The Effects of Low-skilled Migration on Corporate Innovation: Evidence from a Natural Experiment in China[*]

Deqiu Chen
Business School
University of International Business and Economics
No.10 Huixin East Street, Chaoyang District, Beijing, China
nkchendq2006@163.com

Huasheng Gao
Nanyang Business School
Nanyang Technological University
S3-B1A-06, 50 Nanyang Avenue, Singapore 639798
hsgao@ntu.edu.sg

Jiang Luo
Nanyang Business School
Nanyang Technological University
S3-01C-75, 50 Nanyang Avenue, Singapore 639798
luojiang@ntu.edu.sg

This version: June, 2016

Abstract

We identify a negative causal effect of migration of low-skilled workers on corporate innovation in the host areas. Our tests exploit the staggered policy change of city-level household registration system in China, which facilitates rural residents to migrate to the city. 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. This result is more pronounced for firms that rely more on labor and for firms in cities with a larger number of pre-existing migrant workers. Overall, our results support the view that an abundant supply of low-skilled workers increases the benefit of using existing low-skill technology and thus reduces firms' incentive to pursue high-skill new technology.

Keywords: Innovation; Patents; Migration; Low-skilled Worker; Household registration; Hukou

JEL Classification: G38; J24; K31; M14; O31

---

[*] We are grateful for the helpful comments from Bin Ke and the seminar participants Nanyang Technological University. Gao acknowledges the financial support from Singapore Ministry of Education Academic Research Fund Tier 2 (Official Number: MOE2015-T2-1-118) and Academic Research Fund Tier 1 (Official Number: RG66/15). All errors are our own.

1. Introduction

Existing work on the effect of migration on innovation is generally limited to the role played by highly educated migrants, generally migrants with at least tertiary education, and generally finds that such migrants have a positive effect on the innovation in the host places (see, e.g., Chellaraj et al., 2008; Hunt and Gauthier-Loiselle, 2010; Kerr and Lincoln, 2010). However, few studies have investigated the effect of low skilled migration on innovation. This lack of evidence makes it difficult to fully understand the effect of migration on corporate innovation because low-skilled migrants accounts for the majority of migrants.

In this paper, we fill this gap and document a negative effect of low-skilled migration on firms' innovation in the host areas, using a quasi-natural experiment in China. Our empirical identification strategy is based on the staggered policy changes of China city-level household registration system that relaxes the restriction for farmers in the rural areas to migrate to the cities. We use these policy changes to capture an exogenous increase in the inflow of low-skilled migration 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 as urban citizens and abolish the rural-urban divide. As these policies changes were not made with the intention of hindering innovation, potential effects on innovation are likely to be an unintended consequence of these policies. Second, the staggered policies changes in several China cities 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 and Whited, 2013).

We expect the migration of low-skilled workers to decrease corporate innovation because companies are less likely to adopt new technology or innovate when there is an abundant supply of low-skill labor (Lewis, 2011; Peri, 2012). Suppose that a firm is currently using a pre-existing low-skill technology operated by low-skilled workers. The firm is considering to invest in some risky R&D projects to develop a high-quality technology operated by high-skilled workers. The likelihood of making such R&D investment depends on the cost of the R&D expenditure and as well as the relative profit of using new technology versus the existing one. An abundant supply of low-skilled workers in the labor market increases the benefit of using the existing low-skill technology, and thus enhances the hurdle for the firm to pursue the high-quality new technology, which in turn hinders corporate innovation. Anecdotal evidence does support this view. For example, Habakkuk (1962) claim that technological progress was slower in Britain than in the U.S. in the nineteenth century because of a large supply of low-skilled cheap workers in the former country. Elvin (1972) suggests that a sophisticated spinning wheel used for hemp in fourteenth-century China was later abandoned and was not used for cotton largely because the cheap and abundant Chinese labor made it unprofitable.

Using a panel of 18,481 Chinese public firms from 1999 to 2011 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 16%, 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 law changes. Our tests show that their pre-treatment trends are indeed indistinguishable. Moreover, most of the impact of household registration policy on innovation occurs two years after the policies' 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 GDP, population, education, and investment in R&D. 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 household registration policies allows us to difference away any unobserved confounding factors as long as they affect both the treated cities and its neighbors. By comparing treated firms to their immediate neighbors, we can better identify how much of the observed innovation change is due to household registration policies 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 policies, relative to their neighboring firms. These results indicate that the observed decrease in innovation following the changes in household registration policies is not driven by local economic shocks.

Finally, to provide further evidence that the effects of household registration relaxation on innovation are indeed tied to migrant workers, we apply a double 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 a larger number of pre-existing migrants. These cross-sectional variations in the treatment effects further increase our confidence that the impact of household registration policy change on innovation is indeed related to migrant workers.

This paper contributes to at least two strands of literature. First, our paper adds to the studies that examine the drivers of corporate innovation. This strand of literature is important for the economy, because innovation is widely believed to be crucial for sustainable growth and economic development (Solow, 1957; Romer, 1990, Porter, 1998). Current research on this topic has focused on factors such as incentive compensation for management (Manso, 2011), institutional ownership (Aghion et al., 2013), anti-takeover provisions (Atanassov, 2013), access to the equity market (Hsu et al., 2013), information environment (He and Tian, 2013), employees' job security (Acharya et al., 2014), etc. Although these studies enhance our understanding of the mechanisms that motivate firms to innovate, the role of the labor market is largely overlooked. This lack of evidence makes it difficult to fully understand the drivers of corporate innovation, given that human capital is emerging as the most crucial asset for an innovative firm (Zingales, 2000). Our paper helps to fill this gap by documenting the labor migration (especially the migration of low-skilled workers) as an important determinant of innovation.

Second, our study sheds light on the real consequences of labor migration, which has recently been at the center of the political and economic agenda. Economists have studied extensively the impact of migration on several economic and social indicators of host areas, such as natives' wages (Borjas, 2003; Ottaviano and Peri, 2012), employment opportunities (Pischke and Velling, 1997 Card, 2005), firm productivity (Peri, 2012), crime rate (Bianchi et al., 2012, Bell et al., 2013), etc. Complementing this strand of literature, we provide evidence the migration of low-skilled workers has a negative causal effect on innovation in the host areas.

