How Does Higher Education Affect Local Entrepreneurship? Evidence from New Campuses in China *

Lu Lin†

Qiong Zhu†

Wei Ha†

November 14, 2024

*Lu Lin: Institute of Economics of Education, Peking University (email: llin@stu.pku.edu.cn); Qiong Zhu: Institute of Economics of Education, Peking University (email: qzhu@gse.pku.edu.cn); Wei Ha: Institute of Economics of Education, Peking University (email: wha@pku.edu.cn). Wei Ha is the corresponding author. We are grateful to the National Natural Science Foundation of China, the National Social Science Fund of China and the Ministry of Education of Humanities and Social Science Project for financial support. We acknowledge comments and suggestions from participants at the 45th Fall Research Conference of the Association for Public Policy Analysis and Management (APPAM), 11th International Workshop on Regional, Urban and Spatial Economics in China (RUSE), the 23rd China Economic Annual Conference (CEAC) and the 48th Annual Conference of the Association for the Study of Higher Education (ASHE). All errors are ours.

†Institute of Economics of Education, Peking University

Abstract

We identify the spillover effects of universities on local entrepreneurship, leveraging temporal and geographical variations in new campus establishments, as well as a unique instrument variable based on China's massive higher education expansion since 1999. Using large-scale firm registration data, we find that the opening of a new campus stimulates firm entry in the host county, especially in Knowledge-Intensive Business Services. These effects follow an inverse U-shaped pattern with geographic distance. Both knowledge spillover and human capital spillover are at play, primarily facilitated by increased firm patenting activities and high-skilled migration.

Key Words: Entrepreneurship; Knowledge Spillover; Higher Education; Innovation

JEL Codes: I25, L10, R11, O30

1 Introduction

Entrepreneurship is a central and enduring topic in economics as it brings both "creative destruction" and fosters regional growth (Aghion & Howitt, 2008; Glaeser et al., 2015; Schumpeter, 1942). Nowhere has entrepreneurship been more vibrant and dynamic than in post-reform China, particularly since its entry into the World Trade Organization (WTO) in 2001, when private enterprises started to mushroom and became the driving force behind the world's second largest economy (Atherton & Newman, 2017; Lerner et al., 2024). This process has created global brand names such as Alibaba, Tencent and Tiktok, along with tens of millions of small and medium enterprises. This is in sharp contrast to an overall stagnation of business dynamism in the U.S. over the same period (Akcigit & Ates, 2023; Hummels & Yue, 2024). In fact, China has surpassed the U.S. in terms of number of new establishments and the total number of establishments per thousand population in 2014 and 2018 respectively[1]. While researchers have explored China's remarkable success through the perspectives of "local state corporatism" (Oi, 1995), strengthened private property rights (S. Wang, 2012) and social environment and networks (Djankov et al., 2006), little attention has been paid to the role of higher education institutions (HEIs) in the rise of entrepreneurship in China. This gap is particularly glaring considering the importance researchers have attached to HEIs in driving economic growth and shaping national and regional innovation systems (Etzkowitz & Leydesdorff, 1995; Jaffe, 1989; Lucas Jr, 1988; Romer, 1986) and the unprecedented expansion of China's higher education sector since 1999.

/ Insert Figure 1 Here/

In 1999, the Chinese government decided to accelerate the expansion of higher education to reverse the downward economic trend due to the restructuring of state-owned enterprises and the Asian financial crisis (Che & Zhang, 2018; J. Li et al., 2020). As a result, annual new entrants at

[1]See Figure 1 for comparisons of total and Knowledge-Intensive Business Services (KIBS) establishments and openings between the U.S. and China from 1995 to 2020.

Chinese higher education institutions (HEIs) increased fivefold from 1.08 million in 1998 to 5.46 million students in 2006 and enrollment continued to grow after 2006 though at a lower speed. This rapid and persistent growth of enrollment put immense pressure on HEIs as most campuses were located in urban areas and could not afford to acquire land nearby. Many had to seek potential sites in neighboring suburban districts or rural counties to construct satellite campuses. The total number of these new campuses established between 1999 and 2020 amounted to 930.

This study takes advantage of the geographical and temporal variation in the number of new campuses in a county to examine the effects of higher education institutions on local entrepreneurship. While we acknowledge that the site selection processes for the new campuses are unlikely to be completely random, we believe the timing of the opening of a new campus is reasonably exogenous given the lengthy bureaucratic processes involving different departments at different levels of the Chinese local government (sometimes involving the central government line ministries as well if the universities fall under their jurisdiction). Specifically, we link the firm registration data from the State Administration for Industry and Commerce of China (SAIC) to a manually collected dataset of 930 new campuses to construct a unique county-level panel data between 1995 and 2020.

We conduct a generalized difference-in-differences (DD) analysis, controlling for baseline characteristics (population density and night-time light) interacted with year dummies, county fixed effects, year fixed effects, as well as province-year fixed effects. Baseline regression results show that the opening of a new campus increases the total firm entry in a treatment county by 6.5 log points (6.7%), equivalent to an average increase of 49 new firms. This effect is driven primarily by a 12.0% rise in the creation of firms in Knowledge-Intensive Business Service (KIBS), followed by a 6.9% increase in firm entry of non-KIBS industries. The effect on manufacturing firms is minimal and is economically and statistically indifferent from zero. To further test whether the site selection and timing of new campus constructions are reasonably exogenous, we perform event study analyses using various heterogeneity-robust DD estimators, which show that the common trend

assumption holds. These results also survive a number of robustness checks, including synthetic DD method, sample restrictions, alternative models and placebo tests.

It is not entirely unreasonable to worry that local governments in China may open new campuses where a booming entrepreneurship sector is in the making, which would lead to overestimating the effect of universities (Andrews, 2023; Lehnert et al., 2024). Therefore, we further leverage an instrumental variable (IV) approach, exploiting both the push factors driving a university to set up a new campus (i.e. distance of old campuses to the city center and time-variant scale of enrollment expansion at the provincial level) and a pull factor that makes a county attractive to universities that prefer sites not too far away from their old campuses (i.e. distance to their old campuses). Since counties have neither the control of the historical sites of old campuses nor the control of degree of university expansion, this instrument will help us further tease out the endogeneity in our estimates. The IV results consistently reveal positive and larger impacts of new campus openings on local entrepreneurship, e.g. a new campus leads to 33.1% more KIBS firms in the treatment counties.

Mechanism analyses show that universities affect local entrepreneurship through both knowledge spillover and human capital spillover. Using patent data from the National Intellectual Property Administration of China (CNIPA) between 1995 and 2015 matched with firm registration data, we show that KIBS firms in counties with new campuses file more patents, especially younger firms and firms in the scientific and technical services (R&D) and information services (ICT) industries. Human capital spillover effect is more challenging to analyze given the paucity of nationally representative longitudinal household data in China. We aggregate the 2015 Chinese population mini-census data to the county level and show that counties with one more new campus boast 10.6% more high-skilled labor (college degree or higher) in their population. In both cases, the indirect channels dominate the direct channels (all other patenting vs. university-firm joint patenting and migrant graduates vs. local graduates). Also, 3-year HEIs are found to play a notable role, at least on par with 4-year HEIs. Following Rosenthal and Strange (2008) and J. Li

et al. (2020), we also conduct a concentric ring analysis to identify the geographic pattern of university spillover effects. We found an inverse U-shaped pattern of the effects for KIBS industries, which peak at around a 15-kilometer radius and disappear beyond 30 kilometers.

Our main contribution is to the literature on the effects of HEIs on the local economy and innovation. Earlier studies have either been descriptive such as the regional innovation system pioneered by Asheim (2007) and Triple Helix model coined by Etzkowitz and Leydesdorff (1995) or correlational studies that are more popular in the field of economic geography that take the university spillover as one of the determining factors in firms' location choices, quite often in a couple of narrowly-defined industries[2]. More recent studies have exploited plausibly exogenous variation in local university research intensity to find that it improves significantly the productivity and innovativeness of incumbent firms (Andrews, 2023; Hausman, 2022; Kantor & Whalley, 2014; Liu, 2015) and it serves as a great catalyst for new economic activity and entrepreneurship (Cantoni & Yuchtman, 2014; Tartari & Stern, 2021). We contribute to this fast-expanding strand of literature by providing a unique instrument that predicts the number of new campuses in a county in a particular year but remains sufficiently exogeneous to the dynamics of entrepreneurship in that county when the popular runner-up strategy in the literature (Andrews, 2023) is not feasible. Given that the constructed instrument stems from pre-determined geographical characteristics and a policy-induced HEI expansion, we argue that our results are picking up a causal relationship.

We also add to this literature by examining the spillover effects of HEIs on local development in the context of rapid higher education expansion in a developing country. Most existing causal studies focus exclusively on western economies, with the exception of Sekhri et al. (2024) and Vu and Vu-Thanh (2022). This oversight is particularly striking given that large developing countries such as Brazil, Mexico, India, Indonesia and China accounted for a significant proportion of universities in the world (Valero & Van Reenen, 2019). UNESCO (2020) shows that all developing

[2]see Audretsch et al. (2005) for a typical co-locational study; see Tartari and Stern (2021) for a comprehensive summary on the pre-existing literature.

regions had doubled or even tripled access to higher education between 2000 and 2018[3]. China experiences the largest increase in higher education enrollment between 2000 and 2020[4]. Our results are informative and relevant to other developing countries in the global south as they contemplate the nexus between universalization of higher education and quest for economic growth.

Our findings provide novel evidence on the role of higher education expansion in explaining the meteoric rise of China's economy and its great transformation from a state-dominated economy to an economy driven primarily by private enterprises in the last four decades. While earlier studies have attributed this transformation in the earlier phase of manufacturing industry development to the "decentralized development state" (Oi, 1995), "transitional institutions" such as township and village enterprises (Qian, 2003), overseas capital inflows from Chinese diaspora (Smart & Hsu, 2004) or privatization of state-owned housing in urban China (S. Wang, 2012), more recent government documents have clearly put more emphasis on building a knowledge economy in China through science, technology and education (Dahlman & Aubert, 2001). Our paper sheds light on how China leverages university-industry linkage to facilitate its transition to a service-based and innovation-driven economy using county panel data over the last two decades, in particular, cultivating the growth of KIBS which is believed to be the key to an economy's competitiveness and innovation (OECD, 2006, 2015).

The rest of the paper is organized as follows. Section 2 explains the institutional context of higher education expansion and new construction in China. Section 3 lays out the empirical strategies used in the quantitative analyses. Section 4 describes the data and the summary statistics. Section 5 reports the analytical results while Section 6 concludes and discusses the implications of our findings.

[3]https://unesdoc.unesco.org/ark:/48223/pf0000375686.

[4]UNESCO Institute for Statistics (UIS) database, Education and literacy, http://data.uis.unesco.org, 2024-10-11.

2 China's Massive Higher Education Expansion and the Construction of New Campuses

Shortly after China's reform and opening-up in 1978, China decided to resuscitate its higher education sector. It first grew at an annual rate of 8.4%, a speed commensurate with its stellar economic growth between 1977 and 1998. In 1999, the Chinese government opted to accelerate the expansion of higher education in response to the declining economic trend resulting from the restructuring of state-owned enterprises and the 1997/1998 Asian financial crisis (Che & Zhang, 2018; J. Li et al., 2020). This unexpected and unprecedented policy shift aimed to kill two birds with one stone, delaying the entry of high school graduates into the labor market and at the same time, boosting the economy by stimulating demand through tuition payments and other student-related consumption expenditures (Kang, 2000). While the initial plan was set to double annual college enrollment in three years, the target was achieved in two years with a 47% increase in 1999 and another 38% increase in 2000. The growth had somewhat attenuated but continued to be in the double digits until 2005. Starting in 2006, the speed of expansion returned to its pre-reform normal of single-digit but the compound growth translated into an unprecedented increase in the total number of enrolled students in Chinese HEIs from 3.4 million in 1998 to 17.4 million in 2006 and then 32.9 million in 2020 (see Figure 2). Previous literature has taken advantage of this expansion as a natural experiment to examine the human capital effect of universities expansion at the industry level on firms' productivity (Che & Zhang, 2018; Feng et al., 2022) or the knowledge spillover effects of universities expansion on local patents at the city level (J. Li et al., 2020).

As J. Li et al. (2020) rightly pointed out, "the enrollment quota was allocated to different cities mainly based on the city-level pre-existing physical and logistics capacity of the higher education sector". We draw on a more refined variation in the expansion of higher education at the county level. Most HEIs were located in urban districts of these cities and the size of their campuses was not sufficient to accommodate such a huge influx of students. Consequently

per capita school facilities dropped from 45 in 1998 to 37 square meters per student in 2000. Therefore, there was an urgent need to construct more new school facilities. However, urban lands were becoming more and more expensive due to the rapid economic growth and a burgeoning demand for commercial and residential housing. Chinese HEIs had to look for potential sites in nearby suburban districts or rural counties to construct new campuses (D. Chen, 2022), echoing a similar trend of decentralization of higher education in Europe (Andersson et al., 2009; Nimier-David, 2023).

Between 1999 and 2020, 930 new campuses were constructed across 418 counties. As Figure 2 shows, the peak of construction happened in 2003, two years after the peak of speed of expansion, with 62 new campuses constructed a year, plateauing at about 40 per year. New campuses tended to concentrate around the periphery of major cities not far from the old sites of the university campuses (See Figure 3 for a map of geographic locations of all new campuses between 1999 and 2020).

/ Insert Figure 2 Here/

/ Insert Figure 3 Here/

To address endogeneity concerns on the opening of new campuses, we include only new campuses built within the same city and exclude two types of institutions in the analysis: (i) campuses built across cities and (ii) brand new universities or colleges. We note that suburban districts or rural county governments certainly have strong economic motives to build or attract HEIs to their locality, which could potentially make the site selection endogenous. We searched for the information on these institutions and identified 124 campuses built across cities and 188 brand new universities between 1999 and 2020, primarily in coastal provinces or provinces with significant economic and population clout but with relatively few higher education resources[5]. We exclude these

[5]Five coastal provinces (Shandong, Guangdong, Jiangsu, Zhejiang and Fujian) account for close to half of the 182 new universities. If we include the other four provinces (Henan, Sichuan, Hubei, Hunan) with GDP ranking among the top 9 in China, the proportion would go up to two-thirds.

institutions from our sample because they reflect the need of these localities for higher education in a context of rapid economic growth. Our analyses draw on the more exogenous variation coming from those new campuses built within the same city (hereinafter referred to as new campuses).

Our identification strategy first leverages the exogeneity from the geographical and temporal variations of new campus openings. First, since HEIs are subject to the jurisdiction of provincial or city governments, the county/district (hereinafter referred to as county) governments only have limited say in their final decisions. When making their locational decisions, HEIs primarily consider the commute time between the old campus and the new site, the characteristics of the land parcel and the associated price offered by various counties among other things. Second, we also leverage another layer of exogenous variation in the actual timing of the opening of new campuses. The process of the construction of a new campus is rather lengthy and bureaucratic and involves stakeholders at different departments and at different levels of governments such as faculty and administrators at the universities, officials at the Education Department, Finance Department, Development and Reform Department, Department of Natural Resources at the county, municipal and provincial level. Any delay or disagreement at any step of the processes could postpone the approval of the new campuses.

Of course, one cannot rule out entirely the possibility that higher-level governments may intentionally locate new campuses in such a way in coordination with local clustering of industries that confounds our DD estimates based on geographical and temporal randomness of new campus openings. Therefore, we further take an instrumental variable (IV) approach and exploit two other sources of exogeneity. First, we construct a distance-based measure of a county's inherent attractiveness as the sites of the new campuses, namely the distance of a county to all the campuses built before the expansion (hereafter referred to as old campus) of universities with varying needs of physical expansion in its city. Second, we incorporate a time-varying factor of the expected change in province-level higher education enrollment into the instrument to capture the degree of policy-induced university expansion. Since counties have neither the control of the site of histori-

cally built campuses nor the control of the expansion policy, this instrument will help improve the causality of our results. See Section 3.2 for more details on the construction of the instrument.

3 Empirical Strategy

3.1 Difference-in-differences and Event Study

Our main goal is to identify the causal impact of universities and colleges on local entrepreneurship by utilizing the temporal and geographic variations associated with new campus openings. In the baseline analysis, we employ a difference-in-differences (DD) approach, estimated using a two-way fixed effect (TWFE) model, with the county as the unit of analysis. This method enables us to compare the level of entrepreneurship in a treated county before and after a new campus is introduced (the first difference) to the changes in control counties during the same period (the second difference). Following Nunn and Qian (2011), we use a continuous measure of the intensity of treatment in the DD framework to estimate the marginal impact of building one additional campus using the following specification:

$$Y_{i,t} = \alpha + \beta \text{ NewCampus }_{i,t-1} \times \text{ Post }_{i,t} + X'_{i,t}\eta + \lambda_i + \mu_t + \varphi_{p,t} + \varepsilon_{i,t} \tag{1}$$

where $Y_{i,t}$ represents the logarithm of one plus the number of firm entries in county i during year t; NewCampus $_{i,t-1}$ equals the cumulative number of new campuses built in county i between 1999 and year $t-1$. Given the focus of the study on the impact of new campus openings within the context of China's 1999 higher education expansion efforts, only new campuses built after 1999 are considered treatment events. Campuses that existed prior to 1999 are classified as old campuses[6]. Post $_{i,t}$ is a binary indicator variable that equals one for treated counties in the years following the

[6]We conducted a thorough search for campuses constructed between 1995 and 1999 using the same method described in Section 4.1. We identified 14 campuses built during this period and present results addressing these campuses in Section 5.4.

construction of their first new campus and zero otherwise. To ensure that sufficient pre-treatment periods are available for counties treated in the early 2000s, our analytical sample covers the years from 1995 to 2020. $X'_{i,t}$ is a vector of baseline county-level covariates (the logarithm of night-time light in 1995 and the logarithm of population density in 1995) interacted with year dummies to control for heterogeneous trends in local demographic and economic conditions; λ_i is a county fixed effects term, capturing time-invariant county characteristics such as geographic location; μ_t is a year fixed effects term capturing time-variant common shock in firm entry; $\varphi_{p,t}$ is a province-by-year fixed effects term that controls for province-specific yearly variation in firm entries, as literature has shown strong regional disparities in entrepreneurial activities, with coastal provinces taking the lion's share (Atherton & Newman, 2017); and $\varepsilon_{i,t}$ is the error term. To address potential heteroscedasticity and serial correlation, standard errors are clustered at the county level.

A valid DD identification relies on parallel trends between treatment and control groups during pre-treatment periods. To confirm this, we use an event study approach to capture the dynamic effects on local entrepreneurship over time. The estimating equation is set as follows:

$$Y_{i,t} = \alpha + \sum_{k=-10}^{15} \gamma_k D_{i,t}^k \times \text{ Post }_{i,t} + X'_{i,t}\eta + \lambda_i + \mu_t + \varphi_{p,t} + \varepsilon_{it} \tag{2}$$

where $D_{i,t}^k$ denotes the relative period of k years to the first year of opening a new campus in county i; other settings are identical to Equation 1. As the effects of new campuses may vary across counties and years over the 26-year span in our observation period, the TWFE estimator could be biased (Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021). We present in Section 5.2 the dynamic effects estimated using several heterogeneity-robust estimators.

3.2 Instrumental Variable Approach

To estimate the causal effect of building a new campus, the popular strategy is to identify "runner-up" sites as valid counterfactuals for selected college locations (Andrews, 2023; Russell et

al., 2024). However, this approach is impractical in the context of China due to the large number of universities building new campuses within a short period. Counties not initially chosen as campus sites may later be selected by other universities. This fact makes it difficult to identify a pure "runner-up" for new campus sites. Therefore, we employ an alternative instrumental variable (IV) strategy that leverages three exogenous variations in the decision-making process for campus site selection. These variations include both the push factors driving a university to set up a new campus and the pull factors that make specific counties appealing to universities.

Push factors. We consider two exogenous variations from push factors motivating new campus openings. The first is the level of higher education expansion in the province where the university is located. As discussed in the Introduction, the primary motivation for universities in China to build new campuses over the past two decades has been the rapid enrollment growth resulting from the national higher education expansion policy since 1999. While this policy affected the entire nation, there is a significant geographic variation in the growth of expansion at the province level stemming from provinces' varying exposure to the expansion policy. Universities facing greater enrollment pressure were more likely to build new campuses. Given that college admission quotas in China are primarily assigned at the university or provincial level, the provincial expansion is not influenced by the economic conditions of individual counties. To avoid endogeneity concerns, we use the province-level expansion level rather than city- or county-level rates to measure the expansion pressure faced by a university. This approach is consistent with strategies used in studies like J. Li et al. (2020) and Ma (2024), which employ province or city-level expansion rates as exogenous shocks to estimate the impact of college expansion.

When measuring the province-level higher education expansion, a drawback of directly using the actual higher education enrollment is that the level of expansion may be endogenous to the economic development of the province. That is, universities and colleges in a province with a higher level of entrepreneurship may be assigned with more quotas of enrollment. To address this concern, we use the potential level of higher education enrollment to proxy for province-level

expansion following H. Li et al. (2017). The potential higher education enrollment expansion of province p in year t is given by:

$$Expansion_{p,t} = TotalEnrollment_t \times EnrollmentShare_{p,1998} \tag{3}$$

where $TotalEnrollment_t$ is total higher education enrollment in year t, which measures expansion at the national level. $EnrollmentShare_{p,1998}$ is the share of province-level enrollment in 1998, which measures each province's exposure to the expansion policy. The use of potential expansion level leverages the pre-determined province-variant exposure to the policy and therefore the estimates are not affected by any post-expansion provincial characteristics. It captures only the expansion in response to the national policy (H. Li et al., 2017). We show in Appendix A the enrollment for 31 provinces in 1998.

