Barcelona School of Economics

WORKING PAPERS

Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation

BSE Working Paper 1467| November 2024

Leonardo D'Amico, Edward Glaeser, Joseph Gyourko, William Kerr, Giacomo A. M. Ponzetto

bse.eu/research

Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation*

Leonardo D'Amico Edward Glaeser Joseph Gyourko
William Kerr Giacomo A. M. Ponzetto

This version: November 14, 2024. First version: December 30, 2023

We document a Kuznets curve for construction productivity in 20th-century America. Homes built per construction worker remained stagnant between 1900 and 1940, boomed after World War II, and then plummeted after 1970. The productivity boom from 1940 to 1970 shows that nothing makes technological progress inherently impossible in construction. What stopped it? We present a model in which local land-use controls limit the size of building projects. This constraint reduces the equilibrium size of construction companies, reducing both scale economies and incentives to invest in innovation. Our model shows that, in a competitive industry, such inefficient reductions in firm size and technology investment are a distinctive consequence of restrictive project regulation, while classic regulatory barriers to entry increase firm size. The model is consistent with an extensive series of key facts about the nature of the construction sector. The post-1970 productivity decline coincides with increases in our best proxies for land-use regulation. The size of development projects is small today and has declined over time. The size of construction firms is also quite small, especially relative to other goods-producing firms, and smaller builders are less productive. Areas with stricter land use regulation have particularly small and unproductive construction establishments. Patenting activity in construction stagnated and diverged from other sectors. A back-of-the-envelope calculation indicates that, if half of the observed link between establishment size and productivity is causal, America's residential construction firms would be approximately 60% more productive if their size distribution matched that of manufacturing.

*D'Amico: Harvard University, damico@g.harvard.edu; Glaeser: Harvard University and NBER, eglaeser@harvard.edu; Gyourko: The Wharton School, University of Pennsylvania and NBER, gyourko@wharton.upenn.edu; Kerr: Harvard Business School and NBER, wkerr@hbs.edu; Ponzetto: CREI, UPF, BSE, IPEG and CEPR, gponzetto@crei.cat. We thank Conor Dougherty, the participants at the 2024 NBER Real Estate Summer Institute Meetings and at the 2024 Stanford Institute for Theoretical Economics Housing and Urban Economics Session, and our discussant Leah Brooks, for their insightful comments. We thank Leonardo Bianchi, Davor Djekic, Kaiser Ke, Aman Mahajan, Louis Maiden, Braydon Neiszner, Vivian Wang, and especially Pietro Dedola for outstanding research assistance. D'Amico and Glaeser thank the generous support of the Harvard Joint Center for Housing Studies. Gyourko thanks the Research Sponsor Program of the Zell/Lurie Real Estate Center at Wharton. Ponzetto thanks the Spanish Ministry of Economy and Competitiveness through its Severo Ochoa Program for Centers of Excellence in R&D (CEX2019-000915-S). Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

1 Introduction

Why are Americans so good at manufacturing traded goods and so bad at producing structures that are fixed in space? Goolsbee and Syverson (2023) document that, while U.S. aggregate productivity doubled between 1970 and 2000, it simultaneously fell by 40% for the construction sector. Foerster et al. (2022) conclude that construction's productivity problems explain more than one fourth of the 2.8 percentage point decline in America's annual GDP growth rate since World War II, which implies $1 trillion lost every five years. Sluggish construction productivity growth also has distributional consequences because it reduces housing affordability.[1] America's construction productivity problem also appears to bedevil the public sector. Brooks and Liscow (2023) show that the real cost of building a mile of highway tripled between the 1960s and the 1980s, and the Transit Costs Database, housed at NYU's Marron Institute, finds that "transit-infrastructure projects in New York cost 20 times more on a per kilometer basis than in Seoul."

In this paper, we formalize and investigate the hypothesis that regulation drives the difference between manufacturing and construction, partly by ensuring that residential construction is built by smaller and less productive firms.[2] Construction sites are both fixed in space and highly visible, while assembly lines can relocate and are typically behind closed doors.[3] Building is regulated ex ante, with local laws and state and federal environmental regulations which may restrict the ability to begin development at all. The approval process can require years of community outreach, and catering to the wishes of incumbent residents can often prove easier for smaller projects.[4] By contrast, federal regulations of manufacturing are typically enforced ex post and size can make compliance easier. If regulation keeps projects small, and small projects mean small

[1] This does not imply that the construction productivity slowdown is the primary explanation for America's growing housing affordability problem. Across space, variation in the cost of delivering a house structure explains only a small fraction of the variation in the price of a home (Gyourko and Saiz 2006). Research on this topic often focuses on the gap between prices and construction costs (Glaeser et al. 2005; Glaeser and Ward 2009a; Gyourko and Krimmel 2021). Recent reviews include Gyourko and Molloy (2015), Glaeser and Gyourko (2018), and Ellickson (2022).

[2] While our empirical focus will be on residential building, our argument applies more generally. We focus on housing not only because it is important economically and socially, but because of superior data quality.

[3] Of course, regulation effectively can change the location of new home building from Massachusetts and California to Texas and Georgia.

[4] In this paper, we will focus on the regulatory burden and its interaction with construction productivity. However, regulation can have additional perverse effects because political economy constraints also affect which types of buildings get built (Hamilton 1975; Fernandez and Rogerson 1996; Calabrese et al. 2007; Krimmel 2021). More generally, regulation can generate policies that cater to the incumbents and disregard externalities that could materially affect aggregate outcomes (Glaeser and Shleifer 2005; Glaeser and Ponzetto 2018; Hsieh and Moretti 2019; D'Amico 2022; Duranton and Puga 2023).

firms, and small firms invest less in technology (Cohen and Klepper 1996; Akcigit and Kerr 2018), then productivity growth will decline, which is exactly what America has experienced.

In Section 2, we build on Goolsbee and Syverson (2023) and document a housing-productivity Kuznets curve. By extending their series on the number of homes per construction worker back to 1900, we find that productivity growth in home-building seems to have soared from 1935 to 1970. During this period, homes produced per construction worker often grew faster than cars produced per worker in the automobile industry or total manufacturing output per industrial worker. This fact rejects the view that there is something unique about homes (or other buildings) that makes it impossible to improve productivity—a view that is also at odds with the many innovations introduced by the great post-WWII builders such as William Levitt. However, the construction sector appears to have begun to stagnate around 1970 and soon lagged behind the rest of the economy. At that same time, an increasingly prosperous America started embracing land-use controls and other regulations (Jackson 2016; Ganong and Shoag 2017).

This Kuznets curve and its timing motivates our model, which we present in Section 3. One explanation for this inverted-U shape is that the rise of ex-ante building regulation caused part of the productivity slowdown in construction. We formalize this link by studying the problem of builders who choose how much to invest in technology and have a limited ability to monitor different projects. We model regulation as a stochastic likelihood of obtaining a building permit. The probability that a project is approved is decreasing in its size, in line with the actual approval process. This project-level regulation causes average project size to shrink as entrepreneurs trade off lower project scale against a higher probability of permitting. The smaller equilibrium size of projects then leads firms to shrink because entrepreneurs cannot monitor one hundred projects with two houses each as easily as two projects with one hundred houses each. As firm size shrinks, so do incentives to invest in technology. The overall productivity of the industry consequently declines.

Our model emphasizes how project-level regulation differs from the regulation of entry because of its impact on firm size and investment. In a competitive market, higher entry costs tend to reduce the number of active firms and increase firm size. If technology can be used across different projects without decreasing returns to scope, then bigger firms invest more in technology. The adverse effects of restricted entry on consumers will thus be partially offset by the lower production costs of the bigger firms remaining in the industry. The adverse effects of project-level regulation on consumers will only be exacerbated by the small scale and limited technological investment of the firms that remain in the industry. This difference can explain why an overall regulatory expansion may lead to productivity growth in manufacturing, but productivity decline in construction.

Our theory accounts for a number of empirical regularities and offers new testable implications that we take to the data. In the cross-section, we expect smaller firms to be less productive, and jurisdictions with more regulation to have smaller and less productive firms. The model also predicts lower output overall in more regulated areas. Over time, we expect shrinking project and firm size to be associated with lower technology investment.

We show in Section 4 that construction projects and firms are, indeed, quite small. While only 2% of manufacturing employees work in firms with fewer than five employees, two-fifths of employment in new single-family housing construction is in such tiny firms. The typical land parcel bought in recent years for single-family development is also quite small. Gyourko and Krimmel (2021) report a median size of seven acres in their sample of 3,600 parcels purchased between 2013 and 2018 across 24 metropolitan areas. In this sample, 94% of parcels have less than 100 acres of land, and there are virtually no cases of a parcel containing more than 1,000 acres.

There is no extant series documenting project size over time, so we develop one using an algorithm that estimates development scale by looking at homes that were built nearby in the same two-year period. Our novel series documents that the biggest post-WWII builders in the 1960s worked with land parcels of more than 5,000 acres and built many thousands of homes on each of them, something that we can confirm from contemporaneous press reports. Since then, the share of housing built in large projects has fallen by over one third.

We test the cross-sectional implications of our model in Section 5 using the Economic Census and micro data from the Longitudinal Business Database (LBD). This Census data set contains establishment- and firm-level data, including direct measures of output such as homes built. We first document the strong connection between firm size and productivity in the construction sector. In housing construction, firms with 500 or more employees produce four times as many units per employee than firms with less than 20 employees. A back-of-the-envelope calculation implies that, if half of the link between size and productivity reflects the causal effect of size, then the construction sector would be nearly 60% more productive if its firm size distribution was the same as in manufacturing.

We then show that areas with stricter land-use regulation (as measured by the Wharton Residential Land Use Regulation Index, WRLURI) have smaller and less productive establishments in the construction industry. The relationship between size and regulation holds across all types of construction firms, but it is especially strong in the subsector involving construction of buildings. In this subsector, a one standard deviation increase in WRLURI, which is approximately the difference between Atlanta's and San Francisco's regulatory strictness, is associated with a 12% reduction in total receipts per establishment and a one-third reduction in the share of employment in large firms. Across all construction firms, a one standard deviation increase in WRLURI is associated with decreases of over 5% in revenues per employee, 3% in payroll per employee, and 5% in capital per employee. We also find much less residential and non-residential construction activity in areas with stricter land-use regulation.

In Section 6, we use data on patents and spending on research and development to test our theory's dynamic implication that shrinking firm size is associated with declining technology investment. Patenting levels for construction, manufacturing and all other industries moved together through the 1950s. After 1970, a permanent wedge appears: patents per employee soared in manufacturing but declined in construction. This finding matches the fact that corporate R&D expenditure as a fraction of revenues is at least ten times higher in manufacturing than in construction. Section 7 concludes.

2 Construction Productivity and Regulation Over Time

In this section, we highlight the facts that motivate our model.[5] We construct a series measuring homebuilding productivity that goes back to 1900. This series displays sluggish productivity growth in the first third of the 20th century that is then followed by sharply higher productivity in 1935-1970, and ends in the post-1970 decline documented by Goolsbee and Syverson (2023). We also note that the beginning of the productivity growth decline coincides with the rise of local land-use regulation. This time series, which we refer to as the construction productivity Kuznets curve, documents that there is nothing inherent about homebuilding that makes productivity growth infeasible.

2.1 House Prices and Building Production Costs

Housing is unusual among physical products because its real price has been rising for at least fifty years. Appendix Figure A1 shows that the prices of cars and houses moved together from 1950 to 1970, but diverged sharply since then. New homes now cost twice as much as they did in 1960 in real terms, but cars are 60% cheaper.

Rising prices have been accompanied by rising costs of production. For many decades, R.S. Means has provided estimates of the costs of building standardized housing units. Appendix Figure A3 uses their data and shows that the real constant quality production cost of a standard

[5]We take as given Goolsbee and Syverson's (2023) conclusions about the dismal productivity performance of the U.S. construction sector since the 1970s. One important contribution of Goolsbee and Syverson's (2023) is to show convincingly that this decline is not an artifact of measurement error, a concern that had been part of the earlier debate on this issue. Goolsbee and Syverson (2023) is the most recent and comprehensive study of low construction-sector productivity, but it is not the first. Academic economists published papers in the 1980s with titles such as "An Examination of the Productivity Decline in the Construction Industry" (Stokes 1981) and "Why Construction Industry Productivity Is Declining" (Allen 1985), both in the *Review of Economics and Statistics*. Government and academic economists have tracked the factors of production used in construction for many decades (e.g., Rothberg 1964; Ball and Ludwig 1971; Swan 1971, and the Bureau of Labor Statistics (BLS) 1972). Two of the government's more recent efforts to measure productivity growth in construction are BLS studies by Sveikauskas et al. (2014, 2018). Finally, Harvard's Joint Center for Housing Studies surveyed home builders themselves on this issue (Colton and Ahluwalia 2019). Concern about increasing construction costs many decades ago led the government to fund research into how to bring high productivity manufacturing processes into the homebuilding sector. One example is a 1969 'special summer session' that resulted in the 1971 MIT Press publication of a series of research articles from that meeting in the volume entitled *Industrialized Building Systems for Housing*. A broader look at this government effort is outlined in the blog post written by Brian Potter (2021), "Operation Breakthrough: America's Failed Government Program to Industrialize Home Production."

house structure was 18% higher in 2021 than in 1985.[6] R.S. Means costs include only the physical costs of construction, not land acquisition or other costs. These rising production costs indicate that homebuilding has been affiliated with the declining productivity that Goolsbee and Syverson's (2023) document for the construction sector as a whole.[7]

2.2 Quantities and Productivity

We now turn to construction output per employee, and we compare (i) newly-started housing units per construction-sector employee, (ii) domestically produced cars per employee in motor-vehicle production, and (iii) manufacturing output per manufacturing employee.[8] Homes per worker is an imperfect measure of productivity, since housing quality changes over time and greater use of capital may lead to more output per worker without technological improvement. Yet, a series on homes per worker is available over a long time period and it provides a particularly tangible picture of the evolution of the construction sector.

Figure 1 illustrates the evolution of new housing units per construction employee per year from 1900 to 2023. This series splices other data sources together with the Construction Census. Output per worker is obviously highly cyclical, but the figure shows a flat trend between 1900 and the mid-30s, a steep increase in output per employee between the mid-30s and early 70s, and then a sharp decline.

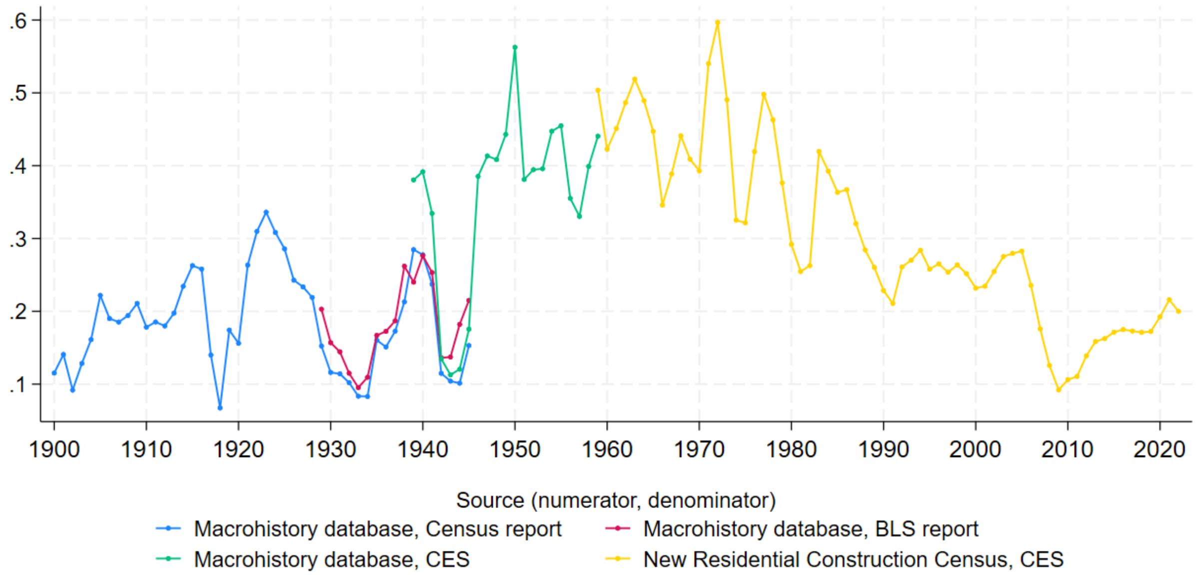
Unfortunately, this series reflects housing units per construction sector employee, not hous-

[6]The structure being priced is an 1,800 square-foot "economy quality" house. The R.S. Means Company defines "economy quality" as a simple, relatively low cost, one-story home, depicted in Appendix Figure A2, which comes from their 2021 data book. The quality of this home does change, but relatively infrequently and in relatively easy-to-measure ways given the range and detail of data provided by R.S. Means. The R.S. Means Company, owned by Gordian, publishes annual information on construction costs, but books from every year are not available. We have information on 24 years since 1985: 1985, 1986, 1989, 1991, 1996, 2002, 2003 and 2006-2021. The reference for 2021 is *Square Foot Costs with RSMeans Data*, 2021, 42nd annual edition (and likewise for earlier years).

[7]While the homebuilding sector is the focus of our empirical work, we collected data on other construction sectors that also show that physical costs in America are high by global standards. Appendix B reports results comparing construction costs in different cities in the United States with other major cities around the world. The key results are as follows. (1) Office construction costs are very high in U.S. coastal markets in particular (e.g., New York City, Los Angeles and San Francisco in our data). (2) Other interior markets such as Chicago and Houston have high office construction costs compared to most global cities in the comparison set (e.g., Paris and Singapore), but they are not expensive once high U.S. input costs are controlled for. (3) Hence, office construction costs are extremely high in U.S. coastal markets even after controlling for higher labor and materials costs. Local building regulation is strictest in the coastal markets.

[8]Cars include automobiles and light trucks.

Figure 1: Housing Units Started per Employee in the Construction Sector



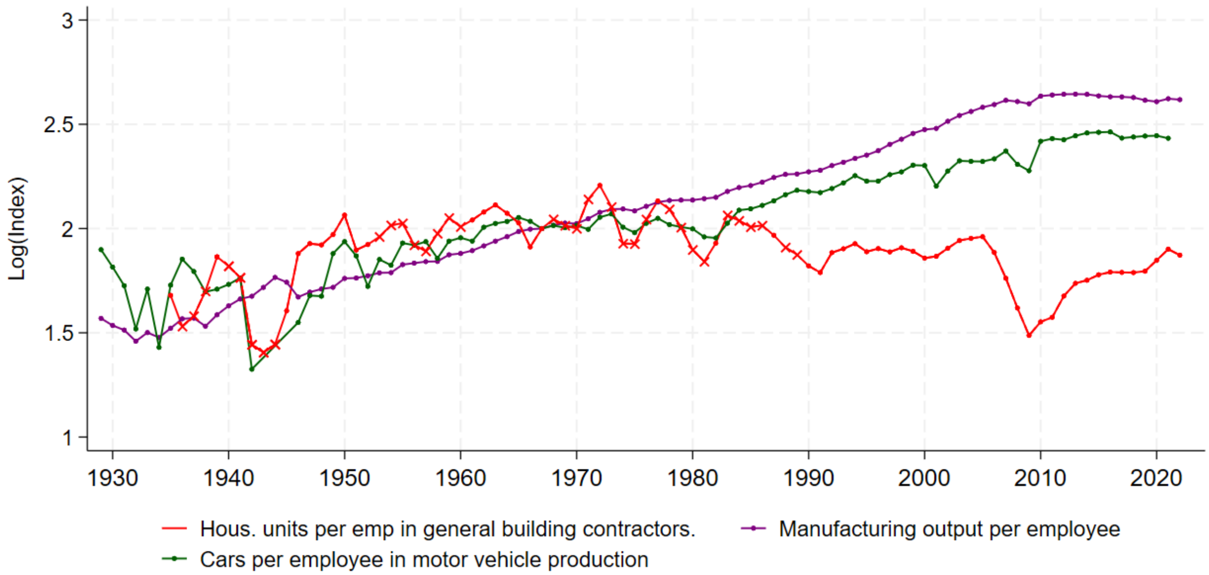
Note. The figure plots housing units started annually per employee in the construction sector, from 1900 to 2023. We take housing units between 1900 and 1959 from the Macrohistory Database, specifically from the sources denominated *US Number of New Private Nonfarm Housing Units Started, One-, Two-*, and *Three-or-more*. From 1959 onward, the data are from the Census's New Residential Construction program, specifically from the *New Privately Owned Housing Units Started* series. Employment data in the construction sector between 1900 and 1945 are from the *Historical Statistics of the United States, 1789–1945*, series D62-76. For the 1929–1945 time period, we also consulted a Bureau of Labor Statistics (BLS) historical report, which corroborates our main series. From 1939 onward, employment in the construction industry is from the BLS's Current Employment Statistics (CES).

ing units per *residential* construction sector employee. If residential construction employment changed dramatically as a share of total construction employment, this figure would misrepresent changes in residential construction productivity. As consistent data on employment in residential construction is not available, we look at the number of employees of general contractors engaged in the construction of buildings, which is available irregularly from 1935 to today.[9] This series should remove employees engaged in heavy construction, such as highways and bridges, because those are not typically built by general contractors. It also removes specialty contractors, such as electricians and plumbers, who mostly work maintaining the existing housing stock.

Figure 2 compares our series of housing units per general contractor employee with our

[9] Henceforth, we refer to general contractors engaged in the construction of buildings simply as general contractors. We have data for 1935, 1939, 1945–1952, 1960–1967, 1972, 1977, 1982, 1987, and from 1990 onward. For missing years between 1935 and 1990, we impute the number of general contractors by assuming a linear trend for the share of general contractors as a share of total employment and multiplying the predicted share by total employment in construction, which we have for all years. Appendix C.3 provides the details.

Figure 2: Housing Units, Manufacturing Output, and Cars per Employee



Note. The figure plots indices of production per employee of houses (red line, bottom in 2022), cars (green, middle), and all manufactured goods (purple, top). Series are indexed to 100 in 1967, and the y-axis uses a base-10 log scale. Cross-shaped markers are used to denote years in which the denominator of the housing series (the number of general building contractors) was estimated through an out-of-sample forecast that assumes a linear trend in the share of general building contractors as a total of all construction employees (see Appendix C.3 for details).

automobile and manufacturing output series.[10] Manufacturing productivity is measured with an index of real manufacturing production divided by total manufacturing employment.[11] We calculate cars per automobile industry employment by weaving together two different series on car production.[12]

This plot uses a base-10 logarithmic scale, and the series are indexed to a value of 100 in

[10]For this exercise, we construct a continuous housing-start series that pastes together our data sources, averaging them for the periods where they are overlapping. Appendix subsection C.1 spells out the details.

[11]In particular, manufacturing goods production is measured by the *Industrial Production and Capacity Utilization - G.17* (IPCC) index for manufacturing output, released by the Board of Governors of the Federal Reserve System, and covers all years from 1919 onward. Data on manufacturing employment come from two data sources: the *Historical Statistics of the United States, 1789–1945* Table D62-76 gives information from 1900 to 1945, and the CES reports employment data from 1939 onward.

[12]For the numerator, between 1929 and 1975 we use data on all automobiles produced in the US from the *Consumer Guide* magazine (2000; 2001a; 2001b; 2004). This data is reported in "US Automobile Production Figures" on Wikipedia. From 1975 onward we use the IPCC index for automobiles and light truck production, which is available from 1972 to 2023. The advantage of using this index is that it accounts also for light-truck production, which was negligible prior to 1975 but now has the lion's share of the market. For the denominator, we take the data on employment in motor vehicles and parts production from the Bureau of Economic Analysis (BEA) *Full-Time and Part-Time Employees by Industry* tables. When the two series on car production overlap (as occurs in 1972–1975), we paste them together, following the same procedure detailed in C.1 for homes per employee.

1967.[13] While US construction firms produce roughly as many houses per employee as they used to almost 90 years ago (e.g., 2022 vs. 1939), manufacturing output per employee grew by ten-fold over the same period, and automobile output per employee rose by 400%. In 1939, we estimate that an individual general contractor produced about 0.96 new homes a year, a similar number to 2022 (0.98). An employee engaged in motor vehicle production in 1939 contributed about 4.82 cars, which increased to around 25 cars in 2020. Houses are more complicated than they were in 1939, but the changes in automobile technology are far more dramatic, as illustrated by the fact that 12% of American homes were built before 1939 (according to the American Community Survey) while 12% of American cars were built more than 20 years ago.

These series are admittedly rough measures of productivity, and the precise size of the productivity between homebuilding and manufacturing remains debatable. There is no doubt, however, that construction has diverged from manufacturing around 1970 (Goolsbee and Syverson 2023).[14] Moreover, this productivity gap emerged at the same time as the regulation of residential land-use became more stringent.

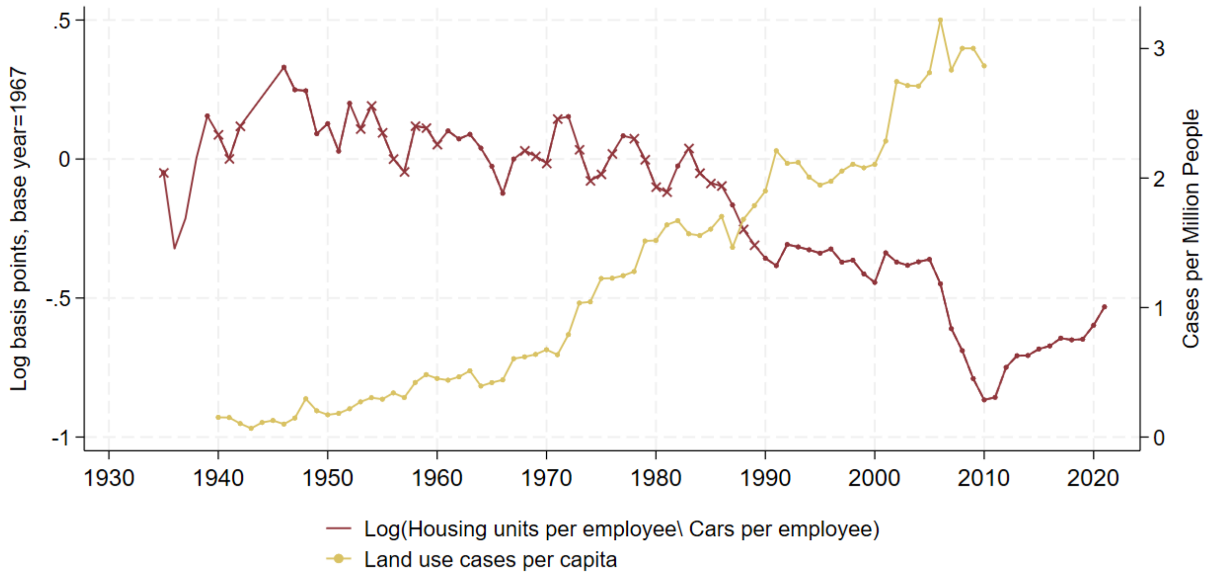
Figure 3 plots the number of land-use legal cases per capita, which is a measure of regulation developed by Ganong and Shoag (2017). The figure also shows the logarithm of the index of housing units per employee divided by the index of car production per employee. Regulation appears to have risen after 1973, roughly when construction and automotive productivity started decoupling.[15] Glaeser and Ward (2009b) document the rising number of types of land use regulations in Greater Boston since the 1970s. Katz and Rosen (1987) investigate the impact of the local growth controls that California towns began imposing during the 1970s.

[13] Thus, the value in 1967 corresponds to $\log_{10}(100) = 2$.

[14] There are potential sources of measurement error that could affect the magnitude of the gap between construction and the rest of the economy since 1970, but reasonable adjustments cannot eliminate an economically meaningful gap. For example, homes increased in size since the 1970s, with the average square footage of living area going from 1,500 square feet in 1975 to 2,500 square feet in 2014, leading us effectively to understate the growth in square footage produced per employee. However, this adjustment is not nearly enough to fully counterbalance the decrease in units produced per employee. One might also be concerned that more employees are now employed in renovations rather than new building, so that we are overstating the growth in the denominator of the ratio. Data from the Census Value of Construction Put in Place Survey shows that 25% of residential construction spending went towards renovations in 1970 and that this grew to 37% in 2024. If the growth in employment is proportional to this growth in spending, it is also not enough to undo the gap between construction and the rest of the economy. Both adjustments together would make productivity in construction productivity constant at best, but never rising at the pace of other sectors.

[15] Appendix Figure C10 shows that the same conclusion holds if we use a more direct measure of regulation, the fraction of municipalities in California with land-use regulations (Jackson 2016). We are grateful to Jacob Krimmel for sharing the data with us.

Figure 3: Productivity Decoupling and Land-Use Regulation



Note. The red line (bottom in 2010) plots the log of the ratio between the index of housing units per employee and the index of cars per employee (reported separately in Figure 2). The dark yellow line plots the number of land-use cases per capita, an index of land-use regulation from Ganong and Shoag (2017). Cross-shaped markers are used to denote years in which the denominator in the housing units per employee series was estimated through an out-of-sample forecast (see Appendix C.3 for details).

3 A Model of Regulated Construction

We now turn to a model that links project size, firm size, and industry productivity. Our theory highlights how the regulation of construction projects stunts both the size and the productivity of construction firms, while the regulation of firm entry does not. Three core assumptions generate this result: (1) developers have a limited span of control, so they cannot manage many small projects as efficiently as fewer larger projects; (2) the benefits of endogenous investment in technological know-how rise with company scale; and (3) developers' entry into each market is driven by heterogeneous preferences, so their average skill does not rise as the number of firms shrink.

Section 3.1 describes our model of construction technology and regulation. Section 3.2 outlines our standard assumptions on housing demand and spatial equilibrium. Section 3.3 provides comparative statics on productivity differences across developers within each location, holding constant local prices and aggregate quantities. Section 3.4 provides comparative statics on the regulatory variables that shape equilibrium differences across locations. Appendix D provides the closed-form solution for the model and proofs of all propositions.

3.1 Market Structure, Technology and Regulation

The economy produces two goods: housing, and a composite good that we take as the numeraire. The numeraire is produced by labor alone with constant returns to scale. In a given location c, each worker produces $w_c > 0$ units of the numeraire. Product and factor markets are perfectly competitive, so w_c is also the prevailing wage. Housing is produced by developers using labor, the numeraire good, and building parcels.

A given location c hosts D_c developers, who are identical ex ante. Each must first incur a setup cost equal to $\Phi_c w_c^{\lambda_\Phi}$, where w_c denotes the prevailing wage and $\lambda_\Phi \in [0,1]$ the labor share in developers' setup costs, while their scale $\Phi_c > 0$ captures, in part, the regulatory barriers to entry that are measured by Djankov et al. (2002).[16] We do not take a stand on whether these barriers were designed to entrench incumbents (Stigler 1971) or serve some larger social purpose, but we distinguish these upstream barriers to entry into the industry from downstream barriers to individual projects.

After incurring the setup cost, developers observe their productivity potential A_i, which is independently drawn across developers from a Pareto distribution with minimum $\underline{A}_c > 0$ and shape $\alpha_c > 1$. After observing their potential, developers choose how much to spend on technology: we denote that spending level by $K_i w_c^{\lambda_K}$, where K_i measures the real amount of investment and $\lambda_K \in [0,1]$ denotes the labor share in technology investment.[17] Developer potential and investment in technology jointly determine the physical costs of construction.

Having learned their productivity potential and invested in technology, developers can develop projects. The building process involves partnerships between risk-neutral developers and risk-neutral owners of building parcels. The developer and landowner first propose a project to the project regulator, which could be a local zoning board. The landowner commits to a payment to the developer conditional upon the project being approved and built. The regulator then either

[16] This cost function implies that setup costs Φ_c must be defrayed according to a Cobb-Douglas production function with constant returns to scale, by employing l_i workers and n_i units of the numeraire such that $(l_i/\lambda_\Phi)^{\lambda_\Phi}[n_i/(1-\lambda_\Phi)]^{1-\lambda_\Phi} = \Phi_c$.

[17] This cost function implies that a real amount K_i of technology investment can be produced according to a Cobb-Douglas production function, by employing l_i workers and n_i units of the numeraire such that $(l_i/\lambda_K)^{\lambda_F}[n_i/(1-\lambda_K)]^{1-\lambda_F} = K_i$.

approves or rejects the project. If the project is approved, then it gets built, its building units are sold, and the developer receives the contracted payment. If the project is rejected, then the parcel has no value to either the developer or the landowner.[18]

More formally, developers can partner with many landowners. A given location is endowed with T_c developable parcels that are identical for developers, but whose owners have heterogeneous costs of preliminary site preparation. Parcel j can be made available for development by incurring a cost $w_c^{\lambda_T}/\tau_j$, where $\lambda_T \in [0,1]$ is the labor share of site-preparation costs, while τ_j is a parcel-specific productivity, independently and identically distributed across parcels according to a Pareto distribution with minimum $\underline{\tau}_c > 0$ and shape $\sigma_c > 0$.[19] We assume that the minimum productivity $\underline{\tau}_c$ is sufficiently low that some inframarginal parcels are not made available for development. As we show in Appendix D, σ_c then equals the supply elasticity of parcels.[20]

We denote by r_c the site preparation cost for the marginal parcel, which in a competitive equilibrium also equals the expected gross return to any parcel owner who participates in partnerships with developers.[21] At the time of contracting, the developer i and landowner j agree upon both the payment conditional on project approval and the size of the proposed project, $b_{i,j}$.[22]

The local regulator then decides whether to approve the project. A project of size b is approved by the regulator of location c with probability:

$$a_c(b) = \min\left\{\left(\frac{\underline{b}_c}{b}\right)^{\rho_c}, 1\right\}. \tag{1}$$

While the scale of fixed costs Φ_c measures the regulation of entry, the tightness of project regulation is captured by $\rho_c \in (0,1)$. This measure is an index because $\rho_c = 0$ implies that all projects are approved and $\rho_c = 1$ implies that it is impossible for the expected number of units in a project to be greater than $\underline{b}_c$. We also assume that the amount of construction $\underline{b}_c > 0$ that would be permitted

[18]This structure reflects the reality that the large majority of developers in the U.S. are contractors building on land they do not own. It is equivalent to an alternative structure in which merchant builders first buy parcels and then propose to develop them, with the parcels becoming worthless to the developer if their project is rejected.

[19]This cost function implies that one parcel can be made available for development through a Cobb-Douglas production function with constant returns to scale, by employing l_j workers and n_j units of the numeraire such that $\tau_j (l_j/\lambda_T)^{\lambda_T} [n_j/(1-\lambda_T)]^{1-\lambda_T} = 1$.

[20]Formally, a corner solution is avoided as long as $\underline{\tau}_c < w_c^{\lambda_T}/r_c$. Otherwise the supply of parcels is perfectly inelastic at T_c.

[21]Likewise, it would equal the market price of parcels if land were bought by merchant builders.

[22]These must be simultaneously agreed upon for the partnership to function smoothly, because the interests of the developer and the parcel owner diverge once the post-development payment to the parcel owner is fixed.

with certainty is negligible, in the sense that it can never be value-maximizing to propose a project that is certain to be permitted.[23]

If the project is approved, $b_{i,j}$ units are developed on the parcel and then sold in a competitive housing market at the equilibrium price p_c. The variable cost of construction is

$$m_{i,j} = \frac{\varepsilon}{z_{i,j}} b_{i,j}^{1+1/\varepsilon} w_c^{\lambda_M}, \tag{2}$$

where $\lambda_M \in [0,1]$ denotes the labor share in the efficient input bundle, while the parameter $\varepsilon > 0$ is an inverse measure of the extent to which costs increase with the scale of each project.[24]

The productivity $z_{i,j}$ of a given project depends on the builder's span of control, namely the number s_i of projects they are supervising. If project (i,j) has rank $s_{i,j}$ in developer i's portfolio, it has productivity:

$$z_{i,j} = A_i K_i^{1/\kappa} s_{i,j}^{-1/\omega}. \tag{3}$$

Overall productivity combines the developer's idiosyncratic productivity potential (A_i) with the developer's investment in technology (K_i), whose returns are governed by the parameter κ. A core assumption in our model is that it is difficult to supervise a large number of projects. The developer's time is limited. If they are constantly shuttling between small projects it is harder for them to keep watch over the costs in any one project. We model the limited span of control by assuming that costs are higher for the second project than the first and for the tenth project than for the fifth, or formally that costs are multiplied by $s_{i,j}^{1/\omega}$ for the $s_{i,j}$-th project.

We assume that $\alpha_c \kappa / (\alpha_c + \kappa) > \omega > \varepsilon (1 - \rho_c)$, which guarantees that a finite number of projects yields a finite amount of building, and that average contractor size is finite. Since the first inequality also implies that $\kappa > \omega$, this assumption also ensures that the returns to technology adoption rise more slowly than the cost of technology adoption, and thus that the technology-choice problem has a unique, interior solution.

[23] Formally, a corner solution is avoided as long as $\underline{b}_c \leq [1 + \varepsilon (1 - \rho_c)] r_c / p_c$.

[24] This cost function implies that $b_{i,j}$ units (if permitted) can be built according to a Cobb-Douglas production function with decreasing returns to scale, by employing $l_{i,j}$ workers and $n_{i,j}$ units of the numeraire such that $\left\{ z_{i,j} \left(l_{i,j} / \lambda_M \right)^{\lambda_M} \left[n_{i,j} / (1 - \lambda_M) \right]^{1-\lambda_M} / \varepsilon \right\}^{\varepsilon/(1+\varepsilon)} = b_{i,j}$.

3.2 Housing Demand and Spatial Equilibrium

The economy consists of a continuum C of locations. A given location c has an exogenous endowment of land parcels T_c. For simplicity, we assume that landowners' earnings are spent entirely on the numeraire and do not contribute to housing demand in the location.[25] This simplifying assumption of absentee landlords ensures that increases in local land values do not translate into increases in local housing demand.[26]

The location also hosts an endogenous population of L_c workers earning w_c each and D_c developers whose average net earnings equal $\bar{\Pi}_c - \Phi_c w_c^{\lambda_\Phi}$.[27] Both workers and developers have the same Cobb-Douglas utility function with budget shares δ for housing and $1-\delta$ for the numeraire. Aggregate housing expenditure in the location is then:

$$H_c = \delta \left[w_c L_c + \left(\bar{\Pi}_c - \Phi_c w_c^{\lambda_\Phi} \right) D_c \right]. \tag{4}$$

Both workers and developers sort in spatial equilibrium according to the continuous-case generalization of the logit model of location choice (Ben-Akiva et al. 1985). Each agent i chooses where to live from a random set of $M_i \in \mathbb{N}$ opportunities, whose locations are uniformly distributed over C. The agent's utility in location c equals $U_{i,c} = y_{i,c} p_c{}^{-\delta} v_{i,c}$, where $y_{i,c}$ denotes expected income, p_c the local price of housing, and $v_{i,c}$ an idiosyncratic taste for local amenities. Such tastes are independent across individuals and locations, and they are distributed according to a Fréchet distribution: $\Pr\left(v_{i,c} \le u\right) = \exp\left(-\Upsilon_c u^{-\mu}\right)$, where $\Upsilon_c > 0$ parametrizes the common appeal of location c, while $\mu > 0$ governs the similarity of tastes across agents.

