Urban Density and Firms' Stock Returns

Mahsa Memarian
INCAE Business School. Campus Walter Kissling Gam La Gartita 1, 960-4050 Alajuela, Costa Rica, mahsa.memarian@incae.edu

Carles Vergara-Alert
IESE Business School. Av. Pearson 21, 08034 Barcelona, Spain, cvergara@iese.edu

Firms located in dense urban areas present higher productivity due to the flow of ideas and innovation in these areas. Through this productivity channel, the urban density characteristics of the areas where firms are located affect the stock returns. We use high-resolution satellite images from Google Earth to develop an exogenous measure of potential density increase (PDI) for the 95 most populated metropolitan statistical areas (MSAs) in the US. This measure represents the proportion of area in the total area within a 1 hour drive from the center of the MSA that could rapidly increase its density. We find that firms located in areas with a high potential density increase present lower stock returns: on average a 10% higher PDI of an MSA results in a 0.29% lower excess stock return of firms located in this MSA.

Key words: productivity; stock returns; urban density; agglomeration

1. Introduction

Agglomeration, that is, the productivity gains that arise from clustering production and workers, is one of the main reasons for the existence of cities. Agglomeration advantages result in firms and workers being more productive in dense urban areas than elsewhere (Marshall (1890); Sveikauskas (1975); Rauch (1993); Rosenthal and Strange (2004); Combes et al. (2012)). These productivity gains in cities are driven by knowledge spillovers that accelerate the adoption of new technologies, the increase in opportunities from specialization, and the existence of economies of scale and low transportation costs (Davis et al. (2014)). As a result, firms invest and grow more (Dougal et al. (2015)) and can generate more revenue (Glaeser et al. (2001)) by locating closer to the center of dense urban areas. Although there is evidence of this location-productivity relationship and the link between firm's productivity and stock returns, there is no study in the asset pricing literature that analyzes whether and how the urban density characteristics of the areas where firms are located affect the stock returns.

In order to address this gap in the literature, in this paper we explore the effect of density characteristics of the urban areas where firms are located on their expected excess stock return.

First, we use high-resolution satellite images from Google Earth to develop an exogenous measure of potential density increase (PDI) for the 95 largest metropolitan statistical areas (MSA) in the US. MSAs are geographic entities, defined by the Office of Management and Budget that contain a core urban area with a population of 500,000 or more. Each MSA consists of 1 or more counties that contain the core urban area as well as adjacent counties that have a high degree of social and economic integration (as measured by commuting to work) with the urban core. We estimate their PDI as the proportion of area in the total area within a 1 hour drive from the center of the MSA that could rapidly increase its density. Furthermore, we control for land availability by developing a measure of the non-potential density increase (NDI) area for each of the 95 MSAs. The NDI for a specific MSA is defined as the proportion of area in the total area within a 1 hour drive from the center of the MSA that cannot rapidly increase its density, either because it is already highly dense or because it is undevelopable due to natural conditions (i.e., highly sloped terrains, lakes, rivers, etc.). PDI as a measure for the potential to build new constructions and NDI as a measure to capture geographical constraints are two dimensions of the density which help us clarifying the effect of urban density as drivers of innovation and productivity. These two measures go beyond the urban density or characterization of geographic space which captured by the total population or the population density in the existing literature (Glaeser et al. (2005); Maantay et al. (2007); Mennis (2003); Sutton et al. (2003)). PDI and NDI distinguish between dense MSAs by pointing out the heterogeneity exists in dense urban areas. Using these two measures we show that the agglomeration impact on corporates' outcome differs in MSAs considered as the ones with high urban density. As an example, Chicago and Los Vegas both known as dense MSAs considering urban density itself. However, the influence of urban density on the outcome of the firms located in these MSAs is different. This difference is explained by the portion of the existed dense areas with potential for new construction in comparison to the fully packed dense areas or the areas with high geographical constraints which are impossible or very costly to construct more. Empirically we found no evidence about the effect of population as a proxy for urban density in the existing literature, on firms' stock returns. Using PDI and NDI we can clarify the effect of urban agglomeration on firm's outcome. Using firm-level data for 2,711 firms from 2010 to 2014, we investigate whether urban density characteristics cause an increase or a decrease in firms' expected stock return.

Our main findings can be described with two sets of results. First, we study the link between the firms' productivity and the cross-section of stock returns following İmrohoroğlu and Tüzel (2014). We provide further evidence relating to this link by showing that firms located in fast-growing MSAs with a high potential density increase present higher productivity. Second, we analyze the relationship between urban density characteristics and the cross-section of stock returns. We present new evidence regarding this relationship. We find that firms located in areas with high PDI present lower stock returns. This result shows the importance of being located in an MSA that can potentially grow quickly. We empirically document that a 10% higher PDI of an MSA results in a 0.29% lower excess stock return of firms located in this MSA. There might be existing alternative mechanisms and channels which this effect could happen through, but the mechanism we explain in this paper is by the excessive flow of ideas and innovation in dense and more vibrant urban areas, which cause higher RD expenditure for firms (Sun et al. (2017)). We show how firms benefit from the knowledge spillover existing in these areas and become more productive. Using our measures of urban density, this mechanism is also shown empirically by the positive and significant influence of firm's RD; which represents the level of innovation in the corresponding area, on their productivity. Through this mechanism, we argue that productivity is one of the channels which drives the causal effect between urban density characteristics and firms' stock returns. Specifically, to confirm our argument, we show that the main part of the effect of firms' productivity on their stock returns is caused by the PDI of their MSA among all other unobservable variables.

Our goal is to identify the causality between the urban characteristics of the areas in which firms are located and the firms' stock return, which is consistent with the idea of the location influence on firms' productivity and expected stock return. Therefore, we require an exogenous source to capture the characteristics of urban density in different areas. Our measure of PDI can be considered exogenous because the urban density characteristics of an MSA do not change as fast as stock prices and returns. However, we adapt the instrumental variable approach developed by Himmelberg et al. (2005) to address any potential concerns about the endogeneity of the PDI measure in our estimations. We instrument our measure of PDI using the interaction of local housing supply elasticity and long-term interest rates to identify changes in the housing demand. To achieve this, we use the local housing supply elasticities provided by Glaeser et al. (2008) and Saiz (2010), which are available for 95 MSAs. These elasticities capture the amount of developable land in each metro area and are estimated by processing satellite-generated data on elevation and the presence of water bodies. As a measure of long-term interest rates for the real estate market, we use the 30-year fixed conventional mortgage rate from the St. Louis Federal Reserve Bank (FRED) website between 2010 and 2014. Our results remain robust to this identification strategy.

Our paper contributes to two growing strands of literature that study the connection between firms' location, their economic activity, and their financial performance. First, we contribute to the agglomeration literature that suggests that the potential for growth in a location becomes more noticeable with geographic proximity and attendant externalities found in specialized workers, suppliers, and infrastructure (Krugman (1991)). In a later work, Burchfield et al. (2006) argue that areas in which about one-half of the land in the immediate vicinity is already built up seem to be the most attractive for new development. Building on that, crowding takes two forms (Duranton and Puga (2004)): the capacity constraint when many people try to use the facility simultaneously and the crowding that occurs as the facility needs to be located somewhere. By expanding the size of the community of users, some of those users will be located further from the facility. Considering the traffic of the users between their residence and the facility, a city is defined as the equilibrium outcome of such a trade-off between the gains from sharing the fixed cost of the facility among a larger number of consumers, which increases the returns through agglomeration economies or localized aggregation, and the costs of urban congestion. Consequently, a larger workforce leads to a more than proportionately higher level of output because of the constant elasticity of substitution aggregation by final producers. This happens in dense areas through more efficient sharing of indivisible facilities (like the local infrastructure), the risks, and the gains from variety and specialization, better matching between employers and employees, buyers and suppliers, partners in joint projects, or entrepreneurs and financiers, and facilitated learning about new technologies, market evolutions, or new forms of organization. Carlino et al. (2007) report that patent intensitythe per capita invention rateis positively related to the density of employment in the highly urbanized portion of metropolitan areas, which suggests that density is a key component of the knowledge spillovers and innovation that power economic development and growth. On the other hand, Davis et al. (2014) study the effect of local agglomeration on aggregate growth by modeling agglomeration as an externality in which the total factor productivity (TFP) at a location increases with the location's output densitythe total output per acre of finished land in production. Denser, higher-productivity acres of land have greater variety, because more intermediate service producers can break even. This connection between density and variety in turn yields an expression for the production of composite services in which labor productivity increases with the variety. With density leading to variety, and variety leading to productivity, the model yields a reduced-form relationship between density and productivity.

Second, our paper is related to the asset pricing literature that investigates the factors that affect firms expected stock return.İmrohoroğlu and Tüzel (2014) is the closest paper to our study. They provide evidence of a negative link between the firm-level TFP and the cross-section of expected stock returns. Building on this body of literature, in this paper, we show the main part of the TFP effect on firms' stock returns is caused by the potential density increase among all the other unobservable variables. In summary, urban density characteristics of the area where the firm is located has influence on firm's innovation and eventually its productivity in this area. Subsequently, through this productivity channel, as one of the drivers, the characteristics of urban density affect the firm's stock returns.

2. Hypotheses Development

The focus of this paper is primarily empirical. To motivate our empirical analysis, we set out a stylized conceptual framework to illustrate the key effects of urban density in firms' stock returns.

The first part of the empirical study builds upon the literature which analyzes the effects of density on innovation and shows that the returns to creative capital on innovation (i.e., creative spillovers) increase with the density of the metropolitan area. For example, Carlino et al. (2007) and Knudsen, Florida, Stolarick, and Gates (Knudsen et al.) document a positive relationship between the density of creative workers and the metropolitan patenting activity. Their results show that density is a key component of the knowledge spillovers and innovation that power economic development and growth. Moreover, recent complex dynamic stochastic general equilibrium models of cities show a positive effect of local agglomeration on aggregate growth. For instance, Davis et al. (2014) model agglomeration as an externality in which the TFP at any location increases with the location's output density.[1]

Building on this literature, we expect that firms located in dense areas with high potential for a density increase have higher TFP and are more productive. Hypothesis 1 outlines this prediction.