The remainder of the paper is organized as follows. Section 2 reviews the background on China's household registration policy. We develop a simple model to illustrate our economic intuition in Section 3. Section 4 describes our sample and key variable construction. Section 5 presents our main empirical results. We implement additional robustness check in Section 6 and conclude in Section 7.

2. Background on China's Household Registration Policy

The household registration system (also known as hukou system) was established in China in 1958, following the Soviet style internal passport. This system was designed to control population migration and labor mobility between rural and urban areas. It strictly divides the population into four types depending on occupational aspects (agricultural versus non-agricultural) and locational aspects (living in urban areas versus rural areas). Such classification is largely based on the place of birth and the household registration status of the parents and the majority of population living in rural areas has an agricultural hukou. This system was introduced to serve as an invisible wall to prevent the rural labor force form moving out of agriculture, and it is closely tied to an exclusive employment system in urban sectors. This hukou system also guarantees basic living and social welfare for urban residents (such as housing, education, medical care, etc.), making the hukou residents' entitlements much better than migrants' entitlements.

The purpose of introducing this system is 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 non-agricultural employment, and subsidized urban housing social benefits.

The hukou system was applied stringently: public security bureaus controlled place-to-place migration and it was almost impossible to move from a rural to an urban area without authorized plans or official agreement. Since the "reform and opening-up" policy was instituted in the 1980s, controls over urban-to-rural labor mobility started to be relaxed. In the middle of 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 the cities with more than 60 million migrants coming to the cities with in the first 10 years of its application. Starting from the late 1990s, China experimented with a variety of reforms to further relax the restriction of hukou system. In 1997, the State Council started a nationwide relaxation of hukou system by permitting transferring the hukou status for family cases (i.e., spouses and children).

The reforms further accelerated in the early 2000s in Chinese municipalities as a result of the devolution of the responsibility for hukou controls to the local government. Several Chinese cities have adopted some hukou reform such as abolishing the distinction between rural and urban hukou or to lower the hurdle for migrant workers to obtain local urban hukou.

We mainly collect the city-level hukou reform information from China City Statistical Yearbook. We record the events of such hukou reform as the first year when the city abolishes the distinction between rual and urban hukou or when the city lowers the criteria for migrant workers to obtain local urban hukou. During our sample period, 30 cities have made such a reform. It is worth pointing out that local economic condition is an important driver of hukou reform. Local governments are more likely to relax hukou restrictions when local economic condition is good (there is a greater demand for labor supply). In contrast, when local economic condition is bad, local government is less likely to relax the restriction to avoid the financial burden of providing social welfare to the new migrants and to secure the employment opportunities for incumbent urban residents (Cai, 2011). This fact actually works against us finding a negative effect of household registration relaxation on firm innovation, considering that good economy condition is likely to be positively associated with innovation outputs.

3. A Simple Model

We develop a simple model to illustrate the relation between the supply of low quality labors and a firm's initiative to innovate. The firm can use two mutually exclusive technologies to produce. Each technology can be interpreted as a specific way to configure the firm's capital stock such as a machine.

First, the firm can use a large quantity, Q_L, of low quality labors, which costs P_L per unit. Then, the firm's total investment is $P_L Q_L$. Its profit is $\Pi_L = \alpha_L (P_L Q_L)^{\gamma}$. α_L is a positive constant, which describes the productivity of this low-skill-labor technology. $\gamma \in (0,1)$ is a positive constant, which is related to the curvature of the production function.

Second, the firm can use a small quantity, $Q_H < Q_L$, of high quality labors, which costs $P_H > P_L$ per unit. Then, the firm's total investment is $P_H Q_H$. Its profit is $\Pi_H = \alpha_H (P_H Q_H)^{\gamma}$. α_H is a positive constant, which describes the productivity of this high-quality-labor technology. We let $\vartriangle \alpha = \alpha_H - \alpha_L > 0$, so high quality labors are more productive than low quality labor. We also let $P_H Q_H = P_L Q_L = I$, so the required total investments of the two technologies are the same.

The low-skill-labor technology is immediately available. The high-quality-labor technology is not. To develop this technology, the firm needs to engage in an R&D, which causes a constant cost, C. The R&D succeeds with probability ρ, and fails with probability $1 - \rho$.

Consider two regimes. In regime 1, there is no abundant supply of low quality labor. Using the low-skill-labor technology is not an option. If the firm engages in R&D, then its NPV equals $\rho\Pi_H - C$. If it doesn't, then its NPV equals 0. The firm engages in R&D if $C < \rho\Pi_H = \rho\alpha_H I^\gamma$.

In regime 2, there is an abundant supply of low quality labor. Using the low-skill-labor technology is an option. If the firm engages in R&D, its NPV equals $\rho\Pi_H + (1-\rho)\Pi_L - C$. If it doesn't, then its NPV equals Π_L. The firm engages in R&D if $C < \rho(\Pi_H - \Pi_L) = \rho \ \triangle \alpha \ I^\gamma$.

The following proposition summarizes the above analysis.

Proposition 1: *(i) If $C < \rho \triangle \alpha I^\gamma$, then the firm engages in R&D in both regimes.*

(ii) If $C \in [\rho \ \triangle \alpha I^\gamma, \rho\alpha_H I^\gamma)$, then the firm engages in R&D in regime 1 (there is no abundant supply of low quality labors), but not in regime 2 (there is an abundant supply of low quality labors). (iii) If $C > \rho \ \alpha_H \ I^\gamma$, then the firm doesn't engage in R&D in either regimes.

An interesting observation is that an abundant supply of low quality labors can have a negative effect on the firm R&D initiative. Particularly, in the parameter range $C \in [\rho \triangle \alpha I^\gamma, \rho\alpha_H I^\gamma)$, if there is no abundant supply of low quality labor (regime 1), the firm will engage in R&D. If there is an abundant supply of low quality labor (regime 2), the firm won't engage in R&D. The intuition is that now the firm has an option of using the low-skill-labor technology to produce. This raises the hurdle rate for the firm to engage in R&D to develop the high-quality-labor technology.

Our main hypothesis follows immediately from Proposition 1.

Hypothesis: *A positive shock to the supply of low quality labors dampens a firm's initiative to innovate.*

4. Sample Formation and Variable Construction

We start with all Chinese A-share listed companies during 1999-2011 obtained from the *China Stock Market & Accounting R*esearch (CSMAR) database, from which we collect the firms' financial information. We start in 1999 because the Chinese patent information is only available from that year.

