INTERNAL MIGRATION LIBERALIZATION AND PRODUCTIVITY*

Ailin Zhang†

July 28, 2026

Abstract

This study investigates the effect of urban size on firm productivity by exploiting the quasi-experimental nature of *hukou* reforms in China. The difference-in-differences estimates reveals that 4 years after the implementation of *hukou* reform in a prefecture-level city, its registered urban population increased by around 4%. The increase in urban size had a significant positive effect on the productivity of manufacturing firms of all ownership types except for foreign firms. The most pronounced impact was observed among private firms, for which a 1% increase in urban population led to around 2.2% higher TFP.

*I am grateful to Gabriel Ahlfeldt, Michael Gmeiner, Christian Hilber, Xavier Jaravel and Michael Pflüger for very helpful discussions and suggestions. I also thank the faculty and students from the Department of Geography and Environment at LSE, participants at the Cambridge Land Economy Early Career Researchers' Conference 2023 and Urban Labor Markets Workshop at IAB Nuremberg 2023. This paper forms the first chapter of my PhD thesis and also served as the summative essay for the course EC475.

†a.zhang12@lse.ac.uk

1 Introduction

The positive effect of agglomeration on productivity has been extensively documented in the literature that was started by Ciccone and Hall (1993). These studies (Au and Henderson, 2006; Bryan and Morten, 2019; Ma and Tang, 2020) suggest that institutional frictions in labour markets that constrain city size and hinder mobility may hamper productivity. While the economy-wide consequences of migration barriers are widely supported by empirical evidence, the firm-level productivity implications of migration remain less clear. On the one hand, Peri (2012) and Ottaviano et al. (2018) find that immigration improves firm productivity. On the other hand, Imbert et al. (2022) show that rural-to-urban migrants in China are likely to be selected into less productive firms and may weaken firms' incentives to adopt capital-intensive technologies due to the presence of cheap labour supply. This suggests that population density may also dampen micro-level productivity in certain contexts.

This study thus aims to contribute to the ongoing debate on the seemingly contradictory effects of population density on firm productivity. A central empirical challenge in this literature is simultaneity: more productive cities attract workers, while worker inflows may themselves affect productivity through agglomeration economies, labour pooling, and market-size effects. I isolate one direction of this simultaneous relationship by exploiting staggered *hukou* reforms that relaxed urban migration barriers and induced changes in urban population. Given that estimates of density economies that exploit plausible quasi-experiments remain novel in developing countries featuring rural-to-urban migration, this paper applies a stacked difference-in-differences (DD) design (Deshpande and Li, 2019; Gu et al., 2021), together with 2SLS, to estimate the elasticity of firm productivity with respect to urban population.

This paper is closely related to Imbert et al. (2022), who also study how rural–urban migration affects Chinese manufacturing firms. Whereas Imbert et al. (2022) combine origin-side agricultural income shocks with historical migration networks to isolate cumulative gross inflows of rural workers, I exploit destination-side *hukou* reforms to isolate reform-induced growth in the urban population stock. Both margins can expand labour supply and may generate factor-substitution as well as agglomeration effects, but they are not the same treatment. Cumulative inflows measure firms' exposure to migrant arrivals, including migrants who may subsequently return, whereas population growth more directly captures the resulting expansion in city size. So the two designs estimate different local responses and their net productivity effects need not coincide. By linking reform-induced urban population growth to firm productivity, this paper provides complementary micro-level evidence on the productivity consequences of migration liberalization.

Understanding the effect of density on firm TFP holds meaningful policy relevance, as it implies the potential benefits of reducing internal migration barriers erected by the governments in some developing countries such as Vietnam, Ethiopia and China. Fig. 1 gives a motivating illustration of the migrating constraints imposed by Chinese *hukou* system[1], plotting the log real output per worker against two measures of log population across Chinese prefecture-level cities in 2000. While there is a positive relationship between permanent population[2] and real output per worker, such a relationship does not exist for *hukou* population. This underscores the institutional barriers that hinder formal migration into more productive cities. Fig. 2 depicts the correlation between mean log TFP and log

[1] More information on the institutional background in Section 3.

[2] In China, permanent population (in some sense equivalent to "real population") denotes those who physically reside in a city for no less than half year, either having an official visa-like *hukou* status or not.

population for Chinese cities in the same year. The slope of the regression line is 0.088, a relatively high estimate as compared to developed countries (eg. 0.025 for France in Combes et al. (2012)). These two figures highlight that although density is positively associated with firm productivity, migration barriers under *hukou* system might have induced labor misallocation, restricting the potential for density economies.

This study finds that the reforms of *hukou* policies between 2001-2005 facilitated internal migration. The increase in urban population positively impacted on the productivity of domestic manufacturing firms, and the effect is larger on private firms than on state-owned or collectively-owned firms.[3] Besides, while the labour supply increased in the reformed cities, both per worker real-value added and per worker real wage were raised at firm level, implying the important role played by density economies. The unique policy implications are that while liberalizing internal migration provides productivity gains, the gains are not equally distributed, and that a more market-oriented economy might maximize the overall density benefits.

The results complement the existing literature in two ways. First, while recent papers implying the productivity losses induced by *hukou* system adopt aggregate analysis, this study yields micro evidence. Second, this study uses *hukou* reform as a quasi-experiment to estimate the density elasticity of firm TFP, shedding light on the seemingly ambiguous density effects on productivity found at firm level.

This article proceeds as follows. The next section reviews related literature and Section 3 introduces the institutional background. Section 4 describes the data. Section 5 presents the empirical strategy. Section 6 discusses the results and robustness. The final section concludes.

2 Related literature

2.1 Density economies and migration barriers

Arising from the spatial concentration of economic activities, agglomeration economies encompass both localization economies, which account for spillover effects within a specific industry (Marshall, 1890; Arrow, 1971; Romer, 1986; Glaeser et al., 1992), and urbanization economies, which highlight cross-industry spillover benefits enjoyed by firms in a more diversified and complementary environment (Jacobs, 1969; Kim, 1991). When viewed in light of urbanization economies, denser[4] cities promote productivity through several mechanisms. First, a larger labor pool facilitates matching, allowing firms to hire more qualified workers with low acquisition costs (Dauth et al., 2022). Second, denser cities enable firms to share local infrastructure more efficiently and access specialized services that may be absent in smaller cities (Eberts and McMillen, 1999), such as logistics networks, specialized suppliers, legal and accounting services, and engineering support. Third, denser cities facilitate interpersonal learning and the transmission of technological knowledge that decays with distance (Jacobs, 1969; Berliant et al., 2006), for example through worker mobility, supplier-client interactions, and informal exchanges among managers, engineers, and skilled workers. Empirical literature widely corroborates these productivity gains and so far has broad consensus on the magnitude of the elasticity

[3] State-owned firms are owned by the central or local state. Collectively-owned firms are formally owned by collectives, often linked to local communities, workers, or local governments, and may operate under different governance constraints from both state-owned and private firms.

[4] As city size (total population) and actual density (population per geographic unit) are highly interdependent and hard to disentangle, especially in studies using between-city comparisons, I follow the literature and use these two concepts interchangeably. See Ahlfeldt and Pietrostefani (2019) for further discussions.

of productivity with respect to city size ranging between 0.02 and 0.10.[5] Based on a citation-weighted average of previous estimates, the meta-analysis conducted by Ahlfeldt and Pietrostefani (2019) indicates that 1% increase in density leads to around 0.04% higher labour productivity and 0.06% higher total factor productivity.

Given the pronounced density economies, it is unsurprising that recent findings suggest migration barriers hinder productivity by limiting population scale. For example, Au and Henderson (2006) find that the internal migration restrictions under *hukou* system prevented Chinese cities from reaching an optimal size that would maximize real income per worker. Using a structural model, Bryan and Morten (2019) estimate the welfare effect of reducing all barriers to internal labour migration in Indonesia, and their counterfactual analysis suggests a 22% increase in labour productivity. Building on an Eaton-Kortum trade model, Tombe and Zhu (2019) imply that the reduction in internal frictions accounts for 20% more of China's growth in aggregate labor productivity than the reduction in external trade costs between 2000 and 2005. However, most of previous findings on productivity consequences of reducing migration frictions are based on simulations under structural frameworks (Bryan and Morten, 2019; Tombe and Zhu, 2019; Ma and Tang, 2020) rather than real-world policies.

Moreover, while most established evidence on the economies of density or diseconomies of migration barriers measures productivity in terms of labor productivity, such as wage or output per worker, only few studies explore the effects on firm-level productivity. Cingano and Schivardi (2004) indicate that in Italy, a move of the city-size from the first to the third quartile increase local productivity growth by 0.6% annually, whereas the effects of urban diversity and local competition are insignificant. In line with this, Combes et al. (2012) use French establishment-level data and show that on average, firms in denser areas have 9.7% higher TFP as a result of urbanization economies rather than selection effect.

Despite the robust evidence on density economies in developed countries, recent studies in developing-country contexts suggest that rural labour inflows may have more ambiguous effects on productivity during structural transformation. Imbert et al. (2022) combine agricultural commodity price shocks with historical migration networks to isolate rural-to-urban migrant inflows in China. They find that these inflows make manufacturing production more labour-intensive, reduce productivity, and shift firms away from capital- and skill-intensive innovation. Similarly, Bustos et al. (2019) show that agricultural technological change in Brazil reallocates less-skilled workers toward low-R&D manufacturing industries and slows aggregate manufacturing productivity growth. Together, these studies show that origin-side shocks releasing less-skilled rural labour can discourage capital deepening and industrial upgrading. By contrast, reforms that lower migration and settlement barriers generate variation in the urban population stock, a distinct treatment that may combine labour-supply responses with the productivity effects of a larger urban market. The limited quasi-experimental evidence on the productivity consequences of relaxing migration barriers, together with the mixed firm-level evidence from developing countries, motivates this study.

Methodologically, a growing literature has focused on obtaining casual estimates of agglomeration economies. One commonly adopted approach among earlier papers is to regress proxies for productivity, such as wages or TFP, on measures of agglomeration scale, such as population density (Ciccone and Hall, 1993; Rauch, 1993). This sort of specifications solely capture a correlation rather than a causal relationship. A typical simultaneity problem is that while population concentration could affect productivity, people might also self-select into more productive areas. To tackle the endo-

[5]See Combes and Gobillon (2015) for a review of empirical literature and theoretical mechanisms behind agglomeration economies.

geneity concerns, some resort to historical or geologic variables such as soil condition as instruments (Redding and Venables, 2004; Rosenthal and Strange, 2008; Combes et al., 2010), while others use quasi-experiments to exploit plausibly exogenous variations in density or market potential (Greenstone et al., 2010; Redding and Sturm, 2008; Ahlfeldt and Feddersen, 2018).

Different from the 'level' instruments like geology and historical population, quasi-experiments are able to capture agglomeration effects over time and give stronger justification for exogeneity. For example, Greenstone et al. (2010) use the outcome of bidding wars for large industrial plants as an exogenous labour demand shock, showing that incumbent plants in the winning county saw a 12% productivity increase after 5 years relative to those in the runner-up county owing to agglomeration spillovers. This study adds to this strand of literature by introducing *hukou* reform as a quasi-experiment that shifted urban migration barriers and generated variation in local population growth, allowing me to estimate the causal effect of urban population on firm TFP. A policy experiment directly targeting internal migration barriers provides a novel setting for estimating density effects in a developing-country context.

2.2 Estimating agglomeration economies with TFP

As discussed in Combes and Gobillon (2015), one way that has been used in recent literature[6] to estimate the effect of agglomeration variables on firm TFP is to adopt a two-stage estimation strategy. In the first stage, without introducing any local variable, the production function as well as TFP as residuals are estimated using either dynamic panel model (Arellano and Bond, 1991) or semiparametric approaches in the spirit of Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015). In the second stage, one regresses the local-time averages of firm TFP on some local agglomeration measures, and the endogeneity of local determinants is further addressed with approaches such as instrumental variables.

However, this literature requires further justification for why agglomeration variables can be excluded from the first-stage estimation of the production function and TFP.[7] If agglomeration affects firms' input choices through channels other than productivity, omitting local agglomeration variables from the control function may bias the estimated input parameters, the recovered TFP, and hence the second-stage relationship between TFP and agglomeration. Using the control function approach in Ackerberg et al. (2015), this study deviates from most related literature by incorporating density into the TFP estimation. While Appendix A gives details on the estimation strategy, here briefly consider the following production function:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it},$$

where y_{it} is value added, k_{it} denotes capital, l_{it} labour, and ω_{it} TFP.[8] In the first step of the ACF

[6] For example, Martin et al. (2011); Combes et al. (2012); Hu et al. (2015); Holl (2016); Cainelli and Ganau (2018); Tao et al. (2019).