The second exogenous variation related to the push factors for new campus openings is the distance from old campuses (i.e., those built before 1999) to the city center. We observe that universities originally located near the city center often faced constraints in expanding their campus space due to limited land availability and high land prices (Zhao, 2010). As a result, these universities were more likely to build new campuses located farther from the city center. We therefore calculate the distance from each old campus to the city center in 1998, expecting this distance to be negatively correlated with the number of new campuses built. Since the locations of old campuses and the city center were historically determined, they are exogenous to the economic characteristics of individual counties[7].

[7]We note that there may be concerns on the endogeneity of counties' geographical distance from the city's old campuses. Counties that locate close to the old campuses may be economically advantaged as they are also closer to the city center. To test whether the distance from old campuses is related to counties' economic performance, we regress the log average distance to the city's old campuses on the county's growth rate of total firm entry, night-time light and population density during pre-expansion periods (1990-1995) on the IV sample. The results indicate insignificant correlation between a county's economic and demographic profiles and the distance from the old campuses. We also provide DD and IV estimates using only suburban counties adjacent to municipal districts to ensure a more balanced sample in terms of geographical proximity to the city center. We present these results in Appendix B.

Pull factor. The final exogenous variation we use reflects the pull factors that attract universities to build new campuses in specific counties. This variation is essential for capturing cross-county differences that is related to the endogenous variable of interest while being exogenous to the county's local economic characteristics. Since the previous two factors do not vary by county, this cross-county variation is crucial for constructing an IV for the number of new campuses opened in each county. To achieve this, we take a distance-based approach and incorporate the geographic distance from each county to all existing old campuses within the city into the instrument. The underlying rationale is that universities were more likely to choose new campus sites with shorter commuting distances to minimize transportation and time costs, given that faculty, staff and students often needed to travel between campuses. Therefore, this measure is expected to be negatively correlated with the number of new campuses built in a county. Importantly, since the locations of old campuses are not affected by the economic characteristics of the destination counties, this distance measure is considered exogenous to the local economic conditions of these counties.

We then combine these three variations from both the push and pull factors to construct the IV as follows:

$$IV_{i,c,p,t} = \sum_{j=1}^{M} \left(\underbrace{\frac{\text{Expansion}_{\ p,t}}{\sqrt{(\text{Distance to City Center})_{\ j,c,p}}}}_{\text{Push factor}} \times \underbrace{\frac{1}{\sqrt{(\text{Distance to Old Campus})_{\ i,j,c,p}}}}_{\text{Pull factor}} \right) \tag{4}$$

where $IV_{i,c,p,t}$ is the instrument variable for the number of new campuses built in county i, located in city c of province p, between 1999 and year t. $\text{Expansion}_{\ p,t}$ represents the potential higher education expansion of province p in year t since 1998. j denotes an old campus in city c and M is the total number of old campuses in city c prior to the expansion. $(\text{Distance to City Center})_{\ j,c,p}$ measures the geographic distance between old campus j and the center of city c. Since this distance

negatively influences the likelihood of an old campus building a new campus, we take the reciprocal of the distance. Additionally, we follow Zacchia (2020) to attribute more importance to old campuses located closer to the city center by applying a square root function to the distance, reflecting the larger pressure they faced to open a new campus. Taken together, $\frac{\text{Expansion}_{p,t}}{\sqrt{(\text{Distance to City Center})_{j,c,p}}}$ captures the overall push factors for existing universities in a city to open new campuses. The pull factor is represented by $\frac{1}{\sqrt{(\text{Distance to Old Campus})_{i,j,c,p}}}$, where $(\text{Distance to Old Campus})_{i,j,c,p}$ is the geographic distance between county i and old campus j in city c. This reflects the geography-related preference for old campus c to build a new campus in county i. The square root form gives greater weight to county-old campus pairs with shorter distance and the reciprocal form facilitates interpretation (Zacchia, 2020). While the potential enrollment is generally positively correlated with the number of new campus openings, both distance measures are negatively correlated with this number. Finally, by taking the product of the push and pull terms and summing over old campus $j = 1, 2, 3, \cdots, M$ in the city, we obtain the IV for the number of new campus openings in county i for year t and run a two-stage least squares (2SLS) regression where the second stage regresses local entrepreneurship outcomes on the predicted number of new campus openings instrumented by the IV. It is worth noting that we include only peripheral counties by excluding 887 central counties (conventionally known as municipal districts), resulting in a reduction of 43.5% county-year observations in the IV sample. This exclusion is necessary because most existing campuses are located in central counties; including them would result in a high proportion of extremely small distance values for $(\text{Distance to Old Campus})_{i,j,c,p}$, which may potentially bias the estimation. We show in Appendix B a further illustration of the intuition of constructing the instrument.

3.3 Concentric Ring Analysis

To identify the geographic pattern of university spillover effects, we follow Rosenthal and Strange (2008) and J. Li et al. (2020) to construct a set of concentric ring variables that mea-

sure entrepreneurship at varying distances from a new campus. Each concentric ring spans 1 kilometer, with 50 rings covering areas up to a 50-kilometer radius around the new campus. Using geocoded firm locations, we calculate the number of firm entries for each concentric ring around the new campus, comparing entrepreneurship within a ring before and after the campus opening and between each ring and the reference ring (within a 1-kilometer radius of the central campus). The specification is as follows:

$$Y_{r,u,t} = \alpha + \sum_{r=2}^{50} \sigma_r \text{ Ring }_{r,u} \times \text{ Post }_{u,t} + \lambda_{u,t} + \varphi_{r,u} + \varepsilon_{r,u,t} \tag{5}$$

where $Y_{r,u,t}$ is the logarithm of one plus the number of firm entries in ring r around new campus u in year t. Ring $_{r,u}$ is a set of dummy variables representing concentric rings covering up areas up to 50 kilometers away from the nearest new campus u (the closest ring is set as the reference group and omitted). Post $_{u,t}$ is an indicator for the post-treatment period for a new campus u. σ_r then estimates the university spillover effect on the nearby rth ring, allowing us to identify the geographic distribution of the university spillovers on entrepreneurship. $\lambda_{u,t}$ is a campus-by-year fixed effects term that captures campus-level variations in factors such as enrollment, revenue and expenditure each year. $\varphi_{r,u}$ denotes campus-specific ring fixed effects, controlling for all time-invariant characteristics such as topographical factors and distances to the city center within a campus-specific ring.

4 Data and Variables

4.1 New Campus Data

One of the major challenges in examining the effects of new campus openings in China is the absence of comprehensive administrative data from government authorities. Unlike the Integrated Postsecondary Education Data System (IPEDS) in the U.S., China lacks a similar data collection

program. To address this, we manually collected information on college and university campus openings between 1995 and 2020 through news reports, university historical archives and personnel interviews.

Starting with the official list of 2,740 universities and colleges, we searched for information on each campus affiliated with these institutions. We first consulted the universities' official websites, which often provide dates and locations for new campus openings. We then used public search engines and the WiseNews database, which contains historical news reports from over 12,000 media outlets dating from 1999 to the present, to find additional information. For campuses with insufficient public data, we sought information from university and government administrative authorities and conducted campus visits and interviews with university staff and local government officials[8]. The final dataset comprises over 4,106 new and old campuses from 2,655 universities and colleges in China[9], featuring key campus-level information, including geographic addresses (used for geocoding to obtain latitude and longitude coordinates) and dates related to various stages, such as approvals, construction and openings.

Our dataset indicates that 877 (32.0% of the total) universities and colleges have built a total of 930 new campuses within their home city between 1999 and 2020. Among these, 495 are new campuses of 4-year universities, while 435 belong to 3-year vocational colleges primarily offering associate degrees.

[8]We visited 48 universities to examine their decision-making processes and site selection for new campus openings. Our site visits include universities and colleges in Beijing, Binzhou, Changsha, Changzhou, Datong, Dongguan, Guangzhou, Lanzhou, Nanjing, Qingdao, Quanzhou, Shanghai, Shenyang, Shenzhen, Suzhou, Tianjin, Wuhan, Xi'an, Xiamen, Yantai and Yibin, representing the eastern, central and western regions of China.

[9]Eighty-five universities and colleges were excluded from this analysis due to closures or mergers with other institutions.

4.2 SAIC Firm Registration Data

Data on local entrepreneurship are obtained from the firm registration records of the State Administration for Industry and Commerce of China (SAIC)[10]. This dataset includes all business establishments registered with SAIC, covering both state-owned enterprises (SOEs) and non-SOEs across all industries. Following conventional practice in studies of local entrepreneurship in China (He et al., 2024; Wei et al., 2024), we exclude sole proprietorships from our analysis. This results in a dataset of 93.12 million firm establishments. We include firms created between 1995 and 2020 to allow the data to straddle across the higher education expansion. Each record includes information on the date of entry, date of exit, two-digit industry based on China Industry Classification Standard (CICS) 2017[11] and ownership at the time of registration. The precise addresses of each registration allow us to use the Baidu Map API to geocode and obtain the longitude and latitude of each firm upon its creation.

In addition to examining overall firm entry, we further categorize the data into four industry groups: KIBS, non-KIBS, manufacturing and others[12]. The proportions of new firms in each group relative to the total number of new firms are 24.2%, 51.2%, 10.5% and 14.1%, respectively. We take firm entries in KIBS, non-KIBS and manufacturing as the primary outcomes of interest. The firm entry of other industries is not examined as a separate outcome due to its relatively weak relevance to universities. When necessary, we further disaggregate the results into four sub-industries

[10]The SAIC firm registration data are newly available data that document the most comprehensive information on Chinese companies. Primary data sources used in previous studies on entrepreneurship and firm innovation are the Annual Survey of Industrial Enterprises (ASIF), Chinese Listed Firms Database (LFD) and Corporate Tax Survey carried out by Chinese State Administration of Tax (SAT). These data share a common disadvantage in the limited representativeness and observations. The ASIF database covers only manufacturing enterprises with annual sales above RMB 5 million ("above scale"), e.g. Che and Zhang (2018). The LFD database includes only information of less than 6,000 listed firms, e.g. C. Lee et al. (2021). The SAT data document only a small proportion of firms that file payroll taxed with the SAT and do not enjoy regional representativeness by the sampling procedure, e.g. Z. Chen et al. (2021). In comparison, the recently available SAIC data are more comprehensive and therefore adopted in several latest studies on China's entrepreneurship and innovation (Cui & Li, 2023; Tian & Xu, 2022).

[11]China Industry Classification Standard (CICS) 2017 is built upon International Standard Industrial Classification (ISIC Rev.4) with mild modifications. See Appendix C for the relationship between China Industry Classification Standard (CICS) and North American Industry Classification System (NAICS).

[12]Other industries include (i) agriculture, forestry, fishing and hunting , (ii) mining, quarrying and oil and gas extraction , (iii) construction and (iv) utilities.

within KIBS: (1) Scientific and Technical Services (R&D) (32.0%); (2) Information Services (ICT) (14.6%); (3) Finance and Insurance (Finance) (3.6%); and (4) Professional services, including management of companies and enterprises (49.7%). For a detailed list of industry classifications and the relationships between CICS and NAICS, please see Appendix C.

4.3 CNIPA Patent Data

To measure local innovation when analyzing the mechanisms of university spillovers on local entrepreneurship, we use patent data from the National Intellectual Property Administration of China (CNIPA). This dataset provides an exhaustive list of patents applied between 1985 and 2015 in China, including details such as patent ID, application date, approval date, inventor name and patent types. Our analysis covers 1,006,441 patent applications filed by firms in KIBS industries between 1995 and 2015.

We link the patent data with firm data using the assignee and firm names. Assignees are classified into categories: universities, firms, individuals and other organizations. We focus on joint patent applications by universities and firms. Furthermore, by identifying whether the co-applicants include an affiliated university of a new campus in the same county where the firm is located, we determine the number of joint patent applications by local universities and firms. More details on processing the patent data are provided in Appendix D.

4.4 Population Census Data

We use population census data to assess the effect of new campus openings on local human capital, one of the mechanisms underlying university spillovers on local entrepreneurship. Our analysis uses a 10% nationally representative sample from the 1% population survey (commonly referred to as the mini-census) conducted by the National Bureau of Statistics of China in 2015, which includes 1.37 million observations. Unfortunately, this data does not provide information on where or when respondents received their postsecondary degrees. However, it does offer detailed

information on the counties of individual's residence one year and five years prior to the survey, which enables us to identify graduates from local universities and colleges as well as migrants with college degrees. Specifically, we limit our sample to recent graduates aged 22 and 30 who have received a postsecondary degree. If a recent graduate resided in the same county at the time of the survey as they did one year and five years ago, we assume that they graduated from the local institutions and categorize them as the "direct human capital spillover" from universities. In contrast, those who moved from other counties are considered "indirect human capital spillover" migrants. For more details on the identification of these groups, see Section 7.2 and Appendix E. All mini-census data are aggregated to the county level and linked with the campus data for further analysis.

4.5 Main Sample

After integrating the various data sources, we obtain our main analytical sample, consisting of county-year level observations from 1995 to 2020[13]. After removing observations with missing values in control variables (night-time light and population density in 1995)[14], we construct an unbalanced panel with 67,420 county-year observations on 2,701 counties for the baseline analysis. Table 1 presents summary statistics of the main sample.

/ Insert Table 1 Here/

For the county-level analysis, there are 411 treatment counties with at least one new campus built between 1999 and 2020 and 2,290 control counties. Among treatment counties, 56.7% have

[13]We use county as the unit of analysis to provide a more accurate estimate of local spillovers from universities, as opposed to the city-level or higher level used in most studies based on China's context (Che & Zhang, 2018; J. Li et al., 2020)

[14]The night-time light data is obtained from "An improved time-series DMSP-OLS-like data (1992-2022) in China by integrating DMSP-OLS and SNPP-VIIRS" archived in the Harvard Dataverse (https://doi.org/10.7910/DVN/GIYGJU). Population density data is sourced from the China Population Spatial Distribution Kilometer Grid Dataset (https://www.resdc.cn/DOI/DOI.aspx?DOIid=32) . There are 341 counties missing night-time light data and 351 counties missing population density data in 1995 among the total of 3,181 counties/districts, resulting in the loss of 480 counties in the analytical data. These counties are mainly located in more remote provinces such as Xinjiang and Tibet and only seven of them have hosted new campuses between 1999 and 2020.

hosted only one new campus during the observation period, 28.5% have hosted two new campuses and 14.8% have hosted more than two new campuses. 60.8% of the treated counties hosted a new 4-year university campus, while 64.9% hosted a new 3-year college campus. Appendix F compares the initial economic and demographic conditions between treatment and control counties. Treatment counties demonstrate a higher level of entrepreneurship and economic development compared to control counties. This is attributable to the fact that most treatment counties are close to urban areas and are generally more economically advantaged than other remote counties. Therefore, we have controlled for the time-varying effects of baseline economic and demographic characteristics and conducted several robustness checks in Section 5.4 to address this issue.

In terms of local entrepreneurship, there were, on average, 984 firms across all industries established for each county-year observation, including 240 KIBS firms, 502 non-KIBS firms and 102 manufacturing firms. Among KIBS firms, new firms in Professional services holds the largest share, followed by R&D, ICT and Finance.

For the concentric ring analysis, we calculate the distance between each firm and the nearest new university or campus, excluding firm-campus pairs that are over 50 kilometers apart. Each firm is assigned to a ring based on its distance (in kilometers) from the new campus. Collapsing the firm-level data by year of creation, new campus and distance to the center, we create a campus-ring-year panel with 1,158,117 observations between 1995 and 2020.

5 Main Results

5.1 The Effects of University Spillover on Local Firm Entry

In line with the primary goal of estimating the spillover effects of universities, we first present the results estimated with Equation 1 in Table 2. Each coefficient captures how a new campus opening would affect the log point of firm entry in the host county. In Column 1, we show that, after controlling for baseline county characteristics and province-year fixed effects, new campus

openings increase overall firm entry by 6.5 log points (6.7%) per campus. Given that the baseline mean of control groups is 730 firms per county, this effect translates to an average increase of 49 new firms per county.

/ Insert Table 2 Here/

The magnitude is considerably large when compared to the benefits of place-based programs in China. Special Economic Zones (SEZs) are considered as a major place-based policy in China (Lu et al., 2019). Among the various types of SEZs, national high-tech industrial development zones (HIDZs) are the most technology-oriented zones aimed at facilitating technology-based innovation (Lu et al., 2019).An HIDZ has been found to increase firm entry in the host city by an average of 12.9 log points (13.8%)[15]. Therefore, the impact of a new campus opening can be translated to approximately 50% of the effect of an HIDZ.

In Columns 2-4, we group firms into KIBS, non-KIBS and manufacturing sectors, excluding firms from the first industry and other secondary industries. We find that the opening of new campuses is associated with an 11.3 log points (12.0%) increase in firm creation in KIBS, corresponding to approximately 19 new firms per campus. Firm creation in non-KIBS industries also increases by 6.7 log points (6.9%) per campus. However, there is no significant effect observed on firm entry in manufacturing.

In Columns 5-8, we further separate KIBS industries by R&D, ICT, finance and professional services. We find that new campus openings have the largest impact on firm creation in the ICT industry with a 17.9 log points (19.6%) increase, translating to approximately 4 new firms per campus. Firm creation in R&D also increases by 10.0 log points (10.5%), equivalent to about 5 new firms per campus. Additionally, firm creation in professional services increases by 8.8 log points (9.2%), amounting to about 8 firms per campus. The magnitude of the effect of new

[15]Tian and Xu (2022) found that based on SAIC data, the establishment of a national HIDZ leads to a 12.9 log points (13.8%) larger increase in the number of new firm registrations one year after the establishment, compared with that of the control cities surrounding the high-tech zone establishments.

campus openings on firm entry in finance appears to be the least with about 0.7 new firms per campus. These results suggest the significant role of universities in fostering local entrepreneurship, particularly in ICT, R&D and professional services, while the impact on the finance industry is comparatively minimal.

5.2 Event Study Estimation with Heterogeneity-robust Estimators

Given that the KIBS sector is crucial in the transition of local economies and the development of innovation, our analysis will focus on effects on KIBS firm entry. Results for other industries are presented in Appendix F and Appendix G. We include as outcomes logged firm entry in the overall KIBS industry as well as in four sub-industries in the following analysis.

The upper left panel of Figure 4 displays the estimated coefficients from Equation 2. First, we check that the parallel trend assumption is satisfied. By examining the coefficients in the pre-treatment period, we find little evidence of statistically or economically significant differences in KIBS firm entry between treated and control counties. Post treatment, we observe a significant and robust upward trend in firm creation within the KIBS sector. The robustness of these findings is underscored by the 95% confidence intervals which consistently remain above zero, indicating a statistically significant increase in KIBS firm entry attributable to the opening of new campuses.

/ Insert Figure 4 Here/

We note that there have been concerns regarding the estimation bias induced by heterogeneous treatment effects in the staggered DD design. Goodman-Bacon (2021) noted that the estimation in TWFE models is equivalent to a weighted average of all possible 2×2 DD estimations in the data. The estimates could be biased if treatment effects vary across units or periods. Treating the already treated group as the control group could result in substantial estimation bias. We therefore further adopt several heterogeneity-robust DD estimators including the double-robust inverse probability weighting (DR-IPW) estimator proposed by Callaway and Sant'Anna (2021),

the interaction weighted (IW) estimator proposed by Sun and Abraham (2021) and the stacked estimator proposed by Cengiz et al. (2019)[16]. Figure 4 plots these heterogeneity-robust DD estimators on the effect on local KIBS firm entry. In general, the findings are in line with those of baseline analysis and lend further support to our analysis. See Appendix H.1 for results of Bacon decomposition.

5.3 2SLS Estimation

Using the instrument constructed in Section 3.2, we estimate the causal effect of new campus openings on local firm entry. Table 3 reports results of the 2SLS estimation. The IV specification follows Equation 4. The instrument is sufficiently strong across all specifications (Cragg-Donald Wald F statistic of 5074.724 and Kleibergen-Papp rk Wald F statistic of 33.516). The estimates confirm significant and larger effects of new campuses on firm entry across KIBS industries. Overall, a new campus opening is associated with an increase of 28.5 log points (33.0%) in KIBS firm entry, amounting to about 43 new KIBS firms in the host county. Specifically, the ICT sector exhibits the most substantial response, with a coefficient of 48.6 log points. This magnitude is translated into 12 new ICT firms per campus, contributing to about 28% of total KIBS entrepreneurship increase. Additionally, significant positive effects are observed in the R&D, finance and professional services sectors, with coefficients of 15.7, 37.9 and 8.0 log points, respectively. Given the larger base of R&D sectors, the 15.7 log points increase amounts to about 7 new R&D firms per campus. The magnitude of increase in the finance sector is comparatively lower with about 3 new firms per campus. The professional services sector shows a relatively lower coefficient of 8.0 log points, which translates into 5 new firms per campus. This variation in the impact across different KIBS sectors suggests that the effects of universities are more pronounced in more technology-related

[16]We notice that there are several other heterogeneity-robust DD estimators including estimators proposed by Borusyak et al. (2024), De Chaisemartin and d'Haultfoeuille (2020) and others. We use the estimators discussed in the main text because the alternatives do not support high-dimensional estimations and, therefore, the results produced by those estimators could not be compared to our baseline results.

KIBS industries (T-KIBS) as opposed to professional KIBS industries (P-KIBS) (H. Lee & Miozzo, 2019; Miles et al., 1995). Overall, the coefficients yielded by 2SLS estimations are about two to four times of the baseline effects. We show in Appendix H.2 the DD estimates using the IV sample.

