000 times.

Figure 2 plots the empirical distribution of the coefficient estimates based on those pseudo-events. The figure clearly shows that the coefficient estimate from column (1) of Table 4 lies well to the left of the entire distribution of coefficient estimates from the placebo test. The coefficient estimate from Table 4, column (1) (–0.132) is approximately three standard deviations (0.036) below the mean (–0.007) of the distribution and is smaller than the minimum coefficient estimate (–0.115) from the placebo test. These results suggest that our findings are indeed driven by relaxation of the *hukou* system and are unlikely to be driven by chance.

6 | CONCLUSIONS

In this paper, we find that rural–urban migration has a negative causal effect on corporate innovation in firms in migrant host cities. We exploit the exogenous shocks from the staggered relaxation of China's city-level household registration system (the *hukou* system), which relaxes the restrictions for rural residents to migrate to urban areas.

Using a difference-in-differences approach, we find a significant decrease in firms' patents following the relaxation of *hukou*-related restrictions, relative to firms in cities that do not implement such policy changes. We then conduct a number of tests in support of a causal interpretation of our findings. Our tests of parallel trends show that there is no time trend difference in innovation output between treated firms and control firms and that the decrease in innovation output occurs several years after the policy changes. Our tests employing the treated firms and their neighboring control firms show that our results are unlikely to be driven by unobservable confounding local economic factors that would have affected both the treated and the control firms equally. Further, we present cross-sectional variations in the treatment effects suggesting that those treatment effects are indeed related to migrant workers: our result is more pronounced for firms that rely more on labor and for firms in cities with a stronger enforcement of such *hukou* relaxation. Overall, our findings support the view that an abundant supply of low-skilled workers hinders corporate innovation because it increases the benefit of using existing low-skilled technology and thus reduces firms' incentive to innovate.

Studying the impacts of low-skilled migrants on host areas is particularly important in the current political and economic environments around the world. In the United States, President Donald Trump is building a wall along the United States–Mexico border to prevent the illegal immigration of Mexicans (who are largely low skilled) and vows to end former president Barack Obama's plan to shield millions of undocumented immigrants from deportation. Concerning the current refugee crisis in Europe, Chancellor Angela Merkel of Germany has adopted an open-door immigration policy. This policy has aroused great controversy, but its real effect is not yet clear. Our results suggest that policies aimed to facilitate rural–urban migration could have an unintended effect of reducing corporate innovation.

ACKNOWLEDGMENTS

We are grateful for the helpful comments from Utpal Bhattacharya (Editor), an anonymous referee, Po-Hsuan Hsu, Ping Jiang, Bin Ke, Xuan Tian, Chen Lin, Rong Wang, Hailong Zhao, and the seminar participants from Nanyang Technological University, Chinese University of Hong Kong (Shenzhen), Sun Yat-sen University, Shanghai University of Finance and Economics, Zhejiang University, Nankai University, University of International Business and Economics, Fudan University, the 2016 China Financial Research Conference, and the 2016 China Finance Review International Conference. Deqiu Chen acknowledges financial support from the National Natural Science Foundation of China (grant numbers 71790604, 71872045, and 71572035). Huasheng Gao acknowledges financial support from the Shanghai Pujiang Program (reference number 18PJC007) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning. Yujing Ma acknowledges financial support from the University of International Business and Economics (reference number 841612). All errors are our own.

REFERENCES

Afridi, F., Li, S. X., & Ren, Y. (2015). Social identity and inequality: The impact of China's hukou system. *Journal of Public Economics*, *123*, 17–29.

Atanassov, J. (2013). Do hostile takeovers stifle innovation? Evidence from antitakeover legislation and corporate patenting. *Journal of Finance*, *68*, 1097–1131.

Beck, T., Levine, R., & Levkov, A. (2010). Big bad banks? The winners and losers from bank deregulation in the United States. *Journal of Finance*, *65*, 1637–1667.

Bell, B., Fasani, F., & Machin, S. (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics*, *95*, 1278–1290.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, *119*, 249–275.

Bianchi, M., Buonanno, P., & Pinotti, P. (2012). Do immigrants cause crime? *Journal of the European Economic Association*, *10*, 1318–1347.

Borjas, G. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, *118*, 1335–1374.

Cai, F. (2011). Hukou system reform and unification of rural-urban social welfare. *China & World Economy*, *19*, 33–48.