[7] To my knowledge, the only exception is Rizov et al. (2012), who account for local economic environment characteristics in their estimation of TFP. They follow Olley and Pakes (1996) and invert the investment demand function to proxy for ω_{it}, which is less likely to satisfy the monotonicity assumption than inverting intermediate input demand, as noted by Levinsohn and Petrin (2003).

[8] All variables are in logarithms.

method, I write intermediate input demand as

$$m_{it} = f_t(k_{it}, l_{it}, \omega_{it}, d_{ct}),$$

where d_{ct} denotes the population size of city c in year t. Including d_{ct} in the control function is useful because local density may affect intermediate input demand not only through firm productivity, but also through local input-market conditions. For example, denser cities may have thicker supplier markets, better logistics, and lower prices or greater availability of specialized intermediate inputs. Conditional on capital, labour, and productivity, firms in denser cities may therefore choose different levels of intermediate inputs. If these local input-market conditions are omitted from the control function, the inversion of intermediate input demand may fail to isolate the productivity term ω_{it}. By incorporating d_{ct}, as long as $\frac{\partial m_{it}}{\partial \omega_{it}} > 0$, then conditional on k_{it}, l_{it}, and d_{ct}, I can invert m_{it} and use $\omega_{it} = f_t^{-1}(k_{it}, l_{it}, m_{it}, d_{ct})$ to proxy for productivity. Writing $\Phi_t(k_{it}, l_{it}, m_{it}, d_{ct}) = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(k_{it}, l_{it}, m_{it}, d_{ct})$, I obtain the first-stage equation to be estimated in ACF:

$$y_{it} = \Phi_t(k_{it}, l_{it}, m_{it}, d_{ct}) + \varepsilon_{it}.$$

3 Institutional background

Legislatively established in 1958[9], the household registration system (*hukou* system) in China divided citizens into agricultural and non-agricultural population[10] that effectively prohibits the internal migration (primarily rural-to-urban). Apart from being an internal visa that regulates population distribution, *hukou* also helped to consolidate the socialist system in other ways such as facilitating the ration of food and services under the planned-economy regime. Essentially, the *hukou* system not only created obstacles to geography mobility, but also constituted a hindrance to social mobility in that compared to agricultural population, non-agricultural population born in urban areas had superior access to employment opportunities, state-provided benefits, housing, education, etc.

After the opening-up of economy in 1978, the Chinese authorities began to legally allow rural migrants to permanently reside in towns (transitional areas between urban and rural) under certain circumstances since 1984[11], but the effect of this reform was insignificant[12]. In 1980s people were legally permitted to work temporarily in other locations without registration, and a limited number of rural-to-urban migrants emerged as a result of both the increased labour demand in cities and

[9]In Jan 1958, greatly influenced by the *propiska* (residential permit) system in the former Soviet Union, the Standing Committee of the National People's Congress (NPC) in China passed *Regulations of the People's Republic of China on Household Registration*.

[10]Notably, there is a dual classification of *hukou* registration that classify *hukou* status alongside two dimensions: domicile (place of regular residence) and social-economic eligibility (type of *hukou* registration). While the first mainly categorizes *hukou* registration into either urban, rural or town settlements and allocate daily necessities accordingly, the second divides people into agricultural and non-agricultural population based on their occupations and determines their socio-economic eligibility. A "formal" rural-to-urban migration requires both a change in regular *hukou* registration place and a shift of *hukou* status to non-agricultural (*nonzhuanfei*), the later of which is the critical step in the *hukou* conversion process (Chan and Zhang, 1999).

[11]According to *Circular on Rural Work in 1984*, province-level governments were authorized to pilot reform programs in towns that allow rural migrants to enter urban sectors and access non-agricultural citizenship.

[12]For one thing, as separated by the place of regular residence (1st dimension of *hukou*), town citizens enjoyed less social welfare (eg. less grain and edible oil rations) as compared to urban citizens. For another, small towns are far less attractive in terms of job opportunities as compared to more developed urban areas.

the rural labour surplus following the mechanization of agriculture. But a “formal" rural-to-urban migration remained arduous in that the conversion of agricultural *hukou* into non-agricultural *hukou* (*nonzhuanfei*) was highly restricted by an annual quota under the control of the central government that set a limit on the amount of qualified people who could eventually succeed (Chan and Zhang, 1999; Meng, 2012). Moreover, upon successfully completing the *nonzhuanfei* process, a migrant had to further meet the criteria set by the local government (usually annual quotas on the number of migrants to be conferred urban citizenship) to convert his/her place of regular residence into urban.

It was not until the 1990s that migration barriers became less of a constraint. By the mid-1990s, the central government had delegated the discretion to impose *nonzhuanfei* quotas to cities and henceforth the management of *hukou* system has been greatly localized. In 1997, the State Council loosen the *hukou* policies in selected small towns and county-level cities by granting non-agricultural *hukou* to rural migrants so long as they had local stable jobs and residence. Yet as a preliminary experiment, migrants were still subject to quota control of higher-level government (Chan and Zhang, 1999; Jin and Zhang, 2022), and this pilot scheme only had small-scale impact. It is documented that until 2002, only 1.39 million *hukou* (0.15% of national agricultural *hukou* population) were converted into non-agricultural ones in these small urban centers nationwide (Tian, 2003).

During 2001–2013, a new round of *hukou* reforms was spontaneously launched, in which prefecture-level governments had great autonomy in deciding whether to reform or not and when to reform. These *hukou* reforms adopted by a number of cities in different years share a similar nature and can be summarized into two moves. First, many local governments replaced the rigid annual quota system with a quota-free system to lower the entry requirements for a local *hukou* status. To be more precise, the strict cap imposed on the amount of migrants eligible to obtain urban citizenship was removed and replaced by basic entry requirements such as stable income and fixed residence. Second, most reformed cities eliminated the distinctions between agricultural and non-agricultural *hukou* and unified both into *jumin hukou*, therefore reducing the dual classification of *hukou* status to single dimension: place of regular residence[13]. It is thus expected that the easier access to urban citizenship and the increased possibility of enjoying equal social benefits within urban areas attracted surplus rural workers, making urban labour supply more elastic[14]. The highly comparable nature behind the *hukou* reforms implemented in different cities enables us to exploit them in a staggered difference-in-difference setting to capture the urban population change in response to the lifting of migration barriers. In 2014, the reform was scaled up nationwide under the “National Urbanization Plan (2014–2020)", almost completely abolishing the restrictions on accessing *hukou* status in all small-to-medium-sized cities. Limited by the firm-level data, this study focuses on the reforms adopted between 2001–2005.

Notably, cities that reformed during 2001–2005 were heterogeneous in size, industrial structure, and baseline economic conditions. The institutional background does not imply that reform timing was literally random. Rather, reform adoption reflected multiple and partly offsetting local considerations, including land conversion, labour retention, urban welfare costs, and local development strategies. This makes it unlikely that reforms were mechanically targeted only to cities with a particular productivity trajectory, although it does not by itself guarantee exogeneity. The empirical strategy therefore relies on conditional parallel trends: absent the reform, treated and control cities would have followed similar trends in population and firm productivity.

[13] Appendix B gives an example on the policy document.

[14] The reform is likely to have a more significant impact on rural migrants than on their urban counterparts because the former needed an additional *nongzhuanfei* process under the previous quota system

Several institutional features support this interpretation. First, mayors' attitudes toward *hukou* reform were mixed, suggesting that reform adoption was not governed by a uniform national rule or a single observable city characteristic. Under the Chinese administrative hierarchy, mayors acted as chief executive officers and possessed substantial autonomy in local development decisions (Henderson et al., 2009). Contemporary reports indicate that *hukou* reform met substantial resistance from many local officials after 2001.[15] At the same time, some mayors, such as Zangshengye in Shijiazhuang, actively promoted liberalization despite opposition from other officials.[16] Second, cities with different industrial structures had reasons to relax migration barriers. As Zhan (2017) argues, both industrial enterprises and agribusiness companies could benefit from *hukou* reform through the acquisition and consolidation of rural land when rural residents were relocated. Third, both large and small cities had potential incentives to reform, although for different reasons. Large cities faced stronger demand for urban and industrial construction land, and the provision of urban *hukou* could facilitate land conversion because giving up rural *hukou* was often associated with the loss of rural land entitlements.[17] This helps explain why some large cities, such as Chengdu, reformed despite the fiscal burden of extending urban welfare. Smaller cities, by contrast, may have reformed to retain workers who might otherwise migrate to nearby reformed cities. Taken together, these institutional features suggest that reform timing was shaped by several competing local incentives rather than by a single city characteristic such as size or industrial composition. I therefore assess the credibility of the identifying assumption using pre-trend tests, placebo timing exercises, and checks on other local outcomes. Table 1 provides descriptive evidence consistent with this institutional discussion. Among cities observed in the 1998 baseline cross-section, treated cities varied substantially in size and economic structure. They were on average larger, richer, and more industrialized than control cities, while the two groups had broadly similar initial urbanization rates and GDP growth rates. The observed baseline differences motivate the use of city fixed effects, group-relative-time fixed effects, controls for city-specific trends, and explicit tests for differential pre-trends.

4 Data

4.1 Hukou reform

To identify the timing of local *hukou* reforms, I compiled information from local government websites, news and law databases following Jin and Zhang (2022), who provide the first prefecture-level dataset on local *hukou* reforms.[18] After collecting reform years for prefecture-level cities between 2001 and 2013 from official documents and law databases, I focus on the 76 treated cities with available city-level data that reformed during 2001–2005, as shown in Appendix Fig. B.1. Because consistent ASIP firm-level data are available only through 2007, this restriction ensures that each treated city contributes

[15] Southern Daily (2012), "More effective measures are needed to advance the *hukou* reform," available at https://news.sina.com.cn/pl/2012-08-21/111325006977.shtml.

[16] Chinanews (2001), "Shijiazhuang implements bold reform and the household registration system is fully liberalized," available at http://news.sohu.com/70/20/news146162070.shtml.

[17] In the 1980s and 1990s, local governments sometimes forced peasants to give up their land in exchange for compensation, which triggered social tension and resistance. To acquire land in a less confrontational way, some local governments used the offer of urban *hukou* and local welfare access to encourage rural households to lease out farmland (Chuang, 2014).

[18] `http://www.law-lib.com` and `http://www.pkulaw.cn`.

at least two post-reform years to the firm-level analysis. Cities reforming between 2006 and 2007 are excluded from the baseline sample because their post-reform firm outcomes cannot be consistently observed within the reliable ASIP sample period. The control group consists of 196 cities that had not implemented a *hukou* reform by the end of 2007.

4.2 Firm-level data

Firm level data between 1998-2007 is from Annual Surveys of Industrial Production (ASIP), an annual census conducted by the National Bureau of Statistics from 1992 onwards. The survey covers all state-owned manufacturing enterprises and all non-state manufacturing firms with sales exceeding RMB 5 million over that period ("above-scale" firms). Observation is the legal unit, not the plant.[19] To my knowledge, 2007 is the latest year with reliable and consistent data, as data in more recent years exhibit drawbacks such as a missing of key variables (e.g., material inputs and fixed assets) or a substantial proportion of missed firms.[20]

The data provides so far the most comprehensive information on Chinese manufacturing firms' balance sheet, statement of profit and loss, statement of cash flow, employment, location, four-digit sector, etc. The aggregated numbers for main variables is consistent with the numbers published annually in China's Statistical Yearbook. Admittedly, it is silent on firms with annual sales below the RMB 5 million threshold. Nevertheless, as Brandt et al. (2012) notes, in 2004, the firms incorporated in the sample account for 90.1% total output and 71.2% industrial workforce. [21]

As I focus on the impact of migration on urban productivity, only manufacturing firms located in urban districts (*shixiaqu*) are kept. Data cleaning and firm-matching methods follow Brandt et al. (2012). After linking firms according to their unique numerical IDs, name, industry sector, county codes, etc., I get an unbalanced panel covering 79,696 firms in 1998 and 169,881 firms in 2007. Table 1 panel B gives a comparison of firm-level descriptive statistics between the treated and the control cities. Notably, the year that a firm first appears (disappears) in the panel should be interpreted as entry (exit) with caution, as due to imperfect business registry, some firms were included by the NBS (National Bureau of Statistics of China) several years after they met the sales threshold.

Firm productivity is estimated as TFP using Ackerberg et al. (2015) methodology, which avoids the functional dependence problem existing in previous literature including Olley and Pakes (1996) and Levinsohn and Petrin (2003). A key feature of ACF method is its inversion of intermediate input demand conditional on labour, which helps tackle the vexing non-identification problem of β_l, allowing for dynamic labour as well as cross-sectional variation in labour compensations. Several studies, including Brandt et al. (2012) and He et al. (2020), apply this approach in the context of China. This study closely follows the procedure in Brandt et al. (2012) to construct key variables for TFP estimation. The heavily borrowed steps include constructing deflators for intermediate input, gross output and investment, backing out the real capital stock to extrapolate real investment, and

[19] In contrast, most countries sample plants or establishments. Auxiliary information in the ASIP suggests 96% of legal units in China are single-plant firms. The small number of firms with multiple plants and the unavailability of plant-level data prevent us from exploiting the variation in TFP across cities within the same firm, claiming that *hukou* policy is as good as randomly assigned from the firm's perspective.