We use patent number 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 patent application date, application ID, publication ID, granting date, and patent ID 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. Chinese patent system classifies patents into three types: invention patents, utility model patents, and design patents. Invention patents refer to the ones granted for a new technical solution to a product or an industrial process. The utility model patents are the ones for new and practical technical solutions relating to some characteristics of the product such as the product's shape and structure. This type of patents demonstrates new functional aspects of a product. The design patents are for the product's new shape, pattern or color that makes the product more pleasing and industrially applicable. It is worth noting that SIPO database does not provide reliable information on patent citation, and thus we are unable to use patent citation to capture the quality of each patent. As pointed out by Tan et al. (2015), invention patents are the most original ones 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, ROA, and Tobin's *Q*. All explanatory variables are lagged by one 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 18,481 firm-year observations from 1999-2011.

Table 2 provides summary statistics. On average, firms in our sample have seven patents filed (and subsequently granted) per year. Out of these patents, three are invention patents and four are utility and design patents. Our average sample firms have book value assets of $3.02 billion and are 10 years old. They hold a sizeable amount of cash with a cash ratio of 18.6% of total assets. The average R&D and capital expenditure account for 0.1% and 6.01% of total assets, respectively. The average firms are moderately levered with a book leverage ratio of 49%, and tangible assets account for 27.5% of total assets in the average firms. In terms of performance, sample firms perform well with an average ROA of 2.8% and Tobin's Q of 2.03.

5. Empirical Results

5.1 Baseline Regression

Several Chinese cities tighten their hukou in different years during the sample period. Thus, we can examine the before-after effect of the change in hukou policies in affected cities (the treatment group) compared to the before-after effect in cities in which such a change 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 Atanassov (2013), Bertrand et al. (2004), and Imbens and Wooldridge (2009). We implement this test through the following regression:

$$Innovation_{i,t} = \alpha + \beta_1\, Relaxation_{s,t-1} + \beta_2 Other\ Firm\ Characteristics_{i,t-1} + Firm\ FE + Year\ FE + \varepsilon_{i,t}\ , \qquad (1)$$

where *i* indexes firm, *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. For the treated group, the indicator variable *Relaxation* equals one for the period after the relaxation of the city-level household registration system, and zero otherwise. For the control group, the indicator variable *Relaxation* always takes the value of zero. We include a set of control variables that may affect a firm's innovation output, as discussed in Section 4. The year fixed effects enable us to control for intertemporal technological shocks. The firm fixed effects allow us to control for time-invariant differences in patenting and citation practices across firms. Given that our treatment is defined at the city level, we cluster standard errors by city.

The coefficient of interest in this model is the $\beta_1$coefficient. As explained by Imbens and Wooldridge (2009), the employed fixed effects lead to $\beta_1$being estimated as the *within-firm* 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 innovation. We can subtract the number of innovations before the policy change from the number of innovations after the policy change for firms headquartered in Beijing. However, economy-wide shocks may occur at the same time and affect corporate innovations in 2002. To difference away such factors, we calculate the same difference in innovations for firms 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 firms in Beijing compared to firms in the control state.

Table 3 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 estimate on the *Relaxation* indicator is -0.154 and significant at the 1% level, suggesting a negative effect of the policy change on corporate innovation. The economic magnitude is also sizeable: the relaxation of household registration system leads to a decrease in the number of patents by approximately 16% $(= e^{0.154} - 1)$.

Examining Ln(1+invention patent) as the dependent variable in column (2), we find that the coefficient on the *Relaxation* indicator is -0.093 and is significant at the 1% level, which implies that the hukou policy relaxation leads to a decrease in the number of invention patent by approximately 10% $(= e^{0.093} - 1)$. We examine Ln(1+ utility and design patent) in column (3). The coefficient on the *Pass* indicator is -0.118 and is significant at the 1% level, indicating a decrease in the number of utility and design patents by 12% $(= e^{0.118} - 1)$.

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

5.3 The Pre-treatment 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 4 investigates the pre-trend between the treated group and control group. In particular, we define seven dummies, *Year −2*, *Year −1*, *Year 0*, *Year +1*, *Year +2*, *Year +3*, and *Year +4 and afterwards* to indicate the year relative to the relaxation of hukou system. For example, year 0 indicates the year in which the hukou relaxation is enacted; 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 re-estimate 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 trend assumption of the difference-in-differences approach is not violated.

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 three years after the enactment: the coefficients on *Year +3* indicator become significantly negative in all the three columns. The coefficients on *Year +4 and afterwards* 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 household registration policy on corporate innovation. This is understandable given that innovation is usually a long-term process.

Overall, Table 4 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. Moreover, Table 4 also indicates that most of the impact of hukou policy on innovation occurs three years *after* they are passed, which suggests a causal effect.

5.4 Confounding Local Business Conditions

Location is one important common factor that likely induces an association between the hukou policy 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 the hukou policy.

Table 5 presents our first test. In addition to our usual set of explanatory variables used in Table 3, we also account for various time-varying, city-level variables in our regressions. Given that richer and larger states 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 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, the city's expenditure for science and technology, and individual input for research and development. These city-level data are collected from the China Statistical Yearbook.

We find that the relaxation of household registration system continues to have a negative and (statistically and economically) significant impact on corporate innovation. Compared to Table 3, the coefficient on the *Hukou Relaxation* dummy becomes a little bigger. Also, we find that city GDP is (weakly) positively associated with innovation output. Other 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 policy 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 hukou policy is driven by unobserved changes in local business conditions, and that it is these changes (not the hukou policy) that influence corporate innovation in reality. 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 and 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 without tightening the hukou policy, 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 untreated firm be within 100 miles.[1] 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 re-estimate Equation (2) by focusing on this sub-sample of firms across the city border.

[1] As a robustness check, 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 6 presents the results. Restricting our sample to the pairs of neighboring treated and control firms reduces the sample to 11,416 firm-year observations; yet, we still find negative and significant coefficients (at the 1% level) on the *Hukou Relaxation* indicator in all three columns. Overall, these results suggest that unobserved local confounds seem not to drive our results.