Hypothesis 1. *Firms located in fast-growing areas with high potential for a density increase are more productive.*

Sun et al. (2017) study the importance of the role of human capital. They show that general human capital affects a firm's success probability directly, and it can affect the firm's success probability indirectly through its R&D-level choice. They also argue that R&D human capital determines innovation directly through the positive significant effect of general human capital and the managerial personnels education on innovation. In this paper, we explain the mechanism by the excessive flow of ideas and innovation in dense and more vibrant urban areas, which cause higher R&D expenditure for firms (Sun et al. (2017)). Therefore, through 2, we show how firms benefit from the knowledge spillover existing in these areas and become more productive. More specifically 2 shows the effect of firm's location characteristics on its productivity through firm-level R&D, which represents the level of innovation in the corresponding area. Consistent with the previous literature, we use TFP as a measure of their productivity and we test whether the addition to a firm's R&D caused by the potential density increase of the area in which the firm is located results in a higher level of productivity for that firm.

[1] Davis et al. (2014) show that denser, higher-productivity acres of land have greater variety, because more intermediate service producers can break even. This connection between density and variety in turn yields an expression for the production of composite services in which labor productivity increases with the variety. With density leading to variety and variety leading to productivity, the model provides a reduced-form relationship between density and productivity.

Moreover, İmrohoroğlu and Tüzel (2014) document a negative relationship between firm productivity and expected returns, indicating that the average implied cost of capital for the low TFP portfolio exceeds that of the high TFP portfolio, which is statistically significant. In the present paper, we argue that firms located in MSAs with a high potential density increase present lower stock returns. The productivity channel is the driver of the causal effect between urban characteristics and stock returns. Hypothesis 2 summarizes this prediction.

HYPOTHESIS 2. *Firms located in fast-growing areas with high potential for a density increase present lower expected stock returns.*

To test this hypothesis, we compute the part of the TFP that comes from urban characteristics of each MSA by creating TFP fitted values on our PDI measure, potential density increase, and we study the effect of the FamaFrench factors on the portfolios of a cross-section of expected excess stock returns, sorted by this part of the TFP. Specifically, we find that the main part of firms' TFP effect on stock returns is caused by the potential density increase among all other unobservable variables.

3. Data

Our variables belong to two different sets of variables as firm level and MSA level. We use firm-level value added, employment, and capital. We use firm-level data on company names and zip codes from Compustat and we link it to the MSA in which each company is located. Using the mapping table between zip codes and MSA codes developed by the U.S. Department of Labor's Office of Workers' Compensation Programs (OWCP), and we match the zip codes from the two files to obtain the company's location. We keep the MSAs that are belongs to our list of 95 most Populated Metropolitan Statistical areas and we end up with the sample of firms located in 82 US MSAs.

To estimate firms' productivity, we use some of the key variables, such as firm-level value added, employment, and capital. We compute the firm-level value added using Compustat data on sales, operating income, and employees, deflated using the output deflator. The stock labor is given by the number of employees (EMP), and firm-level capital stock is given by the gross plant, property, and equipment (PPEGT), both from Compustat. Corporate real estate holding is considered as the sum of building plus capitalized leases, all divided by net property, plant, and equipment (PPE), in accordance with Tuzel (2010). Leverage is computed as long-term debt (DLTT) divided by the sum of DLTT and the market value of equity. Asset growth is considered as the percentage change in the total assets. The hiring rate at time t is the change in the stock of labor (EMP) from t-1 to t. The research and development expenditure is calculated as the research and development expense (XRD) divided by the gross PPE. ROA is computed as net income IB minus dividend on preferred DVP plus income statement deferred taxes (TXDI), all divided by total assets (TA), ROE is calculated as income before extraordinary items over total stockholders' equity, and market to book ratio as firm's market value over its book value, all from Compustat. Finally, we consider firm's age as the number of years since the firm's first year of observation in Compustat.

The monthly stock returns are from the Center for Research in Security Prices (CRSP). In calculating future returns, we match the CRSP stock return data from July of year t to June of year t+1 with accounting information for the fiscal year ending in t-1, as in Fama and French (1992) and Fama and French (1993). We do so in order to ensure that the accounting information is already impounded into the stock prices. Then, we compute the excess expected stock returns, considering four FamaFrench factors: the excess market returns (MKT); the return of the portfolio that is long in small firms and short in big firms (SMB); the return of the portfolio that is long in high-B/M firms and short in low-B/M firms (HML); and the momentum factor (MOM), that is, the return of the portfolio that is long in short-term winners and short in short-term losers. The firm size is computed as the market capitalization by multiplying the number of shares outstanding by the share price.

At MSA level, we use the data on population density computed as the number of MSAs' inhabitants divided by the total MSA square kilometer area from US census Bureau. We obtain the residential home price index (HPI) from the Federal Housing Finance Association (FHFA). Moreover, we use the 30-year conventional mortgage rate from the St. Louis Federal Reserve Bank (FRED) as a measure of the long-term interest rate for the real estate market. And eventually, as urban density and location characteristics data we use our two measures of PDI and NDI. The methodology of developing these two measures is explained in detail in the following section.

Our sample is comprised of all the remaining firms in Compustat that have positive data on sales, total assets, number of employees, gross property, plant, and equipment, depreciation, accumulated depreciation, and capital expenditures. As is standard in the literature (see Chaney et al. (2012); Cvijanović (2014)), we also omit the firms that belong to the finance, insurance, real estate, non-profit, government, construction, or mining industries. This leaves us with an unbalanced panel containing 2,711 distinct firms spanning the years between 2010 and 2014.

4. Measures of Urban Density Characteristics

In this section we develop a computer vision method to measure characteristics of urban density in the US metropolitan areas taking into consideration geographical constraints. We segment images into four geometric classes: non-developable areas, low developed areas, developed areas, and highly developed areas.

First, we define the whole polygon in each MSA, estimated with one hour's driving distance by car, measured from the MSA's center along the main existing roads, allowing for a maximum time variation based on the current traffic information available from Google Maps. This captures the willingness to spend about one hour commuting between work and residential placements, which leads to a distinct polygon for each city based on the traffic congestion or geographic disturbances. Afterwards, we name water bodies, natural reserves and steep-sloped terrain as fully restricted parts for any further construction. The highly dense areas account for the fully packed areas, usually around the central business district of the MSA. Developed areas represent the parts that have already been constructed but still offer more opportunities for further growth in density. Finally, low developed areas are defined within each polygon of the MSAs considered as plain and empty lands.

Our measures build upon previous measures of land availability that have been used in the urban economics literature. Saiz (2010) uses satellite-generated data on terrain elevation to develop a measure of the amount of developable land based on the presence of water bodies and steep-sloped terrain in US MSAs. He demonstrates that topographical constraints correlate positively and strongly with regulatory barriers to development and that both types of constraints negatively affect the supply elasticity. The measure in Saiz (2010) is static. Recent papers propose dynamic measures. Naika, Kominersb, Raskara, Glaeserc, and Hidalgoa (Naika et al.) study the changes in the physical appearance of neighborhoods from street-level imaginary. Their results show how computer vision techniques in combination with traditional methods can be used to explore the dynamics of urban change. Furthermore, Henderson et al. (2016) argue that durable formal-sector buildings can be built high, unlike informal ones, which are malleable. They study this idea through the average height of buildings by grid square in the formal and slum sectors.

The development of our measures of urban density characteristics for 95 US MSAs consists of two main steps: (1) the definition of the areas within the MSA; and (2) the calculation of measures urban density characteristics. We describe these steps below.

4.1. Definition of the areas within the MSA

We first define the area of study for each MSA. We define a polygon surrounding each city by estimating the one-hour driving distance by car, measured from the metropolitan center along the main existing roads, allowing for a maximum time variation based on the current traffic information available from Google Maps. We use high-resolution satellite images from Google Earth in order to define four types of subareas within each MSA as follows.

- Highly developed (HD) areas. These are fully packed urban areas that are characterized by the substantial existence of either tall buildings or residential or even commercial areas, where the observable available space for new developments is negligible.

• Developed (D) areas. These areas correspond to semi-urban areas that are characterized by the existence of residential or commercial areas surrounded by some observable available space for new constructions.

• Low developed (LD) areas. These correspond to zones of empty land that are available for construction, accounting for a large amount of the total space analyzed. This area is characterized mainly by plains.

• Undevelopable (U) areas. Non-developable areas and natural reserves. Protected areas include lakes, rivers, and in many cases mountains as well as national parks.

Figure 1 shows some examples of aerial views of these types of areas.

[Insert Figure 1 around here]

4.2. Calculation of measures urban density characteristics

First, we exclude undevelopable areas from the area of study to obtain an input image for each MSA. We use a computer vision algorithm in Matlab to determine the exact number of square kilometers for each type of area (i.e., HD, D, and LD areas). This algorithm classifies the different types of areas using a segmentation process based on different colors that are recognized in an input image taken from Google Earth. Input images for each MSA are prepared manually by determining the whole polygon within a one-hour drive from the center of the MSAs in Google Earth, excluding undevelopable areas. Figure 2 provides an example of input and output images for the New York.

[Insert Figure 2 around here]

In the next step, we define our main measure of potential density increase (PDI) and a measure of non-potential density increase (NDI) for each MSA. In comparison with the low developed (LD) parts of the MSA, developed areas (D) have more opportunities for fast further construction growth due to the existing infrastructure and services, which make the progress of the land construction process easier and faster. Accordingly, building the development based on the expectation about the further growth of the metropolitan area, we define a measure of PDI. This measure is indicated by dividing the area square meters of the developable land that have already been constructed but is not fully packed and could rapidly increase its density as developed (D) areas, by the sum of developed (D) plus low developed (LD) areas as developable part of the MSA:

$$PDI = \frac{area_D}{area_D + area_{LD}} \tag{1}$$

The measure of PDI quantifies the amount of land with a relatively high urban density that has considerable opportunities for further construction growth. Cities with high PDI, such as Chicago and Los Angeles, present many developable parts that could progress quickly and easily. On the other hand, a low PDI represents cities with a large amount of empty land in their developable parts, like Charlotte and Louisville.