[20] Brandt et al. (2014) gives a thorough discussion of other related challenges in using this data source.

[21] Large firms above-scale are firms of major interest in this study. Notwithstanding the wide coverage of this sample, the estimates might be subject to selection bias if the migrants induced by *hukou* reform has asymmetric effects on the productivity of above-scale and below-scale firms, the latter of which are missing in this sample.

calculating real value added.[22] The estimated $log(TFP)$ and industry-specific inputs coefficients in general align with the existing literature (i.e., He et al., 2020). Details on the method are discussed in the Appendix A.

4.3 City-level data

City-level urban population data for 1998–2007 are collected from the Annual Bulletins of National Economic Statistics issued by prefecture-level statistical bureaus and the China City Statistical Yearbooks (CSY) published by the NBS, which also report a range of annual socioeconomic characteristics. To ensure geographic comparability over time, I exclude city-year observations affected by changes in prefecture-level administrative boundaries. The main population measure is the number of local *hukou* holders registered in municipal districts (*shixiaqu renkou*). This measure captures registered urban population based on the place of *hukou* registration, rather than permanent population, which refers to individuals who physically reside in a city for at least six months regardless of local *hukou* status. I use registered urban population because consistent annual data on permanent population are not available at the prefecture level throughout the sample period. It also has an advantage over the number of non-agricultural *hukou* holders used by Jin and Zhang (2022), because many local reforms eliminated the distinction between agricultural and non-agricultural *hukou*. As a result, the number of non-agricultural *hukou* holders may mechanically change after reform even in the absence of actual migration, whereas *shixiaqu renkou* is based on the urban place of registration. At the same time, registered urban population is an imperfect proxy for actual urban population. Migrants may be induced by the reform to move to cities and work there before obtaining local *hukou*, so this measure is likely to understate the increase in permanent urban population. In that case, the first-stage effect of *hukou* reform on population would be attenuated, and the estimated productivity elasticity with respect to population could be upward biased. I therefore interpret the first-stage estimates as effects on registered urban population and treat them as a lower-bound proxy for broader urban population growth. Census-based permanent population in 2000 and 2005 provides a useful benchmark for assessing this measurement issue, although the lack of annual permanent population data prevents its use in the event-study analysis.

5 Empirical strategy

5.1 Diff-in-Diff

[22] As elaborated in the Appendix of Brandt et al. (2012), the output deflator series is computed by extrapolating the ratio of reported nominal and real sales in the annual NBS industrial surveys (1998-2013) to 2007 using the ex-factory price index from the 2007 CSY. The input deflator is constructed as the product of output deflator and the input shares from 2002 National Input-Output table. The investment deflator is borrowed from Perkins and Rawski (2008). In the absence of data on past net investments, the calculation of real capital stock starts from approximating the real capital stock in the first year that a firm appears in the sample using the perpetual inventory method with an annual depreciation rate of 9%.

The empirical analysis starts from evaluating the effect of *hukou* reform on logarithm urban population of city c in year t with the following stacked DID model:

$$\ln pop_{cgt} = \sum_{\tau=\underline{\tau},\tau\neq -1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau} + \lambda_c + \lambda_{g\tau} + \epsilon_{ct}, \tag{1}$$

where g denotes a cohort-specific stacked sample that includes cities implementing *hukou* reform in the same calendar year and randomly sampled control cities that did not experience a *hukou* reform between 2001 and 2013. T_{cg} is a dummy capturing the treatment status of the city c. $d_{g\tau}$ is a binary variable that takes value of 1 when city c from group g is τ years away from the introduction of *hukou* reform. $[\underline{\tau}, \overline{\tau}]$ is the sample year periods, where $\tau = 0$ implies the year of reform, $\underline{\tau} = -7$ and $\overline{\tau} = 6$. λ_c denotes city fixed effects; $\lambda_{g\tau}$ is the fixed effect of group g in its τ^{th} year relative to the *hukou* reform. Calendar year fixed effect is not included due to its collinearity with $\lambda_{g\tau}$. To alleviate concerns that cities with pre-existing population trends were more likely to reform, I use two approaches: including a vector of city-specific linear time effects $\delta_c time_t$ into the regression; and subtracting a predicted logarithm population term from $\ln pop_{cgt}$.[23] Similar de-trend processes are found in Monras (2019) and Ahlfeldt et al. (2022).

Here $\tau = -1$, the year prior to the reform, is chosen as the base for comparison, where β_{-1} is set to 0. For $\tau >= 0$, β_τ captures the percentage increase in urban population, as a measure of migration, resulting from the *hukou* reform relative to the year before reform. We expect $\beta_\tau > 0$ for $\tau >= 0$ if the reforms are effective in reducing the migration costs and attracting population inflow. Standard errors are clustered at the city level.

To interpret the effects as causal, the crux of the identification is the assumption $\mathbb{E}[T_{cg} \cdot d_{g\tau} \cdot \epsilon_{ct}|c, g\tau, t] = 0, \forall(t, \tau)$. As an explicit test, first we see if the parallel trend assumption holds pre-treatment, which suggests that in absence of the *hukou* reform, the urban population in the treated cities would have evolved the same as the control cities. Empirically, $\beta_\tau = 0$ for all $\tau < 0$ would justify this assumption. Second, $cov(\sum_{\tau=\underline{\tau},\tau\neq -1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau}, \epsilon_{ct}) = 0$ implies that local authorities' decisions about whether and when to implement reforms are orthogonal to unobserved predictors of urban population change in the error term. For instance, if the timing of reform coincides with city-specific labour demand shocks or policies that attract population inflow, then β_τ is biased upward. To test the validity of this problem, I run the regression Eq. 1 on several other outcomes of interest such as local public spending, and the concern could be alleviated empirically if $\beta_\tau = 0$ is not rejected for any $\tau < 0$ in these estimations pre-treatment. Post-reform changes in these outcomes, however, do not necessarily invalidate the design, since local public spending and public-service provision may themselves respond endogenously to population inflows.

5.2 IV Design

The predicted values for $\ln pop_{cgt}$ from Eq. 1 are further exploited as an explanatory variable to estimate the elasticity of firm productivity with respect to city density. In other words, the term $\sum_{\tau=\underline{\tau},\tau\neq -1}^{\overline{\tau}} \gamma_\tau \cdot T_{cg} \cdot d_{g\tau}$ is an arguably exogenous variation to instrument for $\ln pop_{icgt}$ in the following

[23] One concern is that institutional reforms may result from steady internal migration and are thus endogenous. Controlling for different linear evolution paths of urban population across cities helps tackle this issue.

2SLS specification:

$$2^{\text{nd}} Stage: \qquad \ln y_{icst} = \alpha \cdot \ln \widehat{pop_{icgt}} + \eta_i + \eta_{g\tau} + \eta_{st} + u_{icst}, \tag{2}$$

$$1^{\text{st}} Stage: \qquad \ln pop_{icgt} = \sum_{\tau=\underline{\tau},\, \tau\neq -1}^{\overline{\tau}} \gamma_\tau \cdot T_{cg} \cdot d_{g\tau} + \phi_i + \phi_{g\tau} + \phi_{st} + \upsilon_{icst}, \tag{3}$$

where $\ln y_{icst}$ is the logarithm of TFP of firm i in four-digit CIC industry s, city c, and year t, and η_{st} denotes industry-by-year fixed effects that absorb industry-specific shocks common across cities in a given year. Firm TFP is estimated using the methodology proposed in Ackerberg et al. (2015), as discussed in Appendix A.

The validity of staggered policy intervention as an instrumental variable depends on two crucial assumptions. First, the relevance condition requires the *hukou* reform to significantly predict urban population change in the 1st stage: $cov(\sum_{\tau=\underline{\tau},\tau\neq-1}^{\overline{\tau}} \gamma_\tau \cdot T_{cg} \cdot d_{g\tau}, \ln pop_{icgt}) \neq 0$, which is explicitly tested in Sec. 5.1. Second, the exclusion condition requires that, conditional on firm fixed effects and group-relative-time fixed effects, the timing of *hukou* reform is not correlated with unobserved shocks to firm productivity except through its effect on urban population: $cov\left(\sum_{\tau=\underline{\tau},\tau\neq-1}^{\overline{\tau}} \gamma_\tau \cdot T_{cg} \cdot d_{g\tau}, u_{ict}\right) = 0$. This exclusion restriction requires that the implementation of *hukou* reforms is not determined by the prior level or trend of firm productivity in a city. It also requires that the reform affects firms' productivity only through its impact on registered urban population, and not, for instance, through government investment in local infrastructure that directly benefits firms' production independently of population change. Additionally, *hukou* reforms should not systematically coincide with other concurrent city-level events that predict the outcomes of interest.

In support of the exclusion condition above, I first test whether *hukou* reforms have any treatment effects on firm TFP in the pre-reform periods. Second, I perform a placebo timing exercise to examine whether the productivity response is aligned with the timing of the reforms. Specifically, I repeat a static DID specification using placebo reform years ranging from seven years before to six years after the actual reform year, and plot the Wald statistic for the $treat * post$ term against each placebo reform year to see whether the statistic peaks around the actual reform year.[24]

6 Empirical Evidence

In this section, I exploit prefecture-level *hukou* reforms introduced during 2001–2005 to answer two questions. First, by improving access to urban citizenship, to what extent did *hukou* reforms promote internal migration? I use the change in urban population to capture the migration response. Second, given the observed change in urban population, what is the density elasticity of firm TFP?

6.1 Hukou reforms and migration

[24] Admittedly, this placebo timing exercise cannot rule out the possibility that other city-level events occurring at exactly the same time as the reforms affected firm TFP, such as changes in housing policies or road construction. Still, this concern is alleviated if the timing of such events did not systematically coincide with the staggered timing of prefecture-level reforms. Ideally, I would extend the sample to cover more leads and lags around the time of reform, but this is limited by the availability of annual firm-level data.

Using Eq. 1, I first estimate the effect of staggered *hukou* reforms on the internal migration flow (captured by changes in log registered urban population) and Table 3 reports the results. Panel A of Fig. 3 plots $\hat{\beta}_\tau$ and the corresponding 95% confidence intervals under the baseline specification without accounting for city-specific time trends. Registered urban population increases by around 1.26% in the first post-reform year relative to the year immediately preceding the reform ($[e^{0.0125}-1]\times 100\% = 1.26\%$). The estimated effect continues to increase through the fourth post-reform year, reaching approximately 4.02% ($[e^{0.0395}-1]\times 100\% = 4.02\%$). Though there are no significant differential pre-treatment trends between the reformed and the control cities, in Panel B and C I use two alternative specifications to control for their possibly different pre-existing trends. In Panel B, a city-specific linear trend term $\delta_c time_t$ is incorporated into Eq. 1. In Panel C, I first estimate city-specific linear time trends using not-yet-treated observations for cities with at least three observed pre-treatment years, and then remove the pre-existing trends from the whole time series before running a regression analogous to Eq. 1 using de-trended population data. In the de-trended specification reported in Column (4), which I treat as the preferred specification, registered urban population is estimated to be approximately 4.61% higher five years after the reform ($[e^{0.0451}-1]\times 100\% = 4.61\%$). The overall results are robust in showing a gradual post-reform growth in registered urban population.

As a complement to the above "dynamic" specifications, I use two alternative static specifications. First, following Gollin et al. (2021), I replace $\sum_{\tau=\underline{\tau},\tau\neq -1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau}$ in Eq. 1 with $\beta \cdot T_{cg} \cdot \mathbb{1}\{\tau \geq 0\} \cdot \tau$ to capture a linear trend break after the reforms[25]. Notably, $T_{cg} \cdot \mathbb{1}\{\tau \geq 0\}$ equals to one only for actual post-reform observations and takes the value of zero for control cities throughout the sample period. Under the assumption that in absence of the reforms, urban population in treated cities would have exhibited the same trend as that in control cites, coefficient β tells the annual rate at which population increase resulting from the *hukou* reforms. The estimates are reported in Column (1)-(2) of Table 4. Second, to estimate the average treatment effect, I replace $\sum_{\tau=\underline{\tau},\tau\neq -1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau}$ with $\theta \cdot T_{cg} \cdot Post_{g\tau}$, where $Post_{g\tau}$ is a post-reform indicator. Column (3)-(4) in Table 4 report the corresponding results. Though silent on the underlying dynamics, the static specifications give more precise estimates than Eq. 1; and moreover, if the dynamic policy instruments are weak in the 2SLS estimation, a just-identified IV (i.e., $T_{cg} \cdot \mathbb{1}\{\tau \geq 0\} \cdot \tau$ or $T_{cg} \cdot Post_{g\tau}$) is still *approximately* consistent as long as the first-stage coefficient is not zero (Angrist and Pischke, 2009a,b).