5.5 Heterogeneous Treatment Effects

To provide further evidence that the effects of hukou policy 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 and 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, if the decreased innovation after the hukou relaxation is due to low-skilled migrant workers, we expect this treatment effect to be stronger in cities that have a larger number of residents who do not have local hukou. We define the variable *Population without hukou* as the number of residents who do not have a local hukou normalized by total residents in a city. Then, we re-estimate Equation (1) by adding *Population without hukou* and its interaction with *Relaxation*, *Relaxation* × *Population without hukou*. Table 7 presents the results.

The coefficients on *Relaxation* × *Population without 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 larger population without a local urban hukou.

Furthermore, considering that the hukou policy affects productivity associated with human capital, 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 re-estimate Equation (1) by adding the *Labor insensitivity* variable and its interaction with the *Relaxation* indicator. We find that the coefficients on *Relaxation*× *Labor intensity* are negative and significant in all the three columns. Taking column (1) for example (where the dependent variable is Ln (1+ all patents)), the coefficient on *Relaxation*× *Labor intensity* is -0.487 and significant at the 5% level, indicating that the treatment effect is stronger when the firms rely more on labor.

Taken together, the effects of hukou relaxation on corporate innovation are much stronger for firms in cities that have a larger number of migrants (who do not have local urban hukou status), and for firms that rely more on labor. These results suggest that the impact of hukou relaxation on corporate innovation is indeed tied to migrant workers and is unlikely to be spuriously driven by unobserved heterogeneity.

6. Additional Analysis

6.1 City-level Aggregate Innovation

Considering that our treatment effect is at the city level, we conduct a robustness check by investigate the city aggregate level of patents. Based on 3,361 city-year observations, we implement our difference-in-differences estimation using the following regression:

$City\ Aggregate\ Innovation_{i,t} =$

$\alpha + \beta_1\ Relaxation_{s,t-1} + \beta_2 City\ Characteristics_{i,t-1} + City\ FE +\ Year\ FE + \varepsilon_{i,t}$.

(2)

Table 9 reports the results. The coefficient on the *Relaxation* indicator is negative and significant at the 5% level across all the three columns. Taking column (1) for example, the dependent variable is the city-level aggregate patents number, including both invention patent and utility model and design patent, which is defined as the total number of patents in all firms in a city normalized by the total number of firms in the city. The coefficient on the *Relaxation* indicator is -0.165 and is significant at the 5% level, indicating a decrease in the aggregate number of patents by approximately 18% ($= e^{0.165} - 1$). As shown in columns (2) and (3), the number of city-level invention patent and utility model and design patent deceases by 12% ($= e^{0.116} - 1$) and 14% ($= e^{0.131} - 1$), respectively.

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

6.2 City-level Migration

To provide further evidence that hukou relaxation indeed leads to a greater inflow of low-skilled migration, we conduct an additional analysis in Table 10. The difference-in-differences regression specification is similar to that used in equation (1), except that the regression is based on 1,685 city-year observations and we do not include firm-level control variables. In columns (1) and (2), the dependent variable is the number of people who newly switch from their rural hukou status to urban hukou status normalized by the prior-year number of people with an urban hukou in the city. We find that the coefficient on the *Hukou relaxation* indicator is positive and is significant at the 5% level, indicating that hukou relaxation leads to a significant increase in the number of low-skilled migration.

As a robustness check in columns (3) and (4), we examine the number of people who newly switch from their rural hukou status to urban hukou status normalized by the prior-year total number of people (with or without a local urban hukou) in the city. We continue to find a positive and significant coefficient on the *Relaxation* indicator.

Overall, Table 10 provides evidence that hukou relaxation indeed leads to a greater inflow of low-skilled migration workers into the host cities.

6.3 Validating Tests on the Timing of Hukou Relaxation

Our empirical tests are based on the assumption that the cross-city timing of the hukou relaxation policy is unrelated with 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 et al. (2010) to study the state-level banking deregulation.

In particular, we run a city-level regression where the dependent variable, *Ln(T)*, is the expected time to the hukou relaxation based on the 30 event cities. T is the number of years ahead for a city to implement the hukou relaxation policy. Cities are dropped from the sample once they implemented the policy change. The independent variables are the average and changes of innovation outputs of all firms in the event cities. We also control for various city-level variables used in Table 5.

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 magnitude of these coefficients is also close to zero. These results indicate that the timing of the hukou relaxation is not related to the level or change of the pre-existing innovation. It is also worth noting that the coefficient on *Ln (city GDP)* is significantly negative, indicating that cities with good economic growth are more likely to relax its hukou system. This result is consistent with the view that good economic condition mitigates the financial burden of providing social welfare to new migrants, helps to secure employment opportunities for incumbent urban residents, and thus increases the likelihood of local governments relaxing hukou system (Cai, 2011). This positive relation between hukou relaxation and economic condition actually works against us finding a negative effect of hukou relaxation on innovation.

7. Conclusions

In this paper, we find that the migration of low-skilled workers has a negative causal effect on corporate innovation in the host areas. We exploit various exogenous shocks from staggered relaxation of China city-level household registration system (i.e., hukou system), which relaxes the restrictions for rural residents to migrant to the urban areas.

Using a difference-in-differences approach, we find a significant decrease in firms' patents following the policy changes, relative to firms in cities that do not have such a policy change. 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 firms in cities that later relax their hukou system and firms in other cities, and that the improvement 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 immigrant worker: our result is more pronounced for firms in cities with a larger number of pre-existing migrant workers who donot have a local hukou status, and for firms that rely more on labor. 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-skill technology and thus reduces firms' incentive to pursue new innovative technology.

Our paper provides important implications for public policies aimed at fostering innovation. Our results suggest that policies aimed to promote urbanization and equal treatment of residents in the rural and urban areas could have an unintended effect of reducing corporate innovativeness.

Lastly, although our paper focuses on China, the basic mechanism could be applied to other countries with a similar hukou system, including Japan (koseki), Vietnam (Hộ khẩu), Korea (Hoju), etc. Studies of international data on labor segmentation/migration caused by such household registration system could be an interesting area for future research.

Reference

Acharya, Viral, Ramin Baghai, and Krishnamurthy Subramanian, 2014. Wrongful discharge laws and innovation, *Review of Financial Studies* 27, 301–346.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005. Competition and innovation: An inverted-U relationship, *Quarterly Journal of Economics* 120, 701–728.