Moreover, to control for land availability, we create another measure to account for the amount of land where it is very costly or impossible to increase its density. This could happen because the area is either highly dense or undevelopable. We name this variable non-potential density increase (NDI) and we will use it as control variable in our specification. NDI refers to the sum of highly developed areas plus undevelopable areas as a proportion of the total area of the polygon within a one-hour drive from the city center:

$$NDI = \frac{area_{HD} + area_U}{area_{HD} + area_D + area_{LD} + area_U} \tag{2}$$

This measure refers to the proportion of area in the total area within a one-hour drive from the center of the MSA that could not rapidly increase its density, either because it is already highly dense and packed or because it is undevelopable due to natural conditions (i.e., highly sloped terrains, lakes, rivers, etc.). Therefore, cities with lots of highly developed or/and undevelopable areas present high values of NDI as dense cities with substantial geographical constraints to further construction, such as San Francisco. Table 1 reports the values of the two measures of urban density characteristics, PDI and NDI, for all 95 MSAs.

[Insert Table 1 around here]

Table 2 provides the summary statistics of PDI, NDI, various firm characteristics, and the rest of our controls that we described in section 3.

[Insert Table 2 around here]

4.3. Identification of the firm's location

For our firm-level analysis, we identify a firm's location using its headquarters location from Compustat due to the fact that relevant decisions at the firm level are made in the headquarters. Furthermore, this is consistent with the broad body of the literature, as there is no accessible systematic source of information on corporations' true location(s). Overall, defining a firm's location as the location of its headquarters, even if firms' headquarters are often separated from their operations by hundreds or even thousands of miles, may help rather than hinder our ability to identify the types of spillovers that are the focus of this study.

To validate this identification further, we refer to Tuzel and Zhang (2017). They link their CompustatCRSP sample to the ReferenceUSA U.S. Businesses Database and collect employment data for all headquarters, branch, and subsidiary locations of the firms in the sample. This allows them to create an employment map for each of roughly 2000 firms in her linked sample. They find that 63% of the firms in their sample have at least 50% of their employment in the MSA of their headquarters, and, for the median firm in her sample, the headquarter location accounts for 72% of the total employment. Consequently, we use the headquarter location as a proxy for the location of real estate and identify a firm's location with its headquarters location from Compustat. This assumption could be a problem in the case that the majority of a firm's real estate holdings are actually located elsewhere and thus lead to some measuring error(s) for our results. For the firm-level analysis, Chaney et al. (2012) argue that headquarters and production facilities tend to be clustered in the same state and MSA and that headquarters represent an important fraction of corporate real estate assets. They provide hand-collected information on firms' headquarters ownership using their 10K files as evidence supporting this assumption. Therefore, they conclude that headquarters location is a reasonable proxy for firm location. Furthermore, this assumption is validated by Cvijanović (2014) using state-level data on firms' operations from Garcia and Norli (2012), which is obtained by measuring the degree of firm geographic concentration by extracting state name counts from annual reports filed with the SEC on Form 10-K.

A potential concern is about the headquarter moves and the idea that firms with better growth and acquisition opportunities as well as more financial slack, choose to locate in clusters. Pirinsky and Wang (2006) provide evidences document that from 1992 to 1997 less than 2.4% of firms in Compustat moved its headquarters from one MSA to another (i.e., 118 out of 5,000 firms did so). One problem with the COMPUSTAT location data is that COMPUSTAT only reports the current state and county of firms' headquarters. To correct for this deficiency Pirinsky and Wang (2006) cross-check the historic record of firms' headquarters information from Compact Disclosure. Unlike COMPUSTAT, Compact Disclosure provides information on the city and state of a firm's headquarters location on an annual basis over the period from 1988 to 2002. Using the Compact Disclosure data, they identify all firms whose corporate headquarters have moved from one location to another over the period and delete all firm-year observations prior to the relocation from the main sample. However, they specifically examine the effect of headquarters relocation on a firm's co movement for a subsample of relocating firms. They construct their sample of relocating firms as follows. First, they identify firms that report different headquarters locations in two consecutive years in the Compact Disclosure database. Then, they manually verify each move using newspaper reports and wire reports from Factiva (a news report data service provided by Dow Jones News Service and Reuters). The majority of headquarters relocations are a result of corporate mergers and acquisitions or some other forms of major corporate restructuring. They exclude such firms from the relocation sample. They further eliminate firms that moved locally, that is, firms that moved their headquarters to a different city but still remained within the same MSA. In order to allow for 5-year estimation periods before and after relocation, they restrict the sample to corporate relocations occurring during the 1992 to 1997 period. After matching the relocation sample with the CRSP and MSA data, they final sample of relocating firms consists of 118 firms. Most of the relocating firms in the sample are relatively small, although the sample does include some well-publicized moves, such as SBC's relocation from St. Louis to San Antonio and General Dynamics' relocation from St. Louis to Washington DC. The most commonly cited reasons for headquarters relocations by these firms are: to be close to customers; to reduce costs; to move to a more important production base area; and, to capture synergies with other local firms. For example, Southwestern Bell's chairman Edward E. Whitacre Jr. comments on the headquarters relocation of his company: It will put us closer to more of our major growth markets and customers. Similarly, General Dynamics' chairman William Anders argues The company can operate more effectively, more efficiently and be more responsive by having our headquarters and our leadership closer to our principal customers. Unfortunately the historic record of firms' headquarters information from Compact Disclosure has not updated after 2002 and due to the fact that COMPUSTAT only reports the current state and county of firms' headquarters, we have this data limitation on headquarter location move.

5. Empirical Strategy and Empirical Results

5.1. Estimating Total Factor Productivity (TFP)

TFP is a measure of the overall effectiveness with which capital and labor are used in the production process. It provides a broader gauge of firm-level performance than some of the more conventional measures, such as labor productivity or firm profitability.[2] We estimate the production function given in:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \nu_{it} \tag{3}$$

where y_{it} is the log of value added for firm i in period t. Let l_{it} and k_{it} be the log values of labor and capital of the firm, respectively; ω_{it} is the productivity; and ν_{it} is the error term, which is not known by the firm or the econometrician. We employ the semi-parametric procedure suggested by Pakes and Olley (1995) to estimate the parameters of this production function. The major advantage of this approach over more traditional estimation techniques, such as ordinary least squares (OLS), is its ability to control for selection and simultaneity biases and deal with the within-firm serial correlation in productivity that plagues many production function estimates.

Once we have estimated the production function parameters ($\widehat{\beta_0}$,$\widehat{\beta_l}$ and $\widehat{\beta_k}$), we obtain the firm-level (log) TFP from:

[2] Profitability captures only the part of the value added that is distributed to shareholders, and labor productivity can be an adequate measure of overall efficiency, especially in capital-intensive industries. See Lieberman and Kang (2008) for a case study of a Korean steelmaker showing the differences between TFP and profitability in measuring firm performance.

$$w_{it} = y_{it} - \widehat{\beta}_0 - \widehat{\beta}_l l_{it} - \widehat{\beta_k} k_{it} \tag{4}$$

in which firm-level data are supplemented with the price index for the gross domestic product as a deflator for the value added and the price index for private fixed investment as a deflator for investment and capital, both from the Bureau of Economic Analysis and the national average wage index from the Social Security Administration. Value added is computed as (sales materials), deflated by the GDP price deflator. Materials are considered as (total expenses labor expenses), which is equal to (sales operating income before depreciation total stuff expense). Therefore, in the end we compute value added as (operating income before depreciation OIBDP + total stuff expense XLR), all gathered from Compustat. Stock labor, considered as the number of employees (EMP), is available from Compustat. Capital stock is computed as property, plant, and equipment total (gross) PPEGT, deflated by the price deflator for investment, following Hall (1993). In the estimation, we use industry-specific time dummies. Hence, our firms' TFP is free from the effect of industry or aggregate TFP in any given year.

5.2. Urban density characteristics and firms' productivity

We argue that productivity is one of the channels which drives the causal effect between urban density characteristics of the area where the firm is located and its stock returns. There might be existing alternative mechanisms which this effect could happen through, but the mechanism we explain in this paper is by the excessive flow of ideas and innovation in dense and more vibrant urban areas, which cause higher R&D expenditure for firms (Sun et al. (2017)). This mechanism is explained through the large flow of ideas and innovation in dense urban areas and, therefore, it is linked to higher R&D expenditures for firms located in these areas. We confirm that there is a positive and significant influence of firms' R&D expenditure on their productivity. Consequently, we estimate TFP, that is $Productivity^l_{it}$ of firm i with headquarters located in area l at time t as:

$$Productivity^l_{it} = \alpha_i + \beta.PDI/R\&D(PDI) + Controls^l_{it} + i.year + i.industry + i.state + \epsilon_{it} \tag{5}$$

Table 3 provides the summary results for the effects of our urban density measures and RD on firms' productivity. The dependent variable, productivity, is measured as a firm's TFP. $Controls_{it}$ denotes two sets of firm-level controls and MSA level controls. Following the existing literature on productivity, at firm level, we control for (1) leverage; (2) firm size; (3) asset growth; (4) hiring rate; (5) return on assets (ROA); (6) return on equity (ROE); (7) market-to-book ratio; (8) corporate real estate holding; and (9) company age. At MSA level we control for (1) NDI measure; (2) residual housing price index; and (3) population density. We also control for the year fixed effect as well as the industry fixed effect and state fixed effect. Errors are clustered at the MSA level.