To support the claim that the observed changes in urban population were not driven by pre-existing city-level shocks or policies correlated with the timing of *hukou* reforms, I estimate Eq. 1 on several other local outcomes, including local government expenditure, the number of primary schools, the number of high schools, and the number of hospitals.[26] Table C.1 presents the results. I do not detect significant differential pre-trends in these variables before the reforms, which alleviates concerns that treated cities were already on different local policy or public-service trajectories. Post-reform changes in these outcomes should be interpreted with caution, since local spending and public-service provision may themselves respond endogenously to population inflows. Overall, these checks are therefore best viewed as evidence against differential pre-treatment trends rather than as definitive placebo tests.

6.2 Density effect on productivity

[25] The assumed trend break is not taken for granted: it is discernible in Fig. 3.

[26] China implements nine-year pre-college compulsory education, where local education bureaus supervise primary schools and high schools.

First, Fig. 4 presents the reduced-form results for the impact of *hukou* reform on TFP in our 2SLS setting. The post-treatment increase in firm TFP suggests that the relaxation of migration barriers led to agglomeration benefits. Moreover, the insignificant pre-treatment coefficients allow me to fail to reject the parallel trend assumption on the TFP of firms in reformed and non-reformed cities in years preceding the treatment. Appendix Fig. C.2 provides complementary evidence by restricting the sample to firms observed annually throughout 1998–2007. The post-reform increase in TFP remains, particularly among private firms, suggesting that the baseline pattern is not driven solely by changes in firm composition.

Next, Table 5 presents the OLS and IV estimates of the effect of density on TFP, from which we see an overall statistically significant effect of density on TFP in Column (1) – 1% increase in urban population is associated with around 0.9% higher TFP. While we observe a positive effect on firms of all types of ownership except for foreign firms, the effect is most pronounced for private firms in Column (4). This may reflect private firms' greater flexibility in responding to changes in local labour supply and supplier access. Compared with state-owned and collectively-owned firms, private firms may adjust hiring, input mix, and production organization more quickly after *hukou* reforms.[27] The IV point estimate suggests that 1% increase in urban population led to about 2.2% higher TFP of private firms, an elasticity larger than most existing literature.[28] The absence of a significant effect on foreign firms' TFP might be explained by the possibility that the reforms primarily increased the mobility of low-skilled rural workers while foreign firms tend to hire higher-skilled workers. Table 6 exhibits similar patterns using the other two instruments constructed before in Section 6.1.

To understand the dynamics behind this large density effect on private firms' TFP, Fig. 5 provides suggestive evidence on firms' adjustment in input use following the *hukou* reforms. Echoing Imbert et al. (2022), the capital-to-labour ratio declined after the reform, suggesting a shift toward more labour-intensive production, while the intermediate-input-to-labour ratio appears to have increased. Employment among firms already observed in the ASIP sample, however, shows little systematic response. Appendix Fig. C.1 documents a suggestive increase in the number of above-scale manufacturing firms observed in the ASIP sample. This combination is consistent with manufacturing activity being distributed across a larger set of observed firms rather than appearing solely as employment growth among incumbent sampled firms. However, the increase in observed firm counts cannot be interpreted as true firm entry and only serves as indicative evidence, since it may also reflect existing firms expanding sales and crossing the ASIP size threshold, as well as changes in survey coverage.

Both value added per worker and real wage per worker increased following the *hukou*-induced labour-supply shock. This pattern is consistent with higher labour productivity, potentially arising from improved worker–firm matching, stronger labour-market competition, and better access to intermediate inputs. One caveat is that ASIP may not fully capture informal or dispatch workers employed

[27] That Kleibergen Paap rk LM p-value in Table 5 is not significantly low induces unidentification concern, but this is no longer a concern if I use post-reform indicator or trend-break indicator as just-identified IVs. Besides, the low KP rk Wald F-statistics below the Stock & Yogo (2002) 10% maximal bias threshold signals weak IV problem. In light of this, in both Table 5 and Table 6 I report the results of Anderson-Rubin (1949) Wald test, which is robust to weak identification procedure. The test reassuringly rejects the null that the coefficients on the excluded instruments are jointly zero if they are incorporated in place of density in the outcome equation. Moreover, estimates with weak IV are biased toward the OLS estimates (Angrist and Pischke, 2009a), which suggests that the effects of density are likely to be underestimated rather than overestimated given the OLS estimates.

[28] As Moretti (2021) comments, the great variation in the estimates of density elasticity is partly due to the fact that the models, time periods and data selected vary vastly. My estimates echo Greenstone et al. (2010), who also use quasi-experiment method and imply an elasticity in the range of 1.25-3.1.

by formal firms.[29] If the reforms increased the use of workers omitted from reported employment, measured labour input could be understated and estimated TFP effects biased upward. Such bias would, however, require the reforms to generate a differential increase in labour underreporting in treated cities relative to control cities; the general prevalence of informal employment alone would not generate the estimated reform effect. I therefore interpret the TFP estimates with appropriate caution.

6.3 Robustness

As discussed in 5.2, to justify the exogeneity of *hukou* reforms to local productivity, I want to show that the observed changes in TFP were not driven by other concurrent city-level time trends. To test this, I perform a placebo analysis by repeating a static DID specification with placebo reform years ranging between 7 years prior and 6 years posterior to the *de facto* reform years. Fig. 6 plots the Wald statistics related to the $treat * post$ indicator of these regressions against the placebo reform year relative to the true reform year, from which we do find the largest Wald statistic near the actual reform year. Taken together, the lack of differential pre-trends in Fig. 4, the expected peak in Wald Statistics, and the insignificant responses from other local variables in Table C.1 add support to the exogenous nature of *hukou* reforms to urban population and local productivity.

7 Conclusion

Internal migration barriers can generate labour misallocation and limit agglomeration benefits in many developing countries. This paper asks whether relaxing such barriers increases firm productivity by expanding urban labour markets. Using staggered *hukou* reforms in Chinese prefectures, I find that migration liberalization increased registered urban population and raised firm TFP, especially among domestic private firms. The accompanying increase in value added per worker and real wages, together with the decline in the capital-labour ratio, suggests that firms adjusted their production structure in response to an expanded urban labour market. The pattern is consistent with migration liberalization improving productivity through labour pooling, worker-firm matching, and input-market thickening, with stronger effects among private firms that may respond more flexibly to local labour-market changes. The policy implication is therefore nuanced. Relaxing internal migration barriers can raise productivity by expanding urban labour markets and improving worker-firm matching. At the same time, the decline in the capital-labour ratio suggests that firms may also respond by adopting more labour-intensive production. The simultaneous increases in TFP, value added per worker, and wages suggest that, in this setting, agglomeration gains dominate this adjustment margin, particularly among domestic private firms.

This study has several limitations. First, the use of ASIP data restricts the analysis to above-scale manufacturing firms, so the generalizability of the results to smaller firms or firms in other sectors

[29] According to Zhaoyun Xue, the former vice chairman of the All-China Federation of Trade Unions, "informal employment" in China refers not only to employment in informal sectors, but also to informal employment in formal sectors, such as temporary workers, part-time workers and labour dispatch workers (Shaoyun, 2000). For example, many firms signed labour contracts with labour dispatch agencies rather than with workers to reduce labour costs and evade social insurance obligations. One national sample survey in 2004 shows that the labour contract signing rate of migrant workers was only 12.5% (Council, 2006). Low contract coverage, however, does not by itself imply that these workers were excluded from firms' reported employment.

remains uncertain. It is also difficult to fully capture selection through firm entry and exit. Second, the main population measure captures registered urban population rather than permanent population. Since migrants may work in cities before obtaining local hukou, the first-stage population response may understate broader urban population growth. Third, ASIP may not fully capture informal or dispatch workers employed by formal firms. If *hukou* reforms increased such employment, measured labour input may be understated and estimated TFP effects may be biased upward. Fourth, the absence of annual migrant-level data prevents heterogeneity analysis by migrants' characteristics, such as skill level and age. Finally, the reduced-form approach is silent on aggregate general-equilibrium effects and on heterogeneous treatment effects across industries. These limitations point to several directions for future research.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

Ahlfeldt, G. M. and Feddersen, A. (2018). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, 18(2):355–390.

Ahlfeldt, G. M. and Pietrostefani, E. (2019). The economic effects of density: A synthesis. *Journal of Urban Economics*, 111:93–107.

Ahlfeldt, G. M., Roth, D., and Seidel, T. (2022). Optimal minimum wages.

Angrist, J. D. and Pischke, J.-S. (2009a). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Angrist, J. D. and Pischke, J.-S. (2009b). A note on bias in just identified iv with weak instruments. *London School of Economics*, 28.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Arrow, K. J. (1971). *The economic implications of learning by doing*. Springer.

Au, C.-C. and Henderson, J. V. (2006). How migration restrictions limit agglomeration and productivity in china. *Journal of development economics*, 80(2):350–388.

Berliant, M., Reed III, R. R., and Wang, P. (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics*, 60(1):69–95.

Bond, S. and Söderbom, M. (2005). Adjustment costs and the identification of cobb douglas production functions. Technical report, IFS Working Papers.

Brandt, L., Van Biesebroeck, J., Wang, L., and Zhang, Y. (2017). Wto accession and performance of chinese manufacturing firms. *American Economic Review*, 107(9):2784–2820.

Brandt, L., Van Biesebroeck, J., and Zhang, Y. (2012). Creative accounting or creative destruction? firm-level productivity growth in chinese manufacturing. *Journal of development economics*, 97(2):339–351.

Brandt, L., Van Biesebroeck, J., and Zhang, Y. (2014). Challenges of working with the chinese nbs firm-level data. *China Economic Review*, 30:339–352.

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Bustos, P., Castro-Vincenzi, J. M., Monras, J., and Ponticelli, J. (2019). Industrialization without innovation. Technical report, National Bureau of Economic Research.

Cainelli, G. and Ganau, R. (2018). Distance-based agglomeration externalities and neighbouring firms' characteristics. *Regional Studies*, 52(7):922–933.

Chan, K. W. and Zhang, L. (1999). The hukou system and rural-urban migration in china: Processes and changes. *The China Quarterly*, 160:818–855.

Chuang, J. (2014). China's rural land politics: bureaucratic absorption and the muting of rightful resistance. *The China Quarterly*, 219:649–669.

Ciccone, A. and Hall, R. E. (1993). Productivity and the density of economic activity.

Cingano, F. and Schivardi, F. (2004). Identifying the sources of local productivity growth. *Journal of the European Economic association*, 2(4):720–742.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. In *Agglomeration economics*, pages 15–66. University of Chicago Press.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics*, volume 5, pages 247–348. Elsevier.

Council, C. (2006). *Research Report on Migrant Workers in China.* China Yanshi Publishing House.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2022). Matching in cities. *Journal of the European Economic Association*, 20(4):1478–1521.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–48.

Eberts, R. W. and McMillen, D. P. (1999). Agglomeration economies and urban public infrastructure. *Handbook of regional and urban economics*, 3:1455–1495.

Gandhi, A., Navarro, S., and Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6):1126–1152.

Gollin, D., Hansen, C. W., and Wingender, A. M. (2021). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, 129(8):2344–2384.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of political economy*, 118(3):536–598.

Gu, Y., Jiang, C., Zhang, J., and Zou, B. (2021). Subways and road congestion. *American Economic Journal: Applied Economics*, 13(2):83–115.

He, G., Wang, S., and Zhang, B. (2020). Watering down environmental regulation in china. *The Quarterly Journal of Economics*, 135(4):2135–2185.

Henderson, J. V., Quigley, J., and Lim, E. (2009). Urbanization in china: Policy issues and options. *Unpublished manuscript, Brown University.*

Holl, A. (2016). Highways and productivity in manufacturing firms. *Journal of Urban Economics*, 93:131–151.

Hu, C., Xu, Z., and Yashiro, N. (2015). Agglomeration and productivity in china: Firm level evidence. *China Economic Review*, 33:50–66.

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2022). Migrants and firms: Evidence from china. *American Economic Review*, 112(6):1885–1914.

Jacobs, J. (1969). The economy of cities. Technical report, Vintage Books: A Division of Random House.

Jin, Z. and Zhang, J. (2022). Access to local citizenship and internal migration in a developing country: Evidence from a hukou reform in china. *Journal of Comparative Economics.*

Kim, S. (1991). Heterogeneity of labor markets and city size in an open spatial economy. *Regional Science and Urban Economics*, 21(1):109–126.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies*, 70(2):317–341.

Ma, L. and Tang, Y. (2020). Geography, trade, and internal migration in china. *Journal of Urban Economics*, 115:103181.

Marshall, A. (1890). The principles of economics. Technical report, McMaster University Archive for the History of Economic Thought.