As a result, given an economy-wide aggregate endowment of workers L with location-specific incomes $y_{ic} = w_c$, the equilibrium density of workers in each location c equals

$$L_c = \frac{L \Upsilon_c \left(w_c p_c^{-\delta} \right)^\mu}{\int_C \Upsilon_x \left(w_x p_x^{-\delta} \right)^\mu dx}. \tag{5}$$

[25] Almost all our results below would remain unchanged if we assumed instead that an arbitrary share of land earnings is spent on housing in the location where the corresponding land parcel is located. The only exception is that the impact of project regulation on the aggregate value of housing built would become ambiguous.

[26] More precisely, we can derive the same model—up to an immaterial scaling constant—by assuming that all land parcels are owned by profit-maximizing public companies, whose aggregate earnings are rebated to workers and developers in proportion to their incomes.

[27] Appendix D solves for developers' average operating profits $\bar{\Pi}_c > \Phi_c w_c^{\lambda_\Phi}$.

Identically, given an economy-wide aggregate endowment of developers D with location-specific expected incomes $y_{ic} = \bar{\Pi}_c - \Phi_c$, the equilibrium density of developers in each location c equals

$$D_c = \frac{D\Upsilon_c \left[\left(\bar{\Pi}_c - \Phi_c w_c^{\lambda_\Phi} \right) p_c^{-\delta} \right]^\mu}{\int_C \Upsilon_x \left[\left(\bar{\Pi}_x - \Phi_x w_x^{\lambda_\Phi} \right) p_x^{-\delta} \right]^\mu dx}. \tag{6}$$

The taste similarity parameter μ coincides with the elasticity of migration to differences in real income across space.

3.3 Developer Heterogeneity

In equilibrium, all developers operating in the same location (c) choose to undertake projects up to the point where the productivity of the marginal project hits a threshold $\underline{z}_c$ that is homogeneous across builders. Conditional upon undertaking a project, its optimal proposed size is also determined and increasing in project productivity ($z_{i,j}$).

Developers with greater potential (A_i) and technology investment (K_i) can naturally handle a greater number of projects before stretching their span of control to the point at which their marginal project productivity hits the threshold $\underline{z}_c$. Developers also choose optimally their technology investment. Developers with greater intrinsic potential can grow to a larger scale and thus reap greater benefits from their investment. As a consequence, they are incentivized to invest more in technology.

Proposition 1 formally describes the impact of idiosyncratic productivity on developer behavior.

Proposition 1 *Across builders in the same location, the elasticity of technology investment, projects undertaken, units built, and revenues with respect to productivity potential (A_i) is identical and equal to $1/\left(1/\omega - 1/\kappa\right) > 0$. Builders with higher productivity potential also have higher revenues per employee.*

The exogenous driver of differences across developers in a given location is the realization of their idiosyncratic potential (A_i). Developers with greater potential optimally choose greater technology investment and predictably end up being both bigger and more productive. Our measure of their productivity is the ratio of revenues to employment.

Our model features a common equilibrium distribution of the value generated by developers, independent of their idiosyncratic potential. The value of construction (X_i) is split into landowner revenues ($r_c S_i$) and developer revenues (R_i) according to

$$\frac{r_c S_i}{X_i} = 1 - \frac{\varepsilon\,(1-\rho_c)}{1+\varepsilon\,(1-\rho_c)}\frac{1+\omega}{\omega} \text{ and } \frac{R_i}{X_i} = \frac{\varepsilon\,(1-\rho_c)}{1+\varepsilon\,(1-\rho_c)}\frac{1+\omega}{\omega} \tag{7}$$

for every developer i in a given location c. In turn, developer revenues are split into the cost of construction inputs (M_i), the cost of technology investment ($K_i w_c^{\lambda_K}$), and the developer's operating profits according to

$$\frac{M_i}{R_i} = \frac{\omega}{1+\omega},\ \frac{K_i w_c^{\lambda_K}}{R_i} = \frac{\omega}{\kappa}\frac{1}{1+\omega} \text{ and } \frac{\Pi_i}{R_i} = \left(1-\frac{\omega}{\kappa}\right)\frac{1}{1+\omega} \tag{8}$$

for every developer i, irrespective of location c.

As a result, the amount of housing built (X_i/p_c), projects undertaken (S_i), revenues (R_i) and technology investment (K_i) are all equiproportional in equilibrium. They are increasing in productivity potential (A_i) with a common elasticity, which is higher when the returns to technology investment are less rapidly decreasing (lower κ) and the costs of a greater span of control less rapidly increasing (higher ω).

The total cost, payroll and employment associated with technology investment and construction inputs also scale with productivity potential according to the same elasticity. However, setup costs (Φ_i) and their associated payroll and employment are fixed. Developers who draw a higher realization of potential after entering the market will lead to bigger and more productive firms that earn greater revenues for the same setup cost. Thus, they have greater revenues per employee so long as the labor requirement of firm setup is positive ($\lambda_\Phi > 0$).

The link between size and productivity reflects two-way causality. Innately more productive firms take on more projects and build more in each project, but firms that anticipate taking on more projects also invest more in technology. When we simulate the impact of a shift in the firm size distribution on productivity in the construction sector, we will have to make an assumption informed by this model about how much of the observed empirical link between size and productivity reflects the causal effect of productivity on size.

3.4 The Impact of Project Regulation

We now turn to the market-level impact of regulation on firm size and productivity. Propositions 2 and 3 look at the impact of project-level regulation, which is captured by the permitting parameter ρ_c. Proposition 4 looks at the regulation of entry, which is captured by the fixed-cost parameter Φ_c. In these propositions, parameter changes cause equilibrium prices to change, unlike in Proposition 1, which only looked across firms within a given market equilibrium.

Proposition 2 *Tighter project regulation (higher ρ) increases the equilibrium price of buildings, and reduces the equilibrium number of developers and the number and aggregate value of buildings built.*

More restrictive project regulation increases the cost of proposing large projects and reduces profitability in the construction industry, holding prices constant. This decline in profitability reduces the entry of developers into the market. The price of housing rises with reduced entry, moderating but never fully undoing the decline in the number of developers.

The reduction in the equilibrium quantity of housing built reflects a two-fold decline in demand. First, each resident wants less housing as its price rises. Second, some residents leave the city as it becomes less affordable—and also less profitable for developers who live there. With Cobb-Douglas preferences, the overall decline in demand is sufficiently large that the aggregate value of homes built declines, despite the rise in their unit price.

However, the impact of tighter project regulation on land values is ambiguous. Fewer homes get built on the average developed parcel, but each of those homes is more valuable. Which force prevails depends on the extent to which residents leave an increasingly unaffordable (and unprofitable) city.

If the migration elasticity is sufficiently high, spatial equilibrium requires both house prices and developer profits to fall very modestly. The efficiency loss is absorbed through a decline in the value of the immobile factor: land. If instead the migration elasticity is sufficiently low, relatively immobile residents absorb the efficiency loss through a decline in welfare. Workers' expenditure on housing barely declines, causing land to appreciate as construction becomes more land-intensive.

The same forces driving land values shape urban footprint. Tighter project regulation causes fewer houses to be built, but forces them to be built at lower density. If the migration response is high enough, not only do land values decline, but population falls so sharply that both density and the spread of the urban area decline. If instead the migration response is low enough, land values increase and the city sprawls further to accommodate an almost unchanged population at a lower density.

The popularity of zoning with local property owners suggests that in reality mobility is limited and regulatory constraints enrich landowners. In that case, land use regulation provides a counter-example to the Henry George Theorem that land value maximization yields welfare maximization (Henderson 1974; Arnott and Stiglitz 1979). That argument relies on perfect mobility across locations: with imperfect mobility, landowners can instead engage in inefficient rent extraction at the expense of imperfectly mobile residents (Wildasin and Wilson 1996). Project regulation can transfer rents from imperfectly mobile developers and workers to landowners. It is also then a driver of suburban sprawl.

Proposition 3 turns to the link between project regulation and the size and productivity of developers.

Proposition 3 *Tighter project regulation (higher ρ) reduces developers' average technology investment, average revenues, and average revenues per employee. It increases the value of land parcels developed by the average developer.*

More restrictive regulation hinders developers' by forcing them to undertake inefficiently small projects. Deprived of the ability to operate at scale, developers cannot reap the full benefits of their technology investment, and they react by investing less. In Proposition 1, lower ability meant less investment. In Proposition 3, tighter regulation reduces investment. In both cases, reduced technology investment means lower productivity. Consequently, average revenues fall, both per developer and per employee. We will test these implications in the next section.

The last result in Proposition 3 underscores the importance of measuring developer size properly. Contractors build on parcels owned by the investors or home buyers who hire them.

Land value is not included in their revenues nor in their costs. Tighter regulation reduces average contractor revenues, but it increases the value of projects undertaken by the average contractor.

3.5 Regulation of Entry vs. Project Regulation

Land-use regulation is different from the classic forms of regulation discussed in Stigler (1971) or Djankov et al. (2002). Typically, regulations are imposed at the firm level, and are presumably best seen as a fixed cost that the firm must incur. Here we contrast the impact of project-level regulation—discussed above—which leads to overly small projects and too little investment in firm-wide technology, with the impact of a firm-level regulation that acts as a barrier to entry into the industry.

Proposition 4 *Tighter entry regulation (higher Φ) reduces the equilibrium number of developers, the quantity and aggregate value of buildings built, the equilibrium price of land parcels, and developers' average revenues per employee.*

There are two finite thresholds $\check{\Phi}_p > \check{\Phi}_{\bar{\Pi}}$, both of which may be nil. If and only if $\Phi \geq \check{\Phi}_p$ an increase in entry costs increases house prices. If and only if $\Phi \geq \check{\Phi}_{\bar{\Pi}}$ an increase in entry costs increases developers' average technology investment, average revenues, and the value (and a fortiori the number) of land parcels developed by the average developer.

Higher fixed costs require each developer to employ more non-production employees—for instance, to deal with regulators. In equilibrium, this cost is partially compensated by the exit of marginal developers, which tends to increase revenues per developer. Intuitively, however, the direct effect always dominates, and average productivity—as measured by revenues per employee—declines. The decline in developer efficiency then reduces the quantity and value of housing built. Land prices also fall.

Instead, the effect on house prices is not entirely unambiguous, because developers are not only suppliers of housing. They are also local residents who demand housing themselves. Inframarginal developers earn net profits in equilibrium, and an increase in fixed costs acts partly as a tax on those profits. As their wealth declines, so does their demand for housing. In practice, we expect entry costs to play a more decisive role through their impact on developers' entry than on their housing demand. Intuitively, this is the case when fixed costs are sufficiently large, when migration is sufficiently elastic, and when the budget share of housing is low enough.[28]

In this realistic case, higher fixed costs harm consumers through an increase in house prices. The remaining developers, although less profitable, certainly grow larger: they have higher revenues, and each develops so many more land parcels that their aggregate value rises despite falling land prices.

Regulatory barriers to entry do restrict supply, which intuitively harms consumers through higher prices. However, there is at least a countervailing force: larger incumbents have an incentive to invest more in technology. With project-level regulation that reduces firm size, the added effect from technology is negative too, as smaller firms end up being less efficient.

In the next sections, we turn to empirical tests of the implications and predictions of our model.

4 The Small Size of Construction Firms and Projects

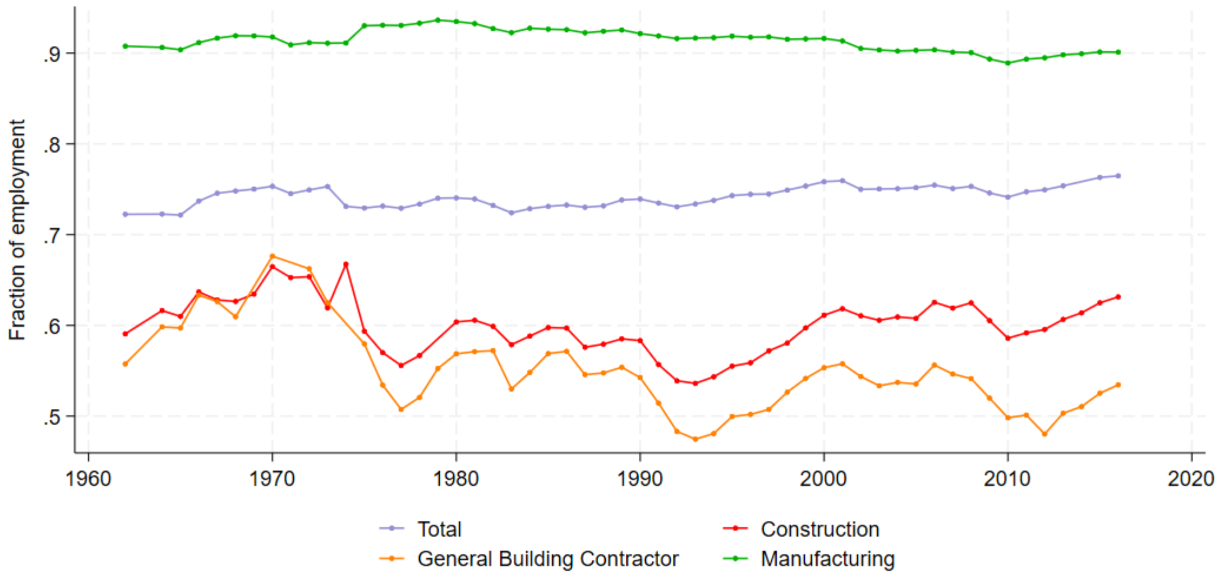
Our model predicts that increasing regulatory tightness should lead to declining project size. We will test whether project size is both small in magnitude and has declined over time. Theoretically, increasing regulation also induces smaller firms. Empirically, firms in the construction sector should then also be small both in absolute terms and relative to other sectors.

4.1 Firm Size Over Time

Figure 4 reports the fraction of employees working in large establishments, which we define as having more than 20 employees, in the U.S. between 1962 and 2016. We construct this measure using data from the Census' County Business Patterns series, which reports employment in each sector broken down across different establishment sizes. Both in construction as a whole (red line) and among general building contractors (orange line), employment in large establishments was growing until 1972. From 1973 onward, this share declined markedly for general building con-

[28]Each of the three parameters suffices alone: the threshold value $\check{\Phi}_p$ reaches zero for large but finite values of μ, or small but positive values of δ.

Figure 4: Share of Employment in Establishments with More than 20 Employees, by Sector



Note. The figure shows the fraction of employees working in establishments with more than 20 employees in the U.S. between 1962 and 2016. Each line represents a different sector, and the purple line represents the total for the U.S. economy.

tractors, which in 2016 report only 53% of employment in large establishments. For construction as a whole, the share also declined until the 1990s, but then started converging to its pre-1973 levels. This growth was mostly led by heavy and civil engineering firms, not by those engaged in homebuilding. The decline in the size of construction firms also began in the 1970s, which was the same time that construction productivity began to decline, as reported in Figures 2 and 3. That is, size and productivity both started shrinking in the mid-70s, the years when the sector started becoming more heavily regulated.

There is no comparable decline in manufacturing or the economy as a whole. The employment share of large establishments has steadily hovered around 91% in manufacturing and 74% for the overall economy.

4.2 Firm Size Distributions in 2012

We use two datasets to measure the distribution of firm sizes. First, we take public data from the national aggregates of the 2012 Statistics of U.S. Businesses (SUSB). This allows us to plot firm size distributions across the publicly-available construction sub-sectors.[29] Second, we use the confidential base responses of the 2012 Census of Construction Industries (CCI) to analyze firm size distributions across fine-grained categories of construction work.

4.2.1 Evidence from Aggregate Data

Figure 5 reports the shares of employment (top panel) and total receipts (bottom panel) across different employment size classes and industrial sectors, including new single-family housing construction, manufacturing, services, tradables, and nontradables (excluding construction).[30] The value of approximately 0.4 for new single-family housing in the 0-4 bracket in the top panel of Figure 5a indicates that approximately 40% of all employees who build single-family homes work in firms with four or fewer employees. The value of approximately 0.8 for manufacturing in the 500+ category indicates that approximately 80% of total revenues in manufacturing were earned by the largest category of manufacturing establishment.

The average single-family residential construction firm is much smaller than the average firm in other industries. More than 63% of employees in New Single-Family Housing Construction work in establishments of firms with less than 10 employees. Fewer than 13% of workers in manufacturing, tradables, services, and non-tradables work in such small firms. Establishments of firms with more than 100 employees are responsible for most of the employment in other industries, but they are rare in the single-family residential construction sector.

Our results on receipts show the same pattern. Approximately 60% of the revenues in the new single-family housing construction subsector accrue to firms with less than 10 employees. Less than 13% of such revenues are generated by firms with more than 100 employees. In manufacturing and tradables, about 80% of revenues are generated by firms with more than 500 employees. More than one half of revenues for services and other non-tradables also went to the largest firms.[31]

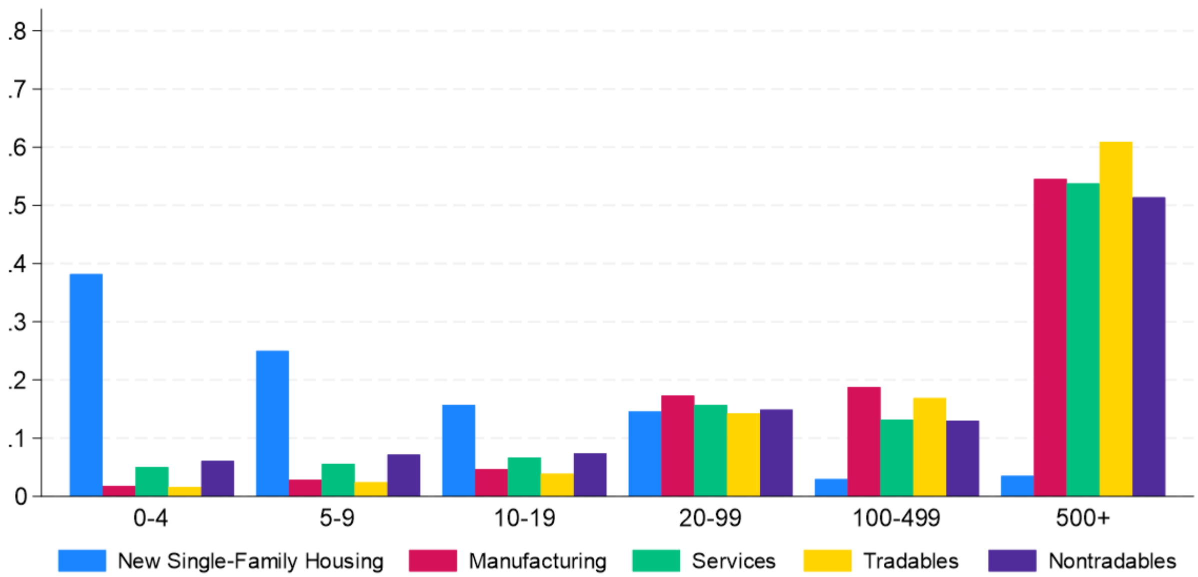
[29] The SUSB is an annual series of the U.S. Census Bureau which supplies both national and subnational data on sector-specific distributions across establishment sizes for many indicators, including employment.

[30] Our list of tradables is: “Agriculture, Forestry and Fishing”, “Mining”, “Manufacturing”, and “Management of Companies and Enterprises”. Our list of nontradables includes: “Retail Trade”, “Real Estate and Rent Leasing”, “Health Care”, “Accomodation and Food Services”, and “Other Services.”

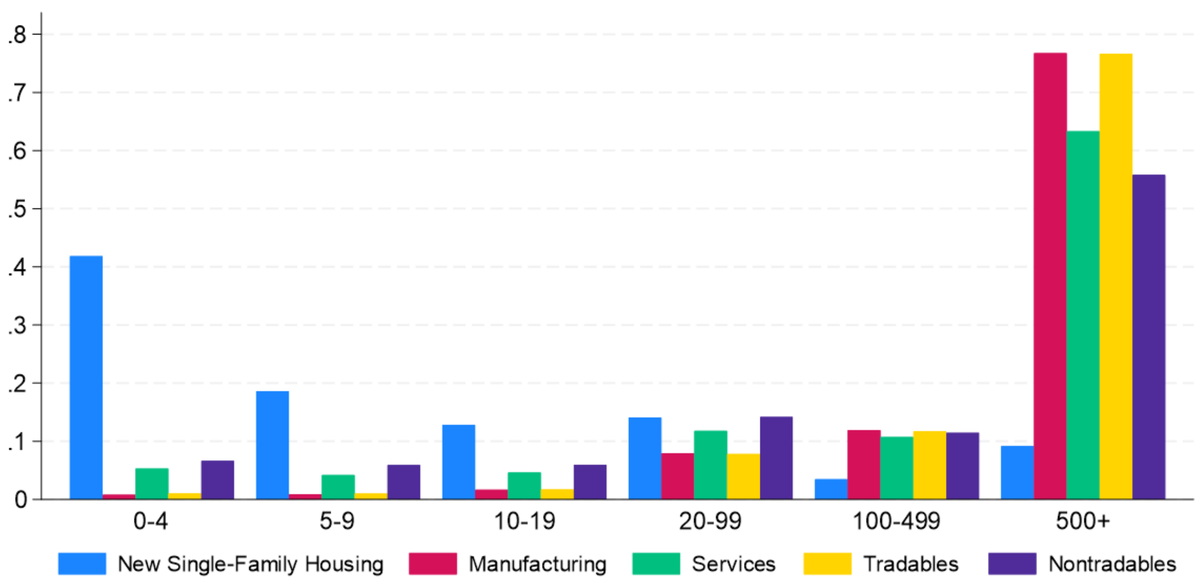
[31] Appendix E.1 shows that these patterns hold also for the number of establishments, number of firms and annual payrolls. Figure E11 in Appendix E.1 shows that multifamily builders are also small. In that sector, the bulk of employment and receipts are in firms between 20 to 99 employees, which is larger than for single-family builders but still considerably smaller than in the other industries.

Figure 5: Firm Size Distributions

(a) Employment Shares



(b) Receipt Shares



Note. The figure plots the share of total employment (panel a) and receipts (panel b) accounted for by firms in different size bins, across different sectors. Data from the 2012 SUSB.

4.2.2 Micro Evidence from the Census of Construction Industries

4.2.2.1 Data Description

The 2012 CCI dataset contains establishment-level operating data from more than 100,000 firms. The vast majority of firms in the construction industry have only a single establishment. We restrict the sample to observations that were used in official CCI tabulations, that possess sample weights, and that link to a firm in the Longitudinal Business Database. We require that establishments in the sample have non-zero employment as reported on the CCI form, non-missing breakouts of revenue by sector (described further below), non-missing values for all operating data used in profit/productivity calculations, and be located in a CBSA with a 2006 WRLURI value. These restrictions produce a sample of approximately 107,000 firms and their establishments, which we use throughout our later analyses.[32] All our results are very similar if these sample choices are relaxed.[33] A key feature of the CCI data is that the establishment-level report of construction revenues is split between 31 types of activities (e.g., "single-family homes, detached," "bridges and elevated highways," "decks, residential types"). This detail allows us to analyze narrowly defined construction sectors. We group the reported variables into the following seven categories: housing, consumer-facing buildings (e.g., restaurants, retail stores), office buildings, warehouses, industrial/manufacturing buildings, other buildings (e.g., dormitories, schools, hospitals), and non-building construction (e.g., bridges, highways, sewage plants).[34]

We calculate the share of revenues across these seven bins for each establishment and firm. The CCI asks firms to report their revenues for activities that exclude the cost of land and other items installed that are not part of the building structure. In graphical analyses, we collapse warehouses and consumer-facing building, and assign firms to a single specialization bin based upon the majority source of revenues. We also introduce an eighth category for firms that do not earn most of their revenues from a single sector. These specialization bins are thus mutually exclusive and collectively exhaustive.[35] These choices mostly follow from disclosure requirements, and we

[32]Here we will present graphical firm size distributions. The next sections present regressions of firm profitability by employment size and sector of operation, and regressions of construction traits at the CBSA level by regulation levels.

[33]Observations counts through this paper are rounded per Census Bureau disclosure requirements.

[34]See pages 8-10 of the CCI form. The housing category is codes 316 to 318. Consumer-facing building is codes 324, 326. Office buildings is code 325. Warehouses is code 327. Industrial/manufacturing buildings is codes 321 and 323. Other buildings is codes 319, 328 to 334, and 338. Non-building construction is all codes within category B in the form.

[35]In regression analyses at the firm level, we will model these revenue shares as continuous variables.

note below robustness to other approaches.

4.2.2.2 Results

Figure 6 presents firm size distributions (FSD) across construction sectors. Different types of construction are in different shades of blue to purple.[36] In green, we show the firm size distributions for non-construction firms based on the 2012 LBD data. The latter sample contains every LBD firm that is not in the CCI sample and has a modal establishment that is not in the construction sector (NAICS 23).

The horizontal axis of each graph groups firms by employment levels, using the same increments used above: 0-4 employees, 5-9, 10-19, 20-99, 100-499, and 500+. The vertical axis shows the share of sector employment and revenues that are accounted for by each employment level (as above, the distribution sums to 1).[37]

Differences between construction firms and non-construction firms remain stark. Around 60% or more of the employment, revenue, and payroll of non-construction firms is accounted for by firms with 500+ employees, even though they are a small share of the number of total firms. This pattern confirms that "the typical firm is small, but the typical employee works in a large firm." The 500+ employee bin for non-construction firms in the LBD often contains 13 to 18 times the activity accounted for by the 0-4 employee bin.

The construction sector is markedly different. Firms specialized in housing construction are the most extreme, with firms of 0-4 employees accounting for the largest share of employees and revenues.[38] In most other building-construction sectors, the activity in the smallest size bins is comparable to the activity in the 500+ employee bin. For example, firms with 0-4 employees account for slightly more employment in office building construction than firms with 500+ employees.

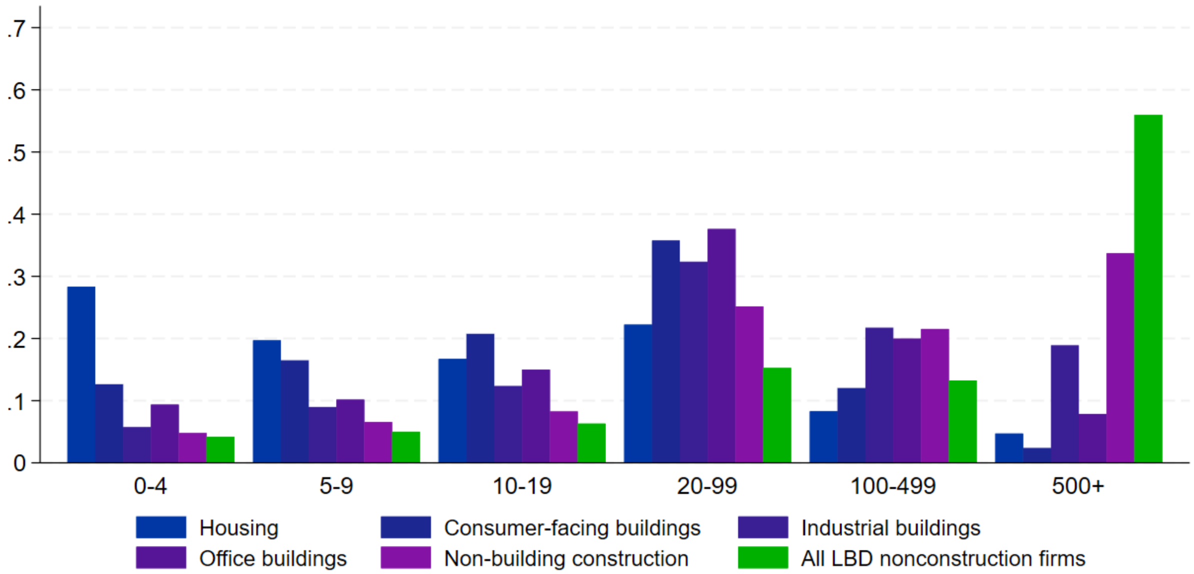
[36] Here we focus on housing, consumer-facing buildings, industrial buildings and warehouses, office buildings, and non-building construction. Appendix Figure E15 reports FSD results that also include the other two subsectors: other buildings construction (dormitories, schools, etc.), and firms with no clear specialization (i.e. that do not have more than 50% of their revenues in one type of construction).

[37] Appendix Figure E14 reports payroll and firm counts. All graphs use sample weights. In the case of multi-establishment firms, we take the minimum weight across the establishments of the firm. Results are very similar under alternative techniques.

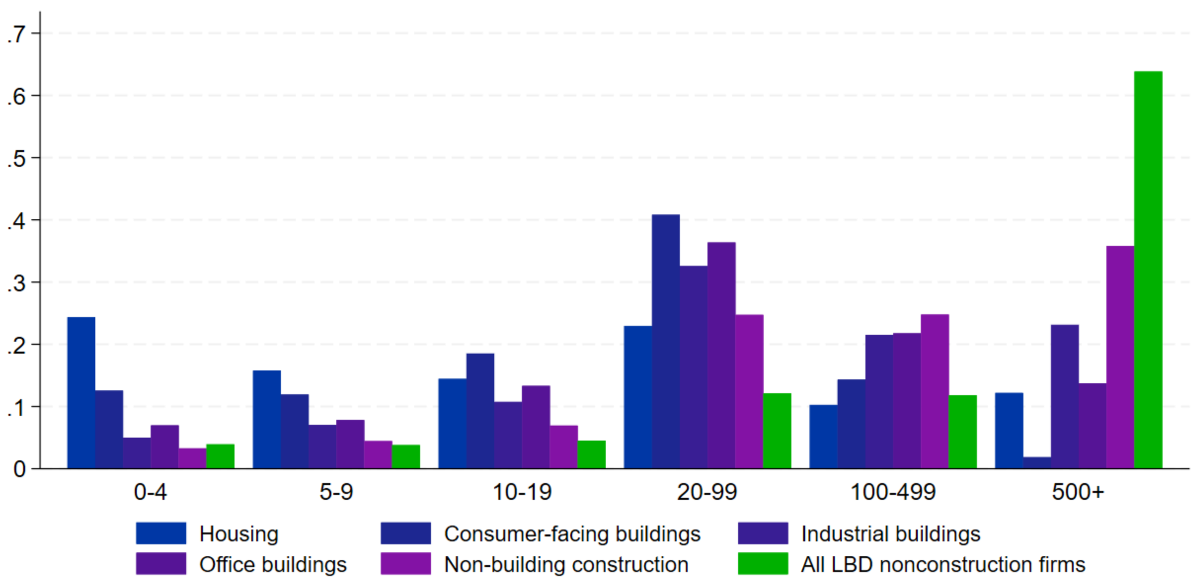
[38] Results on payroll are reported in Appendix Figure E14. There, the 20-99 bin accounts for the highest share (22%), but the 0-4 bin is the second-highest (20%).

Figure 6: Firm Size Distribution

(a) Employment Shares

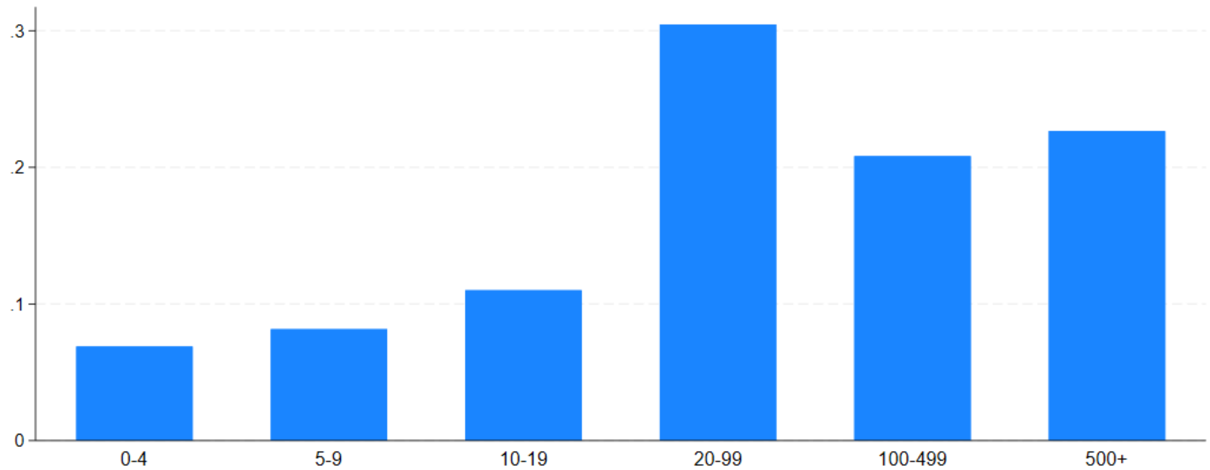


(b) Revenue Shares



Note. The figure plots the share of total employment (panel a) and revenues (panel b) accounted for by firms in different size bins, across different types of construction firms (in shades of blue to purple) and compared also to all non-construction firms (in green). Data for 2012 from the LBD. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

Figure 7: Share of Housing Units Built by Employment Size



Note. The figure plots the share of total housing built accounted for by firms in different size bins. Microdata from the 2012 Census of Construction Industries (CCI). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

The CCI survey asks firms to report units built (single- and multi-unit residential). Most firms specializing in the housing sector based on their revenues are reported as having produced zero housing units. This fact presumably means that the firm builds components of houses and not entire houses (although in some cases, firms may have chosen not to report the data). Figure 7 plots the distribution of units created across the firm size distribution.

4.3 The Size of Residential Construction Projects

We conclude this section by testing whether there is a link between small firms and small projects. We start by describing what is probably the most famous example of large-scale construction projects in the U.S., the case of the post-WWII mega-builder Levitt & Sons. We then compare the size of what the Levitts were building with the cross-section of construction projects today, using data on land-parcel purchases. Finally, to describe the evolution of project size in the U.S. since the 1950s, we construct a new data series on single-family homebuilding project size by identifying the size of construction projects using microdata from Corelogic.

4.3.1 Levitt & Sons[39]

In 1947 Levitt & Sons acquired 1,400 acres of Long Island farmland with the idea of efficiently developing thousands of nearly identical single-family houses. By 1948 the firm was completing more than 35 houses per day or 175 per week. Ultimately, just over 17,000 houses were built and sold for an average price of $7,990 ($100,886 in 2023 dollars). By 1950, Levitt was a household name, who had appeared in major news magazines (e.g., Larrabee 1948 in *Harper's Magazine*; and *Time* magazine, July 3, 1950) in the 1950s they repeated this style of development in Bucks County, PA, building 17,300 homes on 6,000 acres.

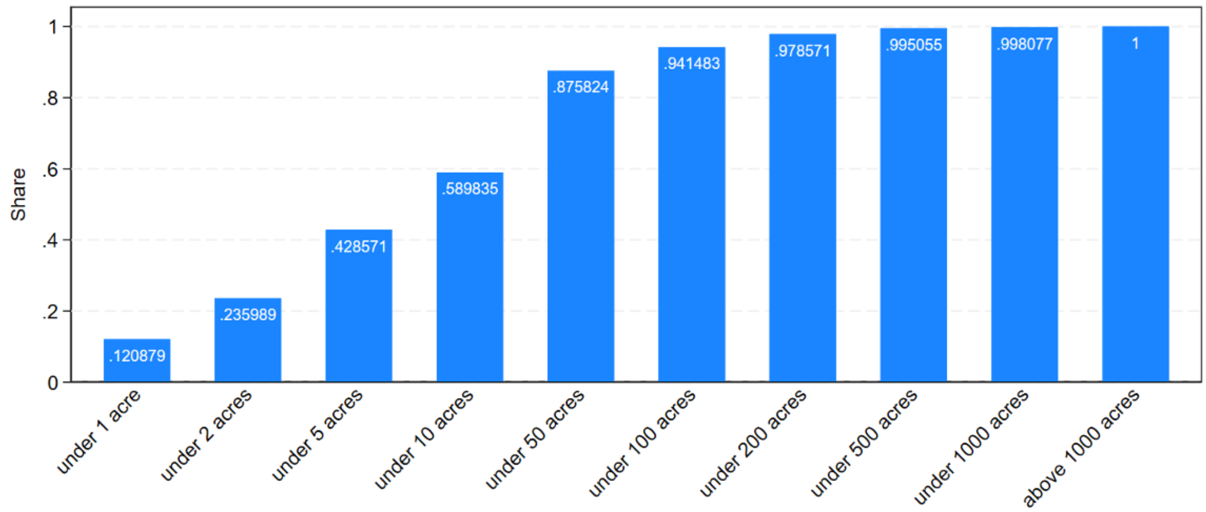
The Levitts produced so much housing at relatively low cost by inventing an assembly line for housing, which moved construction crews along nearly identical homes—thousands of them. They broke down the construction process into 26 specific components and had a team for each of them. They used time and motion study techniques, and brought new processes to homebuilding. They also tried to preassemble as much as possible off-site. Their productivity enabled them to earn profits of about $1,000 per home (roughly $13,800 today; Rybczynski 2017).[40]

Are there any Levitt & Sons today? To describe how projects look today, we first turn to data from Gyourko and Krimmel (2021), who collected information on land parcels purchased for the expressed purpose of single-family development across 24 metro areas over the 2013–2018 period. In their sample, the largest U.S. single-family residential land parcel purchased was a 1,049 acre site north of Denver: one-sixth of what the Levitts worked with in Pennsylvania. Figure 8 reports the cumulative distribution across parcel size. The median project is below 10 acres, and

[39] Much of the information reported in this subsection is taken from Rybczynski (2017).

[40] While the Levitt brothers are the most famous of the early post-WWII homebuilders, they were not unique by any means; rather, they were trend setters. Checkoway (1983) notes other large builders who used Levitt-type production strategies to rapidly construct hundreds or thousands of new homes in the 1950s and 1960s. They existed in a wide range of markets across the country including Baltimore (John Mowbray), Washington, D.C. (Waverly Taylor), Toledo, OH (Don Scholz), Cleveland (Maurice Fishman), Chicago (Irvin Blietz), Kansas City (J.D. Nichols), Phoenix (Del Webb) and San Francisco (Carl Gellert and Ellie Stoneson). Checkoway (1983) also notes that a few builders such as Dave Bohannon, Fritz Burns and James Price actually replicated the strategy across multiple markets. Checkoway (1983) concludes that three factors distinguished this new wave of builders: their size, their lower costs and their suburban focus. He argues that these three traits combined to allow a doubling of the number of new housing starts in the 1950s compared to the 1940s (i.e., 15.1 million starts from 1950-59 versus 7.4 million from 1940-49) without engendering rising real costs that could have made the homes unaffordable. Using data from the San Francisco Bay Area, Maisel (1953) was the first to document both lower costs and higher profits for larger builders who could employ their new production techniques across (potentially) thousands of single-family residential parcels located within a single community on expansive tracts of vacant suburban land. Others such as Herzog (1963) wrote a dissertation on large scale homebuilding. Weiss (1987) describes the rise of this type of large residential developer in his book, *The Rise of Community Builders*.