In Column [1] of Table 3, we find that on average a 10% higher PDI of an MSA results in 2.08% higher productivity of firms located in this MSA. Column [2] shows the result for the specification in column [1], while controlling for land availability using NDI measure. Further analysis explaining the mechanism is reported in column [3]. We find that firm's R&D expenditure has a positive and significant effect on its productivity through our measure of PDI. To show this effect, we calculate the part of a firm's R&D that comes from urban density by estimating the fitted values of the firm's RD on PDI: R&D (PDI). We also control for the fitted values of the firm's R&D on NDI measure: R&D (NDI). Accordingly, we obtain a positive and significant effect of R&D (PDI) on firm's productivity. We find on average that a 10% increase in R&D (PDI) results in a 4.39% increase in the firm's productivity (see column [3]).[3]

[Insert Table 3 around here]

To address a potential concern about the measure of PDI being endogenous to real estate prices, we adapt the instrumental variable approach developed by Himmelberg et al. (2005). We instrument our measure of potential density increase, PDI, using the interaction between the elasticity of supply of the local housing market as in Himmelberg et al. (2005) and Mian and Sufi (2011) and long-term interest rate to pick up changes in the housing demand.[4] We use the local housing supply elasticities provided by Glaeser et al. (2008) and Saiz (2010), which are available for 95 MSAs. These elasticities capture the amount of developable land in each metro area and are estimated by processing satellite-generated data on elevation and the presence of water bodies.[5] As a measure of long-term interest rates for the real estate market, we use the 30-year fixed conventional mortgage rate from the St. Louis Federal Reserve Bank (FRED) website between 2010 and 2014.

$$Productivity^l_{it} = \alpha_i + \beta.PDI/R\&D(PDI) + .Elasticity^l.IR + Controls^l_{it} + i.year + i.industry + i.state + \epsilon_{it} \tag{6}$$

[3] Note that urban density characteristics, geographical constraints, and the existing limitations of an MSA do change rapidly over time. Therefore, we use our two measures, PDI and NPI, as approximately constant proxies for urban density characteristics of each MSA in a reasonable period of five years from 2010 to 2014.

[4] Additionally, as interest rates have a time trend that explains most of the correlation between home prices, we eventually control for the interaction between housing supply elasticity and year in our instrumental regression (Davidoff et al. (2016)).

[5] Davidoff et al. (2016) mentions that the housing supply elasticity in the study by Saiz (2010) could be an invalid instrument, as it reflects both supply and demand factors. To address this concern, we adopt the interaction between this elasticity of supply and the long-term mortgage rate as an instrumental variable, as it is widely used in the literature.

Columns [4] and [5] in Table 3 show the results of the implementation of our IV strategy. Consistently, we find that on average a 10% higher PDI of an MSA results in 3.14% higher productivity of firms located in this MSA (column [4]). Finally, column [5] confirms the positive and significant effect of R&D on firms' productivity through PDI while using our IV strategy. We find that on average a 10% increase in firms' R&D (PDI) results in 6.20% higher productivity.

5.3. Asset-pricing implications of urban density

Our empirical strategy adapts the analyses undertaken by İmrohoroğlu and Tüzel (2014) to study the link between firms' productivity and the cross-section of expected excess stock returns. We create ten portfolios sorted by the part of firms' productivity that comes from urban density, that is, the TFP fitted values on PDI, TFP (PDI).[6] In this section we investigate whether widely used asset-pricing models, such as the capital asset-pricing model (CAPM) and FamaFrench (FF) four-factor (MKT, SMB, HML, and MOM)[7] model capture the variation in excess returns of the TFP fitted value on PDI, TFP (PDI)-sorted portfolios. Table 4 represents the alphas and betas of TFP (PDI)-sorted portfolios for the CAPM and FamaFrench (FF) four-factor models. The betas are estimated by regressing the portfolio excess stock returns on the factors. The alphas are estimated as intercepts from the regressions of excess portfolio returns. The top half of the tables reports the results for the equal- weighted portfolios, and the bottom half reports the value-weighted portfolio results. [Insert Table 4 around here] In the top half of the table, considering equal-weighted portfolios, we find that low-TFP portfolios load heavily on SMB, whereas the loadings of the high-TFP portfolios are low. The loadings on HML are non-monotonic and not always significant. The equal-weighted low portfolios have a significantly lower loading on MKT than the high-TFP portfolios, whereas the value-weighted portfolios vice versa have this effect non-monotonically. Neither the CAPM nor the FF four-factor model completely explain the return spread: the highlow TFP portfolio has a CAPM annualized alpha of around -9.26% and an FF annualized alpha of around -7.31% in the value-weighted portfolios, and both spreads are statistically significant. Overall, these results indicate that, according to the CAPM model, building a portfolio with a long position on firms with low productivity and a short position on highly productive firms gives a positive annualized alpha of 9.26%. Similarly, considering the FF model, building a portfolio with a long position on firms with low productivity and a short position on highly productive firms produces a positive annualized alpha of 7.31%. Return spreads are lower and not significant for equal-weighted portfolios.

[6] The regression analysis of TFP on PDI, considering all the other controls including NDI, is used to predict the TFP (PDI).

[7] MKT is the excess market return; SMB is the return of a portfolio that is long in small and short in big firms; HML is the return of a portfolio that is long in high B/M and short in low B/M firms; and MOM is the average return on two high prior return portfolios minus the average return on two low prior return portfolios (Fama and French (1992) and Fama and French (1993), among others).

Moreover, we run regressions of monthly expected excess stock returns on the lagged firm-level TFP and urban density as well as other control variables. The estimates of the slope coefficients in these regressions allow us to determine the magnitude of the effect of urban density on excess stock returns. In all the specifications, the dependent variable is the residual of regressing excess monthly stock returns on the FamaFrench four factors. First, we run our analysis by considering the population density of each MSA as a proxy for urban density. Accordingly, we run the following specification for a firm's expected excess stock return:

$$\begin{aligned} Firm's\ \ excess\ \ stock\ \ return(Res)^{l}_{it} = \alpha_i + \beta.Population\ \ density + Controls^{l}_{it} + i.year + \\ i.industry + i.state + \epsilon_{it} \end{aligned} \tag{7}$$

$Control_{it}$ denotes two sets of firm-level controls and MSA level controls. At firm level, we control for (1) leverage; (2) firm size; (3) asset growth; (4) hiring rate; (5) return on assets (ROA); (6) return on equity (ROE); (7) market-to-book ratio; (8) corporate real estate holding; and (9) company age. At MSA level we control for (1) residual housing price index. We also control for the year fixed effect as well as the industry fixed effect and state fixed effect. Errors are clustered at the MSA level.

Our main goal here is to show empirically our motivation in using our two measures of urban density instead of population density as the proxy for urban density used in the existing literature. Table 5 in the model shows the individual results for population density on firms' stock returns. In Column [1] of Table 5 we study the effect of TFP as firm's productivity on firm's stock return. In Column [2] we run our main specification using population density as a measure for urban density itself and as our independent variable. The coefficient is shown to be negative and slightly non-significant. Column [3] running the same specification as column [2] controlling for firms' TFP. The results still remain not convincing. Therefore, we found no evidence about the effect of population as a measure of urban density itself in the literature, on firms' stock returns and we empirically confirm our argument that population density is not the optimal measure for urban density in explaining the effect of urban agglomeration on firms' output.

[Insert Table 5 around here]

Afterwards, considering PDI as our dependent variable, controlling for population density, we run the following specification for a firm's expected excess stock return:

$$\begin{aligned} Firm's\ \ excess\ \ stock\ \ return(Res)^{l}_{it} = \alpha_i + \beta.PDI/TFP(PDI) + Controls^{l}_{it} + i.year + \\ i.industry + i.state + \epsilon_{it} \end{aligned} \tag{8}$$

Table 6 study the effect of our measures on firms' stock returns. Here we run same specifications as in Table 5 while using PDI measure as our main dependent variable, controlling for NDI measure, both as proxies for urban density instead of population density. $Control_{it}$ denotes two sets of firm-level controls and MSA level controls. At firm level, we control for (1) leverage; (2) firm size; (3) asset growth; (4) hiring rate; (5) return on assets (ROA); (6) return on equity (ROE); (7) market-to-book ratio; (8) corporate real estate holding; and (9) company age. At MSA level we control for (1) NDI measure; (2) residual housing price index; and (3) population density. We also control for the year fixed effect as well as the industry fixed effect and state fixed effect. Errors are clustered at the MSA level. In Column [1] of Table 6 we study the effect of PDI, controlling for land availability using NDI, while in Column [2] we study the effect of TFP as firm's productivity, separately on firm's stock return. In Column [3] we run our main specification considering PDI as our independent variable controlling for firms' TFP and NDI to capture land availability. The coefficient of PDI is shown to be negatively significant. However, we didn't find significant effect for NDI as our control variable. Therefore, we confirm that our measures better explain the effect of urban agglomeration on firms' output in comparison to population density. Furthermore, here we empirically find evidence that productivity is one of the channels driving this effect. Column [4], shows the effect of PDI on stock returns through the productivity channel. In this column we consider the part of productivity caused by PDI; TFP (PDI), as our independent variable. We estimate this by calculating the fitted values of TFP on PDI. We also control for the fitted values of TFP on NDI measure. The specification for regressing a firm's TFP on PDI and NDI of the MSA in which the firm is located is used to predict the estimated fitted values of TFP on our two measures. Moreover, columns [5], [6] and [7] show that our results are robust using the IV strategy described earlier:

$$Firm's\ excess\ stock\ return(Res)^l_{it} = \alpha_i + \beta.PDI/TFP(PDI) + \gamma.Elasticity^l.IR + Controls^l_{it} + i.year + i.industry + i.state + \epsilon_{it} \tag{9}$$

where $Elasticity^l$ measures the constraints on the land supply at the MSA level and IR measures the nationwide real interest rate at which banks refinance their home loans. $Control_{it}$ are the same as in previous specifications. Columns [5], [6] and [7] of Table 6 report the results when we implement the IV strategy to the regressions in columns [1], [3] and [4] where PDI is instrumented by the interaction of the interest rate and the local constraints on land supply.

[Insert Table 6 around here]

We confirm the important negative and significant effect of PDI on stock returns after implementing the IV strategy. Specifically, we find that on average a 10% higher PDI of an MSA results in a 0.29% lower excess stock return of firms located in this MSA. Similarly, in column [5] we find a significant and negative effect of the part of PDI caused by a potential density increase, TFP (PDI), on firms' excess stock return. The results show that on average a 10% higher TFP (PDI) of a firm results in a 1.76% lower excess stock return of that firm.