Martin, P., Mayer, T., and Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. *Journal of urban Economics*, 69(2):182–195.

Meng, X. (2012). Labor market outcomes and reforms in china. *Journal of Economic Perspectives*, 26(4):75–102.

Monras, J. (2019). Minimum wages and spatial equilibrium: Theory and evidence. *Journal of Labor Economics*, 37(3):853–904.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American*

Economic Review, 111(10):3328–3375.
Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment. *Econometrica*, 64(6):1263–1297.
Ottaviano, G. I., Peri, G., and Wright, G. C. (2018). Immigration, trade and productivity in services: Evidence from uk firms. *Journal of International Economics*, 112:88–108.
Peri, G. (2012). The effect of immigration on productivity: Evidence from us states. *Review of Economics and Statistics*, 94(1):348–358.
Perkins, D. H. and Rawski, T. G. (2008). Forecasting china's economic growth to 2025.
Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of urban economics*, 34(3):380–400.
Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.
Redding, S. J. and Sturm, D. M. (2008). The costs of remoteness: Evidence from german division and reunification. *American Economic Review*, 98(5):1766–1797.
Rizov, M., Oskam, A., and Walsh, P. (2012). Is there a limit to agglomeration? evidence from productivity of dutch firms. *Regional Science and Urban Economics*, 42(4):595–606.
Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, 94(5):1002–1037.
Rosenthal, S. S. and Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2):373–389.
Shaoyun, X. (2000). Several understandings and suggestions on developing informal sector and encouraging informal employment in my country. *"Seminar on Informal Sector Employment", Ministry of Labor and Social Security of the People's Republic of China, Beijing.*
Tao, J., Ho, C.-Y., Luo, S., and Sheng, Y. (2019). Agglomeration economies in creative industries. *Regional Science and Urban Economics*, 77:141–154.
Tian, B. (2003). Zhongguo diyi zhengjian (china's no. 1 document). *Huizhou, China: Guangdong Renmin Chubanshe.*
Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–72.
Zhan, S. (2017). Hukou reform and land politics in china: Rise of a tripartite alliance. *The China Journal*, 78(1):25–49.

Tables

Table 1: Summary Statistics of the baseline year 1998

	Treated Cities				*Control Cities*			
	mean	p25	p50	p75	mean	p25	p50	p75
Panel A. City-level characteristics								
Population (million)	4.35	2.55	4.13	5.75	3.59	1.93	3.14	4.64
Urban population (million)	1.14	0.50	0.89	1.51	0.93	0.46	0.72	1.14
GDP per capita (thousand RMB)	11.77	6.39	8.68	12.74	6.96	4.30	5.61	7.74
Hospital beds (thousand)	11.28	6.48	10.01	14.45	9.00	5.14	7.93	11.29
Government budget (trillion RMB)	2.65	0.83	1.49	2.98	1.46	1.00	1.36	1.71
Urbanization (%)	35.34	15.58	26.00	45.96	33.33	15.62	26.10	43.96
Share of 1st industry (%)	17.64	10.80	15.50	23.20	24.28	16.15	25.40	31.95
Share of 2rd industry (%)	46.95	43.20	46.15	52.40	41.94	33.10	40.95	49.75
Share of 3rd industry (%)	35.57	31.60	35.30	40.00	33.79	29.80	33.35	37.80
GDP growth (%)	9.89	9.35	10.90	13.10	9.00	6.70	9.70	11.90
Panel B. Firm-level variables								
Log(TFP)	2.83	2.12	2.93	3.62	2.52	1.76	2.64	3.41
Real output (thousand)	55.71	7.85	15.83	39.18	49.68	5.84	11.78	28.10
Intermediate input (thousand)	42.56	5.73	11.81	29.46	37.74	4.18	8.78	21.37
Value added (thousand)	13.11	1.56	3.43	8.84	11.86	1.07	2.45	6.16
Employment	415.56	78.00	163.00	373.50	457.18	67.00	154.00	377.00
Real capital stock (thousand)	41.39	2.27	6.68	21.51	49.66	1.70	5.38	18.92

Notes: Data in Panel A are from Chinese Statistical Yearbooks. Panel B displays the variables of all manufacturing firms in the treated and control cities respectively (unbalanced panel). The full city estimation sample contains 76 treated cities and 196 control cities. Because the city panel is unbalanced, this table reports the 1998 cross-section only, for which data are available for 64 treated cities and 157 control cities. Firm-level data are from Annual Surveys of Industrial Production (ASIP), with which several variables are calculated following Brandt et al. (2012). Log(TFP) is calculated using ACF method.

Table 2: Change in Log(TFP) over the sample period

	1 year pre-reform	1 year post-reform	Change
Panel A. All firms			
Reformed cities	3.18	3.35	+0.17
Control cities	3.19	3.29	+0.10
Panel B. Firms active for 10 years			
Reformed cities	3.22	3.33	+0.11
Control cities	3.18	3.25	+0.07

Notes: A comparison of mean firm logarithm TFP in reformed and non-reformed cities in 1st year pre-reform and 1st year post-reform. The placebo treatment year is chosen as 2005 for control cities. Panel B displays the change for firms that remained active throughout the sample period 1998–2007. Firm-level data are from ASIP, and TFP is calculated using the ACF method from production functions estimated separately for each 2-digit CIC industry. In the final specification, log urban population is incorporated as a state variable in the sector-specific production-function estimation.

Table 3: Hukou reform effects on logarithm urban population under dynamic specification

	Two-way FE DID	Stacked DID		
	(1)	(2)	(3)	(4)
7 years pre	0.0061	0.0103		0.0046
	(0.0066)	(0.0165)		(0.0034)
6 years pre	0.0091	0.0089	0.0017	0.0043
	(0.0064)	(0.0117)	(0.0058)	(0.0030)
5 years pre	0.0048	-0.0031	0.0046	0.0014
	(0.0064)	(0.0091)	(0.0062)	(0.0030)
4 years pre	0.0074	-0.0000	0.0037	-0.0001
	(0.0057)	(0.0068)	(0.0054)	(0.0025)
3 years pre	0.0071	-0.0024	0.0024	-0.0007
	(0.0055)	(0.0045)	(0.0040)	(0.0019)
2 years pre	0.0073	-0.0021	-0.0007	-0.0025
	(0.0050)	(0.0023)	(0.0022)	(0.0016)
Reform year	0.0140***	0.0077**	0.0062*	0.0080**
	(0.0053)	(0.0038)	(0.0035)	(0.0037)
1 year post	0.0196***	0.0125**	0.0102**	0.0134***
	(0.0068)	(0.0057)	(0.0049)	(0.0049)
2 year post	0.0270***	0.0192**	0.0155**	0.0211***
	(0.0088)	(0.0079)	(0.0072)	(0.0067)
3 year post	0.0326***	0.0277***	0.0189**	0.0289***
	(0.0112)	(0.0106)	(0.0090)	(0.0086)
4 year post	0.0400**	0.0395***	0.0216*	0.0390***
	(0.0156)	(0.0140)	(0.0111)	(0.0118)
5 year post	0.0320	0.0302	0.0216	0.0451***
	(0.0204)	(0.0194)	(0.0136)	(0.0132)
6 year post	0.0482	0.0415	0.0161	0.0552**
	(0.0344)	(0.0344)	(0.0207)	(0.0240)
General time effects	Yes	No	No	No
Group-specific time effects	No	Yes	Yes	Yes
City-specific time trend	No	No	Yes	No
SE clustered by city	Yes	Yes	Yes	Yes
Obs	2491	2491	2491	2337
R2	0.9978	0.9978	0.9996	0.3975

Notes: Prefecture-level regressions covering 272 cities with available data over 1998–2007 (unbalanced panel) using dynamic specifications. The dependent variable is logarithm urban population except for column (4). Baseline year is chosen as the 1st year before the introduction of hukou policy. Column (1) gives results under traditional two-way fixed effects model. Columns (2)-(4) adopt a stacked DID design, allowing a group-specific time effect, where the group is set based on treated year. Column (2) is baseline specification in Eq. 1. Column (3) allow heterogenous time effects across cities by adding a city-specific linear time trend term $\delta_c time_t$ and dropping leads 1 and 7. The dependent variable in column (4) is detrended log population $\ln pop_{cgt} - \ln \widehat{pop_{cgt}}$ (detrended based on pre-reform data).
Standard errors clustered at city in parentheses.
**** p<.01, ** p<.05, * p<.1*

Table 4: Hukou reform effects on logarithm urban population under two static specifications

	Static version 1		Static version 2	
	(1)	(2)	(3)	(4)
Treated * Treated Years	0.0061***	0.0075***		
	(0.0017)	(0.0019)		
Treated * Post			0.0170***	0.0205***
			(0.0048)	(0.0059)
General time effects	Yes	No	Yes	No
Group-specific time effects	No	Yes	No	Yes
SE clustered by city	Yes	Yes	Yes	Yes
Obs	2337	2337	2337	2337

Notes: Prefecture-level regressions covering 272 cities with available data over 1998–2007 (unbalanced panel) using static specifications. The dependent variable is de-trended logarithm urban population (deviation from predicted log population calculated with pre-reform data). Column (1)-(2) assume that there is a break in population trend after the reform, report the estimates when $\sum_{\tau=\underline{\tau},\tau\neq-1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau}$ in Eq. 1 is replaced with $\beta \cdot T_{cg} \cdot \mathbb{1}\{\tau \geq 0\} \cdot \tau$. Columns (3)-(4) estimate the average treatment effect θ, replacing $\sum_{\tau=\underline{\tau},\tau\neq-1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau}$ with $\theta \cdot T_{cg} \cdot Post_{g\tau}$.
Standard errors clustered at city in parentheses.
**** p<.01, ** p<.05, * p<.1*

Table 5: Estimation of density effect on TFP

	All	State	Collective	Private	Foreign
	(1)	(2)	(3)	(4)	(5)
Panel A. OLS estimates					
ln(urban pop)	0.4048*	0.5270***	0.7069***	0.6931*	0.1190**
	(0.2140)	(0.1322)	(0.1577)	(0.3821)	(0.0553)
Obs	650358	100827	102824	147982	107118
Panel B. Baseline IV estimates					
ln(urban pop)	0.8818***	0.6260**	1.1731***	2.2006**	0.0877
	(0.3264)	(0.2473)	(0.3003)	(0.9180)	(0.1532)
KP rk LM statistic	15.804	23.567	26.255	17.184	18.595
KP rk LM p-value	[0.260]	[0.035]	[0.016]	[0.191]	[0.136]
KP rk Wald F-statistics	2.412	4.717	7.953	1.871	14.902
Anderson-Rubin chi-sq	23.108	35.358	73.111	40.977	59.295
Anderson-Rubin Wald p-value	[0.040]	[0.001]	[0.000]	[0.000]	[0.000]
Obs	650358	100827	102824	147982	107118

Notes: Firm TFP on de-trended logarithm urban population by pre-reform ownership. Panel A displays OLS estimates. Panel B instruments de-trended log urban population with dynamic hukou reform shocks. Column (1) covers all manufacturing firms in urban areas of prefectural cities; Columns (2)–(5) classify each firm using its ownership in the last observed year before its city's actual or pseudo reform year. Both OLS and IV specifications incorporate firm FE, cohort-specific time effects, and industry-by-year fixed effects. See Section 5.2 for details on IV specification.
Standard errors in parentheses are clustered by city.
**** p<.01, ** p<.05, * p<.1*

Table 6: Estimation of density effect on TFP

	All	State	Collective	Private	Foreign
	(1)	(2)	(3)	(4)	(5)
Panel A. Alternative IV: post-reform indicator					
ln(urban pop)	0.9530**	0.5335	1.2764***	3.2177*	0.0865
	(0.4065)	(0.3825)	(0.3497)	(1.9118)	(0.1862)
KP rk LM statistic	3.037	4.787	3.389	2.845	2.686
KP rk LM p-value	[0.081]	[0.029]	[0.066]	[0.092]	[0.101]
KP rk Wald F-statistics	4.345	7.385	7.661	2.815	7.222
Anderson-Rubin chi-sq	8.294	1.282	12.558	16.503	0.200
Anderson-Rubin Wald p-value	[0.004]	[0.258]	[0.000]	[0.000]	[0.655]
Obs	650358	100827	102824	147982	107118
Panel B. Alternative IV: trend-break indicator					
ln(urban pop)	1.0198***	0.6136**	1.3115***	2.5988**	0.1341
	(0.3755)	(0.2638)	(0.3462)	(1.1325)	(0.1601)
KP rk LM statistic	3.240	4.865	3.283	3.549	2.821
KP rk LM p-value	[0.072]	[0.027]	[0.070]	[0.060]	[0.093]
KP rk Wald F-statistics	4.791	7.968	7.405	3.907	7.703
Anderson-Rubin chi-sq	8.317	2.795	11.583	11.580	0.600
Anderson-Rubin Wald p-value	[0.004]	[0.095]	[0.001]	[0.001]	[0.439]
Obs	650358	100827	102824	147982	107118