Figure 8: The Size Distribution of Vacant Land Purchases Intended for Single-Family Housing Development



Note. The figure plots the cumulative distribution function for the share of parcels below given square footage amounts. There are 43,560 square feet in one acre. The underlying data are vacant land purchases intended for single-family housing development for 24 CBSAs over the years 2013-2018. The plot is based on 3,640 observations of vacant land parcel purchases. The individual observations were downloaded from proprietary CoStar files and used in Gyourko and Krimmel (2021). See their paper for more details.

the 99^{th} percentile of the parcel size distribution is 314 acres. Projects with more than 500 acres are essentially non-existent. This pattern also holds if we restrict our attention to places that have large amounts of empty land around—thus casting doubt on the idea that smaller project sizes are due to the fact that land now is scarcer compared to the 1950s. Atlanta is one example of a place with abundant land, and Appendix Figure E16 shows that it has a similar project size distribution.

4.3.2 A New Series of Project Sizes Over Time

When did America start losing its Levitts? Answering this question is challenging because there is no consistent data on the scale of housing developments over time. We address this problem by using CoreLogic's extensive micro data on housing units that include detailed location data (i.e., precise GPS coordinates) along with information on the year a home was built to construct a measure of project or development size over time.

4.3.2.1 Algorithm

We use the sequence described below to group single-family homes (which can be detached or attached) into a housing development whose size is defined by the number of homes in it.[41] For a given county, the process is as follows.

1. Draw a home at random and create a filtering square around the chosen property. This is done by drawing a square that extends 100 yards in every cardinal direction from the GPS coordinates of the housing unit.

2. Count the number of other single-family homes within the square that were built within a 2-year period that begins with the year the focal unit was constructed. The 2-year periods start with 1950-1951 and end with 2018-2019.

   (a) If there are no similarly-aged single-family homes within the 100-yard square surrounding the focal unit, that is counted as a one-unit housing development. I.e., project or development size equals one.

3. If there are other homes within the square that were built within the relevant 2-year window of time, we draw 100-yard squares around each GPS coordinate redoing Step 2.

   (a) This process is repeated until there are no similarly-aged single-family homes within the 100-yard square surrounding any house determined to be in the housing development in a prior iteration of Step 2.

   (b) Project or development size is given by the total number of closely clustered homes found in all rounds of the analysis. If 100 homes were identified, we define the size of that development to be 100.

4. After removing the homes in the development just defined from the county-wide data, the process is repeated until the number of developments equals 10% of the original sample size in the underlying county.

[41] CoreLogic has a product code to identify single-family homes. Other codes allow us to identify specific types of owner-occupied homes such as traditional detached units, townhomes and attached owner-occupied product, and trailer park homes. We exclude the latter from our samples, as they were not produced on site by a home builder. Owner-occupied condominiums in multiple unit structures also are excluded from our analysis. In addition, we restrict our analysis to counties with less than 10% of the observations containing missing values for the year the home was built. The larger the share of such missing values, the greater the potential downward bias in our estimated project size.

We estimated this process on data for 167 counties listed in Appendix Table E7, covering 51% of the US population. Appendix Section E.3.1 reports an example of the algorithm's output by showing maps of Los Angeles County that zoom in on some of the large projects we identified.

4.3.2.2 Results

Our first result is that the typical homebuilding development project is very small, and has been since at least 1950. The median project size always equals one. Even at the 75th percentile of the project size distribution, scale is small. The 75th percentile value for any 2-year period is less than 5 until 2006-2007, when it reached 7. However, it dropped back to 4 in the latest data we have, from 2018-2019. It is not until the 90th percentile of the project size distribution that scale reaches more than ten housing units. The 90th percentile project size value ranges from 12 to 24 depending upon the 2-year time period, and even a 20-house project is not large relative to most manufacturing runs.

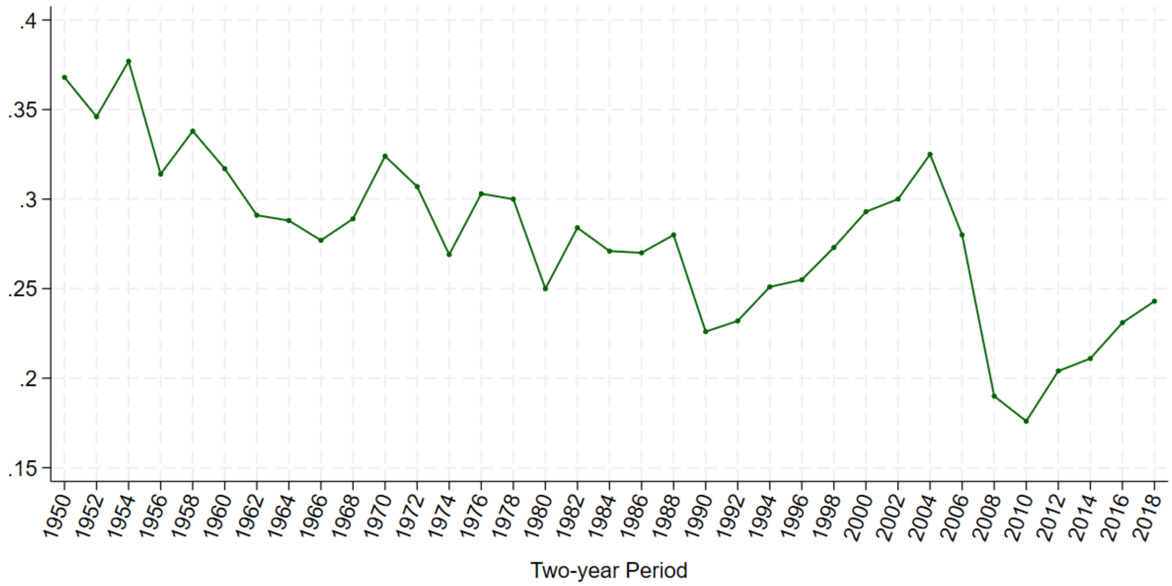
Large projects are relatively rare, but they can be quite big. Figure E20 in Appendix Section E.3.2 shows that the 99th percentile of the project size distribution averages just under 150 homes throughout the 70-year time span of our data. There is a flat trend to this series, as it neither increases nor decreases appreciably over time.[42] Appendix Figure E21 provides more detail on the size distribution of this 1% of largest projects over time. Here, we see some very large projects. The top tenth of these larger projects (i.e., the top 0.1% of all projects of any size) averages about 500 homes per two-year period, or about five completed homes per work week. The largest projects, typically from the 1950s, range from 5,000 to 10,000 units in size.

Figure 9 plots the share of homes from the largest 1% of projects in terms of all homes built in each respective 2-year period sampled by our algorithm.[43] This plot depicts a drop over time in the share of new single-family homes coming from the largest 1% of projects. At the beginning of our time period in 1950-51, 37% of all homes built in our 167 county sample were from these

[42] There is cyclicality about the flat trend, with the amplitude about the latest housing boom and bust marked by the Global Financial Crisis being the largest by far. That period bears closer scrutiny, but that is beyond the scope of this paper.

[43] In Appendix Figure E22 we compute the share of homes from the largest 1% of projects defined as homes built in large projects divided by all homes built, regardless of whether our algorithm sampled them. This is equivalent to assuming that there are no large projects in the unsampled portion of the data, thus providing a lower bound for the share of large projects. The pattern over time is almost identical, because our algorithm ends up sampling most houses. Although we start from a 10% sample of new developments, we then look at all new houses near the initial draw, and all new houses near those latter ones, and so on. This iterative process covers a large fraction of all new developments.

Figure 9: Share of New Housing Stock Built in Largest 1% of Projects



Note. The figure plots the share of homes built in the largest 1% of projects in terms of all homes built in each respective 2-year period.

larger developments. The latest data from 2018-2019 indicate a 24% share. Overall, this represents a one third drop in share of homes being built in large projects, which reflects the fact that the scale of the largest projects has dropped substantially since the 1950s.[44]

These data also enable us to reject the hypothesis that project sizes are declining because an increasingly wealthy population demand a highly personalized home.[45] Appendix Section E.3.3 reports cross sectional hedonic house price regressions using the most recent sale price in the

[44]One concern with our interpretation of this trend is that the measurement error implicit in our algorithm might change over time in a way that drives the trend. The main source of measurement error is that we might pick up many small projects that just happen to be contiguous, but that are not actually one large project. If this bias is larger at the start of the sample, this might explain some of our patterns. To shed some light on this, Appendix Figure E23 computes the ratio of new developments in large projects, but weighting each project by the inverse of the within-project coefficient of variation in the living area of units. If we are capturing many unrelated small projects that happen to be contiguous, we would expect the variance in living area to be high within what we call a large project, and the weighting penalizes such diverse (and likely mismeasured) projects. While the y-axis loses its natural meaning, the pattern over time is even starker. The share of large projects in this case declines by almost two-thirds. This is because, if anything, our measurement error is likely larger in more recent years, where it is more likely to see multiple contiguous small projects getting developed.

[45]This does not mean the homes from large developments are of lower quality. For example, the early Levitt & Sons' mega-projects in Long Island and Bucks County were technologically advanced for their time, as they were the first to include heating through the floor, a feature that was copied by other builders (e.g., Rybczynski 2017).

CoreLogic data, housing and site characteristics and whether the house was from a large project.[46] Unconditionally, homes from relatively large projects are about 10% cheaper than those from elsewhere in the relevant county. After controlling for house age, lot size, living square footage, census tract and year of sale fixed effects, the difference shrinks to less than 2%. We also find that, over time, homes from larger developments do not sell at a significant discount today relative to those that were built in smaller developments.[47]

In sum, project sizes are small and they have shrunk over time, and small construction firms work on those small projects. These are two key features of the model, but they are not evidence that regulation is responsible for the small firms and small projects. We turn to the cross-section correlates of regulation now.

5 Cross-Sectional Evidence

We now test the three main cross-sectional implications of our model: (1) smaller firms are less productive, (2) tighter project regulation is associated with smaller firms, and (3) tighter project regulation is associated with less productive firms.

[46] We used 1,000+ and 100+ unit projects to define "large." Because most really large projects were built in the past, the most recent transaction price is almost never the original sales value. See Appendix Section E.3.3 for the details.

[47] A second analysis also described in Appendix Section E.3.3 examined differences in appreciation over time. Because CoreLogic reports only the most recent sale price, we use mean self-reported price data reported at the census tract level by decade to measure appreciation over time. We are able to do so starting in 1970, as that is the earliest date for which we have consistently defined tract areas over time. The specifications estimated compare the appreciation rates in tracts dominated by large developments compared to the appreciation rates in nearby tracts that contained only small projects of less than five homes, all within a common 2-year time frame. Once again, the results do not indicate that a powerful increase in love of variety has caused homes built in nearby smaller projects to appreciate substantially more than those that were built together with another 100 (or 1,000) probably more similar homes. Unconditionally, Appendix Table E9 shows total appreciation in tracts dominated by large projects to be at least 10% less than in nearby tracts with only small projects. Controlling as best as possible for housing quality differences (at the tract level) reduces the difference, sometimes to zero. The comparison is between appreciation in the tract(s) containing the large project with at least 100 (or 1,000) homes built close together within a 2-year window with that found in tracts within one mile of the focal tract that also only had small projects of fewer than five homes built during the same time period. See Appendix Section E.3.3 for more detail.

5.1 Productivity and Firm Size in Housing Construction

To look at the connection between size and productivity, we estimate housing units per employee and revenues per employee for different firm sizes. We construct revenues per employee in housing construction across different size bins by computing the total of revenues across all firms within a size category and then dividing by the total employment in firms within that size category. We normalize this measure to one for the firms in the 0 to 4 employees category and report the values with the red line in Figure 10.

Due to disclosure requirements, we do not have access to the distribution of housing units constructed by firms specialized in housing. We only have the distribution of housing units built by all firms in the CCI, including those that do not specialize in housing. Consequently, we use the employment and housing built of all CCI firms, but we adjust employment based on the fraction of revenues in each firm that come from housing. For instance, we count all employees of a firm that derives all of its revenues from housing construction, but only half of the employees of a firm that earns half of its revenues from housing construction. We also normalize the measure of housing units per employee to one for the smallest firms. The green line in Figure 10 shows units per employee.[48]

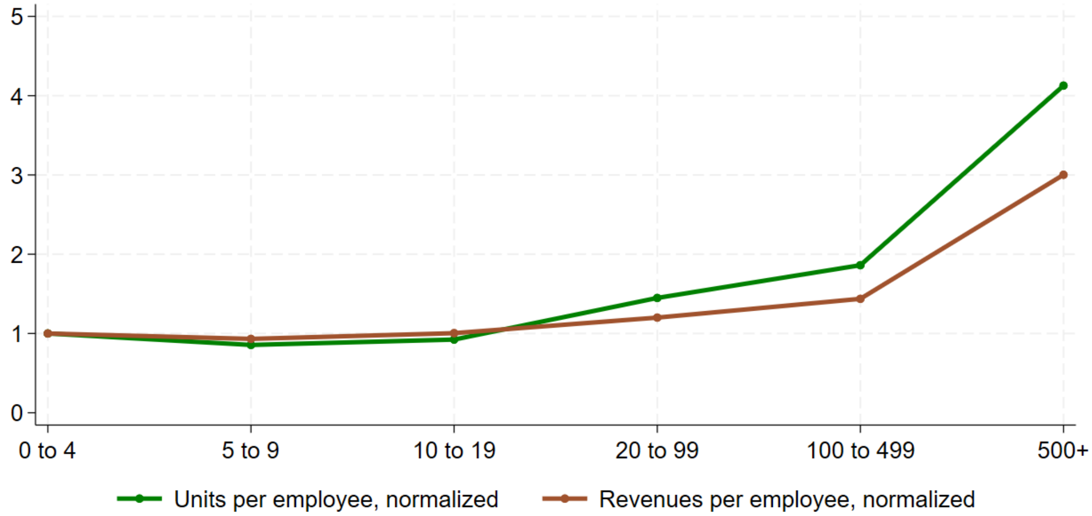
Both revenues and units per employee increase across the size distribution. Firms with 20 to 99 employees produce 45% more units per employee than the smallest firms. Firms with 100 to 499 employees produce almost double the units per employee, while employees in firms with more than 500 employees produce more than four times as much.

We also estimate the cross-sectional relationships between firm size and profits, revenues, and capital per employee. We estimate firm profits by subtracting spending on payroll, benefits, normal depreciation charges, rental payments, materials and supplies, contract labor, fuels, electricity, and reported other purchases from firm revenues.[49] This measure is negative for a significant number of companies. We transform it into a z-score, by subtracting the mean and dividing by the standard deviation of calculated profits. For revenues per employee, which we refer to as labor productivity, we adjust for the total number of subcontractors since subcontracting is responsible for about one-fourth of construction industry revenues in the public tabulations of the 2012 CCI.

[48]Without this adjustment, the gradient of productivity with respect to size would be even steeper because smaller firms derive a larger fraction of revenues from housing construction.

[49]The formula using CCI form numbers is 100 − sum(300, 223, 540, 550, 421, 423, 431–434, 425, 449). The results are robust to dropping non-cash expenses like depreciation.

Figure 10: Output and Revenues per Employee



Note. The figure plots housing units (green) and revenues (red) per employee for construction firms in different size bins. Microdata from the 2012 Census of Construction Industries (CCI). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

Our profit measure includes subcontractor payments as an expense, along with materials purchased on behalf of subcontractors; and thus takes that into account. Our primary labor productivity metric, however, is total revenues divided by employees; the numerator includes some revenues linked to subcontracting while the denominator is own firm employment. This asymmetry is typical of micro-data on firms. To adjust for this, we also report a labor productivity metric where we subtract from revenues the payments that were made to subcontractors. Thus, this measure quantifies firm revenue due to work directly undertaken by the firm divided by the employment in the firm.

Table 1 provides descriptive statistics on our sample. Table 2 reports the regression results of:

$$y_i = \alpha + \beta \times \log(\text{Empl}_i) + e_i, \tag{9}$$

where y_i is profits in column (1); the log of revenues per employee in column (2); the log of revenues per employee, adjusting for subcontractors, in column (3); and the log of capital per employee in column (4). Regressions are unweighted and report robust standard errors. Across firms, we see that a 10% increase in firm employment corresponds to around a 0.014 standard deviations increase in profits, a 1.1% increase in labor productivity, and 0.7% increase in capital

Table 1: Descriptive Statistics for the Firm-level Sample

	Mean (1)	SD (2)
Firm profits, estimated	545.9	5346
Labor productivity	238.7	310.4
Log labor productivity	5.066	0.8543
Employment	32.1	197.4
Log employment	2.325	1.396
Capital (assets) / employee	45.91	81.68
Log capital / employee	3.083	1.231
Labor productivity (with subcontractor adjustment)	178.1	197.8
Log labor prod. w. adjustment	4.875	0.7632
Firm revenue composition		
Share housing	42.38	44.36
Share industrial buildings	6.55	18.85
Share consumer-facing buildings	9.38	21.34
Share office buildings	7.38	18.19
Share warehousing	2.06	9.14
Share other buildings	13.29	25.56
Share non-building construction	18.96	36.86

Note. Microdata from the 2012 Census of Construction Industries (CCI). The sample has 107,000 observations (rounded per Census Bureau disclosure requirements). Firm profits are estimated by subtracting from firm revenues the amount the firm spent on payroll, benefits, normal depreciation charges, rental payments, materials and supplies, contract labor, fuels, electricity, and reported other purchases. Values are in thousands of nominal dollars in 2012; values can be negative. Labor productivity is measured as revenues per employee. The subcontractor adjustment subtracts from revenues the payments that were made to subcontractors. Firm revenue composition is developed by aggregating establishment-level reporting of their construction revenues split out by 31 types of activities (e.g., "single-family homes, detached," "bridges and elevated highways," "decks, residential types"). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

per employee.[50]

[50] In Appendix Tables F10 and F11 we illustrate heterogeneity across different types of construction activity. The link between firm size and productivity is strong and meaningful for housing, but it is even larger for other forms of construction. The only exception is when we look at capital per employee. For firms fully specialized in housing, this displays a weakly negative link with size. Intuitively, the forms of construction where smaller firms are more prevalent (such as housing) are also those that show the weakest link between firm size and productivity. These patterns are robust across many specification variants. In addition, we find comparable results when running regressions for each sector of specialization by itself. The advantages of the integrated model using continuous variables for revenue shares are that (1) it better accounts for mixed revenue streams and (2) it only requires disclosure of a single firm sample. These results are quite robust to many variants like controlling for a firm's location, using sample weights, and similar.

Table 2: Firm Size, Firm Profitability and Labor Productivity

VARIABLES	(1) Profits in unit standard deviations	(2) Log labor productivity	(3) Log labor productivity with adjustment	(4) Log capital per employee
Log employment	0.1375***	0.1094***	0.1017***	0.06882***
	(0.0062)	(0.0020)	(0.0018)	(0.0027)
Observations	107,000	107,000	107,000	107,000
R-squared	0.0368	0.0319	0.0346	0.0061

Note. The table reports results from an OLS regression at the firm level of profitability, capital, and productivity against the log of the number of employees. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Analysis using microdata from the 2012 Census of Construction Industries (CCI). Regressions are unweighted and have 107,000 observations (rounded per Census Bureau disclosure requirements). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

5.1.1 A Counterfactual Productivity in Construction

We now ask how much construction productivity would increase if construction firms were as large as those typically found in either manufacturing or other nontradables. While we observe the relationship between size and productivity, this relationship will reflect both the impact of size on productivity and the impact of exogenous productivity differences on size, just as in our model. Lacking exogenous variation in firm size to identify its effects on productivity, we perform a simple yet transparent exercise and assume that a fraction ϕ of the observed empirical relationship between establishment size and productivity represents the causal effect of size on productivity.

We let j indicate a firm-size-bin (0–4 employees, 5–9, and so on), and denote with $\{N_{0-4}, N_{5-9}, ...\}$ the firm size distribution in construction, where N_j indicates the fraction of employment accounted for by firms in bin j. We let $\{N'_{0-4}, N'_{5-9}, ...\}$ denote the counterfactual firm size distribution, where $N'_{0-4} < N_{0-4}$, for example, means that in the counterfactual we are moving workers out of small firms towards other size bins. Finally, we let $\bar{a}_j$ denote the measure of output per employee for firms in bin j that we estimated in the data.[51] If ϕ represents the fraction of observed productivity differences across firms of different sizes that is causal, then the aggregate change in

[51] These are the units or revenues per employee across the FSD that we reported in Figure 10.

Table 3: Counterfactual Productivity in Construction under Different Assumptions on the Link between Size and Productivity (ϕ)

Counterfactual Size Distribution	Change in Units per Employee (%)			Change in Revenues per Employee (%)		
	$\phi = 1$	$\phi = 0.5$	$\phi = 0.1$	$\phi = 1$	$\phi = 0.5$	$\phi = 0.1$
Manufacturing	+119%	+59.5%	+11.9%	+90%	+45%	+9%
Non-tradables	+107%	+53.5%	+10.7%	+81.6%	+40.8%	+8.2%

Note. Values of ϕ indicate different assumptions on how much of the empirical relationship between productivity, defined as units per employee, and size is causal. $\phi = 1$ assumes that all of the empirical relationship between firm size and productivity is causal, $\phi = 0.1$ assumes that only 10% of it is causal. Analysis using microdata from the 2012 Census of Construction Industries (CCI). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

productivity, Δ, from a shift in the firm size distribution is:[52]

$$\Delta = \sum_j \underbrace{\left(N'_j - N_j\right)}_{\substack{\text{Reshuffling of workers} \\ \text{across firms of} \neq \text{size}}} \times \underbrace{\phi \times \bar{a}_j}_{\substack{\text{Effect of size} \\ \text{on productivity}}} . \tag{10}$$

Table 3 reports these changes for different values of ϕ. The first three columns report changes in average units per employee, while the last three focus on average revenues per employee. If we assume that half of the link between size and productivity is causal, construction firms would produce 59.5% more units per employee if their size distribution matched that of manufacturing. Even if only 10% of the link between productivity and size is assumed to be causal, this estimate would still be 11.9%. This exercise tells us nothing about why construction firms are so small, but suggests that if even a small fraction of the link between firm size and productivity is causal, small firm sizes may be responsible for a significant part of the underperformance of the construction sector. We now turn to the link between regulation and small firm sizes in construction.

[52] See Appendix F.1 for a formal derivation of Equation (10).

5.2 Regulation, Productivity, and Firm Size

To examine the link between regulation and firm size, we use the 2006 version of the Wharton Residential Land Use Regulation Index (WRLURI) described by Gyourko et al. (2008). The survey is at the jurisdiction level (typically at the census place level), and we aggregate it to the CBSA level by averaging across jurisdictions.[53]

The WRLURI measure is available for around 2,600 communities, which aggregate up to 550 CBSAs. Since we are missing data entirely for approximately 300 CBSAs, and in many CBSAs we have only a few jurisdictions, we also replicate our analysis using a "projected" WRLURI. For the CBSAs for which WRLURI is available, we run a regression where WRLURI is predicted based on the (log of) population and population density in a given CBSA, the average school years of male and female residents, and fixed effects at the Census Division level.[54] The predicted WRLURI has a 0.6 correlation with the raw index.[55]

We measure firm size with: (1) the log of total receipts per establishment and (2) the fraction of employment that is in establishments of firms with more than 100 employees.[56] Both variables come from the Census's 2012 Statistics of U.S. Businesses (SUSB). We analyze firm size for construction as a whole and for its subsectors (three-digit NAICS classification): construction of buildings, heavy and civil engineering construction, and specialty trade contractors.

[53] Results are robust to weighting jurisdictions using population weights and land area weights. See Gyourko et al. (2008) for a complete discussion of the index.

[54] Some CBSAs belong to more than one division. In this case, we attribute a CBSA to the state where the highest share of its population resides. For robustness, we run our prediction exercise in three other ways: (1) using our baseline variables interacted by Census Division fixed effects; (2) adding a set of economic variables, the (log of the) self-reported median house values prevailing in the CBSA, residents' self-reported per capita income, the (log of the) median rent, and the (log of the) number of housing units, distinguishing between owner-occupied, renter-occupied and vacant housing units; (3) interacting baseline and economic variables with Census Division fixed effects. Appendix Table F16 reports the results of these different specifications. In the most saturated specification, our R^2 is 0.574. Results are robust to all different specifications, as shown in Appendix Tables F18 to F25.

[55] Our results are robust to restricting the sample to CBSAs where we have more than 5 jurisdictions responding to the survey. To exploit the full granularity of the data, we also predicted WRLURI by running our predictive regressions directly at the place level and then aggregating up the predicted place-level values at the CBSA level. Results (not reported) are robust to this more granular specification and often stronger.

[56] Due to censoring, we cannot retrieve employment counts in large firms in 50 to 100 CBSAs, depending on the particular subsector we focus on. As a consequence, we impute missing employment in large firms using the national mean of employment per firm in each bin, multiplied by the (non-censored) number of firms operating in the CBSA. In Appendix Section F.3 we detail for how many CBSAs we perform the imputation, and show that alternative imputation methods yield almost identical effects.

Using these two measures of firm size and both actual and predicted WRLURI, we estimate:

$$y_i = \alpha + \beta \times \text{WRLURI}_i + X_i'\gamma + \epsilon_i, \tag{11}$$

where y_i is firm size (receipts per establishments or the share of employment in large firms), and X_i denotes a vector of controls including the log of population, population density and the log of (self-reported) median house values, based on the American Community Survey over the period 2008-2012. We run regressions at the CBSA level and we weight by population.

Table 4 reports descriptive statistics for the dependent variables and the WRLURI measures. Table 5 reports results from regressing the log of receipts per establishment on both the raw WRLURI (panel A) and predicted WRLURI (panel B). Areas with stricter land-use regulation display lower revenues per establishment. The results are stronger for the predicted WRLURI measures than for actual WRLURI measures, which is consistent with the view that mismeasurement of land-use regulation leads to attenuation bias. Column (1) reports the results when considering all establishments in the entire construction industry. The coefficient implies that a one standard deviation increase in the raw WRLURI measure is associated with a 10.8% decrease in receipts per establishment. Columns (2) to (4) replicate the same analysis for its subdivisions, and show that construction of buildings is particularly sensitive to regulation in panel A, but land-use regulation has a larger impact on firm size in heavy and civil engineering construction in panel B.

Table 6 reports results on the fraction of employment in establishments of large firms. Stricter land-use regulation is associated with a larger share of employment in smaller firms. Across the entire industry, a one standard deviation increase in land use regulation (using the raw index) is associated with a 2.5 percentage point reduction in the fraction of employment in large firms (14% of the mean). In construction of buildings, again for the raw index, a one standard deviation increase in the regulation index is associated with a 4.3 percentage point reduction in employment in large establishments, which is 38% of the mean.[57]

[57] Appendix Tables F18 to F25 report all coefficients including controls, as well as robustness to different projections of WRLURI.

Table 4: Descriptive Statistics for the CBSA Sample

	Mean	SD
WRLURI	-0.393	0.713
Projected WRLURI	-0.458	0.423
All construction		
Log of receipts per establishment	14.127	0.503
Share of employment in large firms	0.181	0.166
Construction of buildings		
Log of receipts per establishment	14.087	0.684
Share of employment in large firms	0.112	0.163
Heavy and Civil Engineering		
Log of receipts per establishment	15.013	0.902
Share of employment in large firms	0.324	0.285
Specialty Trade Contractors		
Log of receipts per establishment	13.786	.470
Share of employment in large firms	0.133	0.204

Note. Descriptive statistics on WRLURI and SUSB variables.

5.2.1 Micro Evidence Using the Census of Construction Industries

We now turn to the Census of Construction Industries, using the same sample of firms described in Section 4.2.2.[58] We measure the local construction sector's revenues per capita (total, and split by housing vs. other construction), housing units built per capita, revenues and payroll per construction employee and per firm, and the average size of construction establishments compared to the average national size of such establishments.

While we report only the coefficient on the WRLURI value, our model is identical to the one described in the previous section. We always control for log CBSA population, log housing value, and density. Observations are again weighted by CBSA population. Variables are winsorized at their 1% and 99% values. Our CBSA count is modestly higher than the external data given the internal records and the use of the projected WRLURI 2006 value. Our results are robust to using different predictive regressions and using the raw index.

[58]To assign CBSA location accurately, these measures are calculated using establishment-level data. Weights are included in the CBSA collapse.

Table 5: Regulation and Receipts per Establishment

	(1) All Construction	(2) Construction of Buildings	(3) Heavy and Civil Engineering	(4) Specialty Trade Contractors
Panel A: Original WRLURI				
WRLURI	-0.1616***	-0.1810***	-0.1254*	-0.1579***
	(0.0456)	(0.0525)	(0.0657)	(0.0489)
Log of population (2008-2012)	0.2452***	0.3221***	0.3257***	0.2148***
	(0.0320)	(0.0379)	(0.0403)	(0.0249)
Constant	10.6805***	9.3842***	10.9557***	10.0791***
	(1.2267)	(1.2864)	(1.5175)	(1.2303)
Other controls	✓	✓	✓	✓
Observations	545	530	422	543
R-squared	.4224	.4328	.4154	.417
Panel B: Projected WRLURI				
WRLURI	-0.2925***	-0.2899**	-0.3948***	-0.2188***
	(0.0838)	(0.1135)	(0.1143)	(0.0739)
Log of population (2008-2012)	0.2544***	0.3375***	0.3562***	0.2174***
	(0.0302)	(0.0356)	(0.0342)	(0.0246)
Constant	9.9374***	8.6560***	8.5410***	10.0033***
	(1.2057)	(1.3965)	(1.4053)	(1.2205)
Other controls	✓	✓	✓	✓
Observations	852	808	627	852
R-squared	.4456	.4696	.4568	.4228

Note. The table reports results from WLS regressions at the CBSA level of the log of receipts per establishment against WRLURI, both raw (in panel A) and projected using demographic characteristics (panel B). Each CBSA is weighted by population. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. We use the 2006 version of WRLURI (Gyourko et al. 2008). Other controls include population density and the log of self-reported home values (from the 2008-2012 ACS5A). Projected WRLURI is obtained by projecting raw WRLURI on the (log of) population and population density in a given CBSA, the average school years of male and female residents, and fixed effects at the Census Division level. These regressors are from the 2000 Decennial Census.

Table 6: Regulation and the Share of Employment in Large Firms

	(1) All Construction	(2) Construction of Buildings	(3) Heavy and Civil Engineering	(4) Specialty Trade Contractors
Panel A: Original WRLURI				
WRLURI	-0.0354***	-0.0607***	-0.0396*	-0.0328**
	(0.0128)	(0.0146)	(0.0217)	(0.0128)
Log of population (2008-2012)	0.0742***	0.0797***	0.1096***	0.0732***
	(0.0093)	(0.0092)	(0.0111)	(0.0081)
Constant	-0.5453*	-1.1150***	-0.8687*	-0.8772**
	(0.3213)	(0.3390)	(0.4436)	(0.3454)
Other controls	✓	✓	✓	✓
Observations	541	488	381	532
R-squared	.3639	.3251	.3382	.409
Panel B: Projected WRLURI				
WRLURI	-0.0618***	-0.0795***	-0.0963**	-0.0332
	(0.0222)	(0.0268)	(0.0381)	(0.0225)
Log of population (2008-2012)	0.0756***	0.0801***	0.1145***	0.0721***
	(0.0086)	(0.0081)	(0.0098)	(0.0077)
Constant	-0.6836**	-1.1601***	-1.2885***	-0.7820**
	(0.2907)	(0.2943)	(0.4695)	(0.3384)
Other controls	✓	✓	✓	✓
Observations	846	730	541	830
R-squared	.3719	.3302	.3673	.4079

Note. The table reports results from WLS regressions at the CBSA level of the share of employment in large firms against WRLURI, both raw (in panel A) and projected using demographic characteristics (panel B). Each CBSA is weighted by population. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. We use the 2006 version of WRLURI (Gyourko et al. 2008). Other controls include population density and the log of self-reported home values (from the 2008-2012 ACS5A). Projected WRLURI is obtained by projecting raw WRLURI on the (log of) population and population density in a given CBSA, the average school years of male and female residents, and fixed effects at the Census Division level. These regressors are from the 2000 Decennial Census.

Table 7: Regulation, Firm Size, and Construction Output per Capita

(a) Firm Size and Productivity

VARIABLES	(1) Log revenues per firm	(2) Log payroll per firm	(3) Log capital per firm	(4) Log revenues per emp.	(5) Log payroll per emp.	(6) Log capital per emp.	(7) Firm size relative to ave. national size
Projected WRLURI	-0.3235***	-0.2565***	-0.3317***	-0.1301***	-0.0729***	-0.1182**	-0.1426*
	(0.094)	(0.0867)	(0.0859)	(0.0367)	(0.0282)	(0.0479)	(0.0857)
Controls	✓	✓	✓	✓	✓	✓	✓
Mean	7.400	5.307	5.932	5.785	3.696	3.84	0.8652
SD	0.5359	0.2521	0.5472	0.5243	0.1898	0.3569	0.2944
Observations	650	650	650	650	650	650	650
R-squared	0.412	0.373	0.196	0.433	0.491	0.073	0.283

(b) Construction Activity per Capita

VARIABLES	(1) Log of housing units built per capita	(2) Log of total CCI revenues per capita	(3) Log of housing CCI revenues per capita	(4) Log of non-housing CCI revenues per capita
Projected WRLURI	-0.5299*	-0.3718***	-0.2577***	-0.4010***
	(0.2721)	(0.0849)	(0.0901)	(0.0972)
Controls	✓	✓	✓	✓
Mean	-7.361	1.115	-0.320	0.761
SD	1.475	0.584	0.656	0.701
Observation	650	650	650	650
R-squared	0.209	0.335	0.393	0.280

Note. The table reports results from a WLS regression at the CBSA level of different measures of firm size (panel a) and construction activity (panel b) against WRLURI projected using demographic characteristics. Each CBSA is weighted by population. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The analysis uses the 2006 version of WRLURI (Gyourko et al. 2008) and microdata from 2012 Census of Construction Industries (CCI). CBSA regressions control for log population, log housing values, and population density. Variables are winsorized at their 1% and 99% values. The sample has 650 observations (rounded per Census Bureau disclosure requirements). Projected WRLURI is obtained by projecting WRLURI on the (log of) population and population density in a given CBSA, the average school years of male and female residents, and fixed effects at the Census Division level. These regressors are taken from the 2000 Decennial Census. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

Table 7 roughly corresponds to Propositions 2 and 3 in the model. Panel (a) corresponds more closely to Proposition 3. Columns (1), (2), (3) and (7) report the link between regulation and firm size. A one standard deviation increase in the WRLURI index (0.42) is associated with a 12.8%, 10.3%, and 13.9% decline in revenues, payroll, and capital per firm. Firms shrink by 6 percentage points relative to the average construction firm as WRLURI increases by one standard deviation. Columns (4), (5) and (6) turn to the link between regulation and firm productivity, measured by revenues and payroll per employee. The coefficient estimates in those regressions imply that a one standard deviation increase in WRLURI corresponds to a 5.4% decrease in revenues per employee, a 3% decrease in payroll per employee, and a 5% decrease in capital per employee. These results confirm that more restrictive project regulation is associated with lower labor productivity in construction firms, whether measured in average or in marginal terms.[59]

Panel (b) documents the correlation between land use regulation and the size of the construction sector, which corresponds to the predictions of Proposition 2. More regulated places build fewer homes, and total construction-related revenues are lower. A one standard deviation increase in WRLURI is associated with a 20.2% decrease in housing units built per capita and a 14.6% decrease in revenues per capita. Although the index mostly captures residential regulation, this decrease in non-housing revenues is unsurprising both because land-use regulation affects non-housing activity as well, and because lower housing supply hampers population growth, which in turn hampers the growth of other activities. Non-housing-related revenues may decline by more than housing-related revenues, perhaps because the amount of housing built per capita is less flexible than the amounts of other forms of construction per capita.

6 Evidence on Innovative Activity

A final key prediction of our model is that when projects become smaller, firms become smaller too and they invest less in productivity-enhancing technology.[60] We focus on patents, which provide a reliable long-run series on innovative activity across sectors within the U.S. We supplement this evidence with additional data based on spending on research and development. We expect

[59] In Section 3 we assumed for simplicity that the cost of inputs is independent of industry conditions. However, generalizing to a finite elasticity of labor supply at the city-industry level, a decline in construction wages would follow from the contraction in construction employment that our theory predicts as a consequence of tighter project regulation.

[60] Even when firms are larger, some industry professionals claim that "the bespoke nature of each real estate project makes fixed R&D costs impossible to recoup" (Scherr 2024).

to see less innovative activity in the construction sector, and an increasing divergence between construction and manufacturing since 1970, given that project and firm size have both decreased since that point—presumably because of increasingly restrictive land-use regulation.

Patents are, of course, an imperfect measure of innovative activity. Many forms of innovation are not patented at all, and the sheer number of patents may not capture the importance of the patents in any given sector or year. Nonetheless, it is hard to think of any long-term measure that is likely to get more directly at innovative activity than patents. We consider all patents granted since 1836, when the United States Patent and Trademark Office (USPTO) was established, through 2020. We include only patents with one or more U.S.-based inventors. Patents are classified into technologies, and we collect the primary technology of a patent.[61] Figure 11 shows the patent levels for construction, manufacturing, and other industries over time. We normalize the indices to take on a value of one in 1930 to remove time-invariant differences across sectors. Until the 1950s, the series move together tightly. There is a secular increase, with occasional dips, through 1931. After that point, all three series plunge together and reach their nadirs in 1947. It is after that point that the divergence of the series begins.

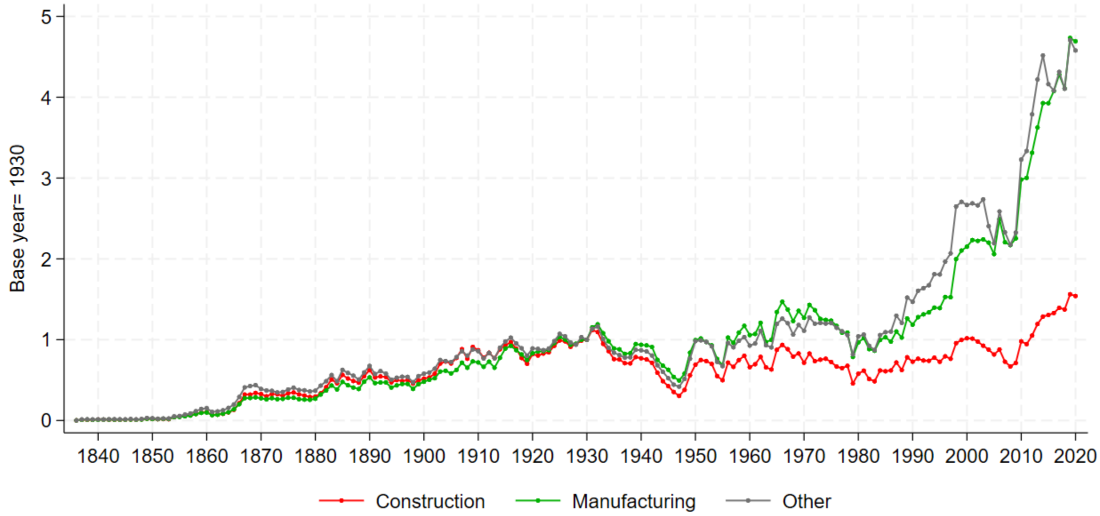
The gaps between the series are relatively small through the 1960s. In 1967, construction-related patenting has almost reached the level seen in 1930, and the other two series are about 50% higher than they were then. But after that point, construction patenting falls more sharply than the other series. A permanent wedge emerges, expanding substantially after 1990 until the present. These series show that construction patents have become a far smaller share of total USPTO patents. In 1909, when construction firms were engaging in pioneering building feats, such as erecting skyscrapers, the construction share of total patents peaked at 8.5%. This share has declined continuously since then and only 2.5% of USPTO patents were in construction in 2020.[62]

[61] In 2013, the USPTO transitioned from the US Patent Classification (USPC) system (which it had maintained from 1836 to 2013) to the Cooperative Patent Classification (CPC) system. We use years prior to 2013 when both the USPC and CPC system were in use by the USPTO to create a common technology series from 1836 to 2020. We then use the mapping concordance developed by Kerr (2008) to probabilistically pair USPC technology codes to "industries of use" for the technology. This concordance is derived from a period in the early 1990s when the Canadian patent office simultaneously assigned patents to both a technology code and a likely industry of use (Silverman 1999).