Eventually in Table 7 we confirm our argument by showing that the main part of the effect of firms' productivity on their stock returns is caused by the PDI of their MSA among all other unobservable variables. In order to do so, the potential influence caused by other unobservable variables on firms' TFP is compared with the effect of a potential density increase. We provide evidence that shows that the main part of firms' TFP effect on their stock returns is caused by the potential density increase among all the other unobservable variables. Accordingly, we run the following specification for a firm's expected excess stock return:

$$\begin{aligned} Firm's\ excess\ stock\ return(Res)^l_{it} = \alpha_i + \beta.PDI/TFP(PDI) + \gamma.TFP(\overline{PDI\&NDI}) + \\ Controls^l_{it} + i.year + i.industry + i.state + \epsilon_{it} \end{aligned} \tag{10}$$

$Controls_{it}$ are the same as in previous specifications. We consider the effect of unobservable variables, $\overline{TFP(PDI\&NDI)}$, by computing the whole estimated TFP minus the TFP fitted values on PDI and NDI. Following the addition of this effect, we report the significant negative effect of the potential density increase. Comparing the corresponding coefficients reported in column [2], we conclude that the main part of TFP's effect on firms' stock return comes from the potential density increase and a comparatively small and negligible part is caused by other unobservable variables. Furthermore, this coefficient is shown to be less significant than the effect of TFP (PDI). Column [3] reports the results implementing the IV strategy where PDI is instrumented by the interaction of the interest rate and the local constraints on land supply:

$$\begin{aligned} Firm's\ excess\ stock\ return(Res)^l_{it} = \alpha_i + \beta.PDI/TFP(PDI) + \gamma.Elasticity^l.IR + \delta.Elasticity^l.IR \\ + Controls^l_{it} + i.year + i.industry + i.state + \epsilon_{it} \end{aligned} \tag{11}$$

This Column shows that our results are robust to this identification strategy.

[Insert Table 7 around here]

6. Robustness Tests

In this section we provide robustness tests for our main results presented in section 5. First, we address the concern about the fact that firms with better growth and productivity choose to locate in more agglomerated or high-tech cities, which could bias our results. We run a second robustness test to address the concern that there is a potential bias for big global firms to have their RD center located in a different area from their headquarters. The results of these two robustness checks are reported in Table 8. Finally, we perform a third robustness check in which we discuss different types of industries that cause different levels of firms' need regarding the innovation, flow of ideas, and spillovers existing in the area where they are located. Table 9 reports the results for different industry classifications.

6.1. Choice of firm location

We address the concern for the potential bias of the idea that firms with better growth and productivity choose to locate in more agglomerated or high-tech cities. Almazan et al. (2010) show that clusters are likely to attract firms with attributes that make them more likely to succeed, which can potentially be an issue for young firms that have recently chosen their location. However, considering that the unobserved characteristics that may influence a firm's location choice become less important over time, the observed effect on the productivity of older firms that chose locations many years ago is unlikely to arise because of a cluster selection effect. For this reason it is interesting to explore whether the relation between firms' stock return and older firms' location is indeed consistent with what we observe for the entire sample.

Columns [1] and [2] of the Table reftable8 show the results of our main specifications for the entire sample, while columns [7] and [8] report these results after implementing our IV strategy as explained earlier. Following Almazan et al. (2010) approach, we report the results of this test in columns [3] and [4] of Table reftable8 for the subsample of firms that are at least 10 years old. The consistency of the effects between the older firms and the entire sample shows that the results are not likely to be due to better-quality firms locating in denser urban areas. Moreover, Columns [9] and [10] report these results after implementing our instrumental variable (IV) strategy. Therefore, consistent with the previous findings, we show that on average a 10% higher PDI of an MSA results in a 0.36% lower excess stock return of firms located in this MSA for the subsample of firms that are at least 10 years old. Similarly, we show the significant and negative effect of the part of PDI caused by the urban density effect, TFP (PDI), on firms' excess stock return.

6.2. Firms' R&D center and their headquarters location

There could be the concern that global companies could locate their RD center in a different area from their headquarters. We define a firm's location as the location of its headquarters. Although a firm's headquarters is often separated from its operations by hundreds or even thousands of miles, this separation may help rather than hurt our ability to identify the types of spillovers that are the focus of this study. Moreover, to confirm our argument, we start by splitting the sample into small and large firms. We consider small firms as those in the lower three quartiles of size from the whole sample. We run regressions for the subsample of small firms and show that the estimated coefficient remains significant and stronger for the subsample of small firms, which can also be considered as the more RD-intensive ones, in comparison with the whole sample. The results are reported in Table reftable8. Columns [1] and [2] report the results for the whole sample, while columns [5] and [6] show the results for the sub sample of small firms. In columns [11] and [12] we implement the instrumental variable (IV) strategy to these regressions and we show that our results remain robust to this identification strategy.

[Insert Table reftable8 around here]

6.3. Industry classification according to innovation

In this section, we study our results for high-tech industries (i.e., innovation firms), which benefit more from the innovation, flow of ideas, and spillovers existed in the area where they are located. Consistent with the literature, we consider firms belong to Electronic Computers, Electrical Machinery, Transportation Equipment, Instruments, Software and Data Processing Services, as high-tech firms (Cortright Mayer, 2001). Here we support our argument by studying the effect of agglomeration and potential density increase on the productivity of high-tech firms and their stock returns, considering the influence of the innovation, flow of ideas, and spillovers in the area. Our results show that this effect is stronger and significant for the subsample of high-tech firms which benefit more from the innovation and spillover effects, in comparison to the whole sample. Specifically, we confirm this by showing that this effect does not work and turns out to be non-significant for the rest of the sample as less-innovative firms. Table reftable9 reports the results of our main specification for different industry classifications. Columns [1] and [2] show our main results for the entire sample, while columns [3] and [4] show the results for the group of firms belong to electronic electronic computers, electrical machinery, transportation equipment, instruments, software and data processing services, (i.e., innovation firms), consistent with the literature. In columns [5] and [6] we see these results for the rest of sample (i.e., less innovation firms). Columns [7], [8], [9] and [10] display the same regressions as columns [1], [2], [3] and [4] plus the implementation of the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and the local constraints on the land supply.

[Insert Table reftable9 around here]

7. Conclusions

The positive effect of agglomeration on productivity has long been documented and quantified by studying spatial patterns in wages and land rents. Such a positive effect on productivity is driven by knowledge spillovers that accelerate the adoption of new technologies, the increase of opportunities from specialization, as well as the existence of economies of scale and low transportation costs. In this paper we present new evidence regarding the relationship between urban density characteristics of the MSA where the firm is located and its stock returns. To do so, we create a measure of potential density increase (PDI) using high-resolution satellite images from Google Earth with a computer vision algorithm. This measure of PDI is estimated as the proportion of the area in the total area within a one-hour drive from the center of the MSA that could rapidly increase its density in each MSA. Our measure of PDI captures the fact that areas with low urban density (i.e., areas with existing facilities and infrastructure, but low density) can potentially increase its density faster than non-developed areas and undevelopable areas.

We find that, on average, a 10% higher PDI of an MSA results in 0.29% lower excess stock returns for firms located in the MSA. We argue that RD is the mechanism behind this effect, and we show that a firm's RD has a positive and significant effect on firm's productivity. In order to confirm these results, we provide evidence that the effect of firms' productivity on stock returns is caused by the potential density increase among all other unobservable variables. Comparing their corresponding coefficients, the main part of TFP's effect on firms' stock return comes from the potential density increase and a small and negligible part in comparison is caused by other unobservable variables. Furthermore, this coefficient is shown to be less significant than the coefficient of TFP (PDI). Our results remain robust to the use of instrumental variables, as well as to several other checks addressing additional potential concerns.

Our results have important implications for managers, entrepreneurs, investors, and local authorities. Managers and entrepreneurs must take into account that the urban density characteristics of the area where they decide to locate their firms have an impact on firms' stock returns. The location of the firm in an MSA that can quickly increase its density is perceived as a low risk when compared to the location in an MSA with a low potential of density increase and, therefore, leads to lower excess stock returns. Investors can optimize their portfolios using measures of PDI in order to improve their performance. Finally, local authorities can develop urban plans to provide areas that can rapidly increase the density of MSAs.

References

Almazan, A., A. De Motta, S. Titman, and V. Uysal (2010). Financial structure, acquisition opportunities, and firm locations. *The Journal of Finance 65*(2), 529–563.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics 121*(2), 587–633.

Carlino, G. A., S. Chatterjee, and R. M. Hunt (2007). Urban density and the rate of invention. *Journal of Urban Economics 61*(3), 389–419.

Chaney, T., D. Sraer, and D. Thesmar (2012). The collateral channel: How real estate shocks affect corporate investment. *American Economic Review 102*(6), 2381–2409.

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

Cvijanović, D. (2014). Real estate prices and firm capital structure. *The Review of Financial Studies 27*(9), 2690–2735.

Davidoff, T. et al. (2016). Supply constraints are not valid instrumental variables for home prices because they are correlated with many demand factors. *Critical Finance Review 5*(2), 177–206.

Davis, M. A., J. D. Fisher, and T. M. Whited (2014). Macroeconomic implications of agglomeration. *Econometrica 82*(2), 731–764.

Dougal, C., C. A. Parsons, and S. Titman (2015). Urban vibrancy and corporate growth. *The Journal of Finance 70*(1), 163–210.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, Volume 4, pp. 2063–2117. Elsevier.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *the Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics 33*(1), 3–56.

Garcia, D. and Ø. Norli (2012). Geographic dispersion and stock returns. *Journal of Financial Economics 106*(3), 547–565.

Glaeser, E. L., J. Gyourko, and A. Saiz (2008). Housing supply and housing bubbles. Technical report, National Bureau of Economic Research.

Glaeser, E. L., J. Gyourko, and R. E. Saks (2005). Urban growth and housing supply. *Journal of economic geography 6*(1), 71–89.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. *Journal of economic geography 1*(1), 27–50.

Hall, B. H. (1993). The stock market's valuation of r&d investment during the 1980's. *The American Economic Review 83*(2), 259–264.

Henderson, J. V., T. Regan, and A. J. Venables (2016). Building the city: sunk capital, sequencing, and institutional frictions.

Himmelberg, C., C. Mayer, and T. Sinai (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives 19*(4), 67–92.

İmrohoroğlu, A. and Ş. Tüzel (2014). Firm-level productivity, risk, and return. *Management Science 60*(8), 2073–2090.

Knudsen, B., R. Florida, K. Stolarick, and G. Gates. 2008), density and creativity in us regions. In *Annals of the Association of American Geographers*. Citeseer.