Notes: Alternative IV estimations of firm TFP on de-trended logarithm urban population by pre-reform ownership. Panel A instruments logarithm urban population with post-reform indicator. Panel B instruments logarithm urban population with a break in trend term as specified in. Column (1) covers all manufacturing firms in urban areas (shixiaqu) of prefectural cities; Columns (2)–(5) classify each firm using its ownership in the last observed year before its city's actual or pseudo reform year. Both OLS and IV specifications incorporate firm FE, cohort-specific time effects, and industry-by-year fixed effects. See Section 5.2 for details on IV specification.
Standard errors in parentheses are clustered by city.
**** p<.01, ** p<.05, * p<.1*

FIGURES

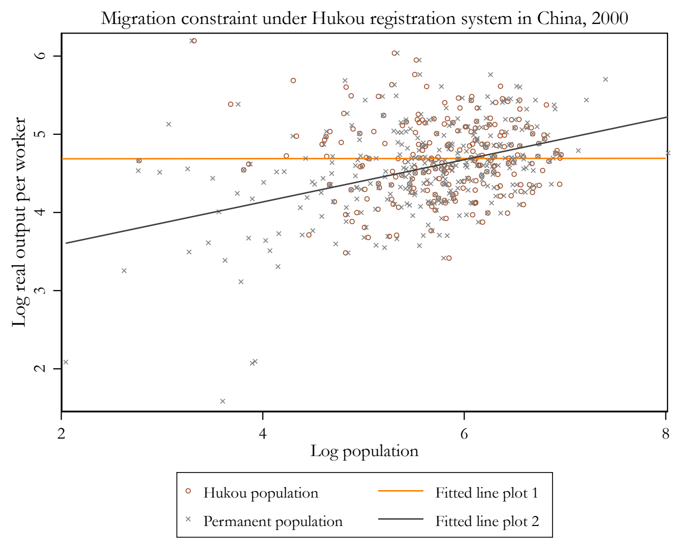
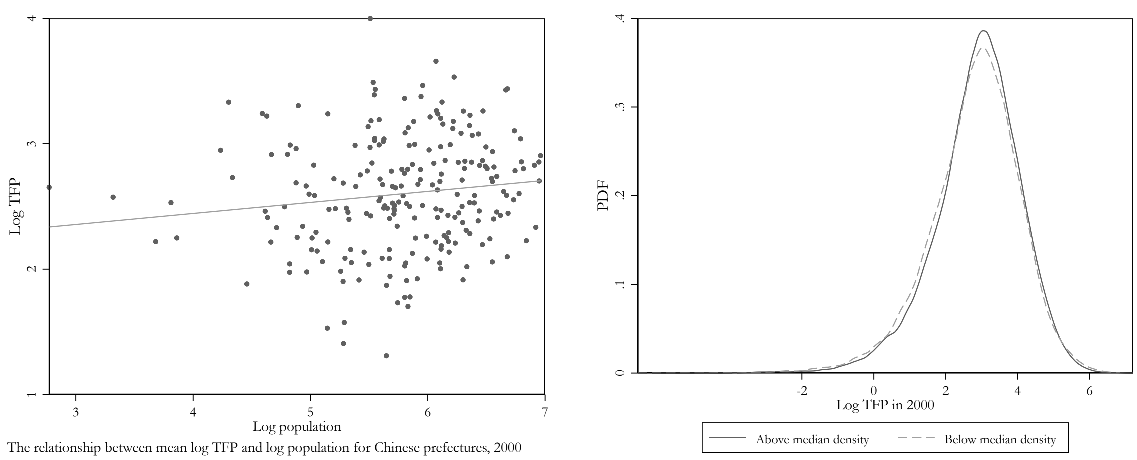


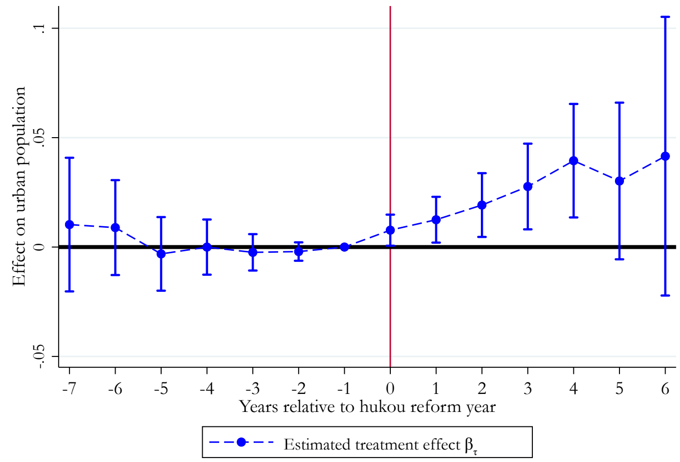
Figure 1: Relationship between the logarithm real output per worker and the logarithm population across Chinese prefecture-level cities as of 2000. Population measured in 2 ways: permanent population (physically locate in a city for $\geq$ 6 months) and *Hukou* population (with local registered hukou status). *Note:* Real output per worker in a city is calculated as the the ratio between the aggregated real output of manufacturing firms and the total number of manufacturing employment in ASIP dataset. Data on *Hukou* population is from CSY, while data on permanent population is from the 5$^{\text{th}}$ population census. Two fitted lines are plotted for each group.



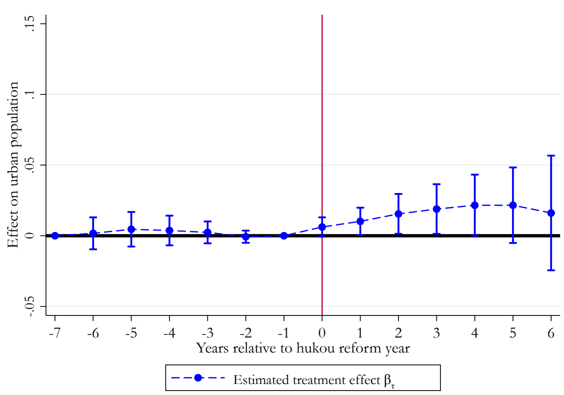
(a) (b)

Figure 2: The density economies of large cities. (a) plots the relationship between mean logarithm firm TFP and logarithm permanent population among Chinese prefectures in 2000. (b) displays the distribution of mean logarithm firm TFP for cities above (solid) versus below (dashed) median density. Data from ASIP and CSY. TFP is computed with ACF method.

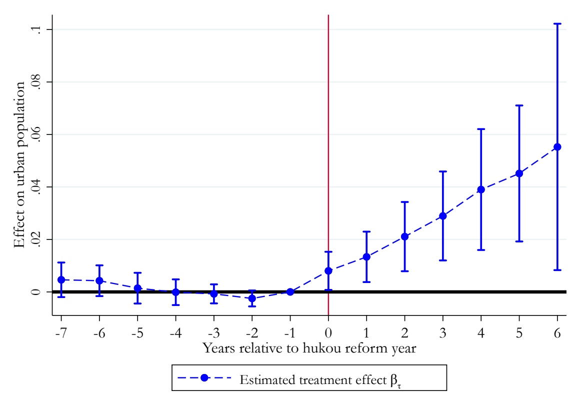
Figure 3: Dynamic effects of *hukou* reform on urban population



Panel A: Stacked DID without city specific trends



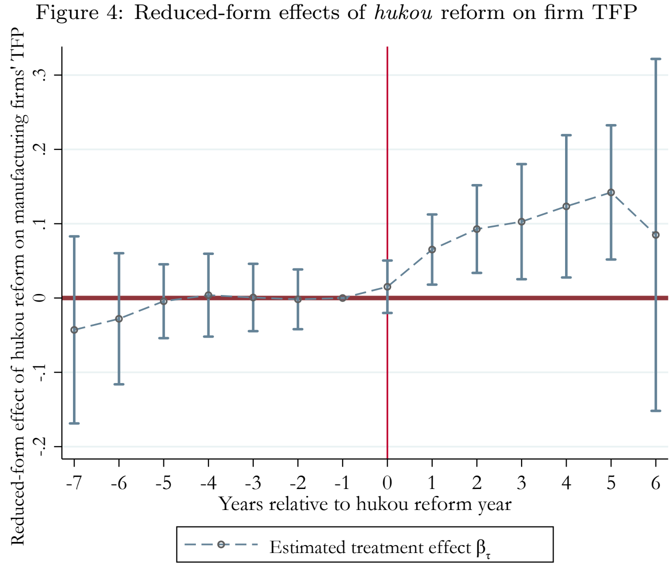
Panel B: Including city-specific trends



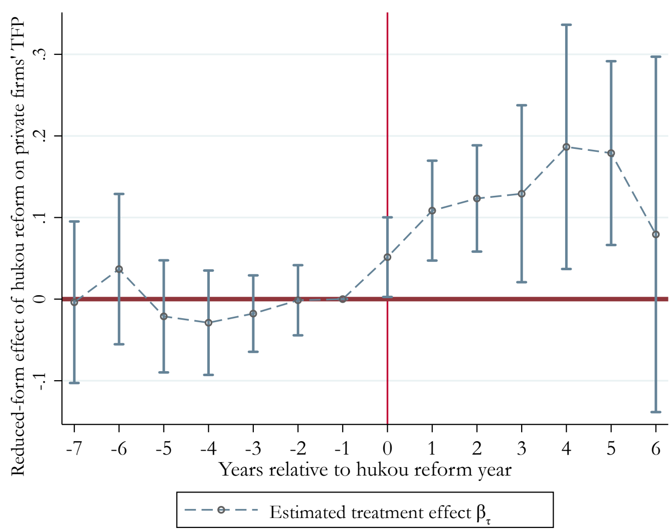
Panel C: Detrend using pre-reform data

Note: Panel A is estimated using Eq. 1, where city-specific linear trend term $\delta_c time_t$ is not incorporated; Panel B is estimated using Eq. 1 with $\delta_c time_t$, dropping leads 1 and 7. Panel C is estimated with the following procedure to purge out city pre-existing linear trend. First, for each city, estimate a linear time trend using years before *hukou* reform: $\ln pop_{cgt} = \delta_c time_t + \epsilon_{ct}$. Then predict $\widehat{\ln pop_{cgt}}$, detrend the whole time series (including years after the reform) for each city and estimate: $\ln pop_{cgt} - \widehat{\ln pop_{cgt}} = \sum_{\tau=\underline{\tau},\tau\neq-1}^{\overline{\tau}} \beta_\tau \cdot T_{cg} \cdot d_{g\tau} + \lambda_c + \lambda_{g\tau} + \epsilon_{ct}$. Similar procedure is used in Monras (2019) and Ahlfeldt et al. (2022).

Figure 4: Reduced-form effects of *hukou* reform on firm TFP



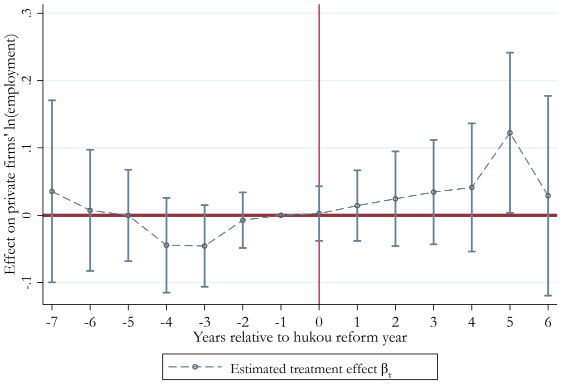
Panel A: For all firms



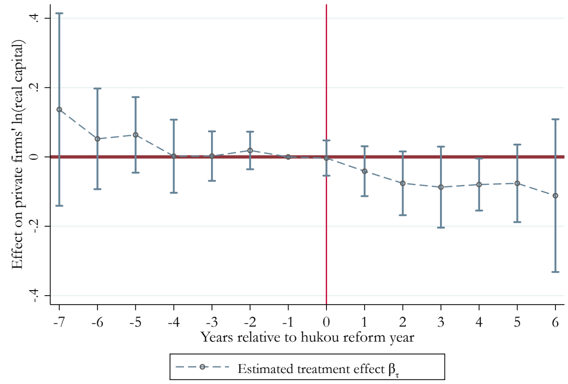
Panel B: For private firms

Note: The reduced-form estimation of the effect of *hukou* reforms on firms' TFP using 2SLS specifications in Section 5.2. Panel A plots the effect for all domestic manufacturing firms during 1998-2007; Panel B plots the effect for private firms only (as table 5 only shows a significant effect on private firms).