/ Insert Table 3 Here/

5.4 Endogeneity Concerns and Robustness

City-Year Fixed Effects. While we have incorporated province-year fixed effects in the baseline regression, we further control for city-year fixed effects to obtain a more conservative estimate by making comparisons within the same city. This exercise also serves to control for any time-variant economic policies and export advantage that may be brought about by WTO accession at the city level (e.g. increase in export volume). Results in Table 4 Column 1 show that while the effects are slightly reduced, they remain robust after a large reduction in the degrees of freedom. Appendix H.3 further shows that the results for event study are robust to this specification.

Comparison within Treatment Group. Counties with new campuses may differ from control counties in terms of key baseline characteristics, such as economic conditions and demographic profiles. This discrepancy raises endogeneity concerns on omitted variable bias in comparing treatment and control groups. To address this issue, we exclude control counties and make comparisons within the treatment group. In doing so, the causality is built on the randomness in the temporal variation of new campus openings only. Results in Table 4 Column 2 suggest that while the estimated magnitudes are slightly reduced, the effects continue to exhibit statistical significance.

Including Counties Adjacent to Municipal Districts. As discussed in the background and in the IV construction, site selection for new campuses typically follows the criteria of maximizing vacant land and minimizing commute time. As a result, most new campuses are located in suburban counties that are in close proximity to the urban area. We believe these counties share a more similar economic condition given their geographical position in the city. To address potential endogeneity problems stemming from the uneven geographical distribution of counties, we further

exclude from the IV sample the remote counties—those not adjacent to municipal districts—and re-estimate the baseline model. Table 4 Columns 3 reports these estimates. The effects remain significantly positive with small changes in magnitudes.

Effects of Existing Universities. The presence of pre-existing universities, especially 4-year institutions, may confound the effects of new campuses. Before the expansion started in 1999, most campuses of 4-year undergraduate universities were situated in urban areas with more developed economies and limited land available for commercial expansion. These factors may distinguish these counties inherently from the counties hosting new campuses. To rule out the bias induced by these universities, we include an interaction term between the number of 4-year university campuses in a county prior to 1999 and the year fixed effects to capture the impact of these campuses over time. Table 4 Columns 4 reports these estimates. The effects remain significantly positive with small changes in magnitudes.

Co-location of Universities and Industrial Clusters. As SEZs account for the majority of industrial clusters in China, the co-location of SEZs and universities may lead to an overestimation of the impact of universities on local economic development (D. Chen, 2022). SEZs often provide various incentives, such as tax deductions and investment benefits (Lu et al., 2019), which may bias our estimates of the universities' effects upward by promoting entrepreneurship in the host counties. To address this concern, we exclude counties that had not hosted any SEZ prior to 2020 and rerun the baseline estimation. Results in Table 4 Column 5 suggest that universities continue to exert a significant impact on local entrepreneurship, even when compared to regions with similar industrial foundations.

Effects of New Campuses Built Before 1999. In our baseline analysis, we focus on new campuses built after the expansion policy in 1999. However, we note that some campuses may have been built prior to this expansion. We are concerned that these earlier campuses could have an ongoing impact on the host counties and therefore bias our estimates. To address this issue, we conducted a thorough search for campuses constructed between 1995 and 1999 and identified 14

such institutions. We exclude the 14 counties hosting these campuses from our analysis. As shown in Column 6 of Table 4, the results, while slightly reduced compared to the baseline estimates, remain robust. The relatively small number of new campuses built before 1999 compared to post-expansion period provides additional support for our argument that the surge in new campuses is primarily driven by the expansion policy.

Model Choice. Recent debates have cast doubts on using "log-like" transformation such as log(1+Y) and arcsinh(Y) in dealing with zeros in the data. The implication of percentage effect becomes poorly defined when the outcome changes from zero to nonzero after receiving treatment, or in cases where the outcome for individual is small before treatment (Campbell & Mau, 2021; J. Chen & Roth, 2024). In comparison, a Poisson model produces more reasonable estimates in this situation (Cohn et al., 2022). Therefore, we use Poisson regression with fixed effects to run the estimation on local firm entry of different industries. Table 4 Column 7 reports results estimated using Poisson regression. The estimates remain statistically and economically significant.

Synthetic DID. The fact that most counties have never built any new campus since 1999 hints at the endogeneity problem in our study. Counties with favorable conditions for building new campuses may differ from control groups in aspects of land availability, distance from urban area and baseline economic conditions. These confounding factors may bias our estimates of the effect on local entrepreneurship. To address this concern, we further introduce the Synthetic Difference-in-Differences (Synthetic DD) method to identify a comparable control group and period.

Synthetic DD can be viewed as a combination of the synthetic control method and DD method. The Synthetic Control Method (SCM), first introduced by Abadie and Gardeazabal (2003), is limited to the analysis of single treatment group samples. While the DD method is applicable to multi-treatment group estimations, the comparability of treatment and control groups is often challenged. In a recent work, Arkhangelsky et al. (2021) proposed to generate a Synthetic DD method as a combination of SCM with DD method. This approach produces weights by identifying more comparable individuals and pre-treatment periods as shown in Equation 6. These weights

are then used to estimate the average effect of the treatment in a regression. Table 4 Column 8 reports results estimated using synthetic DD approach. The presence of any new campus leads to a 32.8 log points (38.8%) increase in the KIBS entrepreneurship. The results are similar with the baseline estimates using a binary indicator of exposure to any new campus presented in Appendix H.4.

$$\left(\hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\tau}^{\text{sdid}}\right) = \arg \min_{\alpha,\beta,\mu,\tau} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau\right)^2 \hat{\omega}_i \hat{\lambda}_t \tag{6}$$

/ Insert Table 4 Here/

The results of the aforementioned robustness checks consistently indicate that the opening of new campuses significantly increases the entry of KIBS firms. These results corroborate the baseline estimates. See Appendix H for results of additional robustness checks including results from sample restrictions, alternative specifications, a matching strategy and placebo tests.

5.5 Heterogeneity Analysis

We then proceed to identify which type of university drives the knowledge-intensive entrepreneurial activities. China's universities are commonly categorized into 4-year and 3-year institutions. Four-year institutions (universities) typically offer 4-year programs with a curriculum that emphasizes general education and academic research. These institutions possess stronger research capabilities, resulting in greater knowledge spillover potential and the production of more advanced human capital. Three-year institutions (colleges) typically offer 3-year programs and primarily train workers for the manufacturing and non-KIBS industries.

We analyze the heterogeneous effects of new campuses by estimating the dynamic effects of counties building the first new campus of a 4-year or 3-year institution. As shown in Figure 5, the magnitudes of the effect are similar for both 4-year and 3-year institutions. Although our hypothesis was that 4-year universities would have a greater influence on local KIBS entrepreneurship, the results suggest that 3-year colleges also play a notable role, at least on par with 4-year universities.

This finding is consistent with the studies on U.S. community colleges, branch institutions and normal schools (Acton, 2023; Howard et al., 2024; Lehnert et al., 2024), which show significant impacts of 3-year HEIs on local economy. Whereas Y. Wang and Tang (2020) show that 4-year university towns generally have a greater impact on promoting human capital agglomeration than 3-year university towns in China, our results indicate that 3-year colleges have a similar effect on promoting entrepreneurship.

/ Insert Figure 5 Here/

6 Geography of Spillover Effects

The spillover effects are often largely attenuated across space due to the increased cost of interactions and reduced chance of unintended encounters, while the exact extent of spillovers remains less known (Audretsch & Feldman, 2004; Eppelsheimer et al., 2022). To understand the geographical distribution of university spillovers, we further examine the pattern of firm entry around new campuses. Figure 6 presents the estimates obtained from Equation 5. Panel A presents the impact across all industries, while Panel B focuses specifically on KIBS industries. The graph shows the impact of the new campus at different distances. Between the scope of 5-30 kilometers, the opening of a new campus has a positive effect on overall firm creation, with this effect peaking at around a 15-kilometer radius. Notably, the new campus opening leads to a significant increase of approximately 15 log points (16.2%) in the number of new firms at the peak. However, beyond the 30-kilometer threshold, the impact of new campuses on firm entry disappears. These results highlight a tendency for firms to agglomerate in close proximity to the new campus, especially within a 30-kilometer radius. These findings show that the university spillover effect increases with proximity to universities rather than remaining homogeneous across geographical space. Notably, the effects are relatively smaller within a close vicinity of new campuses, exhibiting an inverse U-shaped pattern consistent with the findings of Audretsch et al. (2005) and Calcagnini et al.

(2016).

/ Insert Figure 6 Here/

In terms of heterogeneous effects, we find that the overall agglomeration is driven by the entry of KIBS firms and non-KIBS firms. KIBS firms tend to cluster within a 35-kilometer radius of a new campus, while non-KIBS firms prefer sites in a vicinity of 30-kilometer radius. This can be attributed to the necessity for non-KIBS firms to be situated close to consumers. Meanwhile, we observe a significant negative impact of the new campus on the entry of manufacturing firms within a 40-kilometer radius, with the most substantial effect reaching approximately 60 log points (82.2%). These results indicate that manufacturers are less inclined to locate near the campus, instead favoring locations further away. Within the KIBS sector, although the overall pattern of agglomeration is consistent across different KIBS industries, ICT firms show the most substantial response in agglomerating around the university.

7 Mechanisms of University Spillovers

Now that we have established the impact of university spillovers on local entrepreneurship, we proceed to investigate two channels underlying the effects: knowledge spillover and human capital spillover.

7.1 Knowledge Spillover

In a pioneering work, Jaffe (1989) pointed out that local innovation is facilitated by "$\cdots$ the ideas generated by faculty and the high-quality libraries and other facilities of research universities". As primary institutions for knowledge production, universities have long been tasked with advancing research activities, a responsibility that has been integral to their role since the previous century. Numerous empirical studies have also validated the association between the

presence of universities and knowledge-intensive clusters (Baptista et al., 2011; Woodward et al., 2006). Exposure to new universities encourages firms to engage in more research and development collaborations. Intuitively, firms are drawn to cooperate with new universities because knowledge spillovers are fostered by more frequent in-person interactions (Boschma, 2005; Petruzzelli & Murgia, 2020). These collaborations also increase firms' overall innovation output through direct knowledge spillovers.

Most existing studies have intuitively attributed the growth of local innovation to university-affiliated individuals filing an increased number of patents or joint patenting between universities and local firms without empirically observing the connection between universities and firms[17]. The common method for substantiating the assertion of knowledge spillover from universities is to examine the effects of university funding or publications on local outcomes[18]. Nevertheless, this approach is inadequate for discerning the channels through which university knowledge spillover occurs, whether through direct R&D collaborations or other channels. By linking patent data with alumni and faculty lists, Andrews (2023) first showed that while alumni contribute to approximately 5.3% of all patents in the county, university faculty account for a negligible proportion of all patents. This raises the question of the extent to which these new campuses facilitate local entrepreneurship through direct engagement in R&D activities, or through other indirect channels.

To answer this question, we link the campus data to the patent data described in Section 4.3 and distinguish between the applicants of firms, new-campus-related universities and others. This exercise allows us to identify those patents that were filed by either KIBS companies alone or those that were jointly filed by KIBS companies and new-campus-related universities. We reestimate Equation 1 using logged total patent applications filed by local KIBS firms (total patenting) and logged joint patent applications filed by local KIBS firms with new-campus-related universities (joint patenting) as the dependent variable[19]. The observations span from 1995 to 2015 due to

[17]See Abramovsky et al. (2007), Jaffe (1989) for examples.
[18]See Audretsch et al. (2005) and Tartari and Stern (2021) for examples.
[19]Firms of finance and professional services activities are excluded as they file a very small number of patents compared to firms of the R&D and ICT industries.

restrictions by the patent data.

Table 5 reports the estimated effects of new campuses on patenting activities of KIBS firms. In Column 1, total patenting exhibits a substantial increase of 34.8 log points (41.6%) following a new campus opening. This effect amounts to about an increase of 5 patents filed by KIBS firms. Column 3 shows that joint patenting also increases significantly, albeit to a lesser extent of 8.3 log points (8.7%). The magnitude of indirect knowledge spillover is about four times of the direct knowledge spillover. We present in Appendix H.8 the results of an event study on local innovation using the Callaway and Sant'Anna (2021) estimator. See Appendix H.9 for results of 2SLS estimation on knowledge spillover.

/ Insert Table 5 Here/

Our findings are consistent with the findings in Andrews (2023) that universities mainly increase local innovation output through indirect channels. However, our results support a more prominent role for university faculty in fostering local innovation because our results take into account only current faculty, as opposed to the sample of all alumni as well as current and former faculty included in Andrews (2023). While Andrews (2023) found a negligible number of patents filed by both faculty and alumni, our results show that joint patent applications between university faculty and KIBS firms alone contribute to the growth of local innovation.

When we further break it down by divisions, we find that new campus openings promote innovation output in both R&D and ICT industries, while R&D firms appear to file more joint patents with universities. This is primarily attributable to the fact that R&D firms typically have more extensive and direct connections with universities as the form of the knowledge transferred between R&D firms and universities is predominantly basic scientific principles, as opposed to technological solutions (Fernandes & Ferreira, 2013). While ICT firms also seem to benefit from universities with a significant increase in total patent applications, their preference for independent intellectual property may prevent them from sharing patents with universities. ICT firms in China

have been documented with a low propensity of sharing patents, especially with universities (Gao et al., 2014). In general, the proximity to universities and the possibility of technological collaboration with these institutions are of greater significance to firms in the R&D sector than to other types of KIBS firms, a finding consistent with the results in Fernandes and Ferreira (2013).

7.2 Human Capital Spillover

Having identified the knowledge spillovers from universities, we then turn to the analysis of human capital spillovers. In theory, universities generate human capital spillover through both direct and indirect channels. On the one hand, universities produce skilled labors with a strong preference to work near the university, thus reducing search costs for local companies (Valero & Van Reenen, 2019). Graduates working in the neighborhood are themselves a way of disseminating knowledge from universities to local industries (Audretsch et al., 2005; Calcagnini et al., 2016). On the other hand, knowledge-intensive industries clustering around universities generate demand for more skilled labor and result in higher human capital growth in these regions (Abel & Deitz, 2012). As such, more educated population flock to these regions for working opportunities, higher wages and productivity (Glaeser & Saiz, 2003; Moretti, 2004). This is especially true given the higher mobility of educated individuals than others (Schwartz, 1976). The multiplier effect of job creation is particularly large in human-capital-intensive and technology-intensive industries (Moretti & Thulin, 2013). However, most previous studies on universities' effect have been unable to distinguish between these direct and indirect channels due to data limitations. Thus far, most empirical evidence on the human capital spillover from universities is based on the correlation between the size of graduate pool in a region and entrepreneurial activity[20].

To overcome this problem, we link the population mini-census data described in Section 4.4 to the campus data to identify the direct and indirect human capital spillover generated by new campuses. We leverage the information on individual's county of residence one year and five years

[20]See Audretsch et al. (2005) and Bonaccorsi et al. (2014) for examples.

ago to determine whether an individual graduates from a local institution[21] (direct human capital spillover) or migrates to the host county after graduation (indirect human capital spillover) based on the following assumptions. As China's education system implements 9 years of compulsory education starting at the age of 6 and 3 years of high school education, we assume that normally an individual enters university or college at 18 and graduates from a 4-year institution at 22 or from a 3-year institution at 21. To illustrate, a person aged 24 who has completed a 4-year college degree and resided in the county 5 years ago (when one is supposed to be receiving college education) would be identified as a graduate from local institutions. Additionally, the information on the industry of one's professional status enables us to ascertain whether one is employed in KIBS industries. See Appendix E for more details of identification.

Restricted by the observation period of the mini-census data, we only regress each type of human capital stock on the cumulative number of new campus openings a county has received between 1999 and 2014 (one year before observation). We then include logged night-time light in 1995 and logged population density in 1995 and province fixed effects for a more comparable result to the baseline estimates. Counties with zero college graduates aged 22-30 in the 2015 population census sample are excluded from the analysis[22].

Table 6 Panel A Column 1 reports the human capital spillovers generated by new universities and campuses[23]. The estimates show that universities lead to 9.7 log points (10.2%) increase of human capital measured by the number of college graduates employed in KIBS in host counties by 2015. Columns 2-3 report the results for direct spillovers - measured by college graduates from

[21]Unfortunately, we are unable to determine whether an individual is a graduate of a new campus-affiliated institution due to a lack of information on the specific university or college from which the individual graduated. Assuming that the total enrollment of all other local institutions and the share of graduates working in the same county are independent from the new campus openings, we can attribute the increase in graduates from a local institution to the new-campus-related universities and obtain an upper bound estimate. The fact that the upper bound estimate of direct human capital spillover is lower than the indirect spillover lends further support to our hypothesis.

[22]This exclusion is due to uncertainty regarding the reason for the absence of college graduates in these counties, which could stem from: (1) a genuine scarcity of college graduates aged 22-30 in these counties, or (2) these counties being dropped in the sampling procedure. Assigning zero college graduates to these counties could introduce estimation bias if their absence is primarily due to the latter.

[23]See Section 7.2 for details of identification of each population category in the population mini-census data.

local institutions - and indirect spillovers - measured by migrants with college degrees. Results show that the growth of graduates from local institutions (2.3 log points) is smaller than that of migrants with a college degree (7.0 log points). These effects translate to a roughly 38 increase in graduates from local institutions and 128 increase in high-skilled migrants. The indirect channel appears to dominate the spillovers given a considerably large effect of nearly three times of the direct spillover effect.

We then reestimate the same equation using the human capital stock in each sub-industry of KIBS as the dependent variable in Table 6 Panel B. Aligned with the findings on KIBS firm entry, human capital spillovers from new campuses are the most substantial on the ICT industry, estimated at 6.5 log points (6.7%), translating to an increase of 5 high-skilled workers in the host county by 2015. This is followed by a 1.8 log point (1.8%) increase in R&D, which amounts to a 0.5 increase in high-skilled labor. In comparison, the effects of new campuses on human capital accumulation in finance and professional sectors are less significant. It appears that new campus openings primarily stimulate the growth of high-skilled labor in the T-KIBS sector, driven by the increased demand for skilled workers stemming from the entrepreneurial activity within these industries. See Appendix H.10 for results of 2SLS estimation on human capital spillover.

/ Insert Table 6 Here/

Our results reveal a different channel from the Russell et al. (2024) study of the impact of winning colleges on local educational attainment. Using the "runner-up" strategy, Russell et al. (2024) found that counties that had won colleges historically have larger contemporary shares of skilled workers and workers in human capital-intensive sectors, such as information and professional and business services. While this conclusion is similar to our results, they found that these effects were not primarily driven by migration. Our results, however, show that the effects of universities on local human capital occur primarily through the inflow of the high-skilled workers to the host county. The difference in channels may be attributed to a higher level of human capital

agglomeration caused by reduced migration costs and increased wage premium of human capital in China over the past few decades (Fu & Gabriel, 2012; Glaeser & Lu, 2018; Monaco & Yamarik, 2015; Tombe & Zhu, 2019; Zhou & Zhang, 2021).

8 Conclusion

How do higher education institutions stimulate local entrepreneurship? This is a question of both theoretical and practical importance. We answer this question by leveraging a natural experiment of Chinese higher education expansion since 1999 whereby 930 new campuses were constructed within commuting distances from the old campuses. We first use a difference-in-differences strategy that exploits the geographic and temporal variations in the construction of these new campuses. Findings suggest that the spreading of higher education resources into a county generated a substantial increase in local entrepreneurial activities, particularly in the KIBS sector. We also employ an instrumental variable strategy that leverages the exogenous variations in the push and pull factors that are beyond the control of local governments. In the IV models, the presence of a new campus translated to a 33% increase in local KIBS firm entry (or 43 new firms). Concentric ring analyses further show that such stimulative effects on the service sector of the economy tend to concentrate within a 30-kilometer radius of the new campuses with peaks at the 15-kilometer. These results suggest that the presence of universities contributed to the local clustering of knowledge industries. Heterogeneity results reveal a prominent role of both 4-year universities and 3-year vocational colleges in fostering local entrepreneurship.

Mechanism analyses show that these effects of universities went through both the knowledge spillover and the human capital spillover channels. However, we find that indirect spillovers, instead of direct spillovers, drive the effect. New campuses lead to a substantial increase in firm innovation and an inflow of high-skilled labor, while only a small proportion can be attributed to the increase in direct R&D collaboration between universities and firms, or the expansion of the

local graduate pool. These effects are found primarily in the T-KIBS sector, including R&D and ICT.

To our best knowledge, this is the first study that demonstrates the effects of universities on local entrepreneurship and industrial upgrading in the context of a large developing country. From a policy perspective, our results have implications for developed and developing countries alike. For a long time, developing countries have been keen to bank on strengthened university-industry collaborations to move up the global value chain (Mowery & Sampat, 2005), but their attempts were hampered by small enrollment size, weak research capacity and limited funding prevalent in their higher education systems, just to name a few reasons (Weerasinghe et al., 2024). Our findings suggest that emerging economies or transitional economies that have experienced initial economic take-off but lagged in higher education accessibility and innovation could leverage the expansion of higher education to aid their transformation towards a knowledge-based economy. These findings corroborate recent policy suggestions by Gruber and Johnson (2019) to the U.S. that it should take a leaf out of its own playbook during the Cold War and invest heavily in universities and research institutes across a wide array of geographic areas in the country. The spillover effects from such large-scale and geographically dispersed investments in research and development would not only cement America's leadership role in technology innovation but also ensure a vibrant private sector that benefits the lots, instead of benefiting only those in a few big cites.

References

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, *93*(1), 113–132.

Abel, J. R., & Deitz, R. (2012). Do colleges and universities increase their region's human capital? *Journal of Economic Geography*, *12*(3), 667–691.