Krugman, P. (1991). History and industry location: the case of the manufacturing belt. *The American Economic Review 81*(2), 80–83.

Lieberman, M. B. and J. Kang (2008). How to measure company productivity using value-added: A focus on pohang steel (posco). *Asia Pacific Journal of Management 25*(2), 209–224.

Maantay, J. A., A. R. Maroko, and C. Herrmann (2007). Mapping population distribution in the urban environment: The cadastral-based expert dasymetric system (ceds). *Cartography and Geographic Information Science 34*(2), 77–102.

Marshall, A. (1890). " some aspects of competition." the address of the president of section f–economic science and statistics–of the british association, at the sixtiet meeting, held at leeds, in september, 1890. *Journal of the Royal Statistical Society 53*(4), 612–643.

Mennis, J. (2003). Generating surface models of population using dasymetric mapping. *The Professional Geographer 55*(1), 31–42.

Mian, A. and A. Sufi (2011). House prices, home equity-based borrowing, and the us household leverage crisis. *American Economic Review 101*(5), 2132–56.

Naika, N., S. D. Kominersb, R. Raskara, E. L. Glaeserc, and C. A. Hidalgoa. The dynamics of physical urban change.

Pakes, A. and S. Olley (1995). A limit theorem for a smooth class of semiparametric estimators. *Journal of Econometrics 65*(1), 295–332.

Pirinsky, C. and Q. Wang (2006). Does corporate headquarters location matter for stock returns? *The Journal of Finance 61*(4), 1991–2015.

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: evidence from the cities. *Journal of urban economics 34*(3), 380–400.

Rosenthal, S. S. and W. C. Strange (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, Volume 4, pp. 2119–2171. Elsevier.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics 125*(3), 1253–1296.

Sun, X., H. Li, and V. Ghosal (2017). Firm-level human capital and innovation: evidence from china.

Sutton, P. C., C. Elvidge, and T. Obremski (2003). Building and evaluating models to estimate ambient population density. *Photogrammetric Engineering & Remote Sensing 69*(5), 545–553.

Sveikauskas, L. (1975). The productivity of cities. *The Quarterly Journal of Economics 89*(3), 393–413.

Tuzel, S. (2010). Corporate real estate holdings and the cross-section of stock returns. *The Review of Financial Studies 23*(6), 2268–2302.

Tuzel, S. and M. B. Zhang (2017). Local risk, local factors, and asset prices. *The Journal of Finance 72*(1), 325–370.

Table 1 Measures of urban density characteristics.

Rank	MSA	Measure	
		PDI	NDI
1	Chicago, IL	81.60%	16.03%
2	Los AngelesLong Beach, CA	81.09%	56.17%
3	San Diego, CA	57.66%	58.09%
4	RiversideSan Bernardino, CA	53.06%	8.43%
5	San Francisco, CA	51.00%	63.45%
6	New York, NY	48.55%	13.75%
7	Oakland, CA	43.39%	61.91%
8	SeattleBellevueEverett, WA	39.24%	27.41%
9	Ventura, CA	38.88%	80.56%
10	San Jose, CA	36.38%	69.97%
11	Newark, NJ	36.25%	59.67%
12	Salt Lake CityOgden, UT	35.77%	63.35%
13	Fort WorthArlington, TX	32.98%	16.45%
14	Denver, CO	31.74%	26.76%
15	New Orleans, LA	31.23%	35.40%
16	CharlestonNorth Charleston, SC	30.87%	33.75%
17	Tacoma, WA	30.19%	22.92%
18	BostonWorcesterLawrenceLowellBrockton, MA-NH	28.94%	11.47%
19	Fresno, CA	28.52%	46.31%
20	Sarasota-Bradenton, FL	26.35%	19.70%
21	VallejoFairfieldNapa, CA	26.23%	44.82%
22	TampaSt. PetersburgClearwater, FL	26.16%	20.80%
23	PortlandVancouver, OR-WA	25.83%	23.67%
24	New HavenBridgeportStamfordDanburyWaterbury, CT	22.43%	6.97%
25	West Palm BeachBoca Raton, FL	22.22%	32.02%
26	ProvidenceWarwickPawtucket, RI	21.98%	10.57%
27	Omaha, NE-IA	21.39%	6.53%
28	RaleighDurhamChapel Hill, NC	20.03%	7.12%
29	Dallas, TX	19.91%	8.62%
30	Indianapolis, IN	19.76%	3.58%
31	St. Louis, MO-IL	19.48%	6.58%
32	Detroit, MI	19.34%	7.62%
33	WilmingtonNewark, DE-MD	19.12%	14.18%
34	Nashville, TN	18.22%	6.03%
35	Jersey City, NJ	17.79%	12.58%
36	Gary, IN	17.73%	12.97%
37	GreensboroWinstonSalemHigh Point, NC	16.97%	6.72%
38	NorfolkVirginia BeachNew Port News, VA-NC	16.83%	34.77%
39	Atlanta, GA	16.78%	7.02%
40	Columbia, SC	16.72%	10.92%
41	Columbus, OH	16.32%	2.58%
42	Fort Lauderdale, FL	15.83%	12.76%
43	Philadelphia, PA-NJ	15.27%	9.78%
44	San Antonio, TX	14.93%	2.54%
45	Baltimore, MD	14.64%	10.76%
46	Miami, FL	14.40%	17.81%
47	Mobile, AL	14.12%	15.19%
48	PhoenixMesa, AZ	13.61%	43.48%
49	MinneapolisSt. Paul, MN-WI	13.41%	5.93%
50	Springfield, MA	12.88%	2.98%

Table 1 Measures of urban density characteristics (cont.)

Rank	MSA	Measure	
		PDI	NDI
51	Hartford, CT	12.48%	10.80%
52	Tucson, AZ	12.45%	26.60%
53	ClevelandLorain-Elyria, OH	12.41%	5.50%
54	Knoxville, TN	12.34%	9.29%
55	Cincinnati, OH-KY-IN	12.30%	4.39%
56	Akron, OH	11.36%	4.78%
57	HarrisburgLebanonCarlisle, PA	11.17%	3.54%
58	DaytonSpringfield, OH	11.14%	17.65%
59	Birmingham, AL	10.82%	9.90%
60	Kansas City, MO-KS	10.59%	5.01%
61	Memphis, TN-AR-MS	10.47%	11.45%
62	Tulsa, OK	10.26%	6.04%
63	Jacksonville, FL	10.15%	8.55%
64	Rochester, NY	10.07%	3.19%
65	Las Vegas, NV-AZ	10.01%	42.08%
66	AlbanySchenectadyTroy, NY	9.62%	8.65%
67	Grand RapidsMuskegonHolland, MI	9.14%	14.15%
68	Baton Rouge, LA	9.04%	10.75%
69	Washington, DC-MD-VA-WV	8.23%	5.74%
70	AllentownBethlehemEaston, PA	8.10%	2.67%
71	Orlando, FL	7.97%	12.91%
72	Pittsburgh, PA	7.76%	3.75%
73	Austin, San Marcos, TX	7.65%	2.06%
74	Ann Arbor, MI	7.45%	17.35%
75	RichmondPetersburg, VA	7.25%	6.82%
76	El Paso, TX	6.88%	17.80%
77	MilwaukeeWaukesha, WI	6.66%	15.22%
78	Colorado Springs, CO	6.06%	28.87%
79	GreenvilleSpartanburgAnderson, SC	5.92%	2.60%
80	Houston, TX	5.92%	22.92%
81	Wichita, KS	5.90%	2.34%
82	Oklahoma City, OK	5.76%	3.23%
83	Syracuse, NY	5.48%	5.13%
84	StocktonLodi, CA	4.91%	9.72%
85	Bakersfield, CA	4.41%	28.84%
86	Youngstown, Warren, OH	4.23%	3.30%
87	Toledo, OH	4.19%	4.79%
88	Little RockNorth Little Rock, AR	4.05%	8.17%
89	Albuquerque, NM	3.70%	22.79%
90	McAllenEdinburgMission, TX	3.59%	4.92%
91	ScrantonWilkesBarreHazleton, PA	3.19%	5.55%
92	Fort Wayne, IN	2.95%	2.51%
93	Louisville, KY-IN	2.72%	14.46%
94	BuffaloNiagara Falls, NY	2.30%	9.34%
95	CharlotteGastoniaRock Hill, NC-SC	1.47%	2.32%

Note: Measures of potential density increase (PDI) and non-potential density increase (NDI) for metropolitan statistical areas (MSAs) with a population of over 500,000 inhabitants.

Table 2 Summary statistics.

Variable	Mean	Median	Std dev.	25th percent.	75th percent.	Obs.	Definition/Unit
Measures of urban density, geography, regulation, and macroeconomic variables:							
PDI	0.2575	0.1603	0.2216	0.0762	0.3373	110296	Measure of potential density increase: Proportion of area in the total area within a one-hour drive from the center of the MSA that could rapidly increase its density.
NDI	0.3352	0.2616	0.2485	0.1341	0.5306	110296	Measure of non-density increase: Proportion of area in the total area within a one-hour drive from the center of the MSA that cannot rapidly increase its density, either because it is already highly dense or because it is undevelopable.
Saiz	0.3387	0.339	0.2226	0.1269	0.4733	110,296	Measure of geographical constraints from Saiz (2010).
Population density	686.2	511.1	1,134.12	267.2495	739.0379	110,296	Number of MSAs inhabitants divided by the total MSA square kilometer area.
Real estate price index	215.7444	218.69	41.7505	185.25	235.37	107,358	Residential home price index (HPI) from the website of the Federal Housing Finance Association (FHFA).
Firm-level variables:							
Expected excess returns	0.0086	0.0046	0.1376	-0.0555	0.0631	110,296	
Firm size	174050.4	704.9644	5614081	174.3874	2770.277	110,296	Market capitalization is defined as the common share outstanding multiplied by the price bid/ask average in dollars.
Value added	8.335	0.7112	34.1469	0.0579	3.8563	110,296	Value added is defined as the operating income before depreciation plus total staff expenses, all divided by the gross domestic product implicit price deflator.
Labor	13.1178	2.022	40.5669	0.415	8.3455	110,296	Number of employees.
Capital stock	31.8498	2.0731	158493	0.3311	11.0617	110,296	Firm-year PPE total divided by the gross private domestic investment implicit price deflator.
Asset growth	0.147	0.0507	0.6247	-0.0272	0.1585	107,248	Difference in the current and the lagged total assets divided by the lagged total assets.
Leverage	0.2795	0.1866	1.4559	0	0.4466	109,734	Leverage is the total long-term debt divided by the total long-term debt plus the common/ordinary equity total.
ROA	-0.0241	0.0352	0.7372	-0.025	0.0758	107,677	ROA is the income before extraordinary items minus dividends preferred plus income taxes, deferred, all divided by the total assets.
ROE	0.0225	0.076	2.9323	-0.054	0.1547	110,296	ROE is the income before extraordinary items, available for common stock divided by total stockholders equity.
Hiring rate	0.09273	0.0331	0.8045	-0.024	0.1195	107,248	Difference in the current and the lagged number of employees divided by the lagged number of employees.
Inventory growth	0.161	0.06	0.9725	-0.0511	0.2022	84,315	
Market-to-book ratio	3.1339	2.1724	51.4159	1.2966	3.8628	108,114	
Company age	22.6329	18	16.7224	10	30	110,296	Years
Real estate ratio	0.6072	0.4942	0.7351	0.1764	0.8199	110,296	Real estate ratio is defined as the buildings plus capitalized lease divided by the net PPE.
Measures of R&D at the firm level:							
R&D	1.1924	1	5.9705	0.194	1	110,296	R&D is the research and development expense divided by the total PPE.
State R&D	23,579.08	10,954	29,708.80	4,985	17,395	94,591	Expenditure in each US state in millions of dollars.