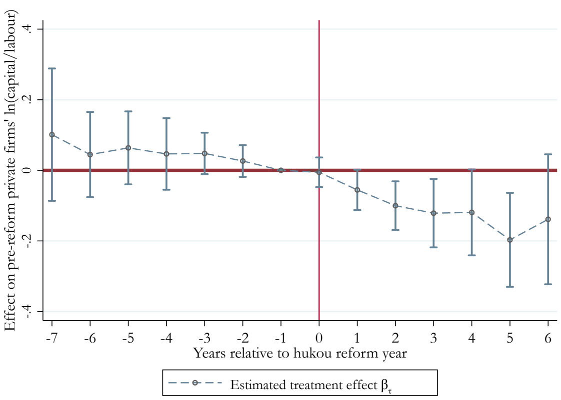
Figure 5: Dynamic effects of *hukou* reform on private manufacturing firms' input mix



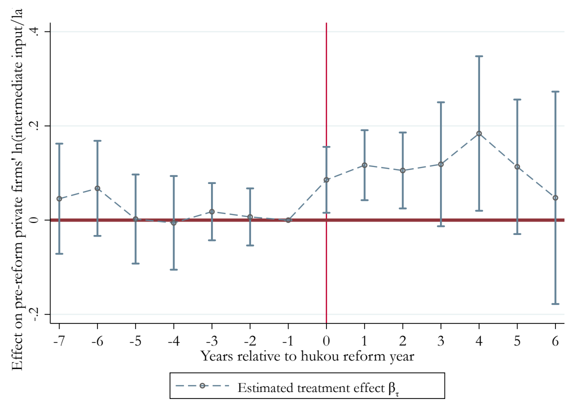
Panel A: Employment



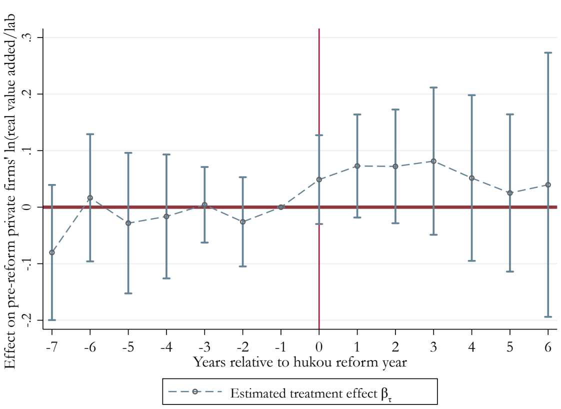
Panel B: Real capital

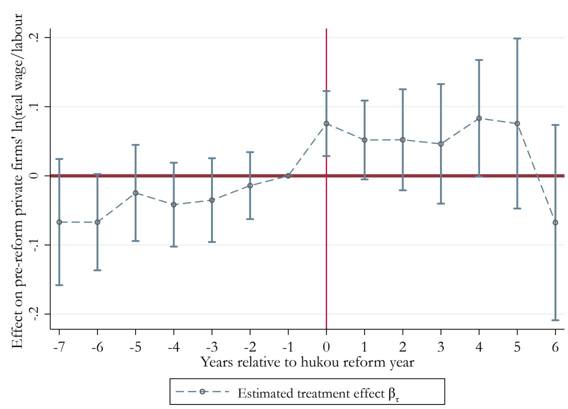


Panel C: Capital to labour ratio



Panel D: Intermediate inputs to labour ratio



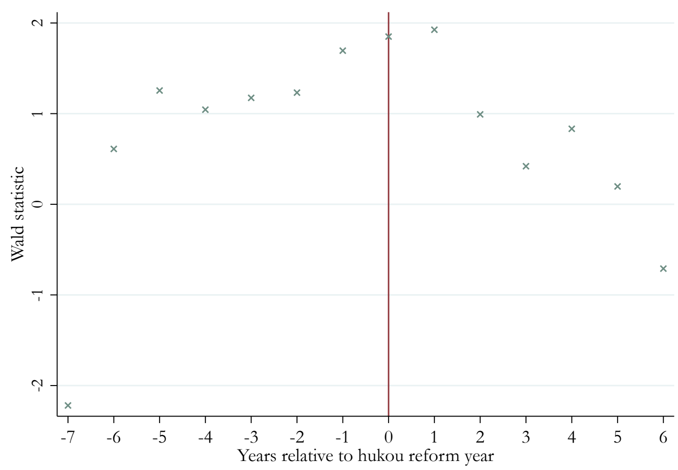


Panel E: Value-added to labour ratio

Panel F: Real wage to labour ratio

Note: Reduced-form estimates of the effect of *hukou* reforms on private manufacturing firms' input mix. All variables are in logarithm form. Reduced-form regressions account for cohort-specific time effects, industry-specific time effects, and firm fixed effects. All regressions are clustered by city.

Figure 6: Wald Statistics from Estimates with Placebo Reform Years



Note: Each dot corresponds to a Wald statistic from the t-test of the key coefficient in a static DID specification where a placebo reform year is chosen. The x-axis suggests the placebo reform year relative to the actual reform year.

Appendix A: Production function estimation and TFP

Estimation strategy following ACF

Assuming a Cobb-Douglas production, the *added value* above material costs is a function of capital k_{it}, labour l_{it}, productivity ω_{it} unobserved by econometricians but observed by the firm at time t, and shocks ε_{it} unobserved by both at time t:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \varepsilon_{it} \tag{A.1}$$

Without data on firm-specific output and input prices, I rewrite this with nominal variables as:

$$\tilde{y}_{it} = \beta_0 + \beta_k \tilde{k}_{it} + \beta_l l_{it} + \tilde{\omega}_{it} + \varepsilon_{it}, \tag{A.2}$$

where $\tilde{y}_{it}$ is deflated revenue $\tilde{r}_{it} = q_{it} + p^Q_{it} - \bar{p}^Q_{it}$ minus deflated costs of intermediate inputs $\tilde{m}_{it} = m_{it} + p^M_{it} - \bar{p}^M_{it}$, $\tilde{k}_{it}$ is the deflated fix assets, l is labour.[30] $\tilde{\omega}_{it}$ denotes the sum of heterogeneous productivity and the deviations of firm-specific output/input prices from the industry-level price deflators: $\tilde{\omega}_{it} = \omega_{it} + (p^Q_{it} - \bar{p}^Q_{it}) - p^M_{it} - \bar{p}^M_{it}$, where $\bar{p}^Q_{it}$ and $\bar{p}^M_{it}$ index industry-level price deflators. Under the assumption that firm-specific deviation in prices is a monotonic function of productivity ω_{it}, we also have $\tilde{\omega}_{it} = g(\omega_{it})$ as a monotonic function. Unlike Brandt et al. (2017), here I exclude intermediate input m_{it} from the production function due to the unidentification problem discussed in Bond and Söderbom (2005) and Gandhi et al. (2020). One possible interpretation is that the gross production function is Leontief: $Y_{it} = \min\{\beta_0' K_{it}^{\beta_k} L_{it}^{\beta_l} e^{\omega_{it}}, \beta_m M_{it}\} e^{\varepsilon_{it}}$, where the intermediate input M_{it} is proportional to the output.

Then following the assumptions:

i) The information set I_{it} includes ω_{is} for $s \leq t$, and $\mathbb{E}[\varepsilon_{it}|I_{it}] = 0$

ii) ω_{it} evolves according to a first order Markov process $p(\omega_{it+1}|I_{it}) = p(\omega_{it+1}|\omega_{it})$ that is known to the firms and increasing in ω_{it}

iii) The state variable k_{it} is decided at time $t-b$. The labour input l_{it} is possibly dynamic chosen at $t-\zeta$, where $0 \leq \zeta \leq 1$. Free variables are chosen at t when ω_{it} is realized

iv) $m_{it} = f_t(\tilde{k}_{it}, l_{it}, \omega_{it}, d_{ict})$ is the proxy variable policy function and is invertible in ω_{it}, where d_{ict} denotes the logarithm of city population.

v) m_{it} is monotonically increasing in ω_{it}

I estimate the production function with the procedure below.

Step 1: Invert the control function $m_{it} = f_t(\tilde{k}_{it}, l_{it}, \omega_{it}, d_{ict})$ to attain $\omega_{it} = f_t^{-1}(\tilde{k}_{it}, l_{it}, m_{it}, d_{ict})$. By defining $\Phi_t(\tilde{k}_{it}, l_{it}, m_{it}, d_{ict}) = \beta_0 + \beta_k \tilde{k}_{it} + \beta_l l_{it} + f_t^{-1}(\tilde{k}_{it}, l_{it}, m_{it}, d_{ict})$, I obtain the first equation to be estimated:

$$\tilde{y}_{it} = \Phi_t(\tilde{k}_{it}, l_{it}, m_{it}, d_{ict}) + \varepsilon_{it},$$

where $\Phi_t(\cdot)$ is approximated with 5th order polynomial. This gives predicted values $\hat{\Phi}_t$ and $\hat{\varepsilon}_{it}$.

[30] The baseline year used for deflating is 1998.

Step 2: Assume that productivity ω_{it} is a first-order Markov process simplified into a linear function of lagged value: $\omega_{it} = \mathbb{E}[\omega_{it}|\omega_{it-1}] + \xi_{it} = g(\omega_{it-1}) + \xi_{it} = \rho\omega_{it-1} + \xi_{it}$, and for simplicity assume that d_{it-1} only impacts on $\mathbb{E}[\omega_{it}|\omega_{it-1}]$ via its impact on ω_{it-1}. It follows that: $\tilde{y}_{it} = \beta_0 + \beta_k\tilde{k}_{it} + \beta_l l_{it} + g(\omega_{it-1}) + \xi_{it} + \varepsilon_{it}$. Then for each coefficient vector $(\beta_0, \beta_k, \beta_l)$ I can plug in for productivity ω_{it-1} with $\Phi_{t-1}(\tilde{k}_{it-1}, l_{it-1}, m_{it-1}, d_{ict-1}) = \beta_0 + \beta_k\tilde{k}_{it-1} + \beta_l l_{it-1} + \omega_{it-1}$ and use the estimated $\hat{\Phi}_{it-1}$ to attain the equation of the second step of estimation:

$$\tilde{y}_{it} = \beta_0 + \beta_k\tilde{k}_{it} + \beta_l l_{it} + g(\hat{\Phi}_{it-1} - \beta_0 - \beta_k\tilde{k}_{it-1} - \beta_l l_{it-1}) + \xi_{it} + \varepsilon_{it}.$$

Then apply GMM method to the observed variables with the following moment conditions to obtain the estimates for 4 parameters $\hat{\beta}_0, \hat{\beta}_k, \hat{\beta}_l, \hat{\rho}$:

$$\mathbb{E}\left[(\tilde{y}_{it} - \beta_0 - \beta_k\tilde{k}_{it} - \beta_l l_{it} - \rho(\hat{\Phi}_{t-1} - \beta_0 - \beta_k\tilde{k}_{it-1} - \beta_l l_{it-1})) \otimes \begin{pmatrix} 1 \\ \tilde{k}_{it} \\ l_{it-1} \\ \hat{\Phi}_{t-1} \end{pmatrix}\right] = 0.$$

Finally I can compute TFP estimates using the residuals $\hat{\varepsilon}_{it}$ from the first stage and the estimated coefficients from the second stage as

$$TFP_{it} = \tilde{y}_{it} - \hat{\beta}_0 - \hat{\beta}_k\tilde{k}_{it} - \hat{\beta}_l l_{it} - \hat{\varepsilon}_{it}.$$

Table A.1: Sector-specific production function parameters

	Labour	Capital
13. Agriculture Food Processing	.4681	.2749
14. Other Food Production	.6034	.4091
15. Beverages	.4940	.3949
16. Tobacco Products	.5957	.8162
17. Textiles	.4488	.2684
18. Textile Wearing Apparel, Footwear and Caps	.5496	.2369
19. Leather, Fur, Feather and Related Products	.5341	.2587
20. Processing of Timber, Articles of Wood, Bamboo, Rattan, Palm and Straw	.4624	.2842
21. Furniture	.6656	.2365
22. Paper and Paper Products	.4114	.3223
23. Printing and Reproduction of Recording Media	.4077	.4851
24. Cultural, Educational, Arts and Crafts, Sports and Entertainment Products	.5522	.2621
25. Processing of Petroleum, Coking and Nuclear Fuel	.3145	.5459
26. Chemicals and Chemical Products	.3613	.3777
27. Pharmaceutical Products	.4568	.4796
28. Man-made Fibres	.4185	.3864
29. Rubber Products	.4768	.4151
30. Plactics Products	.4259	.3044
31. Non-metallic Mineral Products	.3332	.3512
32. Smelting and Processing of Ferrous Metals	.4941	.4372
33. Smelting and Processing of Non-ferrous Metals	.4782	.3204
34. Metal Products	.4153	.3010
35. General-purpose Machinery	.4200	.3865
36. Special-purpose Machinery	.3984	.2806
37. Transport Equipment	.5469	.4042
39. Electrical Machinery and Equipment	.4948	.3716
40. Communication Equipment, Computer and Other Electronic Equipment	.6368	.3478
41. Measuring Instruments and Machinery for Cultural Activity and Office Work	.4533	.2777
42. Artwork and Other Manufacturing	.4778	.2352

Notes: Labour and Capital report the estimated output elasticities from ACF production function estimations run separately for each two-digit CIC manufacturing sector using ASIP firm-level data. The proxy variable is log real intermediate inputs, and log urban population is incorporated as a state variable in the final sector-specific production-function estimations.