Card, D. (2005). Is the new immigration really so bad? *Economic Journal*, *115*, F300–F323.
Chen D., Li, S., Xiao, J. Z., & Zou, H. (2014). The effect of government quality on corporate cash holding. *Journal of Corporate Finance*, *27*, 384–400.
Claessens, S., & Laeven, L. (2003). Financial development, property rights, and growth. *Journal of Finance*, *58*, 2401–2436.
Elvin, M. (1972). The high-level equilibrium trap: The causes of the decline of the invention in traditional Chinese textile industries. In W. Willmott (Ed.), *Economic organization in chinese society*. Stanford, CA: Stanford University Press.
Gao, H., & Zhang, W. (2017). Employment non-discrimination acts and corporate innovation. *Management Science*, *63*, 2982–2999.
Griliches, Z. (1981). Market value, R&D, and patents. *Economics Letters*, *7*, 183–187.
Habakkuk, J. (1962). *American and British technology in the nineteenth century*. London, U.K.: Cambridge University Press.
Heider, F., & Ljungqvist, A. (2015). As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes. *Journal of Financial Economics*, *118*, 684–712.
Imbens, G., & Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 5–86.
Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. *Quarterly Journal of Economics*, *126*, 1029–1069.
Liu, Z. (2005). Institution and inequality: The hukou system in China. *Journal of Comparative Economics*, *33*, 133–157.
Moulton, B. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics*, *32*, 385–397.
Ottaviano, G., & Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, *10*, 152–197.
Peri, G. (2012). The effect of immigration on productivity: Evidence from US states. *Review of Economics and Statistics*, *94*, 348–358.
Pischke, J.-S., & Velling, J. (1997). Employment effects of immigration to Germany: An analysis based on local labor markets. *Review of Economics and Statistics*, *79*, 594–604.
Raddatz, C. (2006). Liquidity needs and vulnerability to financial underdevelopment. *Journal of Financial Economics*, *80*, 677–722.
Roberts, M., & Whited, T. (2013). Endogeneity in empirical corporate finance. In G. Constantinides, M. Harris, & R. Stulz (Eds.), *Handbook of the economics of finance* (Vol. 2). Amsterdam, the Netherlands: Elsevier.
Scherer, F. M. (1965). Firm size, market structure, opportunity, and the output of patented inventions. *American Economic Review*, *55*, 1097–1125.
Tan, Y., Tian, X., Zhang, X., & Zhao, H. (2015). *The real effect of privatization: Evidence from China's split share structure reform*. Working Paper. Dunedin, New Zealand: University of Otago.
Wu, X., & Treiman, D. (2004). The household registration system and social stratification in China, 1955–1996. *Demography*, *41*, 363–384.

How to cite this article: Chen D, Gao H, Luo J, Ma Y. The effects of rural–urban migration on corporate innovation: Evidence from a natural experiment in China. *Financial Management*. 2019;1–25. https://doi.org/10.1111/fima.12280

APPENDIX: VARIABLE DEFINITIONS

Variable	Definition
All patent	Total number of invention, utility model, and design patent applications filed and eventually granted in a given year
Capex	Capital expenditures normalized by the book value of total assets
Cash	Cash and marketable securities normalized by the book value of total assets
City expenditure on science and technology	The expenditure on science and technology normalized by fiscal expenditure in a city
City GDP	City-level GDP
City income per capita	Per capita income of city residents
City housing price index	Price index of residential properties in a city
City-level average number of all patents	The total number of patent applications filed and eventually granted by all firms in a city normalized by the total number of firms in the city
City-level average number of invention patents	The number of invention patent applications filed and eventually granted by all firms in a city normalized by the total number of firms in the city
City-level average number of utility model and design patents	The number of utility model and design patent applications filed and eventually granted by all firms in a city normalized by the total number of firms in the city
City population	Total population in a city
Expected time to *hukou* relaxation implementation	The number of years ahead for a city to implement *hukou* relaxation
Firm age	Number of years since the firm's foundation
Invention patent	Total number of invention patent applications filed and eventually granted in a given year
Firm size	Natural logarithm of total assets
Labor intensity	The expenditure on employee wages normalized by total revenue
Leverage	Total debt normalized by the book value of total asset
Percentage of employees with a bachelor's degree	Percentage of employees with a bachelor's degree among all employees
Percentage of people newly obtaining urban *hukou*	Number of people newly obtaining urban *hukou* normalized by the total number of people with an urban *hukou* in the city
Percentage of technicians	Percentage of technician employees among all employees
R&D	R&D expenditures normalized by the book value of total assets. If R&D expenditures variable is missing, we set the missing value to 0.
Relaxation	An indicator variable that takes the value of 1 for the period after a city relaxed its household registration policy and 0 otherwise. For cities that never relaxed their household registration policy, the indicator variable *Relaxation* always takes the value of 0
ROA	Return on assets, measured as operating income normalized by the book value of total assets
Tangible	Property, plant, and equipment normalized by the book value of total assets
Tobin's *Q*	Market value of equity plus book value of assets minus book value of equity minus balance sheet deferred taxes, normalize by the book value of total assets
Utility model and design patent	Total number of utility model and design patent applications filed and eventually granted in a given year
Number of employees	Total number of employees in a firm
Number of people newly obtaining urban *hukou*	The number of people who were previously holding rural *hukou* and have newly obtained their local urban *hukous*
Number of universities in the city	The number of universities in a city