Note: This table provides the summary statistics for the main variables that we use in the paper with a short description.

Table 3 Urban density and firms' productivity.

	OLS [1]	OLS [2]	OLS [3]	IV [4]	IV [5]
PDI	0.20857***	0. 22241***		0.31403***	
	(0.0322)	(0.0325)		(0.0918)	
R&D (PDI)			0.43935***		0.62035***
			(0.643)		(0.1813)
Controlling for NDI	No	Yes	No	Yes	No
Controlling for R&D (NDI)	No	No	Yes	No	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes
Controlling for Housing supply elasticity*year	No	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
US state FE	Yes	Yes	Yes	Yes	Yes
Observations	103,199	103,199	103,199	103,199	103,199
R^2	0.102	0.102	0.1019	0.1021	0.1021

Note: This table shows the effect of our urban density measures on firms productivity. The dependent variable is the TFP estimated by the production function. While column [1] shows the effect of our PDI measure on firms productivity. Column [2] shows the effect of PDI on firms productivity, controlling for NDI. Column [3] shows the effect of firms R&D caused by urban density on firms productivity. Our main independent variable here is the fitted value of regressing R&D on PDI; we also control for the fitted value of regressing R&D on NDI. Standard errors are reported in parentheses. Columns [4] and [5] represent the same regressions as columns [2] and [3] plus the implementation of the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and local constraints on the land supply. Consistent with the literature, we control for the interaction between housing supply elasticity and year in our instrumental regressions to capture the time trend of interest rates which explains most of the correlation between home prices. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4 Urban density and the cross-section of stock returns.

	Low	2	3	4	5	6	7	8	9	High	HighLow
Equal-weighted portfolios											
Alpha	0.01247***	0.01364***	0.01124***	0.00971***	0.00848***	0.00948***	0.01067***	0.00753***	0.01027***	0.01072***	-0.00175
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0014)	
MKT	0.01305***	0.01598***	0.01497***	0.01968***	0.01568***	0.02006***	0.02114***	0.01673***	0.02117***	0.01818***	
	(0.0027)	(0.0024)	(0.0023)	(0.0021)	(0.0024)	(0.0022)	(0.0021)	(0.0019)	(0.002)	(0.0018)	
Alpha	0.01367***	0.01536***	0.01320***	0.01159***	0.01015***	0.01150***	0.01276***	0.00944***	0.01217***	0.01259***	-0.00108
	(0.0014)	(0.0013)	(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	
MKT	0.01788***	0.02404***	0.02182***	0.02428***	0.02251***	0.02483***	0.02854***	0.01987***	0.02136***	0.01777***	
	(0.0036)	(0.0032)	(0.003)	(0.0028)	(0.0032)	(0.0029)	(0.0028)	(0.0025)	(0.0026)	(0.0024)	
SMB	0.00904***	0.01063***	0.00885***	0.00969***	0.00871***	0.00791***	0.00828***	0.01025***	0.00966***	0.00764***	
	(0.0016)	(0.0014)	(0.0013)	(0.0013)	(0.0014)	(0.0013)	(0.0012)	(0.0011)	(0.0012)	(0.0011)	
HML	0.01363*	0.01945***	0.02040***	0.01167**	0.01600***	0.01580***	0.01890***	0.00384	-0.00042	-0.00186	
	(0.0054)	(0.0047)	(0.0044)	(0.0042)	(0.0048)	(0.0043)	(0.0041)	(0.0037)	(0.0039)	(0.0035)	
MOM	0.00318	0.00586**	0.00479**	0.00278	0.00473*	0.00278	0.00522**	0.00165	-0.00104	-0.00143	
	(0.0021)	(0.0019)	(0.0017)	(0.0016)	(0.0019)	(0.0017)	(0.0016)	(0.0015)	(0.0015)	(0.0014)	
Value-weighted portfolios											
Alpha	0.01511***	0.01111***	0.00861***	0.00951***	0.01111***	0.01247***	0.01260***	0.01022***	0.00645***	0.00704***	-0.00807***
	(0.0018)	(0.0015)	(0.0013)	(0.001)	(0.0009)	(0.0007)	(0.0009)	(0.0013)	(0.0016)	(0.0019)	
MKT	0.00903**	0.01146***	0.01687***	0.01837***	0.02172***	0.01714***	0.01799***	0.01798***	0.02406***	0.01984***	
	(0.0035)	(0.0027)	(0.0022)	(0.0018)	(0.0014)	(0.0011)	(0.0014)	(0.0019)	(0.0024)	(0.0028)	
Alpha	0.01623***	0.01286***	0.01080***	0.01128***	0.01260***	0.01347***	0.01407***	0.01205***	0.00919***	0.00992***	-0.00631**
	(0.0018)	(0.0015)	(0.0013)	(0.001)	(0.0009)	(0.0007)	(0.0009)	(0.0013)	(0.0017)	(0.0019)	
MKT	0.01205**	0.02191***	0.02152***	0.02667***	0.02689***	0.02066***	0.02352***	0.02006***	0.02200***	0.02394***	
	(0.0046)	(0.0036)	(0.0029)	(0.0024)	(0.0018)	(0.0014)	(0.0018)	(0.0024)	(0.0032)	(0.0036)	
SMB	0.00440*	0.00984***	0.01186***	0.01173***	0.00758***	0.00346***	0.00711***	0.00858***	0.01370***	0.01173***	
	(0.0021)	(0.0016)	(0.0013)	(0.0011)	(0.0008)	(0.0006)	(0.0008)	(0.0011)	(0.0014)	(0.0016)	
HML	0.01330*	0.02582***	0.01356**	0.01834***	0.01181***	0.00861***	0.01141***	0.0038	-0.00463	0.01058*	
	(0.0068)	(0.0053)	(0.0043)	(0.0036)	(0.0027)	(0.0021)	(0.0027)	(0.0036)	(0.0047)	(0.0054)	
MOM	0.00177	0.00783***	0.00258	0.00616***	0.00350**	0.00263**	0.00404***	0.00069	-0.00368*	0.0017	
	(0.0027)	(0.0021)	(0.0017)	(0.0014)	(0.0011)	(0.0008)	(0.0011)	(0.0014)	(0.0019)	(0.0021)	

Note: This table presents the regressions of equal-weighted and value-weighted expected excess portfolio returns on various factor returns. The MKT, SMB, HML, and MOM factors are taken from Ken Frenchs website (http://mba.tuck.darmouth.edu). The portfolios are sorted on the part of TFP that comes from urban density: the fitted values of regressing estimated productivity on PDI considering all other control variables including NDI. Returns are measured from July 2011 to June 2015. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 5 The effect of population density on firms' stock returns.

	OLS [1]	OLS [2]	OLS [3]
TFP	-0.00160**		-0.00159**
	(0.0007)		(0.0007)
Log population density		-0.00019	0.00007
		(0.0009)	(0.0009)
Controls	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
US state FE	Yes	Yes	Yes
Cluster error, MSA level	Yes	Yes	Yes
Observation	103,199	103,199	103,199
R^2	0.0102	0.01	0.0102

Note: This table studies the effect of population density as the proxy used in the existing literature for urban density on firms excess stock return. The dependent variable is the residuals of the expected excess stock return, excluding FamaFrench factors. Column [1] shows the effect of TFP as firms productivity on firms stock return. In Column [2] we report the effect of population density on firms stock returns. The coefficient is shown to be negative and slightly non-significant. Column [3] running the same specification as column [2] controlling for firms TFP. Controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; and residual housing price index. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 6 Urban density and firms' stock return.

	OLS [1]	OLS [2]	OLS [3]	OLS [4]	IV [5]	IV [6]	IV [7]
PDI	-0.01026***		-0.00990***		-0.02964**	-0.02916**	
	(0.0034)		(0.0033)		(0.0143)	(0.0144)	
TFP		-0.00159**	-0.00158**			-0.00154**	
		(0.0007)	(0.0007)			(0.0007)	
TFP (PDI)				-0.06088***			-0.17595**
				(0.0203)			(0.0855)
Controlling for NDI	Yes	No	Yes	No	Yes	Yes	No
Controlling for TFP (NDI)	No	No	No	Yes	No	No	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for Housing supply elasticity*year	No	No	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
US state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster error, MSA level	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observation	103,199	103,199	103,199	103,199	103,199	103,199	103,199
R^2	0.0102	0.01	0.0102	0.0101	0.0099	0.0101	0.0099

Note: This table studies the effect of our measures of urban density on firms stock return. The dependent variable is the residuals of the expected excess stock return, excluding FamaFrench factors. Here we run same specifications as in Table V-I while using PDI measure as our independent variable, controlling for population density. Column [1], shows the effect of PDI on stock return, while considering land availability by controlling for NDI. Column [2] separately shows the effect of TFP on stock return. In column [3] we see the results of running the main specification considering PDI and TFP together, controlling for land availability by using NDI. Our independent variable in this column is PDI. Column [4] studies the productivity channel and show the effect of partial productivity caused by urban density on firms expected excess stock return. The independent variable here is the fitted value of regressing TFP on PDI, while we control for the fitted value of regressing TFP on NDI. Columns [5], [6] and [7] represent the same regressions as columns [1], [3] and [4] after implementing the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and the local constraints on the land supply. Consistent with the literature, we control for the interaction between housing supply elasticity and year in our instrumental regressions to capture the time trend of interest rates which explains most of the correlation between home prices. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7 Other unobservable variable and firms' stock return.