Aghion, Philippe, John Van Reenen, and Luigi Zingales, 2013. Innovation and institutional ownership, *American Economic Review* 103, 277–304.

Adnersson, Roland, John Quigley, and Mats Wilhelmsson, 2009. Urbanization, productivity, and innovation: Evidence from investment in higher education, *Journal of Urban Economics* 66, 2-15.

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

Beck, Thorsten, Ross Levine, and Alexey Levkov, 2010. Big bad banks? The winners and losers from bank deregulation in the United States, *Journal of Finance* 65, 1637–1667.

Bell, Brian, Francesco Fasani, and Stephen Machin, 2013. Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics* 95, 1278-1290.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119, 249–75.

Bianchi, Milo, Paolo Buonanno, and Paolo Pinotti, 2012. Do immigrants cause crime? *Journal of the European Economic Association* 10, 13118-1347.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, 2013. Identifying technology spillovers and product market rivalry, *Econometrica* 81, 1347–1393.

Borjas, George, 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, Fang, 2011. Hukou system reform and unification of rural-urban social welfare, *China & World Economy* 19, 33-48.

Card, David, 2005. Is the new immigration really so bad? *Economic Journal* 115, F300-F323.

Claessens, Stijn, and Luc Laeven, 2003. Financial development, property rights, and growth, *Journal of Finance* 58, 2401–2436.

Chellaraj, Gnanaraj, Keith Maskus, and Aaditya Mattoo, 2008. The contribution of international graduate students to US innovation, *Review of International Economics* 16, 444–462.

Cowan, Robin, and Natalia Zinovyeva, 2013. University effect on regional innovation, *Research Policy* 42, 788-800.

Elvin, Mark. 1972. The high-level equilibrium trap: the causes of the decline of the invention in traditional Chinese textile industries, In Economic Organization in Chinese Society, edited by William Willmott. Stanford, CA: Stanford University Press.

Faggian, Alessandra, and Philip McCann, 2009. Human capital, graduate migration and innovation in British regions, *Cambridge Journal of Economics* 33, 317-333.

Fang, Vivian, Xuan Tian, and Sheri Tice, 2014. Does stock liquidity enhance or impede firm innovation? *Journal of Finance* 69, 2085–2125.

Griliches, Zvi, 1981. Market value, R&D, and patents. *Economics Letters* 7, 183–167.

Habakkuk, John, 1962 American and British technology in the nineteenth century. London: Cambridge University Press.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, 2001. The NBER patent citation data file: Lessons, insights and methodological tools, University of California, Berkeley working paper.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, 2005. Market value and patent citations, *RAND Journal of Economics* 36, 16–38.

He, Jie, and Xuan Tian, 2013. The dark side of analyst coverage: The case of innovation, *Journal of Financial Economics* 109, 856–878.

Heider, Florian, and Alexander Ljungqvist, 2015. As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes, *Journal of Financial Economics* 118, 684-712.

Holmstrom, Bengt, 1989. Agency costs and innovation, *Journal of Economic Behavior and Organization* 12, 305–327.

Hsu, Po-Hsuan, Xuan Tian, and Yan Xu, 2014. Financial development and innovation: Cross-country evidence, *Journal of Financial Economics* 112, 116–135.

Hunt, Jennifer, and Marjolaine Gauthier-Loiselle, 2010. How much does immigration boost innovation? *American Economic Journal: Macroeconomics* 2, 31-56.

Imbens, Guido, and Jeffrey Wooldridge, 2009. Recent developments in the econometrics of program evaluation, *Journal of Economic Literature* 47, 5–86.

Kerr, William, and William Lincoln, 2010. The supply side of innovation: H-1b visa reforms and U.S. ethnic invention, *Journal of Labor Economics* 28, 473-58.

Lewis, Ethan, 2011. Immigration, skill mix, and capital skill complementarity, *Quarterly Journal of Economics* 126, 1029-1069.

Manso, Gustavo, 2011. Motivating innovation, *Journal of Finance* 66, 1823–1860.

Merrosty, Peter, 2013. Tolerance of ambiguity: A trait of the creative personality, *Creativity Research Journal* 25, 232–237.

Ottaviano,Gianmarco, and Giovanni Peri, 2012. Rethinking the effect of immigration on wages, *Journal of the European Economic Association* 10,152–197.

Peri, Giovanni, 2012. The effect of immigration on productivity: Evidence from U.S. states, *Review of Economics and Statistics* 94, 348-358.

Porter, Michael, 1998. Competitive Advantage of Nations, Free Press.

Pischke, Jorn-Steffen, and Johannes Velling, 1997. Employment effects of immigration to Germany: An analysis based on local labor markets, *Review of Economics and Statistics* 79, 594-604.

Raddatz, Claudio, 2006. Liquidity needs and vulnerability to financial underdevelopment, *Journal of Financial Economics* 80, 677–722.

Roberts, Michael, and Toni Whited, 2013. Endogeneity in empirical corporate finance, in George Constantinides, Milton Harris, and Rene Stulz, eds. Handbook of the Economics of Finance Volume 2, Elsevier.

Romer, Paul, Endogenous technological change, *Journal of Political Economy* 98, S71-S102.

Scherer, F. M., 1965. Firm size, market structure, opportunity, and the output of patented inventions, *American Economic Review* 55, 1097–1125.

Solow, Robert, 1957. Technical change and the aggregate production function, *Review of Economics and Statistics*, 312-320.

Tan, Yongxian, Xuan Tian, Xinde Zhang, and Hailong Zhao, 2015. The real effect of privatization: Evidence from China's split share structure reform, working paper.

Zingales, Luigi, 2000. In search of new foundations, *Journal of Finance* 55, 1623-1653.

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.
Invention patent	Total number of invention patent applications filed and eventually granted in a given year.
Utility model and design patent	Total number of utility model and design patent applications filed and eventually granted in a given year.
Relaxation	An indicator variable that takes the value of one if the city has relaxed its household registration policy, and zero otherwise.
Capex	Capital expenditures normalized by the book value of total assets.
Cash	Cash and marketable securities normalized by the book value of total assets.
Expenditure on science and technology	The expenditure on science and technology normalized by fiscal expenditure.
Firm age	Number of years since the firm's foundation.
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.
Ln (city GDP)	Natural logarithm of city GDP.
Ln(city population)	Natural logarithm of city population.
Ln(# of University)	Natural logarithm of the number of city university .
Per capita income	Per capita income of city residents.
Personnel of research and development	The expenditure on R&D and technology normalized by city R&D personnel.
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 zero.
ROA	Return on assets, measured as operating income normalized by the book value of total assets.
Tangible	Property, plant & 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.