[62] These conclusions are qualitatively unchanged if we normalize patents by employment in each sector, as shown in Appendix Figure G24. The only difference is that the wedge between construction and the rest of the economy is less dramatic than the wedge between construction and manufacturing. This is due to a relative shrinking of manufacturing employment over time, which amplifies the differences with construction and the rest of the economy in raw patenting.

In Appendix Figure G25 we add patents for the other good producing industries—agriculture and mining—plotted separately from the residual category. This shows that construction is starkly different from all goods producing industries, not only manufacturing. Figure G26 indexes these five industries by employment, and here the difference becomes so stark that we need to use a log scale.

Figure 11: Patents by Industry



Note. The figure plots the relative patenting levels for different industries over time, for U.S.-based inventors. All series are indexed to 1930.

We also look at patenting by manufacturing industries that are primary suppliers for construction, defined as those industries that are responsible for the bulk of construction manufacturing inputs.[63] Perhaps, while construction slowed down its innovative impetus, its suppliers might have picked it up. Appendix Figure G27 shows quite the opposite. Manufacturing industries that are primary suppliers for construction saw their innovative activity—relative to the rest of the sector—decline precipitously after 1980.

Over the long run, there are limited measures of patent quality or other traits. Since 1975, we have measures of patent quality such as citations received, and originality and generality scores. Construction patenting is stagnant, but otherwise invention in the sector is not noticeably different in impact. Chattergoon and Kerr (2022) document that much of the recent growth in U.S. patenting is linked to the software sector. They develop a machine-learning algorithm to classify software-based patents. In patents granted after 2015, 39.7% of non-construction patents have a software component, but only 5.6% of construction patents has such a component.

Other sources also suggest reduced innovation in the construction industry. The OECD Data Explorer (link) enables us to compare U.S. R&D expenditure with such expenditure in other countries. In the U.S., 0.23% of total R&D spending was in the construction sector between 2000

[63]We choose the top 20 manufacturing industries in terms of share of construction inputs from manufacturing. Appendix Section G.2 provides more details.

and 2021. The unweighted average across all countries was 0.79%. Across 39 OECD countries, the US ranked 30$^{\text{th}}$ in the share of total R&D spending that is associated with the construction sector. While these data are imperfect—suffering from partial year coverage for some countries, some values being estimated, etc.—they consistently suggest U.S. R&D expenditures are disproportionately lower in construction compared to other countries.

The 2020 Annual Business Survey surveyed innovation activity by companies across the U.S. during 2017–19. NSF Publication 23-310 (https://ncses.nsf.gov/pubs/nsf23310) provides tabulations of responses by industry. While 6.4% of firms in the construction industry were engaged in product innovation during 2017–19, 11.6% of non-construction firms undertook product innovations. Similarly, while 15.4% of construction firms claim that they had made business-process innovations, 22.6% of non-construction firms reported such innovations. These gaps in innovation were even stronger when companies were asked about "new to market" innovations they were undertaking.

R&D expenditure data from public listed companies, which we take from Compustat, also shows that construction is one of the least innovative sectors in the U.S. In the last fifty years, R&D expenditure as a fraction of total revenues has been 0.04% for the sector, falling to a record low of 0.01% for 2015–2023. In manufacturing, R&D expenditure has been 3.58% of total revenues, rising to a high of 4.5% in 2015–2023. Thus, relative expenditure in R&D in manufacturing has been nearly ninety times higher than in construction. For all other sectors, R&D expenditure has still been much larger than for construction, with an average value of 0.66%, and a high of 0.95% in 2015–2023. While Compustat data is surely not representative of each sector, since it covers only publicly listed companies, the magnitude of these relative comparisons strongly corroborates an overall picture of sluggish innovative activity in construction.[64]

[64]Compustat covers firms from 1950 onward. However, their coverage is poorer before 1973, which is why we focus on 1974 onward in the text. Appendix Table G26 reports sectoral R&D expenditure as a percentage of total revenues at a decennial frequency starting from 1974, and we group earlier observations in one bin. Patterns across sectors are very similar for all decades, including the early data. The spike in R&D expenditure in construction before 1973 is attributable to only one firm, Pullman Inc. (Allegan, MI). The calculations reported in the text impute a value of zero to R&D expenditure to all firms that lack data on this item. In the table, we report the same figures restricted to the sample of firms that have non-missing R&D expenditure data. While the absolute numbers increase, the comparison across sectors is qualitatively identical.

7 Conclusion

In this paper, we formally presented the hypothesis that project-level regulation, as opposed to the regulation of entry, reduces firm size and the incentives to invest in technological innovation. We presented a series of facts that are consistent with that hypothesis, documenting the small size of construction firms, especially in more regulated areas, and the lower productivity of smaller firms. We showed that over time construction firms and projects became smaller at the same time as regulation increased, and that the sector's innovativeness also declined. We also developed a back-of-the-envelope calculation suggesting that firm size alone could explain a significant fraction of the low productivity seen in U.S. residential construction.

We hope that future research does more to understand the links between regulation of process, firm size and productivity. A panel analysis of regulatory changes would be helpful. Future work could further disentangle the complex relationship between firm size and productivity, perhaps focusing on inputs like investment in technology and innovation. Structural work could be helpful in generating more sophisticated counterfactual analyses.

Forty years ago, Mancur Olson's (1984) *The Rise and Decline of Nations* described a process through which insiders enact rules that protect themselves from change. The insiders here are existing homeowners able to block new development in their communities via democratic means because of the highly localized nature of land use controls in the United States. Those rules in turn stymie innovation and maintain the status quo. If the hypothesis described in this paper is correct, then what has happened to construction productivity is a variant of the Olson view. Project-level regulations have been put in place that reduce innovation, not by barring it, but by limiting project and firm size. The small scale of the firms, and the fact that they could not grow dramatically even if they made a breakthrough, then limits innovative activity.

References

Akcigit, Ufuk, and William Kerr. 2018. "Growth Through Heterogeneous Innovations." *Journal of Political Economy* 126 (4): 1374–1443.

Allen, Steven. 1985. "Why Construction Industry Productivity Is Declining." *Review of Economics and Statistics* 67 (4): 661–669.

Arnott, Richard J., and Joseph E. Stiglitz. 1979. "Aggregate Land Rents, Expenditure on Public Goods, and Optimal City Size." *Quarterly Journal of Economics* 93 (4): 471–500.

Ball, Robert, and Larry Ludwig. 1971. "Labor Requirements for Construction of Single-Family Homes." *Monthly Labor Review* 94 (4): 12–14.

Ben-Akiva, Moshe, Nicolaos Litinas, and Koji Tsunokawa. 1985. "Continuous Spatial Choice: The Continuous Logit Model and Distributions of Trips and Urban Densities." *Transportation Research Part A: General* 19 (2): 119–154.

Brooks, Leah, and Zachary Liscow. 2023. "Infrastructure Costs." *American Economic Journal: Applied Economics* 15 (2): 131–158.

Building Construction Cost Data, 1998, 56th annual edition. 1998. Kingston, MA: R.S. Means Company.

Calabrese, Stephen, Dennis Epple, and Richard Romano. 2007. "On the Political Economy of Zoning." *Journal of Public Economics* 91 (1): 25–49.

Chattergoon, Brad, and William Kerr. 2022. "Winner Takes All? Tech Clusters, Population Centers, and the Spatial Transformation of U.S. Invention." PDF available, *Research Policy* 51 (2): 104418.

Checkoway, Barry. 1983. "Large Builders, Federal Housing Programmes and Postwar Suburbanization." Chap. 1 in *Readings in Urban Analysis, Perspectives on Urban Form and Structure,* edited by Robert W. Lake.

Cohen, Wesley M, and Steven Klepper. 1996. "A Reprise of Size and R&D." *The Economic Journal* 106 (437): 925–951.

Colton, Kent, and Gopal Ahluwalia. 2019. *A Home Builder Perspective on Housing Affordability and Construction Innovation.* Technical report. Joint Center for Housing Studies, Harvard University.

Consumer Guide, Auto Editors of. 2004. *Cars of the Classic '30s: A Decade of Elegant Design.* Publications International.

———. 2001a. *Cars of the Fabulous '50s: A Decade of High Style and Good Times.* Publications International.

———. 2001b. *Cars of the Sizzling '60s: a Decade of Great Rides and Good Vibrations.* Publications International.

D'Amico, Leonardo. 2022. *Place Based Policies with Local Voting: Lessons From the EU Cohesion Policy.* Working Paper.

Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117 (1): 1–37.

Duranton, Gilles, and Diego Puga. 2023. "Urban Growth and its Aggregate Implications." *Econometrica* 91 (6): 2219–2259.

Ellickson, Robert C. 2022. *America's Frozen Neighborhoods, The Abuse of Zoning.* New Haven and London: Yale University Press.

Fernandez, Raquel, and Richard Rogerson. 1996. "Income Distribution, Communities, and the Quality of Public Education." *Quarterly Journal of Economics* 111 (1): 135–164.

Flammang, James M., and the Auto Editors of Consumer Guide. 2000. *Cars of the Sensational '70s: A Decade of Changing Tastes and New Directions.* Publications International.

Foerster, Andrew T., Andreas Hornstein, Pierre-Daniel G. Sarte, and Mark W. Watson. 2022. "Aggregate Implications of Changing Sectoral Trends." *Journal of Political Economy* 130 (12): 3286–3333.

Ganong, Peter, and Daniel Shoag. 2017. "Why Has Regional Income Convergence in the US Declined?" *Journal of Urban Economics* 102:76–90.

Glaeser, Edward, and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives* 32 (1): 3–30.

Glaeser, Edward, Joseph Gyourko, and Raven E Saks. 2005. "Why have housing prices gone up?" *American Economic Review* 95 (2): 329–333.

Glaeser, Edward, and Giacomo A. M. Ponzetto. 2018. "The Political Economy of Transportation Investment." *Economics of Transportation* 13:4–26.

Glaeser, Edward, and Andrei Shleifer. 2005. "The Curley Effect: The Economics of Shaping the Electorate." *Journal of Law, Economics, and Organization* 21 (1): 1–19.

Glaeser, Edward, and Bryce Ward. 2009a. "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston." *Journal of Urban Economics* 65 (3): 265–278.

Glaeser, Edward L, and Bryce A Ward. 2009b. "The causes and consequences of land use regulation: Evidence from Greater Boston." *Journal of urban Economics* 65 (3): 265–278.

Goolsbee, Austan D., and Chad Syverson. 2023. *The Strange and Awful Path of Productivity in the U.S. Construction Sector.* Working Paper 2023-04. University of Chicago, Becker Friedman Institute for Economics.

Gyourko, Joseph, Jontahan Hartley, and Jacob Krimmel. 2019. *The Local Residential Land Use Regulatory Environment Across U.S. Housing Markets: Evidence from a New Wharton Index.* Working Paper, Working Paper Series 26573. National Bureau of Economic Research.

Gyourko, Joseph, and Jacob Krimmel. 2021. "The Impact of Local Residential Land Use Restrictions on Land Values across and within Single Family Housing Markets." *Journal of Urban Economics* 126.

Gyourko, Joseph, and Raven Molloy. 2015. "Regulation and Housing Supply." Chap. 19 in *Handbook of Regional and Urban Economics,* edited by J. Vernon Henderson Gilles Duranton and William C. Strange, 5:1289–1337. Elsevier.

Gyourko, Joseph, and Albert Saiz. 2006. "Construction Costs and the Supply of Housing Structure." *Journal of Regional Science* 46 (4): 661–680.

Gyourko, Joseph, Albert Saiz, and Anita Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies* 45 (3): 693–729.

Hamilton, Bruce W. 1975. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies* 12 (2): 205–211.

Henderson, J. Vernon. 1974. "The Sizes and Types of Cities." *The American Economic Review* 64 (4): 640–656.

Herzog, John P. 1963. "The Dynamics of Large-Scale Homebuilding." PhD diss., Real Estate Research Program. Institute of Business and Economic Research, UC-Berkeley.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2): 1–39.

Jackson, Kristoffer. 2016. "Do Land Use Regulations Stifle Residential Development? Evidence from California Cities." *Journal of Urban Economics* 91:45–56.

Katz, Lawrence, and Kenneth T Rosen. 1987. "The interjurisdictional effects of growth controls on housing prices." *The Journal of Law and Economics* 30 (1): 149–160.

Kerr, William. 2008. "Ethnic Scientific Communities and International Technology Diffusion." *The Review of Economics and Statistics* 90 (3): 518–537.

Krimmel, Jacob. 2021. *Reclaiming Local Control: School Finance Reforms and Housing Supply Restrictions.* Working Paper. Federal Reserve Board.

Labor and Material Requirements for Construction of Private Single-Family Houses. 1972. Bulletin 1755. U.S. Department of Labor, Bureau of Labor Statistics.

Langston, Craig. 2014. "Construction Efficiency: A Tale of Two Developed Countries." *Engineering, Construction and Architectural Management* 21 (3): 320–335.

Larrabee, Eric. 1948. "The Six Thousand Homes That Levitt Built." *Harper's Magazine.*

Maisel, Sherman J. 1953. *Housebuilding in Transition.* Berkeley, CA: University of California Press.

Olson, Mancur. 1984. *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities.* Reprint edition. New Haven, Conn.: Yale Univ Press.

Potter, Brian. 2021. *Operation Breakthrough: America's Failed Government Program to Industrialize Home Production.* https://www.construction-physics.com/p/operation-breakthrough-americas-failed.

Rothberg, Herman J. 1964. "Labor and Material Requirements for One-Family Homes." *Monthly Labor Review* 87 (7): 797–800.

Rybczynski, Witold. 2017. "The Pioneering "Levittowner"." *Wharton Real Estate Review,* 73–80.

Scherr, Jonathan. 2024. *Presentation, Harvard Business School.*

Shiller, Robert J. 2015. *Irrational Exuberance.* Princeton University Press.

Silverman, Brian S. 1999. "Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics." *Management science* 45 (8): 1109–1124.

Square Foot Costs with RSMeans Data. 2021, 42nd annual edition (and other years). 2021. Gordian.

Stigler, George J. 1971. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (1): 3–21.

Stokes, H. Kemble. 1981. "An Examination of the Productivity Decline in the Construction Industry." *Review of Economics and Statistics* 63 (4): 495–502.

Sveikauskas, Leo, Samuel Rowe, James Mildenberger, Jennifer Price, and Arthur Young. 2018. *Measuring Productivity Growth in Construction.* Technical report. Monthly Labor Review. U.S. Bureau of Labor Statistics.

———. 2014. *Productivity Growth in Construction.* Technical report 478. BLS Working Papers.

Swan, Craig. 1971. "Labor and Materials Requirements for Housing." *Brookings Papers on Economic Activity* 2.

Turner & Townsend. 2022. *International Construction Market Survey.*

U.S. Bureau of Labor Statistics. 1952. *Construction During Five Decades: Historical Statistics, 1907-1952.* Bulletin of the United States Bureau of Labor Statistics, 1146.

Weiss, Marc A. 1987. *The Rise of Community Builders.* Columbia University Press.

Wildasin, David E, and John Douglas Wilson. 1996. "Imperfect Mobility and Local Government Behaviour in an Overlapping-Generations Model." *Journal of Public Economics* 60 (2): 177–198.

Appendix to "Why Has Construction Productivity Stagnated? The Role of Land-Use Regulation"[1]

Leonardo D'Amico, Edward Glaeser, Joseph Gyourko,
William Kerr, and Giacomo A. M. Ponzetto

November 14, 2024

A House Prices and Costs

A.1 Prices

Figure A1 reports the evolution of prices of houses and cars from 1950 to 2022, where we have indexed all series to 100 in 1960. In red (top line), we use data from decadal censuses and the American Community Survey (ACS) to report a series of new house values controlling for quality and the overall Consumer Price Index (CPI).[2] We correct for quality by estimating a household level regression where the logarithm of CPI-adjusted self-reported home values is regressed on physical attributes. We then report the changing value of houses holding physical characteristics constant.[3] The blue (middle) line reports the CPI for new homes from Shiller (2015).[4] Finally, the green (bottom) line is the CPI for new vehicles collected by the Bureau of Labor Statistics (BLS).

Both housing price series were growing at the same pace as the overall CPI and vehicles

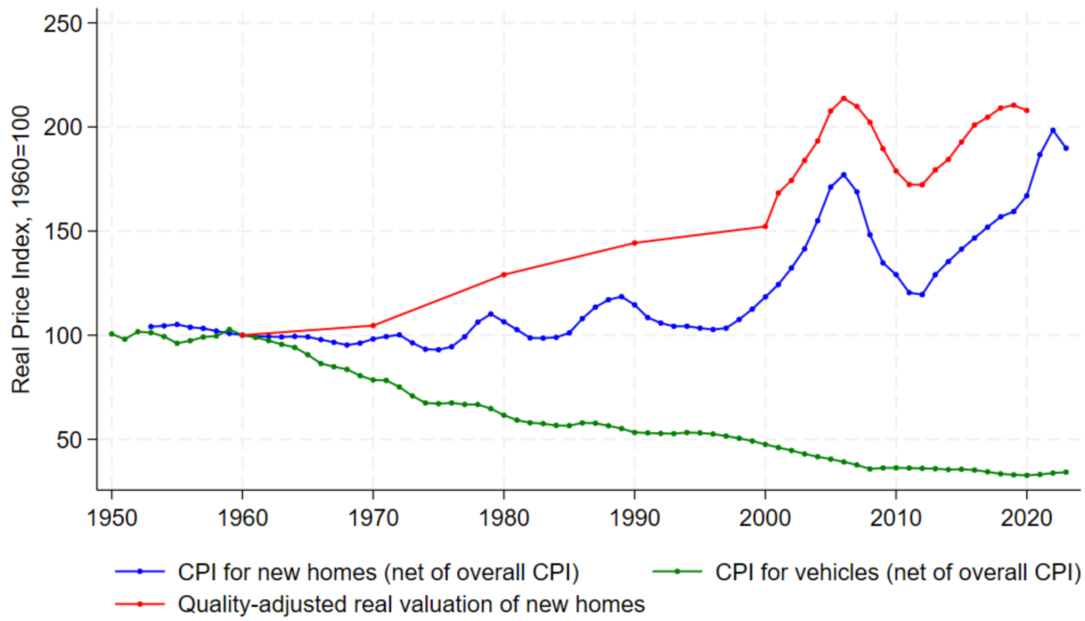
[1]Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

[2]Self-reported value of houses is computed from decadal censuses between 1960 and 2000, and from the ACS 1-yr for 2000 onwards.

[3]The physical characteristics are: the number of rooms in the housing unit, whether the unit has access to a kitchen or to plumbing facilities, and the number of units in the structure. Appendix A.3.1 reports different specifications for this quality adjustment, the associated coefficients, as well as the path of the price as predicted only by the physical characteristics across the different specifications. Results are similar whether or not we put the physical characteristics on their own, or whether we add state-by-year fixed effects or the price of old houses. To be conservative, we chose as baseline the specification that predicts the larger increase in quality over time, which is the one without controls.

[4]As described in Shiller (2015), the 1953 to 1975 data is the home purchase component of the CPI. The Bureau of Labor Statistics collected data on homes constant in age and square footage. The data from 1975 is the Case-Shiller index, which holds quality constant via a repeat sales estimator.

Figure A1: Prices of Cars and Houses, Net of Overall CPI



Note. The figure plots price indexes for homes and cars. Home price indexes are from Census self-valuations (top, red) and Shiller (2015) (middle, blue). The CPI for vehicles (bottom, green) is from the Bureau of Labor Statistics.

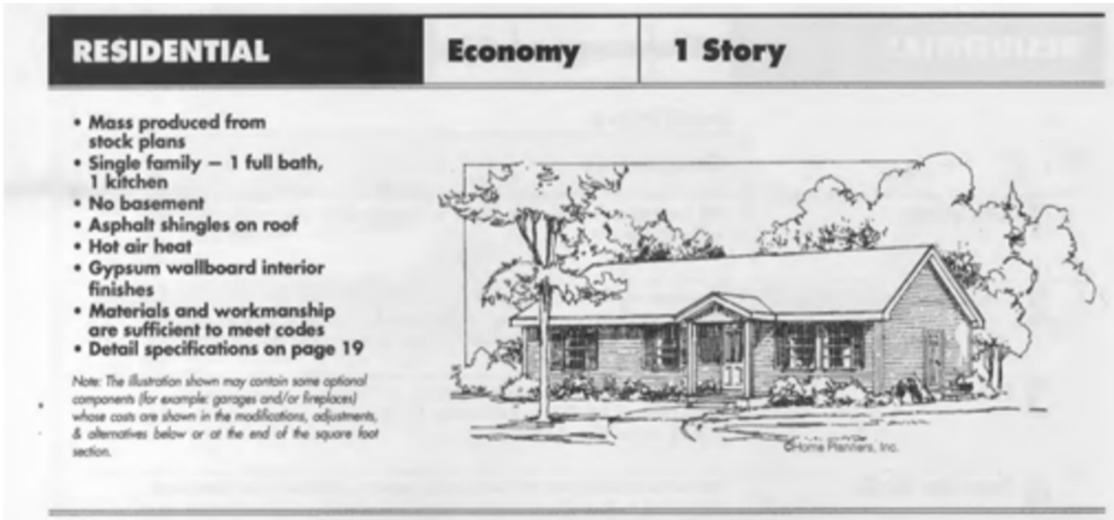
through the 1960s, before diverging substantially in the 1970s. While new homes now cost twice as much as they did in 1960 in real terms, cars are 60% cheaper. This finding mirrors that of Goolsbee and Syverson (2023), who document that the construction output deflator rose much faster than the overall GDP deflator from the 1970s onward.

A.2 Costs

Across space, variation in the costs of producing homes only explains a small fraction of the variation in the costs of buying a home (Gyourko and Saiz 2006). Similarly, much of the rise in housing prices reflects the rising cost of land and the increased difficulty of getting a permit (Glaeser and Gyourko 2018). To focus on construction productivity, we take data from R.S. Means, a private provider of construction cost data, on the real cost of a constant quality 1800 square-foot "economy quality" house.[5] This cost is meant to include only the physical costs of construction,

[5]The R.S. Means Company, now owned by Gordian, publishes annual information on construction costs of different types of structures, but not all books are available. We have information from 24 of those 36 years (1985, 1986, 1989, 1991, 1992, 1996, 2002, 2003, and 2006-2021). Those years are marked with dots in Figure A3. The reference

Figure A2: Depiction of a Constant "Economy Quality" 1,800 Square Foot House



Note. An "Economy Quality" home, used as the basis for the real cost series in Figure A3. From the 2021 R.S. Means company data book.

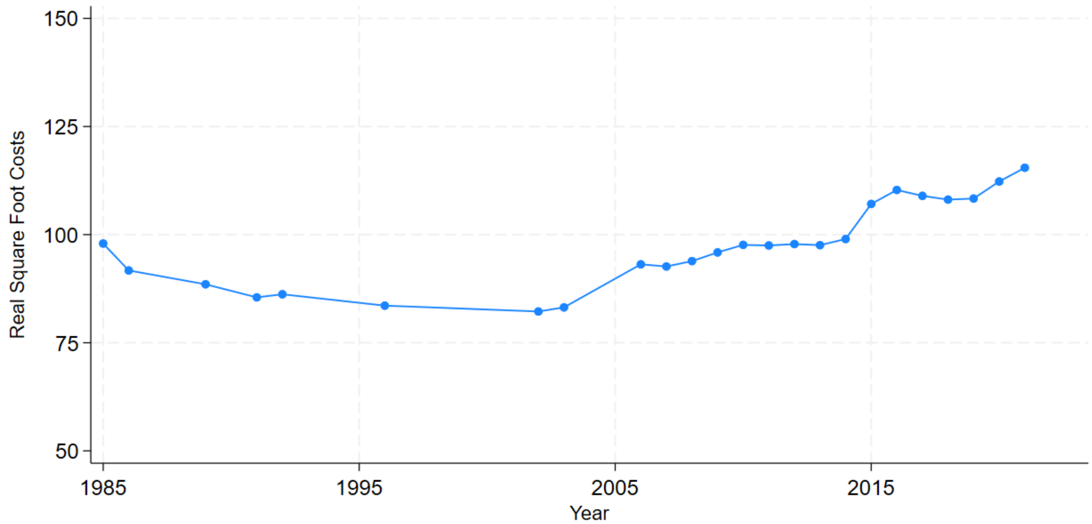
not land acquisition or other costs. R.S. Means defines "economy quality" as a simple, relatively low-cost, one-story home, as depicted in Figure A2, which comes from their 2021 data book. The quality of this home does change, but relatively infrequently and in relatively easy-to-measure ways.[6]

Figure A3 plots the real cost per square foot of supplying that kind of home from 1985 to 2021. Even after accounting for a slight increase in structure quality over time, we find that from 1985 to 2021 the cost of building this modest quality home increased by 18%, or $17.52 per square foot. This fact supports the view that productivity in the homebuilding sector either stagnated or declined, which contrasts sharply with the evidence in Figure A1 that firms have gotten better at producing cars and making them more affordable.

for 2021 is *Square Foot Costs with RSMeans Data. 2021, 42nd annual edition (and other years).* (2021).

[6] Quality is transparently reflected in a set of "traits": whether there is a vapor barrier in the foundation, the number of coats of paint, the type of shingles on the roof, etc. The baseline trait set is constant between 1997 and 2021, and is reported in Appendix Figure A5. Before 1997, however, economy-quality homes had fewer traits. To adjust for this, we reconstruct the cost series before 1997 by keeping the set of traits constant to their 1997-2021 level; which we can do using cost data for each trait, provided by another R.S. Means publication, the *Building Construction Cost Data, 1998, 56th annual edition.* (1998). These costs are inclusive of the materials and the labor cost to apply the trait (e.g. apply an extra coat of paint). Appendix Section A.3.2 provides more details on the procedure and the traits.

Figure A3: Real Cost of a Constant "Economy Quality" 1,800 Square Foot House



Note. The Figure plots the real cost per square foot of supplying a constant "Economy Quality" home, as depicted in Figure A2. The cost comes from our calculations based on R.S. Means company data and only includes the physical costs of construction, not land acquisition or other costs.

A.3 Robustness and Additional Materials

A.3.1 Quality index

In order to construct the quality index, we fit the following model:

$$log(value)_{icst} = \alpha + \lambda_{st} + X'_{icst}\beta + \gamma log(value\_old)_{cst} + \epsilon_{icst}$$

where i,c,s,t stand for individual, county, state and year, respectively. The sample is made by individual respondents that live in a new house, defined as being a house less than (or equal to) 5 years old.

X_{icst} is a vector of physical dwelling characteristics, including the number of rooms, whether the house has plumbing facilities, whether it has a kitchen and which housing structure it belongs to. The term λ_{st} captures state-by-year fixed effects, which we include in two out of four regressions. Finally, $log(value\_old)_{cst}$ is the (log of) the average self-reported value of homes with an estimated age of more than 5 years in the county of the respondent at the time of the response, what we call 'old' houses.

All self-reported values are expressed in 1960 dollars. That is, the dependent variable is the log of self-reported values divided by the CPI normalized to 100 in 1960; and the same normalization is applied to the log of old house prices when we add it as a control.

The next table reports the coefficient of the regressions. Column (1) reports results without controlling for state-by-year fixed effects and price of old houses. Column (2) adds the price of old houses to the controls, while Column (3) adds the state-by-year fixed effects. Column (4) adds both controls.

Table A1: Dwelling Characteristics and the Self-Reported Value of Housing

	(1) Only physical	(2) Physical+old	(3) Physical+FEs	(4) Physical+old+FEs
1-2 rooms (omitted)				
3	0.252***	0.245***	0.259***	0.244***
	(0.00624)	(0.00533)	(0.00538)	(0.00526)
4	0.273***	0.312***	0.322***	0.311***
	(0.00578)	(0.00493)	(0.00499)	(0.00488)
5	0.390***	0.462***	0.472***	0.461***
	(0.00567)	(0.00484)	(0.00490)	(0.00479)
6	0.591***	0.620***	0.638***	0.624***
	(0.00567)	(0.00484)	(0.00490)	(0.00479)
7	0.814***	0.781***	0.812***	0.789***
	(0.00569)	(0.00486)	(0.00491)	(0.00480)
8	1.006***	0.932***	0.978***	0.944***
	(0.00570)	(0.00487)	(0.00493)	(0.00482)
9 or more	1.333***	1.193***	1.263***	1.214***
	(0.00568)	(0.00485)	(0.00492)	(0.00481)
W/out compl. plumb. facilities (omitted)				
With complete plumbing facilities	1.104***	0.763***	0.780***	0.770***
	(0.00553)	(0.00473)	(0.00485)	(0.00475)

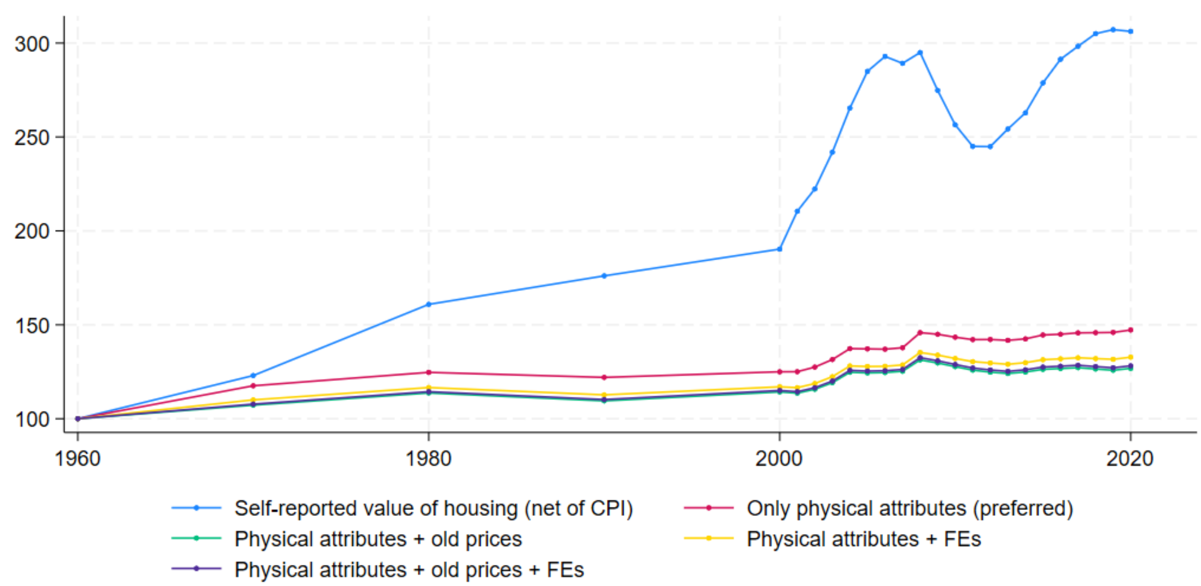
Mobile home or trailer (omitted)				
Boat, tent, van, other	0.395***	0.270***	0.281***	0.277***
	(0.0104)	(0.00892)	(0.00901)	(0.00881)
1-family house, detached	1.103***	1.088***	1.116***	1.099***
	(0.00152)	(0.00130)	(0.00137)	(0.00134)
1-family house, attached	1.308***	1.023***	1.129***	1.051***
	(0.00221)	(0.00191)	(0.00198)	(0.00194)
2-family building	1.559***	1.147***	1.277***	1.163***
	(0.00719)	(0.00615)	(0.00627)	(0.00614)
3-4 family building	1.411***	1.070***	1.188***	1.086***
	(0.00632)	(0.00541)	(0.00548)	(0.00536)
5-9 family building	1.409***	1.043***	1.163***	1.057***
	(0.00604)	(0.00517)	(0.00524)	(0.00513)
10-19 family building	1.446***	1.052***	1.206***	1.087***
	(0.00703)	(0.00601)	(0.00609)	(0.00596)
20-49 family building	1.723***	1.269***	1.450***	1.295***
	(0.00737)	(0.00631)	(0.00639)	(0.00626)
50+ family building	2.029***	1.486***	1.734***	1.531***
	(0.00584)	(0.00501)	(0.00510)	(0.00502)
No kitchen (omitted)				
With kitchen (shared or exclusive use)	-0.0742***	0.0250***	0.0367***	0.0179***
	(0.00693)	(0.00592)	(0.00603)	(0.00589)
Log of old house values (county)		0.742***		0.624***
		(0.000704)		(0.00168)
Constant	7.429***	0.213***	7.634***	1.390***
	(0.00774)	(0.00951)	(0.00672)	(0.0180)
Observations	2993043	2993043	2993043	2993043

Note. The table reports results from a WLS regression at the individual level. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Individual responses are taken from the decadal censuses (1960-2000) and from the ACS 1-yr from 2000 onward. The sample consists of individual responses who claim to be living in new housing units, i.e., housing units built within a 5-year range from the year of the response. Individual responses are weighted by the native household weights.

To compute the quality index, we estimated the predicted values of home according to the coefficients of Table A1 associated to the physical characteristics of the house, i.e. we never predict the quality index using old houses or the fixed effects.

The quality of housing measure used in Figure A1 is the one obtained from column (1) of Table A1. However, the choice of this specification does not affect the pattern of the index and, if anything, only dampens it, as shown in Figure A4.

Figure A4: Alternative Quality Index Specifications



This Figure shows that our preferred specification (the red dotted line) is actually the most optimistic; according to it, quality increased by 47.3% from 1960 to 2020; conversely, in the least optimistic specification, the one controlling only for old house prices, the increase was 26.8%. Controlling for state-by-year fixed effects yields comparable estimates. The yellow line, which only includes fixed effects, depicts an increase in quality by 32.8%; the purple line, which controls both for FEs and the value of old homes, suggests that quality increased by 28.2%. In all cases, the increase in quality as proxied by our index seems insufficient to explain the more than 300% increase in house values.

A.3.2 R.S. Means Traits

As indicated, the baseline trait set is constant from 1997-2021, and traits comprising the home are listed in Figure A5, which is taken from the 2021 book. However, this economy quality home was of lower quality prior to 1997. In 1992 for example, what R.S. Means still labels an economy quality home lacked drip edges, downspouts, a vapor barrier in the foundation, and had one less coat of paint. Structures from earlier years also had inferior roofs that were made of a single-ply roof roll rather than shingles, had a cheaper form of wood siding, contained no aluminum flashing, and had different attic insulation.[7]

To create a constant quality structure, we need to add these traits to the pre-1997 homes. Fortunately, another R.S. Means publication, *Building Construction Cost Data, 1998, 56th annual edition.* (1998), provides per-square-foot costs (including labor and materials) for each of these features of the structure (among many others). The earliest version of this publication available to us is the 1998 book. We use its cost information to 'add' these features to the earlier homes, deflating by the full Urban Workers Consumer Price Index (CPI) to adjust down what each of these costs would have been in an earlier year. We then create an aggregate structure cost based on the full trait set in each year. That nominal cost is then translated into 2021 dollars using the CPI. That real constant quality structure cost is what is plotted in Figure A3.

[7] As shown in Figure A5, the 10 components making up the economy-quality structure can contain multiple individual traits. For example, the Roofing component includes seven traits: 20-year asphalt shingles, #15 felt building paper, aluminum gutters, downspouts, drip edge, and flashings. Each can be priced using the detailed information provided in the Building Construction Cost Data book.

Figure A5: R.S. Means Trait Set Description, Economy Quality, One Story Home

RESIDENTIAL | **Economy** | **Specifications**

	Components
1 Site Work	Site preparation for slab or excavation for lower level; 4' deep trench excavation for foundation wall.
2 Foundations	Continuous reinforced concrete footing, 8" deep x 18" wide; dampproofed and insulated 8" thick reinforced concrete block foundation wall, 4' deep; 4" concrete slab on 4" crushed stone base and polyethylene vapor barrier, trowel finish.
3 Framing	Exterior walls—2 x 4 wood studs, 16" O.C.; 1/2" insulation board sheathing; wood truss roof 24" O.C. or 2 x 6 rafters 16" O.C. with 1/2" plywood sheathing, 4 in 12 or 8 in 12 roof pitch.
4 Exterior Walls	Beveled wood siding and #15 felt building paper on insulated wood frame walls; sliding sash or double hung wood windows; 2 flush solid core wood exterior doors. **Alternates:** • Brick veneer on wood frame, has 4" veneer of common brick. • Stucco on wood frame has 1" thick colored stucco finish. • Painted concrete block has 8" concrete block sealed and painted on exterior and furring on the interior for the drywall.
5 Roofing	20 year asphalt shingles; #15 felt building paper; aluminum gutters, downspouts, drip edge and flashings.
6 Interiors	Walls and ceilings—1/2" taped and finished drywall, primed and painted with 2 coats; painted baseboard and trim; rubber backed carpeting 80%, asphalt tile 20%; hollow core wood interior doors.
7 Specialties	Economy grade kitchen cabinets—6 L.F. wall and base with plastic laminate counter top and kitchen sink; 30 gallon electric water heater.
8 Mechanical	1 lavatory, white, wall hung; 1 water closet, white; 1 bathtub, enameled steel, white; gas fired warm air heat.
9 Electrical	100 Amp. service; romex wiring; incandescent lighting fixtures, switches, receptacles.
10 Overhead and Profit	General Contractor overhead and profit.

Adjustments

Unfinished Basement:
7' high 8" concrete block or cast-in-place concrete walls.

Finished Basement:
Includes inexpensive paneling or drywall on foundation walls. Inexpensive sponge backed carpeting on concrete floor, drywall ceiling, and lighting.

B International Comparisons

We now turn to a comparison of construction costs in US cities and elsewhere. We use data from the construction-consulting firm Turner & Townsend (2022) on building costs for different types of buildings in 31 cities around the world. Like the R.S. Means data, these costs include only the hard costs of physical construction, and exclude land and other costs. This data tells us both the total cost per square meter, as well as the cost of inputs, such as concrete and different types of labor, which allows us to try to separate construction productivity from input costs. We do this by following Langston (2014, 325–26), and aggregating the labor and materials subcomponents to create the cost of standardized baskets of labor and materials.[8] One of the main advantages of this data is that it allows us to look at costs of building offices, whereas the rest of our data mostly concern residential building. Thus, we will focus on commercial office buildings.