Abramovsky, L., Harrison, R., & Simpson, H. (2007). University research and the location of business R&D. *The Economic Journal*, *117*(519), C114–C141.

Acton, R. (2023). *Causes and consequences of the U.S. community college expansion period.* (unpublished)

Aghion, P., & Howitt, P. W. (2008). *The economics of growth*. MIT press.

Akcigit, U., & Ates, S. T. (2023). What happened to US business dynamism? *Journal of Political Economy*, *131*(8), 2059–2124.

Andersson, R., Quigley, J. M., & Wilhelmsson, M. (2009). Urbanization, productivity, and innovation: Evidence from investment in higher education. *Journal of Urban Economics*, *66*(1), 2–15.

Andrews, M. J. (2023). How do institutions of higher education affect local invention? evidence from the establishment of US colleges. *American Economic Journal: Economic Policy*, *15*(2), 1–41.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118.

Asheim, B. (2007). Differentiated knowledge bases and varieties of regional innovation systems. *Innovation*, *20*(3), 223–241.

Atherton, A., & Newman, A. (2017). *Entrepreneurship in China: The emergence of the private sector*. Routledge.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics* (Vol. 4, pp. 2713–2739). Elsevier.

Audretsch, D. B., Lehmann, E. E., & Warning, S. (2005). University spillovers and new firm location. *Research Policy*, *34*(7), 1113–1122.

Baptista, R., Lima, F., & Mendonça, J. (2011). Establishment of higher education institutions and new firm entry. *Research Policy*, *40*(5), 751–760.

Bonaccorsi, A., Colombo, M. G., Guerini, M., & Rossi-Lamastra, C. (2014). The impact of local and external university knowledge on the creation of knowledge-intensive firms: Evidence from the Italian case. *Small Business Economics*, *43*, 261–287.

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, rdae007.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74.

Calcagnini, G., Favaretto, I., Giombini, G., Perugini, F., & Rombaldoni, R. (2016). The role of universities in the location of innovative start-ups. *The Journal of Technology Transfer*, *41*, 670–693.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230.

Campbell, D. L., & Mau, K. (2021). On "trade induced technical change: The impact of Chinese imports on innovation, IT, and productivity". *The Review of Economic Studies*, *88*(5), 2555–2559.

Cantoni, D., & Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, *129*(2), 823–887.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405–1454.

Che, Y., & Zhang, L. (2018). Human capital, technology adoption and firm performance: Impacts of China's higher education expansion in the late 1990s. *The Economic Journal*, *128*(614), 2282–2320.

Chen, D. (2022). *Universities and innovation-driven development: Quasi-experimental evidence from establishing new campuses* (Unpublished doctoral dissertation). Peking University.

Chen, J., & Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, *139*(2), 891–936.

Chen, Z., Liu, Z., Suárez Serrato, J. C., & Xu, D. Y. (2021). Notching R&D investment with corporate income tax cuts in China. *American Economic Review*, *111*(7), 2065–2100.

Cohn, J. B., Liu, Z., & Wardlaw, M. I. (2022). Count (and count-like) data in finance. *Journal of Financial Economics*, *146*(2), 529–551.

Cui, C., & Li, L. S. (2023). Trade policy uncertainty and new firm entry: Evidence from China. *Journal of Development Economics*, *163*, 103093.

Dahlman, C. J., & Aubert, J.-E. (2001). *China and the knowledge economy: Seizing the 21st century*. World Bank Publications.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964–2996.

Djankov, S., Qian, Y., Roland, G., & Zhuravskaya, E. (2006). Who are China's entrepreneurs? *American Economic Review*, *96*(2), 348–352.

Eppelsheimer, J., Jahn, E. J., & Rust, C. (2022). The spatial decay of human capital externalities-a functional regression approach with precise geo-referenced data. *Regional Science and Urban Economics*, *95*, 103785.

Etzkowitz, H., & Leydesdorff, L. (1995). The triple helix–university-industry-government relations: A laboratory for knowledge based economic development. *EASST review*, *14*(1), 14–19.

Feng, Y., Tan, X., & Wang, R. (2022). The value of higher education to entrepreneurial performance: Evidence from higher education expansion in China. *China Economic Review*, *73*, 101789.

Fernandes, C. I., & Ferreira, J. J. (2013). Knowledge spillovers: cooperation between universities and KIBS. *R&D Management*, *43*(5), 461–472.

Fu, Y., & Gabriel, S. A. (2012). Labor migration, human capital agglomeration and regional development in China. *Regional Science and Urban Economics*, *42*(3), 473–484.

Gao, X., Guo, X., & Guan, J. (2014). An analysis of the patenting activities and collaboration among industry-university-research institutes in the Chinese ICT sector. *Scientometrics*, *98*, 247–263.

Glaeser, E. L., Kerr, S. P., & Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, *97*(2), 498–

520.
Glaeser, E. L., & Lu, M. (2018). *Human-capital externalities in China* (Tech. Rep.). National Bureau of Economic Research.
Glaeser, E. L., & Saiz, A. (2003). *The rise of the skilled city.* National Bureau of Economic Research Cambridge, Mass., USA.
Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277.
Gruber, J., & Johnson, S. (2019). *Jump-starting America: How breakthrough science can revive economic growth and the American dream.* Hachette UK.
Hausman, N. (2022). University innovation and local economic growth. *Review of Economics and Statistics*, *104*(4), 718–735.
He, Y., Niu, G., Lu, J., & Zhao, G. (2024). Digital governance and urban entrepreneurial vitality: Evidence from the "National Pilot Policy of Information Benefiting the People" in China. *Journal of Quantitative & Technological Economics*, *41*(01), 47-66.
Howard, G., Weinstein, R., & Yang, Y. (2024). Do universities improve local economic resilience? *Review of Economics and Statistics*, *106*(4), 1129–1145.
Hummels, D., & Yue, K. (2024). *On the ubiquity of declining business dynamism* (Tech. Rep.). National Bureau of Economic Research.
Jaffe, A. B. (1989). Real effects of academic research. *The American Economic Review*, 957–970.
Kang, N. (2000). On educational policy making and institutional innovation - a case study of the 99 college expansion policy. *Journal of Higher Education*(02), 31-38.
Kantor, S., & Whalley, A. (2014). Knowledge spillovers from research universities: evidence from endowment value shocks. *Review of Economics and Statistics*, *96*(1), 171–188.
Lee, C., Lee, C., & Xiao, S. (2021). Policy-related risk and corporate financing behavior: Evidence from China's listed companies. *Economic Modelling*, *94*, 539–547.
Lee, H., & Miozzo, M. (2019). Which types of knowledge-intensive business services firms collaborate with universities for innovation? *Research Policy*, *48*(7), 1633–1646.
Lehnert, P., Dell, M., Backes-Gellner, U., & Bettinger, E. (2024). *The effect of postsecondary educational institutions on local economies: A bird's-eye view.* Universität Zürich, IBW-Institut für Betriebswirtschaftslehre.
Lerner, J., Liu, J., Moscona, J., & Yang, D. Y. (2024). *Appropriate entrepreneurship? The rise of China and the developing world* (Tech. Rep.). National Bureau of Economic Research.
Li, H., Ma, Y., Meng, L., Qiao, X., & Shi, X. (2017). Skill complementarities and returns to higher education: Evidence from college enrollment expansion in China. *China Economic Review*, *46*, 10–26.
Li, J., Liu, S., & Wu, Y. (2020). Identifying knowledge spillovers from universities: quasi-experimental evidence from urban China. *Available at SSRN 3621422*.
Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, *87*, 25–41.
Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy*, *11*(3), 325–360.

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42.

Ma, X. (2024). College expansion, trade, and innovation: Evidence from China. *International Economic Review*, *65*(1), 315–351.

Miles, I., Kastrinos, N., Flanagan, K., Bilderbeek, R., Den Hertog, P., Huntink, W., & Bouman, M. (1995). Knowledge-intensive business services: their role as users, carriers and sources of innovation. *Report to the EC DG XIII Sprint EIMS Programme, Luxembourg*.

Monaco, K., & Yamarik, S. (2015). Are there human capital externalities in US states? evidence from the Current Population Survey. *Economics Bulletin*, *35*(4), 2345–2362.

Moretti, E. (2004). Human capital externalities in cities. In *Handbook of regional and urban economics* (Vol. 4, pp. 2243–2291). Elsevier.

Moretti, E., & Thulin, P. (2013). Local multipliers and human capital in the United States and Sweden. *Industrial and Corporate Change*, *22*(1), 339–362.

Mowery, D., & Sampat, B. (2005). The oxford handbook of innovation. In (chap. Universities in national innovation systems). University Press, Oxford.

Nimier-David, E. (2023). *Local human capital and firm creation: Evidence from the massification of higher education in France.* (unpublished)

Nunn, N., & Qian, N. (2011). The potato's contribution to population and urbanization: evidence from a historical experiment. *The Quarterly Journal of Economics*, *126*(2), 593–650.

OECD. (2006). *Innovation and knowledge-intensive service activities*. OECD Publishing, Paris.

OECD. (2015). *OECD science, technology and industry scoreboard 2015: Innovation for growth and society*. OECD Publishing, Paris.

Oi, J. C. (1995). The role of the local state in China's transitional economy. *The China Quarterly*, *144*, 1132–1149.

Petruzzelli, A. M., & Murgia, G. (2020). University–industry collaborations and international knowledge spillovers: a joint-patent investigation. *The Journal of Technology Transfer*, *45*(4), 958–983.

Qian, Y. (2003). In search of prosperity: Analytic narratives on economic growth. In (chap. How Reform Worked in China). Princeton University Press.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*(5), 1002–1037.

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, *64*(2), 373–389.

Russell, L. C., Yu, L., & Andrews, M. J. (2024). Higher education and local educational attainment: Evidence from the establishment of US colleges. *Review of Economics and Statistics*, *106*(4), 1146–1156.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. Harper.

Schwartz, A. (1976). Migration, age, and education. *Journal of Political Economy*, *84*(4, Part 1), 701–719.

Sekhri, S., Hossain, M. A., & Khosla, P. (2024). Access to colleges, human capital, and empowerment of women. *Journal of Human Resources*, *59*(2), 502–544.

Smart, A., & Hsu, J.-Y. (2004). The Chinese diaspora, foreign investment and economic development in China. *The Review of International Affairs*, *3*(4), 544–566.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Tartari, V., & Stern, S. (2021). *More than an ivory tower: The impact of research institutions on the quantity and quality of entrepreneurship* (Tech. Rep.). National Bureau of Economic Research.

Tian, X., & Xu, J. (2022). Do place-based policies promote local innovation and entrepreneurship? *Review of Finance*, *26*(3), 595–635.

Tombe, T., & Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of China. *American Economic Review*, *109*(5), 1843–1872.

Valero, A., & Van Reenen, J. (2019). The economic impact of universities: Evidence from across the globe. *Economics of Education Review*, *68*, 53–67.

Vu, K., & Vu-Thanh, T.-A. (2022). *Higher education expansion, labor market, and firm productivity in Vietnam.* Working Paper.

Wang, S. (2012). Credit constraints, job mobility, and entrepreneurship: Evidence from a property reform in China. *Review of Economics and Statistics*, *94*(2), 532–551.

Wang, Y., & Tang, W. (2020). Universities and the formation of edge cities: Evidence from China's government-led university town construction. *Papers in Regional Science*, *99*(1), 245–266.

Weerasinghe, R., Jayawardane, A., & Huang, Q. (2024). Critical inquiry on national innovation system: Does NIS fit with developing countries? *Sustainable Technology and Entrepreneurship*, *3*(1), 100052.

Wei, X., Li, H., & Su, Y. (2024). The impact of urban talent policies on new rirm entry: From the perspective of migrants' human capital. *Chinese Journal of Population Science*, *38*(04), 21-35.

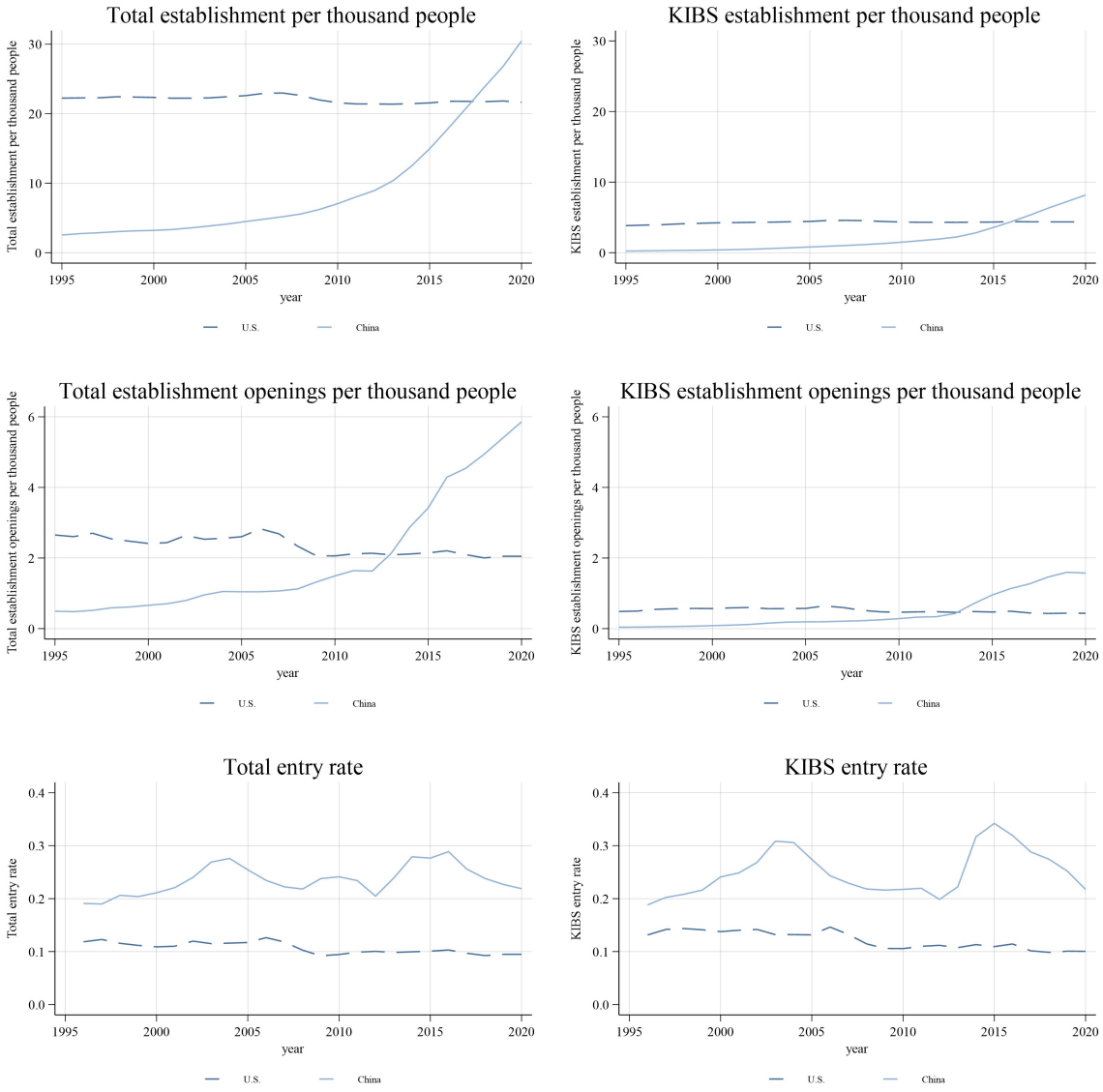
Woodward, D., Figueiredo, O., & Guimaraes, P. (2006). Beyond the Silicon Valley: University R&D and high-technology location. *Journal of Urban Economics*, *60*(1), 15–32.

Zacchia, P. (2020). Knowledge spillovers through networks of scientists. *The Review of Economic Studies*, *87*(4), 1989–2018.

Zhao, Y. (2010). College suburbanization and its influence over urban culture. *Journal of Higher Education (Gaodengjiaoyuyanjiu)*, *31*(08), 64-71.

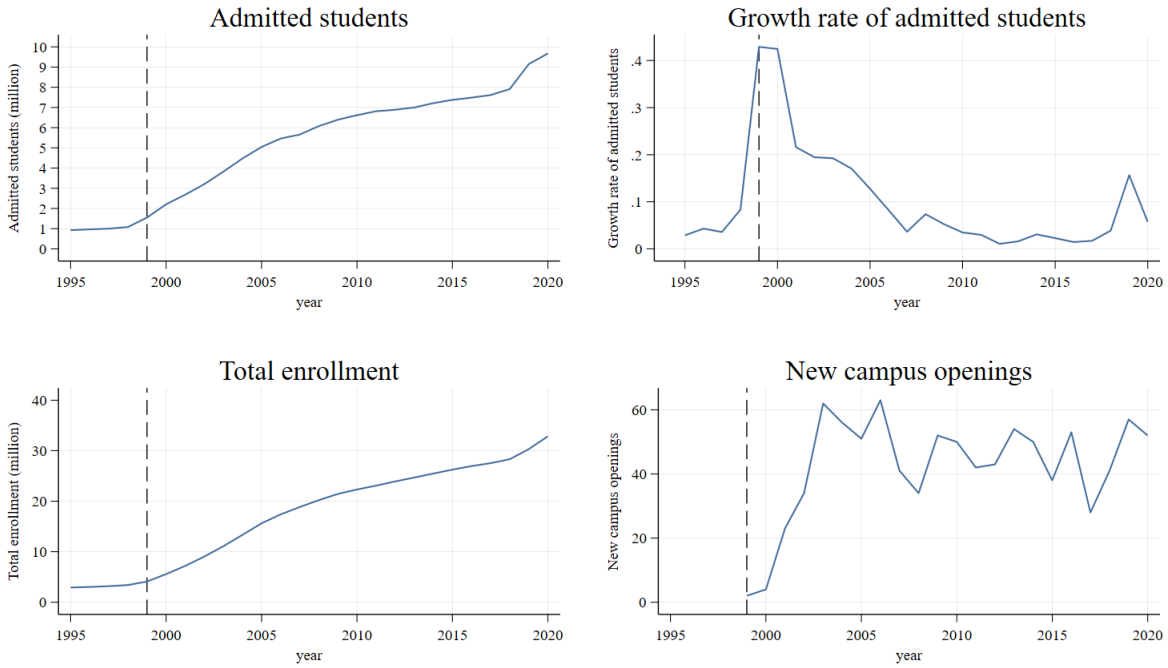
Zhou, K., & Zhang, J. (2021). Trade normalization, export quality, and in-migration of skilled workers: Evidence from China. *Journal of Economic Behavior & Organization*, *184*, 375–387.

Figure 1: **Total Establishments and Openings in the U.S. and China: 1995-2020**



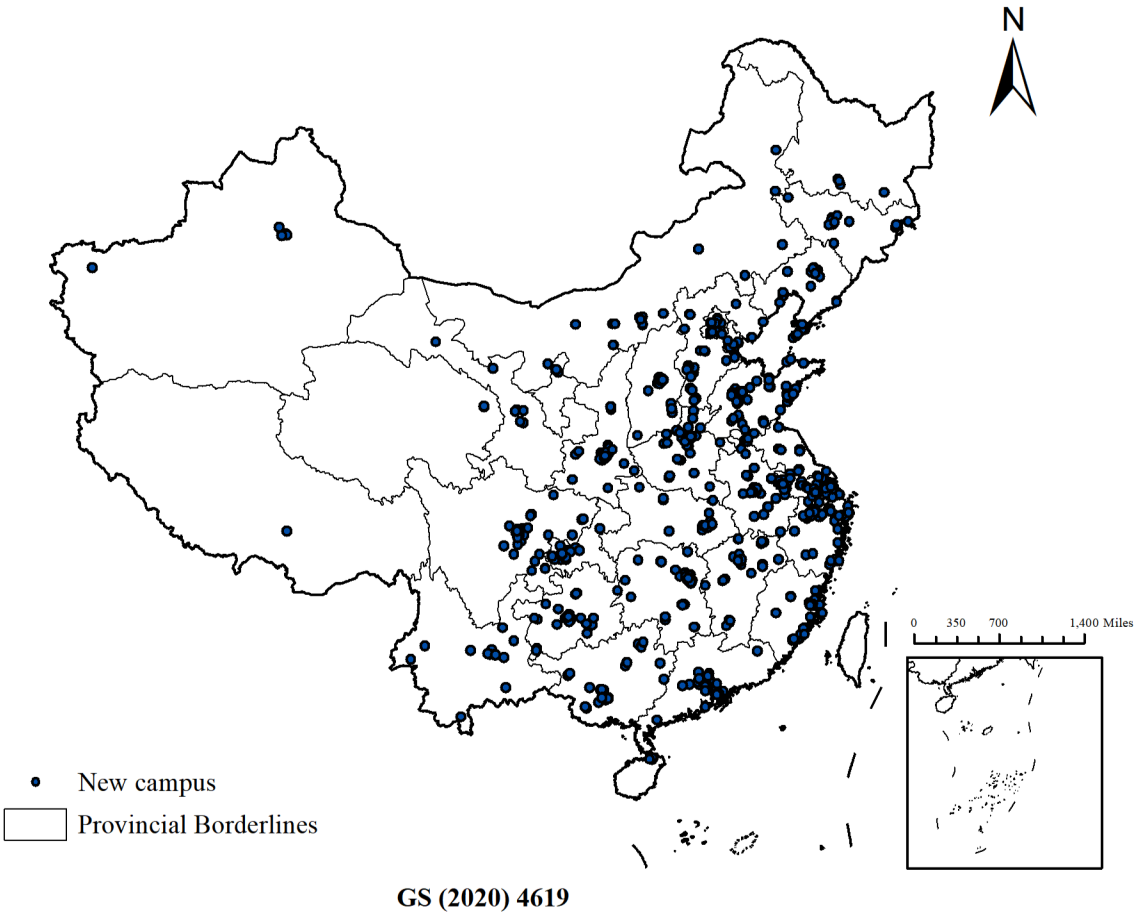
Notes:
1. Figure 1 displays total establishments and openings in the U.S. and China between 1995 and 2020.
2. Entry rate is defined by the ratio of total establishment openings to total establishments at the end of the previous year.
3. Data on total establishments and openings of U.S. are obtained from the U.S. Census Bureau Business Dynamics Statistics (BDS) Program. Data on total establishments and openings of China are obtained from the State Administration for Industry and Commerce of China (SAIC).
4. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.