	OLS [1]	OLS [2]	IV [3]
TFP (PDI)	-0.06088***	-0.06038***	-0.1745**
	(0.0203)	(0.0201)	(0.0853)
TFP($\overline{PDI\&NDI}$)		-0.00158**	-0.00157**
		(0.0007)	(0.0007)
Controlling for TFP (NDI)	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes
Controlling for Housing supply elasticity*year	No	No	Yes
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
US state FE	Yes	Yes	Yes
Cluster error, MSA level	Yes	Yes	Yes
Observation	103,199	103,199	103,199
R^2	0.0101	0.0102	0.0101

Note: This table study the potential effect from other unobservable variables rather than urban density on stock return. The dependent variable is the residuals of the expected excess stock return, excluding FamaFrench factors. Column [1] shows the effect of partial productivity caused by PDI on firms expected excess stock return, while controlling for the part of productivity caused by NDI. The regression used in column [2] confirms our argument that the main part of TFP is caused by urban density due to the small significant effect of other unobservable variables (excluding labor, capital, and urban density), which can capture productivity on firms expected excess stock return. We consider the effect of unobservable variables, (TFP($\overline{PDI\&NDI}$)), by computing the whole estimated TFP minus the TFP fitted values on PDI and NDI. Following the addition of this effect, we compare the corresponding coefficients reported in column [2]. Column [3] represents the same regression as columns [2] plus the implementation of the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and the local constraints on the land supply. Consistent with the literature, we control for the interaction between housing supply elasticity and year in our instrumental regressions to capture the time trend of interest rates which explains most of the correlation between home prices. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8 Choice of firm location; Firms' R&D center and their headquarters location: Robustness tests 1 and 2.

	OLS [1]	OLS [2]	OLS [3]	OLS [4]	OLS [5]	OLS [6]	IV [7]	IV [8]	IV [9]	IV [10]	IV [11]	IV [12]
PDI	-0.00990***		-0.01178***		-0.01160***		-0.02916**		-0.03594***		-0.03171*	
	(0.0033)		(0.0037)		(0.0036)		(0.0144)		(0.0136)		(0.0168)	
TFP(PDI)		-0.06088***		-0.07094***		-0.07055***		-0.1745**		-0.21419***		-0.18989*
		(0.0203)		(0.0182)		(0.0216)		(0.0853)		(0.0808)		(0.0168)
TFP($\overline{PDI\&NDI}$)		-0.00158**		-0.00099		-0.00168**		-0.00157**		-0.00097		-0.00168**
		(0.0007)		(0.0006)		(0.0007)		(0.0007)		(0.0006)		(0.0007)
Controlling for NDI	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Controlling for TFP (NDI)	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for Housing supply elasticity*year	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
US state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster error, MSA level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firms in the sample	Full sample	Full sample	age ¿10	age ¿10	Small firms	Small firms	Full sample	Full sample	age ¿10	age ¿10	Small firms	Small firms
Observations	103,199	103,199	78,438	78,438	77,406	77,406	103,199	103,199	78,438	78,438	77,406	77,406
R^2	0.0102	0.0102	0.0118	0.0118	0.0108	0.0108	0.0101	0.0101	0.0115	0.0115	0.0106	0.0106

Note: This table study the robustness of our results regarding the choice of firm location and firm size. The dependent variable is the residuals of the expected excess stock return, excluding FamaFrench factors. Columns [1] and [2] show our main results for the entire sample. Columns [2] and [3] show the consistency of the relation between a firms stock return and its location for older firms. Capturing the potential bias of the idea that firms with better growth and productivity choose to locate in more agglomerated or high-tech cities, to the extent that the unobserved characteristics that may influence a firms location choice become less important over time, the observed effect on the stock return of older firms that chose locations many years ago is unlikely to arise because of a cluster selection effect. For this reason, we explore whether the relation between a firms stock return and its location for older firms for all the previous different scenarios is indeed consistent with what we observe for the entire sample. These columns report the baseline regressions for the subsample of firms aged at least 10 years. Columns [5] and [6] show the results for the sub sample of small firms after splitting our sample into small and large firms. We consider small firms as the ones in the lower three quartiles of size from the whole sample. Columns [7], [8], [9], [10], [11] and [12] represent the same regressions as columns [1], [2], [3], [4], [5] and [6] plus the implementation of the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and the local constraints on the land supply. Consistent with the literature, we control for the interaction between housing supply elasticity and year in our instrumental regressions to capture the time trend of interest rates which explains most of the correlation between home prices. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 9 Industry classification: Robustness test 3.

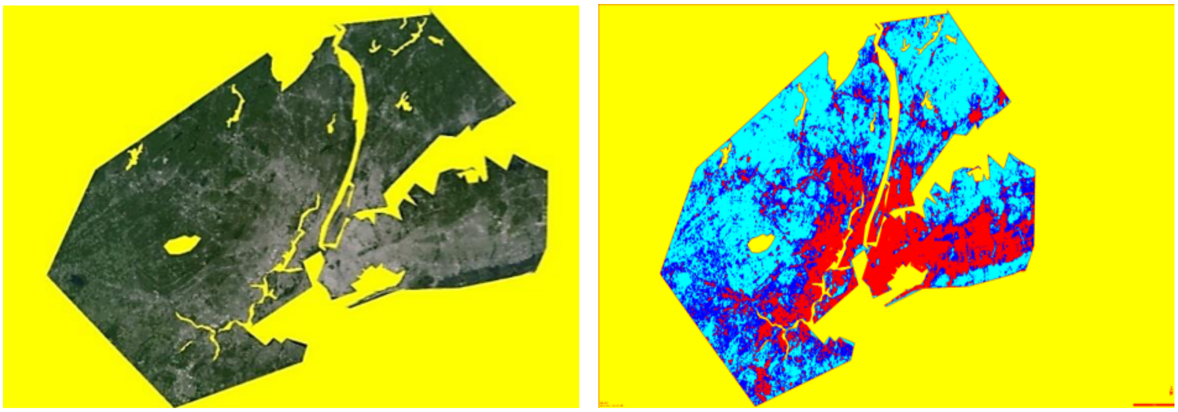
	OLS [1]	OLS [2]	OLS [3]	OLS [4]	OLS [5]	OLS [6]	IV [7]	IV [8]	IV [9]	IV [10]
PDI	-0.00990***		-0.0148***		-0.006		-0.02916**		-0.04426**	
	(0.0033)		(0.0036)		(0.0049)		(0.0144)		(0.0203)	
TFP (PDI)		-0.06088***		-0.0895***		-0.0377		-0.1745**		-0.2640**
		(0.0203)		(0.0214)		(0.0294)		(0.0853)		(0.1211)
TFP($\overline{PDI\&NDI}$)		-0.00158**		-0.0013		-0.0019		-0.00157**		-0.00134
		(0.0007)		(0.0019)		(0.0008)		(0.0007)		(0.001)
Controlling for NDI	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Controlling for TFP (NDI)	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for Housing supply elasticity*year	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
US state FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster error, MSA level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firms in the sample	Full sample	Full sample	High-tech	High-tech	Less-innovative	Less-innovative	Full sample	Full sample	High-tech	High-tech
Observation	103,199	103,199	36,962	36,962	66,237	66,237	103,199	103,199	36,962	36,962
R^2	0.0102	0.0102	0.0129	0.0129	0.0097	0.0097	0.0101	0.0101	0.0124	0.0129

Note: This table study our results for different industry classifications. The dependent variable is the residuals of the expected excess stock return, excluding FamaFrench factors. Columns [1] and [2] show our main results for the entire sample. Columns [3] and [4] show the results for the group of firms belong to electronic computers, electrical machinery, transportation equipment, instruments, software and data processing Services, as high-tech firms consistent with the literature. In columns [5] and [6] we see these results for the sub sample of non-innovative firms. Columns [7], [8], [9] and [10] represent the same regressions as columns [1], [2], [3] and [4] plus the implementation of the instrumental variable (IV) strategy, in which PDI is instrumented using the interaction of the interest rate and the local constraints on the land supply. Consistent with the literature, we control for the interaction between housing supply elasticity and year in our instrumental regressions to capture the time trend of interest rates which explains most of the correlation between home prices. Other controls refer to leverage; firm size; asset growth; hiring rate; return on assets (ROA); return on equity (ROE); market-to-book ratio; corporate real estate holding; company age; residual housing price index; and population density. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Figure 1 Different types of areas according to their urban density.

Note: This figure shows examples of aerial views of highly developed (HD), developed (D), low developed (LD), and undevelopable (U) areas. Source: Google Earth.

Figure 2 Input and output images for the MSA of New York.



Note: The left figure shows the input image for the MSA of New York. It contains the Google Earth aerial view of the area within a maximum of one-hour drive from the center of the city (i.e., Times Square, New York). This area is equal to 6724.35 km^2 in the case of the New York MSA. The left figure shows its corresponding output image after being processed by the computer vision algorithm. The area in red represents highly developed (HD) areas, dark blue corresponds to developed (D) areas with the potential to increase their density, and light blue determine low developed (LD) areas, which is defined almost entirely by free land with some existing facilities. In this specific case of New York, the numeric output of the analysis is as follows: HD areas account for 1540.46 km^2; D areas account for 2868.85 km^2; LD areas account for 1659.48 km^2; and undevelopable (U) areas account for 655.56 km^2.