Appendix B: Hukou reform

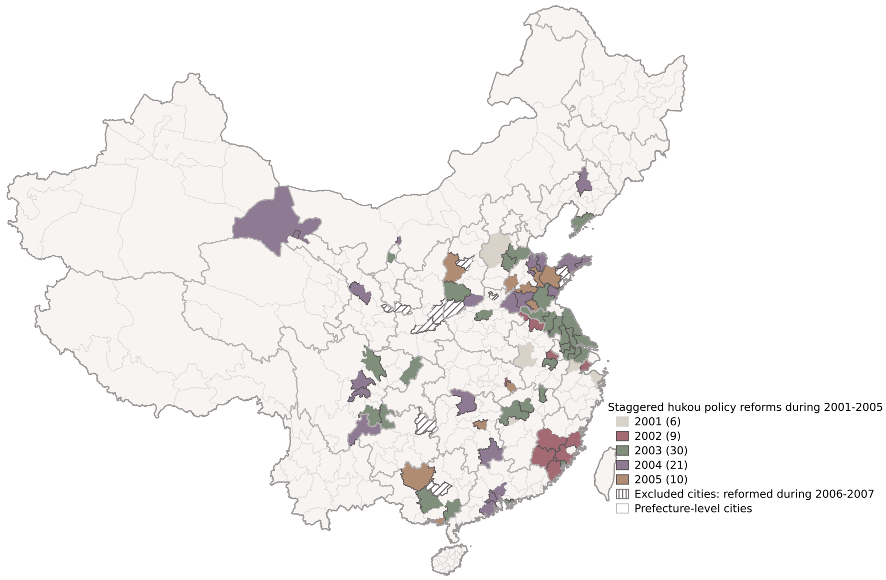


Figure B.1: Geographical distribution of 76 prefecture-level cities that adopted *hukou* reforms during 2001–2005. Numbers of reformed cities in each year are in parenthesis. Cities that reformed after between 2006 and 2007 are excluded from the sample as the firm-level data only covers 1998-2007 and we observe at least 2 post-treatment periods.

Below is a typical example of *hukou* reform notice from *Zhengzhou* municipal government.

Document number: Zheng Political [2009] No.19
Announcement Date: 22/08/2003
Issuing Department: Zhengzhou Municipal People's Government

The people's governments of all counties(cities), districts, all departments of the municipal government:
In accordance with the provincial government's spirits of further deepening the reform of the urban household registration management system, drawing on the successful experiences of other provinces and municipalities, and in light of the situation of our city, we hereby notify you of the issues related to the reform of our household registration management system as follows:

1. Cancel the current classifications of "agricultural hukou", "temporary hukou", "small town hukou", "non-agricultural hukou", change dual household registration system into unified registration system, and collectively referred to "*zhengzhou jumin hukou*". The household registrations of counties and districts can be transferred to each other at the local police station anytime.
2. Citizens from other provinces or cities who purchase housing within the jurisdiction of our city can apply for the residence registration of themselves and their immediate family members with the certificate of housing property rights.
3. Graduates with a diploma above a secondary professional technical school (including a technical school) can apply for a residence registration in *Zhengzhou* with their graduation certificate after filing in the Zhengzhou Talent Center.
4. Anyone who signs a labour contract with an enterprise or public institution in our city and pays a social pooling fund can apply for a registered permanent residence in Zhengzhou...

Source: https://gfxwj.zhengzhou.gov.cn/u/cms/gfxwj/201702/10091138mukr.pdf

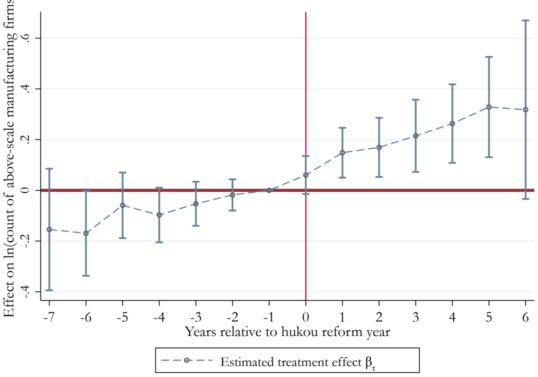
Appendix C: Other Tables

Table C.1: Hukou reform effects on logarithm of other dependent variables

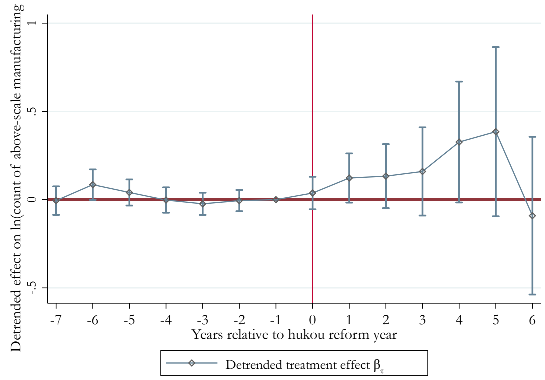
	gov expenditure	#primary	#middle	#hospital
	(1)	(2)	(3)	(4)
7 years pre	-0.0766	-0.0160	-0.1552	-0.0269
	(0.0728)	(0.3636)	(0.2592)	(0.0635)
6 years pre	0.0056	0.1437	-0.1269	-0.0519
	(0.0485)	(0.2585)	(0.1631)	(0.0708)
5 years pre	-0.0311	0.0268	-0.0592	-0.0990
	(0.0407)	(0.1134)	(0.0723)	(0.0894)
4 years pre	-0.0272	0.0815	0.0230	-0.0691
	(0.0316)	(0.0769)	(0.0370)	(0.0649)
3 years pre	0.0062	0.0513	0.0253	-0.1010**
	(0.0317)	(0.0628)	(0.0290)	(0.0482)
2 years pre	0.0063	-0.0056	-0.0017	-0.0162
	(0.0163)	(0.0213)	(0.0092)	(0.0223)
Reform year	0.0170	-0.0058	0.0006	0.0403
	(0.0169)	(0.0254)	(0.0138)	(0.0406)
1 year post	0.0253	-0.0010	-0.0129	0.0164
	(0.0205)	(0.0389)	(0.0165)	(0.0592)
2 year post	0.0292	-0.0103	-0.0139	-0.0011
	(0.0250)	(0.0420)	(0.0173)	(0.0544)
3 year post	0.0041	-0.0101	0.0035	-0.0081
	(0.0313)	(0.0533)	(0.0237)	(0.0586)
4 year post	-0.0033	-0.0434	-0.0269	-0.0546
	(0.0352)	(0.0642)	(0.0299)	(0.0653)
5 year post	-0.0013	-0.0542	-0.0720*	-0.1209*
	(0.0579)	(0.0899)	(0.0409)	(0.0652)
6 year post	-0.0111	-0.1891	-0.1469	-0.0697
	(0.0773)	(0.1270)	(0.0925)	(0.0766)
Group-specific time effects	Yes	Yes	Yes	Yes
SE clustered by city	Yes	Yes	Yes	Yes
Obs	2344	2405	2404	2404

Notes: Prefecture-level regressions covering cities with available data over 1998–2007 (unbalanced panel) using dynamic specifications. The dependent variables in columns (1)–(4) are local government expenditure, the number of primary schools, the number of high schools and the number of hospitals, respectively. Baseline year is chosen as the 1st year before the introduction of hukou reform. All dependent variables are in logarithm form. Data are from Chinese Statistical Yearbooks (CSY).
Standard errors clustered at city in parentheses.
**** p<.01, ** p<.05, * p<.1*

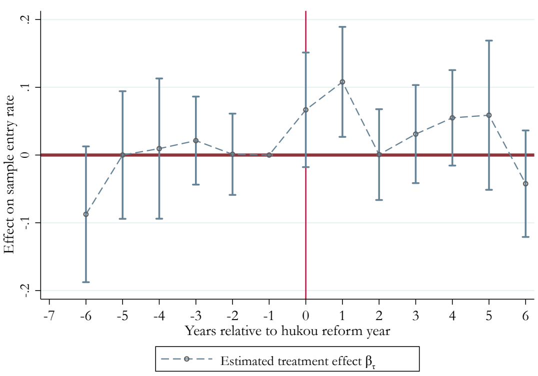
Figure C.1: Reduced-form effects of *hukou* reform on observed above-scale manufacturing firms



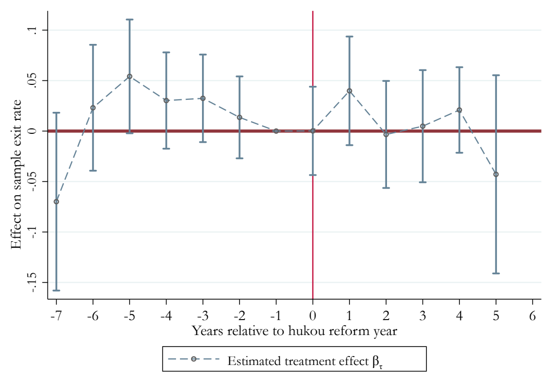
Panel A: Firm count



Panel B: Detrended firm count



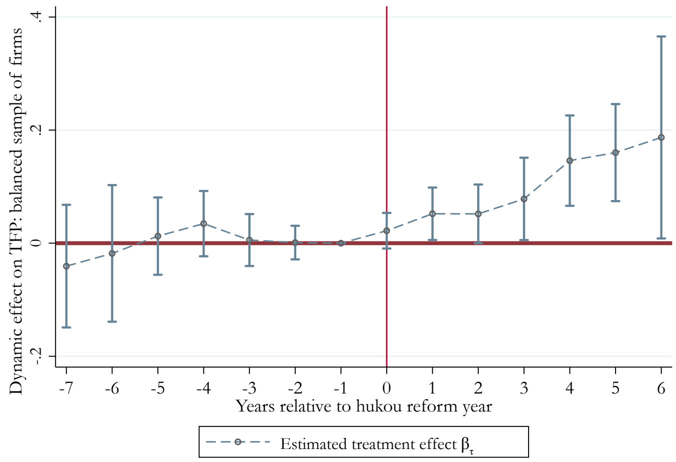
Panel C: Sample entry rate



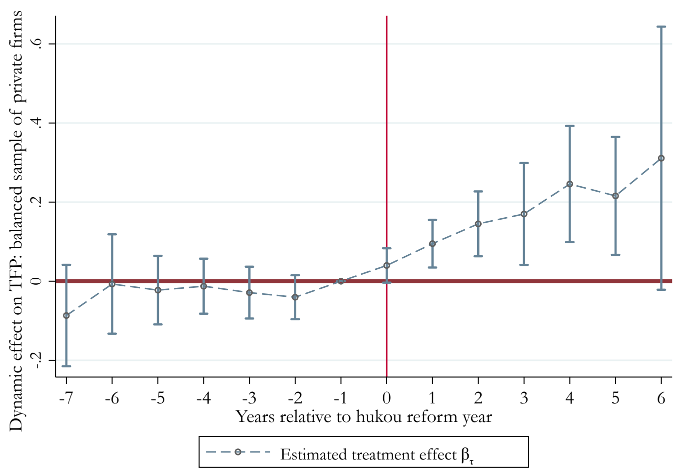
Panel D: Sample exit rate

Note: Panels A and B show dynamic effects on the log number of above-scale manufacturing firms observed in the urban center (*shixiaqu*) of each prefectural city. Panel B residualizes the outcome using city-specific linear trends estimated from prereform observations. Panels C and D define sample entry and exit as a firm's first and last observed year, respectively. The entry rate divides the number of firms first observed in year t by the number observed in $t-1$; the exit rate divides the number last observed in year t by the number observed in t. Entry in 1998 and exit in 2007 are omitted because of censoring. First and last appearance need not represent true firm birth and death, as firms may cross the ASIP size threshold or enter and leave survey coverage. All specifications include city and cohort-by-year fixed effects. Standard errors are clustered at the city level, and the vertical bars indicate 95% confidence intervals.

Figure C.2: Reduced-form effects of *hukou* reform on TFP in the balanced sample



Panel A: Firms surviving throughout 1998–2007



Panel B: Private firms surviving throughout 1998–2007

Note: Dynamic effects of *hukou* reform on firm-level TFP for a balanced sample of manufacturing firms observed annually from 1998 to 2007. Panel A includes firms of all ownership types, while Panel B restricts the sample to private firms classified by prereform ownership status. The omitted reference period is the year preceding the reform. All specifications include firm, cohort-by-year, and industry-by-year fixed effects. Standard errors are clustered at the city level, and the vertical bars indicate 95% confidence intervals.