Figure 2: **Higher Education Expansion and New Campus Openings in China: 1995-2020**



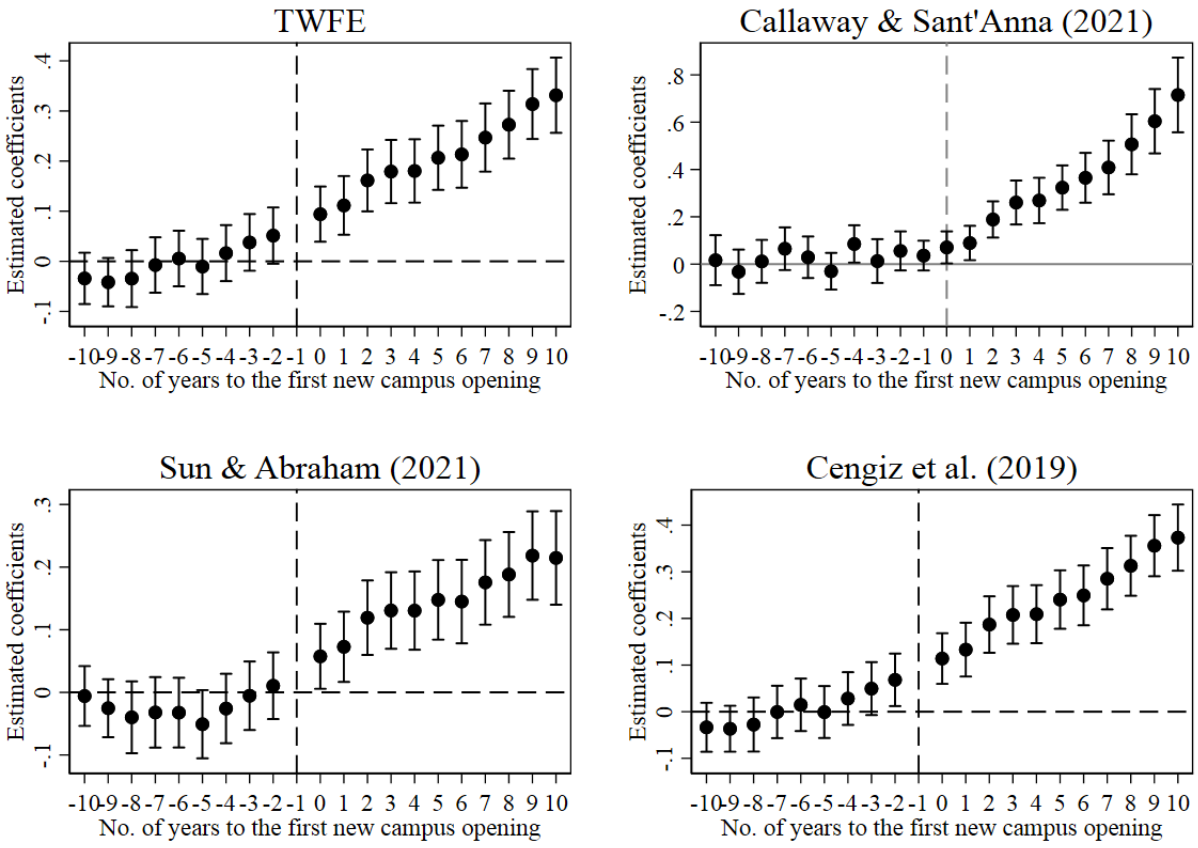
Notes:
1. Figure 2 displays the number of admitted students, the growth rate of admitted students, total enrollment and the number of new campus openings in China between 1995 and 2020. Total enrollment includes only undergraduates in 4-year and 3-year higher education institutions.
2. New campuses include campuses established by 4-year and 3-year higher education institutions.
3. Data on higher education enrollment and admissions are obtained from the Educational Statistics Yearbook of China. Data on new campus openings are collected by the authors.

Figure 3: **Map of New Campus Openings between 1999 and 2020**



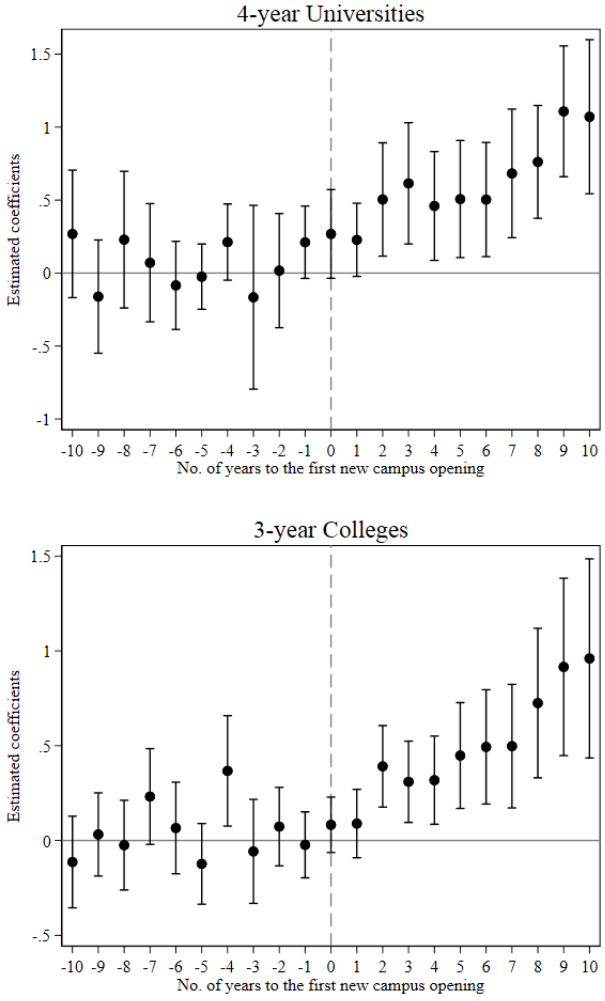
Notes:
1. Figure 3 displays in navy dots the locations of new campuses established between 1999 and 2020.
2. Data on new campus openings are collected by authors.

Figure 4: **Heterogeneity-robust DD Estimators**



Notes:
1. Figure 4 displays in black dots the estimated coefficients of new campuses on local KIBS firm entry using a standard TWFE estimator and three other heterogeneity-robust estimators, including the double-robust inverse probability weighting (DR-IPW) estimator by Callaway & Sant'Anna (2021), the interaction weighted (IW) estimator by Sun & Abraham (2021) and the stacked estimator by Cengiz et al. (2019).
2. The dependent variable is the logarithm of one plus firm entry.
3. Bars represent the estimated 95% confidence intervals.
4. X-axis represents the number of years relative to the first new campus opening in the county since 1999.
5. Each observation is a county-year.
6. All specifications include log night-time light in 1995×Yr, log population density in 1995×Yr, county fixed effects, year fixed effects and province×Yr.
7. Standard errors are clustered at the county level.

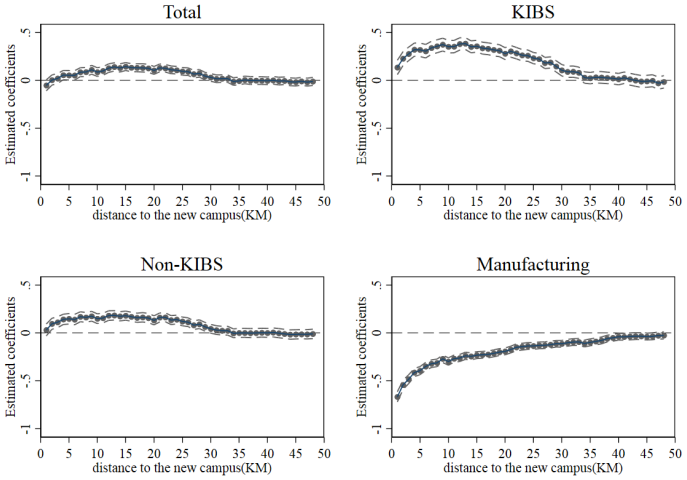
Figure 5: **Heterogeneous Effects of University Spillover by University Type**



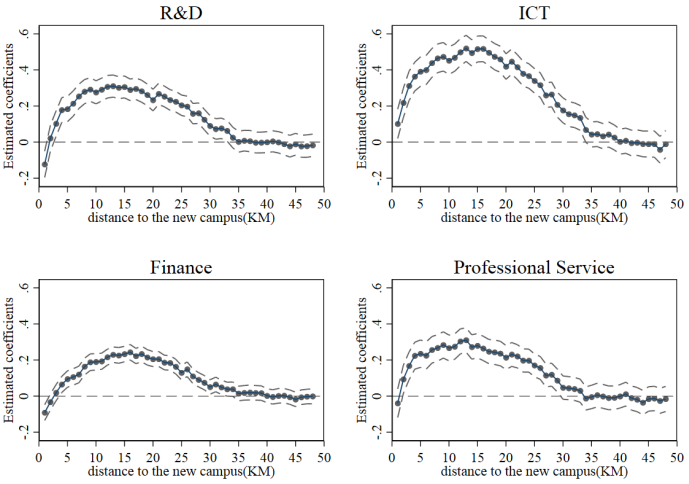
Notes:
1. Figure 5 displays in black dots the estimated coefficients of new campuses on local KIBS firm entry by university type. Dynamic effects are obtained by the Callaway and Sant'Anna (2021) estimator.
2. The dependent variable is the logarithm of one plus firm entry.
3. Bars represent the estimated 95% confidence intervals.
4. X-axis represents the number of years relative to the first new campus opening in the county since 1999.
5. Each observation is a county-year.
6. All specifications include log night-time light in 1995×Yr, log population density in 1995×Yr, county fixed effects, year fixed effects and province×Yr.
7. Standard errors are clustered at the county level.

Figure 6: **Geographical Distribution of University Spillover Effects on Local KIBS Firm Entry**

Panel A: All Industries



Panel B: KIBS Industries



Notes: Figure 6 displays in grey dots the estimated coefficients of new campuses on local KIBS firm entry. The dependent variable is the logarithm of one plus firm entry. The key independent variable is an indicator equal to one if the ring is exposed to the opening of new campus. Gray dashed lines represent the estimated 95% confidence intervals. Each observation is a campus-ring-year. All specifications include campus-ring fixed effects and campus-year fixed effects. Standard errors are clustered at the campus-ring level. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.

Table 1: **Descriptive Statistics**

	Obs	Mean	Std. Dev.	Rounded Min	Rounded Max
Panel A: County-Year Observations					
Dependent Variables					
Total Firm Entry	67,420	984.147	2,906.578	1	128,989
KIBS Firm Entry	67,420	239.678	1,136.338	0	60,886
R&D Firm Entry	67,420	76.810	447.533	0	25,264
ICT Firm Entry	67,420	35.174	182.750	0	10,130
Finance Firm Entry	67,420	8.454	20.748	0	2,228
Professional services Firm Entry	67,420	119.240	604.610	0	59,406
Non-KIBS Firm Entry	67,420	502.998	1,598.306	0	92,836
Manufacturing Firm Entry	67,420	102.039	239.447	0	9,105
Patent Applications by KIBS Firms	53,915	18.329	193.224	0	17,882
Joint Patent Applications by KIBS Firms and Local Universities	53,915	0.023	0.518	0	39
Independent Variables					
Number of New Campus	67,420	0.138	0.722	0	15
Any 4-Year University New Campus	67,420	0.043	0.202	0	1
Any 3-Year College New Campus	67,420	0.037	0.190	0	1
Panel B: County Observations					
Dependent Variables					
Total College Graduates in 2015 (1‰ of population, age 22-30)	2,499	14.009	16.394	0	262
Total College Graduates from Local Institutions in 2015	2,499	5.486	5.508	0	63
Total Migrants with College Degree in 2015	2,499	8.367	11.943	0	199
KIBS College Graduates in 2015 (1‰ of population, age 22-30)	2,499	1.882	3.747	0	89
KIBS College Graduates from Local Institutions in 2015	2,499	0.692	1.205	0	22
KIBS Migrants with College Degree in 2015	2,499	0.878	2.863	0	67
Independent Variables					
Number of New Campus	2,499	0.241	0.929	0	11
Night-time Light (Nit) in 1995	2,701	5.319	10.305	0.000	63.000
Population Density (people per km^2) in 1995	2,701	536.382	846.804	0.158	9686.108
Panel C: Campus-Ring-Year Observations					
Dependent Variables					
Total Firm Entry	1,158,117	593.670	1,557.296	0	91,032
KIBS Firm Entry	1,158,117	190.360	659.716	0	82,715
R&D Firm Entry	1,158,117	68.465	241.940	0	9,867
ICT Firm Entry	1,158,117	27.220	101.258	0	5,509
Finance Firm Entry	1,158,117	3.483	9.926	0	1,916
Professional services Firm Entry	1,158,117	91.193	409.496	0	82,006
Non-KIBS Firm Entry	1,158,117	307.861	855.608	0	79,552
Manufacturing Firm Entry	1,158,117	48.752	112.238	0	3,139
Independent Variables					
New Campus Opening	1,158,117	0.372	0.483	0	1

Notes: Table 1 reports descriptive statistics. All county-year and campus-ring-year observations span from 1995 to 2020 except for patent applications. Data on patent applications is only available between 1995 and 2015. Total firm entry includes KIBS, non-KIBS, manufacturing and other industries in the first and the secondary industry. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport). Firms of finance and professional services activities are excluded from patent analysis as they file a very small number of patents compared to firms of the R&D and ICT industries. Local universities refer to universities and colleges to which local new campuses are affiliated. Counties with zero college graduates aged 22-30 in the 2015 population census sample are excluded from the analysis.

Table 2: **New Campus and Local Firm Entry**

	All Industries				KIBS Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	KIBS	Non-KIBS	Manufacturing	R&D	ICT	Finance	Professional Services
L.Number of New Campuses	0.065***	0.113***	0.067***	0.002	0.100***	0.179***	0.090***	0.088***
	(0.009)	(0.011)	(0.010)	(0.011)	(0.011)	(0.016)	(0.011)	(0.010)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County-Year Observations	67420	67420	67420	67420	67420	67420	67420	67420
Adjusted R-squared	0.936	0.922	0.920	0.888	0.919	0.828	0.571	0.935
Rounded Baseline Mean of Control Group	729.735	157.287	370.707	81.426	46.650	22.829	6.929	80.879
Number of Counties	2701	2701	2701	2701	2701	2701	2701	2701
Start Year	1995	1995	1995	1995	1995	1995	1995	1995
End Year	2020	2020	2020	2020	2020	2020	2020	2020

Notes:
1. Table 2 reports DD estimates of the effects of new campus openings on local firm entry.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Total firm entry includes KIBS, non-KIBS, manufacturing and other industries in the first and secondary industry.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport).
8. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: **New Campus and Local Firm Entry: 2SLS Estimates**

	(1) KIBS	(2) R&D	(3) ICT	(4) Finance	(5) Professional Services
L.Number of New Campuses	0.285***	0.157***	0.486***	0.379***	0.080**
	(0.047)	(0.039)	(0.074)	(0.063)	(0.034)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES
First Stage Cragg-Donald Wald F stat	5074.724	5074.724	5074.724	5074.724	5074.724
First Stage Kleibergen-Papp rk Wald F stat	33.516	33.516	33.516	33.516	33.516
County-Year Observations	38076	38076	38076	38076	38076
Adjusted R-squared	-0.046	-0.023	-0.052	-0.046	-0.016
Rounded Baseline Mean of Control Group	130.367	40.858	19.074	6.613	63.822
Number of Counties	1814	1814	1814	1814	1814
Start Year	2000	2000	2000	2000	2000
End Year	2020	2020	2020	2020	2020

Notes:
1. Table 3 reports 2SLS estimates of the effects of new campus openings on local firm entry. See main text Section 3.2 for details of constructing IV.
2. Central counties are excluded from the regression. See Appendix H.2 for DD estimates using IV sample.
3. The dependent variable is the logarithm of one plus firm entry.
4. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
5. Each observation is a county-year.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: **New Campus and Local KIBS Firm Entry: Robustness Checks**

	(1) City-Yr FE	(2) Excluding Control Counties	(3) Including Counties Adjacent to Municipal Districts	(4) Number of 4-year University Campuses before 1999×Yr	(5) Excluding Counties with No SEZs	(6) Excluding Counties Treated Before 1999	(7) Poisson	(8) Synthetic DD
L.Number of New Campuses	0.094***	0.071***	0.098***	0.113***	0.107***	0.096***	0.071***	
	(0.011)	(0.010)	(0.015)	(0.011)	(0.012)	(0.011)	(0.021)	
L.Any New Campus								0.328***
								(0.022)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	NO
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	NO
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	NO
County-Year Observations	66561	10197	18674	67420	43288	66211	67420	67300
Adjusted R-squared	0.934	0.953	0.928	0.922	0.925	0.933		
Rounded Baseline Mean of Control Group	158.136	202.541	136.317	157.287	188.376	156.481	157.287	157.518
Number of Counties	2667	408	890	2701	1732	2653	2701	2692
Start Year	1995	1995	2000	1995	1995	1995	1995	1995
End Year	2020	2020	2020	2020	2020	2020	2020	2020

Notes:
1. Table 4 reports robust DD estimates of the effects of new campus openings on local firm entry.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. SEZs refer to Special Economic Zones.
5. Each observation is a county-year.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors are in the parentheses. Standard errors are obtained by 100 bootstraps in column (8) and clustered at the county level in columns (1)-(7). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: **New Campus and Local Innovation of KIBS Firms**

	Total Patenting			Joint Patenting		
	(1) All	(2) R&D	(3) ICT	(4) All	(5) R&D	(6) ICT
L.Number of New Campuses	0.348***	0.331***	0.290***	0.083***	0.077***	0.010***
	(0.035)	(0.035)	(0.030)	(0.014)	(0.013)	(0.003)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES
County-Year Observations	53915	53915	53915	53915	53915	53915
Adjusted R-squared	0.803	0.763	0.671	0.342	0.331	0.127
Rounded Baseline Mean of Control Group	12.166	6.466	2.004	0.001	0.001	0.000
Number of Counties	2701	2701	2701	2701	2701	2701
Start Year	1995	1995	1995	1995	1995	1995
End Year	2015	2015	2015	2015	2015	2015

Notes:
1. Table 5 reports DD estimates of the effects of new campus openings on local innovation. Total patenting includes all patent applications filed by KIBS firms. Joint patenting includes patent applications filed by KIBS firms with local new-campus-related universities.
2. The dependent variable is the logarithm of one plus patent applications.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Firms of finance and professional services activities are excluded as they file a very small number of patents compared to firms of the R&D and ICT industries.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: **New Campus and Local Human Capital in KIBS**

Panel A: Direct and Indirect Channels of Human Capital Spillover to KIBS

	(1) Log College Graduates	(2) Log College Graduates from Local Institutes	(3) Log Migrants with College Degree
L.Number of New Campuses	0.097***	0.023**	0.070***
	(0.018)	(0.010)	(0.020)
Log 1995 Night-time Light	YES	YES	YES
Log 1995 Population Density	YES	YES	YES
Prov FE	YES	YES	YES
County Observations	2499	2499	2499
Adjusted R-squared	0.388	0.272	0.356
Rounded Baseline Mean of Control Group	2.751	1.630	1.762

Panel B: Sub-industries of KIBS

	(1) R&D	(2) ICT	(3) Finance	(4) Professional Services
L.Number of New Campuses	0.018**	0.065***	0.026*	0.013
	(0.008)	(0.015)	(0.014)	(0.012)
Log 1995 Night-time Light	YES	YES	YES	YES
Log 1995 Population Density	YES	YES	YES	YES
Prov FE	YES	YES	YES	YES
County Observations	2499	2499	2499	2499
Adjusted R-squared	0.198	0.302	0.249	0.280
Rounded Baseline Mean of Control Group	0.026	0.073	0.114	0.062

Notes:
1. Table 6 reports DD estimates of the effects of new campus openings on local human capital.
2. The dependent variable is the logarithm of population in each category. Data on population of each group is drawn from the 10% representative sample from the 1% population survey conducted by the National Bureau of Statistics of China in 2015.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and 2014.
4. Each observation is a county in 2015.
5. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
6. Counties with zero college graduates aged 22-30 in the 2015 population census sample are excluded from the analysis. This exclusion is due to uncertainty regarding the reason for the absence of college graduates in these counties, which could stem from: (1) a genuine scarcity of college graduates aged 22-30 in these counties, or (2) these counties being dropped in the sampling procedure. Assigning zero college graduates to these counties could introduce estimation bias if their absence is primarily due to the latter.
7. Standard errors in the parentheses are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix

A Higher Education Enrollment by Province

Table A1: **Higher Education Enrollment by Province in 1998**

Province	Enrollment in 1998 (thousand)
Anhui	105
Beijing	213
Chongqing	83
Fujian	85
Gansu	54
Guangdong	185
Guangxi	77
Guizhou	43
Hainan	14
Hebei	144
Heilongjiang	125
Henan	146
Hubei	210
Hunan	157
Inner Mongolia	42
Jiangsu	266
Jiangxi	94
Jilin	118
Liaoning	199
Ningxia	11
Qinghai	9
Shaanxi	149
Shandong	181
Shanghai	165
Shanxi	76
Sichuan	152
Tianjin	79
Tibet	3
Xinjiang	47
Yunnan	62
Zhejiang	113

Notes: Table A1 presents the higher education enrollment (in thousands) of 31 provinces in 1998. Enrollment includes only undergraduates in 4-year and 3-year higher education institutions. Data are obtained from the Educational Statistics Yearbook of China.