We estimate the following regression where Office costs refers to the Turner & Townsend estimate of building an office in the city center with fewer than 20 stories:

$$\text{Office costs}_c = \underset{(.03)}{.171} \times \text{Labor costs}_c + \underset{(.02)}{.085} \times \text{Materials costs}_c + \underset{(.10)}{.366} \times \text{Plant costs}_c + \varepsilon_c \quad (12)$$

The R^2 is 0.86. Table B3 reports the decomposition above for the five US cities in the sample, one "average" city in each continent, and a simple average for the world.[9] The first column reports the raw cost coming from the survey and is followed by the split across the different components as estimated by the regression above, and a residual component based on the same regression. For US cities, we also include a measure of regulatory strictness: the Wharton Residential Land Use Regulation Index (WRLURI), discussed in Gyourko et al. (2019).[10] These data show that US cities

[8] Langston (2014) selected quantities of each input so that they would approximately reflect the typical budget shares of labor (40%), material (50%), and plant costs (10%) for a building of 20 floors or higher at 2012 Sydney prices. See Appendix B.1 for details on the data and on the construction of the bundles.

[9] The average city is defined as the closest to the continent average. The sample does not cover all major cities in the continent; for instance, the African continent is merely represented by the cities of Johannesburg and Nairobi. Predicted costs based on inputs explain quite well the total cost, as shown by Appendix Figure B6; and the predicted and actual cost per square meter have a correlation of 0.93.

[10] The index comes from a survey of over 2,000 primarily suburban jurisdictions, in which the authors ask several questions on the presence of various forms of residential building restrictions that local governments may use (e.g. on the maximum number of permits, zoning laws, density restrictions, etc.). It is standardized, with higher values indicating tighter regulation. See Gyourko et al. (2008) and Gyourko et al. (2019) for a full discussion of the measure. Here we use the 2018 values of the index (Gyourko et al. 2019) since the date of the Turner & Townsend survey is 2022, while in the rest of the paper we use the 2006 value (Gyourko et al. 2008). Either version gives the same relative

Table B3: International Building Costs for an Office Building up to Twenty Floors

	Total Cost	Labor Costs	Material Costs	Plant Costs	Unexplained Residual	WRLURI
	$/m^2	$/m^2	$/m^2	$/m^2	$/m^2	
New York	6,994	3,037	1,500	1,645	813	1.05
San Francisco	6,540	3,100	1,219	1,535	686	1.22
Los Angeles	5,602	2,430	1,232	1,287	653	0.65
Chicago	4,642	1,985	1,275	1,316	67	–0.12
Houston	2,949	1,628	1,275	879	–834	–0.13
Paris	3,107	1,492	983	558	74	
Singapore	2,437	669	1,060	602	106	
Johannesburg	1,006	140	793	313	–241	
São Paulo	751	159	708	546	–662	
World Simple Average	3,004	1,182	1,137	706	–21	

Note. The table reports a decomposition of total building costs across labor costs, material costs, plant costs, and an unexplained residual as described in Equation (12). All costs are expressed in dollars per square meter. The last column adds, for the US cities in the sample, a measure of land-use regulatory tightness from Gyourko et al. (2019).

are expensive places to build offices, and that is particularly true in the most regulated metropolitan areas. However, the residual costs in Chicago and (especially) Houston are not particularly high; these places are expensive because input costs are high in America, even in Texas. Only the highly regulated coastal cities have costs that are much higher than can be explained by the costs of inputs.

We turn to Appendix Section B.1, where we provide more details on the data and replicate this analysis by pooling together the data in the Turner and Townsend reports for the last five years, 2018 to 2023, which shows qualitatively similar patterns.[11] Results are also very similar for offices above 20 floors, which are closer to what Langston (2014) used to define his input bundles. Finally, Appendix Figure B7 abstracts away from the Turner & Townsend bundles and simply presents scatterplots of total raw costs against GDP per capita at the country level, for both offices and apartments, which confirm the results presented here. While GDP per capita explains much of the variation in construction costs, several American cities are notable outliers. Building costs for all American cities except Houston are particularly high.

ranking across cities.

[11] There is no data for 2020, as Turner and Townsend did not publish a yearly report given the pandemic.

B.1 Robustness and Additional Materials

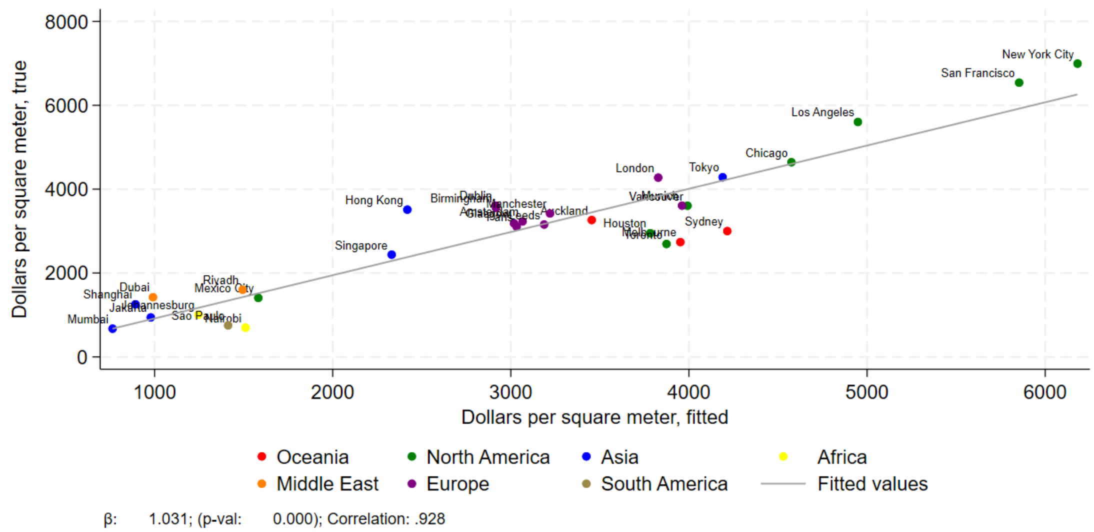
Sample of cities The full list by region includes: Johannesburg, Nairobi, Hong Kong, Jakarta, Mumbai, Shanghai, Singapore, Tokyo, Auckland, Melbourne, Sydney, Amsterdam, Dublin, Munich, Paris, Dubai, Riyadh, Chicago, Houston, Los Angeles, Mexico City, New York City, San Francisco, Toronto, Vancouver, São Paulo, Birmingham, Glasgow, Leeds, London, Manchester.

Definition of building costs. From Turner & Townsend (2022, 11), "building costs per m² [...] are for construction of the building, including preliminaries (or general conditions) costs and substructure, columns, upper floors, staircases, roof, external walls, external doors, internal walls, internal doors, wall finishes, floor finishes, ceiling finishes, fitments, plumbing, HVAC, fire protection, electrical and communication systems and transportation systems." *Exclusions from building costs per m²* are: "External works, landscaping, professional fees, demolition, loose furniture, fittings and equipment, developer's internal costs and finance, local authority fees and headworks charges, land, legal, finance and holding costs, GST or sales taxes, site investigation and test bores, removal of significant obstructions in the ground, abnormal footings. Allowance for underground or onsite car parking is also excluded from the building cost unless stated otherwise."

Labor costs are reported as "the all-inclusive cost to the employer, which includes the basic hourly wage, allowances, taxes, and annual leave costs. Where paid by the employer, this can also include workers' compensation and health insurance, pensions, and travel costs and fares. Labour costs exclude overheads, margins and overtime bonuses."

Definition of inputs baskets The labor basket is an unweighted average of: 150 hours of Group 1 tradesman (e.g. plumber, electrician), 185 hours of Group 2 tradesman (e.g. carpenter, bricklayer), 200 hours of Group 3 tradesman (e.g. carpet layer, tiler, plasterer), 275 hours of general labourer. The materials basket is an unweighted average of: 1,300 m^2 of 13mm plasterboard, 45 m^3 of concrete 30 MPa, 44 m^2 of glass pane 10mm tempered, 2,750 m of softwood timber for framing 100 x 50mm, 6.8 tonnes of structural steel beams. The plant basket is: 5 days of hire of a 50t mobile crane and its operator. We rescale everything by five so to have the plant basket in daily terms.

Figure B6: Total Direct Costs vs. Linear Prediction



These three baskets perform well in explaining costs across cities, as shown in Figure B6, which illustrates the correlation of total direct costs (in dollars per square meter) against the fitted values of regression 12, separating by continent.

Robustness using offices above twenty floors

We replicate our analysis using costs for offices above 20 floors. The linear regression gives:

$$\text{Office costs}_c = \underset{(.05)}{.180} \times \text{Labor costs}_c + \underset{(.02)}{.090} \times \text{Materials costs}_c + \underset{(.24)}{.695} \times \text{Plant costs}_c + \varepsilon_c \quad (13)$$

Table B4 replicates Table B3, for these type of offices:

Table B4: International Building Costs for High-Rise Prestige Offices

	Total Cost	Labor Costs	Material Costs	Plant Costs	Unexplained Residual	WRLURI
	$/m^2	$/m^2	$/m^2	$/m^2	$/m^2	
New York City	9146	4326	1849	1636	1335	1.05
San Francisco	7920	4415	1503	1527	475	1.22
Los Angeles	6602	3461	1519	1280	342	.65
Chicago	5938	2826	1573	1309	230	-.12
Houston	4008	2319	1573	875	-758	-.13
Paris	4039	2125	1212	555	147	
Singapore	2659	952	1308	599	-200	
Johannesburg	1259	200	979	311	-231	
São Paulo	878	227	873	543	-765	
World Simple Average	3762	1683	1403	702	-27	

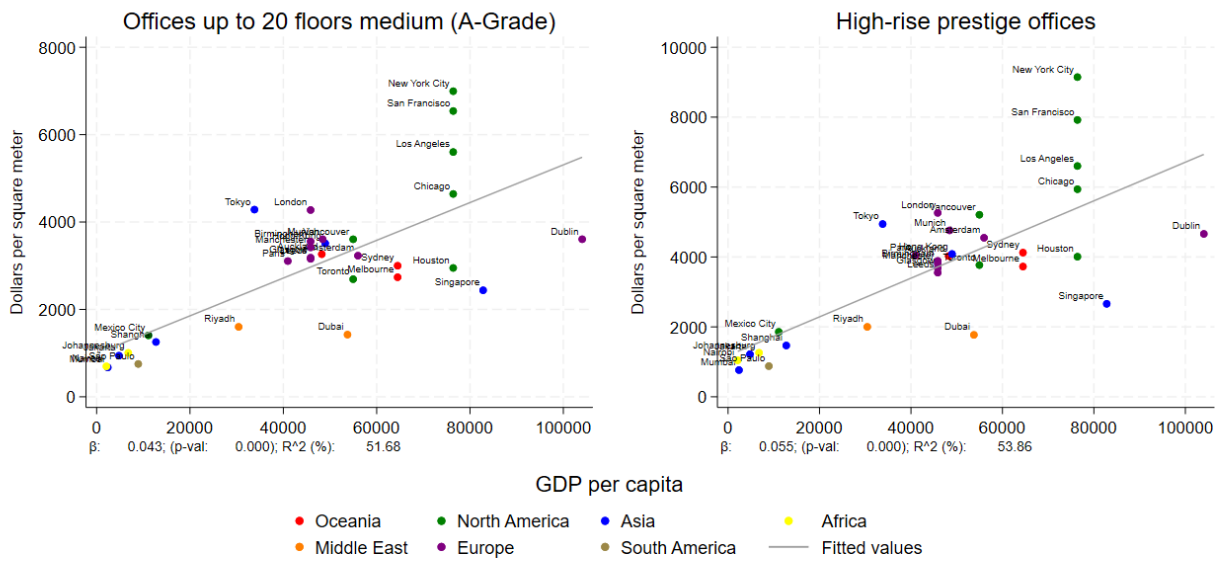
Note. The table reports a decomposition of total building costs across labor costs, material costs, plant costs, and an unexplained residual as described in Equation (13). All costs are expressed in dollars per square meter. The last column adds, for the US cities in the sample, a measure of land-use regulatory tightness from Gyourko et al. (2019).

Costs vs. GDP per capita Figure B7 shows the relationship between cost per square meter and national GDP per capita at the country level.

Robustness using multiple years of data (2018-2023)

We replicate our analysis using multiple Turner & Townsend reports, assembling data from 2018 to 2023, excluding 2020 when Turner & Townsend did not publish a report. We focus on offices below or equal to 20 floors—which is what we show in the main text—but results are qualitatively similar also for offices above 20 floors. The analysis is the same, with the only difference being that costs per square-metre and cost of inputs are now the simple average across years of their values in the reports between 2018 and 2023, rather than their value in 2022 (as in the main text). For several cities, Turner & Townsend do not report data consistently over these

Figure B7: Building Costs and GDP per Capita



years.[12]. For these cities we take the average only across the available years. The linear fit of costs on input bundles for this sample is:

$$\text{Office costs}_c = \underset{(.04)}{.131} \times \text{Labor costs}_c + \underset{(.02)}{.082} \times \text{Materials costs}_c + \underset{(.23)}{.561} \times \text{Plant costs}_c + \varepsilon_c \quad (14)$$

Table B5 replicates Table B3, using multiple years of data.

[12]The cities of Chicago, Mexico City, Riyadh and Vancouver are not reported before 2019. The cities of Birmingham, Glasgow, Leeds, Los Angeles, Manchester and Mumbai are not reported before 2021. The cities of Johannesburg, Nairobi, Hong Kong, Jakarta, Shanghai, Singapore, Tokyo, Melbourne, Sydney, Amsterdam, Dublin, Munich, Paris, Dubai, Houston, New York City, San Francisco, Toronto, Sao Paulo and London are covered throughout the whole time period

Table B5: International Building Costs for an Office Building up to Twenty Floors (2018–2023)

	Total Cost	Labor Costs	Material Costs	Plant Costs	Unexplained Residual	WRLURI
	$/m^2	$/m^2	$/m^2	$/m^2	$/m^2	
New York City	6293	2244	1208	2340	501	1.05
San Francisco	5978	2153	1201	2065	559	1.22
Los Angeles	5511	1822	1291	1950	448	.65
Chicago	4696	1518	1135	1984	59	-.12
Houston	2836	1722	1039	1342	-1267	-.13
Paris	3025	1028	850	849	298	
Singapore	2229	443	943	911	-69	
Johannesburg	932	94	824	455	-441	
São Paulo	971	160	608	814	-611	
World Simple Average	2842	903	978	988	-27	

Note. The table reports a decomposition of total building costs across labor costs, material costs, plant costs, and an unexplained residual as described by Equation (14). All costs are expressed in dollars per square meter. The data averages across years the costs reported in each Turner & Townsend report from 2018 to 2023. The last column adds, for the US cities in the sample, a measure of land-use regulatory tightness from Gyourko et al. (2019).

C Appendix to Section 2

C.1 Pasting Overlapping Series

Figure 1 of section 2.2 shows the evolution of housing units per employee over the course of the past century by juxtaposing data from four different data sources. In particular, as described in the notes to Figure 1, data on housing starts (the numerator) comes from the Macrohistory database until 1959, and from the Census New Residential Construction program from 1959 onward. Data on total employment in construction (the denominator), comes from the Historical Statistics of the United States until 1945, which we validate with historical BLS publications between 1929 and 1945, and from the BLS's Current Employment Statistics from 1939 onward.

In Figure 2 we 'paste' together all these different sources to have a unique series. In particular, we simply link together individual series if they do not overlap, and we take the simple average of multiple sources for periods where they do.

C.2 Choice of Denominator

As anticipated, a possible concern with the analysis in Figure 1 is that our denominator does not adequately capture the evolution of employment in residential construction (since we are interested only in the evolution, we are not concerned of its absolute level). If employment in residential construction shrunk over time as a fraction of total construction employment, this would lead us to understate the growth in output per employee. To assuage this concern, we build three different versions of the denominator by using a mix of the different data sources reported in Table C6.

Table C6: Data on Employment within Construction

Years	Source	Frequency	Subdivisions
1935-1939	Census of Business	Every 4 years	General building contractors, nonbuilding contractors, specialty trade contractors.
1945-1952	U.S. Bureau of Labor Statistics (1952)	Yearly	General building contractors, nonbuilding contractors, specialty trade contractors.
1955-1959	Statistical Abstract Series (Census) – Construction and Housing	Yearly	Nonbuilding contractors
1960-1967	Statistical Abstract Series (Census) – Construction and Housing	Yearly	General building contractors, nonbuilding contractors, specialty trade contractors.
1967-1987	Census of Construction Industries (Economic census)	Every 5 years	Residential and nonresidential building (general) contractors, operative builders, nonbuilding contractors, specialty trade contractors.
1985-2022	Current Employment Statistics (BLS)	Yearly	Residential building (general) contractors
1990-2022	Current Employment Statistics (BLS)	Yearly	Heavy and civil engineering construction
1976-2022	Current Employment Statistics (BLS)	Yearly	Specialty trade contractors

In our preferred version, we divide new houses built by the number of general contractors engaged in the construction of buildings; which allows us to remove employees engaged in heavy construction (highways, bridges, etc.) as well as specialty contractors (electricians and plumbers, f.e., which are mostly employed in the maintenance of the existing housing stock). This data is available in 1935, 1939, from 1945 to 1952 and from 1960 to 1967 at a yearly frequency; from 1967 to 1987, we have data every five years, and from 1990 the data come again with a yearly frequency.[13] However, this denominator still includes general contractors in non-residential

[13]See Table C6 for the different sources used. In some isolated cases, we may compute the shares of general con-

employment. Fortunately, from 1967 to 1987 at a five-year frequency, and from 1985 onward at a yearly frequency, we can distinguish between residential and non-residential general contractors, and we can thus create a series using only general contractors employed in residential building.[14] This would be our preferred series, if not for the case that it only starts in 1967. Another possible concern is that we are still missing specialty contractors (e.g. plumbers and electricians), that are employed in the construction of new building. To the best of our knowledge, there is no data within specialty contractors to discriminate between those who work on new construction vs. those who work on the maintenance of the existing housing stock, and it is likely that such differentiation does not even exist in practice. Thus, we can at least look at what would happen if we constructed a series that includes all specialty contractors.

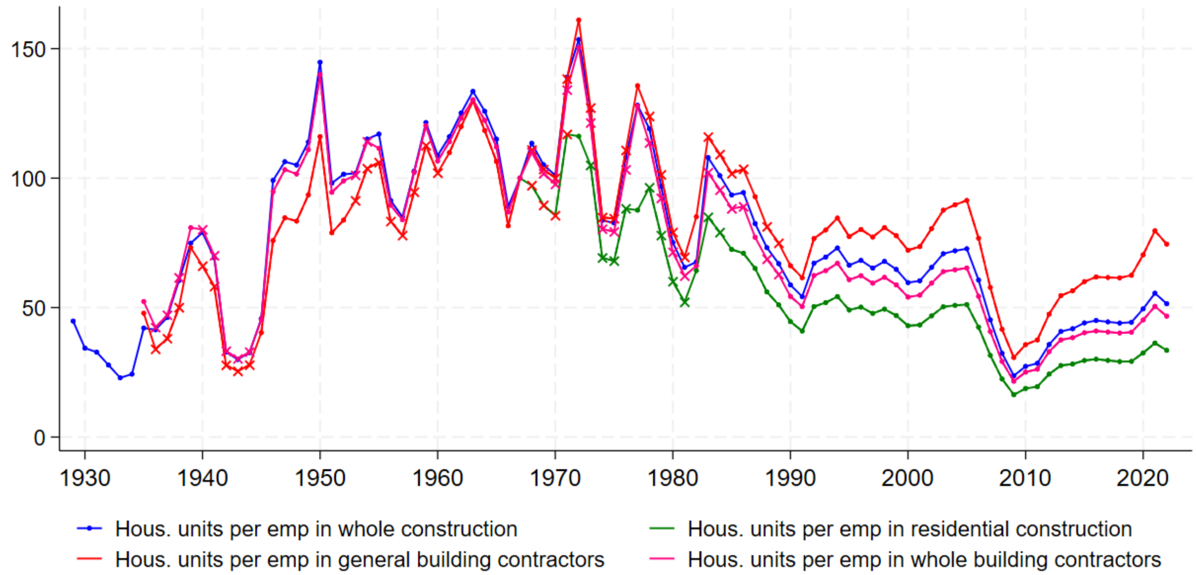
Reassuringly, as shown in Figure C8, all these series show a very similar pattern to the one reported in Figure 1. These similar patterns are due to the fact that share of employment across subsectors within construction moved relatively little over time, as we show in Figures C9a to C9f, which report all the raw employment shares to which we have access. Given the evidence in Figure C8, we feel comfortable in using general contractors as our preferred one because it allows us to go further back in time and because it is the most conservative since, if anything, it suggests a slightly larger productivity gain compared to all other series.

---

tractors employment from two different sources. In 1967, for instance, we both have information from the Statistical Abstract Series and from the Census of Construction Industries. In these cases, we take the simple average of the shares obtained through the different data sources.

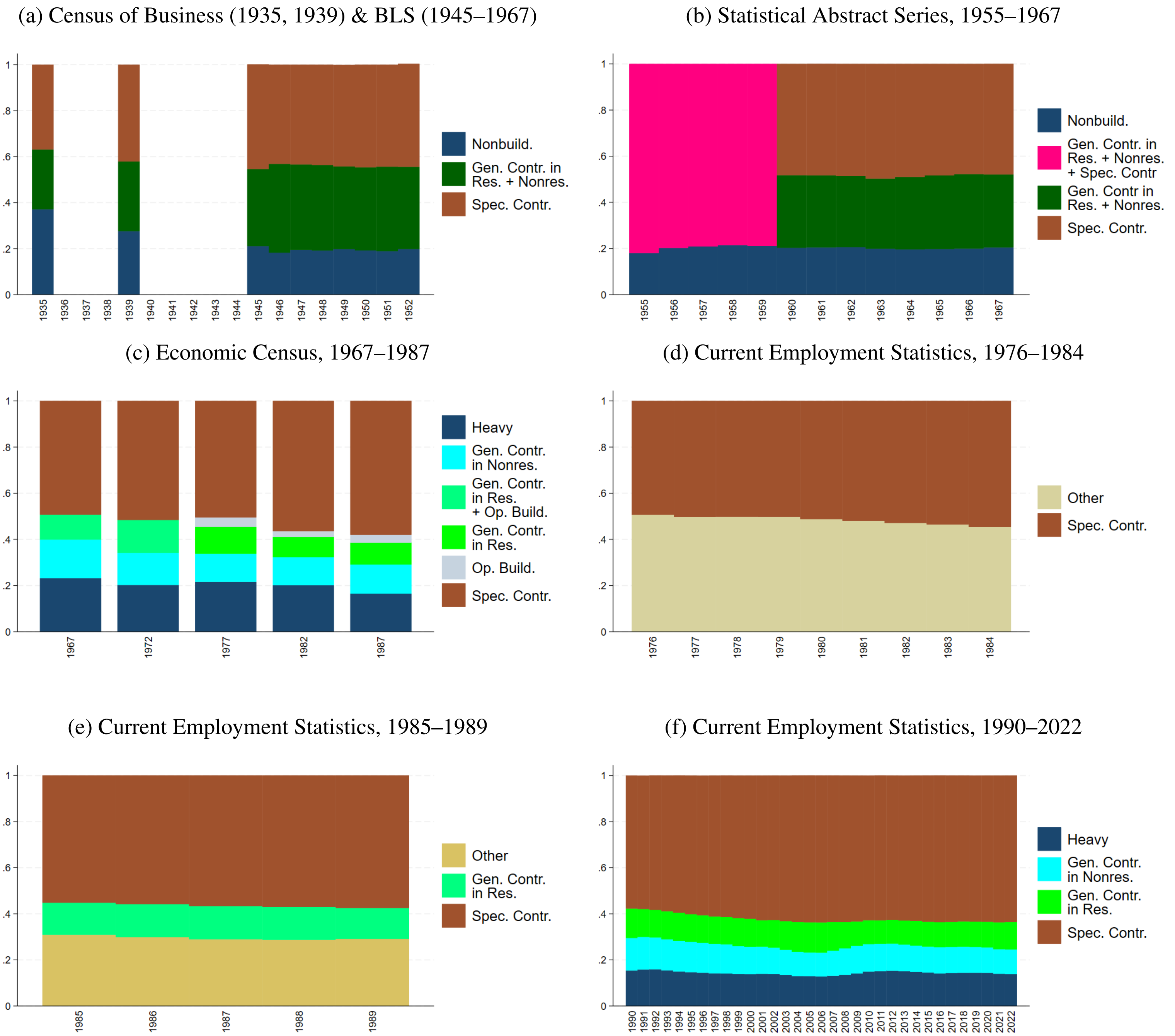
[14]The 1967-1987 Census of Construction Industry data supply a breakdown in residential and nonresidential construction employment. From 1985 to 2022, we rely on the CES series on residential building contractors employment. Employment statistics on non-residential building contractors are only available from 1990 onward.

Figure C8: Housing Units per Employee, Adjusting for Subsector Employment



Note. The circular markers indicate that we take the subsector share from original sources. The cross-shaped markers are used to denote years in which the respective construction employment shares were estimated through an out-of-sample forecast, assuming a linear trend (see Section C.3). All series are indexed to 1967, for which we have original shares in all specifications.

Figure C9: Construction Employment Shares, 1935 to 2022



Note. In the legends: "Nonbuild." stands for *Nonbuilding Construction*; "Gen. Contr." and "Spec. Contr." are short for *General Contractors* and *Specialty Contractors*, respectively; "Res." and "Nonres." stand for *Residential* and *Nonresidential Building Construction*. Finally, "Heavy" abbreviates *Heavy & Civil Engineering Construction*, and *Operative Builders* is abbreviated to "Op. Build."

C.3 Imputing Missing Years

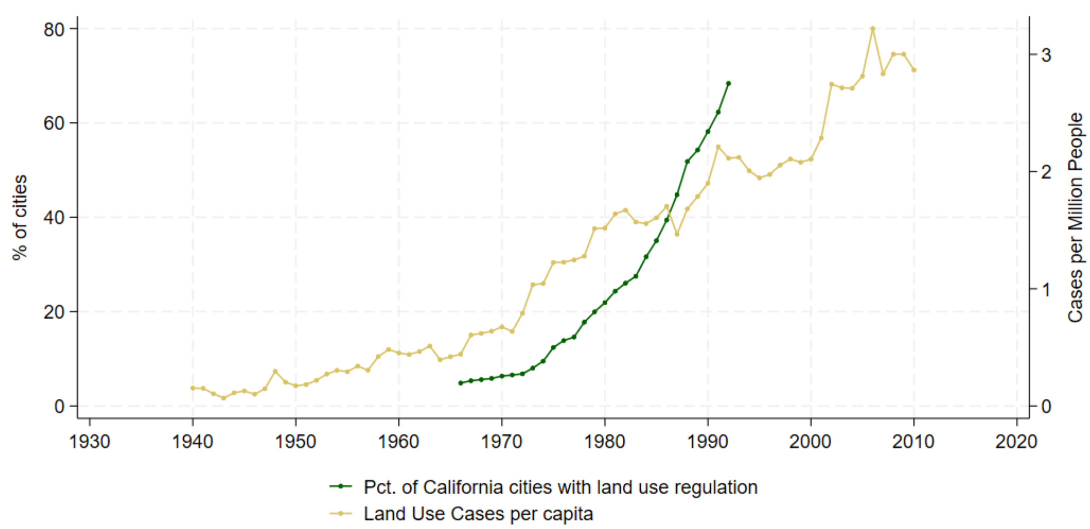
Some of our employment series are not always available at an annual frequency, we assume a linear trend to fill in the series for the years that we miss between the start and the end of each series. In particular, for employment in subsector X, we assume a linear trend in the *share* of construction employment in X, and estimate on the data we have available:

$$\frac{\text{Employment in X}_t}{\text{Emp. in Construction}_t} = \alpha + \beta t + \epsilon_t$$

We then linearly project the subsector share for years in which it is missing and that are between the start and the end of that particular series. This estimated share is then multiplied by overall construction employment (which we always have available) to approximate employment in X. The points that we impute are always indicated by cross-shaped markers.

C.4 Extra Figures

Figure C10: Alternative Measures of Regulation



Note. The dark yellow line plots the number of land use cases per capita from Ganong and Shoag (2017), which we plotted in the main text in Figure 3. Cross-shaped markers are used to denote years in which the denominator in the housing units per employee series was estimated through an out-of-sample forecast (see Appendix C.3 for details).

D Appendix to Section 3

D.1 Proof of Proposition 1

For ease of notation, we omit explicit notation of location c.

D.1.1 Project Size

The earnings-maximizing permit application for a project with productivity $z_{i,j}$ is

$$b\left(z_{i,j}\right) = \arg\max_{b} \left\{ a\left(b\right) \left(pb - \frac{\varepsilon}{z_{i,j}} w^{\lambda_M} b^{1+\frac{1}{\varepsilon}} \right) \right\} = \left[\frac{1-\rho}{1+\varepsilon\left(1-\rho\right)} \frac{p z_{i,j}}{w^{\lambda_M}} \right]^{\varepsilon}, \tag{15}$$

as long as $\underline{b}$ is small enough to ignore corner solutions. The expected value of construction from the project is

$$x\left(z_{i,j}\right) = a\left(b\left(z_{i,j}\right)\right) pb\left(z_{i,j}\right) = p\underline{b}^{\rho} \left[b\left(z_{i,j}\right)\right]^{1-\rho}, \tag{16}$$

its expected input costs are

$$m\left(z_{i,j}\right) = a\left(b\left(z_{i,j}\right)\right) \frac{\varepsilon}{z_{i,j}} w^{\lambda_M} \left[b\left(z_{i,j}\right)\right]^{1+\frac{1}{\varepsilon}} = \frac{\varepsilon\left(1-\rho\right)}{1+\varepsilon\left(1-\rho\right)} x\left(z_{i,j}\right), \tag{17}$$

and its expected earnings for the developer are

$$y\left(z_{i,j}\right) = \max_{b} \left\{ a\left(b\right) \left(pb - \frac{\varepsilon}{z_{i,j}} w^{\lambda_M} b^{1+\frac{1}{\varepsilon}} \right) \right\} - r = \frac{x\left(z_{i,j}\right)}{1+\varepsilon\left(1-\rho\right)} - r. \tag{18}$$

Developers in a given location thus find it profitable to undertake projects so long as their productivity is above the minimum threshold:

$$\underline{z} = \underline{b}^{-\frac{\rho}{\varepsilon(1-\rho)}} \frac{w^{\lambda_M}}{1-\rho} \left[\frac{1+\varepsilon\left(1-\rho\right)}{p} \right]^{\frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)}} r^{\frac{1}{\varepsilon(1-\rho)}}, \tag{19}$$

which corresponds to expected earnings $y\left(\underline{z}\right) = 0$ and an expected value of construction $x\left(\underline{z}\right) = \left[1+\varepsilon\left(1-\rho\right)\right] r$. Assuming that $\underline{b} < \left[1+\varepsilon\left(1-\rho\right)\right] r/p$ is necessary and sufficient for $b\left(\underline{z}\right) > \underline{b}$, which ensures the earnings-maximizing application is never a corner solution.

As regulation (ρ) and prices (p and r) change, the impact on the project threshold is:

$$d\ln \underline{z} = \frac{1}{\varepsilon(1-\rho)} d\ln r - \frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)} d\ln p - \frac{1}{\varepsilon} \ln\left[\frac{b(\underline{z})}{\underline{b}}\right] d\ln(1-\rho). \tag{20}$$

Intuitively, developers are willing to undertake projects with lower productivity when inputs are cheaper, output more valuable, and regulation laxer.

D.1.2 Firm Size

A developer i with potential A_i and technology investment K_i reaches the threshold $\underline{z}$ when undertaking a total amount of projects

$$S_i = \left(\frac{A_i}{\underline{z}} K_i^{\frac{1}{\kappa}}\right)^{\omega}. \tag{21}$$

The expected value of construction in each project is isoelastic in project productivity and sastisfies: $x(z_{i,j})/x(\underline{z}) = (z_{i,j}/\underline{z})^{\varepsilon(1-\rho)}$. As a consequence, the value of buildings built by the developer is

$$X_i = x(\underline{z}) \int_0^{S_i} \left(\frac{z_{i,j}}{\underline{z}}\right)^{\varepsilon(1-\rho)} ds = x(\underline{z}) \int_0^{S_i} \left(\frac{S_i}{s}\right)^{\frac{\varepsilon(1-\rho)}{\omega}} ds = \frac{1+\varepsilon(1-\rho)}{1-\varepsilon(1-\rho)/\omega} rS_i. \tag{22}$$

The developer has input costs

$$M_i = \frac{\varepsilon(1-\rho)}{1+\varepsilon(1-\rho)} X_i = \frac{\varepsilon(1-\rho)}{1-\varepsilon(1-\rho)/\omega} rS_i, \tag{23}$$

and earnings

$$Y_i = \frac{X_i}{1+\varepsilon(1-\rho)} - rS_i = \frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega} rS_i \tag{24}$$

from revenues

$$R_i = X_i - rS_i = \frac{\varepsilon(1-\rho)(1+1/\omega)}{1-\varepsilon(1-\rho)/\omega} rS_i. \tag{25}$$

A developer with potential A_i chooses profit-maximizing technology investment

$$K(A_i) = \arg\max_{K \geq 0} \left\{ \frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega} r \left(\frac{A_i}{\underline{z}} K^{\frac{1}{\kappa}} \right)^{\omega} - w^{\lambda_K} K \right\}$$
$$= \left[\frac{\omega/\kappa}{w^{\lambda_K}} \frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega} r \left(\frac{A_i}{\underline{z}} \right)^{\omega} \right]^{\frac{1}{1-\omega/\kappa}}, \quad (26)$$

which yields operating profits

$$\Pi(A_i) = \max_{K \geq 0} \left\{ \frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega} r \left(\frac{A_i}{\underline{z}} K^{\frac{1}{\kappa}} \right)^{\omega} - w^{\lambda_K} K \right\} = \frac{1-\omega/\kappa}{\omega/\kappa} w^{\lambda_K} K(A_i) \quad (27)$$

with payments to landlords

$$rS(A_i) = r \left(\frac{A_i}{\underline{z}} \right)^{\omega} [K(A_i)]^{\frac{\omega}{\kappa}} = \frac{1-\varepsilon(1-\rho)/\omega}{\varepsilon(1-\rho)/\omega} \frac{\Pi(A_i)}{1-\omega/\kappa}, \quad (28)$$

a value of construction

$$X(A_i) = \frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)/\omega} \frac{\Pi(A_i)}{1-\omega/\kappa}, \quad (29)$$

and input costs, earnings and revenues

$$M(A_i) = \omega \frac{\Pi(A_i)}{1-\omega/\kappa}, \; Y(A_i) = \frac{\Pi(A_i)}{1-\omega/\kappa} \text{ and } R(A_i) = (1+\omega) \frac{\Pi(A_i)}{1-\omega/\kappa}. \quad (30)$$

As a result,

$$\frac{\partial \ln R_i}{\partial \ln A_i} = \frac{\partial \ln X_i}{\partial \ln A_i} = \frac{\partial \ln S_i}{\partial \ln A_i} = \frac{\partial \ln K_i}{\partial \ln A_i} = \frac{\omega}{1-\omega/\kappa} > 0. \quad (31)$$

Taking into account the labor share of each cost, the developer's payroll is $wL_i = \lambda_\Phi w^{\lambda_\Phi} \Phi + \lambda_K w^{\lambda_K} K(A_i) + \lambda_M M(A_i)$, and thus the developer's revenues per employee are

$$P_i \equiv \frac{R_i}{L_i} = w \left[\frac{\lambda_\Phi w^{\lambda_\Phi} \Phi +}{R(A_i)} + \frac{\omega(\lambda_K/\kappa + \lambda_M)}{1+\omega} \right]^{-1}, \quad (32)$$

such that $\partial \ln P_i / \partial \ln A_i > 0$.

D.2 Proof of Propositions 2 and 3

D.2.1 Developer Profitability and Entry

Average investment by developers in a given location equals

$$\bar{K} = \int_{\underline{A}}^{\infty} K(A_i)\frac{\alpha \underline{A}^{\alpha}}{A_i^{1+\alpha}} dA_i = \frac{1-\omega/\kappa}{1-\omega/\kappa-\omega/\alpha} K(\underline{A}) \tag{33}$$

Operating profits, average payments to landowners and the value of construction are all proportional to the cost of investment, at the firm level and therefore on average:

$$\frac{\bar{\Pi}}{w^{\lambda_K}\bar{K}} = \frac{1-\omega/\kappa}{\omega/\kappa}, \ \frac{r\bar{S}}{\bar{\Pi}} = \frac{1-\varepsilon(1-\rho)/\omega}{\varepsilon(1-\rho)/\omega}\frac{1}{1-\omega/\kappa} \text{ and } \frac{\bar{X}}{\bar{\Pi}} = \frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)/\omega}\frac{1}{1-\omega/\kappa}. \tag{34}$$

which corresponds to expected earnings $y(\underline{z}) = 0$ and an expected value of construction $x(\underline{z}) = [1+\varepsilon(1-\rho)]\, r$.

Average developer revenues per employee are

$$\bar{P} = w\int_{\underline{A}}^{\infty}\left[\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{R(A_i)} + \frac{\omega(\lambda_K/\kappa+\lambda_M)}{1+\omega}\right]^{-1}\frac{\alpha\underline{A}^{\alpha}}{A_i^{1+\alpha}}dA_i, \tag{35}$$

and each developer's revenues are proportional to average operating profits

$$R(A_i) = \bar{\Pi}\frac{1+\omega}{1-\omega/\kappa}\frac{1-\omega/\kappa-\omega/\alpha}{1-\omega/\kappa}\left(\frac{A_i}{\underline{A}}\right)^{\frac{\omega}{1-\omega/\kappa}}. \tag{36}$$

As regulation and prices change, the ensuing change in developers' average operating profits is

$$\begin{aligned} d\ln\bar{\Pi} &= \frac{1}{1-\omega/\kappa}\left[\frac{d\ln(1-\rho)}{1-\varepsilon(1-\rho)/\omega} + d\ln r - \omega d\ln\underline{z}\right] \\ &= \bar{\Pi}_p d\ln p - \bar{\Pi}_r d\ln r + \bar{\Pi}_{1-\rho} d\ln(1-\rho), \end{aligned} \tag{37}$$

denoting for ease of notation

$$\bar{\Pi}_p = \frac{1}{1-\omega/\kappa}\frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)/\omega} > 1, \tag{38}$$

$$\bar{\Pi}_r = \frac{1}{1-\omega/\kappa}\frac{1-\varepsilon(1-\rho)/\omega}{\varepsilon(1-\rho)/\omega} \in \left(0, \bar{\Pi}_p - 1\right) \tag{39}$$

and

$$\bar{\Pi}_{1-\rho} = \frac{1}{1-\omega/\kappa}\left[\frac{1}{1-\varepsilon(1-\rho)/\omega} + \frac{\omega}{\varepsilon}\ln\frac{b(\underline{z})}{\underline{b}}\right] > 1. \tag{40}$$

As the price of output rises, or as regulation becomes more permissive, developers expand on three margins: they invest more in technology, undertake more projects, and make each project bigger. Each of these forces alone raises the elasticity of profits to house prices and to regulatory permissivenes $(1-\rho)$ above one. The price of land naturally reduces profits, but the elasticity can be on either side of unity and is certainly lower than the elasticity of profits to house prices. Again the direct effect is amplified by declines in technology investment and in the number of projects, but for the price of land it is also dampened because productivity increases as developers endogenous reduce their span of control.