Table 1: List of the Hukou Relaxation

This table reports the year when each city implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou from 2000 to 2011. Chinese names of the cities are reported in parentheses.

City	Original policy name in Chinese	Year of the policy
Tonghua (通化)	关于推进全市户籍制度改革的实施意见	2000
Urumqi (乌鲁木齐)	关于推进户籍管理制度改革的实施意见	2001
Beijing (北京)	关于推进小城镇户籍管理制度改革的意见	2002
Fuzhou (福州)	关于实施户籍管理制度改革的若干意见	2002
Jiaxing (嘉兴)	关于深化户籍管理制度改革的实施意见	2002
Jincheng (晋城)	关于推进户籍制度改革的实施意见	2002
Haining (海宁)	户籍管理城乡一体化改革实施细则	2003
Nanning (南宁)	户籍管理制度改革的实施细则	2003
Taizhou (泰州)	关于进一步放宽户口迁移准入条件的暂行规定	2003
Tianjin (天津)	居住证管理和证积分指标暂行办法	2003
Xiamen (厦门)	户籍管理若干规定	2003
Zhengzhou (郑州)	关于户籍管理制度改革的通知	2003
Changde (常德)	率先推行一元化户籍改革	2004
Nanjing (南京)	关于全面深化农村综合改革加快城乡一体化发展的意见	2004
Shanghai (上海)	居住证暂行规定	2004
Shenzhen (深圳)	关于推进户籍管理制度改革的实施意见	2004
Chengdu (成都)	关于深化户籍制度改革深入推进城乡一体化的意见	2005
Haerbin (哈尔滨)	建立城乡统一的一元户口登记制度	2006
Xian (西安)	关于进一步推进户籍制度改革的意见	2006
Yunchen (运城)	关于推进新型城镇化的实施方案	2006
Taiyuan (太原)	关于进一步改革户籍管理制度的意见	2007
Anshan (鞍山)	失地农民可落户城市意见	2008
Dalian (大连)	城镇户籍管理制度改革办法	2008
Kunming (昆明)	关于户籍管理制度改革的实施意见	2008
Shenyang (沈阳)	关于进一步改革户口迁移制度若干意见的通知	2008
Zhuhai (珠海)	关于推进户籍制度改革的实施意见	2008
Guangzhou (广州)	关于加快形成城乡经济社会发展一体化新格局的实施意见	2009
Qiqihaer (齐齐哈尔)	取消农业非农业户口性质划分	2009
Chongqing (重庆)	统筹城乡户籍制度改革社会保障实施办法	2010
Yinchuan (银川)	推进新型城镇化重点任务分工方案	2011

Table 2: Summary Statistics

The sample consists of 18,481 firm-year observations from 1999-2011. We obtain patent information from State Intellectual Property Office of China (SIPO) and financial information from 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.

Variable	Mean	SD	P1	Median	P99
All patent	7.148	82.78	0.000	0.000	102
Invention patent	3.220	70.02	0.000	0.000	32
Utility model and design patent	3.928	24.82	0.000	0.000	77
Total assets（Billion）	3.018	48.360	0.024	.230	21.58
Cash	0.186	0.152	0.002	0.143	0.724
Leverage	0.491	0.249	0.054	0.481	1.696
R&D	0.001	0.002	0.000	0.000	0.014
Capex	0.060	0.061	0.001	0.042	0.290
ROA	0.028	0.085	-0.454	0.0354	0.204
Firm age	10.20	4.824	1	10	23
Tobin’s Q	2.033	1.704	0.229	1.530	9.921
Tangible	0.275	0.183	0.003	0.245	0.772

Table 3: Effect of Hukou Relaxation on Innovation

This table reports the difference-in-differences tests that examine the impacts of hukou relaxation on corporate innovation. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Relaxation	-0.154***	-0.093***	-0.118***
	(0.001)	(0.001)	(0.004)
Cash	-0.158	-0.102	-0.075
	(0.143)	(0.171)	(0.471)
Firm size	0.166***	0.116***	0.150***
	(0.000)	(0.000)	(0.000)
Leverage	-0.068	-0.023	-0.042
	(0.221)	(0.548)	(0.377)
R&D	20.578**	18.748**	14.167
	(0.045)	(0.024)	(0.135)
Capex	-0.070	-0.063	-0.001
	(0.612)	(0.530)	(0.993)
ROA	-0.110	-0.095	-0.112
	(0.262)	(0.152)	(0.192)
Firm age	0.057***	0.029***	0.045***
	(0.000)	(0.000)	(0.000)
Tobin's Q	0.008	0.004	0.011
	(0.292)	(0.353)	(0.173)
Tangible	0.152*	0.110	0.111
	(0.083)	(0.113)	(0.130)
Constant	-3.640***	-2.495***	-3.295***
	(0.000)	(0.000)	(0.000)
Observations	18,481	18,481	18,481
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R2	0.691	0.662	0.669

Table 4: Testing for Pre-treatment Trends and Reversals

This table investigates the pre-treatment trends between the treated group and 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 the hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou. For example, the *Year 1* indicator takes the value of one if it is one year after a city adopts such a policy, and zero 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Year −2	0.052	0.041	0.044
	(0.307)	(0.191)	(0.312)
Year −1	-0.010	0.011	0.011
	(0.845)	(0.771)	(0.793)
Year 0 (event year)	-0.028	0.000	-0.010
	(0.624)	(0.999)	(0.850)
Year 1	-0.092	-0.049	-0.055
	(0.163)	(0.321)	(0.299)
Year 2	-0.105	-0.061	-0.068
	(0.138)	(0.187)	(0.289)
Year 3	-0.149**	-0.078*	-0.108*
	(0.030)	(0.090)	(0.096)
Year 4^{+}	-0.280***	-0.181***	-0.213***
	(0.000)	(0.001)	(0.002)
Cash	-0.150	-0.097	-0.068
	(0.168)	(0.204)	(0.509)
Firm size	0.163***	0.114***	0.147***
	(0.000)	(0.000)	(0.000)
Leverage	-0.074	-0.028	-0.047
	(0.175)	(0.456)	(0.328)
R&D	21.237**	19.227**	14.718
	(0.034)	(0.018)	(0.113)
Capex	-0.048	-0.047	0.016
	(0.732)	(0.640)	(0.900)
ROA	-0.118	-0.101	-0.119
	(0.220)	(0.122)	(0.160)
Firm age	0.061***	0.032***	0.048***
	(0.000)	(0.000)	(0.000)
Tobin's Q	0.007	0.004	0.010
	(0.355)	(0.446)	(0.203)
Tangible	0.145	0.105	0.106
	(0.100)	(0.137)	(0.147)
Constant	-3.586***	-2.457***	-3.253***
	(0.000)	(0.000)	(0.000)
Observations	18,481	18,481	18,481
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
R2	0.693	0.664	0.670