B Additional Illustration of the Instrument Variable

B.1 An Example

To illustrate the construction of the instrumental variable discussed in Section 3.2 of the main text, we use Fuzhou (the capital of Fujian province) as an example. Fujian province is situated in southeastern China, contributing 4.4% to the national GDP in 2022. In 1999, there were 46,151 undergraduate students enrolled in the total of 29 Fuzhou's universities and colleges, accounting for about 54% of the total enrollment in Fujian. Before the expansion, Fuzhou housed 36 campuses including 19 affiliated with 4-year universities and 17 affiliated with 3-year colleges. Most of these old campuses were located in the central districts of *Cangshan*(15) and *Gulou*(9), with other campuses scattered in other counties and districts.

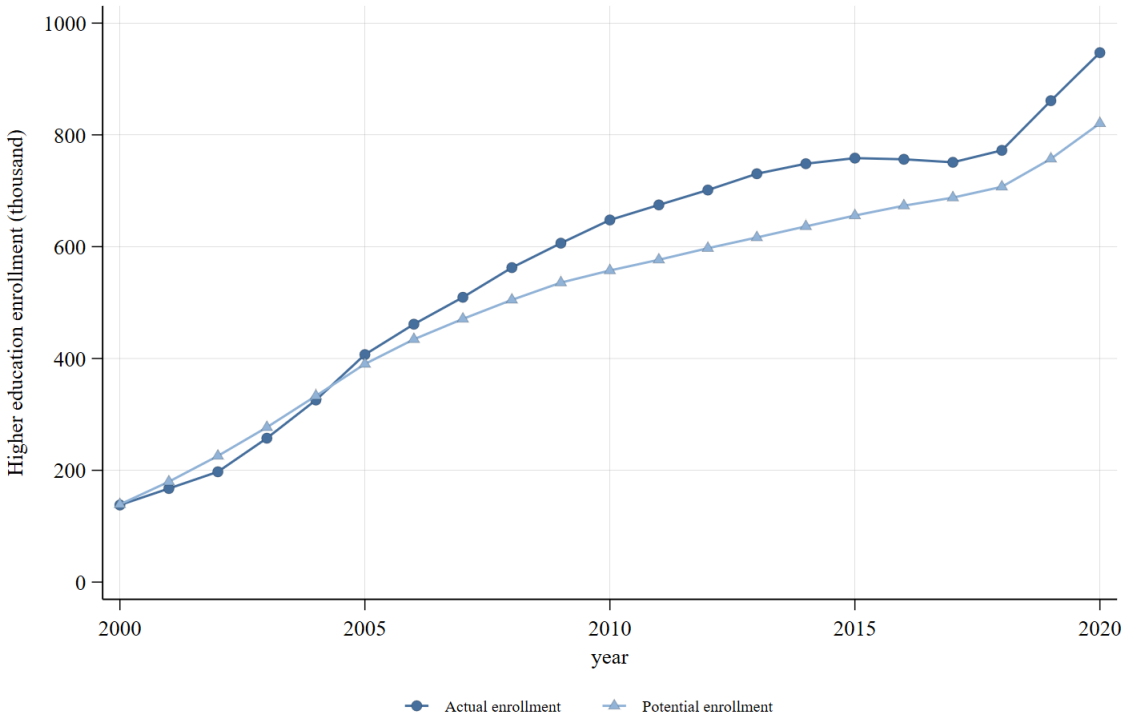
We first present one of the push factors behind the new campus openings in Fuzhou - the province-level potential higher education enrollment $Expansion_{p,t}$. Figure A1 shows both the actual and potential higher education enrollment of Fujian province from 2000 to 2020. The figure shows that the actual and potential higher education enrollment of Fujian province are positively correlated. The actual enrollment increased from 138 in 2000 to 947 thousand in 2020, surpassing the potential enrollment since 2005. While this trend shows that Fujian province was more ambitious in expanding its higher education than the average level, the use of potential enrollment in constructing the IV excludes this factor and incorporates only the expansion in response to the national expansion policy. Furthermore, Fujian's enrollment figure of 85 thousand in 1998 fell below the national average of 110 thousand, underscoring that pre-expansion exposure to the higher education expansion policy was not determined by regional economic conditions.

During the expansion, there were two ways to assign the enrollment quota to universities and colleges. First, universities directly governed by the Ministry of Education (MOE) received quotas from the MOE. In this case, province-, city- and county-level government were takers of the expansion in the universities and had no influence on the quota distribution. Second, the

enrollment quotas of universities and colleges governed by province-level Department of Education (DOE) were assigned by the DOE. The quotas were first assigned to Fujian province by the MOE, which were then distributed by the provincial DOE to each university or college based on its existing infrastructure and capacity. Since all universities and colleges in Fuzhou are governed by the provincial government, this process does not involve any decision-making at the city or county level. As a result, host cities or counties were takers of the effects stemming from the enrollment expansion of the universities they hosted. Both cities and counties had no influence over the province-level enrollment and it is highly unlikely that their economic conditions were considered in the bureaucratic decision-making process for the allocation of enrollment quotas. The higher the expansion target of higher education determined at the provincial level, the more students universities and colleges must accommodate. This situation increases the demand for new campus construction.

To accommodate such an influx of students, Fuzhou began to build new campuses shortly after the expansion policy. In 2003, 4 new campuses were built in *Minhou* county, the closest county to *Cangshan* and *Gulou* the cluster of old campuses in Fuzhou. Another 9 new campuses were then built during 2007 and 2010, with 7 of them also hosted by *Minhou*. Overall, 23 new campuses were built in *Minhou*(14), *Cangshan*(2), *Changle*(2), *Lianjiang*(2), *Pingtan*(1), *Gulou*(1) and *Yongtai*(1) county between 2000 and 2020. With this dynamic showing the exogenous influence of provincial policies on local new campus openings, we include the expansion rate as one of the push factors driving new campus openings in our construction of IV.

Figure A1: **Actual and Potential Higher Education Enrollment of Fujian Province: 2000-2020**



Notes: Figure A1 plots the actual and potential higher education enrollment of Fujian province. The potential enrollment is calculated by the multiplication of national level enrollment growth rate and Fujian's enrollment in 1998. Data are obtained from the Educational Statistics Yearbook of China.

We then present two distance-based components of the instrument, including a pull factor and a push factor. Figure A2 plots the geographic distribution of new campuses built between 1999 and 2020, alongside pre-expansion old campuses, central areas (municipal districts), peripheral areas (counties) and the city center of Fuzhou. The figure aligns with our earlier argument, showing that while most old campuses are concentrated in central areas, the majority of new campuses have been established in peripheral areas, albeit not far from the old campuses. Most new campuses are built in the nearest county *Minhou* from old campuses in peripheral areas, with a few exceptions in *Changle*, *Lianjiang*, *Yongtai*, *Cangshan*, *Gulou* and *Pingtan*.

The solid bidirectional arrow represents the source of the other push factor: the geographical distance between an old campus j and the city center c (with c representing Fuzhou in this case). Due to rapid urbanization in central areas, universities face greater constraints the closer they are to the city center. Consequently, the farther the old campus is situated from the city center, the less incentive a university has to expand its campus space. Since this distance is unaffected by any post-expansion factors, we argue that this measure provides sufficient exogeneity in predicting the number of new campus openings. Given the negative correlation between this measure and campus openings, we use its reciprocal form in constructing the IV.

The dashed bidirectional arrow represents the pull factor: the geographical distance between an old campus j and county i. Since faculty, staff and students frequently travel between old and new campuses, universities prefer locations with shorter commuting distances to minimize transportation and managerial costs. As a result, counties like *Minhou*, *Changle*, *Jin'an* and *Lianjiang*, which are closest to the old campuses—particularly *Minhou*—host the majority of new campuses. Thus, we also expect a negative correlation between this distance and the number of new campus openings in a given county. Accordingly, we incorporate the reciprocal of this measure into the IV.

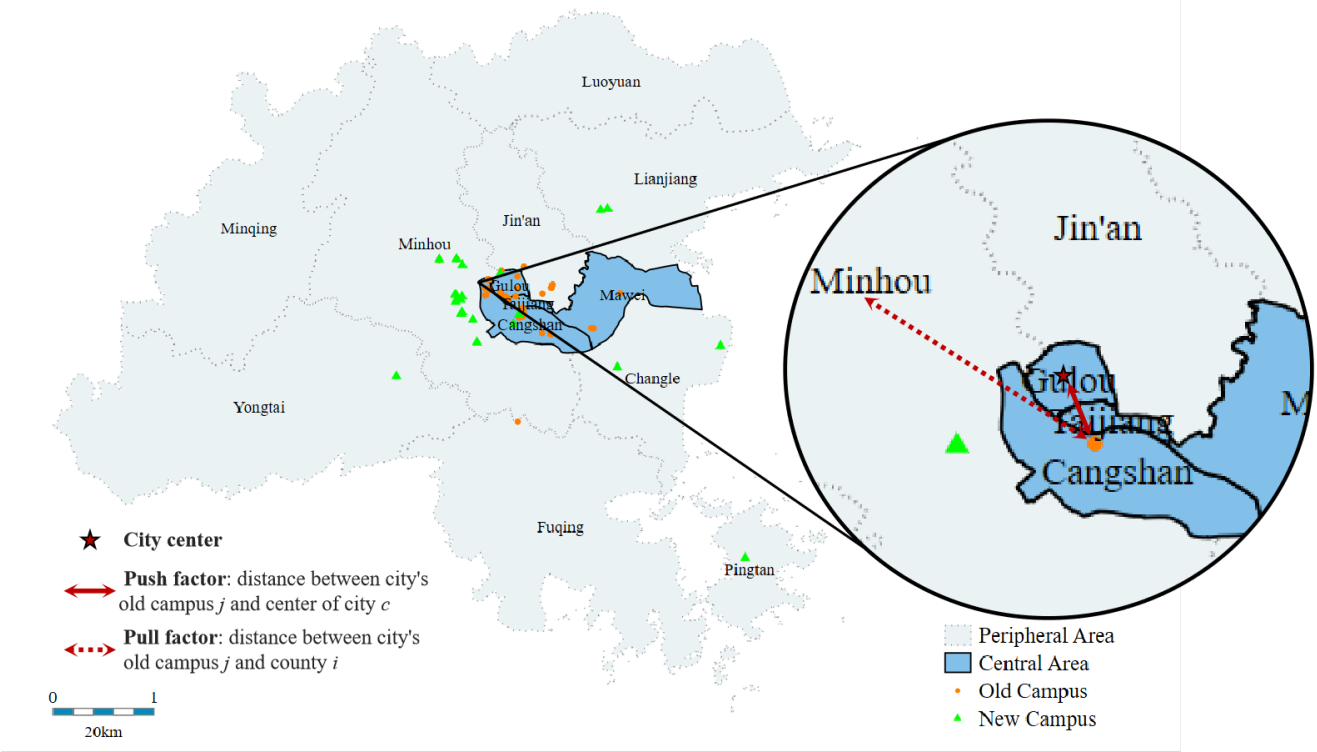
We use the *Qishan* campus of Fujian Normal University (FJNU) as an example. FJNU is a normal university offering mainly 4-year degree programs. Before the expansion, FJNU had only

the *Cangshan* campus situated in *Cangshan* district, one of the municipal districts of Fuzhou. In 1999, about 9,700 undergraduate students were enrolled in FJNU. Against the pressure of expansion, undergraduate enrollment at FJNU increased by approximately 30% by 2001. The influx of students resulted in a sharp decline in campus space per student from 50 m^2 in 1999 to 38 m^2 in 2001[24]. Therefore, it was imperative that FJNU increased its campus space, either by expanding its old facilities or by constructing a new campus. However, with the population density in *Cangshan* district reaching as high as 8,049 people per km^2 in 2000, there was scarce land available for FJNU to construct a new campus in the same district[25]. Therefore, FJNU opted for the neighboring county *Minhou* and built its *Qishan* campus in 2003. While the old *Cangshan* campus covers only 725 acres, the new *Qishan* campus covers approximately 2,800 acres, significantly increasing campus space. The *Qishan* campus is situated approximately 10 kilometers (a 25 minutes drive) from *Cangshan* campus, maintaining a close proximity to the old campus. With the opening of the new campus, the campus space per student of FJNU increased to approximately 95 m^2 in 2004.

[24]The campus space per student is calculated based on the undergraduate enrollment and overall campus space of FJNU.

[25]http://tjj.fuzhou.gov.cn/zz/zwgk/tjzl/ndbg/202105/t20210524_4105019.htm.

Figure A2: **Graphic Illustration of IV: New and Old Campuses in Fuzhou**



Notes: Figure A2 plots the new campus, old campus, city center, central area and peripheral area in Fuzhou city. Fuzhou city is the capital of Fujian province. The star represents the city center c (c =*Fuzhou*). Solid arrows represent the distance between the city's old campus j (j = *FJNU Cangshan*) and the city center. Dashed arrows represent the distance between the city's old campus j (j = *FJNU Cangshan*) and county i (i =*Minhou*).

B.2 Distance to Old Campuses and County's Characteristics

While we have excluded municipal districts from the IV sample to ensure a comparison among the suburban counties, there may be concerns that a county's distance from old campuses are highly correlated to the distance from the city center. This confounding factor could lead to a biased 2SLS estimation as counties situated closer to the city center may be economically advantaged through receiving the spillovers from urban areas (e.g. infrastructure, industrial investment). To address this concern, we first regress log average distance to the city's old campuses on the county's growth

rate of total firm entry, night-time light and population density during pre-expansion periods (1990-1995) on the IV sample. Table A2 presents the results. The coefficients of growth rate of economic and demographic characteristics are all insignificant, suggesting that the distance to old campuses is plausibly exogenous to local economic development, at least on the IV sample. This result lends further support to including the distance to old campuses to our construction of IV.

Table A2: **Distance to Old Campuses and County's Characteristics**

	(1)
Growth Rate of Firm Entry	-0.007
	(0.005)
Growth Rate of Night-time Light	0.016
	(0.014)
Growth Rate of Population Density	-0.067
	(0.046)
City FE	YES
County Observations	1731
Adjusted R-squared	0.310

Notes:
1. Table A2 reports OLS results of pre-expansion growth rate of economic and demographic characteristics on the distance to old campuses.
2. The dependent variable is the logarithm of average distance to each old campus in the city.
3. Growth rate is calculated over the period of 1990-1995 for firm entry and population density and 1992-1995 for night-time light due to data limits.
4. Each observation is a county.
5. Standard errors in the parentheses are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.3 2SLS Estimates on Counties Adjacent to Municipal Districts

To ensure a more balanced distribution of geographical proximity to the city center between the treatment and control counties, we further restrict the sample to suburban counties adjacent to municipal districts (adjacent sample). In the example of Fuzhou city, this exercise removes the remote counties *Minqing*, *Yongtai*, *Luoyuan*, *Fuqing* and *Pingtan* from the sample and leaves only *Minhou*, *Jin'an*, *Lianjiang* and *Changle* in the sample. We believe these counties share a similar weights in terms of their geographical proximity to the city center. In this case, the variations in the instrument are more accrued to the push factor, i.e. the potential province-level expansion and the geographical distance between the old campuses to the city center.

Table A3 reports the 2SLS estimates on the adjacent sample. With a smaller variation in the pull factor, the first stage F stats remain large enough with a smaller magnitude (Cragg-Donald Wald F statistic of 2347.475 and Kleibergen-Papp rk Wald F statistic of 20.404) compared to those in the estimates using the IV sample. The results show that while the number of counties has been reduced from 1,814 to 890, the magnitudes of effects have remained largely consistent with our baseline estimates. The 25.8 log points increase in KIBS entrepreneurship is translated to about 40 new firms in the county. About half of the effects can be attributed to the increase of new firms in the T-KIBS sector, with 5 new firms in R&D and 12 new firms in ICT. The results are in line with the baseline findings.

Table A3: **New Campus and Local Firm Entry: 2SLS Estimates on Adjacent Sample**

	(1) KIBS	(2) R&D	(3) ICT	(4) Finance	(5) Professional Services
L.Number of New Campuses	0.258***	0.126***	0.426***	0.335***	0.061
	(0.056)	(0.047)	(0.085)	(0.070)	(0.041)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES
First Stage Cragg-Donald Wald F stat	2347.475	2347.475	2347.475	2347.475	2347.475
First Stage Kleibergen-Papp rk Wald F stat	20.404	20.404	20.404	20.404	20.404
County-Year Observations	18674	18674	18674	18674	18674
Adjusted R-squared	-0.059	-0.038	-0.068	-0.068	-0.031
Rounded Baseline Mean of Control Group	136.317	39.689	21.843	6.823	67.963
Number of Counties	890	890	890	890	890
Start Year	2000	2000	2000	2000	2000
End Year	2020	2020	2020	2020	2020

Notes:
1. Table A3 reports 2SLS estimates of the effects of new campus openings on local firm entry using only counties adjacent to municipal districts. See main text Section 3.2 for details of constructing IV.
2. Central counties and remote counties are excluded from the regression.
3. The dependent variable is the logarithm of one plus firm entry.
4. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
5. Each observation is a county-year.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C CICS and NAICS

Table A4: **CICS and NAICS**

Sector	NAICS Industry	CICS Industry	Classification
First Industry	Agriculture, forestry, fishing and hunting	Agriculture, forestry, fishing and hunting	Others
Second Industry	Mining, quarrying and oil and gas extraction	Mining, quarrying and oil and gas extraction	Others
Second Industry	Utilities	Utilities	Others
Second Industry	Construction	Construction	Others
Second Industry	Manufacturing	Manufacturing	Manufacturing
Tertiary Industry	Wholesale trade	Wholesale trade	Non-KIBS
Tertiary Industry	Retail trade	Retail trade	Non-KIBS
Tertiary Industry	Transportation and warehousing	Transportation and warehousing	Non-KIBS
Tertiary Industry	Information	Information	KIBS
Tertiary Industry	Finance and insurance	Finance and insurance	KIBS
Tertiary Industry	Real estate and rental and leasing	Real estate and rental and leasing	Non-KIBS
Tertiary Industry	Professional, scientific and technical services	Professional services (including Management of companies and enterprises); Scientific and technical services	KIBS
Tertiary Industry	Management of companies and enterprises		KIBS
Tertiary Industry	Administrative and support and waste management and remediation services	Administrative and support and waste management and remediation services	Non-KIBS
Tertiary Industry	Educational services	Educational services	Non-KIBS
Tertiary Industry	Health care and social assistance	Health care and social assistance	Non-KIBS
Tertiary Industry	Arts, entertainment and recreation	Arts, entertainment and recreation	Non-KIBS
Tertiary Industry	Accommodation and food services	Accommodation and food services	Non-KIBS
Tertiary Industry	Other services (except public administration)	Other services (except public administration)	Non-KIBS

Notes: CICS is China Industry Classification Standard and NAICS is North American Industry Classification System.

D Linking CNIPA Patent Data, SAIC Data and New Campus Data

We first obtain the National Intellectual Property Administration of China (CNIPA) patent data recording the patent ID, application date, full name of assignee and IPC number of each patent. We then reshape the patent data to an assignee-patent structure.

We identify an assignee as a university or a college if it contains "University", "College" or "School" ("daxue", "xueyuan" or "xuexiao" in Chinese). We also include "School" ("xuexiao" in Chinese) because some colleges may use the term "School" in their names. We then exclude "Academy of Sciences" ("kexueyuan" in Chinese) as other organizations since these facilities may be recognized as colleges given their spelling in Chinese. We then identify an assignee as a firm if it contains "Corporation" or "Plant" ("gongsi" or "chang" in Chinese) since these suffixes cover almost every firm (excluding sole proprietorship) in China, regardless of its ownership and size. Other assignees including social organizations, individuals and government bodies are classified into other organizations.

With information on the type of assignees, we then turn to the identification of joint patent applications between universities and firms. Patents are assigned to a joint application if they have at least one university assignee and one firm assignee. In doing so, we exclude those patents applied by universities or firms alone. We then generate a set of variables $univ_1, univ_2, \cdots, univ_n$ to identify the university involved in every patent and collapse the assignee-patent data by assignee, year of application and university involved to form an assignee-year-university panel data. We can then calculate the frequency of joint patenting between a firm and a university.