Developers earn average income $\bar{\Pi} - w^{\lambda_\Phi}\Phi$, which must be positive in equilibrium. Otherwise, no developers would operate and consequently the land rent would be nil; but $\lim_{r\to 0}\bar{\Pi} = \infty$, contradicting the premise. For ease of notation, denote:

$$\tilde{\Phi} \equiv \frac{w^{\lambda_\Phi}\Phi}{\bar{\Pi} - w^{\lambda_\Phi}\Phi} > 0. \tag{41}$$

As regulation (ρ and Φ) and prices (p and r) change, the change in the number of developers is

$$d\ln D = \mu\left[\left(1+\tilde{\Phi}\right)d\ln\bar{\Pi} - \tilde{\Phi}d\ln\Phi - \delta d\ln p\right]. \tag{42}$$

A higher entry cost not only hurts developers' net income $(\bar{\Pi} - w^{\lambda_\Phi}\Phi)$ but also makes it more sensitive to average operating profits.

D.2.2 Land Market Equilibrium

Landowners make their parcels available for development so long as their cost of site preparation w^{λ_T}/τ_i is lower than their expected return r. As long as $\underline{\tau}$ is small enough to ignore corner solutions ($\underline{\tau} < w^{\lambda_T}/r$), the supply of parcels is $T\left(r\underline{\tau}/w^{\lambda_T}\right)^{\sigma}$. Equilibrium in the market for land equates landowners' supply of parcels and developers' demand for parcels $D\bar{S}$, or in terms of value:

$$T\left(\frac{\underline{\tau}}{w^{\lambda_T}}\right)^{\sigma} r^{1+\sigma} = \frac{1-\varepsilon(1-\rho)/\omega}{\varepsilon(1-\rho)/\omega}\frac{\bar{\Pi}D}{1-\omega/\kappa}. \tag{43}$$

As regulation (ρ and Φ) and prices (p and r) change, equilibrium in the land market requires:

$$(1+\sigma)\,d\ln r = d\ln\bar{\Pi} + d\ln D - \frac{1}{1-\varepsilon(1-\rho)/\omega}d\ln(1-\rho), \tag{44}$$

which we can write as

$$\Lambda_r d\ln r = \Lambda_p d\ln p + \Lambda_{1-\rho}\ln(1-\rho) - \Lambda_\Phi d\Phi. \tag{45}$$

As the price of housing rises, each developer demands a greater amount of land, and more developers enter the market, further raising demand for land:

$$\Lambda_p \equiv \bar{\Pi}_p + \mu\left[\left(1+\tilde{\Phi}\right)\bar{\Pi}_p - \delta\right] > \bar{\Pi}_p > 1. \tag{46}$$

As the price of land rises, landowners supply more land, while each developer demands a lower amount and some developers leave the market:

$$\Lambda_r \equiv 1+\sigma+\bar{\Pi}_r + \mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_r > 1. \tag{47}$$

As project regulation becomes laxer, output per developer expands so much each developer demands more land, while more developers enter the market:

$$\Lambda_{1-\rho} \equiv \bar{\Pi}_{1-\rho} - \frac{1}{1-\varepsilon(1-\rho)/\omega} + \mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_{1-\rho} > 0, \tag{48}$$

where the sign of the individual demand response is unambiguous:

$$\bar{\Pi}_{1-\rho} - \frac{1}{1-\varepsilon(1-\rho)/\omega} = \frac{\omega/\kappa}{1-\omega/\kappa}\left[\frac{1}{1-\varepsilon(1-\rho)/\omega} + \frac{\kappa}{\varepsilon}\ln\frac{b(\underline{z})}{\underline{b}}\right] > 0. \tag{49}$$

As the cost of entry rises, some developers leave the market:

$$\Lambda_\Phi \equiv \mu\tilde{\Phi} > 0. \tag{50}$$

D.2.3 Housing Market Equilibrium

Equilibrium in the market for housing equates demand H with supply $D\bar{X}$, namely:

$$\delta\left[wL+\left(\bar{\Pi}-w^{\lambda_\Phi}\Phi\right)D\right]=\frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)/\omega}\frac{\bar{\Pi}D}{1-\omega/\kappa}. \tag{51}$$

As a result, developers' share of housing demand equals:

$$\eta_D\equiv\frac{(\bar{\Pi}-\Phi)D}{wL+(\bar{\Pi}-\Phi)D}=\frac{\delta}{1+\tilde{\Phi}}\frac{(1-\omega/\kappa)\varepsilon(1-\rho)/\omega}{1+\varepsilon(1-\rho)}. \tag{52}$$

As regulation (ρ and Φ) and prices (p and r) change, equilibrium in the housing market requires:

$$\begin{aligned}(1-\eta_D)\,d\ln L+\eta_D\left[\left(1+\tilde{\Phi}\right)d\ln\bar{\Pi}-\tilde{\Phi}d\ln\Phi+d\ln D\right]&\\ =d\ln\bar{\Pi}+d\ln D-\frac{1}{1+\varepsilon(1-\rho)}d\ln(1-\rho),&\end{aligned} \tag{53}$$

which we can write as:

$$\Sigma_r d\ln r=\Sigma_p d\ln p+\Sigma_{1-\rho}d\ln(1-\rho)-\Sigma_\Phi d\ln\Phi. \tag{54}$$

As the price of housing rises, each developer supplies a greater amount, net of own consumption; and some developers enter the market, further increasing housing supply; and some workers leave the market, reducing their demand:

$$\Sigma_p\equiv\bar{\Pi}_p\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right]+\mu\left[\left(1+\tilde{\Phi}\right)\bar{\Pi}_p-\delta\right](1-\eta_D)+(1-\eta_D)\,\mu\delta>0. \tag{55}$$

As the price of land rises, each developer supplies a smaller amount of housing, net of own consumption; and some developers leave the market, further reducing housing supply:

$$\Sigma_r\equiv\bar{\Pi}_r\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right]+\mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_r(1-\eta_D)>0. \tag{56}$$

As project regulation becomes laxer, each developer supplies more housing, net of own consumption; and some developers enter the market, further increasing housing supply:

$$\Sigma_{1-\rho} \equiv \bar{\Pi}_{1-\rho}\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right] - \frac{1}{1+\varepsilon(1-\rho)} + \mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_{1-\rho}\left(1-\eta_D\right) > 0. \tag{57}$$

As the cost of entry rises, some developers leave the market, reducing housing supply; but each developer's net income falls, reducing their housing consumption. The net effect is ambiguous:

$$\Sigma_{\Phi} \equiv \tilde{\Phi}\left[\mu\left(1-\eta_D\right)-\eta_D\right]. \tag{58}$$

All other coefficients are unambiguously signed, because in equilibrium $\left(1+\tilde{\Phi}\right)\eta_D < \delta < 1 < \left(1+\tilde{\Phi}\right)\bar{\Pi}_p$, while the impact of laxer regulation on each developer's housing supply net of own consumption equals:

$$\begin{aligned}\bar{\Pi}_{1-\rho}\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right] - \frac{1}{1+\varepsilon(1-\rho)} &= \left[\frac{1}{1-\omega/\kappa} - \delta\frac{\varepsilon(1-\rho)/\omega}{1+\varepsilon(1-\rho)}\right]\frac{\omega}{\varepsilon}\ln\frac{b\left(\underline{z}\right)}{\underline{b}} \\ &+ \frac{1}{1-\varepsilon(1-\rho)/\omega}\left[\frac{\omega/\kappa}{1-\omega/\kappa} + \frac{\varepsilon(1-\rho)}{1+\varepsilon(1-\rho)} + (1-\delta)\frac{\varepsilon(1-\rho)/\omega}{1+\varepsilon(1-\rho)}\right] > 0. \end{aligned}\tag{59}$$

D.2.4 Comparative Statics for House Prices

The equilibrium impact of regulatory parameters on house prices equals:

$$d\ln p = -\frac{\Lambda_r\Sigma_{1-\rho} - \Lambda_{1-\rho}\Sigma_r}{\Lambda_r\Sigma_p - \Lambda_p\Sigma_r} d\ln(1-\rho) + \frac{\Lambda_r\Sigma_{\Phi} - \Lambda_{\Phi}\Sigma_r}{\Lambda_r\Sigma_p - \Lambda_p\Sigma_r} d\ln\Phi, \tag{60}$$

where the denominator is unambiguously positive:

$$\Lambda_r\Sigma_p - \Lambda_p\Sigma_r = \left[(1+\sigma)\bar{\Pi}_p + \mu\delta\bar{\Pi}_r\right]\left[1-\left(1+\tilde{\Phi}\right)\eta_D + \mu\left(1+\tilde{\Phi}\right)\left(1-\eta_D\right)\right] > 0. \tag{61}$$

This inequality also confirms the equilibrium is stable, in the heuristic sense that self-fulfilling spirals of price changes are impossible. If developers expected $\mathbb{E}\left(d\ln p\right) > 0$ they would bid up land rents by $d\ln r = \mathbb{E}\left(d\ln p\right)\Lambda_p/\Lambda_r$, but the realized change in house prices would then be only $d\ln p = \mathbb{E}\left(d\ln p\right)\left(\Lambda_p/\Lambda_r\right)\left(\Sigma_r/\Sigma_p\right) < \mathbb{E}\left(d\ln p\right)$.

Tighter project regulation increases house prices ($d\ln p/d\ln(1-\rho)<0$) because

$$\begin{aligned}\Lambda_r\Sigma_{1-\rho}-\Lambda_{1-\rho}\Sigma_r &= (1+\sigma)\left\{\bar{\Pi}_{1-\rho}\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right]-\frac{1}{1+\varepsilon(1-\rho)}\right\}\\ &+(1+\sigma)\mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_{1-\rho}(1-\eta_D)+\left[\frac{1-\left(1+\tilde{\Phi}\right)\eta_D}{1-\varepsilon(1-\rho)/\omega}-\frac{1}{1+\varepsilon(1-\rho)}\right]\bar{\Pi}_r\\ &+\mu\left(1+\tilde{\Phi}\right)\left[\frac{1-\eta_D}{1-\varepsilon(1-\rho)/\omega}-\frac{1}{1+\varepsilon(1-\rho)}\right]\bar{\Pi}_r>0,\end{aligned} \tag{62}$$

whose sign is unambiguous because on the first line:

$$\begin{aligned}\bar{\Pi}_{1-\rho}\left[1-\left(1+\tilde{\Phi}\right)\eta_D\right]-\frac{1}{1+\varepsilon(1-\rho)} &= \left[1-\eta_D\left(1+\tilde{\Phi}\right)\right]\frac{\omega/\varepsilon}{1-\omega/\kappa}\ln\frac{b(\underline{z})}{\underline{b}}\\ &+\frac{1}{1-\omega/\kappa}\frac{1}{1-\varepsilon(1-\rho)/\omega}-\left[1+\delta\frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega}\right]\frac{1}{1+\varepsilon(1-\rho)}\\ &>\left[\frac{1}{1-\omega/\kappa}-\frac{1}{1+\varepsilon(1-\rho)}\right]\frac{1}{1-\varepsilon(1-\rho)/\omega}>0,\end{aligned} \tag{63}$$

while on the second and third line:

$$\begin{aligned}\frac{1-\eta_D}{1-\varepsilon(1-\rho)/\omega}-\frac{1}{1+\varepsilon(1-\rho)} &> \frac{1-\left(1+\tilde{\Phi}\right)\eta_D}{1-\varepsilon(1-\rho)/\omega}-\frac{1}{1+\varepsilon(1-\rho)}\\ &=\frac{1}{1-\varepsilon(1-\rho)/\omega}-\frac{1}{1+\varepsilon(1-\rho)}-\delta\frac{1-\omega/\kappa}{1+\varepsilon(1-\rho)}\frac{\varepsilon(1-\rho)/\omega}{1-\varepsilon(1-\rho)/\omega}\\ &>\frac{1+1/\kappa}{1-\varepsilon(1-\rho)/\omega}\frac{\varepsilon(1-\rho)}{1+\varepsilon(1-\rho)}>0.\end{aligned} \tag{64}$$

The impact of tighter entry regulation on house prices is determined by the sign of

$$\frac{\Lambda_r\Sigma_\Phi-\Lambda_\Phi\Sigma_r}{\tilde{\Phi}}=\mu(1+\sigma)-\eta_D(1+\mu)\left(1+\sigma+\bar{\Pi}_r\right). \tag{65}$$

Thus, $d\ln p/d\ln\Phi\geq 0$ if and only if

$$\eta_D\leq\frac{\mu}{1+\mu}\frac{1+\sigma}{1+\sigma+\bar{\Pi}_r}, \tag{66}$$

namely if and only if

$$1+\tilde{\Phi} \geq \frac{1+\mu}{\mu} \frac{\delta}{1+\varepsilon(1-\rho)} \left[\left(1-\frac{\omega}{\kappa}\right) \frac{\varepsilon(1-\rho)}{\omega} + \frac{1-\varepsilon(1-\rho)/\omega}{1+\sigma} \right]. \tag{67}$$

Tighter entry regulation increases house prices if and only if it is above a finite threshold $\check{\Phi}_p$ (which may be nil), because the auxiliary function $\tilde{\Phi}$ is monotone increasing in the exogenous parameter Φ:

$$\frac{d\ln\tilde{\Phi}}{d\ln\Phi} = \left(1+\tilde{\Phi}\right)\left(1-\frac{d\ln\bar{\Pi}}{d\ln\Phi}\right) = \frac{(1+\mu)\left(1+\tilde{\Phi}\right)(1-\eta_D)}{1-\left(1+\tilde{\Phi}\right)\eta_D + \mu\left(1+\tilde{\Phi}\right)(1-\eta_D)} > 0. \tag{68}$$

D.2.5 Comparative Statics for Land Prices

The equilibrium impact of regulatory parameters on land prices equals:

$$d\ln r = -\frac{\Lambda_p\Sigma_{1-\rho} - \Lambda_{1-\rho}\Sigma_p}{\Lambda_r\Sigma_p - \Lambda_p\Sigma_r} d\ln(1-\rho) + \frac{\Lambda_p\Sigma_\Phi - \Lambda_\Phi\Sigma_p}{\Lambda_r\Sigma_p - \Lambda_p\Sigma_r} d\ln\Phi, \tag{69}$$

where the denominator is unambiguously positive, as we proved above.

The impact of tighter project regulation on land prices is determined by the sign of

$$\begin{aligned}\Lambda_p\Sigma_{1-\rho} - \Lambda_{1-\rho}\Sigma_p = &\left[\frac{1-\eta_D\left(1+\tilde{\Phi}\right)}{1-\varepsilon(1-\rho)/\omega} - \frac{1}{1+\varepsilon(1-\rho)}\right]\bar{\Pi}_p \\ &+ \left[\frac{1-\eta_D}{1-\varepsilon(1-\rho)/\omega} - \frac{1}{1+\varepsilon(1-\rho)}\right]\mu\left[\left(1+\tilde{\Phi}\right)\bar{\Pi}_p - \delta\right] + \eta_D\mu\delta\tilde{\Phi}\bar{\Pi}_{1-\rho} \\ &- (1-\eta_D)\mu\delta\left[\bar{\Pi}_{1-\rho} - \frac{1}{1-\varepsilon(1-\rho)/\omega} + \mu\left(1+\tilde{\Phi}\right)\bar{\Pi}_{1-\rho}\right],\end{aligned} \tag{70}$$

which is ambiguous. Intuitively, tighter regulation increases land prices if factor mobility is negligible but reduces them if it is perfect: formally, $\lim_{\mu\delta\to 0} d\ln r/d\ln(1-\rho) < 0 < \lim_{\mu\delta\to 0} d\ln r/d\ln(1-\rho)$.

Tighter entry regulation reduces land prices ($d\ln r/d\ln\Phi < 0$) because

$$\frac{\Lambda_p\Sigma_\Phi - \Lambda_\Phi\Sigma_p}{\tilde{\Phi}} = -\eta_D\left[\bar{\Pi}_p + \mu\left(\bar{\Pi}_p - \delta\right)\right] - \mu(1-\eta_D)\mu\delta < 0. \tag{71}$$

D.2.6 Comparative Statics for Developer Size

Plugging Equation 42 into Equation 53 yields:

$$\left[1-\eta_D\left(1+\tilde{\Phi}\right)\right] d\ln\bar{\Pi} + (1-\eta_D)\, d\ln D - \frac{1}{1+\varepsilon(1-\rho)} d\ln(1-\rho) \\ = (1-\eta_D)\, d\ln L - \eta_D \tilde{\Phi} d\ln\Phi. \quad (72)$$

Plugging in the spatial-equilibrium conditions ($d\ln L = -\mu\delta d\ln p$ and Equation 42):

$$\left[1-\eta_D\left(1+\tilde{\Phi}\right) + \mu\left(1+\tilde{\Phi}\right)(1-\eta_D)\right] d\ln\bar{\Pi} \\ = \frac{1}{1+\varepsilon(1-\rho)} d\ln(1-\rho) + \left[\mu(1-\eta_D) - \eta_D\right] \tilde{\Phi} d\ln\Phi. \quad (73)$$

Tighter project regulation reduces developer profits because

$$\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} = \left[1+\varepsilon(1-\rho)\right]^{-1} \left[1-\eta_D\left(1+\tilde{\Phi}\right) + \mu\left(1+\tilde{\Phi}\right)(1-\eta_D)\right]^{-1} > 0, \quad (74)$$

since developers' equilibrium share of housing demand is $\eta_D < \delta/\left(1+\tilde{\Phi}\right)$ by Equation (52).

The impact of tighter entry regulation on developer profits is determined by the sign of $\mu(1-\eta_D) - \eta_D$. Thus $d\ln\bar{\Pi}/d\ln\Phi \geq 0$ if and only if $\eta_D \leq \mu/(1+\mu)$; namely, plugging in the equilibrium shares of housing demand (Equation 52), if and only if

$$1+\tilde{\Phi} \geq \frac{1+\mu}{\mu} \frac{\delta}{1+\varepsilon(1-\rho)} \left(1-\frac{\omega}{\kappa}\right) \frac{\varepsilon(1-\rho)}{\omega}. \quad (75)$$

Tighter entry regulation increases developer profits if and only if it is above a finite threshold $\check{\Phi}_{\bar{\Pi}} < \check{\Phi}_p$ (which may be nil).

In equilibrium, operating profits and revenues are proportional to the cost of investment, so:

$$\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} = \frac{d\ln\bar{R}}{d\ln(1-\rho)} = \frac{d\ln\bar{K}}{d\ln(1-\rho)} \text{ and } \frac{d\ln\bar{\Pi}}{d\ln\Phi} = \frac{d\ln\bar{R}}{d\ln\Phi} = \frac{d\ln\bar{K}}{d\ln\Phi}. \quad (76)$$

Recall that average developer revenues per employee are:

$$\bar{P} = w(1+\omega)\int_{\underline{A}}^{\infty}\left[\omega\left(\frac{\lambda_K}{\kappa}+\lambda_M\right)+\frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{\underline{A}}{A_i}\right)^{\frac{\omega}{1-\omega/\kappa}}\right]^{-1}\frac{\alpha\underline{A}^{\alpha}}{A_i^{1+\alpha}}dA_i. \tag{77}$$

Thus

$$\begin{aligned}\frac{d\ln\bar{P}}{d\ln(1-\rho)} = {} & \frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\underline{A}^{\frac{\omega}{1-\omega/\kappa}} \\ & \times\frac{\int_{\underline{A}}^{\infty}\left[\omega\left(\frac{\lambda_K}{\kappa}+\lambda_M\right)+\frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{\underline{A}}{A_i}\right)^{\frac{\omega}{1-\omega/\kappa}}\right]^{-2}A_i^{-1-\alpha-\frac{\omega}{1-\omega/\kappa}}dA_i}{\int_{\underline{A}}^{\infty}\left[\omega\left(\frac{\lambda_K}{\kappa}+\lambda_M\right)+\frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{\underline{A}}{A_i}\right)^{\frac{\omega}{1-\omega/\kappa}}\right]^{-1}A_i^{-1-\alpha}dA_i} \\ & \times\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} > 0,\end{aligned} \tag{78}$$

while

$$\begin{aligned}\frac{d\ln\bar{P}}{d\ln\Phi} = {} & \frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\underline{A}^{\frac{\omega}{1-\omega/\kappa}} \\ & \times\frac{\int_{\underline{A}}^{\infty}\left[\omega\left(\frac{\lambda_K}{\kappa}+\lambda_M\right)+\frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{\underline{A}}{A_i}\right)^{\frac{\omega}{1-\omega/\kappa}}\right]^{-2}A_i^{-1-\alpha-\frac{\omega}{1-\omega/\kappa}}dA_i}{\int_{\underline{A}}^{\infty}\left[\omega\left(\frac{\lambda_K}{\kappa}+\lambda_M\right)+\frac{(1-\omega/\kappa)^2}{1-\omega/\kappa-\omega/\alpha}\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{\underline{A}}{A_i}\right)^{\frac{\omega}{1-\omega/\kappa}}\right]^{-1}A_i^{-1-\alpha}dA_i} \\ & \times\frac{\lambda_\Phi w^{\lambda_\Phi}\Phi}{\bar{\Pi}}\left(\frac{d\ln\bar{\Pi}}{d\ln\Phi}-1\right) < 0,\end{aligned} \tag{79}$$

because $d\ln\bar{\Pi}/d\ln\Phi < 1$ as we proved above.

Recall that average developer payments to landowners are:

$$r\bar{S} = \frac{1-\varepsilon(1-\rho)/\omega}{\varepsilon(1-\rho)/\omega}\frac{\bar{\Pi}}{1-\omega/\kappa}, \tag{80}$$

such that $d\ln(r\bar{S})/d\ln\Phi = d\ln\bar{\Pi}/d\ln\Phi$, while

$$\frac{d\ln(r\bar{S})}{d\ln(1-\rho)} = \frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} - \frac{1}{1-\varepsilon(1-\rho)/\omega} \leq 0 \tag{81}$$

because:

$$\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} = \left\{[1+\varepsilon(1-\rho)]\left[1+\mu\left(1+\tilde{\Phi}\right)\right] - \left(1-\frac{\omega}{\kappa}\right)\frac{\varepsilon(1-\rho)}{\omega}(1+\mu)\delta\right\}^{-1}$$
$$\leq \left[1-\frac{\varepsilon(1-\rho)}{\omega}+\varepsilon(1-\rho)\frac{1+\kappa}{\kappa}\right]^{-1} \leq \left[1-\frac{\varepsilon(1-\rho)}{\omega}\right]^{-1}. \quad (82)$$

D.2.7 Comparative Statics for Developer Entry

Since tigther project regulation reduces developer profits and increases house prices, a fortiori it reduces the number of developers:

$$\frac{d\ln D}{d\ln(1-\rho)} = \mu\left[\left(1+\tilde{\Phi}\right)\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} - \delta\frac{d\ln p}{d\ln(1-\rho)}\right] > 0 \quad (83)$$

because $d\ln\bar{\Pi}/d\ln(1-\rho) > 0 > d\ln p/d\ln(1-\rho)$.

Tighter entry regulation reduces the number of developers:

$$\frac{d\ln D}{d\ln\Phi} = \mu\left[\left(1+\tilde{\Phi}\right)\frac{d\ln\bar{\Pi}}{d\ln\Phi} - \tilde{\Phi} - \delta\frac{d\ln p}{d\ln\Phi}\right] = -\frac{\mu\tilde{\Phi}}{1-\eta_D\left(1+\tilde{\Phi}\right)+\mu\left(1+\tilde{\Phi}\right)(1-\eta_D)}$$
$$\times\left\{1+\delta\frac{\mu(1+\sigma)-\eta_D(1+\mu)\left(1+\sigma+\bar{\Pi}_r\right)}{(1+\sigma)\bar{\Pi}_p+\mu\delta\bar{\Pi}_r}\right\} < 0, \quad (84)$$

because

$$\eta_D \leq 1 < \frac{1}{1+\mu}\left[\mu+\frac{1}{\delta}\frac{(1+\sigma)\bar{\Pi}_p}{1+\sigma+\bar{\Pi}_r}\right]. \quad (85)$$

D.2.8 Comparative Statics for Aggregate Housing Value

The aggregate value of housing built equals the total demanded:

$$H = \delta\left[wL+\left(\bar{\Pi}-w^{\lambda_\Phi}\Phi\right)D\right], \quad (86)$$

such that

$$d\ln H = (1-\eta_D)\,d\ln L + \eta_D\left[\left(1+\tilde{\Phi}\right)d\ln\bar{\Pi} - \tilde{\Phi}d\ln\Phi + d\ln D\right]. \quad (87)$$

Thus, tighter entry regulation reduces the aggregate value of housing built:

$$\frac{d\ln H}{d\ln(1-\rho)} = \eta_D (1+\mu)\left(1+\tilde{\Phi}\right)\frac{d\ln\bar{\Pi}}{d\ln(1-\rho)} - \mu\delta\frac{d\ln p}{d\ln(1-\rho)} > 0. \tag{88}$$

Since it also increases the price of housing, a fortiori it reduces the aggregate quantity of housing built.

The aggregate value of housing built also equals the total supplied:

$$H = \frac{1+\varepsilon(1-\rho)}{\varepsilon(1-\rho)/\omega}\frac{\bar{\Pi}D}{1-\omega/\kappa}, \tag{89}$$

such that

$$d\ln H = d\ln\bar{\Pi} + d\ln D - \frac{1}{1+\varepsilon(1-\rho)} d\ln(1-\rho). \tag{90}$$

Equilibrium in the land market (Equation 44) then implies that

$$d\ln H = (1+\sigma)\, d\ln r + \left[\frac{1}{1-\varepsilon(1-\rho)/\omega} - \frac{1}{1+\varepsilon(1-\rho)}\right] d\ln(1-\rho), \tag{91}$$

so tighter entry regulation reduces the aggregate value of housing built:

$$\frac{d\ln H}{d\ln\Phi} = (1+\sigma)\frac{d\ln r}{d\ln\Phi} < 0. \tag{92}$$

Tighter entry regulation also reduces the aggregate quantity of housing built:

$$\begin{aligned}\frac{d\ln H}{d\ln\Phi} - \frac{d\ln p}{d\ln\Phi} &= (1+\sigma)\frac{d\ln r}{d\ln\Phi} - \frac{d\ln p}{d\ln\Phi}\\ &= -\frac{\tilde{\Phi}}{\left[(1+\sigma)\bar{\Pi}_p + \mu\delta\bar{\Pi}_r\right]\left[1-\left(1+\tilde{\Phi}\right)\eta_D + \mu\left(1+\tilde{\Phi}\right)(1-\eta_D)\right]}\\ &\quad \times \left\{\left[(1+\sigma)\left(\bar{\Pi}_p - 1\right) - \bar{\Pi}_r\right]\eta_D + \mu\left[(1+\sigma)\left(\bar{\Pi}_p-\delta\right) - \bar{\Pi}_r\right]\eta_D\right.\\ &\qquad \left. + (1+\sigma)\mu(1+\mu\delta)(1-\eta_D)\right\} < 0\end{aligned} \tag{93}$$

because $(1+\sigma)\left(\bar{\Pi}_p-\delta\right) - \bar{\Pi}_r > (1+\sigma)\left(\bar{\Pi}_p - 1\right) - \bar{\Pi}_r > \bar{\Pi}_p - 1 - \bar{\Pi}_r > 0$.

E Appendix to Section 4

E.1 More on Firm Size Distributions

The SUSB dataset gives us information on additional measures, including establishments, firms[15] and annual payrolls.

Figure E11: Firms, Establishments and Payrolls by Employment Size: New Single-Family vs. Comparison

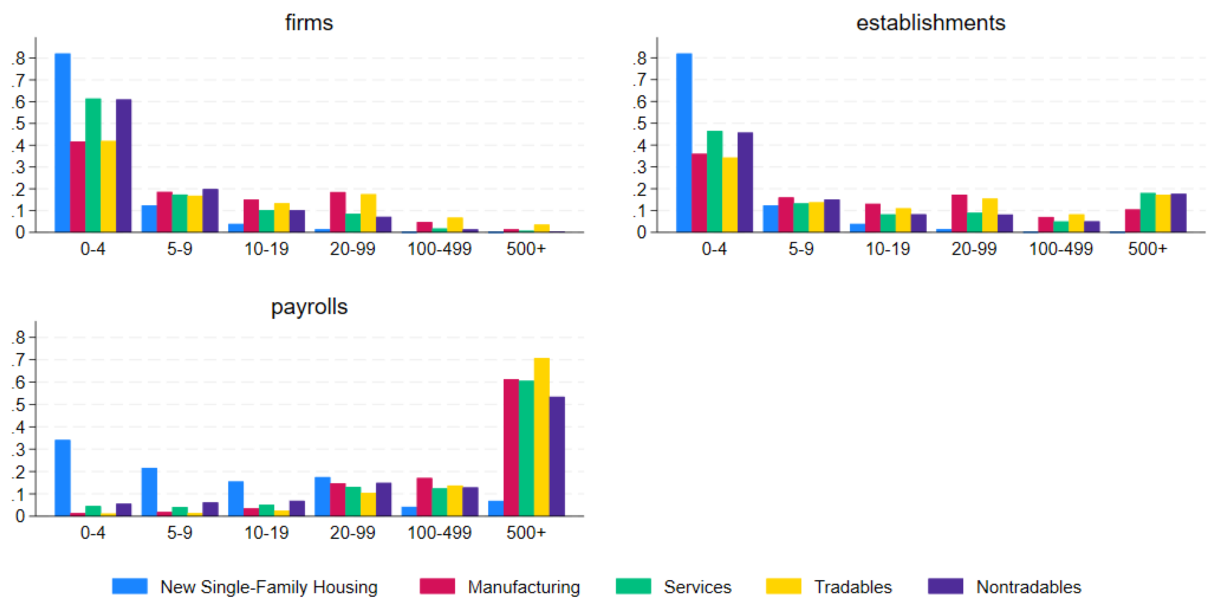


Figure E11 replicates Figures 5a-5b, but for these additional measures. Unsurprisingly, in terms of firms and establishments, the $0-4$ category is the one with the highest share for all industries under examination. However, firms with more than 100 employees are a negligible share for single-family housing, but a nonnegligible one for the other industries. This is in line with the idea that the market for single-family housing is populated by small providers, and is not challenged by any industrial giant; whereas in the other sectors a considerable number of very large firms.

The evidence on payrolls is in line with that on receipts. The bulk of payrolls paid in new

[15] Firms are either single establishments or an agglomerate of establishment identified as a single entity

single-family housing firms is represented by firms with less than 10 employees. This is opposite to comparison industries, where large firms account for the striking majority of payrolls.

Figure E12: Employment and Receipts Share by Employment Size in New Multifamily Construction vs. Comparison

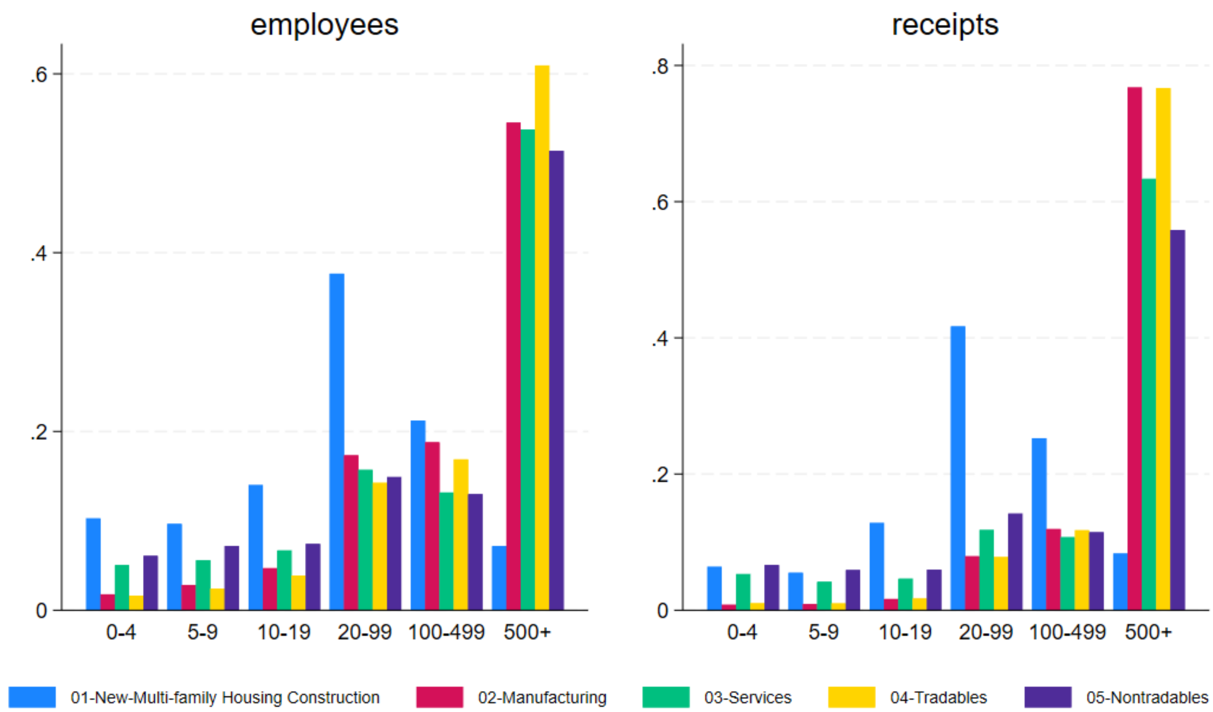


Figure E12 repeats the results in Section 4 for multi-family housing construction firms. Though smaller than the comparison sectors, firms engaged in multi-family housing are larger than those active in the single-family business. In particular, the most abundant employment bracket in the multi-family construction sector is the $20-99$ employees (in the single-family, it was the $0-4$ bracket).

Figure E13: Firms, Establishments and Payroll by Employment Size: New Multifamily vs. Comparison

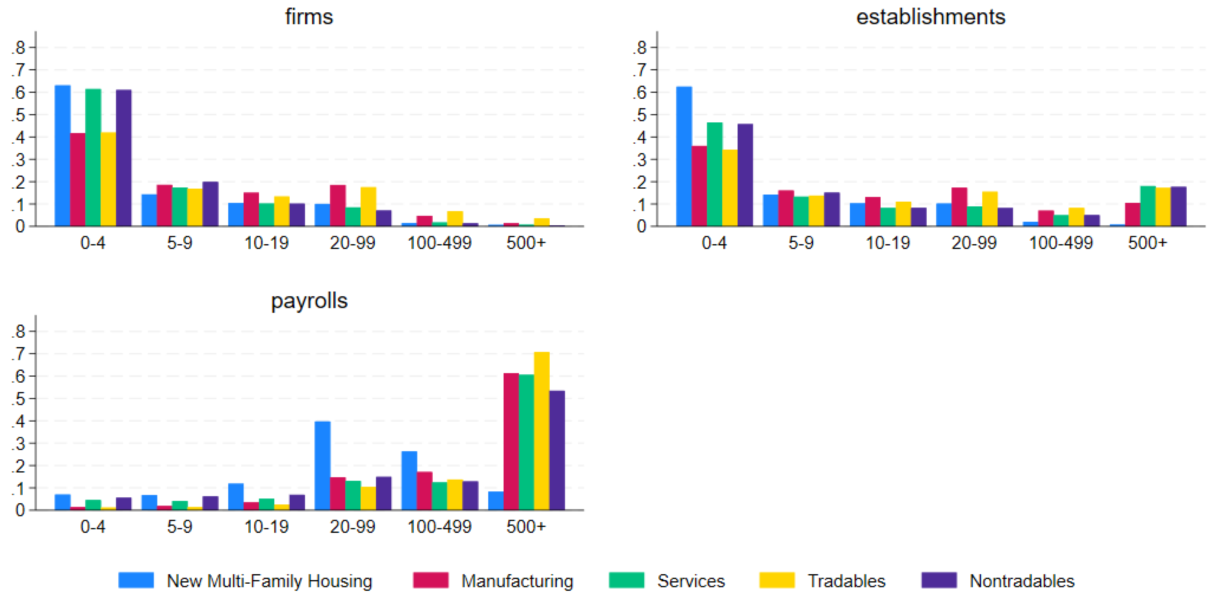
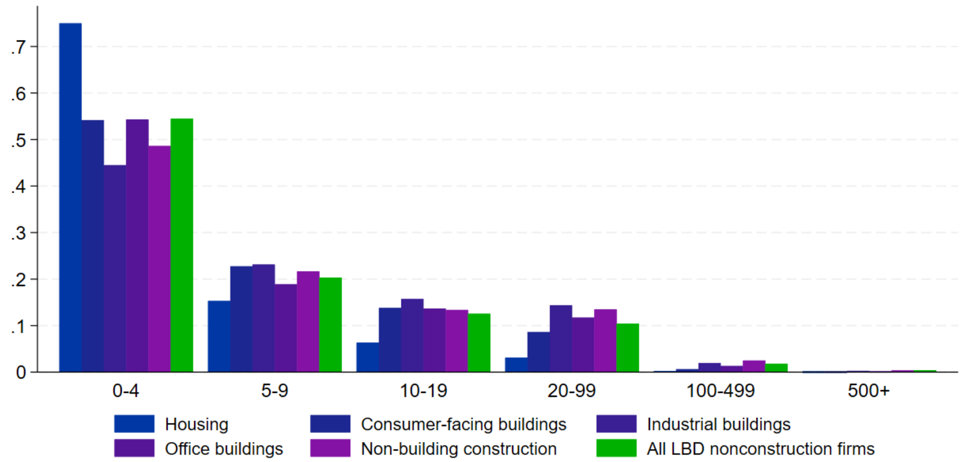


Figure E13 gives information on the distribution of firms, establishments and payrolls in the new multifamily construction sector against comparison industries.

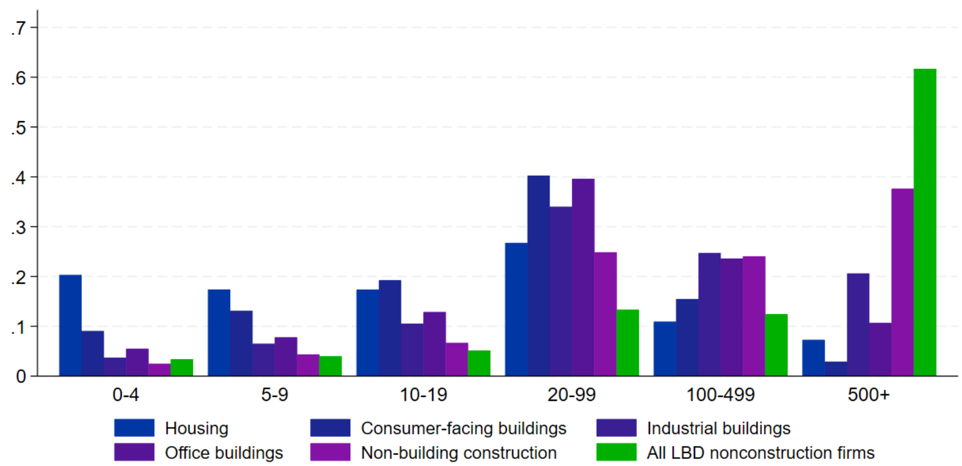
The remainder of this appendix Section reports additional figures using the CCI sample. Figure E14 mirrors the plots in the main text, but focuses on firm counts and payroll. Figure E15 reports results for all types of builders available.