Table 5: Controlling for City-level Characteristics

This table reports the difference-in-differences tests that examine the impacts of hukou relaxation on corporate innovation, controlling for city-level characteristics. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. Variable definitions are provided in the Appendix. All continuous variables are winsorized at the 1^{st} and 99^{th} percentiles. P-values based on standard errors clustered by city are in parentheses. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Relaxation	-0.134***	-0.085***	-0.105**
	(0.006)	(0.007)	(0.016)
Cash	-0.171	-0.110	-0.084
	(0.114)	(0.141)	(0.422)
Firm size	0.165***	0.116***	0.148***
	(0.000)	(0.000)	(0.000)
Leverage	-0.065	-0.021	-0.039
	(0.234)	(0.580)	(0.407)
R&D	19.423*	17.288**	15.013
	(0.059)	(0.037)	(0.115)
Capex	-0.095	-0.072	-0.036
	(0.500)	(0.479)	(0.779)
ROA	-0.110	-0.093	-0.112
	(0.266)	(0.165)	(0.197)
Firm age	0.051***	0.025***	0.038***
	(0.000)	(0.000)	(0.000)
Tobin’s Q	0.008	0.005	0.010
	(0.334)	(0.322)	(0.221)
Tangible	0.156*	0.113	0.114
	(0.070)	(0.105)	(0.112)
Ln (city GDP)	0.032*	0.018	0.030*
	(0.086)	(0.186)	(0.091)
Ln(city population)	-0.014	-0.021	-0.004
	(0.766)	(0.361)	(0.932)
Ln(# of universities in the city)	0.060	0.051	0.069*
	(0.142)	(0.157)	(0.068)
City income per capita	0.006	0.005	0.004
	(0.217)	(0.242)	(0.346)
City expenditure on science and technology	-1.626	-0.557	-1.374
	(0.233)	(0.538)	(0.242)
Constant	-4.167***	-2.790***	-3.827***
	(0.000)	(0.000)	(0.000)
Observations	18,481	18,481	18,481
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R2	0.693	0.663	0.670

Table 6: Treated Firms and Neighboring Control Firms

This table examines whether the effect of hukou relaxation on 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 without adopting the hukou relaxation policy, and closest in distance. To ensure that 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 must be within 100 miles. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Relaxation	-0.232***	-0.131***	-0.195***
	(0.000)	(0.000)	(0.000)
Cash	-0.101	-0.038	-0.058
	(0.507)	(0.675)	(0.681)
Firm size	0.148***	0.090***	0.144***
	(0.000)	(0.000)	(0.000)
Leverage	-0.057	-0.004	-0.026
	(0.428)	(0.939)	(0.667)
R&D	25.271**	18.119**	19.166*
	(0.034)	(0.019)	(0.098)
Capex	-0.113	-0.075	-0.068
	(0.532)	(0.536)	(0.687)
ROA	-0.185	-0.096	-0.173*
	(0.113)	(0.314)	(0.086)
Firm age	0.065***	0.035***	0.051***
	(0.000)	(0.000)	(0.000)
Tobin's Q	0.006	-0.003	0.013
	(0.514)	(0.589)	(0.137)
Tangible	0.179	0.107	0.135
	(0.125)	(0.236)	(0.156)
Constant	-3.289***	-1.953***	-3.183***
	(0.000)	(0.000)	(0.000)
Observations	11,416	11,416	11,416
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R2	0.665	0.612	0.649

Table 7: Heterogeneous Treatment Effects based on Cities' Population without Local Urban Hukou

This table reports the cross-sectional variation of the treatment effects based on the city's population without a local urban hukou. The variable, *Population without hukou*, is the number of residents who do not have the city's urban hukou normalized by the city's total number of residents. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on standard errors clustered by city are in parentheses. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Relaxation × Population without hukou	-1.009**	-0.495*	-1.001**
	(0.044)	(0.087)	(0.035)
Relaxation	-0.036	-0.031	-0.003
	(0.575)	(0.464)	(0.952)
Population without hukou	0.171	-0.132	0.288
	(0.746)	(0.678)	(0.558)
Cash	-0.159	-0.092	-0.084
	(0.176)	(0.245)	(0.465)
Firm size	0.167***	0.119***	0.149***
	(0.000)	(0.000)	(0.000)
Leverage	-0.049	-0.008	-0.028
	(0.369)	(0.840)	(0.553)
R&D	16.928	15.377*	12.508
	(0.104)	(0.055)	(0.192)
Capex	-0.078	-0.072	0.011
	(0.614)	(0.495)	(0.941)
ROA	-0.127	-0.090	-0.124
	(0.209)	(0.186)	(0.179)
Firm age	0.023	-0.002	0.017
	(0.428)	(0.943)	(0.470)
Tobin's Q	0.002	0.003	0.004
	(0.787)	(0.570)	(0.624)
Tangible	0.186**	0.132*	0.138*
	(0.041)	(0.070)	(0.073)
Constant	-2.741***	-2.065***	-2.876***
	(0.002)	(0.000)	(0.000)
Observations	18,481	18,481	18,481
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R2	0.726	0.695	0.705