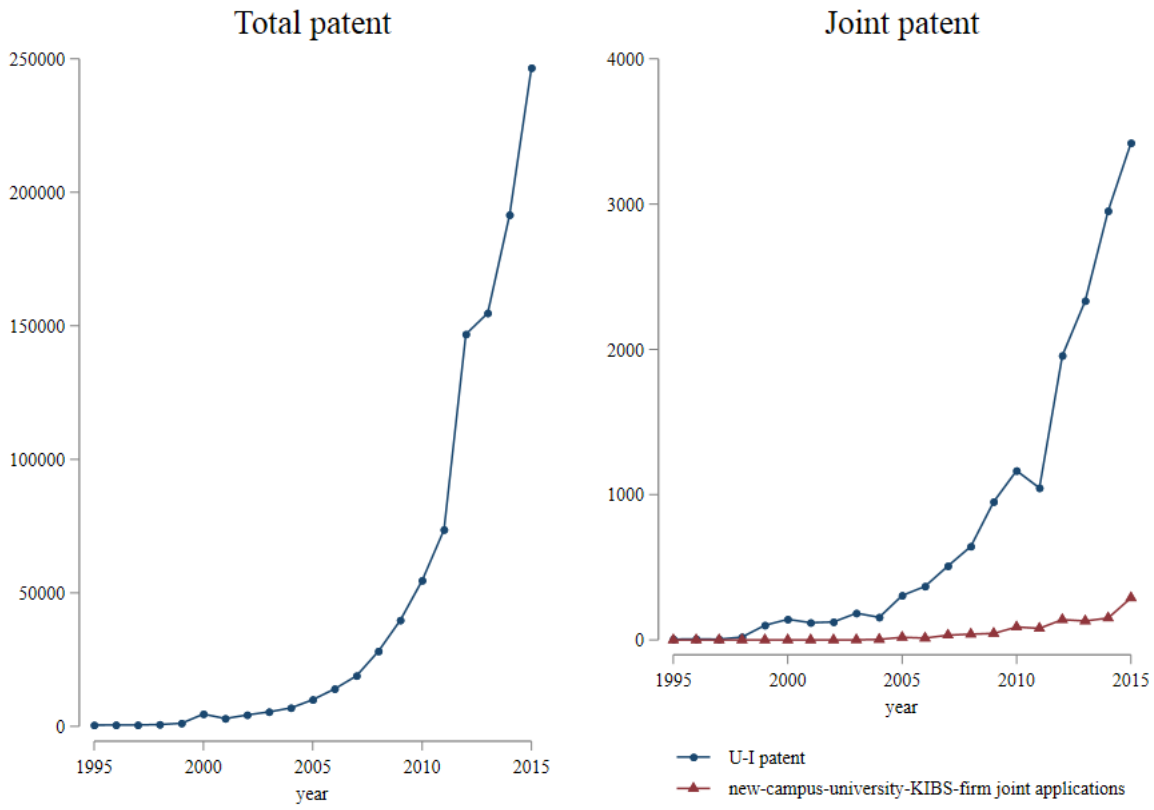
Overall, there are 84,127 joint patent applications by universities and firms between 1985 and 2015 after we sum up the frequency of joint patenting in the data. This is close to but not exactly equal to the total number of joint applications by universities and firms given that under a few circumstances, a patent may be jointly filed by multiple firms and multiple universities. We then

link the patent data to the State Administration for Industry and Commerce of China (SAIC) data using assignee and firm name. We first keep only KIBS firms in SAIC data. We include only firms with unique names in the SAIC data and then merge them with the patent data to ensure the credibility of the matching results[26].

We then collapse the data to the county-year level and merge the data with new campus data to identify those patents filed by firms and the affiliated universities of new campuses that have opened in the same county (hereafter referred to as new-campus-university-KIBS-firm joint patent applications). We do so because the CNIPA data does not reveal the address of each assignee and therefore we could not determine whether the patent comes from a new campus. There are a total of 18,795 patents filed by universities and KIBS firms and 1,185 new-campus-university-KIBS-firm joint patent applications. Of the total 411 counties that host new universities or new campuses, 63 counties have new-campus-university-KIBS-firm joint patent applications. It suggests that in 15.3% of counties, the affiliated universities of new campuses are actually involved in joint patenting with KIBS firms. Figure A3 plots the joint patent applications filed by universities and KIBS firms between 1995 and 2015. The overall U-I joint patent applications appears to grow at a faster pace than local collaborations.

[26]We use an accurate matching method because a fuzzy matching procedure would have brought a number of wrong matches between patent applicants and firms. To determine the overall matching rate between the patent data and the SAIC data, we merged the SAIC data with the patent data. We find that firm names in the SAIC data match 65.8% of the total firm assignees in the patent data.

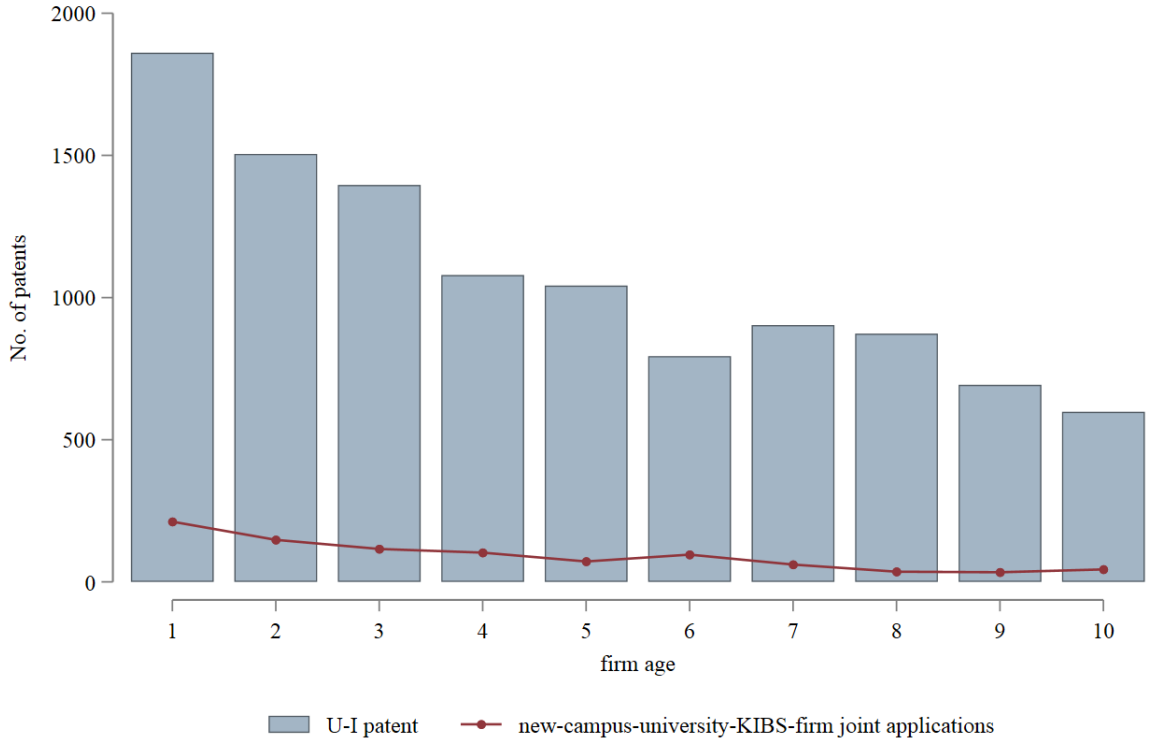
Figure A3: **Total Patent and Joint Patent Applications by Universities and KIBS Firms: 1995-2015**



Notes: Figure A3 displays total patent and joint patent applications by universities and KIBS firms between 1995 and 2015. The left panel plots overall applications filed by KIBS firms. The right panel plots in blue dots the joint patent applications filed by KIBS firms and all universities and in red triangles the joint patents applications filed by KIBS firms and the affiliated universities of new campuses that have opened in the same county. Y-axis is the number of patents and x-axis is year.

The information on the entry year of firms in SAIC data enables us to identify the age when a firm files a patent. Figure A4 plots the number of total patent applications in each firm age group. Overall, firms under the age of five (i.e. young firms) appear to be the most involved in joint patenting activities with universities, whether with new-campus-related universities or not. Joint patent activities peak at the age of one. Firms above five years old file fewer joint patents with universities. This pattern supports the hypothesize that young firms, especially start-ups, are more active in collaborating with external partners such as universities.

Figure A4: **Joint Patent Applications by KIBS firms and Universities: By Firm Age**



Notes: Figure A4 displays the joint patent applications filed by KIBS firms and universities by firm age. Blue bars plot patent applications filed by firms in KIBS and all universities. Red dots plot the joint patent applications filed by KIBS firms and the affiliated universities of new campuses that have opened in the same county. Y-axis is the number of patents and x-axis is firm age.

E Identifying Channels of Human Capital Spillovers

Since 1999, China National Bureau of Statistics has conducted three population censuses in 2000, 2010 and 2020 and two 1% population surveys in 2005 and 2015 (commonly referred to as mini-census). It was not until the 2015 Census that county-level information on individuals' former county of residence was collected. The following information from the 2015 mini-census questionnaire helps us identify graduates from local universities and colleges and migrants with bachelor's degrees: survey county, date of birth, previous year residence (county), previous five-year residence (county), educational attainment and educational status. Table A5 presents details of these questions.

Table A5: **Identifying Channels of Human Capital Spillovers**

Questions in the Survey Questionnaire	**Content**
survey county	including the province, city and county of the respondent
date of birth	including the year, month and date of birth, used to calculate the age in 2015
residence one year ago	including the province, city and county of residence one year ago
residence five years ago	including the province, city and county of residence five years ago
educational attainment	1=uneducated; 2=elementary school; 3=junior high school; 4=senior high school; 5=secondary vocational education (3 years); 6=post-secondary 3-year college; 7=post-secondary 4-year university; 8=postgraduate education
educational status	1=student; 2=graduate; 3=dropout (with study certificate); 4=dropout; 5=other

China's higher education system is composed of 4-year academic universities and 3-year vocational colleges. Meanwhile, China implements nine years of compulsory education starting at the age of six and three years of high school education. Therefore, the majority complete their secondary education at 18 and graduate from 4-year universities at 22 or from 3-year colleges at 21. We use these age cut-offs to identify whether an individual was in the survey county during his/her enrollment in universities or colleges. We then proceed to identify migrants with post-secondary education experience. Those whose residence in the previous year or five years differs from the survey county are identified as migrants. Table A6 presents details of the identification of each group, including age, previous year residence, previous five-year residence, educational attainment and educational status.

Table A6: **Identification of Graduates from Local Institutions and Migrants**

	age	previous year residence	previous five-year residence	educational attainment	educational status
graduates from local 3-year colleges	22	survey county	-	post-secondary 3-year college	graduation/dropout (with study certificate)
	23-26	-	survey county		
graduates from local 4-year universities	23	survey county	-	post-secondary 4-year university	graduation/dropout (with study certificate)
	23-27	-	survey county		
migrants with a 3-year degree	22	other county	-	post-secondary 3-year college	graduation/dropout (with study certificate)
	23-26	-	other county		
migrants with a 4-year degree	23	other county	-	post-secondary 4-year college	graduation/dropout (with study certificate)
	23-27	-	other county		

In Table A7 Columns 1-2, we provide additional results of new campuses on each sub-group of direct human capital spillovers on KIBS, including 3-year college graduates and 4-year university graduates. Results show that the openings of new campuses lead to a significant increase in 4-year university graduates in the county. There appears to be a smaller and not significant increase among graduates of local 3-year colleges. This result is consistent with the main finding that more KIBS firms, which have a greater demand for high-skilled labor, are located in the same county.

We note that using previous five-year residence information could lead to overestimation of

graduates from local institutions because we cannot rule out the possibility of mobility during that five year span. Therefore, we take an alternative identification strategy and include only individuals who stayed in the survey county one year ago to ensure that the identified groups received their highest degree from universities or colleges in the survey county. Table A7 Column 3 reports estimates on migrants with the above specifications. Results show that while the magnitude of effect on local human capital base is reduced compared to the results in Table 6 (2.3 log points), the significance remains robust to excluding individuals identified by previous five-year residence.

Table A7: **Additional Results on Human Capital Spillovers**

	(1) Log Graduates from Local 3-year Colleges	(2) Log Graduates from Local 4-year Universities	(3) Log Graduates from local HEIs (Focusing on One-year Migration)
L.Number of New Campus	0.010	0.018*	0.013**
	(0.009)	(0.010)	(0.006)
Log 1995 Night-time Light	YES	YES	YES
Log 1995 Population Density	YES	YES	YES
Prov FE	YES	YES	YES
County Observations	2499	2499	2499
Adjusted R-squared	0.162	0.258	0.109
Rounded Baseline Mean of Control Group	1.426	1.404	1.318

Notes:
1. Table A7 reports DD estimates of the effects of new campus openings on graduates from different types of local HEIs. Only individuals employed in KIBS industries are included.
2. The dependent variable is the logarithm of 2015 population in each category. Data on population of each group is drawn from the 10% representative sample from the 1% population survey conducted by the National Bureau of Statistics of China in 2015.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and 2014.
4. Each observation is a county in 2015.
5. Counties with zero college graduates aged 22-30 in the 2015 population census sample are excluded from the analysis. This exclusion is due to uncertainty regarding the reason for the absence of college graduates in these counties, which could stem from: (1) a genuine scarcity of college graduates aged 22-30 in these counties, or (2) these counties being dropped in the sampling procedure. Assigning zero college graduates to these counties could introduce estimation bias if their absence is primarily due to the latter.
6. Standard errors in the parentheses are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F Difference in Baseline Characteristics Between the Treatment and Control Counties

To show the difference in baseline characteristics between the treatment and control counties, we conduct a two-sample T test on baseline characteristics of the treatment and control counties. We show in Table A8 that treatment counties are significantly more advantaged in economic and demographic profiles compared to control counties. This is primarily because most of the treated counties are situated closer to urban districts, making them geographically advantageous as potential sites for new campuses. While the economic conditions of these counties may not directly influence the opening of new campuses, they exhibit stronger economic activity overall.

While we acknowledge that the site selection of new campus could be endogenous to these baseline characteristics, we provide evidence for a relatively constant gap of the outcome variables between the treatment and control groups across all pre-treatment periods in the main text Section 5.2 and Figure A5 using an event study. The results show that the level of entrepreneurship remains consistent across all pre-treatment periods and industries, suggesting that the treatment counties did not exhibit increasing entrepreneurial activity prior to the treatment. Following the opening of the first new campus, the gap between the treatment and control groups expanded. This provides further evidence for the causal effect of new campuses on stimulating local entrepreneurship.

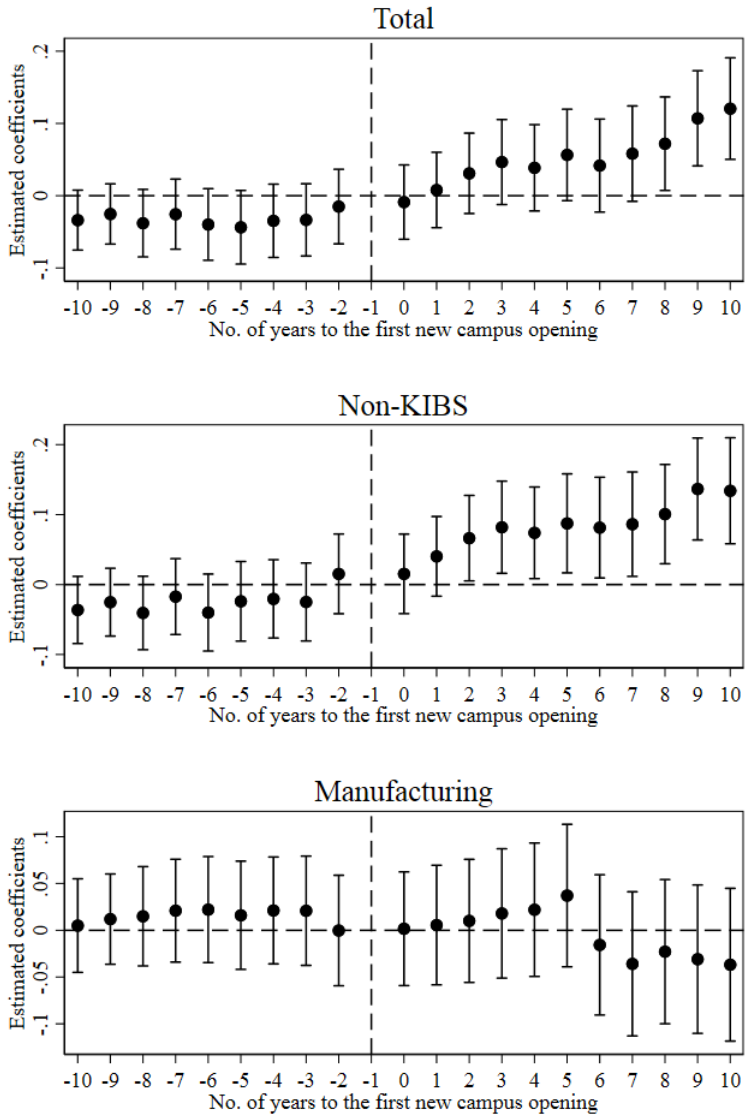
We also present several robustness checks to address other relevant endogeneity concerns in the main text Section 5.4. The 2SLS estimation also supports a reliable baseline estimate using a policy- and distance-based instrument for new campus openings.

Table A8: **Two-sample T Test on Baseline Characteristics of the Treatment and Control Counties**

Variables	Obs.(Control)	Mean (Control)	Obs.(Treatment)	Mean (Treatment)	Diff. (Treatment-Control)
Total Firm Entry	2290	170.369	411	455.659	285.291***
KIBS Firm Entry	2290	14.320	411	39.533	25.213***
Non-KIBS Firm Entry	2290	108.495	411	276.591	168.096***
Manufacturing Firm Entry	2290	37.901	411	115.039	77.138***
Night-time Light	2290	4.678	411	8.894	4.216***
Population Density	2290	483.580	411	830.582	347.002***

Notes: Table A8 reports results from a two-sample T test on baseline characteristics of the treatment and control counties. Variables used include total firm entry, KIBS firm entry, non-KIBS firm entry, manufacturing firm entry, night-time light and population density in 1995. Each observation is a county in 1995.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure A5: **Event Study Results for Total, Non-KIBS and Manufacturing Firm Entry**



Notes: Figure A5 displays in black dots the estimated coefficients of new campuses on local firm entry using a standard TWFE estimator. The dependent variable is the logarithm of one plus firm entry. Bars represent the estimated 95% confidence intervals. X-axis represents the number of years relative to the first new campus opening in the county since 1999. Each observation is a county-year. All specifications include log night-time light in 1995×Yr, log population density in 1995×Yr, county fixed effects, year fixed effects and province×Yr. Standard errors are clustered at the county level.

G Additional 2SLS Results on Total, Non-KIBS and Manufacturing

Table A9: **2SLS Results for Total, Non-KIBS and Manufacturing Firm Entry**

	(1) Total	(2) Non-KIBS	(3) Manufacturing
L.Number of New Campuses	0.137***	0.119***	-0.230***
	(0.031)	(0.033)	(0.057)
Log 1995 Night-time Light×Yr	YES	YES	YES
Log 1995 Population Density×Yrr	YES	YES	YES
County FE	YES	YES	YES
Year FE	YES	YES	YES
Prov×Yr	YES	YES	YES
First Stage Cragg-Donald Wald F stat	5074.724	5074.724	5074.724
First Stage Kleibergen-Papp rk Wald F stat	33.516	33.516	33.516
County-Year Observations	38076	38076	38076
Adjusted R-squared	-0.027	-0.021	-0.065
Rounded Baseline Mean of Control Group	721.789	341.828	100.272
Number of Counties	1814	1814	1814
Start Year	2000	2000	2000
End Year	2020	2020	2020

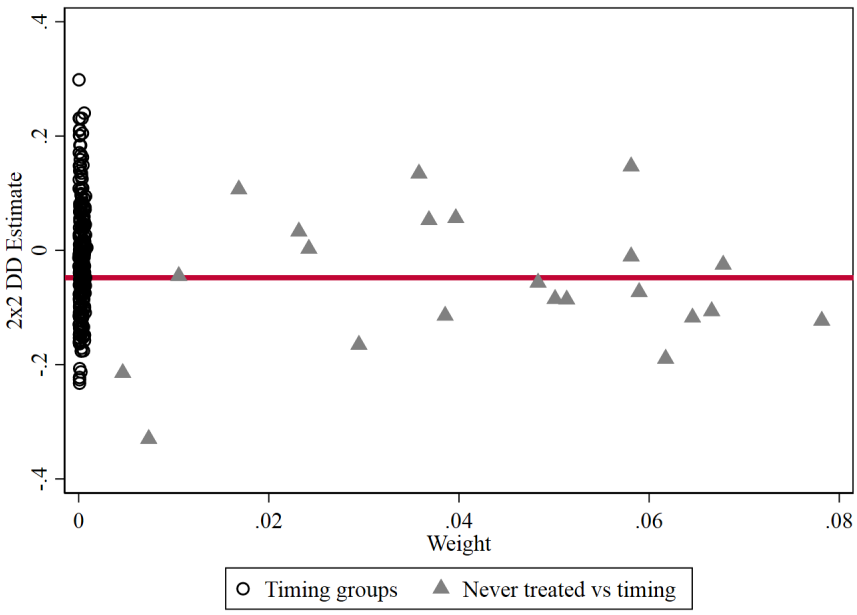
Notes:
1. Table A9 reports 2SLS estimates of the effects of new campus openings on local firm entry.
2. Central counties are excluded from the regression.
3. The dependent variable is the logarithm of one plus firm entry.
4. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
5. Each observation is a county-year.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H Additional Robustness Checks

H.1 Bacon Decomposition

Figure A6 presents the results from a Bacon decomposition on the DD estimation. The comparison between timing groups accounts for 6.9% of total DD estimates and the comparison between never treated groups and timing groups accounts for 93.1% of total DD estimates. These results suggest that the comparisons between the timing groups and the never treated groups consist of the majority of DD groups in the regression. Such comparison causes no estimation bias. Therefore, the estimation bias arising from treating the already treated group as the control group is relatively small. This is primarily due to the fact that, among all the counties, over 80% have never hosted any new campus during the entire observation period.

Figure A6: **Bacon Decomposition of Effects of New Campuses on Local Firm Entry**



Notes: Figure A6 plots the results from a Bacon decomposition. The dependent variable is the logarithm of one plus total firm entry. Each observation is a county-year. Timing groups include counties that were not treated in the first observation period and the treatment occurs during the observation period. Never Treated groups include the counties that received no intervention at all during the entire observation. No counties were treated in the first observation period and therefore there is no always-treated group.