Figure E14: Firm Size Distributions, Firm Counts and Payroll Shares

(a) Firm Counts

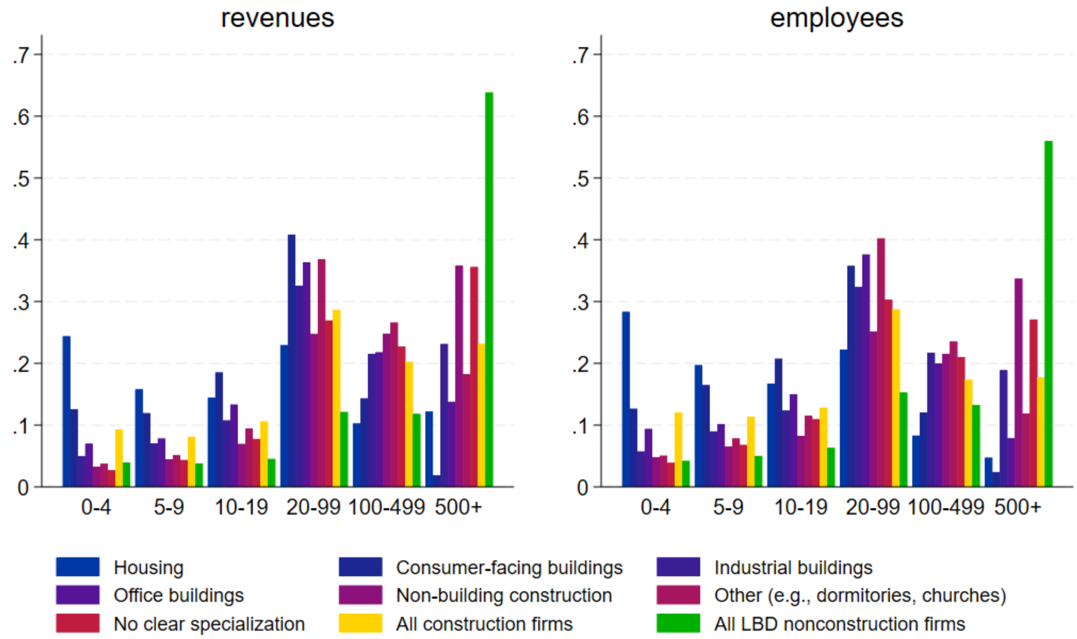


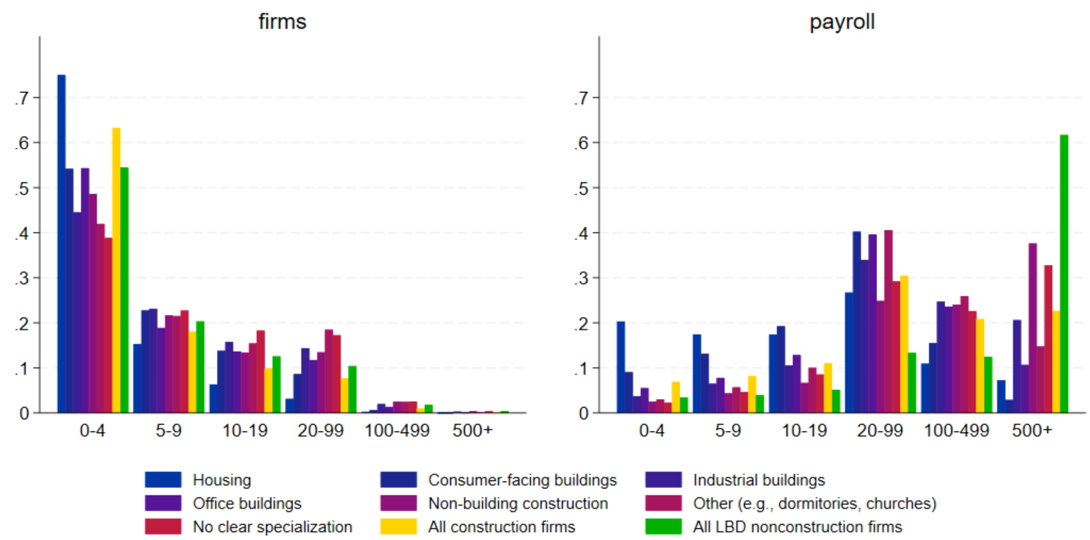
(b) Payroll



Note. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

Figure E15: Distribution in All Available Subsectors

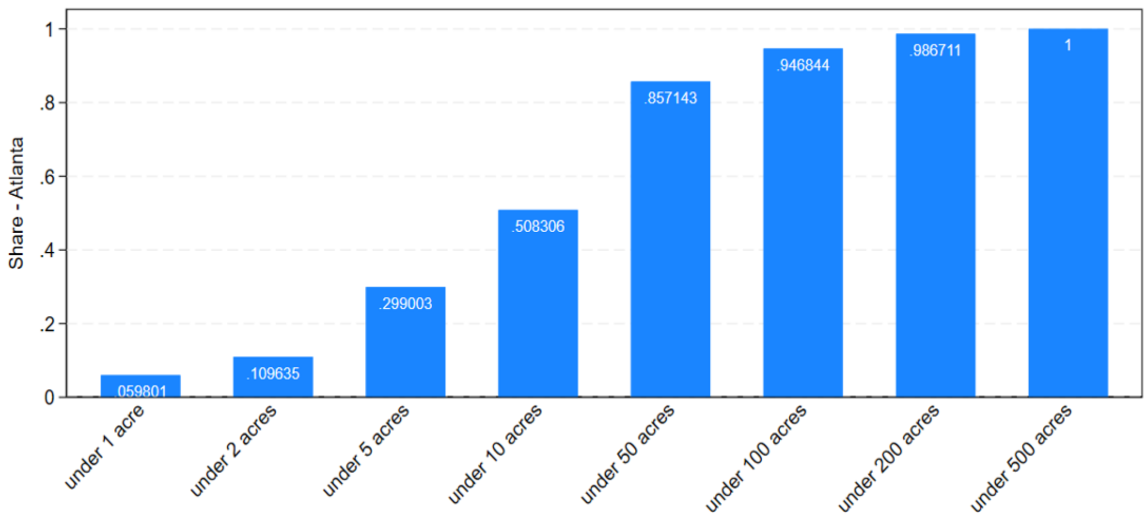




Note. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396 (CBDRB-FY24-P2396-R11004, R11417).

E.2 More on the Size Distribution of Vacant Land

Figure E16: The Size Distribution of Vacant Land Purchases Intended for Single-Family Housing Development in the Atlanta CBSA (Cumulative Distribution Function for the Share of Parcels below Given Square Footage Amounts)



Note. The underlying data are vacant land purchases intended for single-family housing development for 24 CBSAs over the years 2013-2018. The plot is based on 3,640 observations of vacant land parcel purchases. The individual observations were downloaded from proprietary CoStar files and used in Gyourko and Krimmel (2021). See their paper for more details. There are 43,560 square feet in one acre.

E.3 More on Project Size

Table E7: CoreLogic CBSAs and Counties

CBSA	**Counties**
Akron OH MSA	Summit
Albuquerque NM MSA	Bernalillo
Atlanta-Sandy Springs-Roswell GA MSA	DeKalb, Gwinnett, Cobb, Fulton
Austin-Round Rock TX MSA	Travis, Williamson
Bakersfield CA MSA	Kern
Baltimore-Columbia-Towson MD MSA	Baltimore, Anne Arundel, Baltimore city
Baton Rouge LA MSA	East Baton Rouge Parish

Table E7: CoreLogic CBSAs and Counties

CBSA	Counties
Birmingham-Hoover AL MSA	Jefferson
Boise City ID MSA	Ada
Boston-Cambridge-Newton MA-NH MSA	Middlesex, Essex, Suffolk, Norfolk, Plymouth
Bridgeport-Stamford-Norwalk CT MSA	Fairfield
Brownsville-Harlingen TX MSA	Cameron
Buffalo-Cheektowaga-Niagara Falls NY MSA	Erie
Cape Coral-Fort Myers FL MSA	Lee
Charleston-North Charleston SC MSA	Charleston
Charlotte-Concord-Gastonia NC-SC MSA	Mecklenburg
Chicago-Naperville-Elgin IL-IN-WI MSA	Cook, Will, Kane, DuPage, Lake (IL), Lake (IN)
Cincinnati OH-KY-IN MSA	Hamilton
Cleveland-Elyria OH MSA	Cuyahoga
Columbia SC MSA	Richland
Columbus OH MSA	Franklin
Dallas-Fort Worth-Arlington TX MSA	Collin, Dallas, Denton, Tarrant
Dayton OH MSA	Montgomery
Deltona-Daytona Beach-Ormond Beach FL MSA	Volusia
Denver-Aurora-Lakewood CO MSA	Jefferson, Denver, Arapahoe, Adams
Detroit-Warren-Dearborn MI MSA	Wayne, Oakland, Macomb
Durham-Chapel Hill NC MSA	Durham
El Paso TX MSA	El Paso
Flint MI MSA	Genesee
Grand Rapids-Wyoming MI MSA	Kent
Greensboro-High Point NC MSA	Guilford
Greenville-Anderson-Mauldin SC MSA	Greenville
Hartford-West Hartford-East Hartford CT MSA	Hartford
Houston-The Woodlands-Sugar Land TX MSA	Fort Bend, Harris, Montgomery

Table E7: CoreLogic CBSAs and Counties

CBSA	Counties
Indianapolis-Carmel-Anderson IN MSA	Marion
Jacksonville FL MSA	Duval
Kansas City MO-KS MSA	Wyandotte, Jackson, Platte, Clay, Johnson
Knoxville TN MSA	Knox
Lakeland-Winter Haven FL MSA	Polk
Lancaster PA MSA	Lancaster
Las Vegas-Henderson-Paradise NV MSA	Clark
Los Angeles-Long Beach-Anaheim CA MSA	Los Angeles
Madison WI MSA	Dane
Manchester-Nashua NH MSA	Hillsborough
McAllen-Edinburg-Mission TX MSA	Hidalgo
Memphis TN-MS-AR MSA	Shelby
Miami-Fort Lauderdale-West Palm Beach FL MSA	Broward, Palm Beach
Milwaukee-Waukesha-West Allis WI MSA	Waukesha
Minneapolis-St. Paul-Bloomington MN-WI MSA	Dakota, Ramsey, Hennepin
Mobile AL MSA	Mobile
Modesto CA MSA	Stanislaus
Nashville-Davidson–Murfreesboro–Franklin TN MSA	Davidson
New Haven-Milford CT MSA	New Haven
New Orleans-Metairie LA MSA	Orleans Parish, Jefferson Parish
New York-Newark-Jersey City NY-NJ-PA MSA	Hudson (NJ), Middlesex (NJ), Ocean (NJ), Morris (NJ), Monmouth (NJ), Richmond (NY), Passaic (NJ), Union (NJ), Kings (NY), Nassau (NY)
North Port-Sarasota-Bradenton FL MSA	Sarasota, Manatee
Oklahoma City OK MSA	Oklahoma
Omaha-Council Bluffs NE-IA MSA	Douglas

Table E7: CoreLogic CBSAs and Counties

CBSA	Counties
Orlando-Kissimmee-Sanford FL MSA	Seminole
Oxnard-Thousand Oaks-Ventura CA MSA	Ventura
Palm Bay-Melbourne-Titusville FL MSA	Brevard
Philadelphia-Camden-Wilmington PA-NJ-DE-MD MSA	Burlington (NJ), Montgomery (PA), New Castle (DE), Camden (NJ), Delaware (PA), Philadelphia (PA), Chester (PA), Bucks (PA)
Phoenix-Mesa-Scottsdale AZ MSA	Pinal, Maricopa
Pittsburgh PA MSA	Allegheny
Portland-Vancouver-Hillsboro OR-WA MSA	Multnomah (OR), Washington (OR), Clark (WA), Clackamas (OR)
Providence-Warwick RI-MA MSA	Bristol (MA), Providence (RI)
Provo-Orem UT MSA	Utah
Raleigh NC MSA	Wake
Reading PA MSA	Berks
Reno NV MSA	Washoe
Riverside-San Bernardino-Ontario CA MSA	Riverside, San Bernardino
Rochester NY MSA	Monroe
Sacramento–Roseville–Arden-Arcade CA MSA	Sacramento
St. Louis MO-IL MSA	St. Charles (MO), St. Louis (MO)
Salinas CA MSA	Monterey
Salt Lake City UT MSA	Salt Lake
San Antonio-New Braunfels TX MSA	Bexar
San Francisco-Oakland-Hayward CA MSA	Contra Costa, Alameda, San Mateo, Marin, San Francisco
San Jose-Sunnyvale-Santa Clara CA MSA	Santa Clara
Santa Cruz-Watsonville CA MSA	Santa Cruz
Santa Maria-Santa Barbara CA MSA	Santa Barbara
Santa Rosa CA MSA	Sonoma
Seattle-Tacoma-Bellevue WA MSA	Pierce, King, Snohomish
Spokane-Spokane Valley WA MSA	Spokane

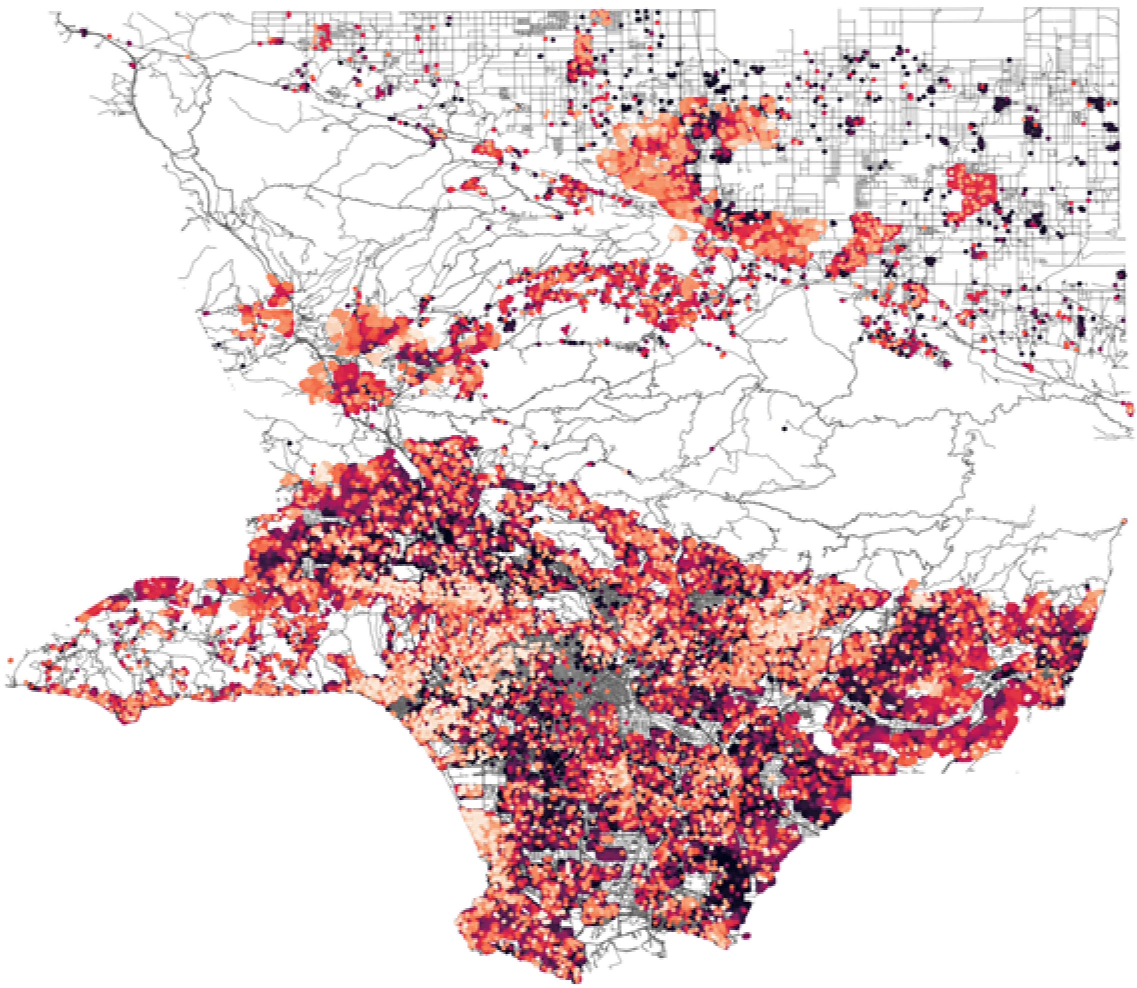
Table E7: CoreLogic CBSAs and Counties

CBSA	**Counties**
Springfield MA MSA	Hampden
Syracuse NY MSA	Onondaga
Tampa-St. Petersburg-Clearwater FL MSA	Hillsborough, Pasco, Pinellas
Toledo OH MSA	Lucas
Tucson AZ MSA	Pima
Tulsa OK MSA	Tulsa
Urban Honolulu HI MSA	Honolulu
Vallejo-Fairfield CA MSA	Solano
Virginia Beach-Norfolk-Newport News VA-NC MSA	Virginia Beach city
Visalia-Porterville CA MSA	Tulare
Washington-Arlington-Alexandria DC-VA-MD-WV MSA	Prince George's (MD), Fairfax (VA), District of Columbia, Montgomery (MD), Loudoun (VA), Prince William (VA)
Wichita KS MSA	Sedgwick
Worcester MA-CT MSA	Worcester
York-Hanover PA MSA	York

E.3.1 Algorithm to Determine Housing Project Size

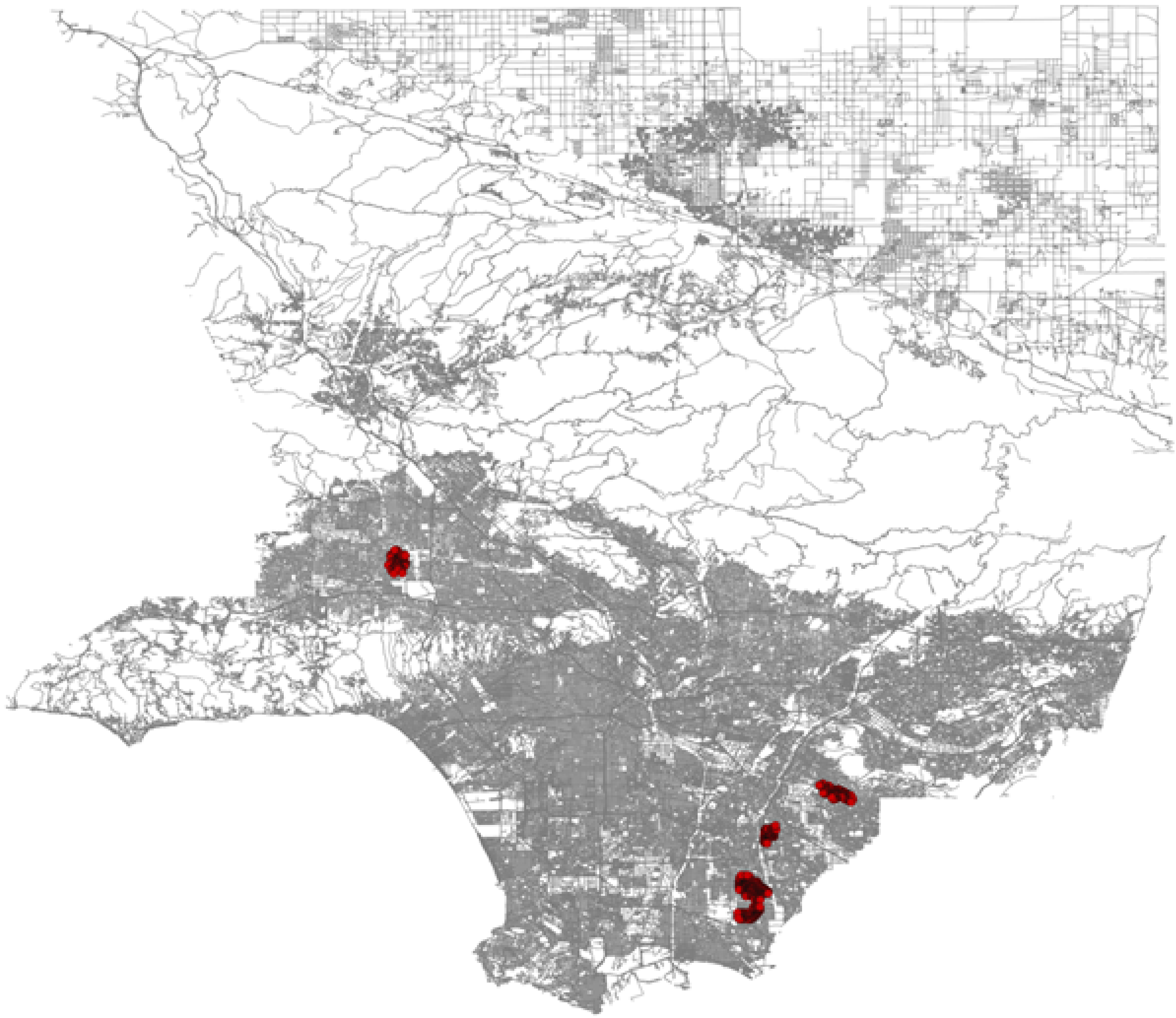
In this section, we illustrate an example of our process for determining housing project size over time. Figure E17 shows every project we defined with the process outlined in the main text between 1950-51 and 2018-19 in Los Angeles County. The different shades of red reflect different project sizes, with the darker colors indicating larger projects. Los Angeles County is a very large area that includes a densely populated southern and coastal region, a much less sparsely developed northern region, and a largely empty mountain region in between.

Figure E17: Single-Family Home Projects Built in Los Angeles County
2-year periods, 1950-51 through 2018-19



It is also easy to see the path of housing development over time, although space considerations prevent us from showing pictures for each 2-year period. Figure E18 shows only the five largest projects from the 1950-51 period. Note that they exist in very different parts of the county. The largest is what we call Lakewood in the southeastern corner of the county. We impute that it comprised 6,344 homes that were built in that two-year period, with each of those homes within 100 yards of at least one other home in the project.

Figure E18: Locations of the Five Largest Projects in 1950-51, Los Angeles County

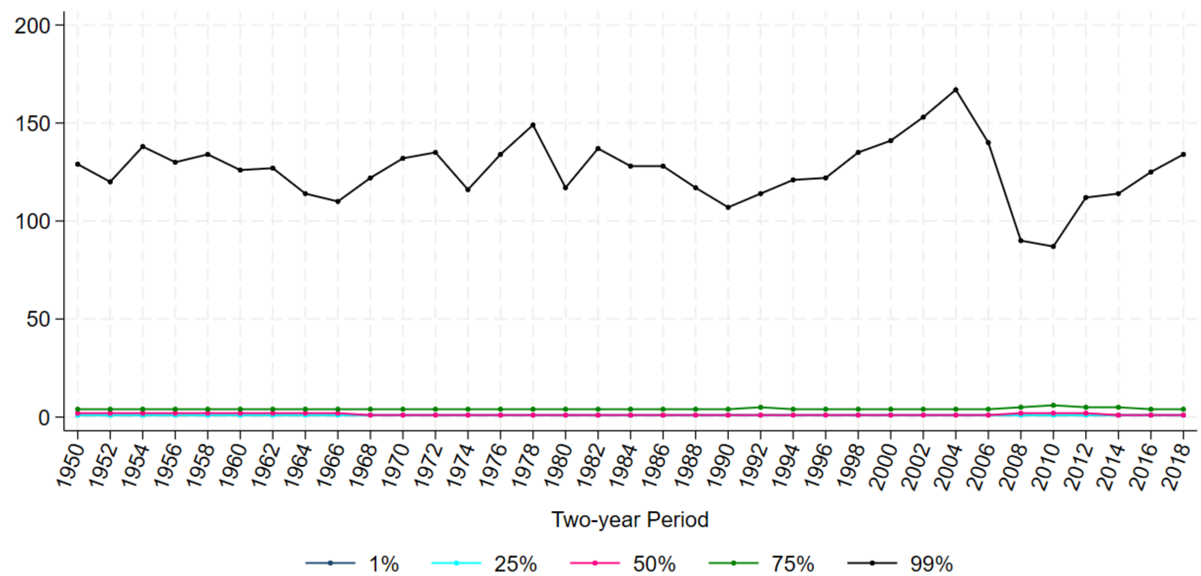


Finally, Figure E19 picture zooms into the so-called Lakewood project. We can confirm that this was an actual project, not some random output of our spatial measurement algorithm. A firm called the Lakewood Park Corporation (hence, our name) spearheaded the development, and there is a rich history of its activities catalogued by the City of Lakewood, California. The detail from zooming into the project illustrates that the GPS coordinates from CoreLogic are accurate, as the homes line up along finely detailed streets from a map that was overlaid, with homes also surrounding what other maps show to be park or other public areas.

Figure E19: 'Lakewood'–The Largest Development Project in Los Angeles County 1950-51 Period

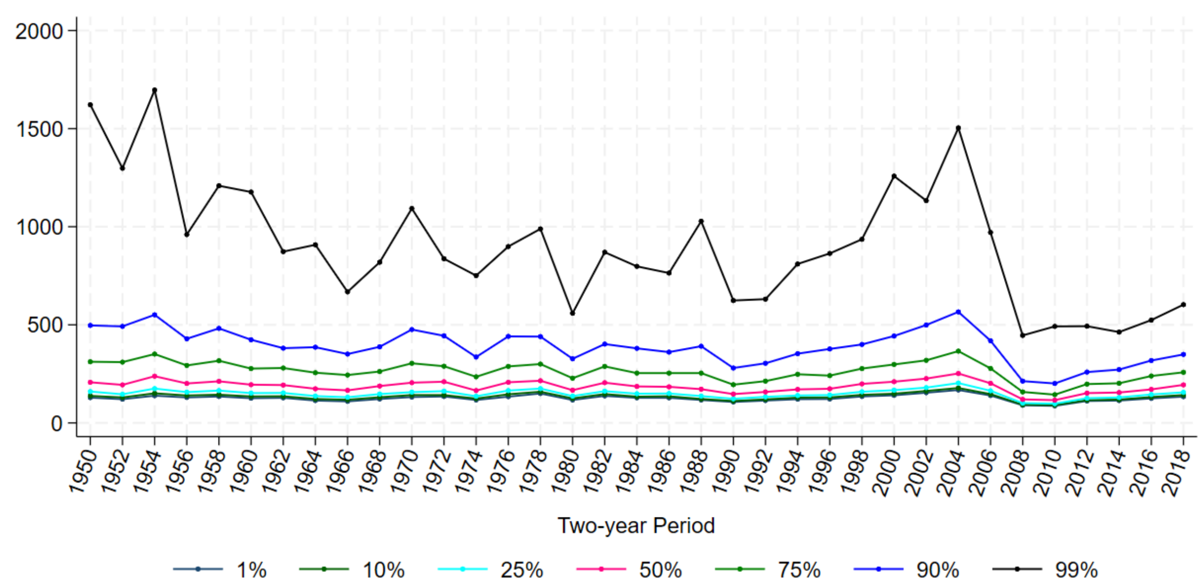
E.3.2 Additional Figures on Project Size

Figure E20: Size Distribution of New Housing Developments over Time: All Counties 1st, 25th, 50th, 75th and 99th Percentiles



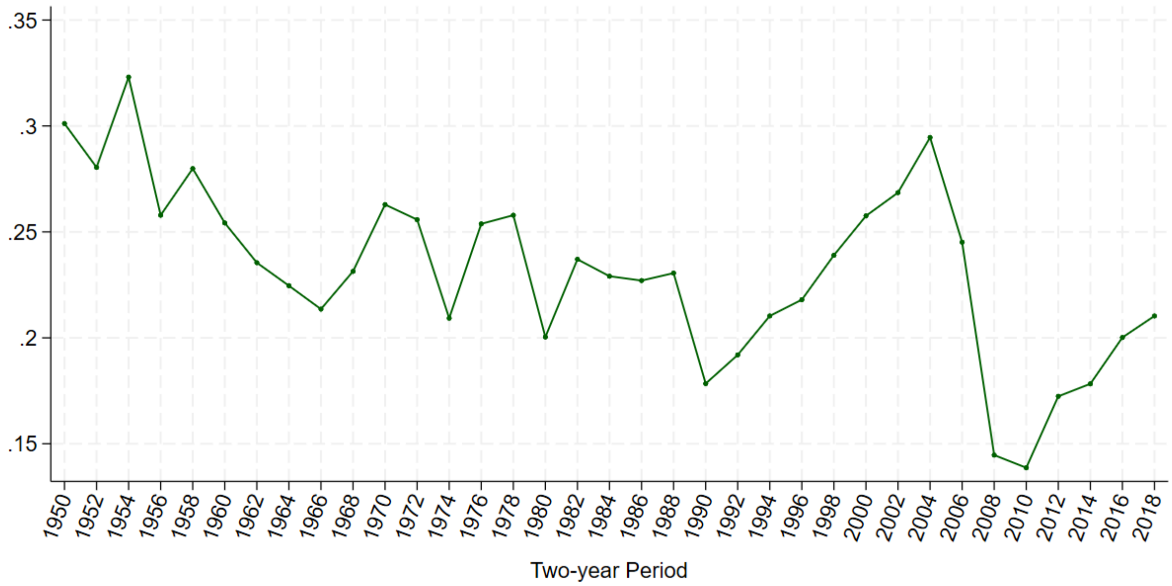
Note. The figure illustrates the percentiles of the project size distribution of all new housing develpments in each respective 2-year period.

Figure E21: Size Distribution of the 1% Largest New Housing Developments over Time: All Counties 1st, 10th, 25th, 50th, 75th, 90th and 99th Percentiles of Largest 1%



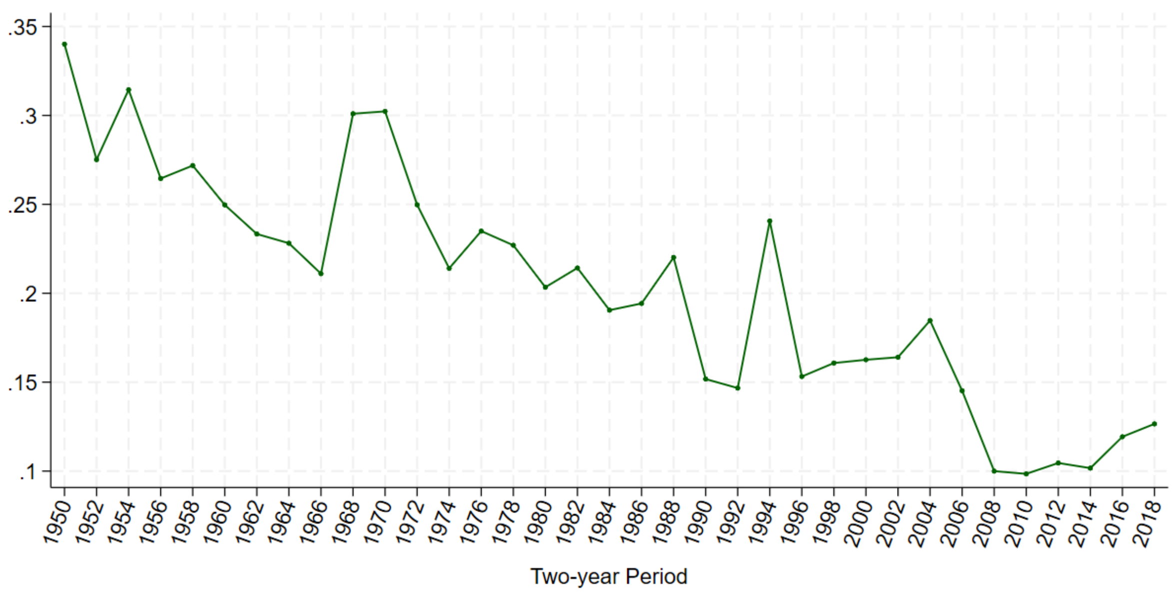
Note. The figure illustrates the percentiles of the project size distribution of the 1% of largest new housing developments in each respective 2-year period.

Figure E22: Share of Top 1% of New Developments in Total New Housing Built (Full CL Data)



Note. The figure plots the share of homes built in the largest 1% of projects in terms of all homes built in each respective 2-year period. The denominator uses all housing construction in CoreLogic, regardless of whether the unit is sampled by our algorithm. See Section 4.3.2 for details.

Figure E23: Share of Top 1% of New Developments in Total New Housing Built, Inversely Weighted by Within-Project Coefficient of Variation in Living Area



Note. The figure plots the weighted share of homes built in the largest 1% of projects in terms of all homes built in each respective 2-year period. Each project is weighted inversely for the within-project coefficient of variation across individual units' living areas. See Section 4.3.2 for details.

E.3.3 Large Projects and Love of Variety

Table E8: Cross-Sectional Hedonic Regressions

	Real Sale Amount (log)			
	(1)	(2)	(3)	(4)
Project Size Dummy (1,000+ units)	-0.108***	-0.0146***		
	(0.00148)	(0.00103)		
Project Size Dummy (100+ units)			-0.110***	-0.0211***
			(0.000506)	(0.000292)
Obs	7,332,624	7,331,891	7,332,624	7,331,891
R^2-Adj	0.0007	0.7757	0.0065	0.7759
Census Tract FE		✓		✓
Sale Year FE		✓		✓
Control - Log Land Square Footage		✓		✓
Control - Living Square Footage		✓		✓
Control - House Age		✓		✓

Note. The table reports results from an OLS regression at the transaction level of the reported real transaction price against a *large* and *non-large project* dummy. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Individual transaction data are taken from the CoreLogic's microdata.

Table E8 uses CoreLogic transactions data to estimate the differential impact of being a home that was built in a large development project (with large defined as at least 100 or 1,000 homes being built within a 2-year period, as described in the text) compared to a home not built in such a large development. The dependent variable is the log of the reported transactions price in constant 2020 dollars. Columns 1 and 3 include only a project size dummy and show the unconditional differences in prices of homes in large versus not-large projects being just over 10% (cheaper in this case). Columns 2 and 4 control for differences in house and site quality. The independent variables in these specifications also include three traditional house quality controls from the hedonic literature–lot size, living area size and house age—as well as census tract and year of sale fixed effects. Price differences narrow to 1%–2% in these specifications.

Table E9: Price Appreciation over Time

	100 homes Threshold		1000 homes Threshold	
	(1)	(2)	(3)	(4)
Large Project Flag	-0.133***	-0.0807***	-0.0969***	-0.0256
	(0.00339)	(0.00345)	(0.0239)	(0.0287)
Obs	29,260	29,256	330	330
R^2-Adj	0.949	0.957	0.969	0.974
Cluster ID FE		✓		✓
Control - D. Avg Bedroom		✓		✓
Control - D. Avg other room		✓		✓
Control - D. Avg House age		✓		✓
Control - House age Square		✓		✓
Control - D. Avg Land Square Footage		✓		✓
Control - D. Avg Living Area		✓		✓

Note. The table reports results from an OLS regression at the census tract level of the real price appreciation against an indicator of whether a census tract contains a large project. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Data on prices are taken from the Decennial census from the 1970s onward. House quality traits data are taken either from the Decennial Census or are built based on the CoreLogic's microdata.

Table E9 uses decennial census data at the tract level dating back to 1970. Tracts are consistently defined over time from that start date. The regressions compare appreciation in the tract(s) containing the large project with at least 100 (or 1,000) homes built in a given 2-year window to appreciation in nearby tracts that only contained projects of fewer than five homes over the same time span. The price data are self-reported values from the decennial censuses. We convert these values into constant 2020 dollars before creating appreciation rates over time.

Because the census reports only minimal house quality traits at the tract level (e.g., the mean number of bedrooms and the mean number of total rooms, which allows us to create a new variable measuring the mean number of other rooms), we exploit the CoreLogic micro data to create tract level measures of house age, lot size and living area square footage that we map onto the census tract data. For example, consider a tract that we know contains hundreds of homes from a very large development that we know was built in the 1950-51 time period. We use CoreLogic's micro data to compute the mean age of homes in the tract during that time period, as well as the typical lot size and living space of all the homes that were in existence before 1951. These averages are used to reflect average house quality in the resulting hedonic estimation.

That specification compares appreciation in large tracts versus nearby (within one mile of the focal large project) tract(s) that only had small projects of fewer than five homes built during the same time span. Columns (1) and (3) estimate unconditional differences using two different measures of being a large project—100+ and 1,000+ homes. Those differences range from 10% to just over 13%. Controlling for house quality as best we can with self-reported (not transactions) data sometimes reduces the gap to close to zero. When it does not, the coefficient is not economically large. For example, the -0.0807 coefficient from Column (2) is just over one-tenth the sample mean appreciation and is well under one-tenth of the standard deviation about that sample mean.

F Appendix to Section 5

F.1 More on Counterfactual Productivity

This section discusses formally the derivation of Equation (10). Let the productivity of worker i in firm j (defined either as the number of units built per employee or revenues per employee) be represented as $a_{ij} = \alpha + f(b_j) + \varepsilon_{ij}$, where b_j is the employment bin (micro, small, medium firms, etc..) of firm j and f is a flexible function that captures the effect of firm size on individual productivity. ε_{ij} is instead the individual-and-firm-level productivity component that does not depend causally on the size of the firms (but might as well be correlated with it). b_j can be interpreted as being an indicator for micro firms (0-4 employees), small firms (5-9), and so on. $f(b_j)$ captures the fact that micro-firms, small firms, etc.. may all have different degrees of productivity because of their size—so that the same worker in a different type of firm would be more productive *because* that firm is larger. ε_{ij} is instead the individual-and-firm-level productivity component that does not depend causally on the size of the firms (but might as well be correlated with it).