Table 8: Heterogeneous Treatment Effects based on Human Capital Intensity

This table reports the cross-sectional variation of the treatment effects based on the firm's human capital intensity. The variable, *Labor intensity*, is the expenditure on employee wages normalized by total revenue. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on standard errors clustered by city are in parentheses. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1) Ln (1+all patent)	(2) Ln (1+invention patent)	(3) Ln (1+utility model and design patent)
Relaxation × Labor intensity	-0.487**	-0.297**	-0.619***
	(0.013)	(0.034)	(0.000)
Relaxation	-0.078***	-0.041**	-0.035
	(0.006)	(0.043)	(0.169)
Labor intensity	0.167	0.203**	0.171
	(0.195)	(0.028)	(0.144)
Cash	-0.110*	-0.063	-0.035
	(0.084)	(0.165)	(0.544)
Firm size	0.175***	0.124***	0.153***
	(0.000)	(0.000)	(0.000)
Leverage	-0.054	-0.024	-0.027
	(0.183)	(0.406)	(0.470)
R&D	22.472***	20.259***	15.832***
	(0.000)	(0.000)	(0.000)
Capex	-0.056	-0.049	0.012
	(0.605)	(0.531)	(0.906)
ROA	-0.091	-0.065	-0.097
	(0.291)	(0.288)	(0.212)
Firm age	0.054***	0.026***	0.043***
	(0.000)	(0.000)	(0.000)
Tobin's Q	0.012**	0.007*	0.014***
	(0.027)	(0.072)	(0.005)
Tangible	0.127**	0.097**	0.090*
	(0.025)	(0.017)	(0.080)
Constant	-3.935***	-2.764***	-3.502***
	(0.000)	(0.000)	(0.000)
Observations	18,481	18,481	18,481
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
R2	0.692	0.663	0.669

Table 9: City-level Aggregate Innovation

This table reports the difference-in-differences tests that examine the impacts of hukou relaxation on corporate innovation, using city-level aggregate innovation. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	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	-0.165**	-0.116**	-0.131**
	(0.016)	(0.014)	(0.032)
Ln (city GDP)	0.153***	0.081***	0.129***
	(0.000)	(0.000)	(0.000)
Ln(city population)	0.064	0.034	0.046
	(0.168)	(0.288)	(0.268)
Ln(# of universities in the city)	0.076*	0.052*	0.061
	(0.076)	(0.075)	(0.111)
City income per capita	0.067***	0.042***	0.048***
	(0.000)	(0.000)	(0.001)
City expenditure on science and technology	-0.032	1.180	-0.022
	(0.980)	(0.193)	(0.985)
Constant	-2.765***	-1.513***	-2.239***
	(0.000)	(0.000)	(0.000)
Observations	3,361	3,361	3,361
Year FEs	Yes	Yes	Yes
City FEs	Yes	Yes	Yes
R2	0.536	0.500	0.516

Table 10: Effects of Hukou Relaxation on Migration

This table reports the difference-in-differences tests that examine the impacts of hukou relaxation on migration. The dependent variable in columns (1) and (2) is the number of people who newly obtain their urban hukou normalized by the total number of people with an urban hukou. The dependent variable in columns (3) and (4) is the number of people who newly obtain their urban hukou normalized by the total number of people in the city. For the cities that have implemented hukou relaxation that relaxes the restrictions for migrant workers to obtain local urban hukou, the indicator variable *Relaxation* takes the value of one for the period after the policy change, and zero for the period prior to the policy change. For the cities that have never implemented such hukou relaxation in our sample period, *Relaxation* always takes the value of zero. The sample consists of 1,685 city-year observations. 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
	# of people who newly obtain their urban hukou normalized by the total #of people with an urban hukou		# of people who newly obtain their urban hukou normalized by the total population with and without an urban hukou	
Hukou Relaxation	0.026**	0.026**	0.009**	0.009**
	(0.040)	(0.049)	(0.014)	(0.020)
Ln (city GDP)		0.013***		0.004***
		(0.009)		(0.006)
Ln(city population)		0.049***		0.013***
		(0.001)		(0.005)
Ln(# of universities in the city)		-0.004		-0.002
		(0.608)		(0.418)
City income per capita		0.001		-0.001
		(0.964)		(0.740)
City expenditure on science and technology		0.575*		0.277***
		(0.067)		(0.003)
Constant	0.155***	-0.302**	0.132***	-0.036
	(0.000)	(0.011)	(0.000)	(0.312)
Observations	3,361	3,361	3,361	3,361
Year FEs	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes
R2	0.229	0.242	0.330	0.344

Table 11: Timing of Hukou Relaxation and Pre-existing Corporate Innovation

The model is a Weibul hazard model here the dependent variable is the Ln (expected time to hukou relaxation). The sample period is from 1999 to 2010 and the sample consists of 30 cities that relaxed their hukou policies after 2000. Cities drop from the sample once they relax their hukou policy. 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. The superscript ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	1	2	3	4	5	6
	Ln(expected time to hukou relaxation)					
Ln (city-level average number of all patent)	-0.085					
	(0.129)					
Ln(city-level average number of invention patent)		-0.159				
		(0.133)				
Ln(city-level average number of utility model and design patent)			-0.061			
			(0.132)			
Change in Ln (city-level average number of all patent)				0.030		
				(0.450)		
change in Ln(city-level average number of invention patent)					0.018	
					(0.726)	
change in Ln(city-level average number of invention patent)						0.038
						(0.342)
Ln (city GDP)	-1.071***	-1.070***	-1.094***	-0.703***	-0.709***	-0.700***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.004)	(0.004)
Ln(city population)	-0.084	-0.088	-0.083	0.242	0.217	0.245
	(0.179)	(0.158)	(0.188)	(0.424)	(0.471)	(0.416)
Ln(# of universities in the city)	-0.143**	-0.123*	-0.143**	-0.116	-0.122	-0.113
	(0.038)	(0.071)	(0.039)	(0.119)	(0.101)	(0.127)
City income per capita	-0.028	-0.032	-0.028	0.199**	0.200**	0.200**
	(0.412)	(0.351)	(0.419)	(0.020)	(0.019)	(0.019)
City expenditure on science and technology	-0.237	-0.108	-0.253	0.467	0.382	0.493
	(0.887)	(0.948)	(0.880)	(0.787)	(0.825)	(0.775)
Constant	19.094***	19.078***	19.444***	8.877**	9.111**	8.791**
	(0.000)	(0.000)	(0.000)	(0.018)	(0.015)	(0.019)
Observations	219	219	219	188	188	188
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.859	0.859	0.863	0.874	0.874	0.874