H.2 DD Estimation using IV Sample

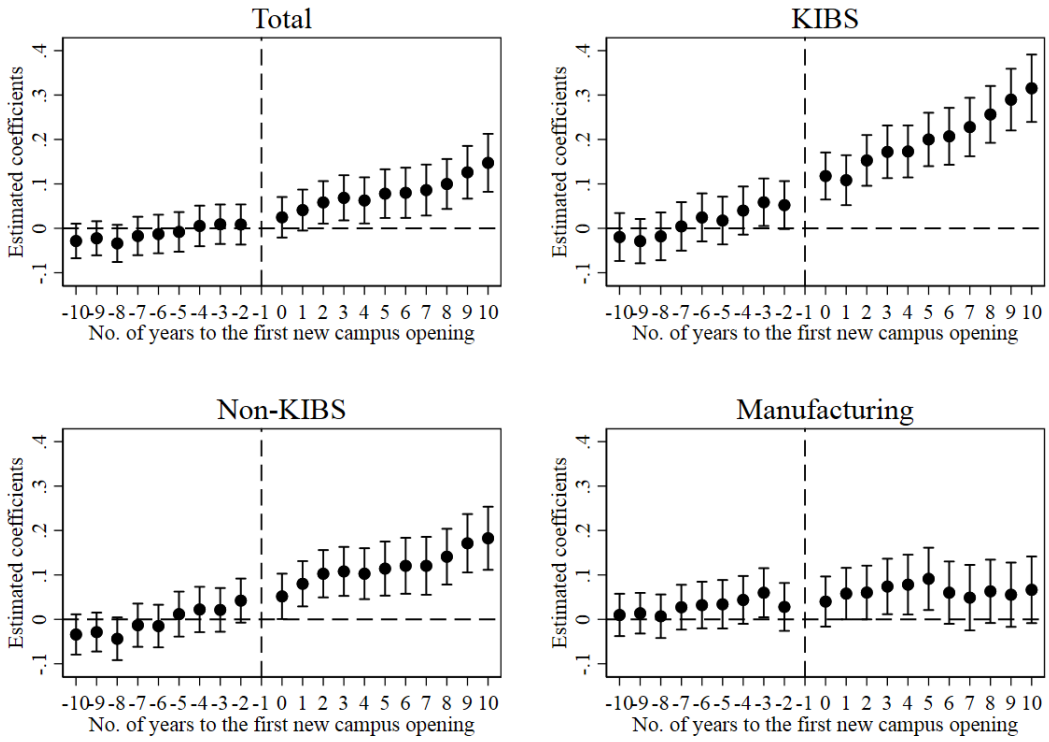
Table A10: **New Campus and Local Firm Entry: IV Sample**

	All Industries				KIBS Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	KIBS	Non-KIBS	Manufacturing	R&D	ICT	Finance	Professional Services
L.Number of New Campuses	0.050***	0.091***	0.050***	-0.014	0.058***	0.160***	0.083***	0.056***
	(0.011)	(0.013)	(0.011)	(0.013)	(0.012)	(0.019)	(0.014)	(0.012)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County-Year Observations	38076	38076	38076	38076	38076	38076	38076	38076
Adjusted R-squared	0.942	0.917	0.924	0.903	0.916	0.776	0.529	0.929
Rounded Baseline Mean of Control Group	721.789	130.367	341.828	100.272	40.858	19.074	6.613	63.822
Number of Counties	1814	1814	1814	1814	1814	1814	1814	1814
Start Year	2000	2000	2000	2000	2000	2000	2000	2000
End Year	2020	2020	2020	2020	2020	2020	2020	2020

Notes:
1. Table A10 reports DD estimates of the effects of new campus openings on local firm entry using the IV sample.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Total firm entry includes KIBS, non-KIBS, manufacturing and other industry classes in the first and secondary industry.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport).
8. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.3 Event Study with City-Year FE

Figure A7: **Event Study with City-Year FE**



Notes: Figure A7 displays in black dots the estimated coefficients of new campuses on local firm entry using a standard TWFE estimator. The dependent variable is the logarithm of one plus firm entry. Bars represent the estimated 95% confidence intervals. X-axis represents the number of years relative to the first new campus opening in the county since 1999. Each observation is a county-year. All specifications include log night-time light in 1995×Yr, log population density in 1995×Yr, county fixed effects, year fixed effects and city×Yr. Standard errors are clustered at the county level.

H.4 New Campus and Local Firm Entry: Binary Indicator

Table A11: **New Campus and Local Firm Entry: Binary Indicator**

	All Industries				KIBS Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	KIBS	Non-KIBS	Manufacturing	R&D	ICT	Finance	Professional Services
L.Any New Campus	0.128***	0.270***	0.145***	-0.012	0.268***	0.494***	0.193***	0.208***
	(0.023)	(0.024)	(0.025)	(0.028)	(0.026)	(0.029)	(0.023)	(0.025)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County-Year Observations	67420	67420	67420	67420	67420	67420	67420	67420
Adjusted R-squared	0.936	0.921	0.920	0.888	0.919	0.828	0.570	0.935
Rounded Baseline Mean of Control Group	729.735	157.287	370.707	81.426	46.650	22.829	6.929	80.879
Number of Counties	2701	2701	2701	2701	2701	2701	2701	2701
Start Year	2000	2000	2000	2000	2000	2000	2000	2000
End Year	2020	2020	2020	2020	2020	2020	2020	2020

Notes:
1. Table A11 reports DD estimates of the effects of new campus openings (1=new campus opening) on local firm entry.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable equals to 1 if the county has built any new campus and 0 otherwise.
4. Each observation is a county-year.
5. Total firm entry includes KIBS, non-KIBS, manufacturing and other industry classes in the first and secondary industry.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport).
8. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.5 Early Stage of Expansion

As noted in the background, the shortage of campus space, coupled with a policy-driven surge in enrollments, has contributed to a substantial campus boom over the past two decades. While we argue that both factors provide sufficient exogeneity in the establishment of new campuses, there are reasons to believe that in the earlier stage of this expansion, the tension between limited campus space and rapidly growing student numbers was particularly acute. As a result, the development of new campuses during this period was driven more by the immediate needs of the higher education sector and less by local economic conditions. This dynamic lends greater exogeneity to both the geographic location and the timing of new campus openings. In 2006, the State Council emphasized that the expansion rate should be controlled and that the focus of expansion policy should shift towards improving the quality of universities and colleges. Figure 2 in the main text shows that the growth rate of admitted students also decreased to lower than 10% in 2006. The opening of new campuses also experienced a sharp increase between 2000 and 2004 and fell back to around 40 in 2007. We therefore rerun the baseline regression using only observations during 1995-2007 to provide an estimation for short-term effects of new campuses built in this period.

Table A12 reports the results. New campuses built in the early stage of expansion appear to have a smaller impact on local entrepreneurship by stimulating the creation of KIBS firms. The magnitudes in T-KIBS industries — 7.1 log points on R&D and 8.8 log points on ICT — are smaller than those in the baseline results. This may be due to the shorter length of observations or simply the smaller effects of campuses built within this period. Nevertheless, the effects remain significant.

Table A12: **New Campus and Local Firm Entry: 1995-2007**

	All Industries				KIBS Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	KIBS	Non-KIBS	Manufacturing	R&D	ICT	Finance	Professional Services
L.Number of New Campuses	0.044***	0.058***	0.056***	0.006	0.071***	0.088***	0.002	0.067***
	(0.013)	(0.011)	(0.015)	(0.016)	(0.014)	(0.018)	(0.018)	(0.013)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County-Year Observations	32310	32310	32310	32310	32310	32310	32310	32310
Adjusted R-squared	0.920	0.855	0.896	0.908	0.850	0.744	0.510	0.890
Rounded Baseline Mean of Control Group	278.696	41.487	160.978	55.536	9.545	7.842	5.887	18.213
Number of Counties	2700	2700	2700	2700	2700	2700	2700	2700
Start Year	1995	1995	1995	1995	1995	1995	1995	1995
End Year	2007	2007	2007	2007	2007	2007	2007	2007

Notes:
1. Table A12 reports DD estimates of the effects of new campus openings on local firm entry between 1995 and 2007.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Total firm entry includes KIBS, non-KIBS, manufacturing and other industry classes in the first and secondary industry.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport).
8. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.6 Propensity Score Matching

As a robustness check to our baseline estimates, we use a nearest neighbor propensity score matching procedure to identify a more comparable control group. We use the baseline characteristics (log 1995 night-time light and log 1995 population density) to predict likelihood of becoming a new campus county. Each treated county with non-missing values on the above measures is paired with a control county that has the most similar performance. The difference in the control variables between the control and treatment group becomes insignificant after matching. Table A13 shows that the baseline mean of the control group is higher than that in the main analysis.

Table A13 reports the results. Overall, the findings align with the conclusions in the main analysis that new campuses stimulate entrepreneurship in total and in the service sector and they have no impact on the creation of manufacturers. Columns 5-8 show that the effect on KIBS entrepreneurship remains significant with slightly smaller magnitudes across each category. However, we note that the increase in the number of firms created becomes larger than the baseline results. The 10.1 log points (10.6%) increase is translated to about 23 new firms in KIBS, which is 21% greater than the baseline effects.

Table A13: **New Campus and Local Firm Entry: Matching Strategy**

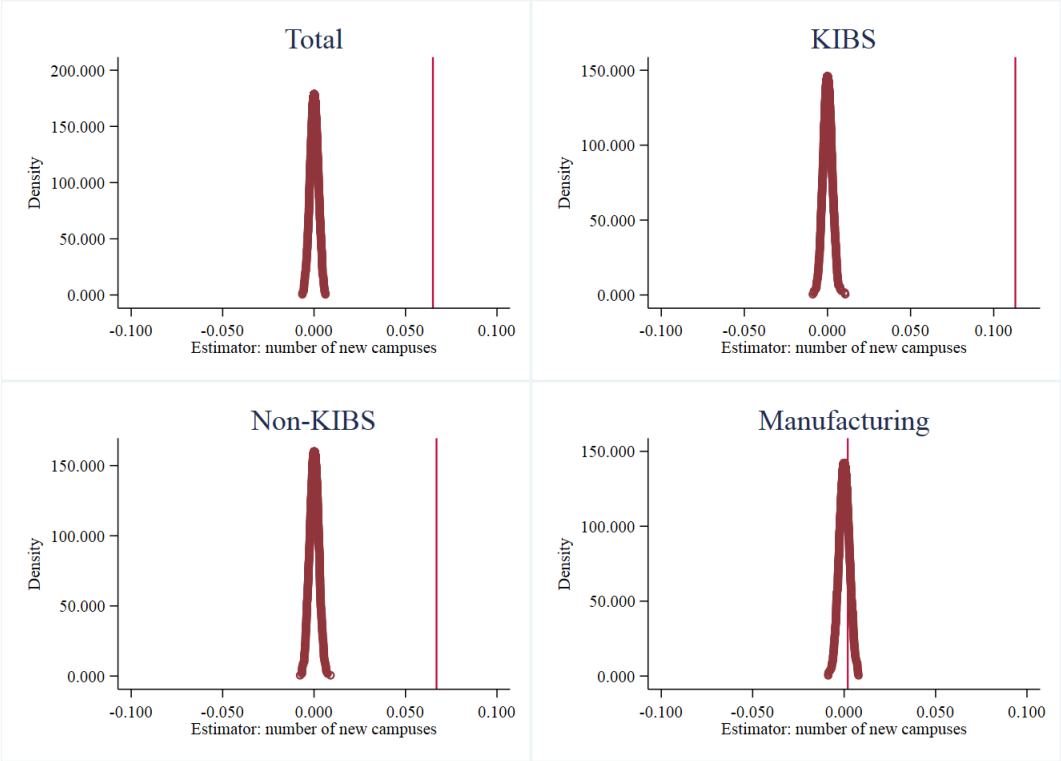
	All Industries				KIBS Industries			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total	KIBS	Non-KIBS	Manufacturing	R&D	ICT	Finance	Professional Services
L.Number of New Campuses	0.065***	0.101***	0.069***	0.014	0.094***	0.150***	0.083***	0.086***
	(0.011)	(0.013)	(0.013)	(0.014)	(0.013)	(0.019)	(0.015)	(0.013)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES	YES	YES
County-Year Observations	16172	16172	16172	16172	16172	16172	16172	16172
Adjusted R-squared	0.944	0.943	0.930	0.886	0.938	0.869	0.601	0.949
Rounded Baseline Mean of Control Group	1007.048	214.585	533.796	126.589	63.674	34.961	9.027	106.923
Number of Counties	647	647	647	647	647	647	647	647
Start Year	2000	2000	2000	2000	2000	2000	2000	2000
End Year	2020	2020	2020	2020	2020	2020	2020	2020

Notes:
1. Table A13 reports DD estimates of the effects of new campus openings on local firm entry using a matching strategy.
2. The dependent variable is the logarithm of one plus firm entry.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Total firm entry includes KIBS, non-KIBS, manufacturing and other industry classes in the first and secondary industry.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Non-KIBS includes all services other than KIBS (e.g. wholesale, retail trades, transport).
8. Standard errors in the parentheses are clustered at the county level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.7 A Placebo Test

We conduct a placebo test to investigate whether our estimates are biased from unobservable factors. We randomly assign both the location and timing of the treatment. We then repeat this process 1000 times. If the outcome is affected by other unobservables, we would expect to see a skewed distribution of estimates with a non-zero mean. Figure A8 plots the estimated coefficients of the effect on local firm entry from the permutation for firm entry across each industry classification. The distribution of all estimates conforms to a normal distribution with zero mean across each industry classification. We take this as a further indication of the validity of the baseline results.

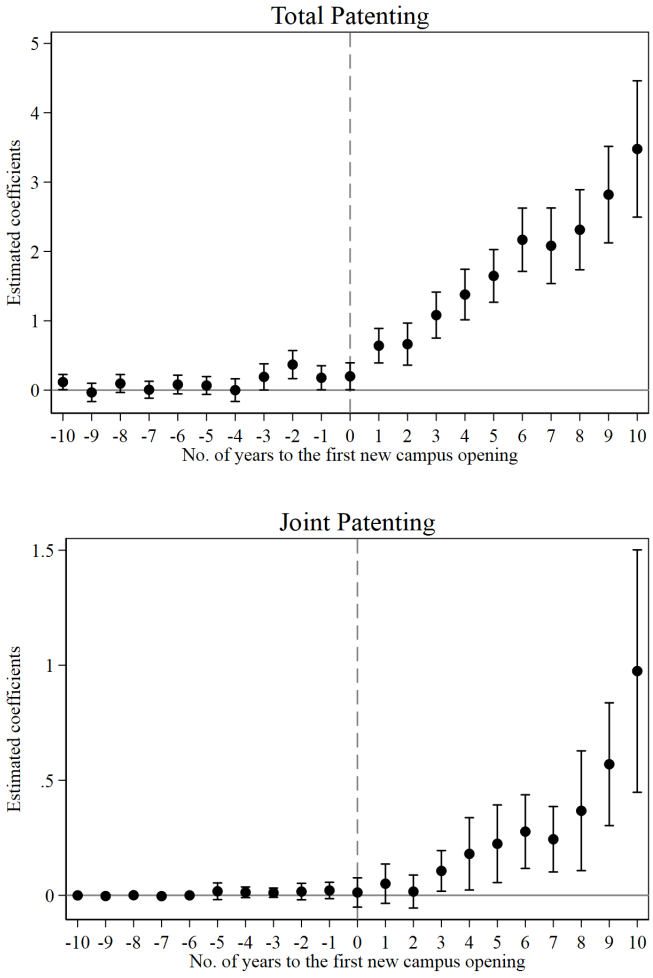
Figure A8: **Results from A Placebo Test**



Notes: Figure A8 plots the estimated coefficients produced by a placebo test. Red dots represent estimated coefficients. Red lines represent baseline estimates. The dependent variable is the logarithm of one plus firm entry. Each observation is a county-year. Standard errors are clustered at the county level.

H.8 Event Study on Local Innovation

Figure A9: **Event Study Results on Local Innovation**



Notes:
1. Figure A9 displays the estimated coefficients of new campuses on patenting activities by KIBS firms. Total patenting includes all patent applications filed by KIBS firms. Joint patenting includes patent applications filed by KIBS firms with local new-campus-related universities. Dynamic effects are obtained by the Callaway and Sant'Anna (2021) estimator.
2. The dependent variable is the logarithm of one plus patent applications.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation in heterogeneity analyses.
4. The observation period spans from 1995 to 2015.
5. Firms of finance and professional services activities are excluded as they file a very small number of patents compared to firms of the R&D and ICT industries.
6. Bars represent the estimated 95% confidence intervals.
7. X-axis represents the number of years relative to the first new campus opening in the county since 1999.
8. Each observation is a county-year.
9. All specifications include log night-time light in 1995×Yr, log population density in 1995×Yr, county fixed effects, year fixed effects and province×Yr.
10. Standard errors are clustered at the county level.
11. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.

H.9 2SLS Results on Knowledge Spillover

Table A14: **2SLS Estimation of the Effects of New Campus on Local Innovation of KIBS Firms**

	Total Patenting			Joint Patenting		
	(1) All	(2) R&D	(3) ICT	(4) All	(5) R&D	(6) ICT
L.Number of New Campuses	1.044***	1.056***	1.108***	0.147***	0.132***	0.020
	(0.216)	(0.198)	(0.230)	(0.038)	(0.035)	(0.013)
Log 1995 Night-time Light×Yr	YES	YES	YES	YES	YES	YES
Log 1995 Population Density×Yr	YES	YES	YES	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Prov×Yr	YES	YES	YES	YES	YES	YES
First Stage Cragg-Donald Wald F stat	4084.046	4084.046	4084.046	4084.046	4084.046	4084.046
First Stage Kleibergen-Papp rk Wald F stat	25.602	25.602	25.602	25.602	25.602	25.602
County-Year Observations	29006	29006	29006	29006	29006	29006
Adjusted R-squared	-0.189	-0.207	-0.374	-0.015	-0.011	-0.023
Rounded Baseline Mean of Control Group	8.633	6.366	1.462	0.001	0.001	0.000
Number of Counties	1814	1814	1814	1814	1814	1814
Start Year	2000	2000	2000	2000	2000	2000
End Year	2015	2015	2015	2015	2015	2015

Notes:
1. Table A14 reports 2SLS estimates of the effects of new campus openings on local innovation. Total patenting includes all patent applications filed by KIBS firms. Joint patenting includes patent applications filed by KIBS firms with local new-campus-related universities.
2. The dependent variable is the logarithm of one plus patent applications.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and one year before observation.
4. Each observation is a county-year.
5. Firms of finance and professional services activities are excluded as they file a very small number of patents compared to firms of the R&D and ICT industries.
6. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
7. Standard errors in the parentheses are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

H.10 2SLS Results on Human Capital Spillover

Table A15: **2SLS Estimation of the Effects of New Campus on Local Human Capital in KIBS**

Panel A: Direct and Indirect Channels of Human Capital Spillover to KIBS

	(1) Log College Graduates	(2) Log College Graduates from Local Institutes	(3) Log Migrants with College Degree
L.Number of New Campuses	0.328***	0.128**	0.285***
	(0.086)	(0.054)	(0.079)
Log 1995 Night-time Light	YES	YES	YES
Log 1995 Population Density	YES	YES	YES
Prov FE	YES	YES	YES
First Stage Cragg-Donald Wald F stat	332.338	332.338	332.338
First Stage Kleibergen-Papp rk Wald F stat	29.867	29.867	29.867
County Observations	1746	1746	1746
Adjusted R-squared	-0.049	-0.073	-0.188
Control Baseline Mean	1.235	0.532	0.431

Panel B: Sub-industries of KIBS

	(1) R&D	(2) ICT	(3) Finance	(4) Professional Services
L.Number of New Campuses	0.064***	0.242***	0.105*	0.089**
	(0.022)	(0.071)	(0.058)	(0.037)
Log 1995 Night-time Light	YES	YES	YES	YES
Log 1995 Population Density	YES	YES	YES	YES
Prov FE	YES	YES	YES	YES
First Stage Cragg-Donald Wald F stat	322.424	322.424	322.424	322.424
First Stage Kleibergen-Papp rk Wald F stat	28.076	28.076	28.076	28.076
County Observations	1746	1746	1746	1746
Adjusted R-squared	-0.060	-0.207	-0.022	-0.088
Control Baseline Mean	0.138	0.350	0.434	0.314

Notes:
1. Table A15 reports 2SLS estimates of the effects of new campus openings on local human capital.
2. The dependent variable is the logarithm of population in each category. Data on population of each group is drawn from the 10% representative sample from the 1% population survey conducted by the National Bureau of Statistics of China in 2015.
3. The key independent variable is the cumulative number of new campus openings a county has received between 1999 and 2014.
4. Each observation is a county in 2015.
5. Knowledge-Intensive Business Services (KIBS) include the following CICS industries: (1) Scientific and technical services (R&D), (2) Information services (ICT), (3) Finance and insurance (Finance), (4) Professional services including management of companies and enterprises (Professional services). See Appendix C for the relationship between CICS and NAICS.
6. Counties with zero college graduates aged 22-30 in the 2015 population census sample are excluded from the analysis. This exclusion is due to uncertainty regarding the reason for the absence of college graduates in these counties, which could stem from: (1) a genuine scarcity of college graduates aged 22-30 in these counties, or (2) these counties being dropped in the sampling procedure. Assigning zero college graduates to these counties could introduce estimation bias if their absence is primarily due to the latter.
7. Standard errors in the parentheses are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.