Our counterfactual exercise consists in estimating the change in total average productivity per employee that a planner would achieve by changing the firm size distribution to match the FSD of a sector such as manufacturing or nontradables. That is, we are interested in understanding how much more productive would the sector be—because of the effect of size on productivity—if employees were to be more concentrated in larger firms. Letting N_j be the fraction of employment accounted for by firms in bin j in the data, and N_j' be the employment share in that same bin according to the counterfactual FSD, we want to estimate:

$$\Delta A := \sum_j \underbrace{\left(N_j' - N_j\right)}_{\substack{\Delta \text{ in number} \\ \text{of workers in bin } j}} \underbrace{f(b_j)}_{\substack{\text{Size-dependent} \\ \text{productivity}}}$$

The key issue is that we do not observe the causal impact of size on productivity, $f(b_j)$. Instead, we can only estimate the bin-specific productivities reported in Figure 10. In particular, we estimate $\bar{a}_j$: the average output per employee for firms in bin j, where the 0 subscript indicates that the estimate is conditional on the firm structure at time 0 (i.e. the one we observe in the data). If $\bar{a}_j$ only reflected the size of the firm, then we would be able to run our counterfactual exercise since we would effectively be estimating $f(b_j)$. However, productivity may also be the artifact of the sorting of high-ability workers in large firms or of large firms being more productive for other reasons unrelated to size. To discipline this omitted variable bias, we parametrize it using a parameter ϕ that linearly controls how much of the link between size and productivity is causal. Denote with $\Delta\bar{a}_j$ the difference in productivity between firms in bin j and firms in the lowest bin, i.e. $\Delta\bar{a}_j = \bar{a}_j - \underline{a}$, where $\underline{a}$ is the productivity of micro firms (those with the lowest $b_j = \underline{b}$, where the reference group is taken without loss of generality). $\phi \in [0,1]$ is defined as:

$$\phi \times \underbrace{\Delta\bar{a}_j}_{\substack{\text{Observed} \neq \text{in} \\ \text{productivity}}} = \underbrace{\Delta f_j}_{\substack{\text{Causal} \neq \text{in prod.} \\ \text{due to firm size}}}$$

where $\Delta f_j = f(b_j) - f(\underline{b})$. Under this definition, we have (10). To see this, note that adding and subtracting $f(\underline{b})$ in the definition of ΔA, one has:

$$\Delta A = \sum_j \underbrace{\left(N_j' - N_j\right)}_{\substack{\Delta \text{ in number} \\ \text{of workers in bin } j}} \underbrace{f(b_j)}_{\substack{\text{Size-dependent} \\ \text{productivity}}} = \sum_j \left(N_j' - N_j\right) \left(f(b_j) - f(\underline{b}) + f(\underline{b})\right) \tag{94}$$

$$= \sum_j \left(N_j' - N_j\right) \left(f(b_j) - f(\underline{b})\right) \tag{95}$$

where the last equality follows from the fact that $\sum_j N_j' = \sum_j N_j = 1$ since N_j and N_j' are employment shares in bin j, so that $\sum_j \left(N_j' - N_j\right) f(\underline{b}) = 0$. Finally, substituting (F.1) and noting again that $\sum_j \left(N_j' - N_j\right) \underline{a} = 0$, one has Equation (10).

F.2 More on CCI Size Regressions

Columns (1) and (3) of Appendix Table F10 replicates the uninteracted coefficient of Columns (1) and (4) in Table 2. In the remaining columns, we introduce six regressors for the firm's revenue share by type of construction b (omitting housing), $RS(b)_i$, and six interactions of these shares with the log employment variable. Appendix Table F11 does the same for the dependent variables in Columns (2) and (3) of Table 2. The regression model is:

$$\text{Profit}_i = \alpha + \beta_0 * \log(\text{Empl}_i) + \sum_b (\gamma_b \times RS(b)_i + \beta_{1,b} \times RS(b)_i * \log(\text{Empl}_i)) + e_i$$

In this model, the beta on log employment measures the scaling elasticity for a firm that is fully specialized in housing construction. As shares are measured in whole numbers (an $RS(b) = 100$ corresponds to full specialization in sector b), the interacted coefficients can be interpreted as the additional scaling elasticity for 1% of the firm's revenue instead coming from specified sector. Thus, the estimate for a firm fully specialized in office buildings, for example, would come from multiplying the interaction term for office buildings times 100 and adding to the base term (β_0 + 100*$\beta_{1,\text{offices}}$).

Table F10: Firm Size, Firm Profitability, and Capital per Employee

VARIABLES	(1) Profits in unit standard deviations	(2) Profits in unit standard deviations	(3) Log capital per employee	(4) Log capital per employee
Log employment	0.1375***	0.07904***	0.06882***	-0.0107**
	(0.0062)	(0.0086)	(0.00273)	(0.0048)
Log employment interacted with firm revenue composition				
Share industrial buildings		0.001343***		0.0001439
		(0.00036)		(0.00016)
Share consumer-facing buildings		0.0004744**		0.0002448*
		(0.00014)		(0.00014)
Share office buildings		0.0008268***		-0.0002462
		(0.00031)		(0.00015)
Share warehousing		-0.0004776		0.0002288
		(0.00049)		(0.00035)
Share other buildings		0.0007339***		-0.00000825
		(0.00024)		(0.00011)
Share non-building construction		0.001955***		0.0006862***
		(0.00022)		(0.00008)
(Share housing omitted category)				
Main effects for firm revenue composition				
Share industrial buildings		-0.003917***		0.002548***
		(0.00084)		(0.00049)
Share consumer-facing buildings		-0.001419***		0.000717*
		(0.00037)		(0.00038)
Share office buildings		-0.002532***		-0.0000783
		(0.00064)		(0.00045)
Share warehousing		0.0001325		0.002679***
		(0.00094)		(0.00102)
Share other buildings		-0.002806***		0.001861***
		(0.00055)		(0.00035)
Share non-building construction		-0.00521***		0.009609***
		(0.00048)		(0.00025)
Observations	107,000	107,000	107,000	107,000
R-squared	0.0368	0.0475	0.0061	0.1094

Note. The table reports results from a firm-level OLS regression of profit and log capital per employee on employment interacted by firm revenue composition. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Analysis uses microdata from 2012 Census of Construction Industries (CCI). Regressions are unweighted and have 107,000 observations (rounded per Census Bureau disclosure requirements). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396. (CBDRB-FY24-P2396-R11004, R11417).

Table F11: Firm Size and Labor Productivity

VARIABLES	(1) Log labor productivity	(2) Log labor productivity	(3) Log labor productivity with adjustment	(4) Log labor productivity with adjustment
Log employment	0.1094***	0.0605***	0.1017***	0.06617***
	(0.0020)	(0.0040)	(0.0018)	(0.0036)
Log employment interacted with firm revenue composition				
Share industrial buildings		0.0001993*		0.0002062**
		(0.00011)		(0.00010)
Share consumer-facing buildings		0.0000683		-0.0001603
		(0.00012)		(0.00010)
Share office buildings		0.0002567**		0.00017
		(0.00012)		(0.00011)
Share warehousing		0.0000946		0.0001938
		(0.00026)		(0.00023)
Share other buildings		0.0008441***		0.0002791***
		(0.00009)		(0.00008)
Share non-building construction		0.00074***		0.000501***
		(0.00006)		(0.00006)
(Share housing omitted category)				
Main effects for firm revenue composition				
Share industrial buildings		0.000052		0.0008289**
		(0.00036)		(0.00033)
Share consumer-facing buildings		0.0009285***		0.001295***
		(0.00032)		(0.00029)
Share office buildings		0.0009451**		0.000504
		(0.00037)		(0.00034)
Share warehousing		0.0002044		0.0005828
		(0.00075)		(0.00066)
Share other buildings		0.0002715		0.0008163***
		(0.00029)		(0.00026)
Share non-building construction		-0.000471***		0.0009115***
		(0.00017)		(0.00017)
Observations	107,000	107,000	107,000	107,000
R-squared	0.0319	0.393	0.0346	0.0442

Note. The table reports results from a firm-level OLS regression of log labor productivity with and without subcontractor adjustment on employment interacted by firm revenue composition. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Analysis uses microdata from 2012 Census of Construction Industries (CCI). Regressions are unweighted and have 107,000 observations (rounded per Census Bureau disclosure requirements). This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2396. (CBDRB-FY24-P2396-R11004, R11417).

F.3 Details on Censoring

In section 5.2, we introduce the share of employment in establishments of large firms (i.e., firms with more than 100 employees) as a proxy for local firm size. To construct the employment share in large firms, we subtract from the total amount of employees in a CBSA the number of employees working in establishments of firms with < 20 and $20 - 99$ employees. Employment data is sometimes censored, but, importantly, the total count of establishments is not—which allows us to impute the missing data. We miss employment data on 41 CBSAs for all construction, 153 CBSAs for the Construction of Buildings subsector, 175 and 71 CBSAs for the Heavy and Civil Engineering and Specialty Contractors subsectors, respectively. To address this issue, we adopt three different imputation schemes. The first one, which we refer to as 'national midpoints' consists of finding the national average number of employees in establishments with < 20 and $20 - 99$ employees in each subsector, and multiplying such averages by the noncensored number of establishments in such bins operating in the CBSA.[16] The second and third imputation techniques are based on interpreting the suppression flags coming with the SUSB data. Whenever employment in a given CBSA is censored, the data gives us a lower bound and an upper bound for the suppressed figures. We use this information to impute data in missing CBSAs, setting employment in the $100 - 499$ and in the $500+$ categories equal to their respective upper and lower bounds.

[16]For a handful of cases, this procedure leads to significantly exaggerate the number of employees, as national averages might well be greater than local averages. As a result, we drop those CBSAs for which our imputed employment in the < 100 employment bin exceeds by more than 10% the total employment in the CBSA. This leads us to exclude 2 CBSAs in the all construction sample, as well as 10 (Construction of Buildings), 21 (Heavy and Civil Engineering) and 10 (Specialty Trade Contractors) CBSAs for the subsector samples. In the projected WRLURI samples, this very exclusion leads us to drop 5, 17, 38 and 13 CBSAs, for the All Construction, Construction of Buildings, Heavy and Civil Engineering and Specialty Contractors subsectors, respectively.

Table F12: Employment in Large Firms in All Construction, with Different Imputation Schemes

	(1)	(2)	(3)	(4)
VARIABLES				
WRLURI	-0.0367***	-0.0356***	-0.0346***	-0.0354***
	(0.0130)	(0.0128)	(0.0129)	(0.0128)
Log of population (2012)	0.0750***	0.0746***	0.0715***	0.0742***
	(0.00948)	(0.00925)	(0.00929)	(0.00928)
Population density (2012)	-255.4***	-253.4***	-242.6***	-252.1***
	(36.96)	(36.35)	(36.10)	(36.33)
Log of house values (2012)	-0.00918	-0.0108	-0.0143	-0.0113
	(0.0304)	(0.0302)	(0.0306)	(0.0303)
Constant	-0.581*	-0.557*	-0.471	-0.545*
	(0.324)	(0.321)	(0.325)	(0.321)
Observations	502	543	543	541
R-squared	0.365	0.368	0.336	0.364
Imputation	Not imp.	Lower bound	Upper bound	National midpoints

Note. The table reports results from a CBSA-level WLS regression of the share of employment in large firms against WRLURI. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. CBSAs are weighted by the resident population. Column (1) does not impute employment data in missing CBSAS. Column (2) imputes data in the 100-499 and 500+ employment beans based on the lower bounds of employment as summarized by the suppression flags. Column (3) performs the imputation as in column (2), but using the upper bounds of employment as summarized by the suppression flags. Column (4) imputes missing values based on national midpoints of employment per firm in the < 20 and $20 - 99$ bins. We drop CBSA for which this imputed employment exceeds total employment by more than 10%.

Table F13: Employment in Large Firms in Construction of Buildings, with Different Imputation Schemes

VARIABLES	(1)	(2)	(3)	(4)
WRLURI	-0.0700***	-0.0615***	-0.0588***	-0.0607***
	(0.0152)	(0.0143)	(0.0150)	(0.0146)
Log of population (2012)	0.0888***	0.0837***	0.0741***	0.0797***
	(0.0101)	(0.00904)	(0.00948)	(0.00919)
Population density (2012)	-338.3***	-319.2***	-296.5***	-308.9***
	(56.10)	(54.50)	(54.78)	(54.01)
Log of house values (2012)	0.0314	0.0273	0.0249	0.0243
	(0.0303)	(0.0290)	(0.0293)	(0.0292)
Constant	-1.327***	-1.208***	-1.041***	-1.115***
	(0.361)	(0.336)	(0.340)	(0.339)
Observations	343	498	498	488
R-squared	0.391	0.366	0.269	0.325
Imputation	Not imp.	Lower bound	Upper bound	National midpoints

Note. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. See Table F12 for more information.

Table F14: Employment in Large Firms in Heavy and Civil Engineering, with Different Imputation Schemes

	(1)	(2)	(3)	(4)
VARIABLES				
WRLURI	-0.0463*	-0.0398*	-0.0382*	-0.0396*
	(0.0246)	(0.0217)	(0.0226)	(0.0217)
Log of population (2012)	0.127***	0.123***	0.0874***	0.110***
	(0.0139)	(0.0108)	(0.0120)	(0.0111)
Population density (2012)	-211.5***	-212.7***	-140.2**	-185.7***
	(68.21)	(60.01)	(60.56)	(60.72)
Log of house values (2012)	-0.0159	-0.00853	-0.00298	-0.00945
	(0.0395)	(0.0356)	(0.0379)	(0.0361)
Constant	-1.037**	-1.080**	-0.629	-0.869*
	(0.515)	(0.441)	(0.461)	(0.444)
Observations	227	402	402	381
R-squared	0.387	0.410	0.206	0.338
Imputation	Not imp.	Lower bound	Upper bound	National midpoints

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. See Table F12 for more information.

Table F15: Employment in Large Firms in Specialty Trade Contractors, with Different Imputation Schemes

VARIABLES	(1)	(2)	(3)	(4)
WRLURI	-0.0354***	-0.0334***	-0.0316**	-0.0328**
	(0.0130)	(0.0127)	(0.0127)	(0.0128)
Log of population (2012)	0.0745***	0.0740***	0.0698***	0.0732***
	(0.00841)	(0.00805)	(0.00812)	(0.00811)
Population density (2012)	-259.9***	-256.8***	-242.7***	-253.2***
	(34.62)	(33.67)	(33.53)	(33.78)
Log of house values (2012)	0.0183	0.0155	0.0106	0.0136
	(0.0316)	(0.0312)	(0.0317)	(0.0314)
Constant	-0.952***	-0.911***	-0.794**	-0.877**
	(0.350)	(0.342)	(0.349)	(0.345)
Observations	471	542	542	532
R-squared	0.413	0.422	0.365	0.409
Imputation	Not imp.	Lower bound	Upper bound	National midpoints

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. See Table F12 for more information.

Tables F12-F15 replicate the WRLURI regressions in the noncensored samples (column (1)) and for each imputation scheme. National midpoints estimates are typically very close to the noncensored sample, and lay in between the upper and lower bound estimates. Nonetheless, all coefficients are qualitatively identical and quantitatively very similar.

F.4 Projected WRLURI

As described in Section 5.2, we do four different WRLURI projections, depending on which CBSA-level characteristics we use to predict WRLURI. Table F16 reports the results for the different specifications. Table F17 reports pairwise correlations across the different imputed measures and the raw index. Tables F18 to F25 report regression results of Section 5.2 across all the projections.

Table F16: CBSA Characteristics and WRLURI

VARIABLES	(1) WRLURI (2006)	(2) WRLURI (2006)	(3) WRLURI (2006)	(4) WRLURI (2006)
Log of population	0.0664***	0.0177	-0.438	0.676
	(0.0151)	(0.0436)	(0.391)	(0.948)
Log of population density	0.0461	0.177	-0.0510	0.124
	(0.0666)	(0.215)	(0.0636)	(0.256)
Log of housing units			1.313	-2.988
			(0.950)	(3.095)
Log of (median) house value			0.597***	0.477
			(0.211)	(0.596)
Log of per capita income			-0.924**	0.167
			(0.388)	(1.021)
Average school years completed by females	0.263*	0.872**	0.0613	0.485
	(0.149)	(0.340)	(0.139)	(0.424)
Average school years completed by males	-0.158	-0.406	-0.120	-0.0861
	(0.118)	(0.279)	(0.112)	(0.323)
Log of owner-occupied housing units			-0.264	1.337
			(0.485)	(1.375)
Log of renter-occupied housing units			-0.335	0.481
			(0.339)	(1.276)
Log of vacant housing units			-0.252*	0.447
			(0.129)	(0.294)
Log of median contract rent			1.038***	1.108
			(0.263)	(0.885)
Constant	-2.341***	-5.342**	-7.367***	-11.49**
	(0.812)	(2.598)	(1.337)	(4.700)
Observations	550	550	550	550
R-squared	0.364	0.438	0.473	0.574
Census Division FE	✓	✓	✓	✓
Census Division Interactions	✗	✓	✗	✓

Note. The table reports results from a CBSA-level WLS regression of the WRLURI on CBSA's characteristics. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. We use 2006's WRLURI vintage (see Gyourko et al. (2008) for reference) collapsed at the CBSA level taking simple means. All of the regressors listed above belong to 2000's Decadal Census. Observations are weighted by population.

Table F17: Correlation across WRLURI Measures

	Raw WRLURI	WRLURI hat (1)	WRLURI hat (2)	WRLURI hat (3)	WRLURI hat (4)
Raw WRLURI	1.00				
WRLURI hat (1)	0.60***	1.00			
WRLURI hat (2)	0.66***	0.88***	1.00		
WRLURI hat (3)	0.69***	0.86***	0.79***	1.00	
WRLURI hat (4)	0.76***	0.66***	0.76***	0.77***	1.00
Observations	942				

Note. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Table F18: WRLURI Projections on the Log of Total Receipts per Establishment by CBSA - All Construction

	Dependent variable: log of total receipts per establishment					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.162***	-0.293***	-0.292***	-0.174***	-0.368***	-0.224***
	(0.0456)	(0.0916)	(0.0838)	(0.0650)	(0.114)	(0.0834)
Log of population (2012)	0.245***	0.251***	0.254***	0.242***	0.249***	0.242***
	(0.0320)	(0.0343)	(0.0302)	(0.0294)	(0.0292)	(0.0268)
Population density (2012)	-725.6***	-746.7***	-742.4***	-722.1***	-686.6***	-674.1***
	(130.1)	(126.0)	(117.1)	(124.5)	(113.8)	(114.1)
Log of house values (2012)	0.0500	0.111	0.0990	0.0470	0.249*	0.122
	(0.111)	(0.111)	(0.102)	(0.0966)	(0.135)	(0.119)
Constant	10.68***	9.834***	9.937***	10.74***	8.203***	9.838***
	(1.227)	(1.319)	(1.206)	(1.116)	(1.607)	(1.415)
Observations	545	545	852	852	852	852
R-squared	0.422	0.421	0.446	0.431	0.450	0.452
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. The table reports results from a CBSA-level WLS regression of the log of total receipts per establishment against different WRLURI measures. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Observations are weighted by population. We use 2006's WRLURI vintage (see Gyourko et al. (2008) for reference). Set 1 is obtained projecting WRLURI on the (log of) population and population density in a given CBSA, the average school years of male and female residents, and fixed effects at the Census Division level. Set 2 is obtained by augmenting the regressors in Set 1 with their interactions with Census Division fixed effects. Set 3 contains the regressors of Set 1, with the addition of self-reported house prices, per capita income, median rent and number of housing units (broken down by owner-, renter-occupied and vacant). Set 4 comprises the regressors of set 3, with their interactions with Census Division fixed effects. All of the regressors listed above are taken from the 2000's Decadal Census. *Original* indicates that we restrict the sample to the original WRLURI 2006 sample of CBSAs, whereas *Full* suggests that the sample is that of all SUSB nonmissing CBSAs.

Table F19: WRLURI Projections on the Share of Employment in Large Firms by CBSA - All Construction

	Dependent variable: share of employment in large firms					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.0354***	-0.0602**	-0.0618***	-0.0366**	-0.0749***	-0.0497**
	(0.0128)	(0.0241)	(0.0222)	(0.0172)	(0.0290)	(0.0210)
Log of population (2012)	0.0742***	0.0754***	0.0756***	0.0730***	0.0742***	0.0730***
	(0.00928)	(0.00964)	(0.00856)	(0.00843)	(0.00829)	(0.00798)
Population density (2012)	-252.1***	-255.5***	-253.6***	-249.2***	-241.5***	-239.6***
	(36.33)	(34.48)	(32.26)	(33.90)	(32.57)	(33.07)
Log of house values (2012)	-0.0113	-0.000675	-0.00187	-0.0130	0.0269	0.00556
	(0.0303)	(0.0287)	(0.0264)	(0.0260)	(0.0341)	(0.0298)
Constant	-0.545*	-0.694**	-0.684**	-0.510*	-1.014***	-0.736**
	(0.321)	(0.317)	(0.291)	(0.278)	(0.390)	(0.347)
Observations	541	541	846	846	846	846
R-squared	0.364	0.361	0.372	0.365	0.373	0.377
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. For a description, consult the notes of Table F18.

Table F20: WRLURI Projections on the Log of Total Receipts per Establishment by CBSA - Construction of Buildings

	Dependent variable: log of total receipts per establishment					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.181***	-0.283**	-0.290**	-0.177*	-0.271*	-0.172**
	(0.0525)	(0.123)	(0.113)	(0.0933)	(0.145)	(0.0870)
Log of population (2012)	0.322***	0.326***	0.337***	0.326***	0.328***	0.323***
	(0.0379)	(0.0405)	(0.0356)	(0.0348)	(0.0345)	(0.0325)
Population density (2012)	-1,098***	-1,107***	-1,137***	-1,119***	-1,071***	-1,063***
	(159.7)	(158.3)	(148.8)	(155.5)	(147.4)	(148.6)
Log of house values (2012)	0.0883	0.125	0.129	0.0812	0.185	0.100
	(0.110)	(0.121)	(0.110)	(0.106)	(0.152)	(0.122)
Constant	9.384***	8.862***	8.656***	9.408***	8.102***	9.207***
	(1.286)	(1.537)	(1.397)	(1.354)	(1.914)	(1.502)
Observations	530	530	808	808	808	808
R-squared	0.433	0.425	0.470	0.461	0.463	0.465
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. For a description, consult the notes of Table F18.

Table F21: WRLURI Projections on the Share of Employment in Large Firms by CBSA - Construction of Buildings

	Dependent variable: share of employment in large firms					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.0607***	-0.0887***	-0.0795***	-0.0746***	-0.107***	-0.0597**
	(0.0146)	(0.0289)	(0.0268)	(0.0197)	(0.0319)	(0.0245)
Log of population (2012)	0.0797***	0.0804***	0.0801***	0.0779***	0.0788***	0.0767***
	(0.00919)	(0.00911)	(0.00807)	(0.00819)	(0.00783)	(0.00781)
Population density (2012)	-308.9***	-309.1***	-305.6***	-312.4***	-291.4***	-287.3***
	(54.01)	(46.86)	(45.30)	(48.19)	(45.80)	(48.84)
Log of house values (2012)	0.0243	0.0321	0.0270	0.0372	0.0742**	0.0323
	(0.0292)	(0.0279)	(0.0257)	(0.0250)	(0.0356)	(0.0324)
Constant	-1.115***	-1.225***	-1.160***	-1.253***	-1.716***	-1.175***
	(0.339)	(0.322)	(0.294)	(0.279)	(0.411)	(0.390)
Observations	488	488	730	730	730	730
R-squared	0.325	0.311	0.330	0.336	0.336	0.334
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. For a description, consult the notes of table F18.

Table F22: WRLURI Projections on the Log of Total Receipts per Establishment by CBSA - Heavy and Civil Engineering Construction

	Dependent variable: log of total receipts per establishment					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.125*	-0.401***	-0.395***	-0.252***	-0.597***	-0.274**
	(0.0657)	(0.122)	(0.114)	(0.0937)	(0.156)	(0.114)
Log of population (2012)	0.326***	0.342***	0.356***	0.341***	0.353***	0.339***
	(0.0403)	(0.0392)	(0.0342)	(0.0345)	(0.0338)	(0.0325)
Population density (2012)	-337.1*	-410.2***	-441.1***	-426.4***	-381.7**	-349.2**
	(181.8)	(157.2)	(148.1)	(162.1)	(147.8)	(159.0)
Log of house values (2012)	0.0204	0.193	0.183	0.131	0.481***	0.186
	(0.131)	(0.129)	(0.119)	(0.121)	(0.175)	(0.145)
Constant	10.96***	8.621***	8.541***	9.394***	4.958**	8.760***
	(1.517)	(1.521)	(1.405)	(1.416)	(2.080)	(1.760)
Observations	422	422	627	627	627	627
R-squared	0.415	0.434	0.457	0.448	0.470	0.457
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. For a description, consult the notes of table F18.

Table F23: WRLURI Projections on the Share of Employment in Large Firms by CBSA - Heavy and Civil Engineering Construction

	Dependent variable: share of employment in large firms					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.0396*	-0.0969**	-0.0963**	-0.0656**	-0.107**	-0.0700***
	(0.0217)	(0.0403)	(0.0381)	(0.0303)	(0.0448)	(0.0228)
Log of population (2012)	0.110***	0.113***	0.115***	0.111***	0.112***	0.110***
	(0.0111)	(0.0111)	(0.00978)	(0.00978)	(0.00938)	(0.00935)
Population density (2012)	-185.7***	-199.2***	-200.6***	-198.2***	-181.0***	-177.8***
	(60.72)	(60.40)	(57.99)	(57.40)	(59.79)	(59.20)
Log of house values (2012)	-0.00945	0.0243	0.0191	0.00974	0.0542	0.0227
	(0.0361)	(0.0400)	(0.0378)	(0.0391)	(0.0523)	(0.0350)
Constant	-0.869*	-1.326***	-1.288***	-1.121**	-1.683***	-1.270***
	(0.444)	(0.504)	(0.470)	(0.472)	(0.632)	(0.422)
Observations	381	381	541	541	541	541
R-squared	0.338	0.345	0.367	0.363	0.366	0.369
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. For a description, consult the notes of Table F18.

Table F24: WRLURI Projections on the Log of Total Receipts per Establishment by CBSA - Specialty Trade Contractors

	Dependent variable: log of total receipts per establishment					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.158***	-0.224***	-0.219***	-0.132**	-0.362***	-0.218***
	(0.0489)	(0.0814)	(0.0739)	(0.0556)	(0.101)	(0.0748)
Log of population (2012)	0.215***	0.217***	0.217***	0.208***	0.216***	0.209***
	(0.0249)	(0.0278)	(0.0246)	(0.0238)	(0.0229)	(0.0206)
Population density (2012)	-598.9***	-598.8***	-589.9***	-575.3***	-558.0***	-546.3***
	(113.1)	(113.6)	(104.9)	(110.8)	(97.89)	(96.70)
Log of house values (2012)	0.101	0.116	0.103	0.0653	0.299**	0.173
	(0.110)	(0.112)	(0.103)	(0.0983)	(0.122)	(0.115)
Constant	10.08***	9.849***	10.00***	10.59***	7.644***	9.269***
	(1.230)	(1.343)	(1.220)	(1.136)	(1.437)	(1.354)
Observations	543	543	852	852	852	852
R-squared	0.417	0.400	0.423	0.413	0.443	0.445
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. For a description, consult the notes of Table F18.

Table F25: WRLURI Projections on the Share of Employment in Large Firms by CBSA - Specialty Trade Contractors

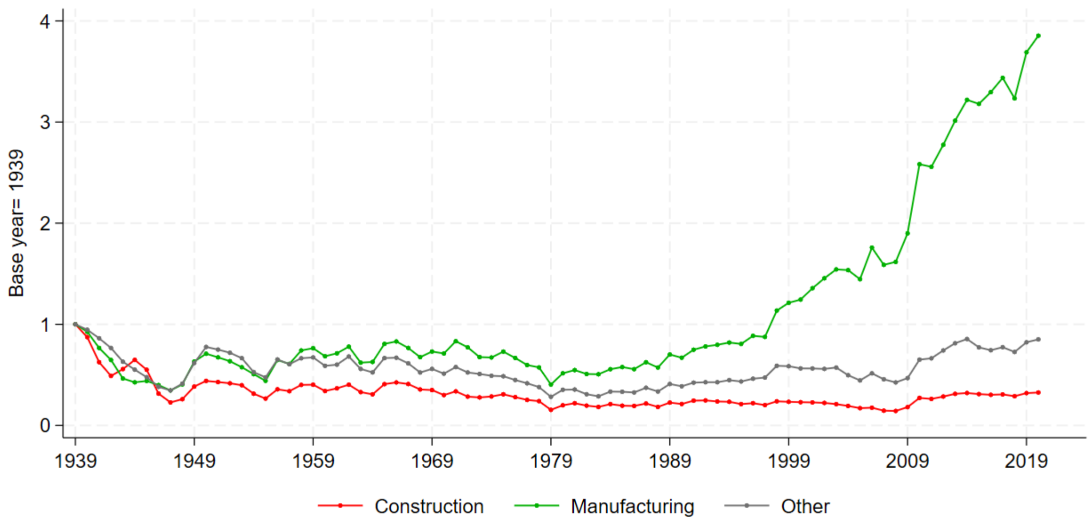
	Dependent variable: share of employment in large firms					
	(1)	(2)	(3)	(4)	(5)	(6)
WRLURI	-0.0328**	-0.0304	-0.0332	-0.0130	-0.0469	-0.0381*
	(0.0128)	(0.0244)	(0.0225)	(0.0178)	(0.0300)	(0.0196)
Log of population (2012)	0.0732***	0.0725***	0.0721***	0.0705***	0.0717***	0.0711***
	(0.00811)	(0.00865)	(0.00771)	(0.00750)	(0.00745)	(0.00706)
Population density (2012)	-253.2***	-247.8***	-245.7***	-240.5***	-240.0***	-239.8***
	(33.78)	(34.08)	(31.73)	(33.20)	(30.34)	(30.39)
Log of house values (2012)	0.0136	0.00545	0.00667	-0.00521	0.0286	0.0223
	(0.0314)	(0.0321)	(0.0293)	(0.0281)	(0.0360)	(0.0321)
Constant	-0.877**	-0.772**	-0.782**	-0.614*	-1.041**	-0.957**
	(0.345)	(0.373)	(0.338)	(0.317)	(0.425)	(0.374)
Observations	532	532	830	830	830	830
R-squared	0.409	0.396	0.408	0.404	0.410	0.416
Sample	Original	Original	Full	Full	Full	Full
WRLURI projection	Raw	Set 1	Set 1	Set 2	Set 3	Set 4

Note. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. For a description, consult the notes of Table F18.

G Appendix to Section 6

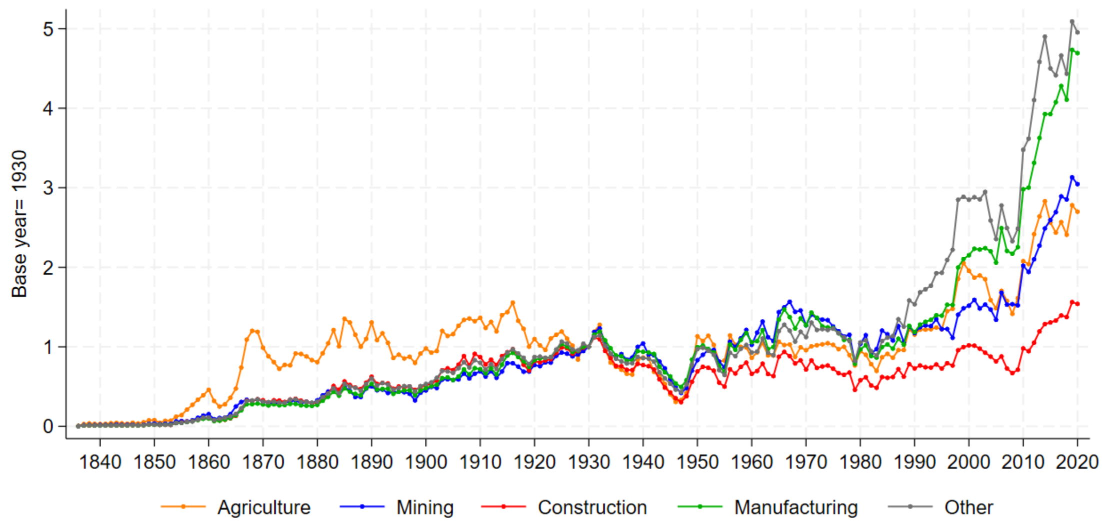
G.1 Patents per Employee

Figure G24: Patent Levels Normalized by Industry Employment, by Industry



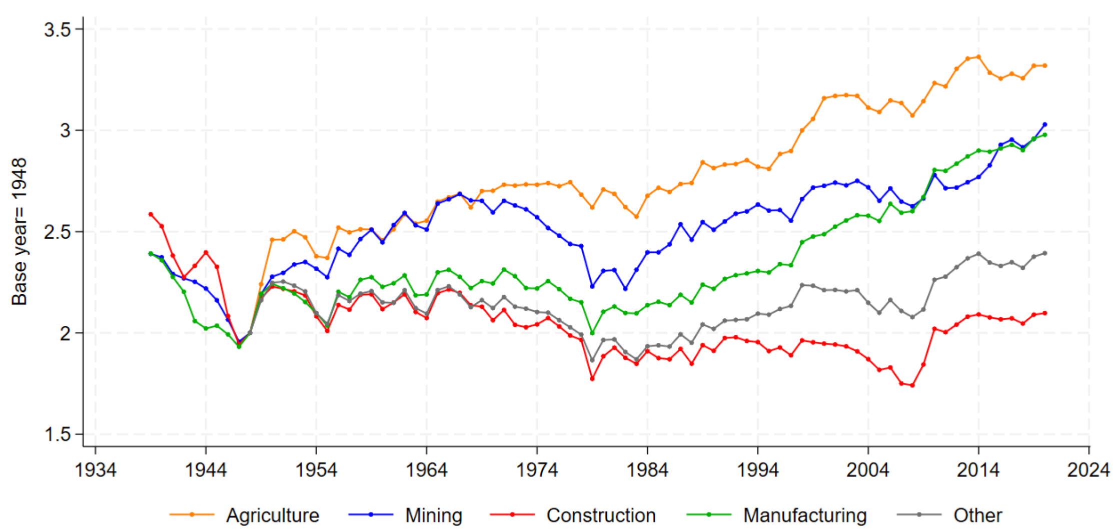
Note. The figure plots by industry the relative patent levels per employee over time for US-based inventors, with the series for the construction sector, the manufacturing sector, and other industries indexed to 1939.

Figure G25: Patent Levels by Industry, Including Mining and Agriculture Separately



Note. The figure plots by industry the relative patent levels over time for US-based inventors, indexed to 1939.

Figure G26: Patent Levels Normalized by Industry Employment, by Industry Including Mining and Agriculture Separately (Log Scale)



Note. The figure plots by industry the relative patent levels per employee over time for US-based inventors. Series are indexed to 100 in 1939, and the y-axis uses a base-10 log scale.

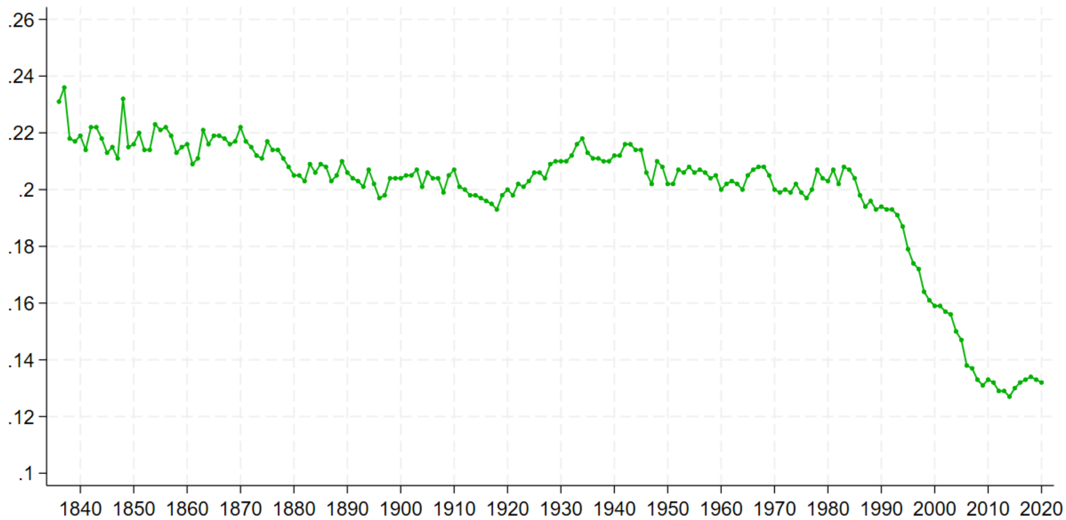
G.2 Innovation by Upstream Suppliers

Here we focus on patenting in the upstream manufacturing industries that supply key inputs to construction. To define the manufacturing subindustries that are most relevant for construction we use the 1997 BEA input-output matrix and look at inputs to construction from manufacturing. Within industries that provide these inputs, we compute the fraction of inputs that construction sources from each of these industries, as a share of total inputs sourced from all manufacturing industries. We define an industry as a primary input industry to construction if that industry is in the top 20 industries in terms of this fraction. These 20 industries account collectively for 80.1% of manufacturing inputs of the construction sector and each accounts for more than 1.3%.[17] To apportion patents to each of these subsectors of manufacturing we use the mapping concordance developed by Kerr (2008).

Having defined primary manufacturing input industries for construction, as well as the patents relating to each manufacturing subindustry, we plot the fraction of patents of all primary input industries as a share of total manufacturing patents. Figure G27 shows that this share started declining precipitously after 1980. This is not due to a decline in the raw patents of these industries. Their count grows somewhat but does not keep pace with the broad rate of patenting in manufacturing.

[17] The top 5 industries (43.5% of inputs) are Millwood, Veneer, Plywood, & Structural Wood Members; Miscellaneous Plastic Products, NEC; Concrete, Gypsum and Plaster Products; Fabricated Structural Metal Products; Electric Lighting and Wiring Equipment. The next 10 (29.5% of inputs) are: Sawmills and Planing Mills; Miscellaneous Wood Products, Paints and Allied Products; Petroleum Refining, Asphalt Paving and Roofing Materials; Miscellaneous Nonmetallic Mineral Products, Blast Furnace and Basic Steel Products; Nonferrous Rolling and Drawing; Miscellaneous Fabricated Metal Products; Refrigeration and Service Machinery. The top 15 to top 20 industries are: Electric Distribution Equipment; Cutlery, Hand Tools and Hardware; Construction and Related Machinery; Wood Buildings and Mobile Homes; Plumbing and Heating, Except Electric.

Figure G27: Share of Manufacturing Patents by Manufacturing Industries That Are Primary Suppliers for Construction, Relative to All Manufacturing Industries



Note. The figure plots the share of patents by manufacturing industries that are primary suppliers for the construction sector, as a share of total patenting in manufacturing.

G.3 Compustat R&D Data

Table G26: R&D as a Percentage of Total Revenues

		Firms with non-missing R&D expenditure			All firms, imputing R&D=0 if missing		
		R&D expenditure (% of revenues)	# of firms	Observations	R&D expenditure (% of revenues)	# of firms	Observations
Before 1974	Construction	.52	14	71	.11	31	611
	Manufacturing	2.41	674	7014	.85	1427	31035
	All other	1.31	329	2532	.24	1065	25866
1974–1983	Construction	.34	33	234	.07	82	1054
	Manufacturing	1.91	1485	17348	1.51	1761	28879
	All other	1.46	1305	10946	.36	2440	37689
1984–1993	Construction	.69	28	201	.05	69	1014
	Manufacturing	3.18	1563	20352	2.54	2089	31380
	All other	2.39	1558	14107	.58	3841	51845
1994–2003	Construction	.66	10	133	.06	91	1079
	Manufacturing	4.22	2157	26495	3.5	2910	38384
	All other	2.34	2568	23147	.59	6313	80398
2004–2013	Construction	.65	11	98	.04	45	696
	Manufacturing	4.02	1760	23163	3.47	2101	30260
	All other	1.66	1585	17072	.46	4799	72520
2014–2023	Construction	.47	7	51	.02	56	598
	Manufacturing	4.99	2175	20652	4.37	2466	25454
	All other	2.94	1640	13935	.91	4623	58708

Note. Authors' calculations based on Compustat data. R&D as a percentage of revenues is computed on the sector-period as a whole.